Copyright Notice

IIUM ENGINEERING JOURNAL EDITORS

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 https://journals.iium.edu.my/ejournal. Further correspondence on the status of the paper could be done through the journal website.

INTERNATIONAL ADVISORY COMMITTEE

IIUM Engineering Journal

Volume 24, Issue 2, July 2023
https://doi.org/10.31436/iiumej.v24i2

Table of Contents

MATERIALS AND MANUFACTURING ENGINEERING

MECHANICAL AND AEROSPACE ENGINEERING

MECHATRONICS AND AUTOMATION ENGINEERING

BIODEGRADATION OF FATS, OIL AND GREASE USING MICROORGANISMS ISOLATED FROM PALM OIL MILL EFFLUENT

MA'AN FAHMI RASHID AL KHATIB[*], **FADI ALQEDRA AND MD. ZAHANGIR ALAM**

Biotechnology Engineering Research Center (BERC),
Department of Chemical Engineering and Sustainability, Kulliyyah of Engineering,
International Islamic University Malaysia, Kuala Lumpur, Malaysia

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

ABSTRACT: The biodegradation of fat, oil, and grease (FOG) is important in water pollution control and wastewater management. In this study, the viability of FOG-degrading microorganisms on palm oil biodegradation was assessed. Seven strains capable of degrading FOG were isolated from palm oil mill effluent (POME). The potential bacterial strains were selected based on Tween-80-degrading ability. *Micrococcus lylae* strain DSM 20315 showed the highest growth compared to the other strains. Hence, it was selected for FOG degradation test. The biodegradability was performed as a function of pH (6, 7, 8), initial oil concentration (1, 3, 5% v/v), and inoculum concentration (2, 6, 10% v/v). Optimization of these parameters of palm oil degradation was studied using 2-level factorial design. The maximum oil degradation was 68%, obtained at pH 6, initial oil concentration 1 % v/v, and bacterial inoculum concentration of 10 % v/v. The lowest oil degradation obtained was 22%. The initial oil concentration followed by bacterial inoculum concentration enhanced the removal efficiency of FOG, but the pH level did not significantly promote the degradation rate. As a result, the optimum process conditions for maximizing oil degradation were at pH 6, initial oil concentration 1 % v/v, and bacterial inoculum concentration of 10 % v/v.

ABSTRAK: Biodegradasi lemak, minyak, dan gris (FOG) adalah penting dalam kawalan pencemaran air dan rawatan air buangan. Kajian ini adalah berkenaan kebolehhidupan organisma pengurai-FOG dalam biodegradasi minyak kelapa sawit. Tujuh strain berkeupayaan mendegradasi FOG diasingkan daripada cairan buangan minyak kelapa sawit (POME). Strain bakteria yang berpotensi telah dipilih berdasarkan keupayaan degradasi-Geladak-80. Strain *Mikrokokus lilae* DSM 20315 menunjukkan pertumbuhan tertinggi berbanding strain lain. Oleh itu, ia dipilih bagi ujian degradasi FOG. Keupayaan biodegradasi telah dihasilkan berdasarkan fungsi pH (6, 7, 8) ketumpatan awal minyak (1, 3, 5% v/v) dan ketumpatan inokulum (2, 6, 10% v/v). Parameter optimum degradasi minyak kelapa sawit dikaji menggunakan reka bentuk faktorial 2-tahap. Nilai maksimum degradasi minyak adalah sebanyak 68%, terhasil pada pH 6, berketumpatan awal 1% v/v, dan ketumpatan inokulum bakteria 10% v/v. Degradasi minyak terendah pula adalah sebanyak 22%. Ketumpatan awal minyak diikuti ketumpatan bakteria inokulum meningkatkan kecekapan penyingkiran FOG, tetapi level pH tidak ketara dalam membantu kadar degradasi. Sebagai kesimpulan, keadaan optimum bagi degradasi minyak maksimum adalah pada pH 6, ketumpatan awal minyak 1% v/v dan ketumpatan bakteria inokulum sebanyak 10% v/v.

1. INTRODUCTION

Wastewater with high fat, oil, and grease (FOG) concentrations is attracting more attention because of population growth and industrialization. FOG accumulation in sewer systems has a negative impact on both health and the environment, thus effective strategies for controlling FOG deposition are critical. Triglycerides, or triglycerides linked to glycerol molecules, are the building blocks of fat, oil, and grease. The physical properties of fat, oil, and grease vary depending on the type of fatty acids that comprise the triglycerides and their physical state.

FOG is produced by the dairy industry, slaughterhouses, food processing plants, and other industries that process fatty substances, in addition to private residences [1]. When wastewater carrying fat, oil, and grease flow in sewer pipes, FOG will deposit in the pipes by forming layers of lipid-rich materials [2]. Sewer blockage causes over 25,000 flooding events in the United Kingdom each year. FOG is thought to account for more than half of all sanitary sewer overflows (SSOs) [3]. The American Environmental Agency (EPA) estimates that between 10,350 and 36,000 sanitary sewer spills occur in the United States each year, with FOG accounting for approximately 47% of the total. Furthermore, Indah Water Konsortium (IWK), the wastewater municipality in Malaysia, reported that up to 45% of SSOs may be caused by FOG [4]. Due to their prevalence in sewer systems (such as pipelines, pump stations, wet wells, etc.) and subsequent wastewater treatment facilities, FOG deposits have an impact on health and the environment. They often result in sanitary sewer overflows (SSOs) by reducing sewer diameter or totally blocking the pipelines. The ensuing sewage discharge increases water contamination and pathogen exposure. FOG can also attract vermin such as rats, and sloughed deposits can clog pumping stations and sewage treatment plants [5]. To implement effective control measures, it is necessary to have a complete understanding of the properties and impacts of FOG.

Approaches to controlling FOG can be categorized as physical, chemical, or biological. Long chain fatty acids are introduced to wastewater during the chemical hydrolysis of FOG, which may limit microbial activity and disrupt the wide variety of microorganisms present [6], resulting in a decline in wastewater treatment efficacy and recurrent production of intolerable odors.

Physical methods use FOG's reduced density in comparison to water. Grease traps (also known as grease interceptors) are one of the most frequently used physical-based FOG removal devices used prior to entering wastewater systems. When FOG-containing wastewater enters a grease trap, lower density FOG floats to the surface, allowing virtually FOG-free water to exit the grease trap and enter the sewer system. The FOG that has floated to the surface will be manually removed. Tilted plates (TP) are a modified grease trap. TP are parallel gravity separators that offer a large surface area while taking up less than 10% of the volume of a typical grease trap [7]. During dissolved air foliation, micro-bubbles bind to FOG particles, allowing them to float to the water's surface and be swiftly removed. To clear the accumulating FOG layers, physical approaches require human intervention. The increased demand for human resources raises the expense of FOG cleaning and the maintenance of related facilities, which already cost millions of dollars per year. Furthermore, these techniques are ineffective at reducing FOG, which can eventually limit the oxygen transfer rate, compromising biological treatment [8]. Additionally, it was shown that microbial activity in grease traps generated large amounts of long-chain fatty acids, which, when released into sewer systems, aid in the production of FOG deposits [9].

Biological FOG treatment is gaining traction due to its ability to treat FOG at the source and prevent it from entering sewer systems. FOG is biologically treated by bacteria capable

of secreting enzymes that speed up the hydrolysis process or by direct use of enzymes such as lipases and commercial additives. Isolating such bacterial strains and determining the optimal parameters and overall process. Teixeira et al. isolated bacteria from activated sludge and wastewater effluent to biodegrade triolein and oleic acid [10]. *Aeromonas* sp. and *Staphylococcus cohnii* were shown to be the best degraders. However, in a 7-day assessment of FOG content, the elimination of oleic acid and/or triolein was 37.9% and 19.1%, respectively. Phong et al. identified bacteria from wastewater collected from restaurants and canteens containing vegetable oil [11]. Because of its strong lipid degradation ability, *Acinetobacteria soli* strain AL3 was selected as a potential for FOG biodegradation. Similarly, strains belonging to the genera *Acinetobacter*, *Bacillus*, and *Pseudomonas* demonstrated the ability to efficiently degrade FOG [12]. The FOG removal efficiency of *Pseudomonas* sp. strain D2D3 were 94.5% and 94.4% for olive oil and animal fat, respectively, whereas sunflower oil had the lowest percentage at 62% [2]. It is beneficial to isolate the bacteria from a medium that is rich in FOG and where the microorganisms have had time to acclimate to ensure a high degree of biodegradation.

Raw palm oil mill effluent (POME) contains high concentrations of organic substances, including oil and grease, which lead to elevated levels of COD and BOD. POME treatment in open oxidation ponds will eventually reduce these amounts significantly [13]. This suggests that POME would accommodate and sustain bacteria capable of metabolizing palm oil during the natural decomposition process in the ponds. Many studies have demonstrated the potential of isolating lipolytic bacteria from POME [14]. Some strains, however, were unable to use palm oil directly as a carbon source. As a result, the isolation and screening procedures were modified to select bacterial strains capable of using palm oil as a carbon source. POME is a preferred source for isolating lipid-degrading bacteria due to its nature and potential.

The objective of this study is to isolate bacteria from palm oil mill effluent and screen them for lipolytic activity on Tween 20. The high-performing bacteria will then be used to biodegrade cooking palm oil, which is a popular oil with a fatty acid composition similar to that of FOG. The selected bacteria's growth conditions will then be optimized.

2. MATERIALS AND METHODS

2.1 Screening and Isolation of Lipolytic Bacteria

POME that was serially diluted from 101 to 107 was used to isolate bacterial strains. A sample of each dilution was plated on LB agar and incubated overnight at 37 °C [15]. To obtain the pure strains, the colonies that formed on the agar plates were sub-cultured on new agar plates. More than 20 pure strains were obtained. The lipase activity was seen using Kanmani et al.'s agar plate assay [16]. Tween 20 (Sigma-Aldrich) was utilized as a substrate since it is suitable for the simple and rapid detection of microbial lipolytic activity on an agar plate. The pure colonies were streaked on Tween 20 agar plates. Lipase producers have a zone of white-like precipitate surrounding the lipase-active colony. The first 7 strains that appeared on agar were selected for further screening.

2.2 Characterization of the Isolated Strains

Individual bacteria isolates were identified beforehand using classical tests such as cell shape, Gram staining, and colony morphology on solid nutrient media. First BASE Laboratories Sdn. Bhd., Malaysia, performed the genetic identification of isolates by determining the nucleotide sequence of the 16S rRNA gene.

2.3 Growth Profile of the Isolated Strains

The lipolytic strains were added to a 250 mL Erlenmeyer flask containing 1 mL of palm oil as a carbon source and 100 mL of minimal salt medium (MSM) media. The author of [15] describes how to make an inoculum in LB broth medium. The inoculum was added to 100 mL of MSM containing 2.5 g/L NaCl, 4.74 g/L K_2HPO_4, 0.56 g/L KH_2PO_4, 0.5 g/L $MgSO_4.7H_2O$, 0.1 g/L $CaCl.6H_2O$, and 0.5 g/L NH_4NO_3 in dH_2O. The media (100 mL) were poured into a 250 mL Erlenmeyer flask and autoclaved for 15 minutes at 121 °C. All samples were inoculated with a starter culture (1% v/v) produced in LB medium (OD600=1.0) [15]. Before inoculation, the starter culture was centrifuged (13,000 rpm, 10 minutes, 4°C) and rinsed a couple of times in a physiological salt solution to avoid the transfer of LB organic compounds into fresh media. Bacteria were omitted from the control samples. All the samples underwent a two-week incubation period in a rotary shaker (150 rpm, 25 °C). On a UV-Visible spectrophotometer every 6 hours, the growth was measured at absorbance of 600 nm [17].

2.4 Optimization of FOG-Degradation

Biodegradation of palm oil was performed to determine the highest concentration of palm oil that could be degraded using bacterial strains in 250 mL Erlenmeyer flasks that were stoppered with cotton wool to promote oxic conditions. Each flask held 100 mL of MSM, supplemented with a predetermined oil concentration, inoculum concertation, and pH level. 1N NaOH or HCl were used to adjust the pH. Before inoculation, samples were autoclaved. The experiment lasted two weeks at 37 °C and 150 rpm. Experiments with bacterial samples and controls were carried out in duplicate.

Using Design Expert 6.0.8 software, biodegradation optimization runs were designed applying 2-Level Factorial Design (2LFD) with one replicate and three center points. The optimized parameters were the pH of the MSM medium (A), the concentration of oil (B), and the inoculum concentration (C). These variables were assessed at 3 different levels: pH (6, 7, 8), oil concentration (1, 3, 5% v/v), and inoculum concentration (2, 6 and 10% v/v). The software developed a matrix containing 11 experiments. Under predetermined conditions, oil degradation was carried out as designed. Following the incubation period, the oil was degraded using gravimetric methods. The software was used to enter and analyze these data. ANOVA (standard analysis of variance) and a contour plot were produced.

The partition gravimetric technique was implemented to estimate the medium's residual oil content [18] The residual oil content was strained numerous times with 5 mL of hexane such that there was no oil layer in the aqueous phase and the solvent phase was clear. In a water bath at 70°C, the mixed solvent extract was completely evaporated, and the amount of residual oil was calculated to determine the percent decrease [3].

3. RESULTS AND DISCUSSION

3.1 Growth Profile of the Isolated Strains

To construct an efficient oil-containing wastewater treatment system, oil-degrading microorganisms were isolated from POME. A sample of palm oil mill effluent was serially diluted and screened. In the initial screening, twenty POME isolates were obtained. The oil-degrading ability of the selected strains was then tested using Tween 20 agar medium. Seven strains that displayed the highest precipitate formation around the colonies in Tween 20 agar plates were selected for subsequent FOG biodegradation experiments. The growth profiles of the seven strains were screened using optical density, and the strain with the highest cell

growth rate was investigated for oil degradation optimization. Figure 1 depicts the growth curve for the selected seven. Strain X3 demonstrated the highest growth, hence it was selected for further investigation.

3.2 Characterization of Isolated Bacteria

Based on gram staining of the isolated strains, Table 1 indicates that two strains belong to gram negative category, while the other five strains belong to gram positive bacteria. Table 2 shows the phenotypical characteristics of the seven isolates when grown on LB agar.
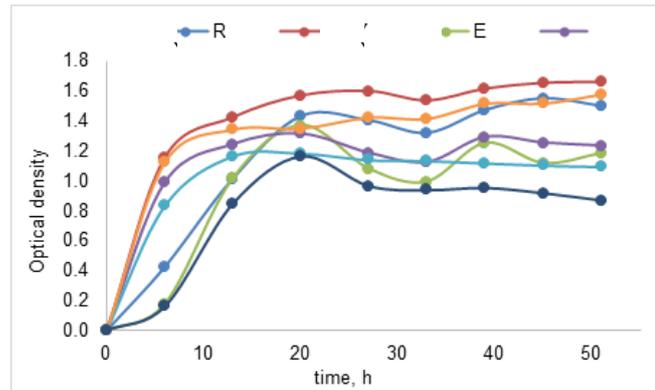
Fig. 1: The growth curve for the seven selected strains.

Table 1: Identification of the isolated strains

Sample	Gram staining	Identification of the strains
(X1, X2,X3)	Gram positive	*Micrococcus lylae* strain DSM 20315
X4	Gram positive	*Corynebacterium aurimucosum* strain H2456
X5	Gram negative	*Lysinibacillus boronitolerans* strain 10a
X6	Gram positive	*Staphylococcus hominis* subsp. novobiosepticus strain GTC 1228
X7	Gram negative	*Bacillus drentensis* strain LMG 21831

Table 2: Phenotypical characteristics of the seven isolates when grown on LB agar

Isolate	(X1, X2, X3)	X4	X5	X6	X7
Form of colony	Irregular	Circular	Circular	Circular	Circular
Elevation	Convex	Convex	Flat	Flat	Flat
Margin	Entire	Entire	Entire	Entire	Entire
Pigmentation	Yellow	White	Red	White	Red
Texture	Rough	Smooth	Smooth	Smooth	Smooth
Optical properties	Opaque	Opaque	Opaque	Opaque	Opaque
Appearance	Dull	Shiny	Shiny	Shiny	Shiny

Strain X3, which showed the highest degradation capability, was identified using morphological tests and 16S RNA sequence analysis (Fig. 2). Strain X3 was gram-positive bacteria capable of producing lipase. The 16S ribosomal RNA (16S rRNA) gene was sequenced and the generated sequences were analyzed using BLAST to reveal that X3 belongs to the genus *Micrococcus*. The 16S rRNA (500bp) sequence analysis of strain X3 revealed a significant similarity (more than 99%) to that of the genus *Micrococcus*. Based on these traits and its 16S rRNA sequence, strain X3 was identified as *Micrococcus lylae* strain DSM 20315.
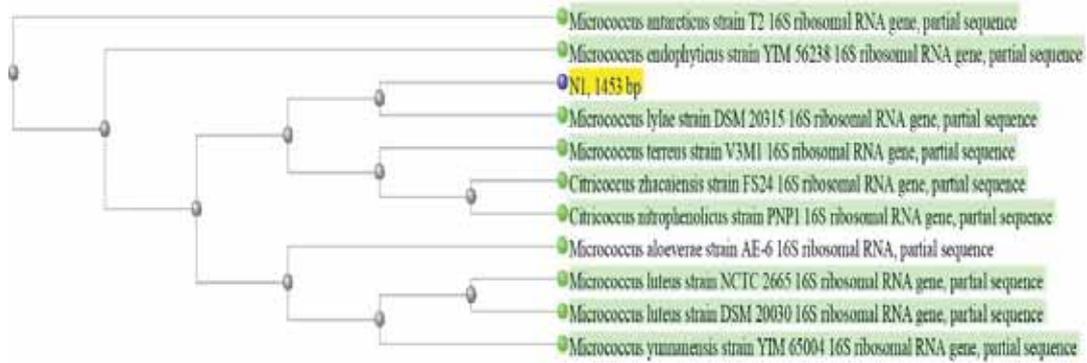
Fig. 2: Phylogenetic tree of X3 strain.

In the present study, *Micrococcus lylae* strain DSM 20315 was shown to have lipolytic activity. To the knowledge of the researchers, this study was the first time *Micrococcus lylae* strain DSM 20315 was used as a potential lipid-degrading strain. This supports the hypothesis that POME is a potential source for oil-degrading bacteria [5].

Table 3: Two LFD Experimental design setup and response

Run no.	A pH	B Oil conc. % (v/v)	C Inoculum conc. % (v/v)	Response Oil biodegradation %		Std
				Exp'l	Pred.	
1	7.0	3.0	6.0	35.3	44.9	6.79
2	8.0	5.0	10.0	25.4	22.8	1.84
3	8.0	1.0	2.0	58	55.4	1.84
4	6.0	5.0	2.0	38.6	36.0	1.84
5	7.0	3.0	6.0	33	44.9	8.41
6	8.0	1.0	10.0	53	55.6	1.84
7	7.0	3.0	6.0	37	44.9	5.59
8	6.0	1.0	10.0	68	65.4	1.84
9	6.0	5.0	10.0	29.2	31.8	1.84
10	8.0	5.0	2.0	37	39.6	1.84
11	6.0	1.0	2.0	50	52.6	1.84

3.3 Statistical Optimization of FOG-Degradation

The highest biodegradation achieved following the experimental design setup (Table 3) was 68% at pH 6, 1% (v/v) oil concentration, and 10% (v/v) inoculum concentration (run no. 8). Lipid-degrading bacteria such as *Bacillus thermoleovorans* IHI-91 was found to be able to degrade palmitic acid, steric acid, lanoline, olive oil, sunflower seed oil, soya oil, and fish oil as sole carbon source [19]. [2] isolated and identified *Pseudomonas* sp. strain D2D3 that is capable of degrading fat, oil, and grease. They also found that the degradation rate ranged from (62.0% -94.5%) based on the substrate. [20] investigated the lipolytic activity of *Raoultella planticola* bacterial strain and found its efficiency to degrade a mixture of lipids substances with maximum degradation level of 91.9% for oleic acid.

ANOVA for the 2LFD model (Table 4) revealed an F-value of 11.84, indicating that the model is significant and that there is only a 3.4% likelihood that such a big F-value could occur due to noise. Model terms are significant when "Prob > F" is less than 0.05. In this case, B (oil concentration) and BC are significant model terms. Values larger than 0.10

suggest that the model terms are not significant. Generally, large magnitudes of t and F, as well as smaller p-values, indicate that the associated coefficient terms are significant. R^2 indicates how much of the observed response's variability may be attributed to the experimental factors and their interactions. R^2 is a measure of how well a model can predict a response, and the closer it is to 1, the stronger the correlation between experimental and predicted values are. A ratio larger than 4 is preferred when measuring the signal to noise ratio with adequate precision. With a ratio of 10.975, it is clear that there are sufficient signals and models to help one move about the design space. Lack of fit > 0.05 is insignificant, and because the lack of fit test reflects how well the model fits the experimental data, this model could be used to make predictions.

The final equation terms of coded factors are:

Oil biodegradation = +44.9 -1.55 * A -12.35 * B – 1.00 * C +0.2 * A * B – 3.15 * A * C -4.25 * B * C   (1)

The equation can be used to make predictions about the outcomes, for given amounts of each of the factors. It can also be used to ascertain the relative influence of factors by comparing the factor coefficient.

<div align="center">Table 4: ANOVA of oil biodegradation</div>

source	Sum of squares	Df	Mean square	F-value	p-value Prob >F	
Model	1471.6	6	245.2667	11.839	0.034	significant
A-pH	19.22	1	19.22	0.928	0.407	
B-Oil conc.	1220.18	1	1220.18	58.896	0.005	
C-Inoculum conc.	8	1	8	0.386	0.578	
AB	0.32	1	0.32	0.015	0.909	
AC	79.38	1	79.38	3.832	0.145	
BC	144.5	1	144.5	6.975	0.078	
Pure Error	8.0726	2	4.0363			
Lack of Fit	54.08	1	54.08	13.398	0.067	not significant

Figure 3 shows the predicted versus actual plot values. The data points are evenly distributed across the 45° line, indicating agreement between the actual data and that produced by the model.
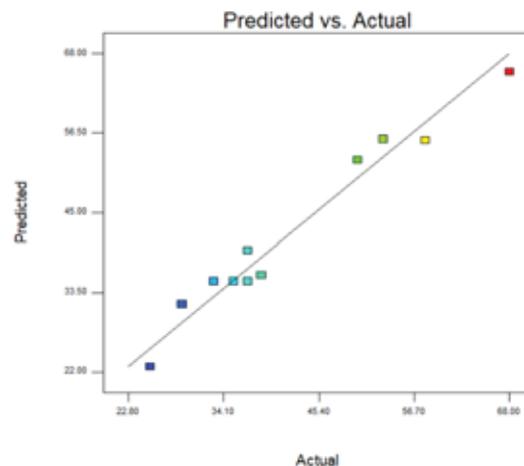
Fig. 3: Predicted versus actual plot of oil biodegradation percentage.

(a)

(b)

(c)
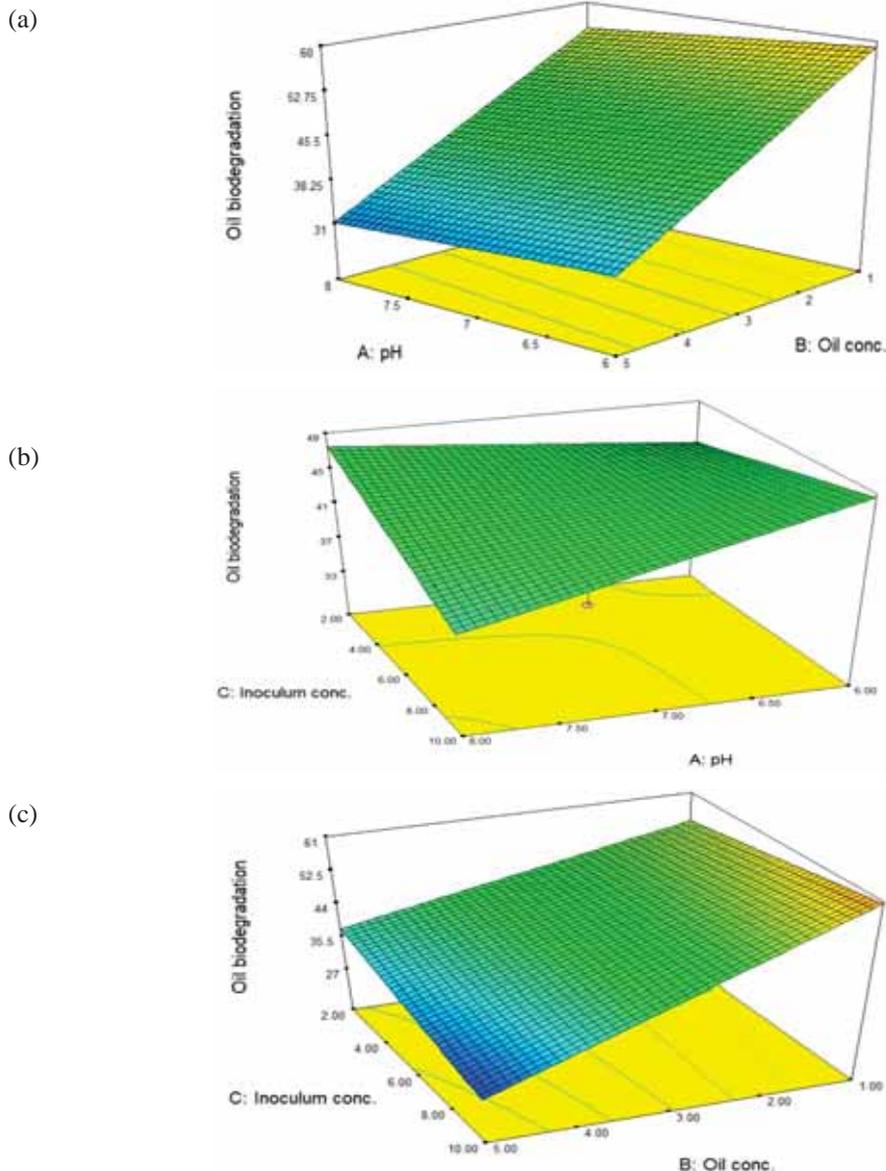
Fig. 4: 3D response surface showing the effects of (a) pH and oil concentration (b) pH and inoculum concentration and (c) oil concentration and inoculum concentration on oil degradation.

The 3D response surface plots are a graphical approach to expressing the regression equation. They illustrate how the variables interact and are used to identify the optimal value of each factor for a suitable response. The plots for the interactions between pH and oil concentration (A-B), pH and inoculum concentration (A-C), and oil concentration and inoculum concentration (A-C) are shown in Fig. 4 a, b, and c, respectively (B-C).

Figure 4a shows a decrease in oil biodegradation with an increase in oil concentration while pH did not show any effect. Figure 4b depicts the interaction between pH and inoculum concentration. Increased pH resulted in a slight increase in oil biodegradation, whereas increased inoculum concentration resulted in an increase in oil biodegradation at lower pH. In Fig. 4c, the interaction between oil concentration and inoculum concentration shows that oil degradation is highest when the oil concentration is lowest, and the inoculum concentration is highest. This is consistent with what is stated by [21]. Increased bacterial enzyme accessibility to substrate will result in improvement of the hydrolysis reaction rate.

This study revealed that while microbial concentration does not have significant effect as a single factor (Table 4) its interaction with oil concentration has a significant effect on the amount of degraded oil. Microbial concentration is important as it provides more bacterial cells and thus more enzyme secretion to degrade the oil and therefore increase the rate of degradation [22].

By applying the 2-Level Factorial Design of Oil degradation, a maximum oil degradation was obtained at a pH of 6, oil conc. 1% (v/v), and 10% (v/v) of bacterial inoculum.

4. CONCLUSIONS

Seven microorganisms isolated from POME were studied for the degradation of FOG. *Micrococcus lylae* strain DSM 20315 was identified as the largest growing microorganism through morphological tests and 16S rRNA sequence analysis. This strain was used to optimize the degradation process. The optimum condition that achieved the highest biodegradation was at pH 6, oil concentration 1% (v/v), and 10 % (v/v). The results revealed that microorganisms isolated from POME are capable of biodegrading FOG, which could be attributed to the two wastes' similar properties.

ACKNOWLEDGMENT

REFERENCES

[1] Brooksbank AM, Latchford JW, Mudge SM. (2007) Degradation and modification of fats, oils and grease by commercial microbial supplements. World Journal of Microbiology and Biotechnology 23 (7): 977-985. doi: 10.1007/s11274-006-9323-1.

[2] Shon HK, Tian D, Kwon DY, Jim CS, Lee TJ, Chung WJ. (2002) Degradation of fat, oil, and grease (FOGs) by lipase-producing bacterium *Pseudomonas* sp. strain D2D3. Journal of Microbiology and Biotechnology, 12(4): 583-591.

[3] Williams JB, Clarkson C, Mant C, Drinkwater A, May E. (2012) Fat, oil and grease deposits in sewers: Characterisation of deposits and formation mechanisms. Water Research, 46(19): 6319–6328. doi: 10.1016/j.watres.2012.09.002.

[4] Husain IAF, Alkhatib MF, Jammi MS, Mirghani MES, Zainudin Z, Hoda S. (2014) Problems, Control, and treatment of fat, oil, and grease (FOG): A Review. Journal of Oleo Science, 63(8): 747-752. doi: 10.5650/jos.ess13182.

[5] Del Mundo DMN, Sutheerawattananonda M. (2017) Influence of fat and oil type on the yield, physico-chemical properties, and microstructure of fat, oil, and grease (FOG) deposits. Water Research, 124: 308-319. doi: 10.1016/j.watres.2017.07.047.

[6] Ma J, Zhao QB, Laurens LLM, et al. (2015) Mechanism, kinetics and microbiology of inhibition caused by long-chain fatty acids in anaerobic digestion of algal biomass. Biotechnology for Biofuels, 8(1): 1-12. doi: 10.1186/s13068-015-0322-z.

[7] Willey R. (2001) Fats, oils, and greases: The minimization and treatment of wastewaters generated from oil refining and margarine production. Ecotoxicology and Environmental Safety, 50 (2): 127–133. doi: 10.1006/eesa.2001.2081.

[8] Abomohra AEF, Elsayed M, Esakkimuthu S, Sheekh ME. (2020) Potential of fat, oil and grease (FOG) for biodiesel production: A critical review on the recent progress and future perspectives. Progress in Energy and Combustion Science, 81: 100868. doi: 10.1016/j.pecs.2020.100868.

[9] He X, de los Reyes FL, Ducoste JJ. (2017) A critical review of fat, oil, and grease (FOG) in sewer collection systems: Challenges and control. Critical Reviews in Environmental Science and Technology, 47(13): 1191-1217. doi: 10.1080/10643389.2017.1382282.

[10] Teixeira PD, Silva VS, Tenreiro R. (2020) Inegrated selection and identification of bacteria from polluted sites for biodegradation of lipids. International Microbiology, 23(3): 367-380. doi: 10.1007/s10123-019-00109-w.

[11] Hong NT, Duyen NT, Diep CN. (2014) Isolation and Characterization of Lipid-Degrading Bacteria in Wastewater of Food Processing Plants and Restaurants in Can Tho City, Vietnam. American Journal of Life Sciences, 2(6): 382-388.
doi: 10.11648/j.ajls.20140206.18.

[12] Bhumibhamon O, Koprasertsak A, Funthong S. (2002) Biotreatment of High Fat and Oil Wastewater by Lipase Producing Microorganisms. Kasetsart J(NatSci), 36: 261-267.

[13] Alias Z, Tan IKP. (2005) Isolation of palm oil-utilising, polyhydroxyalkanoate (PHA)-producing bacteria by an enrichment technique. Bioresource Technology, 96(11): 1229-1234.
doi: 10.1016/j.biortech.2004.10.012.

[14] Affandi IE, Suratman NH, Abdullah S, Ahmad WA, Zakaria ZA. (2014) Degradation of oil and grease from high-strength industrial effluents using locally isolated aerobic biosurfactant-producing bacteria. International Biodeterioration & Biodegradation, 95: 33-40. doi: 10.1016/j.ibiod.2014.04.009.

[15] Kis Á, Laczi K, Zsíros S, Rákhely G, Perei, K. (2015) Biodegradation of animal fats and vegetable oils by *Rhodococcus erythropolis* PR4. International Biodeterioration and Biodegradation, 105: 114-119. doi: 10.1016/j.ibiod.2015.08.015.

[16] Kanmani P, Kumaresan K, Aravind J. (2015) Utilization of coconut oil mill waste as a substrate for optimized lipase production, oil biodegradation and enzyme purification studies in Staphylococcus pasteuri. Electronic Journal of Biotechnology, 18(1): 20-28.
doi: 10.1016/j.ejbt.2014.11.003.

[17] Iqbal SA, Rehman A. (2015) Characterization of lipase from *Bacillus subtilis* I-4 and its potential use in oil contaminated wastewater. Brazilian Archives of Biology and Technology 58(5): 789-797. doi: 10.1590/S1516-89132015050318.

[18] Latha R, Kalaivani R. (2012) Bacterial Degradation of Crude Oil by Gravimetric Analysis. Analysis 3(5): 2789-2795.

[19] Markossian S, Becker P, Märkl H, Antranikian G. (2000) Isolation and characterization of lipid-degrading *Bacillus thermoleovorans* IHI-91 from an icelandic hot spring. Extremophiles : life under extreme conditions 4 (6): 365–371. doi: 10.1007/s007920070006.

[20] Sugimori D, Watanabe M, Utsue T. (2013) Isolation and lipid degradation profile of *Raoultella planticola* strain 232-2 capable of efficiently catabolizing edible oils under acidic conditions. Applied Microbiology and Biotechnology 97(2): 871-880. doi: 10.1007/s00253-012-3982-7.

[21] Tzirita M, Papanikolaou S, Chatzifragkou A, Quilty B. (2018) Waste fat biodegradation and biomodification by Yarrowia lipolytica and a bacterial consortium composed of *Bacillus* spp. and *Pseudomonas putida*. Engineering in Life Sciences 18(12): 932-942.
doi: 10.1002/elsc.201800067.

[22] Campbell MK, Farrell SO (2016) Biochemistry. Cengage Learning.

EXPERIMENTAL AND COMPUTATIONAL ANALYSIS FOR OPTIMIZATION OF SEAWATER BIODEGRADABILITY USING PHOTO CATALYSIS

NAYEEMUDDIN MOHAMMED[1*], PUGANESHWARY PALANIANDY[1], FEROZ SHAIK[2] AND HIREN MEWADA[2]

[1]*School of Civil Engineering, Universiti Sains Malaysia, Penang, Malaysia*
[2]*Prince Mohammad Bin Fahd University, Al Khobar, Kingdom of Saudi Arabia*

[*]Corresponding author: mnayeemuddin@student.usm.my

ABSTRACT: Seawater pollution is a significant global environmental problem. Various technologies and methods have been used to remove the contaminants found in saltwater. This experimental study investigates the degradation of contaminants present in seawater using solar photocatalysis, where a combination of TiO_2 and ZnO was used. The effects of catalyst dosage, pH, and reaction duration were assessed using percentage removal efficiencies of total organic carbon (TOC), chemical oxygen demand (COD), biological oxygen demand (BOD), and biodegradability (BOD/COD). Biodegradability is essential for removing pollutants from saltwater and plays a vital role. The higher the biodegradability, the more efficient the treatment procedure will be. The most effective percentage reduction rates from the experimental data obtained were TOC=59.80%, COD=75.20%, BOD=23.94%, and biodegradability=0.055. For modeling, optimizing, and assessing the effects of parameters, the Design Expert based on Box *Behnken* design (RSM-BBD) and a predictive model based on the MATLAB adaptive neuro-fuzzy inference system (ANFIS) tools were used. The coefficient of determination R^2 was found to be 0.977 for the RSM-BBD model and 0.99 for the ANFIS model. According to the RSM-BBD design, the maximum percentage pollutant elimination efficiencies were found to be TOC=55.4, COD=73.4, BOD=23.70%, and BOD/COD=0.054, but for the ANFIS model, they were TOC=59.4, COD=75.4, BOD=24.1%, and BOD/COD=0.055. It was discovered that the ANFIS model outperformed RSM-BBD in process optimization.

ABSTRAK*: :* Pencemaran air laut adalah masalah alam sekitar global yang ketara. Pelbagai teknologi dan kaedah telah digunakan bagi menyingkirkan pencemaran yang dijumpai dalam air laut. Kajian eksperimen ini menilai degradasi pencemaran yang hadir dalam air laut menggunakan fotopemangkin, di mana kombinasi TiO_2 dan ZnO digunakan. Kesan dos pemangkin, pH, dan tempoh reaksi dipantau menggunakan peratus kecekapan penyingkiran jumlah karbon organik (TOC), keperluan kimia oksigen (COD), keperluan biologi oksigen (BOD), dan kebolehdegradasian (BOD/COD). Kebolehdegradasian adalah sangat penting bagi menyingkirkan bahan cemar dari air laut dan berperanan penting. Semakin tinggi kebolehdegradasian, semakin cekap prosedur rawatan. Peratus kadar pengurangan yang paling berkesan daripada data eksperimen adalah didapati pada TOC=59.80%, COD=75.20%, BOD=23.94%, dan biodegradasi=0.055. Bagi mengkaji kesan parameter terhadap model, kadar optimum, dan memantau keberkesanan parameter, kaedah Pakar Reka Bentuk pada rekaan Kotak Behnken (RSM-BBD) dan model ramalan berdasarkan sistem pengaruh menggunakan sistem MATLAB iaitu Inferens Neural-Fuzi Boleh Suai (ANFIS) digunakan. Pekali penentu R^2 terhasil pada 0.977 bagi model RSM-

BBD dan 0.99 pada model ANFIS. Berdasarkan reka bentuk RSM-BBD, peratus maksimum keberkesanan penyingkiran bahan cemar dijumpai pada TOC=55.4, COD=73.4, BOD=23.70%, dan BOD/COD=0.054, tetapi bagi model ANFIS, TOC=59.4, COD=75.4, BOD=24.1%, dan BOD/COD=0.055. Model ANFIS adalah lebih berkesan daripada model RSM-BBD dalam proses pengoptimuman.

1. INTRODUCTION

The requirement for fresh and clean water has been increasing day by day. The demand for or availability of drinkable water is a major concern due to the global population rise. Seawater is a major source of freshwater generation, but it also contains several organic, inorganic, and biological contaminants in addition to salt. Before seawater is given to the primary desalination process, these contaminants must be treated. Many traditional techniques are employed to treat the contaminants found in seawater [1], but recently, sophisticated oxidation techniques have been used to remediate the contaminants in saline water.

Natural solar light and artificial UV radiation sources were utilized in heterogeneous photocatalysis using TiO_2 as a photocatalyst. Utilizing photocatalytic reactors, organic contaminants, such as phenol and benzoic acid, were removed from seawater. A semiconductor, such as TiO_2, SnO_2, ZnO, and PbO, served as a photocatalyst in this process and was exposed to artificial and natural light [2]. Seawater contaminants can be degraded by solar photocatalysis, which has been well-defined by the process of solar photocatalysis mechanism [3,4].

Numerous studies have used the solar photocatalytic degradation mechanism to remove contaminants found in saltwater. A batch recirculation reactor system was used to study the photo degradation of contaminants found in saltwater. By assessing different parameters such as TOC, total dissolved solids (TDS), COD, and total inorganic carbon, the performance of photocatalytic degradation was assessed combining TiO_2 photo catalyst and polyamide. The solar photo degradation phase saw a significant parameter drop [5]. Using a Yb-TiO_2-rGO photo catalyst, phenol in saline water was significantly reduced throughout the sun photocatalytic degradation process. To prevent salt ions from adhering to the photocatalyst, ethylene glycol was grafted onto the catalyst [6]. A photocatalyst (TiO_2) that had been immobilized in plug flow reactors was used with an artificial light source (UV lamp) to investigate the breakdown of the model of organic molecule benzoic acid to gauge the photocatalytic performance. Significant benzoic acid decomposition was seen using plug flow reactors of various diameters. Under various circumstances, photocatalysis has been used and shown to be a viable technique for purifying phenolic wastewater. In a batch reactor with a recycle stream, photocatalysis was used to completely mineralize the phenol [7]. Researchers looked at the efficiency of photocatalytic degradation of diesel impurities present in saltwater by studying the sun photocatalytic destruction of those pollutants. The initial concentration of diesel contaminants, pH, catalyst ratio, and dosage were changed. The visible photocatalytic degradation mechanism resulted in a 78.7% degradation of diesel pollutants [8]. The water produced by the oil industry, which contained both organic and inorganic contaminants, underwent nano-solar photocatalysis. Using a TiO_2 catalyst and free

solar energy, experiments were carried out in batch and continuous reactors. It was found that the amount of contaminants in oil-produced water had significantly decreased [9].

A thorough overview of numerous photocatalytic research using saltwater and saline industrial effluent, including oil-generated water, was presented by Nayeem et al. [10]. Photocatalysis refers to the acceleration of a photochemical reaction in the presence of a catalyst. The photocatalytic efficiency in photo-generated catalysis depends on the catalyst's ability to generate electron-hole pairs and free radicals capable of secondary reactions. Furthermore, it is employed in numerous processes known as the Advanced Oxidation Process (AOP) [11]. The photocatalytic reaction breaks down the harmful molecules without leaving any residue, obviating the need to transport sludge to a landfill. An additional advantage is that the catalyst has a long lifespan, and there is no requirement for chemicals in the process, keeping the operation simple and financially viable [12]. An oxidant that absorbs UV light and reacts with water to generate highly reactive OH radicals could be introduced during UV irradiation to optimize elimination efficiency. The most common oxidants are hydrogen peroxide (H_2O_2) and ozone. During the experiments, it was revealed that the absorbance accuracy of H_2O_2 is dependent on its concentration. Therefore, the efficiency of the water treatment system increases with H_2O_2 concentration [13].

For low boron concentrations in genuine desalinated seawater, $UiO-66-NH_2/GO/Fe_3O_4$ shows good adsorption uptake (22.46 mg/g). A novel technique of magnetic composite with a metal-organic frame was employed in the desalination of salt water to remove boron. The single-factor design model was carried out to study the ideal parameter settings using the response surface approach and the MATLAB program ANN tool. The maximal temperature, 318 K, pH 3.38, and dose of 99.1 mg/L were discovered. Both approaches have demonstrated a good ability to anticipate the adsorption process [14] accurately. For the statistical analysis, Hashemi et al. [15] used pipette-tip solid phase extraction with molecularly imprinted polymer and an RSM-BBD model with seven variables at three levels to determine the presence of methyl red in seawater. The computed average recoveries ranged from 84.0 to 98.0 percent, with a mean of 2.5 - 6.7 percent. For the breakdown of tetracycline, a TiO_2 photo catalyst in powdered form was used to treat contaminants found in seawater. Due to the strong visible light reaction, it is observed that after 20 hours, the removal of tetracycline from seawater by PU sponge-filled spheres had achieved 80%, offering the greatest performance [16]. The necessary studies to investigate the impacts of yeast extract, whey, heating temperature, and Caspian Sea water were designed using a central composite design.

According to the response surface method, the maximal specific urease activity (16.50 mm urea.min^{-1}. OD^{-1}) could be attained with 9.94 gL^{-1} of yeast extract, 23.43 gL^{-1} of whey, 128.6 °C of heating, and no seawater [17]. The ANN has the advantage of not requiring the central composite design before experimentation and being a continuously improving prediction approach as more data becomes available. Still, the RSM has a modest advantage in model accuracy [18]. A three-factor and three-level Box Behnken design was used to investigate the novel core-shell micro-structured nanocomposites with inner seawater for syngas/air fire prevention. Two mathematical models were created to examine the interactions between the parameters for salt rejection and to permeate flow in RO membranes [19]. Ions in seawater physically and chemically interact with mineral species, making the flotation and thickening processes harder to operate. ANOVA and RSM-based modeling are used to identify the variables that affect the variable of interest, which is then

optimized. The design of experiments (DOE) establishes the number of experiments needed and the values of the independent and dependent variables [20]. During training, the developed ANN model demonstrated a good agreement between prediction and experimental data, with good statistical metrics values (RMSE, MAE, and AARD). According to ANN, permeate conductivity, flow rate, and recovery were predicted with coefficients of determination of 0.969, 0.942, and 0.963, respectively [21].

The optimization of experimental variables to ascertain the impact of independent factors on the responses has frequently used response surface methods. The Box Behnken tool was used to reduce the number of experiments in RSM with the best response. The pH concentration was found to have the best COD elimination percentage [22, 23]. The establishment of a linear or nonlinear relationship between water salinity and its controlling factors (such as water table, evaporation, and distance to saltwater bodies) and the use of those relationships for the prediction of water salinity for regions with low data points have shown the capacity of ML models to model groundwater salinity [24]. The independent variables considered in ANN were pH, POME concentration, pressure, time, and one hidden layer and output layer. It was noted that in the filtration process, the ANN accurately displays the projected optimum values near the experimental data [25,26]. Response surface approach and a genetic algorithm tool were used to undertake statistical modeling in order to obtain the desirability function and forecast values. It was found that the RSM-GA predicted the experimental values accurately [27,28]. Utilizing RSM (Design Expert) and ANN-LM, a statistical model with three independent variables was created. The removal elimination of COD, TOC, BOD, and biodegradability were predicted by this model [29].

Numerous studies have used the MATLAB tool to predict response variables using ANN, ANFIS, FIS, and other models. A fuzzy system is the first stage in building a network of fuzzy systems and can be created by "if-then" rules. The mathematical techniques are known as the recurrent network (RN), time-lagged recurrent network (TLRN), ANN, and ANFIS are directly derived from the workings of the human brain [30]. These methods are promising for simulating response variables and can also be used with nonlinear systems due to their simplicity. When the relationship between the input and output variables is unknown, and it is impossible to identify the system by mathematical problem, the robust application known as the fuzzy inference system was utilized [31].

Utilizing RSM-BBM and ANN-ANFIS, a few studies have been published on optimizing the removal of contaminants from saltwater using solar photocatalytic degradation. To find the ideal values, polynomial quadratic models were constructed. It was discovered that the central composite design was a reliable and accurate forecast. With five input variables, a Genetic Algorithm was used as a tool in ANN and RSM [32]. The term "biodegradability" refers to the BOD/COD ratio. The elimination process of contaminants from seawater depends on biodegradability and a higher biodegradability can increase the effectiveness of the treatment approach. In addition, there will be a significant decrease in fouling characteristics during the seawater treatment process due to the high biodegradability. However, no research has reported the natural sunlight-based biodegradability of saltwater combining photocatalysts with RSM-BBD and ANFIS applications. The proposed approach uses the combination of TiO_2 and ZnO photocatalysis processes to examine the biodegradability of saline water. The process was validated using RSM-BBD and ANFIS to assess the model and optimize and investigate the relationship between the parameters.

2. MATERIALS AND METHODS

A 5 L sample of salt water was taken at a depth of 10 m from the water's surface, 2.3 km from Khobar Beach in the Kingdom of Saudi Arabia. Commercial TiO_2 Degussa P-25 (80% A-20% R) from Evonik Industries in Germany, which is 99.9% pure, and ZnO (99.9 percent pure, APS:20 nm) from mkNano in Canada were used as the catalysts total carbon analyzer (Shimadzu), thermal scientific Orion COD 125, and AQ 400 were used to measure COD and TOC, respectively. Thermo Fisher Scientific's BOD incubator with a complete water analysis kit (Eutech PCD 650) was used to measure the DO and BOD estimation, and a pH meter (JENWAY 3520) was used to measure the pH value. Table 1 displays the initial properties of seawater.

2.1 Experimental Procedure

Figure 1 depicts an experimental setup sketch for the batch investigations. As a batch reactor, a 1500 mL glass beaker with a magnetic stirrer was utilized. A 1000 mL sample of seawater was obtained in a batch reactor setup, and catalysts TiO_2 and ZnO were added. Between 10:00 AM to 2:00 PM, the photocatalytic reaction was carried out in the open air while stirring with a magnetic stirrer. At regular intervals, the samples were brought out for examination. The measured parameters were calculated by percentage removal efficiencies using equation 1.

$$\text{Percentage removal} = \frac{(P_o - P_f)}{P_o} \times 100 \qquad (1)$$

where, P_0 and P_f are the starting and final concentrations in mg/L.

Table 1: Initial parameters for a seawater sample

Parameter	Values
TOC (mg/L)	2.74
COD (mg/L)	111
BOD (mg/L)	1.75
pH	8.665
Biodegradability, BOD/COD	0.0157
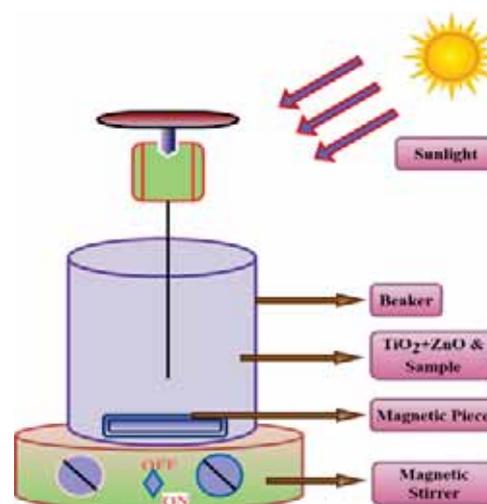
Fig. 1: The batch reactor's experimental setup.

2.2 Statistical Analysis Theoretical Model Description

2.2.1 RSM-Box Behnken Statistical Modelling

Design Expert response surface methodology is one of the major strategies for detecting and illustrating the cause-and-effect relationship between genuine average responses and input control variables influencing responses. For the statistical experimental designs, modelling approaches, regression modelling, prediction, and optimization, the Design-Expert version-11.1-2.0, State-Ease Inc. MN, USA-based Box Behnken Design of response surface methodology was used. RSM combines mathematical methods with statistical analysis to fit the second-order or quadratic model. The Box Behnken Design only needs three levels of each component, such as lower, mid, and upper level, with randomized type, in contrast to the central composite design, which needs five levels of factors. BBD offers greater refining, optimization, and precision than CCD because it is almost rotatable and requires fewer experimental runs. After essential factors are found during the screening of factorial trials, it is also utilized to investigate the effects of quadratic factors. The phases in the RSM-BBM method's strategy are shown in Fig. 2. In the RSM explained by equation 2, the Box Behnken Design utilized the generalized model equation for the second-order polynomial [33].

$$Y = (A_0) + \sum (A_i x_i) + \sum (A_{ii} \, x_i^2) + \sum (A_{ij} x_i x_j) \qquad (2)$$

The input variables (i) and (j) are coded values in this case. The parameter or quadratic value is X_i^2. The linear, interaction terms, intercept, linear regression, and quadratic coefficients for intercept are A0, Ai, Aij, and Ajj. Y stands for the output variables, which include percentages of TOC, COD, BOD, and biodegradability removal efficiencies. To ascertain variable interaction, the statistical significance of models, and the impacts of research variables, ANOVA (analysis of variance) was performed. The equations in the quadratic-polynomial model were determined by the coefficient of determination (R^2), and the significance of regression coefficients using the F-test was evaluated at probabilities (P) 0.001 and 0.01. The 3-dimensional surface plots represent the maximum or ideal extraction conditions. An additional confirmation experiment was conducted to verify the statistical experimental procedures [34].

2.2.2 ANN-ANFIS Statistical Modelling

In this approach, an adaptive neuro-fuzzy inference system was created using MATLAB programming to discover the best parameters for the percentage removal elimination of input and output pairs (ANFIS). The ANFIS is a standard mathematical tool based on artificial intelligence and is used to model complicated nonlinear problems utilizing artificial neural network learning and neuro-fuzzy interference techniques based on the Sugeno first-order system [35]. The system's behavior can be analyzed with large datasets. Therefore, there is a need for system modeling to estimate the output for unseen datasets which fall within range or outside range. ANFIS is one of the powerful tools for such nonlinear time series analysis. This work uses ANFIS modeling to obtain a nonlinear input and output mapping.

Additionally, it employed a hybrid learning algorithm that included the least squares and backpropagation learning algorithm approaches. Figure 3 shows the first-order Sugeno-type inference system's complete ANFIS architecture was built using three

inputs, four outputs, and five layers. The models were constructed using 70% training, 30% testing, and 40% model validation [36].
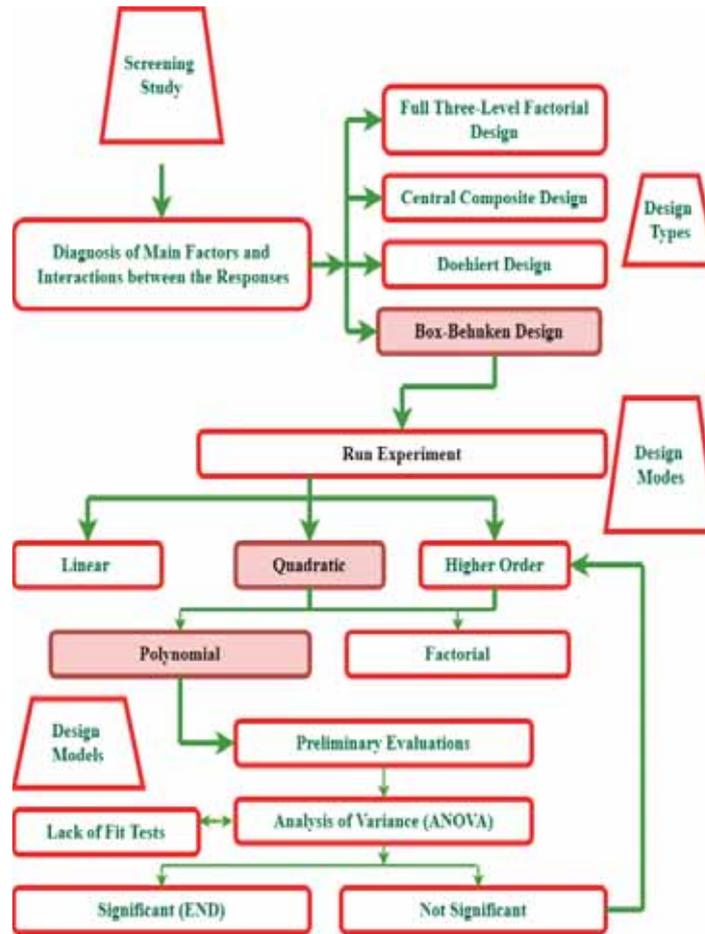
Fig. 2: Procedure for RSM - Box-Behnken [34].
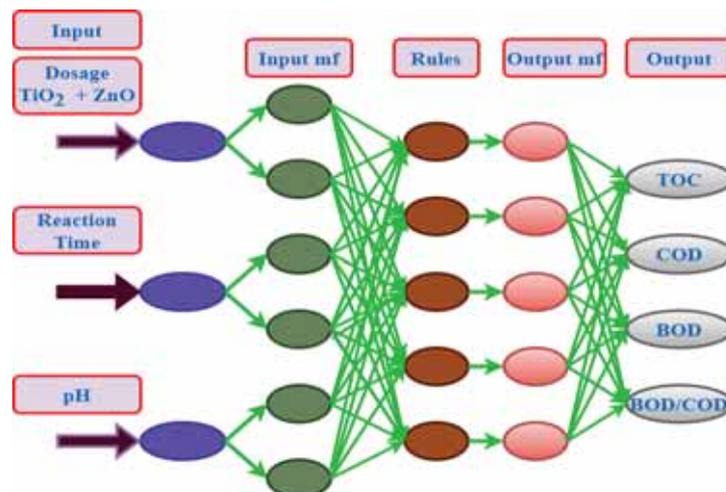
Fig. 3: Schematic diagram of Architectural Network of ANFIS model [36].

Model correctness was measured using several error estimates, such as RMSE and R^2. This ANFIS m was created using the first-order Sugeno model with fuzzy IF-THEN techniques. The roles of each of the five layers were described in terms of two rules (equations 3 and 4) in the section that follows rules 1 and 2 [37].

<u>Rule 1:</u> If (x) is (A_1) and (y) is (B_1), then (f_1) = p_1 (x) + q_1 (y) + (r_1)   (3)

<u>Rule 2:</u> If (x) is (A_2) and (y) is (B_2), then (f_2) = p_2 (x) + q_2 (y) + (r_2)   (4)

where p_1, p_2, q_1, q_2, r_1, and r_2 are the output coefficient functions that will be computed during testing, training, and validation and where x' and y' are the inputs and A_1, A_2, B_1, and B_2 are the fuzzy sets.

The input variables are passed from layer 1 to layer 2, also called the independent layer or input layer. Node I has an adaptive node output expressed by Eq. 5 and Eq. 6.

O_i = μ_{ai} (x),   for (i) = 1, 2   (5)

O_i = μ_{bi-2} (y),   for (i) =3, 4   (6)

Rule nodes are the second tier. Each input's independent values are multiplied by one another and are designated by a circle with the label "fixed nodes, non-adaptive", where w_i is given as weights and is as follows by Eq. 7:

O_i^2 = (w_i) = (μ_{ai}) (x) (μ_{bi}) (y)   for (i) =1, 2   (7)

The third layer with the letter N, is referred to as the layer of average nodes or the normalization layer by Eq. 8.

O_i^3 = (w_i) = (w_i)/(w_1 + w_2) for (i) =1, 2   (8)

Layer 4 is where defuzzification occurs; it is also referred to as the layer of subsequent nodes because, in this layer, the output layer is coupled with the preceding layer using the Sugeno fuzzy function by Eq. 9.

O_i^4 = w_i (f_i) = w_i [pi (x) + qi (y) + r_i]   for i =1, 2   (9)

The output layer is the fifth and final layer. It is denoted by the node with the name "$\sum$," which ultimately determines total output and does so by equation 10.

O_i^5 = f_{out} = $\sum w_i$ (f_i)   (10)

The final output Eq. 11 can be written as:

f $_{out}$ = w_1 (x) p_1 + w_1 (y) (q_1) + w_1 (r_1) + w_{2+} (x) (p_2) + w_2 (y) (q_2) + w_2 (r_2)   (11)

It can be seen from the aforementioned Eq. 3 to Eq. 11 that the ANFIS model depends on the performance of specific parameters, such as the center, and must be appropriately calibrated.

    The design model contains 27 datasets with measured values of dosage, time, and pH. The ANFIS modeling has two phases: training and testing. Therefore, the dataset was divided into training and testing sets. The division of the dataset was obtained with random permutation, and the model was repeated multiple times to get the best modeling. Fuzzy C-mean clustering with Sugeno-rule is used in the ANFIS modeling. Three individual models were developed to estimate three properties, i.e., TOC, COD, and BOD. The modeling needs an appropriate selection of various parameters, including the number of memberships,

clusters, iterations, and epochs. The optimum parameters are obtained by repeating the experiment and analyzing the error obtained in predicted values. For choosing the optimum model, Mean Square Error (MSE), Root Square Error (RMSE), and Correlation Coefficient (R^2) were utilized as performance parameters. A lower value of MSE, RMSE, and a higher value of R^2 indicates superior system modeling. Let $Xp(i)$ is the predicted value from ANFIS modeling and $Xt(i)$ is the actual experimental response. If the total data is N, then quantitative parameters can be calculated using the Eqs. 12-14. The independent or input layer in this statistical modeling of the ANFIS technique comprises three neurons, and the output layer has one [38]. The three ANFIS models will be created and examined for four output responses in this investigation, including TOC, COD, BOD, and BOD/COD, respectively. All of the model's statistical parameters were compared, and the following equations below [39] can be used to assess the trained network's ultimate performance:

$$\text{Root mean square error (RMSE)} = \sqrt{\frac{1}{N}\sum_{i'=1}^{N}[Xp,i' - Xt,i]^2} \tag{12}$$

$$\text{Mean square error (MSE)} = \frac{1}{N}\sum_{i'=1}^{N}[(Xp,i' - Xt,i')]^2 \tag{13}$$

$$\text{Standard error (SE)} = S = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}[(Xp - \widehat{Xt})]} \tag{14}$$

3. RESULTS AND DISCUSSION

Equation 1 was employed to calculate the experimental data's percentage elimination efficiencies and biodegradability.

3.1 Effect of Combined Catalyst TiO₂ and ZnO Dosage

Figure 4 shows the combined TiO_2 and ZnO photocatalyst dose variation for the percentage elimination efficiency of reactions like TOC, COD, BOD, and biodegradability. The reactants were charged into the batch reactor, thoroughly mixed, and the photocatalyst concentration was gradually increased to 4 g/L maximum. Every hour, samples were collected, and the parameters were examined.
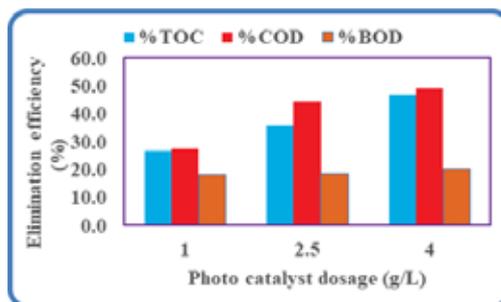
Fig. 4: Percent (%) elimination efficiency versus photocatalyst dosage.
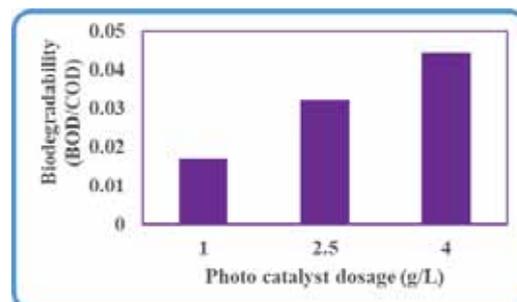
Fig. 5: Biodegradability (BOD/COD) versus photocatalyst dosage.

The parameter's percentage increases were analyzed, and the final data was plotted [3]. It was discovered that there was a significant increase in TOC, COD, and BOD up to 4 g/L of dosage of combination photocatalyst due to an increase in photonic energy availability from 10:00 AM to 2 PM. Furthermore, it was discovered that the degradation of pollutants

appears on the greater side at higher dosages of combination catalysts, which could be attributed to the availability of more active sites on the catalyst [1]. The highest percentage removal efficiency was reported to be 46.98, 49.3, and 20.23 percent for TOC, COD, and BOD at a dosage of 4 g/L, a reaction time of 300 min, and a pH of 9. Nevertheless, as shown in Fig. 5, biodegradability was found to be 0.045.

3.2 Effect of Reaction Time

Figure 6 shows the output variable's percentage removal efficiencies, depicted with reaction times. The reaction time steadily grew from 0 to 180 minutes before rapidly decreasing to 300 minutes [13]. TOC and COD increase sharply up to 180 min and decrease at 300 min [2]. The highest percentage efficiency for TOC, COD, and BOD were determined to be 60, 75.2 and 24% respectively, for reaction times of 180 minutes, dosages of 4 g/L and pH values of 9 and 0.05 for biodegradability Fig. 7.
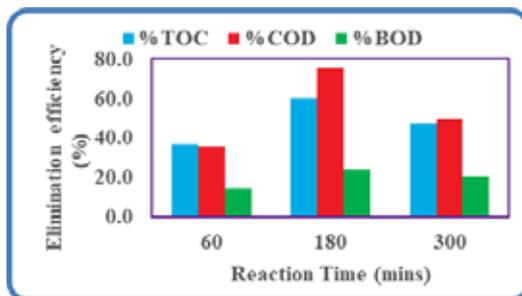
Fig. 6: Percent (%)% elimination efficiency versus reaction time.
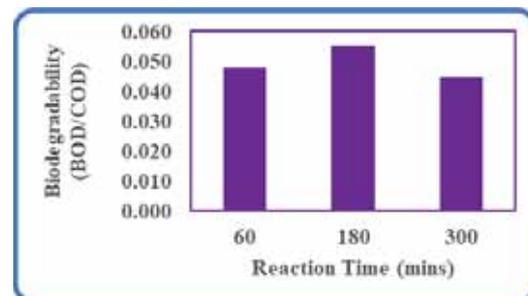
Fig. 7: Biodegradability (BOD/COD) versus reaction time.

3.3 Effect of Percentage Removal Efficiency versus pH

Figure 8 displays the output response's percentage elimination efficiencies concerning pH value at various variations [4]. It was reported that the maximum removal efficiencies were found to be TOC = 47, COD = 49.2, and BOD = 20.23 %, respectively [13]. Whereas biodegradability = 0.045 at pH = 9, reaction time = 300 minutes and dosage of combined catalyst = 4 g/L as shown in Fig. 9.
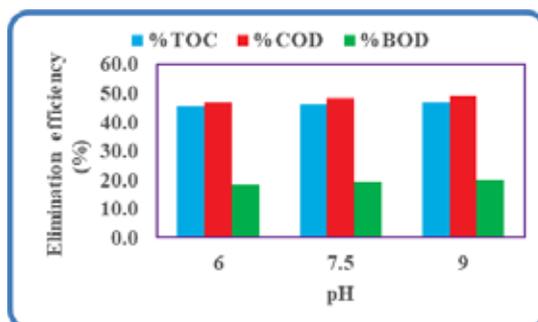
Fig. 8: % elimination efficiency versus pH.
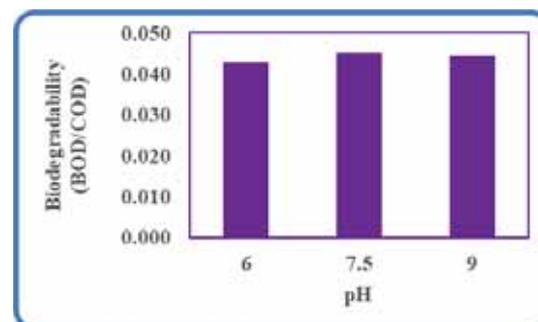
Fig. 9: Biodegradability (BOD/COD) versus pH.

3.4 RSM-Box Behnken Design Model Studies

In this method, Stat-Ease Inc.'s Design-Expert statistical software program, which is solely dedicated to doing the design of experiments (DOE), was used in the research. We

can perform comparative testing, screening, characterization, optimization, resilient parameter design, mixture designs, and integrated designs. A combination of three independent parameters T = pH, R = Dosage and S = Time, based on the Box Behnken Design model technique, was utilized. A quadratic polynomial surface methodology model statistical analysis was used according to Eq. 15 of BBD and a total of 15 experiments were found under the randomized subtype, as shown in Table 2.

$$N = 2 (K') (K' - 1) + (C_0) \tag{15}$$

where (K') = number of independent variables such as combined dosage, reaction time, and pH and Co = number of center points.

The highest and lowest variable's coded values were assigned as positive 1 and negative 1, as indicated in Table 3. These coded equations are extremely useful for identifying response factors by comparing factor coefficients.

Table 2: Experimental factors RSM-BBD method

Ver	11.1.1.0		
Type of Study	RSM-BBD	Subtype	Randomized type
Type of Design	Box-Behnken Design	Exp. Runs	15
Model	Quadratic polynomial	Blocks	No Blocks
Time (mins)	4.00		

Table 3: Independent variables employed

Factors	Independent Variable	Units	Type	Min.	Max.	Coded Low	Coded High	Mean	Std. Dev.
R	T : pH		Numeric	6	9	$-1 \leftrightarrow 6$	$+1 \leftrightarrow 9$	7.5	1.13
S	R : Dosage	(g/L)	Numeric	1	4	$-1 \leftrightarrow 1$	$+1 \leftrightarrow 4$	2.5	1.13
T	S : Time	(Minutes)	Numeric	60	300	$-1 \leftrightarrow 60$	$+1 \leftrightarrow 300$	180	90.71

The findings of the analysis of variance (ANOVA) for all the model responses, such as TOC, COD, BOD and biodegradability, are presented in Table 4-7. Using the ANOVA together with additional data like F-value, acceptable precision, coefficient of variance, probability value (Prob>F), and lack of fit, the model's validity and statistical significance ($p<0.05$) were assessed. The F-values of dependent variables TOC, COD, BOD and BOD/COD calculated were 14.2, 30.0, 43.3 and obtained P-values were 0.0047, 0.0001, 0.0003, respectively. The F-value and Prob>F values for the models and their independent parameters were significant because their p-values were less than 0.05, as indicated in Tables 4-7. Additionally, there was a 0.01 percent probability that noise could cause such a high F-value. Because they are not necessary to maintain the model hierarchy, model terms with p-values > 0.05 (not significant) were not taken into account via model reduction to improve the models [22]. The mathematical statistic model's coefficient of determination R^2 predicted R^2, adjusted R^2, and determined R^2 values for the desired responses of TOC, COD, BOD, and biodegradability were found to be 0.9623, 0.995, 0.9818 and 0.9873, respectively. According to adequate precision, the required responses were determined to be 12.80, 37.91, 17.86 and 19.95, all of which are more than 4. These values are utilized to calculate the signal-to-noise ratio and indicate that the empirical model has an appropriate signal and may be used to explore the design space. The R^2 value is more than 0.977, which is close to 1,

indicating that the model fits the data well and shows a strong correlation between the experimental and anticipated values [25].

Table 4: Analysis of Variance results for TOC elimination

Source	Sum of Squares	Degree of freedom	Avg. Square	F-value	P-value	
Design Model	3725.83	9	413.98	14.18	0.0047	significant
R: Dosage	2086.23	1	2086.23	71.44	0.0004	
S: Time	823.43	1	823.43	28.20	0.0032	
T: pH	54.19	1	54.19	1.86	0.2313	
RS	30.30	1	30.30	1.04	0.3551	
RT	0.0012	1	0.0012	0.0000	0.9951	
ST	0.0942	1	0.0942	0.0032	0.9569	
R²	3.81	1	3.81	0.1305	0.7327	
S²	661.21	1	661.21	22.64	0.0051	
T²	82.13	1	82.13	2.81	0.1544	
Residual	146.01	5	29.20			
Lack of Fit	137.34	3	45.78	10.56	0.0877	not significant
Pure Error	8.67	2	4.33			
Cor Total	3871.84	14				

Table 5: Analysis of Variance results for COD elimination

Source	Sum of Squares	Degree of freedom	Avg. Square	F-value	P-value	
Design Model	3813.46	9	423.72	121.92	< 0.0001	significant
R: Dosage	2314.12	1	2314.12	665.88	< 0.0001	
S: Time	0.0189	1	0.0189	0.0054	0.9441	
T: pH	42.76	1	42.76	12.31	0.0171	
RS	68.14	1	68.14	19.61	0.0068	
RT	0.6833	1	0.6833	0.1966	0.6760	
ST	1.16	1	1.16	0.3346	0.5880	
R²	180.73	1	180.73	52.01	0.0008	
S²	1267.51	1	1267.51	364.72	< 0.0001	
T²	2.46	1	2.46	0.7078	0.4385	
Residual	17.38	5	3.48			
Lack of Fit	16.71	3	5.57	16.71	0.0570	not significant
Pure Error	0.6667	2	0.3333			
Cor Total	3830.84	14				

Table 6: Analysis of Variance results for BOD elimination

Source	Sum of Squares	Degree of freedom	Avg. Square	F-value	P-value	
Design Model	295.92	9	32.88	29.95	0.0008	significant
R: Dosage	28.42	1	28.42	25.89	0.0038	
S: Time	84.98	1	84.98	77.42	0.0003	
T: pH	25.63	1	25.63	23.35	0.0047	
RS	1.02	1	1.02	0.9277	0.3797	
RT	0.2868	1	0.2868	0.2613	0.6310	
ST	0.0001	1	0.0001	0.0001	0.9928	
R^2	2.56	1	2.56	2.33	0.1876	
S^2	154.89	1	154.89	141.11	< 0.0001	
T^2	0.2232	1	0.2232	0.2034	0.6709	
Residual	5.49	5	1.10			
Lack of Fit	4.82	3	1.61	4.82	0.1766	not significant
Pure Error	0.6667	2	0.3333			
Cor Total	301.41	14				

Table 7: Analysis of Variance results for biodegradability

Source	Sum of Squares	Degree of freedom	Avg. Square	F-value	P-value	
Design Model	0.0021	9	0.0002	43.28	0.0003	significant
R: Dosage	0.0018	1	0.0018	338.13	< 0.0001	
S: Time	3.510E-07	1	3.510E-07	0.0656	0.8081	
T: pH	0.0000	1	0.0000	2.81	0.1542	
RS	8.959E-07	1	8.959E-07	0.1674	0.6994	
RT	6.455E-06	1	6.455E-06	1.21	0.3222	
ST	6.172E-08	1	6.172E-08	0.0115	0.9187	
R^2	0.0000	1	0.0000	2.86	0.1517	
S^2	0.0002	1	0.0002	40.62	0.0014	
T^2	0.0000	1	0.0000	7.95	0.0371	
Residual	0.0000	5	5.352E-06			
Lack of Fit	0.0000	3	3.364E-06	0.4037	0.7684	not significant
Pure Error	0.0000	2	8.333E-06			
Cor Total	0.0021	14				

To accommodate the experimental findings from the design runs carried out in connection with the configured Box Behnken Design, the response surface technique designs were either adopted or modified. Table 8 shows the responses of percentage elimination efficiencies of dependent variables. The RSM-BBD coded and actual components in terms of model equations were quadratic models [27], where R stands for dosage, S for reaction time, and T for pH level. The other terms, such as RS, RT, and TS, are the interaction terms. The square terms of dosage, reaction time, and pH were designated by R^2, S^2, and T^2, respectively. These equations can be used to anticipate responses for a given amount of each component, and coded factor equations help to determine the relative

effect factors by comparing factor coefficients. The actual coded equations, however, cannot be used for determination since the coefficients are scaled to accommodate the units of each factor rather than being at the center of the design space [28].

Table 8: Coded and Actual equations for all the responses

	Equations with coded factors	Equations with actual factors
TOC elimination (%)	$= + 40.33 + 16.15 (R) + 10.15 (S) + 2.60 (T) - 2.75 (RS) + 0.0173 (RT) - 0.1534 (ST) + 1.02 (R^2) - 13.38 (S^2) - 4.72 (T^2)$.	$= - 167.89300 + 11.20248 (R) + 0.463710 (S) + 33.31105 (T) - 0.015289 (RS) + 0.007694 (RT) - 0.000852 (ST) + 0.451526 (R^2) - 0.000929 (S^2) - 2.09612 (T^2)$
COD elimination (%)	$= + 62.33 + 17.01 (R) - 0.0486 (S) + 2.31 (T) - 4.13 (RS) - 0.4133 (RT) - 0.5392 (ST) - 7.00 (R^2) - 18.53 (S^2) - 0.8162 (T^2)$	$= - 76.83403 + 32.39084 (R) + 0.542583 (S) + 7.98120 (T) - 0.022929 (RS) - 0.183687 (RT) - 0.002996 (ST) - 3.10948 (R^2) - 0.001287 (S^2) - 0.362761 (T^2)$
BOD elimination (%)	$= + 21.33 + 1.88 (R) + 3.26 (S) + 1.79 (T) - 0.5045 (RS) - 0.2678 (RT) - 0.0050 (ST) - 0.8320 (R^2) - 6.48 (S^2) - 0.2459 (T^2)$	$= - 22.13268 + 4.50268 (R) + 0.195879 (S) + 3.12508 (T) - 0.002803 (RS) - 0.119007 (RT) + 0.000028 (ST) - 0.369756 (R^2) - 0.000450 (S^2) - 0.109285 (T^2)$
BOD/COD	$= + 0.0417 + 0.0150 (R) + 0.0002 (S) + 0.0014 (T) - 0.0005 (RS) + 0.0013 (RT) - 0.0001 (ST) - 0.0020 (R^2) - 0.0077 (S^2) - 0.0034 (T^2)$	$= - 0.089911 + 0.010788 (R) + 0.000205 (S) + 0.022264 (T) - 2.62921E-06 (RS) + 0.000565 (RT) - 6.90090E-07 (ST) - 0.000905 (R^2) - 0.5.32887E-07 (S^2) - 0.001509 (T^2)$
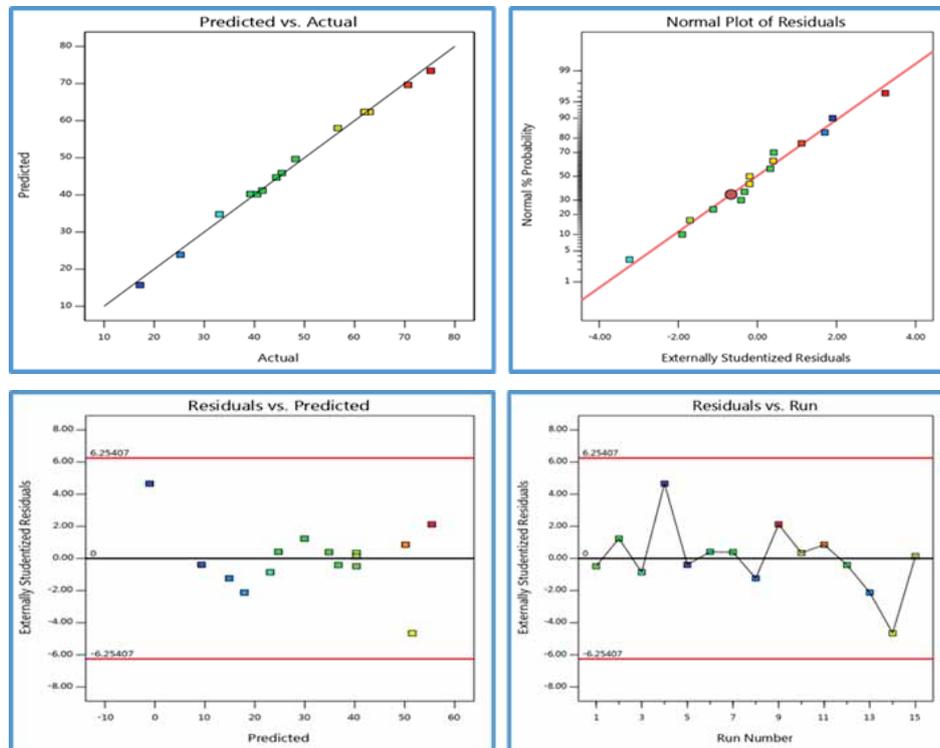
Fig. 10: RSM-BBD Predicted vs Actual and Normal plot of Residuals.

The plot of the predicted response against the internally studentized residuals in Fig. 10 provides graphic confirmation of the continuous variance assumption. The internally studentized residual values were obtained by dividing the residual values by the appropriate standard deviation. The sample points were dispersed randomly between the negative 6.25407 and positive 6.25407 outlier detection limits. Additionally, there were only slight differences between the actual and expected responses, indicating reasonable agreement between the predicted model and observed values. The described prediction model equations are thus considered adequate [29].

3.5 Interaction of 2-D Contour Plots and 3-D Surface Plots

Figure 11 depicts the cross-factor interaction effects between the independent predictors for eliminating dependent variables in photocatalytic seawater treatment using a combination using three-dimensional (3-D) surfaces and two-dimensional (2-D) contour plots. The maximum percentage elimination efficiencies as per RSM-BBD such as TOC = 55.4%, COD = 73.4%, BOD = 23.7%, and BOD/COD = 0.054 with a maximum effect of combined photocatalyst as 4 g/L, reaction time = 180 minutes and pH = 9.
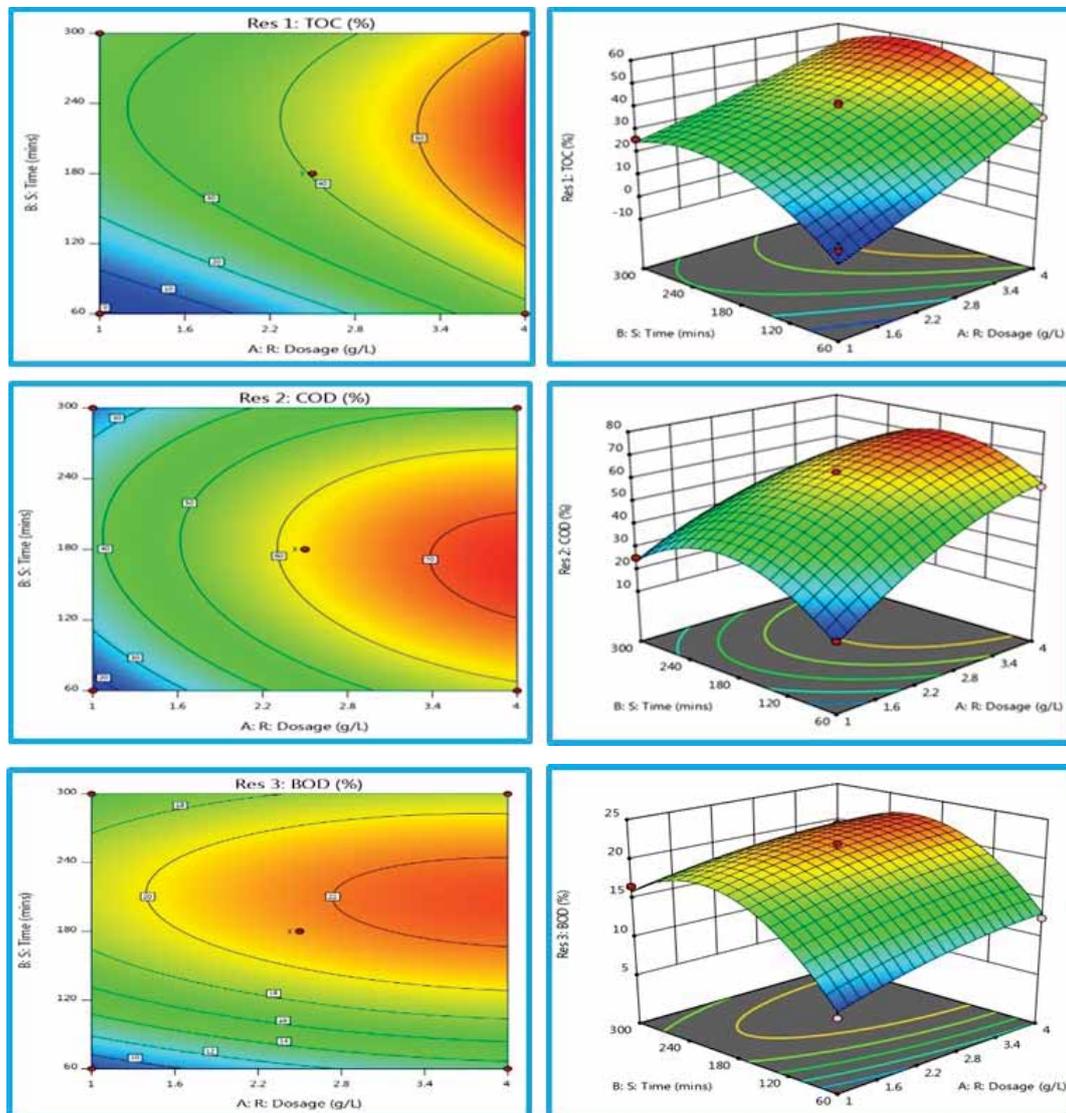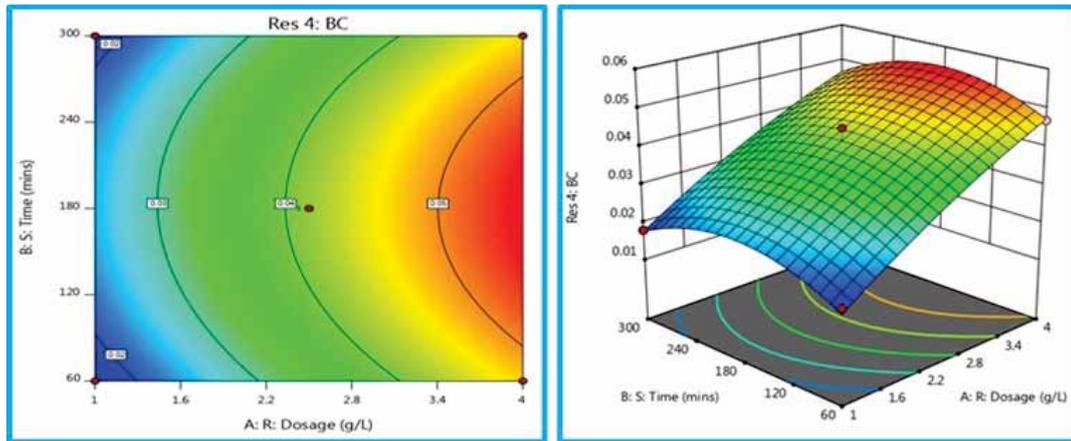
Fig. 11: 3D surface plots for all the responses.

3.6 Numerical Optimization of Model

The TiO$_2$ and ZnO combined photocatalyst load was set to 1, 2.5 and 4 g/L, the reaction duration was set to 60, 180 and 300 minutes and pH was set to 6, 7.5 and 9. These parameters were used in the RSM-BBD Design Expert software numerical optimization to maximize pollutant abatement. In addition, the output variables were set to optimum with a 95% confidence level. The desirability function approach was used, and the optimum circumstances obtained are depicted in the ramp function [32]. It is a function that uses a mathematical transformation to convert a multiple-response problem to a single-response problem. Finally, these function programs look for the option with the highest overall desirability. According to Fig. 12, the ramp function graph shows the desirability function of 0.972, 1 out of 36 solutions which are close to 1. The selected maximum values for the output variables of TOC = 58.1%, COD = 72.6%, BOD = 23.3% and biodegradability = 0.055, respectively, with an effect of combined catalyst as 4 g/L, pH = 8.266 and reaction time = 190.4 minutes.
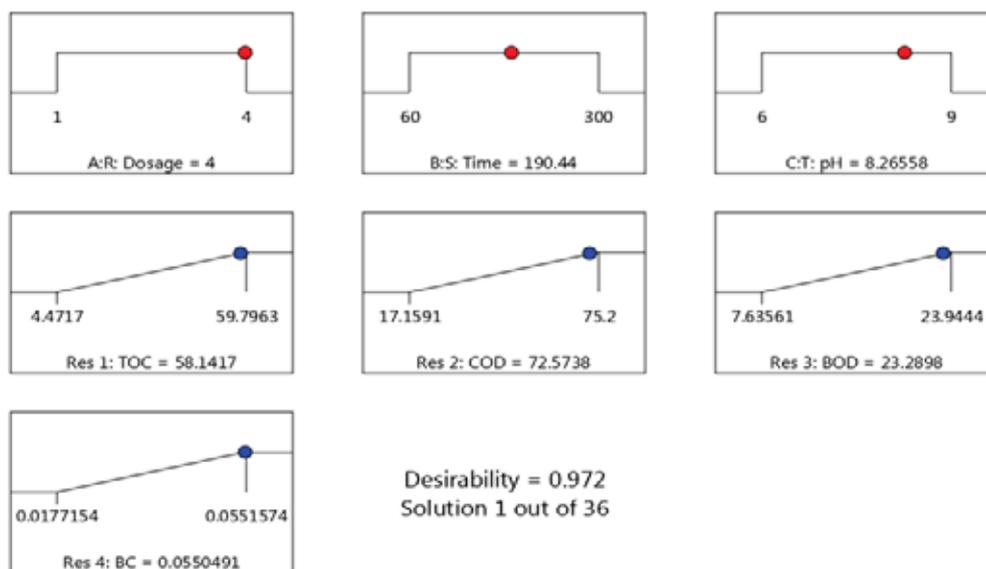
Fig. 12: RAMP function shows the desirability of the selected solution.

4. MATLAB ANFIS Based Statistical Modelling and Analysis

This section utilized the adaptive neuro-fuzzy inference system to optimize the responses. The first stage of this model is to initialize and optimize the (FIS) fuzzy inference system to represent the experimental data precisely, which can be accomplished by many procedures such as loading the experimental data sets, developing the FIS, training, and testing the (FIS) model. The ANFIS architectural network structure is depicted in Fig. 13.
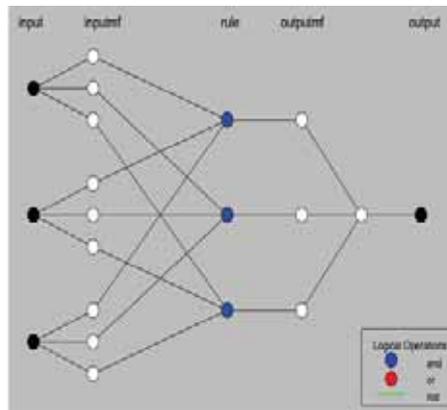
Fig. 13: Output network architecture for ANFIS model.

This ANFIS training approach incorporates the gradient descent and least square method. In this method, around 25% of the dataset was designated for testing, while 75% was used for training [35]. High correlation coefficients were used in its development for training, testing, and validation. The ANFIS model was developed with a high degree of correlation coefficient (i.e., 0.990) for the training, testing, and validation between the values. Using various membership functions for various combinations of the variables, the higher the R^2 ANFIS model was filtered. The models' effectiveness was assessed for both training and testing.

Table 9 displays the best clustering results, coefficient of determination (R^2), and root mean square error (RMSE). Figure 14 represents the RMSE analysis of training, testing, and overall datasets. It can be observed that RMSE for TOC and BOD are significantly less in comparison to COD data. However, the TOC and BOD's maximum value in the experiment set is 2.7 and 1.7. In contrast, COD ranges between 27 mg/L to 95 mg/L, and a deviation of 1.82 is minimal in corresponds to large values of COD. Figure 15 shows a correlation between the predicted and actual value in terms of R^2 value. Sample points closure to diagonal line gives a better R^2 value [37]. The simulation suggests that a small number of clustering in ANFIS modeling compared to a large number of clustering gives a better R^2 value.

Table 9: RMSE and R^2 values for the responses

Responses	No. of clusters	Training RMSE	Testing RMSE	All data RMSE		R^2 Analysis		
						Train data	Test data	All data
TOC	2	0.018	0.0374	0.023	0.998	0.999		0.998
BOD	2	0.009	0.017	0.010	0.992	0.988		0.991
COD	3	1.940	1.194	1.825	0.994	0.997		0.995
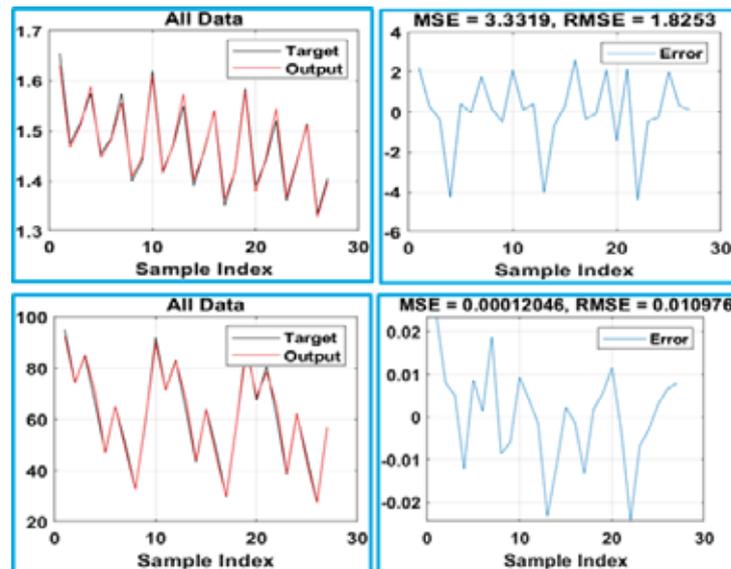
Fig. 14: Model predictions on training-testing data and its RMSE analysis.

The impact of Input parameters on a particular property can be analyzed by studying a 3D surface plot. Figure 16 shows the surface plots for three inputs, pH, dosage, and time. The experimental values or independent variables and the desired responses appeared to interact strongly, according to the surface plots [38]. Table 11 of the ANFIS statistical model shows that at a photocatalyst dose of 4 g/L, reaction time of 180 minutes and pH of 9, the maximum percentage elimination efficiency was found to be TOC=59.4%, COD=75.4%, BOD=24.2%, and biodegradability=0.055. Polynomial equation 16, using the mathematical formulation, is established to express the relationship between inputs and output [39]. Each polynomial equation correlates three inputs, i.e., dosage, time and pH, to one of the outputs, i.e., TOC, COD, BOD and BOD/COD. The polynomial equation is expressed below, where C is the coefficient, as listed in Table 10.
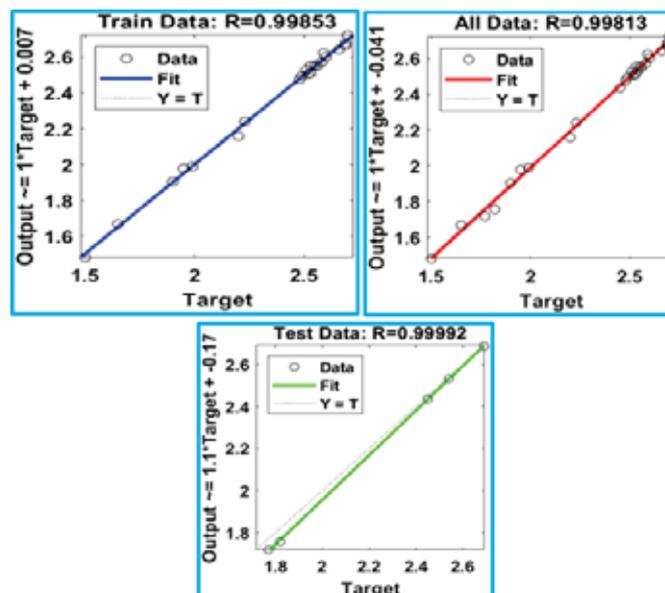
Fig. 15: Regression value for output responses.

$$O(D, t, P) = C_1 D^3 + C_2 D^2 t + C_3 D^2 P + C_4 D^2 + C_5 Dt^2 + C_6 DtP + C_7 Dt + C_8 DP^2 + C_9 DP + C_{10} D + C_{11} t^2 - C_{12} t^2 P + C_{13} t^2 + C_{14} tP^2 + C_{15} tP + C_{16} t + C_{17} P^3 + C_{18} P^2 + C_{19} P + C_{20} \qquad (16)$$
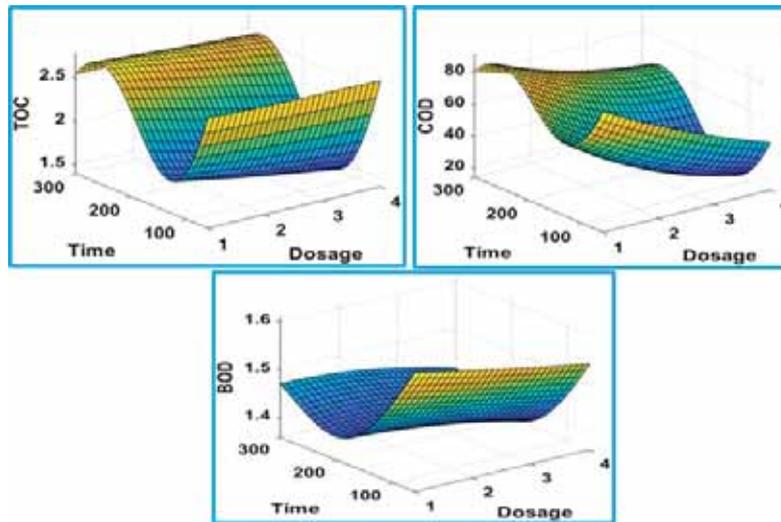
Fig. 16: 3D graph of regression for the responses.

Table 10: Coefficient value of polynomial equation expressed in the equation

Coefficient	TOC	COD	BOD
C1	-6.09154E+12	-3.54E+13	3.92837E+11
C2	0.000398538	7.92E-04	-7.58767E-05
C3	0.009997724	0.2850	0.001698481
C4	4.56866E+13	2.655E+14	-2.94628E+12
C5	9.31171E-06	0.000152428	5.78312E-08
C6	6.09955E-05	0.000706241	6.00877E-06
C7	-0.005512508	-0.038514773	0.000348171
C8	0.006051481	0.131168367	0.001380013
C9	-0.152814172	-3.4851686	-0.027966879
C10	-1.0051E+14	-5.84101E+14	6.48181E+12
C11	-682353.5565	-41651152.73	-1142667.33
C12	2.01832E-06	6.21677E-05	8.01073E-07
C13	368470920.5	22491622477	617040358.4
C14	6.72591E-05	0.000527089	1.9375E-07
C15	-0.001958112	-0.030065374	-0.000311682
C16	-56498874475	-3.44872E+12	-94612854956
C17	-3.65342E+11	9.39214E+11	68372164146
C18	8.22019E+12	-2.11323E+13	-1.53837E+12
C19	-6.08294E+13	1.56379E+14	1.1384E+13
C20	2.1109E+14	1.08568E+14	-2.79169E+13

5. CONCLUSION

The photocatalyst used in this study works well as a photocatalytic oxidizing agent for efficient seawater treatment. The experimental results showed that the optimum percentage elimination efficiency occurred at 4 g/L TiO₂ and ZnO doses, 180 minutes of reaction time

and pH 9. In addition, the biodegradability was found to be 0.055. The most significant removal efficiencies in terms of percentage were determined to be TOC=59.8, COD=75.2, BOD=23.9%, and biodegradability=0.055, respectively. According to RSM-BBD, the highest possible removal efficiencies were reported to be 55.4, 73.4, and 23.7% for TOC, COD, BOD and biodegradability as 0.054. Based on optimization criteria, a total of 36 solutions were discovered using RSM-BBD statistical modeling, and all of the response factors had a maximum desirability of 0.972 which is less than 1.0. The ANFIS was more precise with the prediction of TOC, COD, BOD, and biodegradability equal to 59.4, 75.4, 24.1, and 0.055, correspondingly with a dosage of 4 g/L, 180 mins reaction time and pH 9. ANFIS was also superior to RSM-BBD modeling's predictions, with an average R^2 value of 0.999 vs. 0.977. This research also indicated that the ANFIS model could be a helpful tool and a reliable alternative to the RSM-BBD model.

REFERENCES

[1] Feroz S, Siyabi F, Dumaran JJC. (2014) Application of solar nano photo catalysis in treatment of seawater. Int. Con. On Sust. Arch. and Environ. Eng. 1:1-4.

[2] Feroz S, Anna Jesil C. (2012) Treatment of organic pollutants by heterogeneous photo catalysis. J. of the Inst. of Eng. (India): Ser. E, 1: 45-48. https://doi.org/10.1007/s40034-012-0001-6

[3] Jabri HA, Feroz S. (2015) Effect of combining TiO_2 and ZnO in the pretreatment of seawater reverse osmosis process. Inter. J. of Environ. Sci. and Dev. 6(5): 348-351.

[4] Feroz S, Bawain M.S, Saadi S, Joy VM. (2015) Experimental studies for treatment of seawater in a recirculation batch reactor using TiO_2 P25 and polyamide. Inter. J. of app. Eng. Res., 10: 26259-26266.

[5] Feroz S, Raut NB, Maimani R. (2011) Utilization of solar energy in degrading organic pollutant – A case study. Inter. J. of COMADEM. 14(3): 33-37.

[6] Xu H, Hao Z, Feng W, Wang T, Li Y. (2021) Mechanism of photo degradation of organic pollutants in seawater by TiO_2 – based photo catalysts and improvement in their performance, American Chem. Soc., 6(45): 30698-30707. https://doi.org/10.1021/acsomega.1c04604

[7] Azevedo EB, Torres AR, Neto FRA, Dezotti M. (2009) TiO_2 – photo catalyzed degradation of phenol in saline media in an annular reactor: hydrodynamics, lumped kinetics, intermediates and acute toxicity, Brazilian J. Chem. Engr., 26(1): 75-87. https://doi.org/10.1590/S0104-66322009000100008

[8] Qiuyi J, Xiaocai Y, Jian Z, Xinyang Q. (2017) Photocatalytic degradation of diesel pollutants in seawater by using ZrO_2 (Er^{3+})/TiO_2 under visible light. J. Environ. Chem. Engr., 5(2): 1423-1428. https://doi.org/10.1016/j.jece.2017.01.011

[9] Al Deen D, Aljuboury A, Feroz S. (2021) Assessment of TiO_2/ZnO/H_2O_2 Photo catalyst to treat wastewater from oil refinery within visible light circumstances. South African J. of Chem. Eng. 35: 69-77. https://doi.org/10.1016/j.sajce.2020.11.004

[10] Nayeemuddin M, Puganeshwary P, Feroz S. (2021) Pollutants removal from saline water by solar photo catalysis: a review of experimental and theoretical approaches. Inter. J. of Environ. Anal. Chem, 1-21. https://doi.org/10.1080/03067319.2021.1924160

[11] Mashari MA, Varghese MJ, Feroz S, Rao LN. (2016) Characterization and photocatalytic treatment of oil produced water-using TiO2. Int. J. of Apl of nanotechnology, 3(1): 1-10.

[12] Araimi M, Sangeetha BM, Feroz S, Saadi S. (2016) Treatment of seawater using solar energy in a recirculation pipe reactor. Inter. J. for Inn. Res. in Sci & Tech. 2(12): 2349-6010.

[13] Jabri H, Hudaifi A, Feroz S, Marikar F, Baawain M. (2015) Investigation on the effect of TiO_2 and H_2O_2 for the treatment of inorganic carbon present in seawater, Int. J. of Eng and Sci. 5: 50-55.

[14] Wang Z, Jia Y, Liu X, Liao L, Wang Z, Zheng. (2022) Removal of boron in desalinated seawater by magnetic metal-organic frame-based composite materials: Modeling and optimizing based on methodologies of response surface and artificial neural network. J. of Mol. Liqs, 349: 118090. https://doi.org/10.1016/j.molliq.2021.118090

[15] Hashemi SH, Kaykhaii M, Keikha AJ, Sajjadi Z. (2018) Application of Box-Behnken design in response surface methodology for the molecularly imprinted polymer pipette-tip solid phase extraction of methyl red from seawater samples and its determination by spectrophotometry. Mar. Pol. Bul. 137: 306-314. https://doi.org/10.1016/j.marpolbul.2018.10.037

[16] Xu H, Hao Z, Feng W, Wang T, Fu X. (2022) The floating photo catalytic spheres loaded with weak light-driven TiO_2 –based catalysts for photodegrading tetracycline in seawater, 144: 106610. https://doi.org/10.1016/j.mssp.2022.106610

[17] Kahani M, kalantary F, Soudi M.R, Pakdel L, Aghalizadeh S. (2020) Optimization of cost effective culture medium for sporosarcina pasteurii as biocementing agent using response surface methodology: Up cycling dairy waste and seawater. J. of Clean. Prod. 253: 120022. https://doi.org/10.1016/j.jclepro.2020.120022

[18] Altayb HN, Kouidhi B, Baothman AS, Abdulhakim AJ, Ayed L, Hager M, Chaieb K. (2021) Mathematical modeling and optimization by the application of full factorial design and response surface methodology approach for decolourization of dyes by a newly isolated Photo bacterium ganghwense. J. of Wat. Proc. Eng. 44: 102429 https://doi.org/10.1016/j.jwpe.2021.102429

[19] Vatanpour V, Sheydaei M, Esmaeili M. (2017) Box-Behnken design as a systematic approach to inspect correlation between synthesis conditions and desalination performance of TFC RO membranes. Desal, 420: 1-11. https://doi.org/10.1016/j.desal.2017.06.022

[20] Peng X, Yang G. Shi Y, Zhou Y, Zhang M, Li S. (2020) Box-Behnken design based statistical modeling for the extraction and physicochemical properties of pectin from sunflower heads and the comparison with commercial low-methoxyl pectin. Sci. Reports. 10: 3595. https://doi.org/10.1038/s41598-020-60339-1

[21] Mi H, Yi L, Wu Q, Xia J, Zhang B. (2021) Preparation and optimization of a low-cost adsorbent for heavy metal ions from red mud using fraction factorial design and Box-Behnken response methodology. Cols. and Sur. A: Phy. and Eng. Asp. 627:127198. https://doi.org/10.1016/j.colsurfa.2021.127198

[22] Jaafarzadeh N, Zarghi MH, Salehin M, Roudbari A Zahedi A. (2020) Application of Box-Behnken Design (BBD) to optimizing COD removal from fresh leachate using combination of ultrasound and ultraviolet. J. Environ. Treat. Tech., 8(3): 861-869.

[23] Hout S, Salem Z, Tassalit D, Tigrine Z, Aburideh H. (2020) Assessing desalination pretreatment conditions towards pilot scale-up using Box-Behnken experimental design. Wat. And Environ. J. Prom. Sus. sol. 35(2): 473-485. https://doi.org/10.1111/wej.12644

[24] Jamei M, Karbsi M, Malik A, Abualigah L, Islam ART, Yaseen ZM. (2022) Computational assessment of groundwater salinity distribution within coastal multi-aquifers of Bangladesh. Sci. Rep. 12:11165. https://doi.org/10.1038/s41598-022-15104-x

[25] Boutra B, Sebti A, Trari M. (2022) Response surface methodology and artificial neural network for optimization and modeling the photo degradation of organic pollutants in water. Inter. J. of Environ. Sci and Tech. 19: 11263-11278. https://doi.org/10.1007/s13762-021-03875-1

[26] Nayeemuddin M, Palaniandy P, Feroz S. (2022) Solar photocatalytic biodegradability of saline water: Optimization using RSM and ANN. AIP Conf. Proc. 2463, 1-11. https://doi.org/10.1063/5.0080297

[27] Joy VM, Feroz S, Dutta S. (2020) TiO_2 / Photo-Fenton process for seawater pretreatment: modelling and optimization using response surface methodology (RSM) and artificial neural networks (ANN) coupled genetic algorithm (GA). J. Indian Chem. Soc., 2:1:10. https//:doi.org/10.5281/zenodo.5657210

[28] Joy VM, Feroz S, Dutta S. (2021) Solar nano photocatalytic pretreatment of seawater: process optimization and performance evaluation using response surface methodology and genetic

algorithm. App. Water Sci. 11:18. https://doi.org/10.1007/s13201-020-01353-6

[29] Nayeemuddin M, Puganeshwary P, Feroz S. (2021) Optimization of solar photocatalytic biodegradability of seawater using statistical modelling. J. the Indian Chem. Soc. 98(12): 100240. https://doi.org/10.1016/j.jics.2021.100240

[30] Miao T, Huang H, Guo J, Li G, Zhang Y, Chen N. (2022) Uncertainty analysis of numerical simulation of seawater intrusion using deep learning-based surrogate model. Wat. 14(18): 2933. https://doi.org/10.3390/w14182933

[31] Kim CM, Parnichkun M. (2017) Prediction of settled water turbidity and optimal coagulant dosage in drinking water treatment plant using a hybrid model of k-means clustering and adaptive neuro-fuzzy inference system. Appl. Water Sci., 7: 3885-3902. https://doi.org/10.1007/s13201-017-0541-5

[32] Khayet M, Cojocaru C, Essalhi M. (2011) Artificial neural network modelling and response surface methodology of desalination by reverse osmosis. J. of Mem. Sci. 368: 202-214. https://doi.org/10.1016/j.memsci.2010.11.030

[33] Jia Y, Li J, Wang Z, Wu X, Xu K, Wang Z. (2020) Adsorption performance and response surface optimization of boron in desalinated seawater by UiO-66-NH$_2$. Des. and Wat. Treat. 197:157-169. https://doi.org/10.5004/dwt.2020.25999

[34] Sibiya NP, Amo-Duodu G, Kweinor Tetteh E, Rathilal S. (2022) Response surface optimization of a magnetic coagulation process for wastewater treatment via Box-Behnken. Mat. Today: Proc, 62: S122-S126, https://doi.org/10.1016/j.matpr.2022.02.098

[35] Mahadeva R, Kumar M, Manik G, Patole SP. (2021) Modeling, simulation, and optimization of the membrane performance of seawater reverse osmosis desalination plant using neural network and fuzzy based soft computing techniques. Des. and Wat. Treat. 229:17-30. https://doi.org/10.5004/dwt.2021.27386

[36] Ehteram M, Ahmed AN, Kumar P, Sherif M, Shafie AE. (2021) Predicting freshwater production and energy consumption in a seawater greenhouse based on ensemble frameworks using optimized multi-layer preceptron. Energy Reports. 7: 6308-6326. https://doi.org/10.1016/j.egyr.2021.09.079

[37] Najah A, El-Shafie A, Karim OA, El-Shafie AH. (2013) Performance of ANFIS versus MLP-NN dissolved oxygen prediction models in water quality monitoring. Environ. Sci. Pollut. Res. 21:1658-1670. https://doi.org/10.1007/s11356-013-2048-4

[38] Jain A, Toussi IB, Mohammadian A, Bonakdari H, Sartaj M. (2022) Applications of ANFIS-type methods in simulation of systems in marine environments. Math. Comput. Appl. 27(29): 1-25. https://doi.org/10.3390/mca27020029

[39] Mojarradi V, Sahraei S. (2020) Prediction of RO membrane performances by use of adaptive network-based fuzzy interference systems. J. of Chem. and Petr. Eng. 54(1): 99-110. https://doi.org/10.22059/jchpe.2020.292454.1300

Table 11. Shows the experimental, RSM and ANFIS values

Run	Input factors			TOC elimination (%)			COD elimination (%)			BOD removal (%)			BOD/COD		
	R : Dos (mg/L)	S : RT (min)	T : pH	Expt val.	BBD Pred. val.	Anfis Pred. val.	Expt val.	BBD Pred. val.	Anfis Pred val.	Expt val.	BBD Pred. val.	Anfis Pred. val.	Expt val.	BBD Pred. val.	Anfis Pred. val.
1	2.5	180	7.5	38.000	40.330	39.580	62.000	62.330	62.310	21.000	21.330	21.300	0.040	0.042	0.039
2	2.5	300	6	33.104	29.930	33.519	41.589	41.170	41.264	15.677	16.070	15.143	0.030	0.030	0.028
3	1	180	9	20.692	23.070	22.900	39.223	40.240	37.962	20.771	20.430	20.520	0.021	0.021	0.020
4	1	60	7.5	4.472	4.300	4.519	17.159	15.720	18.929	7.636	8.380	7.917	0.018	0.016	0.018
5	2.5	60	6	8.178	9.330	7.900	40.523	40.190	36.404	10.589	9.570	8.345	0.029	0.029	0.025
6	1	300	7.5	25.938	24.720	24.348	25.232	23.880	25.428	16.583	15.900	15.562	0.018	0.018	0.018
7	2.5	300	9	35.986	34.830	38.310	44.377	44.710	43.840	18.642	19.660	18.610	0.032	0.032	0.023
8	2.5	60	9	11.673	12.850	12.240	45.468	45.890	41.253	13.534	13.140	11.869	0.031	0.032	0.024
9	4	180	9	59.796	55.400	59.386	75.200	73.430	75.392	23.944	23.660	24.138	0.055	0.054	0.055
10	2.5	180	7.5	42.000	40.330	35.580	62.000	62.330	60.496	21.000	21.330	20.864	0.040	0.042	0.032
11	4	180	6	52.539	50.160	53.740	70.646	69.630	70.672	20.275	20.620	18.948	0.049	0.049	0.044
12	4	60	7.5	35.501	36.720	35.454	56.641	57.990	58.731	12.475	13.150	12.297	0.047	0.047	0.034
13	1	180	6	13.504	17.900	15.005	33.015	34.790	33.271	16.031	16.310	15.390	0.020	0.021	0.020
14	4	300	7.5	45.959	51.510	45.282	48.205	49.640	47.776	19.405	18.660	18.191	0.045	0.047	0.025
15	2.5	180	7.5	41.000	40.330	38.580	63.000	62.330	61.900	22.000	21.330	20.864	0.045	0.042	0.032

EFFECTS OF HIVE SIZES AND MESH MATERIALS ON THE PRODUCTIVITY OF PROPOLIS AND HONEY PRODUCED BY *TETRAGONULA LAEVICEPS*

MUHAMMAD YUSUF ABDUH[1,2*], ANNISA SHABRINA[1], ANDREAS RADEN CAMAN[1],
ARSY ELIA PRATIWI[1] AND MUHAMMAD IHSANU[1]

[1]*School of Life Sciences and Technology, Institut Teknologi Bandung, Bandung, Indonesia*
[2]*University Center of Excellence for Nutraceuticals,*
Bioscience and Biotechnology Research Center,
Institut Teknologi Bandung, Bandung, Indonesia

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

ABSTRACT: Propolis and honey are bioproducts produced by *Tetragonula laeviceps*, a stingless bee species. Both products have anti-inflammatory, antimicrobial, and antibacterial properties. Modular Tetragonula hives of different sizes were used in the cultivation of *T. laeviceps* to increase the productivity of propolis while maintaining the sustainability of the bee colonies. This study was carried out in Jatinangor, West Java, Indonesia, with three size variations: small (21 × 18 × 14 cm), medium (26 × 22 ×17 cm), and large (30 × 26 × 17 cm). Each hive was equipped with a mesh of a different material. The results obtained in this study showed that increasing the hive's size increases the productivity of propolis. The highest productivity of propolis (2.53 ± 0.37 g/colony/week) was obtained when *T. laeviceps* was cultivated in large hives equipped with a nylon-based mesh. The highest productivity of honey (0.78 ± 0.18 g/colony/week) was obtained when *T. laeviceps* was cultivated in medium hives equipped with an aluminum-based mesh. The harvested propolis was extracted using a maceration method. Total flavonoid and phenolic content of the propolis solution lies in the range of 1.77 ± 0.86 to 3.18 ± 1.43 mg QE/g propolis and 32.23 ± 14.09 to 112.13 ± 47.64 mg GAE/g propolis, respectively. The harvested honey had a water content of 21.86%, with 72.86% reducing sugar, 192.86 µg/mL vitamin C, and 2613.41 µg/mL of antioxidant content.

ABSTRAK: Propolis dan madu adalah produk bio yang dihasilkan oleh *Tetragonula laeviceps*, spesies lebah kelulut. Kedua-dua produk mempunyai ciri anti-radang, anti-mikrob dan antibakteria. Sarang Tetragonula modular dengan saiz yang berbeza digunakan dalam mengusahakan *T. laeviceps* bagi meningkatkan produktiviti propolis sambil mengekalkan kemampanan koloni lebah. Kajian ini dijalankan di Jatinangor, Jawa Barat, Indonesia, dengan tiga variasi saiz: kecil (21 × 18 × 14 cm), sederhana (26 × 22 × 17 cm), dan besar (30 × 26 ×17 cm). Setiap sarang dilengkapi dengan jaringan bahan yang berbeza. Dapatan kajian yang diperolehi melalui kajian ini menunjukkan bahawa penambahan saiz sarang dapat meningkatkan produktiviti propolis. Produktiviti tertinggi propolis (2.53 ± 0.37 g/koloni/minggu) diperoleh apabila *T. laeviceps* dibela dalam sarang besar yang dilengkapi dengan jaring berasaskan nilon. Produktiviti madu tertinggi (0.78 ± 0.18 g/koloni/minggu) diperoleh apabila *T. laeviceps* diusahakan dalam sarang sederhana yang dilengkapi dengan jaring berasaskan aluminium. Propolis yang dituai telah diekstrak menggunakan kaedah maserasi. Jumlah kandungan flavonoid dan kandungan fenolik larutan propolis berada dalam julat 1.77 ± 0.86 hingga 3.18 ± 1.43 mg QE/g propolis dan 32.23 ± 14.09 hingga 112.13 ± 47.64 mg GAE/g propolis, masing-masing. Madu yang dituai mempunyai kandungan air 21.86%, dengan 72.86% penurunan gula, 192.86 µg/mL vitamin C, dan 2613.41 µg/mL kandungan antioksidan.

1. INTRODUCTION

Bees produce various beneficial products such as honey, pollen, propolis, beeswax, and bee bread. Bees make propolis by mixing plant resin, which contains flavonoids, polyphenols, saponins, amino acids, and minerals, with wax and their saliva secretion [1]. Propolis consumption can reduce inflammation, inhibit scar tissue contraction, promote wound healing, and enhance stamina and overall health. Antibacterial and antifungal properties of propolis allow for its extensive use in the health and cosmetic industries [2]. Moreover, bees also produce honey from flower nectars, which comprise flavonoids and other phenolic compounds, contributing to antimicrobial, antioxidant, anti-inflammation, and anticancer activities [3].

Tetragonula laeviceps is a stingless bee species that produces propolis to seal cracks in their hive, to protect their hive from pests and diseases, and to build honey and pollen pots. *T. laeviceps* can produce propolis more abundantly compared to honeybees of genus *Apis* with productivities of 3 kg/colony/year and 20–30 g/colony/year, respectively [4]. Total flavonoid content in stingless bee propolis with quercetin equivalent (QE) reaches 4% (mg QE/g propolis extract), higher than *Apis* bee with only 1.5% [5].

However, honey productivity of *T. laeviceps* is relatively small compared with its propolis only, around 6.5 kg/year/colony [6]. The physicochemical properties in *T. laeviceps* honey include a pH of 3.15–4.66, acidity of 5.9–109 meq/kg, anti-microbial activity of 0.9–23 DN, 58–75.7 g/100 g of reducing sugar, 1.1–4.8 g/100g of sucrose, and water content (19.9–41.9 g/100 g) [7]. Therefore, *T. laeviceps* is widely cultivated by stingless beekeepers for its propolis.

There are several stingless beekeepers still beekeeping with bamboo, the bees' natural hive. Beekeeping with bamboo hives yields propolis at a rate of 2.36 g/colony/week [8]. However, the application of bamboo hives requires beekeepers to damage the hive during the harvesting process, which harms the colony and results in high contamination of pollen and honey in the propolis [9]. The hive design that is currently under development is a wooden box equipped with a mesh as a propolis trap.

A Modular *Tetragonula* Hive (MOTIVE) is a stingless beehive invention that combines a main wooden box and a perforated mesh to optimize the productivity of propolis [8,10,11]. The mesh in the hive contributes to an increase in propolis production compared to a simple wooden box without a mesh [8,10]. A MOTIVE can also sustain the colony and improve the quality of propolis harvested because they do not mix with non-propolis products [10]. MOTIVE beekeeping has been accomplished with MOTIVEs the size of 21×14×18 cm which resulted in propolis productivity of 2–2.8 g/colony/week [8,10]. A MOTIVE is constructed of three main parts: the wooden box for the colony, a mesh holder comprised of two wooden frames above the wooden box, and the top cover that is legged to create a ventilation space above the mesh [8,10].

The mesh in the MOTIVE is a perforated material that serves as a place for the bees to attach propolis. The mesh requires properties such as rust resistance, acid resistance, odorless, and processability [12]. Aluminum wire, nylon, and polyethylene are mesh materials that are widely used. Nylon membranes have high pH, high-temperature resistance, and a small pore size distribution [13]. Polyethylene membranes have a melting point of 115 °C and are flexible, whereas aluminum materials have high-temperature- and corrosion-resistant properties [14].

In this study, *T. laeviceps* cultivation was carried out using MOTIVEs with various hive sizes and mesh materials. According to Abd-Elmawgood et al. [15], the smaller hive results in a higher temperature inside the hive. Internal conditions, such as temperature, will affect the condition and activity of the colony in the foraging process. In addition, the size of the hive, which is proportional to the mesh size, will affect the propolis production by the bees. Therefore, this study was conducted to determine the size of MOTIVE and the mesh material that generates optimum productivity of propolis and honey.

2. MATERIALS AND METHODS

2.1 Preparations of Colony

Twenty *Tetragonula laeviceps* colonies were obtained from a breeder in Kebumen, Central Java, Indonesia, in September–November 2018. The *T. laeviceps* colonies were in bamboo hives and acclimatized at the cultivation site for one to two months. The MOTIVE was sterilized with ultraviolet (UV) light for 30 minutes. Eighteen colonies were moved to the MOTIVE, and two colonies remained in the bamboo hive as control treatments. The colony inside the MOTIVE was re-acclimatized for two weeks.

2.2 Variation of Hive Size and Mesh Type

The MOTIVE used had three size variations: small (21 × 18 ×14 cm), medium (26 × 22 × 17 cm), and large (30 × 26 × 17 cm). Each hive size had three variations of mesh materials: nylon, polyethylene, and aluminum wire placed between the wooden box and the cover. The pore size of the nylon mesh was 1 × 1 mm, while the pore size of the polyethylene and aluminum wire mesh was 2 × 2 mm.

2.3 Harvesting of Propolis and Honey

The beekeeping process of *T. laeviceps* was carried out for eight weeks at the Wood Laboratory, Institut Teknologi Bandung, Jatinangor Campus, Sumedang, West Java. Propolis was harvested every two weeks, while honey was harvested at the end of the cultivation period. The honey harvesting process was accomplished using a syringe and glass pipette.

The mesh that was filled with propolis was weighed to determine the propolis weight. Propolis productivity of *T. laeviceps* per week can then be determined by dividing the propolis weight by two (for the two-weeks period). The harvested honey was then measured for its weight and volume. Honey productivity per week was determined by dividing the weight of the honey by eight (for the eight-weeks period).

2.4 Extraction of Propolis

The propolis was extracted according to the method of Machado et al. [16] with a slight modification. Harvested propolis was separated from the mesh using liquid nitrogen. The ratio of propolis to ethanol was 1:15 (w/v). The extraction was carried out on a shaker incubator (IKA KS 4000i Control) with a rotational speed of 500 rpm and a temperature of 40 °C for one hour in a dark room. The extract was filtered with a glass funnel and filter paper Whatman no. 2. The filtrate obtained was stored in a dark vial bottle. Then, 0.5 mL of the propolis filtrate sample was transferred to evaporating dishes to evaporate the solvent using a hot plate at 40 °C. During this process, the evaporating dish was covered with aluminum foil to reduce air and light exposure.

2.5 Determination of Total Flavonoid Content

Total flavonoid content (TFC) was determined using a quercetin calibration curve, as described by Machado et al. [16]. Standard quercetin (0.05 g) was dissolved in 50 mL ethanol to obtain 1000 µg/mL of quercetin stock solution. The stock solution was then diluted to obtain quercetin solutions with concentrations of 100 µg/mL, 50 µg/mL, 25 µg/mL, 12.5 µg/mL, 6.25 µg/mL, and 0 µg/mL, for the blank solution. 2 mL of standard solution was mixed with 2 mL of 2% $AlCl_3$ solution and incubated for 30 minutes. Next, the absorbance was measured at a wavelength of 415 nm with an ultraviolet-visible (UV-VIS) spectrophotometer (Shimadzu UV-1800). Before analyzing the sample, propolis extract from the evaporation process was dissolved in 10 mL 80% methanol. The propolis-methanol solution was added to 2% $AlCl_3$, and the absorbance was measured. Quercetin contents in the sample solution were considered as TFC in units of mg of quercetin equivalent per mL (mg QE/mL), determined by the following equation:

$$TFC_{sample\ solution} = \left(\frac{A - 0.0296}{0.035}\right) \times N, \tag{1}$$

where A is the absorbance value of the sample, N is the dilution factor of the sample solution, and 0.0296 and 0.035 are the constant value and the coefficient of the quercetin calibration curve, respectively. Meanwhile, the TFC in propolis extract was determined by the following equation:

$$TFC_{propolis\ extract}\ (\text{mg QE/g}) = \frac{TFC_{sample}\left(\frac{mg}{mL}\right) \times V_{solution}\ (mL)}{m_{extract}\ (g)}, \tag{2}$$

where $V_{solution}$ (10 mL) is the total volume of propolis methanol extract solution, and $m_{extract}$ is the mass (g) of propolis extract (mass of remaining solids from the evaporation of propolis filtrate).

2.6 Determination of Total Phenolic Content

The total phenolic content (TPC) of propolis was determined according to the Folin-Ciocalteu method, as explained by Machado et al. [16]. A calibration curve was constructed by first dissolving 0.1 g of standard gallic acid in 10 mL ethanol and then diluting with 100 mL distilled water to obtain a stock solution with a concentration of 1000 µg/mL. The stock solution was diluted with distilled water to obtain standard solutions with concentrations of 750 µg/mL, 500 µg/mL, 250 µg/mL, 100 µg/mL, 50 µg/mL, and 0 µg/mL, for the blank solution. 0.5 mL of each standard solution was added to 2 mL of Folin-Ciocalteu reagent and 2.5 mL of 7.5% Na_2CO_3 solution. The mixture was then incubated for five minutes in a water bath at 50 °C. The absorbance of each standard solution was measured with a UV-VIS spectrophotometer at a wavelength of 765 nm. Sample analysis was performed by dissolving evaporated propolis extract in 5 mL ethanol 80%. Gallic acid contents in the sample solution were considered as TPC with a unit of mg of gallic acid equivalent per mL (mg GAE/mL). TPC in the sample was determined with the following equation:

$$TPC_{sample\ solutione}\ (\text{mg/mL}) = \left(\frac{A - 0,1351}{0,0111}\right) \times N \tag{3}$$

where A is the absorbance value of the sample, N is the dilution factor of the sample solution, and 0.1351 and 0.0111 are the constant value and the coefficient of the gallic acid calibration curve, respectively. Meanwhile, the TPC in the propolis extract was determined with Eq. (2), with a $V_{solution}$ of 5 mL.

2.7 Characterization of Honey

Honey characterization was carried out at Sibaweh Laboratories, Bandung. The moisture content of the honey was measured according to Indonesian National Standard 01-3545-2004 (19- SNI, 2004) using a refractometer. Reducing sugar content was determined with the Luff-Schoorl method according to the Indonesian National Standard 01-2892-1994 (20-SNI, 1994). Antioxidant activity was determined with the DPPH (2,2-diphenyl-1-picrylhydrazyl) activity method as suggested by Sadeli et al. [17]. Vitamin C content was measured using a titration method according to Silva et al. [18].

3. RESULTS AND DISCUSSIONS

3.1 Productivity of Propolis Produced by *T. laeviceps*

T. laeviceps colonies that were cultivated in bamboo hives produced propolis at a rate of 0.67 ± 0.01 g/colony/week. The propolis productivity of the bees in small, medium, and large MOTIVEs with nylon, polyethylene, and aluminum meshes ranged from 1.12–2.52 g/colony/week (Table 1). *T. laeviceps* bees produce propolis to cover the holes in the hive for self-protection [19]. A MOTIVE is equipped with a perforated mesh to stimulate more propolis production by the bees. The productivity of propolis of the colonies in bamboo hives was lower than the colonies in the MOTIVEs. This phenomenon is related to the difference in cavity size in the hives. Bamboo hives have smaller cavity gaps, which extend vertically, resulting in less propolis produced and difficulty in harvesting.

Table 1: Productivity of propolis and honey (g/colony/week) produced by *T. laeviceps* in various hive sizes and mesh materials

Hive sizes	Mesh materials	Propolis productivity [g/colony/week]	Honey productivity [g/colony/week]
Small	Nylon	1.24 ± 0.21^a	0.09 ± 0.02^c
	Polyethylene	1.12 ± 0.58^a	0.08 ± 0.01^c
	Aluminum	1.12 ± 0.54^a	0.02 ± 0.01^c
Medium	Nylon	$1.51 \pm 0.63^{a,b}$	0.09 ± 0.07^d
	Polyethylene	$1.60 \pm 0.27^{a,b}$	0.12 ± 0.17^d
	Aluminum	$1.81 \pm 0.80^{a,b}$	0.78 ± 0.18^d
Large	Nylon	2.52 ± 0.37^b	$0.12 \pm 0.02^{c,d}$
	Polyethylene	2.32 ± 0.39^b	$0.35 \pm 0.35^{c,d}$
	Aluminum	1.31 ± 0.50^b	$0.03 \pm 0.01^{c,d}$

*Different letters indicate statistical differences (P-value < 0.05, Duncan's test).

The mesh sizes in this study were proportional to the MOTIVE size, such that the large MOTIVE has a larger mesh surface area than the small MOTIVE. Thus, larger MOTIVE results in higher propolis productivity as the bees need to cover more areas. This phenomenon was shown through the propolis productivity in the large MOTIVEs with nylon and polyethylene meshes, which were higher than that of the small and medium MOTIVE. However, the MOTIVE with an aluminum mesh did not show the same pattern. Propolis productivity in the large MOTIVE with an aluminum mesh was lower than that of the medium MOTIVE.

Propolis production by the bees was influenced by several factors, such as the colony condition during the beekeeping process [20]. In this study, the number of bees per colony varied with bee colonies. The large MOTIVEs with nylon, polyethylene, and aluminum meshes had 859, 1120, and 937 bees, respectively. These population numbers were enough for the bees to cover the mesh with propolis, as shown by the highest productivity of propolis in the large MOTIVEs. Similarly, the medium MOTIVE with nylon and aluminum meshes also had enough populations with 774 and 1788 bees, respectively. Interestingly, the medium MOTIVE with a polyethylene mesh only had 189 bees, but the colony could still meet the needs of propolis production to cover the mesh. This phenomenon is related to the task divisions of worker bees, which are determined based on anatomical and physical conditions, environmental stimuli, and age [21]. The bee colony was considered young, signified by the abundant brood cells and bee breads in the hive. Thus, the high needs of the colony affected the activity of the forager bees in collecting resin [22]. In this study, the age of colony and number of bees per colony were not pre-determined prior to the study but were calculated after the harvesting process, which resulted in different numbers of bees in the colonies in each treatment.

In the research conducted by Agussalim et al. [23], the highest productivity of propolis was produced by the bees in a wooden box hive sized $35 \times 20 \times 17.5$ cm with a value of 3.85 ± 3.66 g/colony/week. This value was higher than the colony from a wooden box hive sized $40 \times 20 \times 20$ cm with a productivity rate of 2.30 ± 1.82 g/colony/week. These results were similar to this study, where the productivity of propolis in the MOTIVEs with aluminum meshes was not linearly proportional to the size of the hive. This could be due to factors such as the number of worker bees, flight activities of the forager bees, and sources and availability of resins.

Table 1 shows that propolis productivity was more affected by hive size because propolis productivity differences were not significant in mesh material variation. The highest productivity of propolis was from the colonies in the large MOTIVE with a nylon mesh at 2.52 ± 0.37 g/colony/week. According to Tsakgarakis et al. [24], propolis production is higher with smaller mesh pore size, which has a larger total surface area. Pore sizes of the aluminum, nylon, and polyethylene meshes were 2×2 mm, 1×1 mm, and 2×2 mm, respectively. The small pore size of the nylon mesh results in more holes to cover but allows for easier propolis placement by the bees.

3.2 Productivity of Honey Produced by *T. laeviceps*

In addition to propolis, *T. laeviceps* bees produce honey at a rate of 0.02–0.78 g/colony/week. The honey tends to have a dark brown color with a distinctly sour taste. The results show that honey productivity is directly proportional to the size of MOTIVE with a polyethylene mesh, but not for the MOTIVE with nylon and aluminum meshes. According to Bushal et al. [25], honey productivity is not affected by the size of the hive but is affected by the size of the colony, the number of populations in the colony, and the bees' flight activity. Honey production tends to increase when the population increases owing to the growing need for bees to forage, which results in more forager bees. Based on the research conducted by Salatnaya [26], honey productivity using a $25 \times 15 \times 15$ cm wooden box hive was 0.02–0.19 g/colony/week, less than the honey productivity of small MOTIVEs.

In addition to the colony population, honey productivity is influenced by nectar flow, or the availability of nectar in the environment [27]. The bees will collect nectar in warm weather for supply during the rainy season. However, honey production for stingless bees, in general, is not as high as the *Apis* honeybees owing to their smaller body sizes, which results in shorter flight distances [28]. The maximum flying distance of *Tetragonula*

stingless bees is approximately 712 m [29]. However, its small body prevents the stingless bee from collecting nectar and pollen on small flowers [30]. Several factors that can cause low honey production are the focus of worker bees on hive construction, waste disposal activity from the hive, and low productivity of the queen bee to produce worker bees [31].

In this study, honey productivity was directly proportional to propolis productivity in MOTIVEs with polyethylene and aluminum meshes. This was because large amounts of honey can provide energy for the bees to forage other resources, such as pollen and resin. Stingless bees, which are eusocial organisms, are interdependent to fulfil their needs and maintain hive conditions. These activities improved the sustainability of the colony, and the colony remained healthy and protected from pests and diseases.

3.3 Total Flavonoid and Total Phenolic Content of Propolis Extract

The harvested raw propolis was extracted. The resulting propolis extracts had a wide range of TFC and TPC. Propolis from the bamboo hive had a TFC and TPC of 3.42 ± 1.15 mg QE/g and TPC 284.62 ± 39.46 mg of GAE/g, respectively. The complete TFC and TPC of each MOTIVE size and mesh material are shown in Table 2. The propolis extraction process will remove the inert materials and keep the polyphenol fraction [32]. The results of this study indicate that there is no correlation between hive size and propolis quality. The difference in propolis quality is caused by numerous factors, such as bee genetics (species), different sources of plant resins, and microclimates [26]. This is supported by the research of Hakim & Abduh [8] and Abduh et al. [10], in which the colony of *T. laeviceps* in a MOTIVE sized $21 \times 14 \times 18$ cm that was cultivated in the Cileunyi Wetan Village had a TFC of 6.64–11.42 mg QE/g extract, whereas a MOTIVE in Cibodas Village had a TFC amounting to 14.8 mg QE/g extract.

The extraction process of propolis also affects the dissolved flavonoids and phenolics content. The antioxidant content and activity of the propolis extract depend on the extraction method and the type of solvents used [33]. Propolis extracted by the maceration method for three days resulted in TFC and TPC of 52.17 mg QE/g extract and 26.13 mg GAE/g extract, respectively, whereas five days of maceration resulted in TFC and TPC of 68.04 mg QE/g extract and 46.68 mg QE/g extract, respectively [34]. These results indicate that the duration of the maceration process can affect the solubility and the yield of flavonoid and phenolic compounds associated with the mass transfer process of the compounds into the solvent. The solid-to-solvent ratio, temperature, and the maceration speed can also affect the TFC and TPC.

TFC and TPC in the propolis from the bamboo hive were also measured with the result of 3.42 ± 1.15 mg QE/g extract and 284.62 ± 39.46 mg GAE/g extract. TPCs in the propolis from MOTIVEs were lower than TPC in the propolis from the bamboo hive. The position of propolis in the hive may play a role in the chemical composition of propolis. According to Pratami et al. [35], propolis located inside the hive had a higher TFC and TPC compared with propolis located outside the hive, which was exposed to the external environment. Propolis in the MOTIVEs was more exposed to the external environment because of the ventilation space between the mesh and the top cover compared with propolis in the bamboo hive. Therefore, the difference in propolis position in the hive may be a factor that caused differences in the TPC of the propolis. Exposure to light, air, and high temperatures can change the structure of phenolic compounds affecting their degradation process [35].

Moreover, the chemical composition of propolis is also influenced by the source of the plant resins. During the beekeeping process, the bees produce propolis from plant resins obtained near the beekeeping sites, such as mahogany (*Swietenia macrophylla*), velvet apple

(*Diospyros blancoi*), gmelina (*Gmelina arborea*), guanacaste (*Enterolobium cyclocarpum*), and kassod tree (*Senna siamea*). The colors of propolis produced varied, light to dark brown, red, as well as greenish yellow. Color differences in the propolis indicate that the chemical composition and the active compounds in each propolis might also be different [10].

Table 2: Total flavonoid and total phenolic content of propolis produced by *T. laeviceps* in various hive sizes and mesh materials

Hive sizes	Mesh materials	Total Flavonoid Content (TFC) [mg QE/g extract]	Total Phenolic Content (TPC) [mg GAE/g extract]
Small	Nylon	3.18 ± 1.43^a	112.13 ± 47.64^b
	Polyethylene	2.60 ± 1.66^a	32.23 ± 14.09^b
	Aluminum	2.46 ± 1.21^a	80.45 ± 53.93^b
Medium	Nylon	1.95 ± 1.02^a	39.10 ± 15.91^b
	Polyethylene	1.78 ± 1.02^a	39.12 ± 23.35^b
	Aluminum	3.02 ± 1.50^a	54.58 ± 26.76^b
Large	Nylon	1.77 ± 0.86^a	77.61 ± 34.61^b
	Polyethylene	2.84 ± 2.01^a	52.44 ± 25.01^b
	Aluminum	2.84 ± 1.00^a	71.41 ± 27.30^b

*Different letters indicate statistical differences (P-value < 0.05, Duncan's test).

3.4 Characteristics of Honey Produced by *T. laeviceps*

The honey harvested from *T. laeviceps* was measured for its moisture, reducing sugars, antioxidants, and vitamin C contents. Table 3 shows that the moisture content, reducing sugars, and active antioxidants of *T. laeviceps* honey from MOTIVEs were lower than in the honey from the bamboo hive. However, the vitamin C content of honey from the MOTIVEs was higher than in honey from the bamboo hive.

High moisture content can reduce the quality of honey because the presence of water can accelerate the rate of yeast growth so that fermentation occurs. The type of yeast that causes fermentation in honey is *Zygosaccharomyces* [36]. The higher moisture content in the honey from the bamboo hive could be caused by the differences in the bamboo hive and MOTIVE structures. Bamboo hives are narrower and have less ventilation space and, hence, higher humidity compared to the MOTIVEs. Honey is also hygroscopic, able to absorb water from the environment. Therefore, the moisture content in honey could be influenced by environmental humidity [37].

According to Souza et al. [7], the moisture content in *Apis* bee honey is only 17%, lower than in honey from *Tetragonula* bees, which is 24%. The difference is caused by the more open structure of honey storage pots in *Apis* bees, so the rate of water evaporation in *Apis* honey will be faster than in *Tetragonula* honey. Honey from the MOTIVEs met the maximum moisture content requirement of 22% outlined by the Indonesian National Standard SNI 01-3545-2004 [38]. Yeast in honey will degrade sugars, such as glucose and fructose, into alcohol and CO_2. This will increase the reducing sugar content, which is the total amount of fructose and glucose in honey [37]. Table 3 shows that the reducing sugars in honey from the bamboo hive was higher than in honey from the MOTIVEs. This shows the correlation between moisture content and reducing sugar content in honey. Bees from the Meliponini family, such as *T. laeviceps*, generally have a reducing sugar content of 66% (w/w) with a sucrose content of 2.3% [9]. Based on the standard quality of honey from the

Indonesian National Standard SNI 01-3545-2004 [38], the minimum of total reducing sugar is 65%, and thus, honey from the bamboo hive and MOTIVEs still meets this requirement.

Table 3: Characteristics of honey produced by *T. laeviceps*

Parameters	Bamboo hive	MOTIVE
Moisture content [%]	24.08	21.86
Reducing sugar [%]	74.70	72.84
Active antioxidant/IC_{50} [µg/mL]	3544.88	2613.41
Vitamin C [µg/mL]	184.50	192.86

Honey contains many antioxidant compounds such as catalase, glucose oxidase, peroxidase, and phenolic compounds. The active antioxidant of *T. laeviceps* honey from a bamboo hive is higher than in honey from a MOTIVE. However, both honeys have higher antioxidant content compared to honey researched by Abduh et al. [10]. The antioxidant content of honey cultivated from this study in Cibodas and Cileunyi Wetan, West Java, were 1188.2 µg/mL and 1,341.88 µg/mL, respectively. The lower active antioxidant content in honey from this study could be caused by the location of cultivation, sources of nectar, and the age of honey.

Vitamin C (ascorbic acid) is one of the numerous compounds present in honey. The vitamin C content in honey from MOTIVEs is higher than in honey from bamboo hives, which are 192.86 µg/mL and 184.5 µg/mL, respectively. Different sources of nectar could cause the difference in vitamin C content [7]. Therefore, both the content of vitamin C and antioxidants in honey is strongly influenced by the plant nectar sources around the cultivation site.

4. CONCLUSION

In brief, increasing the hive's size increases the productivity of propolis. The highest productivity of propolis (2.53 ± 0.37 g/colony/week) was obtained when *T. laeviceps* was cultivated in large hives equipped with a nylon-based mesh. The highest productivity of honey (0.78 ± 0.18 g/colony/week) was obtained when *T. laeviceps* was cultivated in medium hives equipped with an aluminum-based mesh. The harvested propolis was extracted using a maceration method to determine the total flavonoid and phenolic content of the propolis which lies in the range of 1.77 ± 0.86 to 3.18 ± 1.43 mg QE/g propolis and 32.23 ± 14.09 to 112.13 ± 47.64 mg GAE/g propolis, respectively. Relevant properties of the harvested honey were also characterized and shown to meet the standard quality of honey. The result of this study provides an insight for larger scale cultivation of *T. laeviceps* to produce higher productivity of propolis from stingless bees.

ACKNOWLEDGMENT

REFERENCES

[1] Nakajima Y, Tsuruma K, Shimazawa M, Mishima S, Hara H. (2009) Comparison of bee products based on assays of antioxidant capacities. BMC Complementary and Alternative Medicine, 9: 1-9.
[2] Silva CR, Baltazar F, Almeida AC. (2015) Propolis: A complex natural product with a plethora

of biological activities that can be explored for drug development. Evidence-Based Complementary and Alternative Medicine, 1-29. https://doi.org/10.1155/2015/206439

[3] Rao P.V, Krishnan KT, Salleh N, Gan SH. (2016) Biological and therapeutic effects of honey produced by honey bees and stingless bees: a comparative review. Revista Brasileira de Farmacognosia, 26: 657-664.

[4] Borror DJ, Triplehorn LA, Johnson NF. (1989) An introduction to the study of insects. 6th edition. Saunders College Publishing, Philadelphia.

[5] Fadhilah R, Rizkika K. (2015) Laba Lebah Tanpa Sengat, My Trubus Potential Business. Trubus Swadaya, Jakarta.

[6] Singh S, Ketan V, Saurabh S, Jain N, Tawale P. (2018) Propolis market research report - forecast to 2023. Market Research Future.

[7] Souza BA, Roubik DW, Barth OM, Heard TA, Enríquez E, Carvalho CA, Villas-Bôas J, Marchini L, Locatelli J, Persano-Oddo L, Almeida-Muradian L, Bogdanov S, Vit P. (2006) Composition of Stingless Bee Honey: Setting Quality Standards. Interciencia, 31: 867-875.

[8] Hakim MN, Abduh MY. (2019) Produksi propolis dari lebah *Tetragonula laeviceps* menggunakan sarang MOTIVE yang dilengkapi dengan sistem instrumentasi. Jurnal Otomasi Kontrol dan Instrumentasi, 10: 133-148.

[9] Hilmi M, Bradbear N, Mejia D. (2011) Beekeeping and sustainable livelihood, Food and Agriculture Organization of the United Nations, Rome.

[10] Abduh MY, Adam A, Fadhlullah M, Putra RE, Manurung R. (2020) Production of propolis and honey from Tetragonula laeviceps cultivated in Modular Tetragonula Hives. Heliyon, 6: 1-8.

[11] Adam A, Hakim MN, Oktaviani L, Inderaja BM, Manurung, R, Putra RE, Abduh MY. (2020) Techno-economic evaluation for integrated cultivation of coffee and stingless bees in West Java, Indonesia. Biological And Natural Resources Engineering Journal, 3: 28-36.

[12] Budiaman, Rahman A. (2006) Uji efektivitas empat variasi propolis trap terhadap produksi propolis lebah madu *Apis mellifera* L. Jurnal Perennial, 2: 1-4.

[13] Huang L, Bui N, Meyering MT, Hamlin TJ, Mccutcheon J. (2013) Novel hydrophilic nylon 6,6 microfiltration membrane supported thin film composite membranes for engineered osmosis. J. Memb. Sci., 437: 141-149.

[14] Surdia T, Saito S. (1995) Pengetahuan bahan teknik. Pradnya Paramita, Jakarta, Indonesia.

[15] Abd-Elmawgood BH, Al-Rajhi MA, El-Ashhab AO. (2015) Effect of the internal size and thermal insulation of the hive on bee colonies strength and productivity. Egyptian Journal of Agricultural Research, 93: 185-195.

[16] Machado BA, Silva RP, Barreto GD, Costa SS. Silva DF, Brandao HN, Rocha JLC, Dellagostin OA, Henriques JAP, Umsza-Gues MA, Padilha FF. (2016) Chemical composition and biological activity of extracts obtained by supercritical extraction and ethanolic extraction of brown, green and red Propolis derived from different geographic regions in Brazil. PLoS ONE, 11: 1-26.

[17] Sadeli RA. (2016) Uji Aktivitas Antioksidan dengan Metode DPPH (1,1-diphenyl-2-picrylhydrazyl) Ekstrak Bromelain Buah Nanas (*Ananas comosus* (L.) Merr.). Faculty of Pharmacy, Universitas Sanata Dharma, Yogyakarta.

[18] Silva CR, Simoni JA, Collins CH, Volpe PL. (1999) Ascorbic acid as a standard for iodometric titrations: an analytical experiment for general chemistry. Journal of Chemical Education, 76: 1421-1422.

[19] Sanpa S, Popova M, Bankova V, Tunkasiri T, Eitssayeam S, Chantawannakul P. (2015) Antibacterial compounds from propolis of *Tetragonula laeviceps* and *Tetrigona melanoleuca* (Hymenoptera: Apidae) from Thailand. PLoS ONE, 10: 1-11.

[20] Bhusal S. Thapa R. (2006) Response of colony strength to honey production: Regression and correlation analysis. Journal of the Institute of Agriculture and Animal Science, 27: 133-137.

[21] Akratanakul, P. (1990) Beekeeping in Asia, Food and Agriculture Organisation of the United Nation, Rome.

[22] Leonhardt SD, Dworschak K, Eltz T, Bluthgen N. (2007) Foraging loads of stingless bees and utilisation of stored nectar for pollen harvesting. Apidologie, 38: 125-135.

[23] Agussalim AA, Umami N, Erwan E. (2015) Production of stingless bees (*Trigona* sp.) propolis in various bee hives design. The 6th International Seminar on Tropical Animal Production, 20-22 October 2015, Universitas Gadjah Mada.

[24] Tsagkarakis AE, Katsikogianni T, Gardikis K, Katsenios I, Spanidi E, Balotis GN. (2017) Comparison of traps collecting propolis by honey bees. Advances in Entomology, 5: 68-74.

[25] Bushal SJ, Kafle L, Thapa RB, Shih CJ. (2011) Effect of Colony Strength on the Performance of Honeybees (*Apis mellifera*) in Nepal (Hymenoptera:Apidae). Sociobiology, 58: 435-447.

[26] Salatnaya, H. (2012) Produktivitas lebah Trigona sp. sebagai penghasil propolis perkebunan pala monokultur dan polikultir di Jawa Barat, Institut Pertanian Bogor, Bogor.

[27] Jager AJ, Taylor GJ, Greeff P, Lishman, AW. (2001) The effect of commercial propolis production on hive profitability. Proceedings of The 37th International Apicultural Congress.

[28] Faheem M, Aslam M, Razaq M. (2004) Pollination Ecology with Special to Insects: A Review, J. Res. (Sci), 15: 359-409.

[29] Smith, J. P., Heard, T. A., Beekman, M., Gloag. R., 2016. Flight range of the Australian stingless bee *Tetragonula carbonaria* (Hymenoptera: Apidae): Flight range of the bee *Tetragonula carbonaria*. Austral Entomology, 56: 1-4.

[30] Jalil AM, Kasmuri AR, Hadi H. (2017) Stingless Bee Honey, The Natural Wound Healer: A Review. Skin Pharmacology and Physiology, 30: 66-75.

[31] Agussalim AA, Umami N, Budiasatra IG. (2017) The Effect of Daily Activity Stingless Bees of Trigona sp. on Honey Production. The 7th International Seminar on Tropical Animal Production, 12-14 September, Universitas Gadjah Mada; pp. 223-227.

[32] Paviani LC, Fiorito G, Sacoda P, Cabral FA. (2013) Different Solvents For Extraction of Brazilian Green Propolis: Composition and Extraction Yield of Phenolic Compounds, III Iberoamerican Conference on Supercritical Fluids. Colombia.

[33] Narimane S, Demircan E, Salah A, Ozcelik B, Salah R. (2017) Correlation between antioxidant activity and phenolic acids profile and content of Algerian propolis: Influence of solvent. Pakistan Journal Pharmaceutical Sciences, 30: 1417-1423.

[34] Zin NB, Azemin A, Rodi MM, Mohd KS. (2018) Chemical Composition and Antioxidant Activity of Stingless Bee Propolis from Different Extraction Methods. International Journal of Engineering & Technology, 7: 90-95.

[35] Pratami DK, Mun'im A, Sundowo A, Sahlan M. 2018. Phytochemical profile and antioxidant activity of propolis ethanolic extract from *Tetragonula* Bee. Pharmacognosy Journal, 10: 128-135.

[36] Salas, PG, Soto AM, Carretero AM, Gutierrez AF. (2010) Phenolic-Compound-Extraction Systems for Fruit and Vegetable Samples. Molecules, 15: 8813-8826.

[37] Nanda PB, Radiati LE, Rosyidi D. (2015) Perbedaan Kadar Air, Glukosa, dan Fruktosa pada Madu Karet dan Madu Sonokeling, Universitas Brawijaya, Malang.

[38] Standar Nasional Indonesia (SNI). (2004) Madu (SNI 01-3545-2004). Badan Standardisasi Nasional, Jakarta.

PARAMETRIC OPTIMIZATION OF OIL PALM MESOCARP FIBER VALORIZATION WITH HYBRID OZONATION-ULTRASONIC PRETREATMENT METHOD

DIDI DWI ANGGORO, INDAHSARI KUSUMA DEWI*, **LUQMAN BUCHORI AND AJI PRASETYANINGRUM**

[1]*Department of Chemical Engineering, Faculty of Engineering, Universitas Diponegoro, Jl. Prof. H. Soedarto S.H., Semarang 50275, Indonesia*

*Corresponding author: ndah.dewi17@gmail.com

ABSTRACT: Oil palm mesocarp fiber is a promising lignocellulosic biomass as a raw material for valorizing biomass into more valuable products such as second-generation biofuels, biocomposites, or bioenergy. However, the lignin composition present in lignocellulosic biomass provides resistance to the valorization process and protects the cellulose composition, thereby limiting the conversion of cellulose into more valuable products. The hybrid ozonation-ultrasonic method as a lignin-degrading method is starting to be considered an effective method. Additionally, a Box-Behnken Design (BBD) was employed to investigate each independent variable's effect on pretreatment process conditions using the response surface methodology (RSM), namely reaction time (30-90) min, reaction temperature (20 -40) $^{\circ}$C and ozone flow rate (1-3) L/min to the response of the percentage of lignin degradation (%). The optimum condition of the pretreatment process is determined using the desirability function graph. The results showed that reaction time, reaction temperature, and ozone flow rate had a significant effect on lignin degradation (p <0.05). The optimum conditions obtained the highest percentage of lignin degradation, namely 92.08% at a reaction temperature of 30 $^{\circ}$C with an ozone flow rate of 2 L/min for 60 minutes reaction time. The decrease in lignin absorption peaks at 1638 cm^{-1} and 1427 cm^{-1} was supported by the results of the analysis of increased crystallinity in the sample after the pretreatment of lignin degradation to 80.20% and was validated by changes in the morphology of the mesocarp fiber after the pretreatment process indicating that the lignin compound had been successfully degraded from cellulose products of mesocarp fibers.

ABSTRAK: Sabut gentian kelapa sawit berpotensi sebagai bahan mentah biojisim lignoselulosa bagi menambah nilai produk biojisim seperti bahan bio api generasi kedua, biokomposit atau biotenaga. Walau bagaimanapun, komposisi lignin yang wujud dalam biojisim lignoselulosa menentang proses tambah nilai dan melindungi komposisi selulosa, dengan itu mengehadkan penukaran selulosa kepada produk yang lebih berharga. Kaedah hibrid ozonasi-ultrasonik sebagai kaedah merendahkan lignin, mula mendapat perhatian sebagai kaedah berkesan. Selain itu, Reka Bentuk Kotak-Behnken (BBD) telah digunakan bagi menyiasat setiap kesan pembolehubah bebas pada keadaan proses prarawatan menggunakan kaedah permukaan tindak balas (RSM), iaitu masa tindak balas (30-90) min, suhu tindak balas (20 -40) $^{\circ}$C dan kadar aliran ozon (1-3) L/min terhadap tindak balas pada peratusan degradasi lignin (%). Keadaan optimum bagi proses prarawatan ditentukan menggunakan graf fungsi keboleh inginan. Dapatan kajian menunjukkan bahawa masa tindak balas, suhu tindak balas, dan kadar aliran ozon mempunyai kesan yang signifikan terhadap degradasi lignin (p<0.05). Keadaan optimum peratusan degradasi lignin tertinggi adalah pada 92.08% pada suhu tindak balas 30 $^{\circ}$C dengan kadar aliran ozon 2 L/min selama 60 minit masa tindak balas. Penurunan puncak penyerapan lignin pada 1638 cm^{-1} dan 1427 cm^{-1} disokong

oleh keputusan analisis peningkatan kehabluran sampel selepas prarawatan degradasi lignin sebanyak 80.20% dan telah disahkan oleh perubahan morfologi sabut gentian selepas proses prarawatan menunjukkan bahawa sebatian lignin telah berjaya didegradasi daripada produk selulosa sabut gentian.

KEY WORDS: *Hybrid ozonation-ultrasonic lignin degradation, oil palm mesocarp fiber, RSM, AOP$_S$ pretreatment*

1. INTRODUCTION

Mesocarp fiber is one of the by-products produced in large quantities by the palm oil industry in the palm oil production process. Total palm oil production continues to increase yearly due to the increasing world market demand for vegetable oil. In 2020, the total world palm oil production reached 73 million tonnes per year, then it will increase in the following year 2021 to reach 75 million tonnes per year, and now the total worldwide palm oil production has reached 77 to 79 million tonnes per year [1,2,3]. Then, from the total amount of palm oil production, it produces around 12-15% by-products in the form of mesocarp fiber which, if not handled seriously, can become a problem in waste management and endanger the surrounding environment [4]. In line with the problems that can be caused, oil palm mesocarp fiber is also one of the interesting potential lignocellulosic biomasses because of its abundant availability at a low cost. It has a high main component of cellulose which is around 40.12% by weight, hemicellulose 20.12% by weight, and 30.33% lignin by weight, so if utilized properly, it can be promising in terms of the economy [5,6,7,8].

The valorization of oil palm mesocarp fiber biomass, which entails converting lignocellulosic biomass into second-generation biofuels or other valuable products such as biocomposites, has become a current trend. However, the complex structure of lignocellulosic biomass has received special attention because of the lignin component that hinders the accessibility of cellulose, thereby inhibiting the valorization of lignocellulosic biomass [9]. Therefore, the pretreatment process is an important step before valorizing lignocellulosic biomass. The pretreatment process's primary goal is to degrade the complex structure of lignocellulosic biomass into simpler fibrous components and remove some of the lignin components in what is commonly referred to as the biomass delignification process [10]. Ozone-based advanced oxidation technology is one of the oxidation methods that has recently gained popularity for degrading complex compounds in waste and biomass [11]. In degrading complex compounds effectively and optimally, advanced ozone-based oxidation technology is based on the use of radical generation such as the hydroxyl radical (OH^*) as an oxidizing agent which has a strong oxidation potential value and is non-selective so that it shows good potential in the treatment of lignocellulosic biomass [12,13].

Ozone-based advanced oxidation processes (AOPs) can degrade complex compounds by two mechanisms that depend on the pH of different solutions, namely directly and indirectly. The indirect oxidation of ozone involves the hydroxyl radical (OH^*) as an oxidizing agent, while the direct oxidation of ozone represents oxidation with ozone molecules [14]. Both ozone-based advanced oxidation processes rely on different pH solutions to react with complex compounds. Ozone oxidation directly oxidizes complex compounds when the pH of the solution is acidic (< 5), while indirect ozone oxidation can take place when the pH of the solution is alkaline (> 8) [15]. The oxidation potential of ozone is 2.07 eV, which makes it a selective and strong oxidizer [11,14]. The ozonation process as a pretreatment method for lignocellulosic biomass has several advantages, such as a strong oxidation potential value and a characteristic of ozone molecules that they react readily with conjugated double bonds found

in complex compounds and functional groups with high density, which makes ozone able to oxidize lignin which has a high C=C bond [16]. In addition, ozone can be easily decomposed into compounds with greater oxidizing abilities, such as hydroxyl radicals (OH*), which have a potential oxidation value of 2.80 eV and are also non-selective so that they can easily be better in degrading complex compounds of lignocellulosic biomass [11].

Ozonation is also a green pretreatment method because, in the process of oxidizing lignocellulosic biomass, it does not produce side products that can harm the surrounding environment. Ozone can decompose into oxygen molecules, so it is safe if directly discharged into the surrounding environment. The filtrate produced from the lignin degradation process can be utilized for lignin recovery [17]. Praptyana and Budiyono [18] reported that the pretreatment of mahogany sawdust biomass using ozone in alkaline conditions could degrade lignin and increase biohydrogen production up to 86.9% over other pretreatments. Similar results were also reported by Hassan et al. [19], that ozone pre-treatment of green algae Ulva lactuca could degrade complex biomass compounds, thereby significantly increasing biogas production. In addition, ozone has a few drawbacks that limit its industrial use, which is related to the insufficient solubility and stability of ozone molecules in solution and their limited selectivity in the oxidation process resulting in the low mass transfer of ozone in solution [20,21].

In recent years, researchers have become interested in finding ways to improve ozone mass transfer. The ultrasonic process provides several advantages compared to other technologies in increasing the efficiency of ozone-based advanced oxidation processes (AOPs), which include a simple operational process that can increase the solubility of ozone and does not produce harmful pollutants [22]. Several studies have found that ultrasonic cavitation has the ability to increase the mass transfer of ozone in water and also increase the production of free radicals such as hydroxyl radicals (OH*). This is because the cavitation phenomenon in the ultrasonic process can cause the formation, growth, and collapse of acoustic bubbles, which can produce great energy with very high temperatures (> 4000 K) and local pressure (> 500 bar) as a result of the hotspot phenomenon that is formed when the rupture cavitation bubbles to increase the surface area of contact between ozone and solution and can cause dissociation of water molecules which leads to an increase in the mass transfer of ozone and an increase in the formation of radical compounds such as hydroxyl radicals (OH*) [23]. The hybrid of ozonation and ultrasonic methods can accelerate the oxidation process of lignin by hydroxyl radical compounds (OH*) into oxyaromatics compounds, which then turn into short-chain aliphatic acids. Additionally, the collapse of cavitation bubbles in the ultrasonic process also produces hydromechanical shear forces in solution which can interfere with Surface morphology of lignocellulosic biomass [24]. The synergistic effect between the ozonation and ultrasonic processes can be described as follows [25].

Shen et al. [26] investigated the effectiveness of using ozonation and ultrasonic processes simultaneously to provide a significant degradation effect of X-3B reactive red dye up to 99.2%. In addition, Weavers and Hoffmann [27] also reported that the mass transfer of ozone was much larger at a frequency of 20 kHz, which resulted in a mass transfer coefficient value ($K_L a$ 0.28 – 0.41 min^{-1}), while at a frequency of 500 kHz, the value of the displacement coefficient was mass is only equal to ($K_L a$ 0.05 – 0.1 minutes^{-1}). Therefore, the combination of ozonation and ultrasonic processes is an approach that has great potential and promise. The synergistic effect of the combination of the two processes increases the efficiency of the degradation of complex compounds such as lignocellulosic biomass. In line with the synergistic effect of the hybrid ozonation-ultrasonic method, process variables also affect the effectiveness of the lignocellulosic biomass pretreatment process. Optimization of the hybrid ozonation-

ultrasonic process variables using statistical and mathematical approaches such as the Response Surface Methodology (RSM) method in the biomass pretreatment process have not been reported so far.
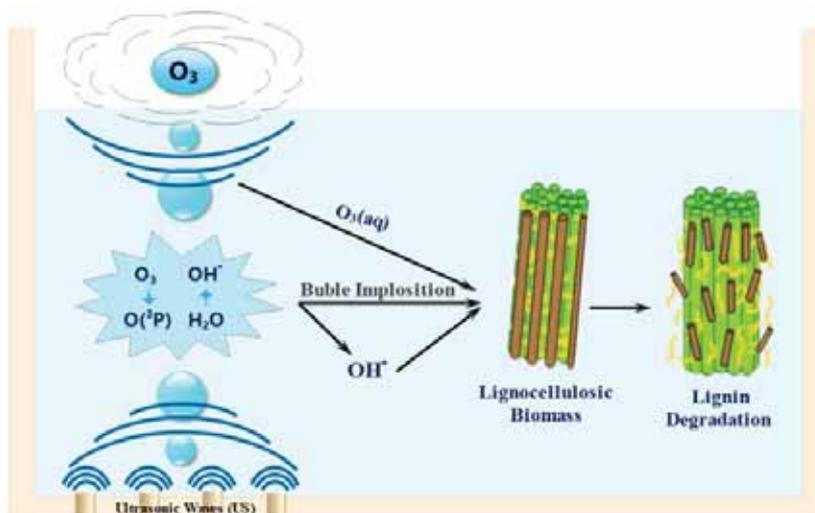
Fig. 1. Synergism of hybrid ozonation-ultrasonic pretreatment

Thus, it is necessary to optimize the process variables to provide information about the effect of the pretreatment process variables on the response, namely lignin degradation, and to save processing time and pretreatment costs [5]. This study applied a hybrid ozonation-ultrasonic process as a pretreatment method for lignocellulosic biomass in oil palm mesocarp fibers. Three different operational variables, operating temperature, operating time, and ozone flow rate were evaluated using a three-factor box-behnken design (BBD) approach, which is one type of design of experiment (DOE) in the response surface methodology (RSM) method to determine the most influential variable on the response and to obtain optimal operating conditions in the biomass pretreatment process.

2. EXPERIMENT

2.1. Material and Methods

The oil palm mesocarp fibers in this study were obtained from Indonesian palm oil companies. Mesocarp fiber raw materials were washed with water to remove particulates and dried in an oven at 105 ºC for 24 hours. Then the size was made uniform at 60 mesh and stored in a closed container at room temperature. The chemicals used in this study were H_2SO_4 with a purity of 95% - 97% (E. Merck Cat No. 100731), NaOH with a purity of 99% (E. Merck Cat. No. 106498), ozone gas obtained from an ozone generator (Dipo Technology Indonesia) and ultrasonic irradiation was carried out using an ultrasonic bath type KLS 303365 equipped with a thermostatic water bath with a frequency of 42 kHz.

2.2. Oil Palm Mesocarp Fiber (OPMF) Characterization

The characterization of the composition of cellulose and hemicellulose in oil palm mesocarp fibers was determined by the cellulose isolation and bleaching methods that were slightly modified from Candido [28]. While the characterization of lignin composition in oil palm mesocarp fibers was determined using the Chesson Datta method, namely by weighing 1 g of dry sample (weight a), then reacting with 150 mL of distilled water for 1 hour at a temperature of 100 ºC. The resulting residue was then washed with hot aquadest, dried in an

oven for 24 hours, and weighed to a constant (weight b). Then the solid was reacted again with 150 mL H_2SO_4 1N for 1 hour at 100 °C, filtered and washed with distilled water, dried in the oven for 24 hours at 100 °C, and then weighed to a constant (weight c). The weighed solid was then reacted again with 10 mL of 72% H_2SO_4 for 4 hours at room temperature, then added to 150 mL of 1N H_2SO_4 and the mixture was refluxed for 1 hour. The residue was washed, dried, and weighed to a constant weight (weight d). Following that, the final solid is subjected to dry ashing using a muffle furnace at 525 °C, and the dry ash resulting from that process is weighed until constant (weight e). The percentage composition of cellulose, hemicellulose and lignin can be calculated by equations (1 - 4). Changes in the mesocarp fiber surface morphology were confirmed using Scanning Electron Microscopy (SEM). In contrast, the crystallinity value of the OPMF before pretreatment and after pretreatment was confirmed by X-Ray Diffraction (XRD) rays, and the change in the chemical structure after the pretreatment process was confirmed by using Fourier-Transform Infrared Spectroscopy (FTIR).

$$Cellulose \quad (\%) \quad = \left[\frac{c-d}{a} \right] \times 100\% \tag{1}$$

$$Hemicellulose \quad (\%) \quad = \left[\frac{b-c}{a} \right] \times 100\% \tag{2}$$

$$Lignin \quad (\%) \quad = \left[\frac{d-e}{a} \right] \times 100\% \tag{3}$$

$$Lignin\ Degradation(\%) \quad = \left[\frac{lignin\ before\ pretreatment - lignin\ after\ pretreatment}{lignin\ before\ pretreatment} \right] \times 100\% \tag{4}$$

In which *a* represents the initial dry weight (g) of the oil palm fiber sample, *b* represents the dry weight of the sample residue after refluxing with hot aquadest (g), *c* represents the weight of sample residue after refluxing with H_2SO_4 1 N (g), *d* represents the weight of sample residue after 72% H_2SO_4 treatment (g), and *e* represents the weight of ash from sample residue (g).

2.3. Hybrid Ozonation-Ultrasonic Pretreatment

The hybrid ozonation-ultrasonic pretreatment process for oil palm mesocarp fibers was carried out using a glass reactor equipped with a gas sparger and an ultrasound bath, as shown in Fig 2.
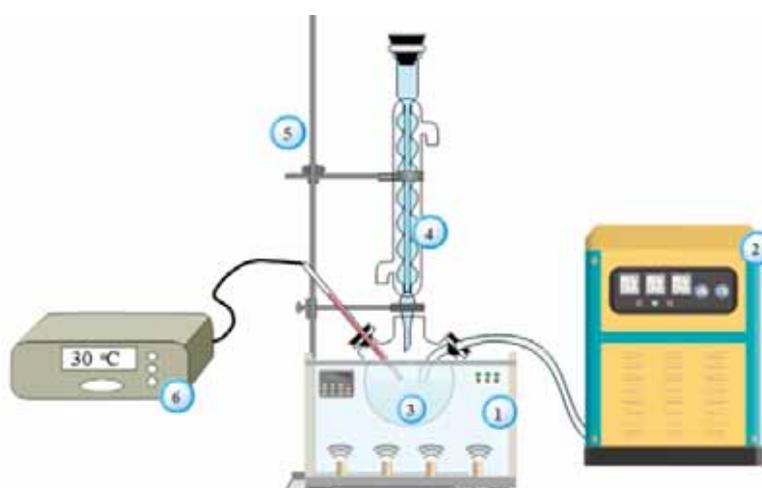
Fig. 2. Hybrid ozonation-ultrasonic process equipment, (1) ultrasound bath; (2) ozone generator; (3) flask; (4) reverse cooling; (5) stative clamp; (6) thermocouple

Oil palm mesocarp fibers weighing 15 g were placed into a glass reactor connected to an ozone generator and ultrasound bath, then 225 mL of distilled water was added, which was adjusted to pH 9 by adding 0.1 M NaOH solution, and the reaction process was carried out by flowing ozone gas and ultrasonic irradiation at an intensity of 42 kHz. After the reaction process is complete, the ultrasound bath and ozone generator were turned off, and the reaction products were filtered and then washed with distilled water until the pH was neutral. The filtered residue was dried using an oven at 105 ºC for 24 hours and then stored for analysis of the composition of lignin, hemicellulose, and cellulose using the Chesson-Datta method.

2.4. Process Variable Optimization

Experimental analysis of the mesocarp fiber pretreatment optimization process was carried out using statistical software 6.0 with RSM and BBD design, which has three factors, one block, and fifteen trials, was chosen to identify the optimal parameters that have a significant influence on the response and investigate the optimum conditions for the pretreatment process of oil palm mesocarp fibers [29]. Three independent variables in the pretreatment process, namely reaction time (X_1), reaction temperature (X_2), and ozone flow rate (X_3), each of which has three levels of value, were selected based on Abdurrahman [25]. The ranges of independent variables in the experimental process are summarized in Table 1, while the experimental designs given by the RSM are presented in Table 2.

Table 1: Process independent variable range

| Variable | Variable Code | | | | Response (%) |
	Code	-1	0	+1	
Reaction time (min)	x_1	30	60	90	Lignin Degradation
Reaction Temperature (ºC)	x_2	20	30	40	
Ozone Flow Rate (L/min)	x_3	1	2	3	

Table 2: Independent variable from design of experiment (DOE)

| Run | Variable | | | Code | | | Response |
	Time (min)	Temperature (ºC)	Ozone Flow Rate (L/min)	X_1	X_2	X_3	Lignin Degradation (%)
1	30	20	2	-1	-1	0	Y_1
2	90	20	2	1	-1	0	Y_2
3	30	40	2	-1	1	0	Y_3
4	90	40	2	1	1	0	Y_4
5	30	30	1	-1	0	-1	Y_5
6	90	30	1	1	0	-1	Y_6
7	30	30	3	-1	0	1	Y_7
8	90	30	3	1	0	1	Y_8
9	60	20	1	0	-1	-1	Y_9
10	60	40	1	0	1	-1	Y_{10}
11	60	20	3	0	-1	1	Y_{11}
12	60	40	3	0	1	1	Y_{12}
13	60	30	2	0	0	0	Y_{13}
14	60	30	2	0	0	0	Y_{14}
15	60	30	2	0	0	0	Y_{15}

The data obtained were analyzed using analysis of variance (ANOVA) and the response surface to determine the functional relationship between the independent process variables and the desired response [30]. The second-order polynomial regression equation describes the significance of the relationship between the independent variables in the response. The polynomial equation is described as follows:

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i X_i + \sum_{i=1}^{k} \beta_{ii} X_i^2 + \sum_{i=1}^{k} \sum_{j=i+1}^{k} \beta_{ij} X_i X_j \tag{5}$$

Where Y is the desired response, β_O represents a constant coefficient, X_i and X_j represent independent variables, k represents the number of independent variables, and β_i, β_{ii}, and β_{ij} represent linear coefficients, quadratic coefficients, and interaction coefficients. Using the RSM method in the experimental process can help summarize the number of trials needed to optimize the system and response so that the optimization process can be more efficient and save the cost of the experiments used [31].

2.5. Analysis of Cellulose Crystallinity with X-Ray Diffraction (XRD)

Crystallinity index (CrI) values of oil palm mesocarp fiber biomass without pretreatment and pretreatment and commercial cellulose were analyzed using X-rays at an angle of 2θ ranging from 5º to 100º. The detector detects the recorded X-ray diffraction spectrum results to form a diffraction pattern with peak intensity, which can be used to calculate cellulose crystallinity. The crystallinity index (CrI) value can be calculated using the following methods [32].

$$CRI\ (\%) = \left(\frac{I_{002} - I_{am}}{I_{002}}\right) \times 100 \tag{6}$$

Where (I_{002}) denotes the crystal peak intensity, (CrI) the crystallinity index, and (I_{am}) the amorphous phase peak intensity.

2.6. Functional Group Analysis with Fourier-Transform Infrared Spectroscopy (FTIR)

Fourier-transform infrared spectroscopy (FTIR) analysis aims to determine the chemical composition of lignocellulosic biomass of oil palm mesocarp fiber before pretreatment and after pretreatment with commercial cellulose composition as the standard value. Infrared spectroscopy (FTIR) examination of lignocellulosic biomass was performed at wavelengths between 400 cm^{-1} and 4000 cm^{-1}. The results of all formed spectra are reported in reflection mode at a predetermined wavelength range.

2.7. Cellulose Morphology Analysis by Scanning Electron Microscopy (SEM)

In order to study how the oil palm mesocarp fiber changed morphologically before and after pretreatment with lignin degradation, scanning electron microscopy (SEM) was conducted. Micrographs of oil palm mesocarp fiber biomass samples were observed at 7500 x magnification.

3. RESULTS AND DISCUSSION

3.1. Oil Palm Mesocarp Fiber Composition

The main composition of oil palm mesocarp fiber consists of carbohydrates, namely holocellulose (cellulose and hemicellulose) and lignin. Oil palm mesocarp fiber composition can be determined using the Chesson-Datta method. The results of measurements using the

Chesson Datta method show that the composition of the oil palm mesocarp fiber in this study is in accordance with previous studies that have been reported in the literature (Table 3). The lignin composition in mesocarp fiber, which is relatively high, gives toughness and stiffness to the cell walls of biomass so that it can become an obstacle in the valorization process of oil palm mesocarp fiber biomass both in the bioenergy and biocomposite fields [33].

Table 3. Composition of treated and untreated oil palm mesocarp fiber

Raw material	Oil Palm Mesocarp Fiber Composition (%)			Reference
	Cellulose	Hemicellulose	Lignin	
Untreated Mesocarp Fiber	37.12	29.72	28.70	This study
Untreated Mesocarp Fiber	43.70	34.20	24.00	[34]
Untreated Mesocarp Fiber	28.28	32.70	32.40	[35]
Untreated Mesocarp Fiber	27.85	24.04	31.30	[36]
Treated Mesocarp Fiber	90.01	4.56	2.27	This study

From the data reported, the composition of lignin in oil palm mesocarp fibers is relatively different. It is influenced by differences in the background from plantation areas and types of oil palm as well as the level of maturity of oil palm fruit when the sample is used [37]. Therefore, it is very important to analyze the initial composition of lignocellulosic biomass to determine the composition of the raw materials to be used, considering that cellulose, hemicellulose, and lignin are the main compositions that can affect the efficiency in the valorization of oil palm mesocarp fiber biomass. Additionally, the results showed that the lignin and hemicellulose content of palm mesocarp fiber decreased after the hybrid ozonization-ultrasonic pretreatment, as lignin and hemicellulose had been degraded.

3.2. Statistical Analysis and Empirical Models for Lignin Degradation

In this study, the results of the lignin degradation pretreatment and the predicted values generated by the experimental design of the statistical software are shown in Table 4.

The percentage of lignin degradation in mesocarp fiber obtained from this study showed varying values. The lowest percentage of lignin degradation was obtained at 79.27%, and the highest percentage of lignin degradation was obtained at 92.08%. While the process of lignin degradation using the ozonation and ultrasonic methods used individually has been reported by previous studies and shows that lignin degradation is only 87.9% in the ozonation process and 20.11% in the ultrasonic process, respectively [5,38]. Thus, the pretreatment of lignocellulosic biomass using the hybrid ozonation-ultrasonic method showed an increase in the lignin degradation process. Regression analysis uses second-order polynomial equations to explain the relationship between the independent and dependent variables. The polynomial equation is presented in Eqn. (1):

$$Y = 62.3300 + 0.32694\ X_1 + 1.11200\ X_2 + 2.94083\ X_3 - 0.00263\ X_1^2 - 0.01848\ X_2^2$$
$$- 0.68833\ X_3^2 + 0.00007\ X_1X_2 - 0.00017\ X_1X_3 - 0.003\ X_2X_3 \tag{1}$$

Where Y is the percentage of lignin degradation (%), X_1 is the time (min), X_2 is the temperature (°C), and X_3 is the ozone flow rate (L/min). The polynomial equation shows that time, temperature, and ozone flow rate can significantly affect the response, namely the percentage of lignin degradation indicated by a positive constant value in the model. Meanwhile, the interaction between the independent variables did not significantly affect the

percentage of lignin degradation. It was reflected in the negative value of the model. The significance of the polynomial equation can be validated by the resulting regression coefficient (R^2) and statistical analysis such as analysis of variance (ANOVA). The regression coefficient value formed from this study is 99.86%. It proves that the predicted and observed values are close together, so the model can be represented in predicting lignin degradation with an error value of less than 5%. The significance of the independent variables of the hybrid ozonation-ultrasonic pretreatment consisting of reaction time, reaction temperature, and ozone flow rate to the response, namely the percentage of lignin degradation, is shown in Table 5.

Table 4: Experimental data of oil palm mesocarp fiber

Run	Variable			Response (%)	
	Time (min)	Temperature (ºC)	Ozone Flow Rate (L/min)	Experiment	Prediction
1	30	20	2	87.55	87.60
2	90	20	2	88.24	88.25
3	30	40	2	87.51	87.50
4	90	40	2	88.12	88.07
5	30	30	1	88.64	88.61
6	90	30	1	89.20	89.21
7	30	30	3	88.82	88.81
8	90	30	3	89.40	89.43
9	60	20	1	89.48	89.46
10	60	40	1	89.34	89.38
11	60	20	3	89.78	89.74
12	60	40	3	89.52	89.54
13	60	30	2	92.08	92.07
14	60	30	2	92.06	92.07
15	60	30	2	92.07	92.07

P values (probability) can be used to differentiate statistical analysis results in ANOVA tables. Where if the value of P (probability) obtained is less than 5% or ($P < 0.05$), then it is considered significant because it has a probability level of 95%. It can be inferred that the independent variable and response have a greater significance if the P value (probability) is smaller [39]. Statistical analysis in this study showed that the linear function and quadratic function of the variable reaction time (X_1), reaction temperature (X_2), and ozone flow rate (X_3) showed a significant effect on the response of the percentage of lignin degradation because it had a P value <0.05. The interaction function between the reaction time variable and the ozone flow rate (X_1X_3) showed similar P value. Meanwhile, the interaction function between the variable reaction time and reaction temperature (X_1X_2) and the reaction temperature variable and ozone flow rate (X_2X_3) did not have a significant effect on the response ($P > 0.05$). This hypothesis is strengthened by the F-test (Fisher), where the results obtained from the F-test show that the independent variables (reaction time, reaction temperature, and ozone flow rate), both in linear and quadratic form, have a higher F value than the F table (F-value > 4.772) so that it can be said that the independent variables in the linear and quadratic forms have a significant effect on lignin degradation. In the interaction between variables, the F-value has a smaller value than the F-table (F-value < 4.772), so the interaction between variables does not have a significant effect on lignin degradation.

Table 5: ANOVA for quadratic equation model in hybrid ozonation-ultrasonic pretreatment

Source	df	SS	MS	Reg Coeff	P-Value	F-Value	
Model	9	32.3804	3.5978	62.3300	0.00000	1484.660	Significant
Time (X_1)	1	0.7442	0.7442	0.32694	0.00001	307.098	
Temperature (X_2)	1	0.0392	0.0392	1.11200	0.01010	16.176	
Ozone Flow Rate (X_3)	1	0.0925	0.0925	2.94083	0.00162	38.150	
Time (X_1^2)	1	20.6228	20.6228	-0.00263	0.00000	8510.101	
Temperature (X_2^2)	1	12.6142	12.6142	-0.01848	0.00000	5205.295	
Ozone Flow Rate (X_3^2)	1	1.7494	1.7494	-0.68833	0.00000	721.909	
Time-Temperature (X_1X_2)	1	0.0016	0.0016	-0.00007	0.45341	0.660	Not Significant
Time- Flow Rate (X_1X_3)	1	0.0001	0.0001	0.00017	0.84703	0.041	
Temperature- Flow Rate (X_2X_3)	1	0.0036	0.0036	-0.00300	0.27727	1.486	
Error	5	0.0121	0.0024				
Lack-of-Fit	3	0.0117	0.0039		0.0572		
Pure Error	2	0.0005	0.0002				
Total	14	32.3925					
$R^2 = 0.9996$ dan $R^2_{Adj} = 0.9989$							

The residual mean square (MS) value in the analysis of variance (ANOVA) is used to describe the difference between the experimental data and the predicted value of the model [40]. This study produced a residual mean square (MS) value of 0.0173, so it can be concluded that the model used is good enough and accurate in explaining the closeness between the experimental results and the predicted values obtained from the experimental design. The closeness between the experimental data and the predicted value is clarified by the coefficient
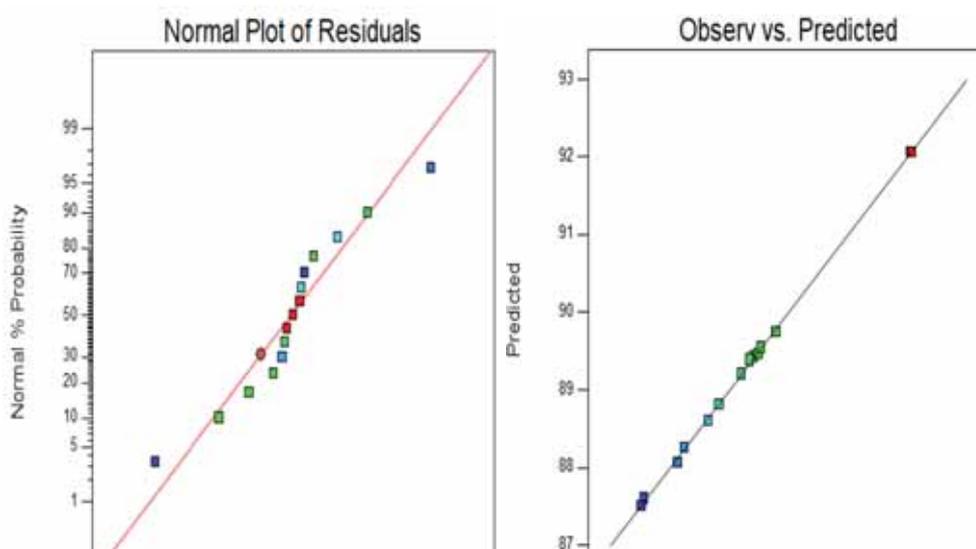
Fig. 3. Probability plots (a) normal vs internally studentized residual, (b) predicted vs observer

of determination (R^2) obtained, which is equal to 0.9996, so it can be stated that 99.96% of the model can explain the experimental data. The reliability and significance of the model can be validated by looking at the average probability percentage in Figure 3a and the relationship between the experimental results obtained, and the predicted value of the model can also be validated by the observer vs predicted diagram illustrated in Figure 3b.

The diagonal line on the normal probability plot shows the data predicted by the experimental design. In contrast, the dots around the diagonal line show the value of the experimental results obtained. The closer the points, which are the experimental data, to the diagonal line, the residual values are normally distributed [41]. This normal distribution analysis aims to observe the magnitude of the deviation from the model. In this study, the normal probability plot shows dots that spread and approach the diagonal line, so it can be said that the model given by the experimental design can fulfill the assumption of normality.

3.3. Effect and Interaction of Variables Process on Lignin Degradation

The influence of process variables and their interactions during the pretreatment of lignin degradation with the ultrasonic hybrid ozonation method can be explained by employing response surface plots. Response surface analysis can provide a 3-dimensional profile that explains the relationship between the independent variables and the response. The response surface plot of the percentage of lignin degradation is presented in Figure 4.
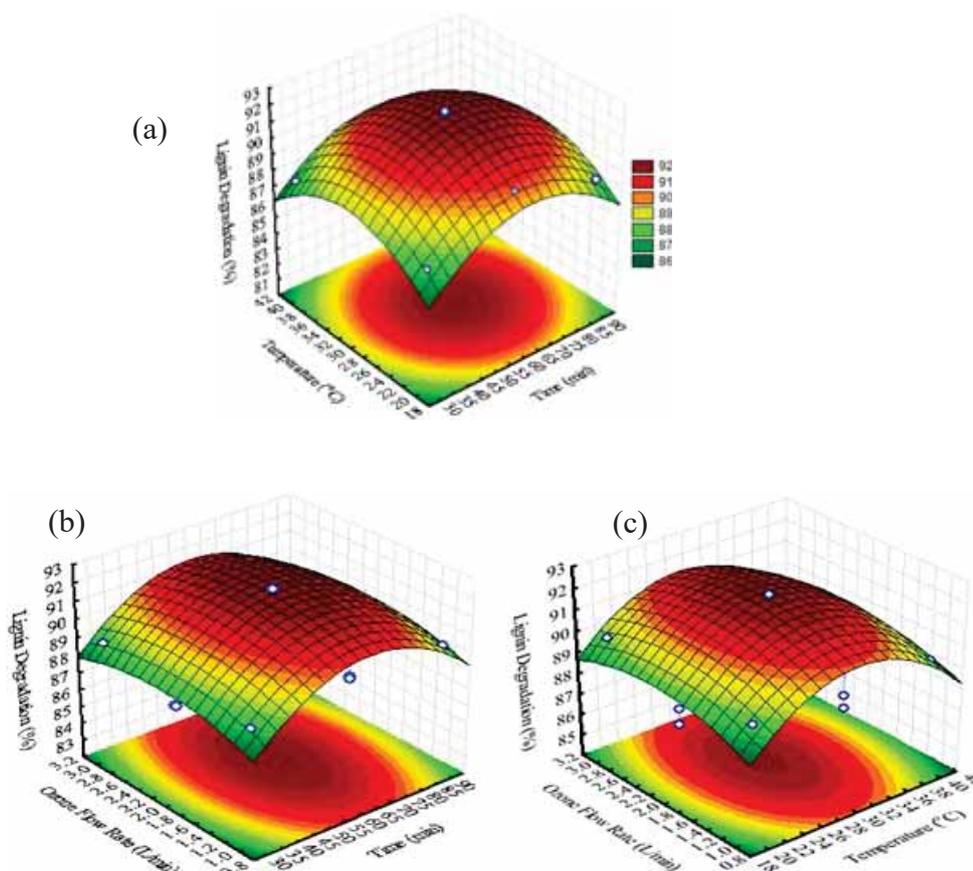
Fig. 4. Response surface methodology of an hybrid ozonation-ultrasonic pretreatment
(a) X_1 & X_2, (b) X_1 & X_3, (c) X_2 & X_3

Based on the response surface plot images, the percentage of lignin degradation increases with increasing reaction time, reaction temperature, and ozone flow rate up to a certain critical point. The highest percentage of lignin degradation is represented by the contour plot area of the response surface, which has a solid red color. In contrast, the smallest percentage of lignin degradation depicted on the contour plot is colored green. In this study, the contour plots showed that more than 90% of lignin degradation was obtained when it was reacted at a temperature of 24-31 °C, with an ozone flow rate of 1-2 L/minute, for 50-65 minutes. According to Ramados et al. [42] and Pujakaroni et al. [43], the highest percentage of lignin degradation occurs when the temperature is 40 °C for 60 minutes. Increasing the reaction time can extend the contact time between radical compounds and a little ozone with lignocellulosic biomass, thus increasing lignin degradation. However, using a reaction time of more than 60 minutes does not have a significant effect because most of the lignin has been degraded so that it opens the surface of lignocellulosic biomass, and ozone could attack the holocellulose component [44]. Likewise, an increase in temperature causes an increase in the solubility and diffusion coefficient of the lignin component. On the other hand, although the concentration and solubility of ozone in water decrease with an increase in the reaction temperature, more hydroxy radical compounds are also produced under the same conditions, which have a greater oxidation potential value than ozone so that they are more oxidative which leads to an increase in the rate of degradation [45].

According to Chiha et al. [46], the degradation products increased from 20°C to 30°C as temperature increased and tended to be constant and even decreased when the temperature was higher than 30 °C. The decrease in the efficiency of the organic compound degradation process using the hybrid ozonation-ultrasonic method is caused by a decrease in ozone solubility and a decrease in the formation of radical compounds due to an increase in water vapor in the bubbles, which provides a cushioning effect on the collapsed bubbles so that the bursting of cavitation bubbles only produces low energy [47]. A flow rate of 2 L/min of ozone has a significant influence on the results, as shown by Wang et al. [48], that the ozone flow rate is increased, and degradation efficiency increases because cavitation bubbles form under ultrasonic irradiation, accelerating the mass transfer of ozone and causing radical compounds to form that can accelerate organic compound degradation. However, excessive ozone flow rates can cause greater turbulence, decreasing the contact time between ozone and water and shorter biomass, thereby reducing the degradation percentage [49].

3.4. Optimization of The Hybrid Ozonation-Ultrasonic Pretreatment Process Variables

Optimizing the lignin degradation pretreatment process using the hybrid ozonation-ultrasonic method was conducted by analyzing the desirability function represented by the response desirability profiling contained in statistical software 6. This tool consists of a series of graphs that represent every independent variable to determine whether the independent variables provide significant responses to the desired outcome. The results of optimizing the lignin degradation pretreatment process with the hybrid ozonation-ultrasonic method are shown in Figure 5.

The results of the analysis of the desirability function for each independent variable and the predicted response at optimal conditions can be observed from the response desirability profiling graph. The desirability profile graph shows that the percentage of lignin degradation in the hybrid ozonation-ultrasonic pretreatment has increased along with the increasing values of independent process variables such as reaction time, reaction temperature, and ozone flow rate until they reach a certain critical value. The optimal condition of the lignin degradation process can be seen from the intersection of the independent variables with the highest response values. In this study, the optimum conditions were obtained at 30 °C, with an ozone flow rate

of 2 L/min, and for 60 minutes of reaction time, the optimum response was in the form of a lignin degradation percentage of 92.08%.
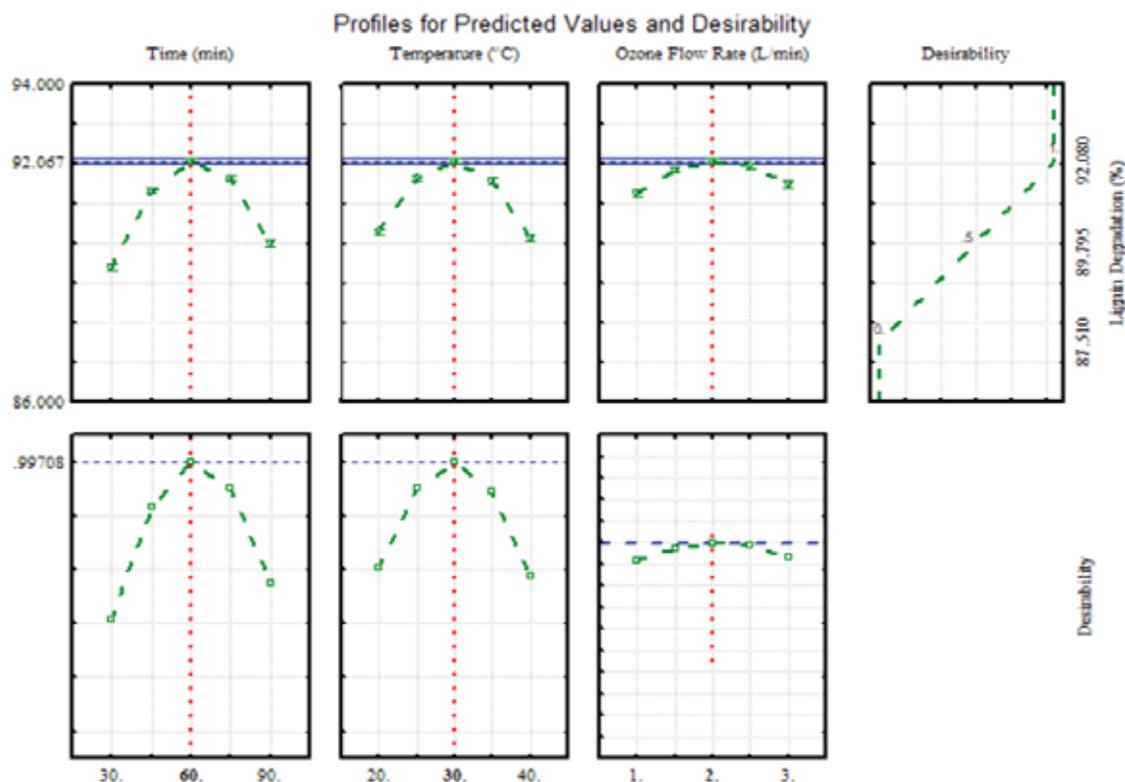
Fig. 5. Profiles desirability of hybrid ozonation-ultrasonic pretreatment

3.5. Proposed Hybrid Ozonation-Ultrasonic Reaction (AOPS) Mechanism

The process of degradation of organic compounds such as lignocellulosic biomass using a combined method of ozonation and ultrasonic processes provides a higher degradation rate than the ozonation and ultrasonic processes when carried out individually. The possible reaction mechanism in the hybrid ozonation-ultrasonic process in degrading lignocellulosic biomass is shown in Figure 6.

$$O_3 \xrightarrow{us)))} O(^3P)^* + O_2 \quad (7)$$
$$H_2O \xrightarrow{us)))} OH^* + H^* \quad (8)$$
$$O(^3P)^* + H_2O \longrightarrow 2OH^* \quad (9)$$
$$O_2 + H^* \longrightarrow HO_2^* \quad (10)$$
$$HO_2^* \longrightarrow H^* + {}^*O_2^- \quad (11)$$
$${}^*O_2^- + O_3 \longrightarrow {}^*O_3^- + O_2 \quad (12)$$
$${}^*O_3^- + H^* \longrightarrow HO_3^* \quad (13)$$
$$HO_3^* \longrightarrow OH^* + O_2 \quad (14)$$
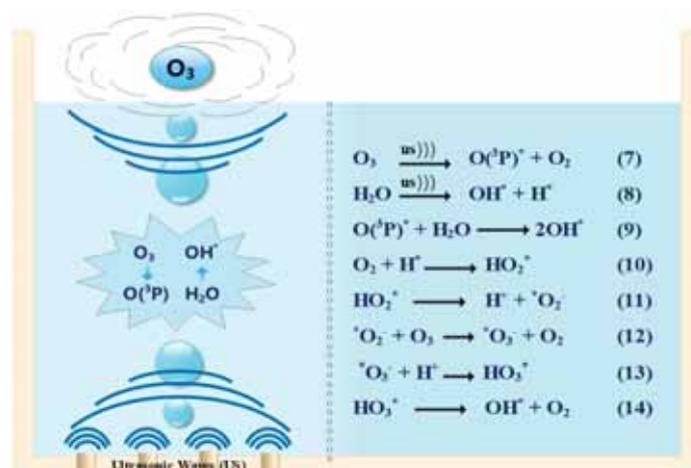
Fig. 6. Reaction mechanism of the hybrid ozonation-ultrasonic pretreatment process for oil palm mesocarp fiber

Decomposed ozone molecules initiate the mechanism of the hybrid ozonation-ultrasonic reaction, and water molecules dissociate in the presence of cavitation bubbles under ultrasonic irradiation to produce the free radical compounds shown in equations (7) and (8), then oxygen atoms produced from the ozone decomposition process. It will react directly with steam from water molecules in the liquid phase to form radical hydroxy compounds, as in equation (9). In addition, a small part of the ozone dissolved during the ultrasonic process can react with lignin directly on lignocellulosic biomass as shown in Figure 7 and then the lignin is degraded. The hybrid ozonation-ultrasonic process generally provides high efficiency in the lignocellulosic biomass pretreatment process because each ozone molecule will produce twice the hydroxyl radicals, which can better oxidize lignin and also the products resulting from the dissociation of water molecules and ozone decomposition will react with each other in the bubble interface and then will diffuse into the aqueous phase (Eq. (10)-(14)) [25]. Zhao et al. [50] reported that the enhanced decomposition of the ozone molecule in the cavitation bubbles under ultrasonic irradiation is due to mechanical effects such as the collapse of the cavitation bubbles.
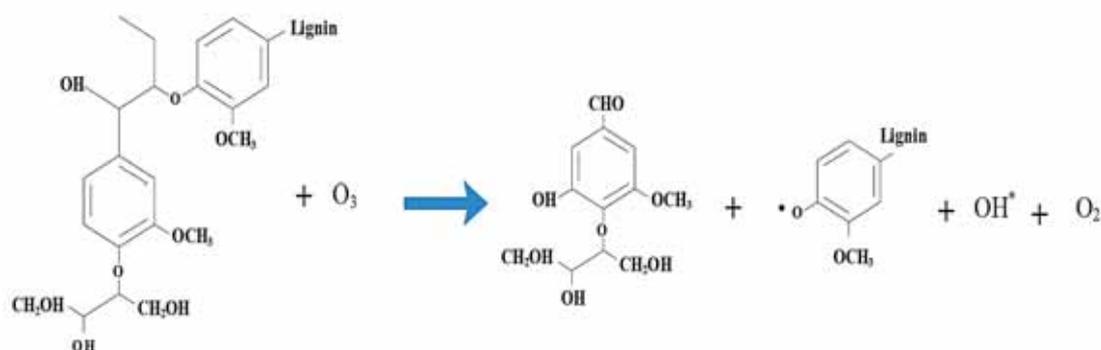
Fig. 7. Mechanism of lignin degradation reaction with ozone

The reaction mechanism for the degradation of lignin compounds by hydroxyl radicals formed during the decomposition process of ozone and the dissociation of water molecules due to the effects of the acoustic cavitation phenomenon in the hybrid ozonation-ultrasonic pretreatment process is shown in Figure 8. However, it is important to note that the mechanism of the proposed lignin degradation reaction by hydroxyl radical compounds is only a speculative mechanism based on the reaction mechanism that has been proposed previously in the literature [51,52]. Pretreatment of lignin degradation with hydroxyl radical compounds involves several processes, such as demethylation or demethoxylation, side chain oxidation, and aromatic hydroxylation processes. Lignin degradation begins with breaking aryl ether bonds in lignin and forms water-soluble phenolic groups [53]. Furthermore, the highly electrophilic hydroxyl radical compounds will also attack other lignin groups rich in electrons through side chain oxidation processes and aromatic hydroxylation.

3.6. X-Ray Diffraction (XRD) Characterization of Lignocellulosic Biomass

XRD analysis of oil palm mesocarp fiber aims to determine the crystal structure of the biomass and structural changes both before and after the lignin degradation process. Cellulose in lignocellulosic biomass has crystalline and amorphous fractions, while the hemicellulose and lignin components have amorphous fractions. The crystallinity index (CrI) can be determined by referring to the XRD pattern, which is the diffraction peak of the cellulose I structure at around 20º and the lowest diffraction peak at around 18º, which is an amorphous cellulose region. Most industries have used analysis of cellulose crystallinity to determine elasticity, absorption, and other physical properties, which are important parameters in the production of biocomposites and bioenergy [54]. The diffraction patterns of non-pretreated and

pretreated oil palm mesocarp fiber biomass and the diffraction patterns of commercial alpha cellulose are presented in Figure 9.
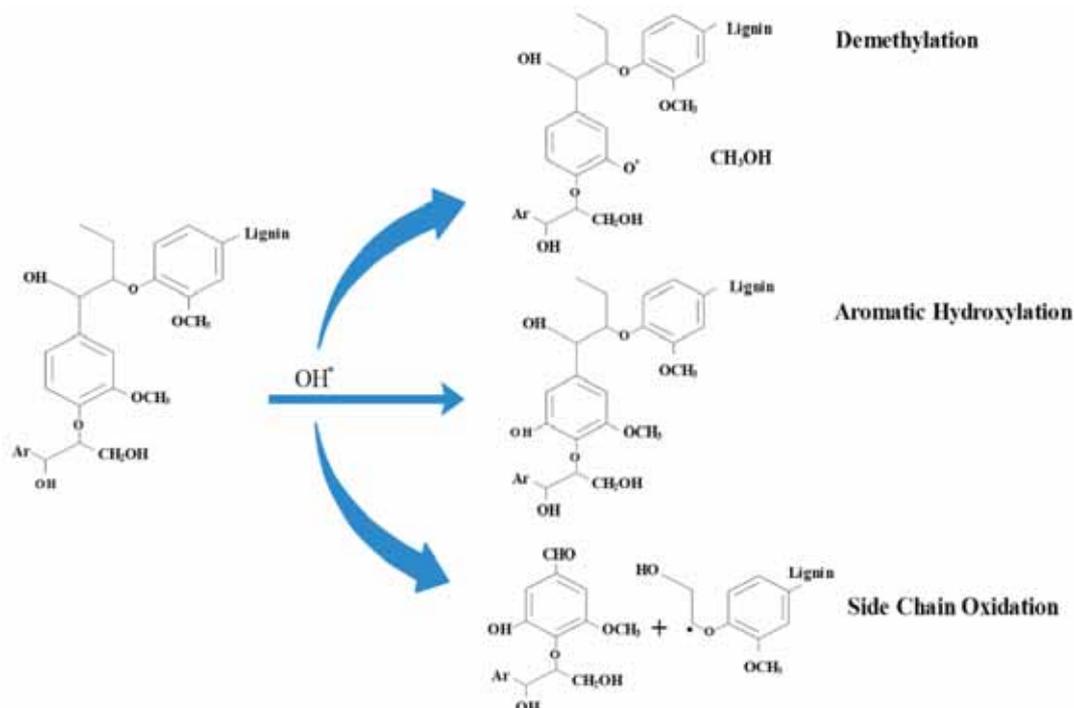
Fig. 8. Mechanism of lignin degradation reaction with hydroxyl radicals
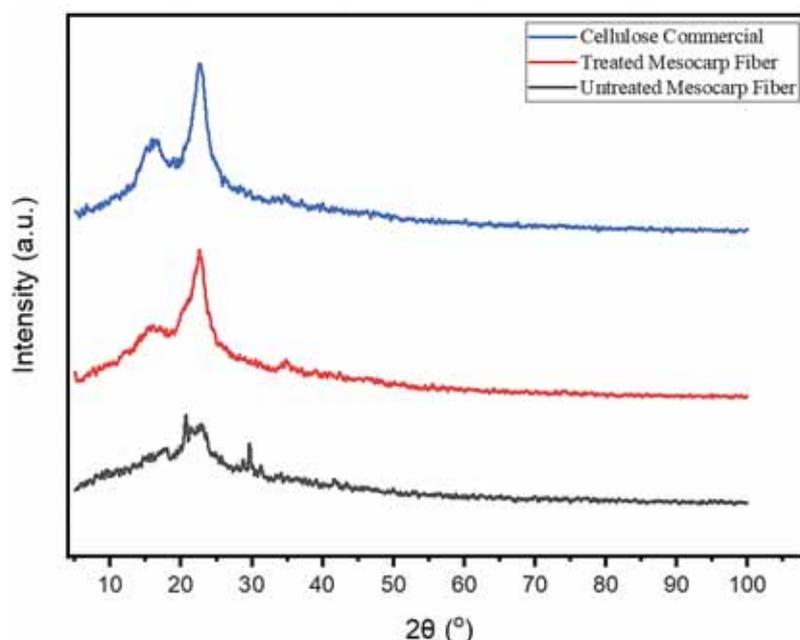
Fig. 9. XRD pattern from untreated and treated oil palm mesocarp fiber

The diffraction peaks detected in the XRD pattern show high peaks at an angle of 2θ, namely at $15^{\circ} - 16^{\circ}$ and $2\theta = 20^{\circ} - 22.5^{\circ}$, which indicates that the oil palm mesocarp fiber biomass has a crystal structure of cellulose I and is a natural plant cellulose [55]. The diffraction peak at angle 2θ, about 22.5°, is associated with the typical peak portion of the cellulose crystal

structure. Whereas at an angle of 2θ, which is 15º, is part of the amorphous region in the biomass. Rosli [56] stated that the diffraction peak of the cellulose crystal structure is at an angle of 2θ, which is around 22-23º. In the biomass that was not pretreated, it showed a broad peak which was associated with the amorphous cellulose area. The crystallinity index (CrI) in mesocarp fibers can be calculated using the Segal method, which is summarized in Table 6.

Table 6: Crystallinity index of oil palm mesocarp fiber and cellulose commercial

Samples	2θ (Amorphous)($^{\circ}$)		2θ (002) ($^{\circ}$)		CrI (%)
	Degree	Intensity (I_{am})	Degree	Intensity (I_{002})	
Mesocarp Fiber	20.75	1080	22.67	2332	53.69
Cellulose Hybrid O_3/US	20.33	1250	22.52	6314	80.20
Cellulose Commercial	16.00	2184	22.64	8205	73.38

The crystallinity index (CrI) value of mesocarp fibers undergoing hybrid ultrasonic ozonation pretreatment increased due to the loss of amorphous structures associated with lignin and hemicellulose components. The crystal structure of cellulose is recalcitrant, so it is not easy to remove in the pretreatment process. The low crystallinity index (CrI) value of the raw material for palm oil mesocarp fiber is due to the lignin and hemicellulose components which are cross-linked and have an amorphous structure. Thus, increasing the crystallinity index (CrI) value may imply that pretreatment with the hybrid ozonation-ultrasonic method can degrade components with amorphous structures such as lignin and hemicellulose properly and efficiently. On the other hand, increasing the crystallinity index (CrI) can also encourage an increase in the mechanical properties of cellulose for advanced industrial use [37].

3.7. Fourier-Transform Infrared Spectroscopy (FTIR) Characterization of Biomass

Fourier-transform infrared spectroscopy (FTIR) analysis observed functional groups of oil palm mesocarp fiber before and after hybrid ozonation-ultrasonic pretreatment. The characteristic features of the FTIR spectrum of oil palm mesocarp fiber biomass are presented in Figure 10.
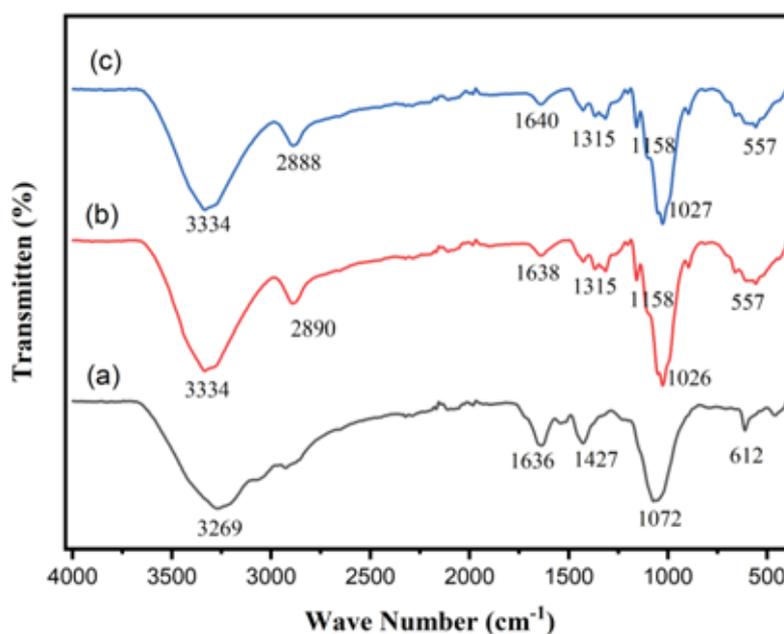
Fig. 10. FTIR spectra of (a) untreated mesocarp fiber; (b) treated mesocarp fiber; (c) commercial cellulose.

The FTIR spectrum of cellulose in oil palm mesocarp fiber both before pretreatment and after pretreatment of lignin degradation produced different types and the number of absorption peaks indicating that the hybrid ozonation-ultrasonic pretreatment could destroy the structure of lignocellulosic biomass. The broad absorption peaks in the region from 3200 cm^{-1} to 3400 cm^{-1} are caused by the stretching of the O-H functional groups, which is the main characterization of cellulose [57,58]. Pretreated cellulose and commercial cellulose samples showed absorption peaks in the area of 2890 cm^{-1}, which indicated sp^3 stretching carbon in the methyl (CH_3) and methylene (CH_2) groups. In contrast, in that area, there were no absorption peaks for non-pretreated oil palm mesocarp fibers [59]. The absorption peaks in the 1638 cm^{-1} area and 1427 cm^{-1} area, respectively, represent the structure of the aromatic ring in lignin according to the C=C and -CH_2 bonding functional groups. The intensity of this absorption peak was reduced in pretreated cellulose and commercial cellulose samples, so it can be concluded that the hybrid ozonation-ultrasonic pretreatment process can damage the structure of lignocellulosic biomass and degrade lignin [60,61]. The absorption peak in the 1315 cm^{-1} area represents a typical cellulose absorption peak with the CH_2 rocking functional group. In this area, the intensity of the absorption peak increases in pretreated cellulose [54,62]. Increased intensity of absorption peaks in pretreated cellulose samples and commercial cellulose also occurs in the areas 1027 cm^{-1} and 557 cm^{-1}, which are the stretching of the β-glycosidic bond, namely C-O-C stretching and the lowest C-H vibration, which refers to an increase in the value of cellulose crystallinity [63,64]. An analysis of Fourier-transform infrared spectroscopy (FTIR) results confirmed that the lignin composition and amorphous areas in lignocellulosic biomass were successfully degraded because similar trends were also shown in commercial cellulose.

3.8. Scanning Electron Microscopy (SEM) Characterization of Biomass

Changes in surface morphology and roughness of the lignocellulosic biomass of oil palm mesocarp fiber before and after pretreatment can be identified through scanning electron
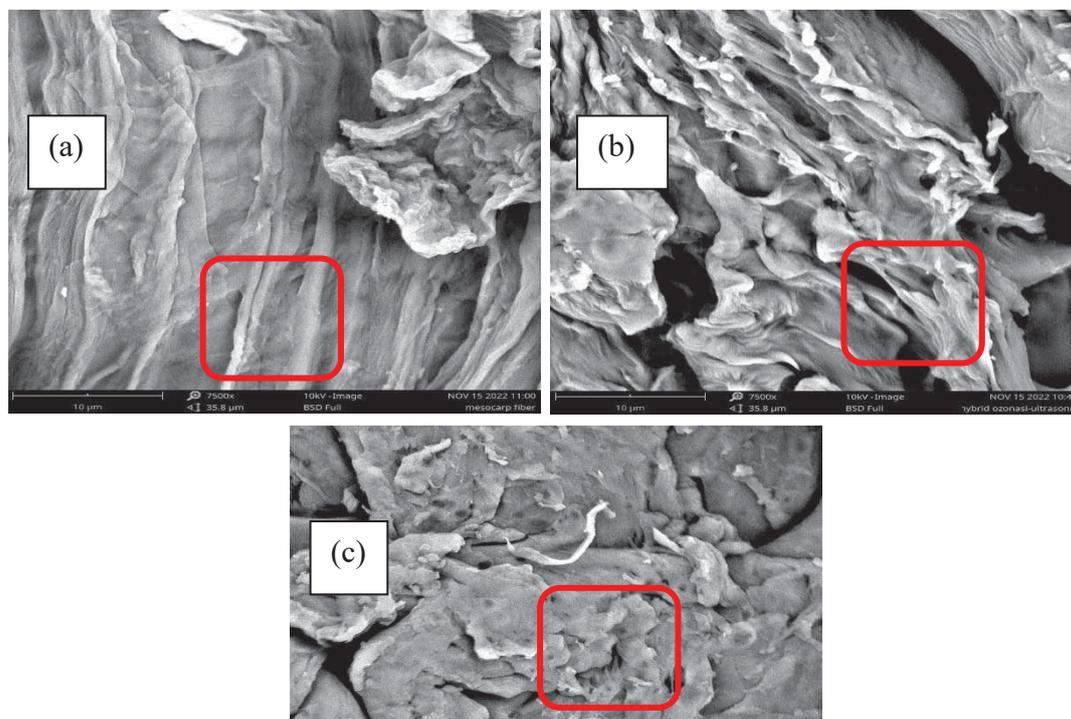
Fig. 11. SEM micrographs of (a) untreated mesocarp fiber, (b) hybrid ozonated-ultrasonic, (c) commercial cellulose

microscopy (SEM) analysis. Micrographs of oil palm mesocarp fibers are presented in Figure 11.

The SEM micrographs demonstrated that the surface of untreated mesocarp fiber lignocellulosic biomass had a dense and smooth structure without any cell wall cracks or fiber damage. It is shown by the red square, which indicates that the lignin structure was still present on the surface of the lignocellulosic biomass. However, in the biomass that was pretreated using the hybrid ozonation-ultrasonic method, as indicated by the red square in section (b), shows that the surface structure of the biomass experienced perforations such as damage caused by some of the degraded lignin structure and also due to shear forces and the sudden burst of cavitation bubbles that collided with the structure. Lignin in the ultrasonic process causes lignocellulosic biomass surface damage [65]. In commercial cellulose, it is also seen that the structure of lignocellulosic biomass is no longer dense and smooth but has been damaged, as shown by the red square in section (c). It is because lignin degradation has occurred, and some of the structure has been removed.

4. CONCLUSION

The hybrid ozonation-ultrasonic pretreatment process for lignin degradation in oil palm mesocarp fiber biomass shows the potential for application, with the acquisition of a high percentage of lignin degradation, up to 98.02%. The RSM was used based on the BBD to obtain the highest percentage of lignin degradation. The optimum conditions were obtained at 30°C, with a flow rate of 2 L/min for 60 minutes. A decrease in the absorption peak of the aromatic ring structure in the FTIR spectrum confirms the occurrence of a lignin degradation process. It was validated by an increase in the crystallinity index (CrI) of mesocarp fiber biomass subjected to a hybrid ozonation-ultrasonic pretreatment process and reinforced by changes in the surface structure of lignocellulosic biomass as shown in the results of SEM micrographs. Therefore, the results of this study can be used as a reference to increase the percentage of lignin degradation in lignocellulosic biomass.

ACKNOWLEDGEMENTS

REFERENCES

[1] Megashah LN, Ariffin H, Zakaria MR, Hassan MA. (2018). Properties of Cellulose Extract from Different Types of Oil Palm Biomass. IOP Conference Series: Materials Science and Engineering, 368. https://doi.org/10.1088/1757-899X/368/1/012049

[2] Dirkes R, Neubauer PR, Rabenhorst J. (2021). Pressed Sap from Oil Palm (Elaeis guineensis) Trunks: a Revolutionary Growth Medium for the Biotechnological Industry. Biofuels, Bioproducts and Biorefining, 15, 931–44. https://doi.org/10.1002/bbb.2201

[3] OECD/FAO. (2018). OECD FAO Agricultural Outlook 2018 2027 Oilseeds and Oilseed Products. 127–38. https://doi.org/10.1787/agr_outlook-2018-7-en

[4] Hossain MA, Jewaratnam J, Ganesan P. (2016). Prospect of Hydrogen Production from Oil Palm Biomass by Thermochemical Process – A Review. International Journal of Hydrogen Energy, Elsevier Ltd. 41, 16637–55. https://doi.org/10.1016/j.ijhydene.2016.07.104

[5] Wan Omar WNN, Amin NAS. (2016). Multi Response Optimization of Oil Palm Frond Pretreatment By Ozonolysis. Industrial Crops and Products, Elsevier B.V. 85, 389–402. https://doi.org/10.1016/j.indcrop.2016.01.027

[6] Travaini R, Martín-Juárez J, Lorenzo-Hernando A, Bolado-Rodríguez S. (2016). Ozonolysis: An Advantageous Pretreatment for Lignocellulosic Biomass Revisited. Bioresource Technology, Elsevier Ltd. 199, 2–12. https://doi.org/10.1016/j.biortech.2015.08.143

[7] Kabir G, Mohd Din AT, Hameed BH. (2017). Pyrolysis of Oil Palm Mesocarp Fiber and Palm Frond in a Slow-Heating Fixed-Bed Reactor: A Comparative Study. Bioresource Technology, Elsevier Ltd. 241, 563–72. https://doi.org/10.1016/j.biortech.2017.05.180

[8] Okolie JA, Epelle EI, Tabat ME, Orivri U, Amenaghawon AN, Okoye PU. (2022). Waste Biomass Valorization for The Production of Biofuels and Value-Added Products: A Comprehensive Review of Thermochemical, Biological and Integrated Processes. Process Safety and Environmental Protection, Elsevier. 159, 323–44. https://doi.org/10.1016/j.psep.2021.12.049

[9] Lee KM, Quek JD, Tey WY, Lim S, Kang HS, Quen LK. (2022). Biomass Valorization by Integrating Ultrasonication and Deep Eutectic Solvents: Delignification, Cellulose Digestibility and Solvent Reuse. Biochemical Engineering Journal, Elsevier B.V. 187, 108587. https://doi.org/10.1016/j.bej.2022.108587

[10] Baksi S, Saha S, Birgen C, Sarkar U, Preisig HA, Markussen S. (2019). Valorization of Lignocellulosic Waste (Crotalaria juncea) Using Alkaline Peroxide Pretreatment under Different Process Conditions: An Optimization Study on Separation of Lignin, Cellulose, and Hemicellulose. Journal of Natural Fibers, Taylor & Francis. 16, 662–76. https://doi.org/10.1080/15440478.2018.1431998

[11] Chauhan R, Dinesh GK, Alawa B, Chakma S. (2021). A critical Analysis of Sono-Hybrid Advanced Oxidation Process of Ferrioxalate System for Degradation of Recalcitrant Pollutants. Chemosphere, Elsevier Ltd. 277, 130324. https://doi.org/10.1016/j.chemosphere.2021.130324

[12] Li S, Yang Y, Zheng H, Zheng Y, Jing T, Ma J. et al. (2022). Advanced Oxidation Process Based on Hydroxyl and Sulfate Radicals to Degrade Refractory Organic Pollutants in Landfill Leachate. Chemosphere, Elsevier Ltd. 297, 134214. https://doi.org/10.1016/j.chemosphere.2022.134214

[13] Heebner A, Abbassi B. (2022). Electrolysis Catalyzed Ozonation for Advanced Wastewater Treatment. Journal of Water Process Engineering, Elsevier Ltd. 46, 102638. https://doi.org/10.1016/j.jwpe.2022.102638

[14] Nagels M, Verhoeven B, Larché N, Dewil R, Rossi B. (2022). Corrosion Behaviour of Lean Duplex Stainless Steel in Advanced Oxidation Process (AOP) Based Wastewater Treatment Plants. Engineering Failure Analysis, 136. https://doi.org/10.1016/j.engfailanal.2022.106170

[15] Kiyanmehr K, Moussavi G, Mohammadi S, Naddafi K, Giannakis S. (2022). The Efficacy of The VUV/O3 Process Run in A Continuous-Flow Fluidized Bed Reactor for Simultaneous Elimination of Favipiravir and Bacteria in Aqueous Matrices. Chemosphere, Elsevier Ltd. 304, 135307. https://doi.org/10.1016/j.chemosphere.2022.135307

[16] Barrera-Martínez I, Guzmán N, Peña E, Vázquez T, Cerón-Camacho R, Folch J. et al. (2016). Ozonolysis of Alkaline Lignin and Sugarcane Bagasse: Structural Changes and Their Effect on Saccharification. Biomass and Bioenergy, 94, 167–72. https://doi.org/10.1016/j.biombioe.2016.08.010

[17] Suzuki H, Araki S, Yamamoto H. (2015). Evaluation of Advanced Oxidation Processes (AOP) Using O3, UV, and Tio2 for The Degradation of Phenol in Water. Journal of Water Process Engineering, Elsevier Ltd. 7, 54–60. https://doi.org/10.1016/j.jwpe.2015.04.011

[18] Praptyana IR, Budiyono. (2022). Biohydrogen Production from Wood Dust Mahogany (Swietenia Mahagony) By Dark Fermentation Using Enterobacter Aerogenes: Effect of Ozone Pretreatment Time and Ph. Materials Today: Proceedings, Elsevier Ltd. 63, S203–9. https://doi.org/10.1016/j.matpr.2022.02.406

[19] Hassaan MA, Nemr AEl, Elkatory MR, Eleryan A, Ragab S, Sikaily AEl et al. (2021). Enhancement of Biogas Production from Macroalgae Ulva Latuca Via Ozonation Pretreatment. Energies, 14, 1–16. https://doi.org/10.3390/en14061703

[20] Heidari Z, Pelalak R, Eshaghi MR, Pishnamazi M, Rezakazemi M, Aminabhavi TM. et al. (2022). A New Insight Into Catalytic Ozonation of Sulfasalazine Antibiotic by Plasma-Treated Limonite Nanostructures: Experimental, Modeling and Mechanism. Chemical Engineering Journal, Elsevier B.V. 428, 131230. https://doi.org/10.1016/j.cej.2021.131230

[21] Ngo Thi TD, Nguyen LH, Nguyen XH, Phung HV, The Vinh TH, Van Viet P et al. (2022). Enhanced Heterogeneous Photocatalytic Perozone Degradation of Amoxicillin by ZnO Modified TiO2 Nanocomposites under Visible Light Irradiation. Materials Science in Semiconductor Processing, Elsevier Ltd. 142, 106456. https://doi.org/10.1016/j.mssp.2022.106456

[22] Barik AJ, Gogate PR. (2018). Hybrid Treatment Strategies for 2,4,6-Trichlorophenol Degradation Based on Combination of Hydrodynamic Cavitation and AOPs. Ultrasonics Sonochemistry, Elsevier. 40, 383–94. https://doi.org/10.1016/j.ultsonch.2017.07.029

[23] Ran J, Duan H, Srinivasakannan C, Yao J, Yin S, Zhang L. (2022). Effective Removal of Organics from Bayer Liquor Through Combined Sonolysis and Ozonation: Kinetics and Mechanism. Ultrasonics Sonochemistry, Elsevier B.V. 88, 106106. https://doi.org/10.1016/j.ultsonch.2022.106106

[24] Fan X, Cao T, Yu X, Wang Y, Xiao X, Li F et al. (2020). The Evolutionary Behavior of Chromophoric Brown Carbon During Ozone Aging of Fine Particles from Biomass Burning. Atmospheric Chemistry and Physics, 20, 4593–605. https://doi.org/10.5194/acp-20-4593-2020

[25] Abdurahman MH, Abdullah AZ. (2020). Mechanism and Reaction Kinetic of Hybrid Ozonation-Ultrasonication Treatment for Intensified Degradation of Emerging Organic Contaminants in Water: A Critical Review. Chemical Engineering and Processing - Process Intensification, Elsevier. 154, 108047. https://doi.org/10.1016/j.cep.2020.108047

[26] Shen Y, Xu Q, Wei R, Ma J, Wang Y. (2017). Mechanism and Dynamic Study of Reactive Red X-3B Dye Degradation By Ultrasonic-Assisted Ozone Oxidation Process. Ultrasonics Sonochemistry, Elsevier B.V. 38, 681–92. https://doi.org/10.1016/j.ultsonch.2016.08.006

[27] Weavers LK, Hoffmann MR. (1998). Sonolytic Decomposition of Ozone in Aqueous Solution: Mass Transfer Effects. Environmental Science and Technology, 32, 3941–7. https://doi.org/10.1021/es980620o

[28] Candido RG, Gonçalves AR. (2019). Evaluation of Two Different Applications for Cellulose Isolated from Sugarcane Bagasse in a Biorefinery Concept. Industrial Crops and Products, 142. https://doi.org/10.1016/j.indcrop.2019.111616

[29] Montgomery. (1991). Debris Flow Hazard Mitigation for Colluvium-Filled Swales.

[30] Dagnino EP, Felissia FE, Chamorro E, Area MC. (2017). Optimization of The Soda-Ethanol Delignification Stage for a Rice Husk Biorefinery. Industrial Crops and Products, Elsevier B.V. 97, 156–65. https://doi.org/10.1016/j.indcrop.2016.12.016

[31] Wu Y, Yao S, Narale BA, Shanmugam A, Mettu S, Ashokkumar M. (2022). Ultrasonic Processing of Food Waste to Generate Value-Added Products. Foods, 11. https://doi.org/10.3390/foods11142035

[32] Segal L, Creely JJ, Martin AE, Conrad CM. (1959). An Empirical Method for Estimating the Degree of Crystallinity of Native Cellulose Using the X-Ray Diffractometer. Textile Research Journal, 29, 786–94. https://doi.org/10.1177/004051755902901003

[33] Vanholme R, Morreel K, Ralph J, Boerjan W. (2008). Lignin Engineering. Current Opinion in Plant Biology, 11, 278–85. https://doi.org/10.1016/j.pbi.2008.03.005

[34] Anuar TATY, Ariffin H, Norrrahim MNF, Hassan MA. (2017). Factors Affecting Spinnability of Oil Palm Mesocarp Fiber Cellulose Solution for The Production of Microfiber. BioResources, 12, 715–34. https://doi.org/10.15376/biores.12.1.715-734

[35] Megashah LN, Ariffin H, Zakaria MR, Ando Y. (2018). Characteristics of Cellulose from Oil Palm Mesocarp Fibres Extracted by Multi-Step Pretreatment Methods. IOP Conference Series: Materials Science and Engineering, 368. https://doi.org/10.1088/1757-899X/368/1/012001

[36] Rizal NFAA, Ibrahim MF, Zakaria MR, Bahrin EK, Abd-Aziz S, Hassan MA. (2018). Combination of Superheated Steam with Laccase Pretreatment Together with Size Reduction to Enhance Enzymatic Hydrolysis of Oil Palm Biomass. Molecules, 23. https://doi.org/10.3390/molecules23040811

[37] Azlan N, Mohd, Lin C, Gan S, Basyaruddin M, Rahman A. (2022). Materials Today : Proceedings Effectiveness of Various Solvents In The Microwave-Assisted Extraction of Cellulose from Oil Palm Mesocarp Fiber. Materials Today: Proceedings, Elsevier Ltd. https://doi.org/10.1016/j.matpr.2021.12.086

[38] Olughu, Onu Onu, Tabil, Lope G, Dumonceaux, Dan, T. (2021). Ultrasonic Delignification and Microstructural Characterization of Switchgrass (In Indonesia).

[39] Mukherjee A, Banerjee S, Halder G. (2018). Parametric Optimization of Delignification of Rice Straw Through Central Composite Design Approach Towards Application in Grafting. Journal of Advanced Research, Cairo University. 14, 11–23. https://doi.org/10.1016/j.jare.2018.05.004

[40] Prasetyaningrum A, Ratnawati R, Jos B. (2017). Optimization of the Ozonation Process on Depolymerization of κ-Carrageenan with the Response Surface Method (In Indonesia). Reaktor, 17, 1. https://doi.org/10.14710/reaktor.17.1.1-8

[41] Mostapha M, Azamkamal F, Mohd Salleh K, Amran UA, Gan S, Zakaria S. (2021). Parameter Optimization for Esterified Palm Oil Empty Fruit Bunches (EFB) Cellulose (In Indonesia). Sains Malaysiana, 50, 3719–32. https://doi.org/10.17576/jsm-2021-5012-21

[42] Ramadoss G, Muthukumar K. (2016). Mechanistic Study on Ultrasound Assisted Pretreatment of Sugarcane Bagasse Using Metal Salt With Hydrogen Peroxide for Bioethanol Production. Ultrasonics Sonochemistry, 28, 207–17. https://doi.org/10.1016/j.ultsonch.2015.07.006

[43] Pujokaroni AS, Ohtani Y, Ichiura H. (2020). Ozone Treatment for Improving The Solubility of Cellulose Extracted from Palm Fiber. Journal of Applied Polymer Science, 138, 1–11. https://doi.org/10.1002/app.49610

[44] Cubero MTG, Palacín LG, González-Benito G, Bolado S, Lucas S, Coca M. (2012). An Analysis of Lignin Removal in A Fixed Bed Reactor By Reaction of Cereal Straws With Ozone. Bioresource Technology, 107, 229–34. https://doi.org/10.1016/j.biortech.2011.12.010

[45] Ruan XC, Liu MY, Zeng QF, Ding YH. (2010). Degradation and Decolorization of Reactive Red X-3B Aqueous Solution By Ozone Integrated with Internal Micro-Electrolysisa. Separation and Purification Technology, Elsevier B.V. 74, 195–201. https://doi.org/10.1016/j.seppur.2010.06.005

[46] Chiha M, Hamdaoui O, Baup S, Gondrexon N. (2011). Sonolytic Degradation of Endocrine Disrupting Chemical 4-Cumylphenol in Water. Ultrasonics Sonochemistry, Elsevier B.V. 18, 943–50. https://doi.org/10.1016/j.ultsonch.2010.12.014

[47] Gągol M, Przyjazny A, Boczkaj G. (2018). Effective Method of Treatment of Industrial Effluents under Basic pH Conditions using Acoustic Cavitation – A Comprehensive Comparison with Hydrodynamic Cavitation Processes. Chemical Engineering and Processing - Process Intensification, 128, 103–13. https://doi.org/10.1016/j.cep.2018.04.010

[48] Wang T, Le T, Hu J, Ravindra AV, Xv H, Zhang L et al. (2022). Ultrasonic-Assisted Ozone Degradation of Organic Pollutants in Industrial Sulfuric Acid. Ultrasonics Sonochemistry, Elsevier B.V. 86, 106043. https://doi.org/10.1016/j.ultsonch.2022.106043

[49] Xiong X, Wang B, Zhu W, Tian K, Zhang H. (2019). A Review on Ultrasonic Catalytic Microbubbles Ozonation Processes: Properties, Hydroxyl Radicals Generation Pathway and Potential in Application. Catalysts, 9. https://doi.org/10.3390/catal9010010

[50] Zhao L, Ma W, Ma J, Wen G, Liu Q. (2015). Relationship Between Acceleration of Hydroxyl Radical Initiation and Increase of Multiple-Ultrasonic Field Amount in the Process of Ultrasound Catalytic Ozonation for Degradation of Nitrobenzene in Aqueous Solution. Ultrasonics Sonochemistry, Elsevier B.V. 22, 198–204. https://doi.org/10.1016/j.ultsonch.2014.07.014

[51] Arantes V, Jellison J, Goodell B. (2012). Peculiarities of Brown-rot Fungi and Biochemical Fenton Reaction with Regard to Their Potential as a Model for Bioprocessing Biomass. Applied Microbiology and Biotechnology, 94, 323–38. https://doi.org/10.1007/s00253-012-3954-y

[52] Maqsood HS, Bashir U, Wiener J, Puchalski M, Sztajnowski S, Militky J. (2017). Ozone Treatment of Jute Fibers. Cellulose, Springer Netherlands. 24, 1543–53. https://doi.org/10.1007/s10570-016-1164-y

[53] Martínez AT, Rencoret J, Nieto L, Jiménez-Barbero J, Gutiérrez A, Del Río JC. (2011). Selective Lignin and Polysaccharide Removal in Natural Fungal Decay of Wood as Evidenced by in Situ Structural Analyses. Environmental Microbiology, 13, 96–107. https://doi.org/10.1111/j.1462-2920.2010.02312.x

[54] Chieng BW, Lee SH, Ibrahim NA, Then YY, Loo YY. (2017). Isolation And Characterization Of Cellulose Nanocrystals From Oil Palm Mesocarp Fiber. Polymers, 9, 1–11. https://doi.org/10.3390/polym9080355

[55] Okahisa Y, Furukawa Y, Ishimoto K, Narita C, Intharapichai K, Ohara H. (2018). Comparison of Cellulose Nanofiber Properties Produced from Different Parts of The Oil Palm Tree. Carbohydrate Polymers, 198, 313–9. https://doi.org/10.1016/j.carbpol.2018.06.089

[56] Rosli NA, Ahmad I, Abdullah I. (2013). Isolation and Characterization of Cellulose Nanocrystals from Agave Angustifolia Fibre. BioResources, 8, 1893–908. https://doi.org/10.15376/biores.8.2.1893-1908

[57] Bogolitsyn K, Parshina A, Aleshina L. (2020). Structural Features of Brown Algae Cellulose. Cellulose, Springer Netherlands. 27, 9787–800. https://doi.org/10.1007/s10570-020-03485-z

[58] Gonzalez M, Pereira-Rojas J, Villanueva I, Agüero B, Silva I, Velasquez I et al. (2022). Preparation and Characterization of Cellulose Fibers from Meghatyrsus Maximus: Applications in its Chemical Derivatives. Carbohydrate Polymers, Elsevier Ltd. 296, 119918. https://doi.org/10.1016/j.carbpol.2022.119918

[59] Reddy KO, Maheswari CU, Dhlamini MS, Mothudi BM, Kommula VP, Zhang J et al. (2018). Extraction and Characterization of Cellulose Single Fibers from Native African Napier Grass. Carbohydrate Polymers, Elsevier Ltd. 188, 85–91. https://doi.org/10.1016/j.carbpol.2018.01.110

[60] Zhang K, Si M, Liu D, Zhuo S, Liu M, Liu H, et al. (2018). A Bionic System with Fenton Reaction and Bacteria as A Model for Bioprocessing Lignocellulosic Biomass. Biotechnology for Biofuels, BioMed Central. 11, 1–14. https://doi.org/10.1186/s13068-018-1035-x

[61] Valladares-Diestra KK, Porto de Souza Vandenberghe L, Zevallos Torres LA, Nishida VS, Zandoná Filho A, Woiciechowski AL, et al. (2021). Imidazole Green Solvent Pre-Treatment as A Strategy for Second-Generation Bioethanol Production from Sugarcane Bagasse. Chemical Engineering Journal, 420. https://doi.org/10.1016/j.cej.2020.127708

[62] Rasid NSA, Zainol MM, Amin NAS. (2021). Synthesis and Characterization of Carboxymethyl Cellulose Derived from Empty Fruit Bunch. Sains Malaysiana, 50, 2523–35. https://doi.org/10.17576/jsm-2021-5009-03

[63] Zulnazri Adha, muhammad ikhwanul, Dewi R. (2022). Experimental Study of Cellulose Extraction from Oil Palm Empty Fruits Bunches. 7–14. https://doi.org/10.31284/j.iptek.2022.v26i1.2

[64] Lu H, Liu S, Shi Y, Chen Q. (2022). Efficient Delignification of Sugarcane Bagasse by Fenton Oxidation Coupled with Ultrasound-Assisted NaOH for Biotransformation from Agaricus Sinodeliciosus Var. Chaidam. Chemical Engineering Journal, Elsevier B.V. 448, 137719. https://doi.org/10.1016/j.cej.2022.137719

[65] Ren X, Zhang X, Liang Q, Hou T, Zhou H. (2017). Effects of Different Working Modes of Ultrasound on Structural Characteristics of Zein and ACE Inhibitory Activity of Hydrolysates. Journal of Food Quality, 2017. https://doi.org/10.1155/2017/7896037

CORONARY HEART DISEASE CLASSIFICATION USING IMPROVED PENGUIN EMPEROR OPTIMIZATION-BASED LONG SHORT TERM MEMORY NETWORK

RAJESHWARI MARAMGERE RAMAIAH[1*] AND KAVITHA KUNTAEGOWDANALLI SRIKANTEGOWDA[2]

[1]*Department of Computer Science & Engineering, Visvesvaraya Technological University, Belagavi-590018, Karnataka, India*
[2]*Department of Computer Science and Engineering, Dayananda Sagar college of Engineering, Bengaluru, Karnataka 560078, India*

[*]*Corresponding author: rajeshwari_maya1@yahoo.com*

ABSTRACT: Ventricular fibrillation (VF) is the most life-threatening and dangerous type of Cardiac Arrhythmia (CA), with a mortality rate of 10-15% in a year. Therefore, early detection of cardiac arrhythmia is important to reduce the mortality rate. Many machine learning algorithms have been proposed and have proven their usefulness in the classification and detection of heart problems. In this research manuscript, a novel Long Short Term Memory (LSTM) classifier with Improved Penguin Optimization (IPEO) is implemented for VF classification. The IPEO is used in finding optimal hyperparameters that overcome the overfitting problem. The presented model is tested, trained, and validated using two standard datasets that are available publicly: Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) and the China Physiological Signal Challenge (CPSC) 2018 dataset. Both of them consist of ECG recordings for five seconds of coronary heart disease (CHD) patients. Furthermore, Fuzzy C-Means and Enhanced Fuzzy Rough Set method (FCM-ETIFRST) are used for feature selection to extract informative features and to cluster membership degree, non-membership degree, and hesitancy degree. On the MIT-BIH dataset, the proposed model achieved accuracy, sensitivity, specificity, precision, and Matthews's correlation coefficient (MCC) of 99.75%, 98.29%, 98.39%, 98.35%, and 97.79% respectively. On the CPSC 2018 dataset, the proposed model achieved accuracy of 99.79%, sensitivity of 99.11%, specificity of 98.20%, precision of 99.43%, and MCC of 98.57%. Hence, the results proved that the proposed method provides better results in the classification of VF.

ABSTRAK: Pemfibrilan Ventrikel (VF) adalah ancaman nyawa nombor satu dan jenis Aritmia Jantung (CA) berbahaya dengan kadar kematian 10-15% setahun. Oleh itu, pengesanan awal Aritmia Jantung sangat penting bagi mengurangkan kadar kematian. Terdapat banyak algoritma pembelajaran mesin yang telah dicadangkan dan terbukti berkesan dalam pengelasan dan pengesanan sakit jantung. Kajian ini mencadangkan kaedah baru pengelasan Memori Ingatan Jangka Panjang Pendek (LSTM) dengan Pengoptimuman Penambahbaikan Penguin (IPEO) yang dilaksanakan bagi klasifikasi VF. IPEO digunakan bagi mencari hiperparameter yang dapat mengatasi masalah padanan berlebihan. Model yang dicadangkan diuji, dilatih dan disahkan menggunakan dua dataset piawai yang dapat diperoleh secara terbuka: Institut Teknologi Hospital Massachusetts-Beth Israel (MIT-BIH) dan Cabaran Signal Psikologi Cina 2018 (CPSC). Kedua-dua data ini mempunyai rakaman ECG selama lima saat daripada pesakit Penyakit Jantung Koronari (CHD). Malah, kajian itu turut menggunakan Purata-C Kabur dan Kaedah Set

Kasar Kabur Dipertingkat (FCM-ETIFRST) untuk pemilihan bagi mengekstrak ciri-ciri dan mengelaskan kelompok tahap keahlian, bukan ahli dan tahap keraguan. Bagi dataset MIT-BIH, model yang dicadangkan mencapai ketepatan, tahap sensitif, tahap spesifik, kejituan dan pekali kaitan Matthews (MCC) sebanyak 99.75%, 98.29%, 98.39%, 98.35%, dan 97.79% masing-masing. Bagi dataset CPSC 2018 pula, model yang dicadangkan mencapai ketepatan sebanyak 99.79%, 99.11% tahap sensitif , 98.20% tahap spesifik, 99.43% kejituan dan 98.57% MCC. Oleh itu, dapatan kajian membuktikan kaedah yang dicadangkan menunjukkan keputusan lebih baik dalam pengelasan VF.

KEYWORDS: *coronary heart disease; fuzzy c-means and enhanced fuzzy rough set method; improved penguin emperor optimization; long short term memory; ventricular fibrillation*

1. INTRODUCTION

As per the World Health Organization (WHO), coronary heart disease (CHD) is a substantial global epidemic and the most common cause of death worldwide, accounting for one in every three deaths. Coronary heart disease (CHD) is a substantial global epidemic and the most common cause of death worldwide, accounting for one in every three deaths. In 2002, 16.7 million cardiac disease (CD)-related deaths were reported, and this is expected to rise to 23.3 million by 2030 [1]. Age, obesity, smoking, exercise, hypertension, diabetes, and high blood cholesterol are all risk factors for CD. Heart disease diagnosis is conducted using a variety of medical tests, a medical history, and an examination of the patient's lifestyle. There are many variables to consider when diagnosing heart diseases, thus a specialist is usually involved [2-3]. The contraction of the ventricles that exceeds 100 bpm causes ventricular tachycardia (VT), and when the ventricles' contraction exceeds 500 bpm, it is referred to as ventricular fibrillation (VF), which is the most abnormal cardiac arrest. VT and VF are the two subsets of ventricular arrhythmias (VAs) [4]. Nearly all countries face a significant coronary artery disease (CAD) burden, due to limited resource availability for providing comprehensive health care of CAD patients, and insufficient awareness campaigns on such diseases [5]. Various machine learning methods are developed to detect CHD at an early stage using different software libraries and platforms to extract huge amounts of information from large datasets. The predictive analysis of CHD includes pre-processing the data collected from datasets, preparing the data for implementing suitable algorithms, training, testing, and validating the chosen model, and making final predictions with fine-tuned parameters [6,7]. Atrial fibrillation (AF) detection has been done in many types of research depending on the activities of the atrial region of the heart, which also includes the detection of P-wave, the entropy of wavelet samples, and the detection of F-waves [8]. To detect P-wave or R-R wave irregularities with manual instructions on long-term ECG recordings, AF uses currently available algorithms. Any change to P waveform in shape or blunt from Lead 2 and 4 results in irregularity, and it directly relates to the condition of atrial activity since it is known as a depolarization wave of the atrial region [9-10]. In this work, a design of methodology to detect and classify CHD is developed by using ECG signals of MIH-BIT and CPSC 2018 dataset. The major contributions of the paper are described as follows:

- To resolve the problem of class imbalance and dimensionality curse, the method of FCM-ETIFRST has been included in the process of feature selection.

- For accurate classification of MIT-BIH arrhythmia and CPSC 2018 datasets with reduced overfitting issues, an efficient LSTM classification has been considered.

- An Improved Penguin Emperor Optimization (IPEO) has been proposed in this research for finding the optimal hyperparameters to operate LSTM without overfitting issues. This is done by improving the exploitation and exploration stage of PEO by levy flight and Gaussian mutation mechanism.

2. LITERATURE REVIEW

Li et al. [11] developed a novel feature fusion framework that used phonocardiogram (PCG) signals of 175 subjects to identify CAD features. A sum of 110 features from different domains were extracted, reduced, and selected. The obtained images were taken as CNN input for feature learning. The feature selection and the features of deep learning were then combined and led to a multilayer perceptron for classification purposes. This proposed framework outperformed deep learning features and multi-domain features with the highest accuracy, sensitivity, and specificity. Too many features inhibit the model learning process and result in poor generalization, which is a limitation of this work. Future work can be focused on improving the detection accuracy of CAD by using dynamic features and multi-channel PCG signals. Nguyen et al. [12] developed a stacking method to predict atrial fibrillation (AF) from ECG signals and perform statistical segment-based feature recognition, which was done by Support Vector Machine (SVM), the segmentation units produced by a convolutional neural network. To validate this method, the ECG dataset from the PhysioNet/Computing in Cardiology Challenge 2017 was used, which contained 8528 ECG recordings. The proposed method applied to the same dataset and metric outperformed state-of-the-art methods with a high F1 score. The limitation of this method was the nature of the dataset that was chosen. The AF signal segment was not included with AF due to the absence of information about AF or the occurrence of other ECG signals. The proposed method was applied to other medical signal-related problems. Jahmunah et al. [13] proposed an automated system (AS) using automated categorization of ECG signals with CNN, which classified signals into normal, CAD, myocardial infarction (MI), and congestive heart failure (CHF) classes. The proposed GaborCNN achieved better performance with 98.5% accuracy. The use of GaborCNN resulted in low computational complexity which was an added advantage for this work. Furthermore, the system was validated with a large database and was seen to exhibit the highest potential to assist clinicians in screening for CVDs using ECG signals. Only a few subjects and smaller datasets were used for CAD and CHF, which was the limitation of this work. To improve the classification accuracy of GaborCNN, more data needs to be trained to the network in the future, so that the onset of CAD is detected early and prevented from progressing to MI or CHF.

Tseng et al. [14] developed a method for VF prediction which was a novel deep-learning method. For training and validation of the data, ECGs from MIT-BIH datasets were used. The estimated results demonstrated that the proposed two-dimensional short-time Fourier transform (2D STFT)/continuous wavelet transform (CWT) convolutional neural network (CNN) model achieved the highest recall and accuracy. The proposed model was also compared with 1D CNN and 2D time-domain models and again achieved the highest accuracy. The database did not identify concomitant diseases, and the influence of these diseases on ECG signals was unknown, which was one of the limitations of this work. In the future, larger datasets should be used to build a more accurate model. Additional data from conventional and portable hospital vital sign monitors, as well as wearable devices should be collected. Panigrahy et al. [15] proposed a novel approach to detect the VF rhythm that involved algorithms namely: Support Vector Machine (SVM), Adaptive Boosting (AdaBoost) and Differential Evolution (DE). These algorithms were implemented with the help of optimal variable combinations. The training and test datasets were taken and

validated from three databases namely the arrhythmia database, CUDB database, and MIT-BIH malignant ventricular arrhythmia databases for this methodology. The proposed approach achieved the highest accuracy, sensitivity, and specificity in detecting VF rhythm. Further, the proposed method was suitable for implementation during real time detection of VF rhythm. Çınar et al. [16] developed a deep learning architecture with SVM for the classification of Normal Sinus Rhythm (NSR) which was based on hybrid Alex-net along with ECG signals of Abnormal Arrhythmia (ARR) and Congestive Heart Failure (CHF). This approach was implemented on 192 ECG signals, out of which 96 were arrhythmias, 30 were CHFs, and 36 were NSRs. SVM and KNN algorithms were implemented on ARR, CHR, and NSR signals for classification purposes. Later on, these signals were classified using Long Short-Time Memory (LSTM). Chen et al. [17] proposed a CNN model for CA classification from ECG signals of the China Physiological Signal Challenge (CPSC) 2018 dataset. The dataset provided 6,877 recordings with a 12-lead ECG dataset, in which different types of diagnoses for 476 patients were predicted. The proposed technique obtained an overall highest F1-score, and first rank in the classification-and-challenge competition. This approach still needed to be filtered to achieve an equal level of performance on other datasets, which is a limitation of this work. Various CA types and wearable ECG devices are possible to be developed for use with this approach in the future.

3. METHODOLOGY

The framework for the proposed work is a combination of feature selection with FCM-ETIRST and LSTM classification with Improved Emperor Penguin Optimization (IEPO) hyperparameter optimization for reducing the problems of dimensionality curse and overfitting in CHD analysis and it is illustrated in Fig. 1. Because of the nature of time-step order in CNN with its several hidden layers, it is used to extract the features from the complex data of ECG signals. The characteristics of time will be taken by the layers of LSTM while classification is skipped at the stage of feature extraction. To optimize the computational resources and make more efficient detection of CHD, the proposed IPEO algorithm is utilized to optimize hyperparameters of LSTM. The steps of the proposed methodology workflow are designed as shown in Fig. 1.
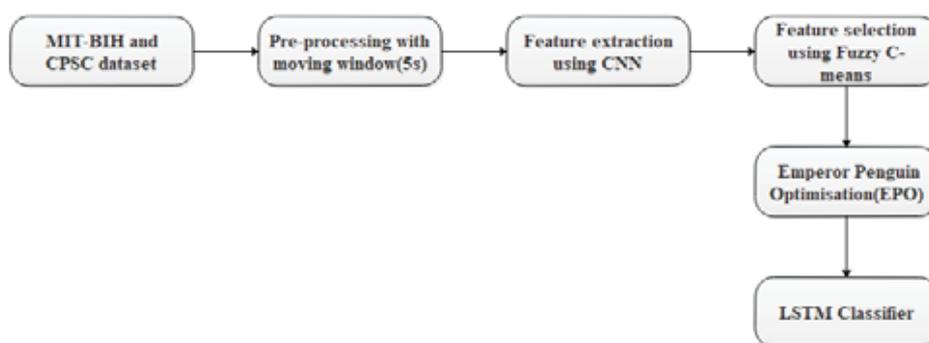
Fig. 1: Flowchart of the proposed IEPO based LSTM model.

3.1 Data Pre-processing

ECG recordings of VF patients are considered from two datasets namely MIT-BIH and CPSC 2018. The MIT-BIH dataset consists of 23 publicly available ECG recordings which include four types of rhythms: AF, VF, AV junction, and Normal. However, in this work, only VF signals are classified and sampled at a frequency of 250Hz. The proposed algorithm is implemented for detecting VF by pre-processing ECG signals in a window length of 5

seconds. The CPSC 2018 also consists of ECG signals in which all forms of arrhythmias are available. All signals are segmented into a number of episodes, where each episode is labeled by the clinician and divided into training, test, and validation datasets. Two leads are selected for this work from the dataset, and the data is divided into VF and non-VF categories. The ECG signals from the CPSC dataset are sampled at a frequency of 500Hz and segmented into 5-second episodes. A lead 1 sample signal from the CPSC dataset is considered, as shown in Fig. 2, and is segmented into normal signal and VF signal as shown in Fig. 3(a) and Fig. 3(b). Lead 2 normal signal and VF signal are shown in Fig. 4(a) and Fig. 4(b) respectively.
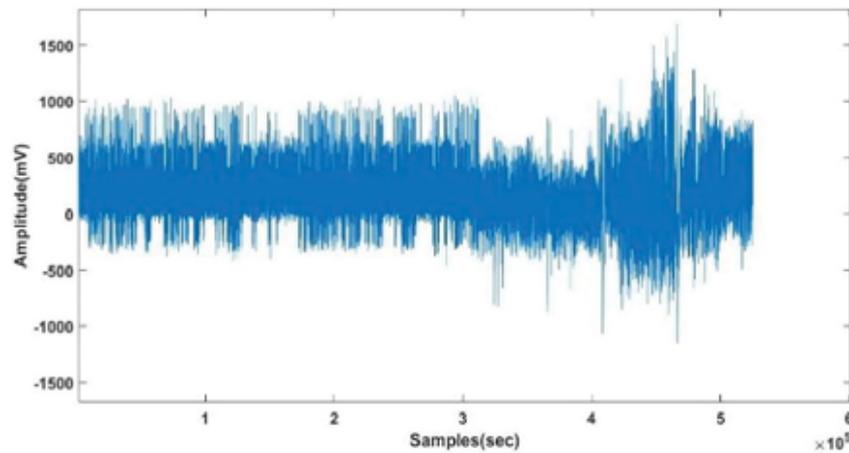
Fig. 2: Sample signal from MIT-BIH dataset.
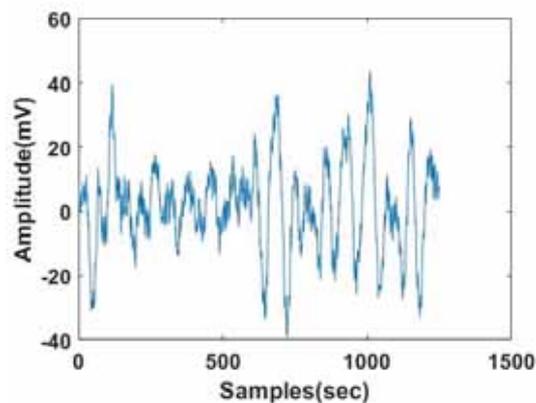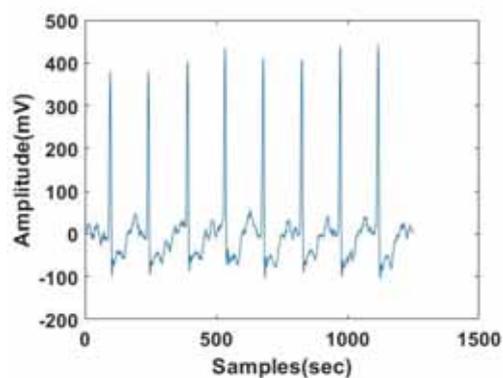
(a)   (b)

Fig. 3. Lead 1 sample signal (a) Normal, (b) VF.
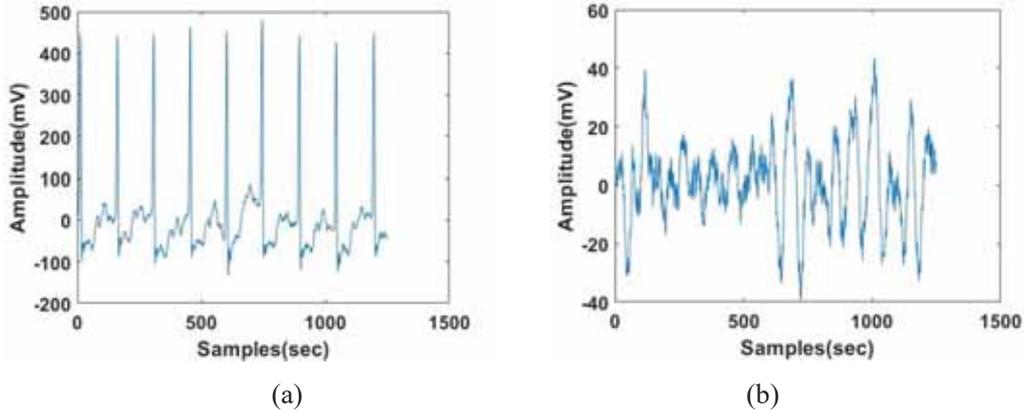
Fig. 4: Lead 2 sample signal (a) Normal, (b) VF.

3.2 Feature Extraction

Here, the extraction of features is done using CNN layer architecture which is based on the 1-D framework. Timing characteristics and sequence of data can be retained with the representation of CNN-related features in the proposed framework. The following features are to be extracted from the proposed CHD framework.

3.2.1 1-D Approximation and Details Coefficients

By use of low pass and high pass filtering operations on the transform coefficients group, a set of coefficients namely, 'approximation from low pass' and 'detail from high pass' is obtained. Consider the parameters of coefficients for approximation and details to be CA_j and CD_j correspondingly, in case of a stationary signal, the transform their properties for the level of j as mentioned in Eq. (1) and Eq. (2).

$$E(CA_j) = \frac{\sqrt{2^{-j}}}{N} \int f(t) \tag{1}$$

$$E(CD_j) = \sqrt{2^{-j}} \int f(t) E_n[\varphi(2^{-j}t - n)]dt \tag{2}$$

Where CA_j and CD_j are parameters of coefficients at the level of j, E_n represents the wavelet transform of the signals at n-level coefficients, $\varphi(t)$ and $f(t)$ represents the wavelet coefficient transformation functions. N represents the total number of wavelets obtained from the signal.

3.2.2 Hjorth Activity

Signal power is represented using an activity parameter that is time function variance, as given in Eq. (3). This activity denotes the power spectrum surface in the frequency domain.

$$Activity = var(y(t)) \tag{3}$$

where the signal is denoted using $y(t)$.

3.2.3 Hjorth Mobility

The mean frequency or power spectrum of standard deviation proportion is represented as a mobility parameter. A first derivative variance of the square root of signal $y(t)$ divided by signal variance $y(t)$, which is explained in Eq. (4).

$$Mobility = \sqrt{\frac{var\left(\frac{dy(t)}{dt}\right)}{var(y(t))}} \tag{4}$$

3.2.4 Hjorth Complexity

The frequency change is denoted using the complexity parameter. The signal's similarity and a pure sine wave are compared using the parameter, where the value converges to 1 if the signal is similar, as given in Eq. (5).

$$Complexity = \frac{Mobility\left(\frac{dy(t)}{dt}\right)}{Mobility(y(t))} \tag{5}$$

3.2.5 Mean Curve Length (MCL)

Katz fractal dimensions are measured using MCL and this measures EEG signals activity, as given in Eq. (6).

$$CL[n] = \frac{1}{N}\sum_{m=k-N+1}^{k}|x[m] - x[m-1]| \tag{6}$$

where the epoch last sample is denoted as k, the window length is denoted as N, and the EEG time series is denoted as $x[m]$.

3.2.6 Mean Teager Energy (MTE)

EEG research was widely carried out using MTE feature, as given in Eq. (7).

$$MTE[k] = \frac{1}{N}\sum_{m=k-N+3}^{k}(x[m-1]^2 - x[m]x[m-2]) \tag{7}$$

where the last epoch sample is denoted as k, the window length is denoted as N, and the EEG time series is denoted as $x[m]$.

3.2.7 Zero Crossing Rate

The rate of signal changes from negative to zero to positive or positive to zero to negative is denoted as Zero-Crossing Rate (ZCR), as in Eq. (8).

$$ZCR = \frac{1}{T-1}\sum_{t=1}^{T-1}1_{R<0}(S_t S_{t-1}) \tag{8}$$

where the indicator function is denoted as $1_{R<0}$ and the length of the signal is denoted as S.

The above-mentioned features are extracted and forwarded to the process of feature selection for further enhancement of dimensionality and to enable better output on the overfitting of classifiers.

3.3 Feature Selection

The FCM – ETIFRST feature selection method is applied to the extracted features to select relevant features for arrhythmia classification.

3.3.1 Fuzzy C Means

Feature selection on the proposed method is done based on the clustering of Fuzzy C-Means and Enhanced Fuzzy Rough Set method. The FCM method has the advantage of producing better results for overlapped datasets than the k-means method. The data point is an assigned membership to every cluster to form the appropriate cluster [18]. The data point is applied with membership based on the distance between the data point and cluster center c. If more data is near to c, then the cluster has more data point membership.

The estimated membership is increased for every iteration. An enhanced fuzzy C-means algorithm is implemented in the proposed method. The algorithm is given as follows:

<u>Step 1:</u> Clusters number c varies from 2 to c_{max}. A certain value c is considered and initial class prototypes are selected and use $\epsilon > 0$.

<u>Step 2:</u> Image ς of new linearly weighted sum in terms of ς_k is calculated using Eq. (9).

$$\varsigma_k = \frac{1}{1+\alpha}\left(x_k + \left(\frac{\alpha}{N_R}\right)\sum_{j \in N_k} x_j\right) \tag{9}$$

where k^{th} pixel gray value of ς are ς_k and x_k (8-bit resolution is usually encoded for gray value) respectively. The neighbor x_j and the set of neighbours N_k is present in a window which is nearby x_k. The mean-filtered pixel value is $\sum_{j \in N_k} x_j$ where neighbor terms' effect controls α with a value approximately near to 1.

<u>Step 3:</u> Image matrix ς is updated using Eq. (10).

$$f_{il} = \frac{(\varsigma_l - v_i)^{\frac{-2}{s-1}}}{\sum_{j=1}^{c}(\varsigma_l - v_j)^{\frac{-2}{s-1}}} \tag{10}$$

The ith cluster prototype is v_i and f_{il} is gray value l of fuzzy membership that is with respect to cluster i. The l value varies from 1 to the number of gray levels denoted as d (gray level has a maximum value of 256 levels) of the image and each fuzzy membership of s weighted component parameter that estimates the final classification of fuzziness.

<u>Step 4:</u> Prototype v_i is calculated using Eq. (11).

$$v_i = \frac{\sum_{i=1}^{d} x_l f_{il}^S \varsigma_l}{\sum_{i=1}^{d} x_l f_{il}^S} \tag{11}$$

Convergence criteria are measured as $|v_{new} - v_{old}| < \epsilon$ by repeating step 3 and 4, cluster prototypes hold vectors v that is represented as $v = [v_1, v_2 \ldots \ldots v_c]$.

3.3.2 Enhanced Fuzzy Rough Set Feature Selection

The FCM has a wide range of applications and is significantly applied in numerous fuzzy clustering methods. Every sample point of membership degree to the class center is used as objective function optimization [19]. Euclidean distance is applied for the objective function of the sample point and clustering center. Solving the non-similarity index value function is the minimum value of every clustering center. The generalization is given in Eq. (12).

$$J(U, c_1, c_2, \ldots, c_c) = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c} \sum_{j}^{n} u_{ij}^m d_{ij}^2 \tag{12}$$

where the weighted index number is denoted as m, the jth sample point of ith clustering center of Euclidean distance is $d_{ij} = k_{ci} - x_{jk}$, the fuzzy set of clustering center, and u_{ij} is between 0 and 1. The constraint formula of the Lagrangian multiplier is constructed based on the derivate of input parameters in Eq. (13) and Eq. (14) to reach the minimum.

$$c_i = \frac{\sum_{j=1}^{n} u_{ij}^m x_j}{\sum_{j=1}^{n} u_{ij}^m} \tag{13}$$

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{d_{ij}}{d_{kj}}\right)^{\frac{2}{m-1}}} \tag{14}$$

The FCM outputs are centers c and membership matrix U. Every object degree contains U, which belongs to centers c. A fuzzy rough set of lower approximations is given in Eq. (15).

$$sim_a(x_i, x_j) = 1 - \frac{}{\sqrt{\alpha\left(\mu_a(x_i) - \mu_a(x_j)\right)^2 + \beta\left(v_a(x_i) - v_a(x_j)\right)^2 + \gamma\left(\pi_a(x_i) - \pi_a(x_j)^2\right)}}$$ (15)

where object x_i of hesitancy degree, non-membership, and membership are denoted as $\pi_a(x_i)$, $v_a(x_i)$ and $\mu_a(x_i)$, respectively. The weighted factors are α, β and γ.

Each pixel membership degree is normalized and initialized as in Eq. (16).

$$\mu_D(x_{ij}) = \frac{x_{ij} - (x_{ij})_{min}}{(x_{ij})_{max} - (x_{ij})_{min}}$$ (16)

where the image pixel of minimum and maximum intensity is denoted as $(x_{ij})_{max}$ and $(x_{ij})_{min}$. Complexity is reduced in the normalization process, which calculates in between 0 and 1.

Non-membership value is present in the new method in uncertain presence. A high grade of certainty is measured based on observations when the membership value is near 0 or 1 and the membership value is near 0.5 for a high grade of uncertainty. Non-membership value is measured in Eq. (17)

$$v_D(x_{ij}) = \left(1 - \mu_D(x_{ij})\right)\left(exp\left(\frac{-\mu_D(x_{ij})}{a\sigma}\right)\right)^{\frac{1}{a}}$$ (17)

where $(\mu_D(x_{ij}))$ membership value measures standard deviation σ, which is in the range of 0.39 – 0.41. The hesitation degree is measured using $\pi_D(x_{ij}) = 1 - \mu_D(x_{ij}) - v_D(x_{ij})$.

Every feature of fuzzy clustering gets every object membership degree, as given in Eq. (18).

$$\mu R_a(x) = U_{FCM(a)}[x]$$ (18)

Model equality is based on equivalence relations; approximate similarity or equality is measured based on fuzzy equivalence. Fuzzy equivalence relation R is used, as in Eq. (19).

$$R(x, y) = \frac{a - a|x - y| + bmin(x,y)}{a - (a-1)|x - y| + bmin(x,y)}$$ (19)

The size of the opening on one side is denoted as value b. If b is equal to 0, then the function makes sure to balance two objects, otherwise it is not able to balance the two objects.

3.4 Improved Emperor Penguin Optimization for Hyperparameter Optimization

In IEPO, three strategies are incorporated into EPO to maintain a suitable balance between exploration and exploitation to solve the drawback of the standard EPO [20]. These strategies are opposition-based learning, Levy flight, and Gaussian mutation. Equation (20) contains the fundamental formula for the EPO algorithm, which moves the penguins toward the center to locate the ideal value by measuring their distance from one another in the population. However, it is simple to get caught in the local optimal problem when the EPO algorithm optimizes the hyperparameters of LSTM. The strategy strengthens the bonds between penguins and improves each penguin's capacity for jumping to address this issue. The hyperparameters taken for LSTM have been given in the following Table 1.

The distance between the penguins is represented mathematically, as shown in Eq. (20), with the term Depth. The proposed method incorporates the Gaussian Mutation (GM) into Eq. (20) because the EPO's unpredictability is a vulnerability.

Table 1: Hyperparameters for Long Short Term Memory

Hyperparameters	Range
Number of hidden layers	3-10
Dropout	0.2 - 0.5
Decay rate	0.2-0.9
Learning rate	0.0 - 0.1
Momentum	0.5 - 0.9
Batch size	32,64,128,256
EPOCHS	10-100

The GM has excellent randomization capabilities, which it can use to strengthen the penguin relationships and accomplish the goal of leaving the local optimal. The following formula can be used to represent the updated mathematical model:

$$P_{eq}(x + 1) = P(x) - A \times Dep \oplus G(\alpha) \qquad (20)$$

The individual penguins are then observed to move slowly while readily guiding the swarm to the regional ideal. The suggested strategy enhances the individual penguin's update position formula by utilizing the Levy flight's strong stochastics. The Levy flight expands the swarm's search area while simultaneously improving each penguin's capacity to jump. The best penguin can therefore swiftly arrive at the greatest answer. This is how it is formulated:

$$P \, levy \, eq(x + 1) = P(x) \times Levy - A \times Dep \oplus G(\alpha) \qquad (21)$$

where, $P \, levy \, eq$ represents the updated position of the ith solution. Through the Levy flight mechanism, a new candidate solution is generated, which increases the knowledge diffusion across objects and yields a better solution, as Levy flight is a random process and the jump size leads to the Levy probability distribution function. The best penguin may also leave the search arena due to the addition of Levy flight and the strong randomness of Gaussian mutation. By doing this, the swarm might avoid falling into the infinite cycle and also the best solution from being updated. The Opposition Based Learning (OBL) is added to the EPO algorithm to reduce the upgraded algorithm's ability to jump. The OBL broadens the search domain's exploration, which in turn diversifies the swarm. It can be created using Eq. (22)

$$P_{eq} = P_{max} + P_{min} - P_{best} + r(P_{best} - P_{eq}) \qquad (22)$$

where P_{eq} is the i_{th} opposite penguin's location within the search domain. The i_{th} variables lower and upper bounds are P_{max} and P_{min} respectively. P_{best} is the location of the best penguin, and r is an element-filled random vector within the range$[0,1](0,1)$. P is the ith penguin in the population's position vector. Then, based on the fitness of the opposing sites, the best penguin is modified. To make the penguin change more quickly and discover the best value for the chosen hyperparameters, the new technique delivers greater performance and higher convergence velocity.

3.5 Classification using LSTM

Due to its outstanding performance in the extraction of temporal and spatial variables, the classification network used in the framework is an LSTM and its architecture is shown in Fig. 5. According to Fig. 5, the model is made up of an input layer, a single LSTM cell, a hidden layer, a normalization layer, and a classification layer.

Performance is improved by the idea of integrating the most recent data along with older data to predict the subjects one step ahead. With the benefits of a hidden layer self-feedback mechanism, the LSTM model solves the long-term dependence problem [21]. Gates of input state, output state, and forget state in the LSTM model are known to be three unique gates that are used to update data that are meant to be stored in the memory cell.
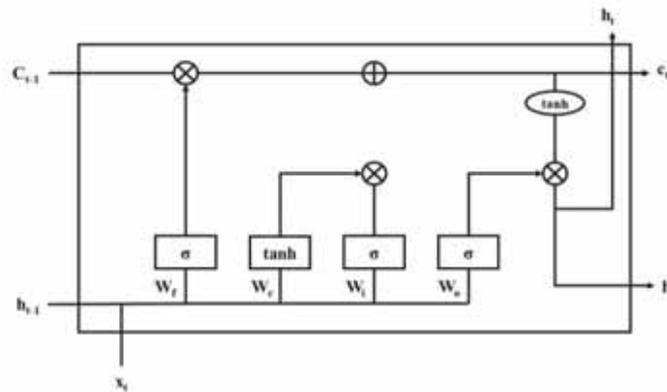
Fig. 5: LSTM architecture.

The input feature for LSTM is denoted and it contains the characteristics obtained via FCM-ETIRST. The previous instant output of the LSTM cell is marked as h_t, a cell of the memory unit is represented, and the output of the LSTM cell is indicated for the unit of LSTM input data at a time. The following steps illustrate the LSTM unit's process.

1) As shown in Eq. (1), the cell of memory for the candidate, the matrix of weight and the bias are assessed in (23).

$$\widetilde{c}_t = tanh\ (W_c.\,[h_{t-1}, x_t] + b_c) \tag{23}$$

2) The gate of the input state modifies the memory state value depending upon the data input being used at the time; the function activation of sigmoid, matrix of weight, and bias are shown in Eq (24).

$$i_t = \sigma(W_i.\,[h_{t-1}, x_t] + b_i) \tag{24}$$

3) By using past data, the gate of forget state modifies the memory state value; the bias gate is indicated by the weight matrix which is given in Eq (25).

$$f_t = \sigma(W_i.\,[h_{t-1}, x_t] + b_f) \tag{25}$$

4) As shown in Eq (26), the value of the most recent LSTM unit is calculated by measuring the value of the current memory state.

$$c_t = f_t * c_{t-1} + i_t * \widetilde{c}_t \tag{26}$$

where the dot product is represented by "*". Depending upon the state values of the last and candidate cells, the input state and forget state gates regulate the updated memory cell.

5) The matrix of weight in the layer W_0, the bias b_0, and the output value of the gate controls the output of the memory cell value, as shown in Eq. (27)

$$Eq.o_t = \sigma(W_0.\,[h_{t-1}, x_t] + b_0) \tag{27}$$

6) As shown in Eq. (28), the output of the LSTM is measured as:

$$h_t = o_t * tanh\,(c_t) \tag{28}$$

According to the 3 gates of control and memory state, the LSTM continuously stores, reads, resets, and updates the data. Because the LSTM internal parameter is shared, the weight matrix's dimensions are changed to regulate the size of the output unit. The LSTM's input and feedback are separated by a significant amount of time. The internal state of the memory cell's architecture maintains a constant error flow, and the gradient doesn't explode or vanish [22]. The LSTM classifier predicts the type of arrhythmia present in the CHD and those results are evaluated in the following section.

4. RESULTS

In this work, the proposed modified model is replicated using enhanced LSTM with the system requirements. The performance of the modified model is analyzed by means of sensitivity, specificity, accuracy, F-score, and MCC on the actual educational dataset. The feature length of 506 is obtained after feature selection. Specificity is defined as the test to appropriately recognize the areas that are negatively less in learning skills. Sensitivity is defined as the test to correctly identify the regions that are positively high in learning skills. Further, accuracy is the most important performance measure that is utilized in the LMS, where it is the ratio of correctly predicted observations to the total observations. Accuracy, sensitivity, specificity, and MCC are mathematically defined in the Eqs. (29) to (33),

$$Specificity = \frac{TN}{FP+TN} \times 100 \tag{29}$$

$$Sensitivity = \frac{TP}{FN+TP} \times 100 \tag{30}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \tag{31}$$

F-score is determined as the harmonic mean of the model's recall and precision, and its mathematical equation is followed in Eq. (32).

$$F - score = \frac{2TP}{2TP+FP+FN} \times 100 \tag{32}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \times 10 \tag{33}$$

4.1 Quantitative Analysis

The performance analysis of enhanced LSTM without augmentation in accuracy, sensitivity, specificity, F-score, and MCC, with and without feature selection, are shown in Tables 2, 3, 4 and 5 respectively. Table 2 demonstrates the quantitative analysis of different classifiers with and without feature selection for the MIT-BIH dataset. Table 3 demonstrates the quantitative analysis of different classifiers with and without feature selection for the CPSC 2018 dataset.

The performance measure of classifiers on the MIT-BIH dataset with feature selection is highest compared to those without feature selection. The accuracy, sensitivity, specificity, precision, and MCC of Neural Network (NN), RNN (Recurrent Neural Network), GAN (Generative Adversarial Network), and DNN (Deep Neural Network) are measured and compared with the proposed IEPO-based LSTM. The obtained results show that the proposed IEPO based LSTM achieves the highest values with the accuracy of 99.75%, the

sensitivity of 98.29%, specificity of 98.29%, the precision of 98.35%, and MCC of 97.79% while considering feature selection compared to the other classifiers. The simulation results of the proposed method by varying different classifiers on MIT-BIH dataset, with and without feature selection are specified in Fig. 6 and Fig. 7.

Table 2: Simulation results of IEPO-based LSTM by varying the classifiers for the MIT-BIH dataset

Without Feature Selection Results				
Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	MCC (%)
NN	88.99	87.53	89.44	88.91	86.5
RNN	90.49	88.63	90.84	90.71	89.4
GAN	91.99	90.53	93.74	92.81	90.4
DNN	93.39	91.73	95.64	94.81	92.1
IEPO based LSTM	96.29	95.83	95.74	95.51	94.3
With Feature Selection Results				
NN	90.69	89.53	92.14	91.31	87.7
RNN	92.32	94.18	95.61	92.71	92.34
GAN	92.25	92.28	89.98	91.43	89.91
DNN	90.86	94.53	93.21	93.04	90.56
IEPO based LSTM	99.75	98.29	98.29	98.35	97.79
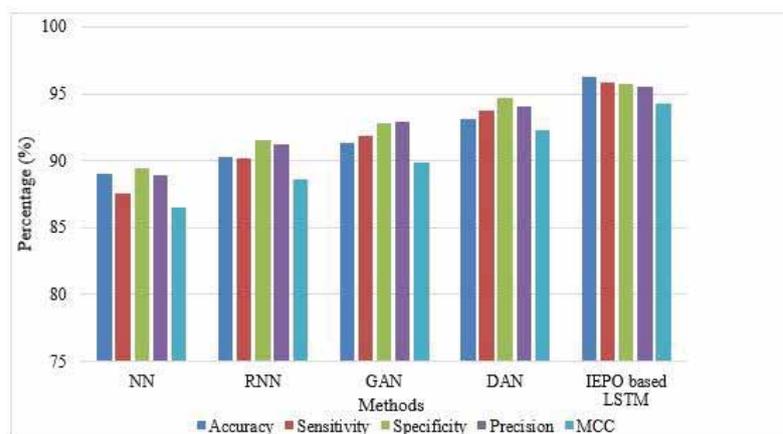
Fig. 6: Performance of MIT-BIH arrhythmia classification without feature selection.
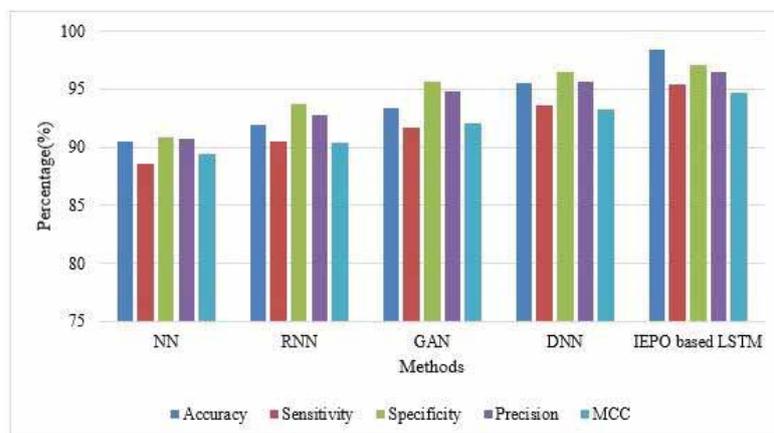
Fig. 7: Performance for MIT-BIH Arrhythmia classification with feature selection.

Table 3: Simulation results of IEPO based LSTM by varying the classifiers for CPSC 2018 Dataset

Without feature selection results for CPSC 2018 Dataset					
Methods	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**Precision (%)**	**MCC (%)**
NN	85.49	86.03	87.04	90.71	88.6
RNN	88.02	88.58	89.61	92.71	90.44
GAN	89.15	90.38	90.28	91.03	91.71
DNN	92.16	93.63	92.71	93.94	92.16
IEPO based LSTM	95.75	94.09	94.19	95.35	94.49
With Feature selection results					
Methods	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**Precision (%)**	**MCC (%)**
NN	90.28	90.05	91.13	92.09	88.43
RNN	92.03	94.68	96.56	93.48	91.90
GAN	90.40	92.58	89.59	91.55	88.81
DNN	50.28	51.65	52.75	53.23	51.65
IEPO based LSTM	99.79	99.11	98.20	99.43	98.57

The performance measure of classifiers on CPSC dataset with feature selection is highest when compared with the values of performance metrics without feature selection. The accuracy, sensitivity, specificity, precision, and MCC of Neural Network (NN), RNN (Recurrent Neural Network), GAN (Generative Adversarial Network), and DNN (Deep Neural Network) are measured and compared with the proposed IEPO-based LSTM. The obtained results show that the proposed IEPO based LSTM achieves highest values with accuracy of 99.79%, sensitivity of 99.11%, specificity of 98.20%, precision of 99.43% and MCC of 98.57% while considering with feature selection compared to the other classifiers. The simulation results of the proposed method by varying different classifiers on CPSC dataset with and without feature selection are specified in Fig. 8 and Fig. 9.
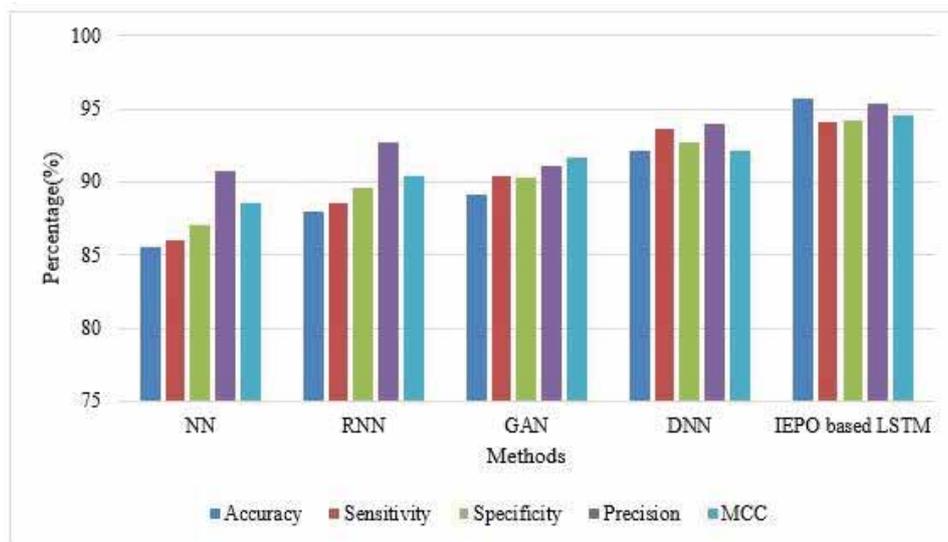
Fig. 8: Performance for CPSC Arrhythmia without feature selection.

The FCM – ETIFRST method selects the unique features to represent the class that solves the problem of imbalance and overfitting. The modified non-membership function helps to provide exploitation, and the modified membership function performs exploration. The similarity measure helps to maintain the exploration and exploitation of the feature

selection process. LSTM shows higher sensitivity than other classifiers due to its efficiency in handling the features.
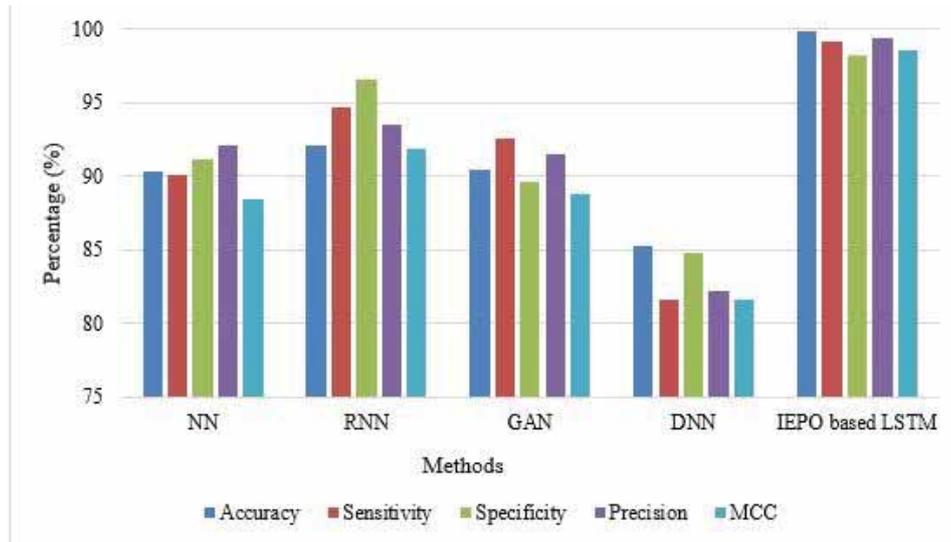
Fig. 9: Performance for CPSC Arrhythmia without feature selection.

Table 4 describes the performance of optimizing the hyperparameters of LSTM using various multi-objective optimization methods like Particle Swarm Optimization (PSO), Fruit Fly Optimization (FFO), Anti Colony Optimization (ACO), Salt Swarm optimization Algorithm (SSA) with the selected features obtained from FCM-ETIRST on MIT-BIH arrhythmia dataset. In the observation, the results infer that the proposed IEPO algorithm has chosen the best set of hyperparameters for LSTM to perform at its better stage, when compared to the resultant parameters obtained from other optimization algorithms. It also shows an improvement of around 3% in accuracy compared to the existing optimization algorithm SSA. The performance comparison of different hyperparameter optimization techniques on MIT-BIH dataset, with and without feature selection is graphically represented in Fig. 10 and Fig. 11.

Table 5 represents the performance analysis of K-fold validation in terms of performance metrics like accuracy, sensitivity, specificity, precision, and MCC on the MIT-BIH dataset and its graphical representation is specified in Fig. 12.

Table 4: Comparison of Hyper parameter optimization method on MIT-BIH dataset

Without Feature Selection Hyper Parameter Optimization method comparison					
Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	MCC (%)
PSO	85.89	84.69	86.14	80.59	86.38
FFO	83.33	84.23	84.06	83.24	86.00
ACO	96.17	92.03	94.83	97.08	97.54
SSA	95.31	90.67	96.00	94.04	96.30
IEPO based LSTM	98.95	99.49	99.49	98.45	97.99
With Feature Selection Hyper Parameter Optimization method					
PSO	85.89	84.69	86.14	80.59	86.38
FFO	83.33	84.23	84.06	83.24	86.00
ACO	96.17	92.03	94.83	97.08	97.54
SSA	95.31	90.67	96.00	94.04	96.30
IEPO based LSTM	98.95	99.49	99.49	98.45	97.99

Ramaiah and Srikantegowda

Fig. 10: Performance of hyper parameter optimization with feature selection.
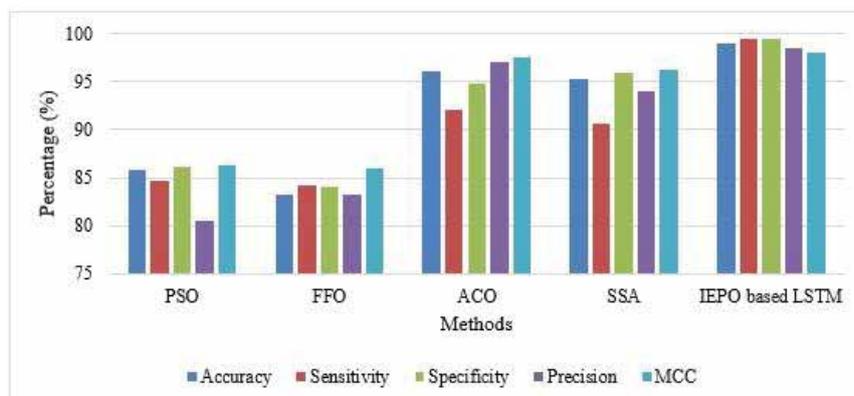
Fig. 11: Performance of hyper parameter optimization without feature selection.

Table 5: Performance analysis of K-fold validation on MIT-BIH dataset [13]

K-fold values	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	MCC (%)
4-fold	97.99	97.08	97.55	97.19	96.65
5-fold	98.95	99.49	99.49	98.45	97.99
7-fold	97.83	96.09	97.19	97.95	96.74
9-fold	97.47	96.28	96.15	96.37	95.45
10-fold	98.67	95.47	96.33	95.93	90.61
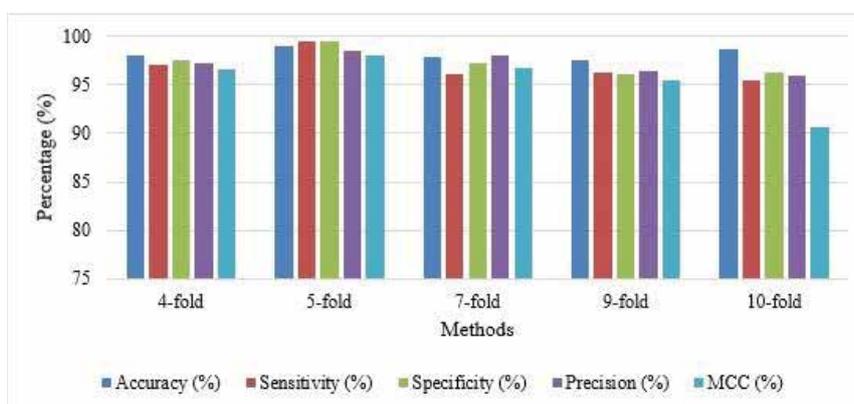
Fig. 12: K-fold validation performance.

5. COMPARATIVE ANALYSIS

The comparative analysis of the proposed VF classification using IPEO based LSTM to the existing Boosted SVM based DE algorithm and Hybrid CNN-SVM deep neural networks using the LSTM model in terms of accuracy, specificity, sensitivity, and precision is shown in Table 6. The proposed model is trained, tested, and validated using two publicly available standard datasets namely the Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) and the China Physiological Signal Challenge CPSC 2018 dataset and both consist of ECG recordings for 5 seconds of coronary heart disease (CHD) patients. The accuracy was improved to 99.75%, with a specificity of 98.29%, a sensitivity of 98.39% and the precision held was 98.35%. The proposed IPEO was evaluated for efficient classification using the VF technique by overcoming overfitting problems and dataset classification.

Table 6: Comparative analysis of existing SVM classifier to the LSTM classifier

Technique	Dataset	Accuracy (%)	Specificity (%)	Sensitivity (%)	Precision (%)
Boosted SVM based DE algorithm[15]	MIT-BIH	98.20	98.18	98.25	-
Hybrid CNN-SVM deep neural networks using LSTM [16]	CPSC	98.69	99.99	97.74	-
IEPO based LSTM	MIT-BIH	99.75	98.29	98.39	98.35
	CPSC	98.11	97.21	97.79	98.29

6. CONCLUSION

In this paper, a novel classification method for classifying VF rhythm has been developed on the MIT-BIH and CPSC datasets consisting of ECG recordings of CHD patients. The method uses an LSTM classifier to classify VF and FCM- ETIFRST for feature selection and performs clustering of membership, non-membership, and hesitancy degrees. The method uses a window size of 5 seconds that requires less memory and obtained the highest accuracy of 99.75%. The proposed algorithm, on the other hand, combines traditional machine learning algorithms and CNN algorithms. The deep CNN was able to successfully extract useful deep features for ECG classification from raw ECG signals without prior knowledge of ECG signals or cardiac rhythm disorders. The proposed method outperformed the existing classification models on CHD by improving the efficiency of arrhythmia classification with suitable hyperparameter optimization in the LSTM algorithm. The improved performance of the proposed method on different test datasets suggests that it has the potential for clinical application. Future work can be focused on improving classification accuracy for both VF and AF by considering a window size of more than 5 seconds and it will be analyzed in real-time events.

DATA AVAILABILITY STATEMENT

Dataset Links:

MIT BIH Arrhythmia Dataset : https://archive.physionet.org/cgi-bin/atm/ATM
CPSC 2018 Dataset: http://2018.icbeb.org/Challenge.html

REFERENCES

[1] Nilashi M, Ahmadi H, Manaf AA, Rashid TA, Samad S, Shahmoradi L, Aljojo N, Akbari E. (2020) Coronary heart disease diagnosis through self-organizing map and fuzzy support vector machine with incremental updates. Int. J. Fuzzy Syst., 22(4): 1376-1388. https://doi.org/10.1007/s40815-020-00828-7

[2] Khan MU, Aziz S, Iqtidar K, Zaher GF, Alghamdi S, Gull M. (2022) A two-stage classification model integrating feature fusion for coronary artery disease detection and classification. Multimedia Tools and Applications, 81(10): 13661-13690. https://doi.org/10.1007/s11042-021-10805-3

[3] Shah SMS, Shah FA, Hussain SA, Batool S. (2020) Support vector machines-based heart disease diagnosis using feature subset, wrapping selection and extraction methods. Comput. Electr. Eng., 84: 106628. https://doi.org/10.1016/j.compeleceng.2020.106628

[4] Mohanty M, Dash M, Biswal P, Sabut S. (2021) Classification of ventricular arrhythmias using empirical mode decomposition and machine learning algorithms. Prog. Artif. Intell., 10(4):489-504. https://doi.org/10.1007/s13748-021-00250-6

[5] Muhammad LJ, Algehyne EA. (2021) Fuzzy based expert system for diagnosis of coronary artery disease in Nigeria. Health and Technology, 11(2): 319-329. https://doi.org/10.1007%2Fs12553-021-00531-z

[6] Nasarian E, Abdar M, Fahami MA, Alizadehsani R, Hussain S, Basiri ME, Zomorodi-Moghadam M, Zhou X, Pławiak P, Acharya UR, Tan RS, Sarrafzadegan S. (2020) Association between work-related features and coronary artery disease: A heterogeneous hybrid feature selection integrated with balancing approach. Pattern Recognit. Lett., 133: 33-40. https://doi.org/10.1016/j.patrec.2020.02.010

[7] Apostolopoulos ID, Groumpos PP. (2020) Non-invasive modelling methodology for the diagnosis of coronary artery disease using fuzzy cognitive maps. Comput. Methods Biomech. Biomed. Eng., 23(12): 879-887. https://doi.org/10.1080/10255842.2020.1768534

[8] Ma C, Wei S, Chen T, Zhong J, Liu Z, Liu C. (2020) Integration of results from convolutional neural network in a support vector machine for the detection of atrial fibrillation. IEEE Trans. Instrum. Meas., 70(2504610): 1-10. https://doi.org/10.1109/TIM.2020.3044718.

[9] Chen X, Cheng Z, Wang S, Lu G, Xv G, Liu Q, Zhu X. (2021) Atrial fibrillation detection based on multi-feature extraction and convolutional neural network for processing ECG signals. Comput. Methods Programs Biomed., 202: 106009. https://doi.org/10.1016/j.cmpb.2021.106009

[10] Kumari CU, Murthy ASD, Prasanna BL, Reddy MPP, Panigrahy AK. (2021) An automated detection of heart arrhythmias using machine learning technique: SVM. Mater. Today Proc., 45(2): 1393-1398. https://doi.org/10.1016/j.matpr.2020.07.088

[11] Li H, Wang X, Liu C, Zeng Q, Zheng Y, Chu X, Yao L, Wang J, Jiao Y, Karmakar C. (2020) A fusion framework based on multi-domain features and deep learning features of phonocardiogram for coronary artery disease detection. Comput. Biol. Med., 120: 103733. https://doi.org/10.1016/j.compbiomed.2020.103733

[12] Nguyen QH, Nguyen BP, Nguyen TB, Do TT, Mbinta JF, Simpson CR. (2021) Stacking segment-based CNN with SVM for recognition of atrial fibrillation from single-lead ECG recordings. Biomed. Signal Process. Control, 68: 102672. https://doi.org/10.1016/j.bspc.2021.102672

[13] Jahmunah V, Ng EYK, San TR, Acharya UR. (2021) Automated detection of coronary artery disease, myocardial infarction and congestive heart failure using GaborCNN model with ECG signals. Comput. Biol. Med., 134: 104457. https://doi.org/10.1016/j.compbiomed.2021.104457

[14] Tseng LM, Tseng VS. (2020) Predicting ventricular fibrillation through deep learning. IEEE Access, 8: 221886-221896. https://doi.org/10.1109/ACCESS.2020.3042782

[15] Panigrahy D, Sahu PK, Albu F. (2021) Detection of ventricular fibrillation rhythm by using boosted support vector machine with an optimal variable combination. Comput. Electr. Eng., 91: 107035. https://doi.org/10.1016/j.compeleceng.2021.107035

[16] Çınar A, Tuncer SA. (2021) Classification of normal sinus rhythm, abnormal arrhythmia and congestive heart failure ECG signals using LSTM and hybrid CNN-SVM deep neural networks. Comput. Methods Biomech. Biomed. Eng., 24(2): 203-214. https://doi.org/10.1080/10255842.2020.1821192

[17] Chen TM, Huang CH, Shih ES, Hu YF, Hwang MJ. (2020) Detection and classification of cardiac arrhythmias by a challenge-best deep learning neural network model. Iscience, 23(3): 100886. https://doi.org/10.1016/j.isci.2020.100886

[18] Chang KC, Hsieh PH, Wu MY, Wang YC, Chen JY, Tsai FJ, Shih ES, Hwang MJ, Huang TC. (2021) Usefulness of machine learning-based detection and classification of cardiac arrhythmias with 12-lead electrocardiograms. Can. J. Cardiol, 37(1): 94-104. https://doi.org/10.1016/j.cjca.2020.02.096

[19] Cai W, Chen Y, Guo J, Han B, Shi Y, Ji L, Wang J, Zhang G, Luo, J. (2020) Accurate detection of atrial fibrillation from 12-lead ECG using deep neural network. Comput. Biol. Med., 116: 103378. https://doi.org/10.1016/j.compbiomed.2019.103378

[20] Rai HM, Chatterjee K. (2022) Hybrid CNN-LSTM deep learning model and ensemble technique for automatic detection of myocardial infarction using big ECG data. Appl. Intell., 52(5):5366-5384. https://doi.org/10.1007/s10489-021-02696-6

[21] Baral S, Alsadoon A, Prasad PWC, Al Aloussi S, Alsadoon OH. (2021) A novel solution of using deep learning for early prediction cardiac arrest in Sepsis patient: enhanced bidirectional long short-term memory (LSTM). Multimedia Tools and Applications, 80(21-23): 32639-32664. https://doi.org/10.1007/s11042-021-11176-5

[22] Sun L, Wang Y, He J, Li H, Peng D, Wang Y. (2020) A stacked LSTM for atrial fibrillation prediction based on multivariate ECGs. Health Inf. Sci. Syst., 8(1): 19. https://doi.org/10.1007/s13755-020-00103-x

TREBLE SEARCH OPTIMIZER: A STOCHASTIC OPTIMIZATION TO OVERCOME BOTH UNIMODAL AND MULTIMODAL PROBLEMS

PURBA DARU KUSUMA*AND **ASHRI DINIMAHARAWATI**

Computer Engineering, Telkom University, Bandung, Indonesia

*Corresponding author: purbodaru@telkomuniversity.ac.id

ABSTRACT: Today, many metaheuristics have used metaphors as their inspiration and baseline for novelty. It makes the novel strategy of these metaheuristics difficult to investigate. Moreover, many metaheuristics use high iteration or swarm size in their first introduction. Based on this consideration, this work proposes a new metaheuristic free from metaphor. This metaheuristic is called treble search optimizer (TSO), representing its main concept in performing three searches performed by each member in each iteration. These three searches consist of two directed searches and one random search. Several seeds are generated from each search. Then, these searches are compared with each other to find the best seed that might substitute the current corresponding member. TSO is also designed to overcome the optimization problem in the low iteration or swarm size circumstance. In this paper, TSO is challenged to overcome the 23 classic optimization functions. In this experiment, TSO is compared with five shortcoming metaheuristics: slime mould algorithm (SMA), hybrid pelican komodo algorithm (HPKA), mixed leader-based optimizer (MLBO), golden search optimizer (GSO), and total interaction algorithm (TIA). The result shows that TSO performs effectively and outperforms these five metaheuristics by making better fitness scores than SMA, HPKA, MLBO, GSO, and TIA in overcoming 21, 21, 23, 23, and 17 functions, consecutively. The result also indicates that TSO performs effectively in overcoming unimodal and multimodal problems in the low iteration and swarm size.

ABSTRAK: Dewasa ini, terdapat ramai metaheuristik menggunakan metafora sebagai inspirasi dan garis dasar pembaharuan. Ini menyebabkan strategi baharu metaheuristik ini susah untuk dikaji. Tambahan, ramai metaheuristik menggunakan ulangan berulang atau saiz kerumunan dalam pengenalan mereka. Berdasarkan penilaian ini, kajian ini mencadangkan metaheuristk baharu bebas metafora. Metaheuristik ini dipanggil pengoptimum pencarian ganda tiga (TSO), mewakilkan konsep utama dalam pemilihan tiga pencarian yang dilakukan oleh setiap ahli dalam setiap ulangan. Ketiga-tiga carian ini terdiri daripada dua pencarian terarah dan satu pencarian rawak. Beberapa benih dihasilkan dalam setiap carian. Kemudian, carian ini dibandingkan antara satu sama lain bagi mencari benih terbaik yang mungkin berpotensi menggantikan ahli yang sedang digunakan. TSO juga direka bagi mengatasi masalah pengoptimuman dalam ulangan rendah atau lingkungan saiz kerumunan. Kajian ini TSO dicabar bagi mengatasi 23 fungsi pengoptimuman klasik. Eksperimen ini TSO dibandingkan dengan lima kekurangan metaheuristik: algoritma acuan lendir (SMA), algorithma hibrid komodo burung undan (HPKA), Pengoptimum Campuran berdasarkan-Ketua (MLBO), Pengoptimuman Carian Emas (GSO), dan algoritma jumlah interaksi (TIA). Dapatan kajian menunjukkan TSO berkesan menghasilkan dan lebih baik daripada kelima-lima metaheuristik dengan menghasilkan pemarkahan padanan terbaik berbanding SMA, HPKA, MLBO, GSO, dan TIA dalam mengatasi fungsi 21, 21, 23, 23, dan 17, secara

berurutan. Dapatan kajian juga menunjukkan TSO turut berperanan efektif dalam mengatasi masalah modal tunggal dan modal ganda dalam iterasi rendah dan saiz kerumunan.

KEYWORDS: *optimization; metaheuristic; swarm intelligence; unimodal; multimodal*

1. INTRODUCTION

Metaheuristics is a popular tool extensively used in various optimization problems. Many optimization studies from a wide range of subjects use metaheuristics, such as in smart farming [1], path planning for autonomous robots [2], traffic forecasting [3], power systems [4], electric vehicle charge scheduling [5], and so on. Today, hundreds of metaheuristics exist and are ready to be used in any optimization problem. This circumstance becomes one of several reasons why metaheuristic is so popular. Moreover, there are optimization studies that hybridize a metaheuristic with other methods, whether they are metaheuristics or exact methods. Because metaheuristics are flexible in overcoming various optimization problems and easy to modify, there are many studies on hybridizing metaheuristics.

In general, this massive development of metaheuristics comes from two reasons. The first reason is that various things can be used as inspiration for searching mechanisms, especially nature. Many metaheuristics use nature, especially animal behavior, as their inspiration and transform it into an optimization or searching strategy. Several shortcoming metaheuristics that use animal behavior as their inspiration, such as the Komodo mlipir algorithm (KMA) [6], northern goshawk optimizer (NGO) [7], marine predator algorithm (MPA) [8], hybrid pelican Komodo algorithm (HPKA) [9], coati optimization algorithm (COA) [10], cheetah optimizer (CO) [11], chameleon swarm algorithm (CSA) [12], and so on. Several metaheuristics used the term leader to represent the reference during the directed search, such as mixed leader-based optimizer (MLBO) [13], random selected leader-based optimizer (RSLBO) [14], hybrid leader-based optimizer (HLBO) [15], and so on. Meanwhile, several metaheuristics declared their main concept or strategy for their name rather than using metaphors, such as total interaction algorithm (TIA) [16], golden search optimizer (GSO) [17], average and subtraction-based optimizer (ASBO) [18], and so on. The second reason is that no metaheuristic is suitable or superior in overcoming any optimization problem, as stated in the no-free-lunch theory. Each strategy has its strengths and weaknesses. In other words, no metaheuristic can accommodate all strategies.

There are several critiques following the massive development of new metaheuristics. First, many metaphor-based metaheuristics use their metaphor as a novelty or contribution. However, through the investigation of the algorithm and mathematical model, their method is slightly different from the previous ones [19]. Second, many studies proposing new metaheuristics exploited their ability to outperform the previous metaheuristics rather than highlighted their distinct mechanics in a clear explanation [19]. Besides, the performance of many metaheuristics is investigated in the high iteration or swarm in their first appearance. Moreover, these circumstances need to be clarified in several studies proposing a new metaheuristic. For example, NGO uses the behavior of the northern goshawk as metaphor and the maximum iteration is set to 1,000 during the evaluation [7]. The maximum iteration is also set to 1,000 in the first introduction of ASBO [18]. In the first introduction of GSO, the maximum iteration is set to 1,000 while the swarm size is set to 30 [17]. Unfortunately, the performance of these metaheuristics has not been investigated in the low swarm and low iteration circumstance.

The objective of this work is to promote a new simple and metaphor-free swarm-based metaheuristic that works effectively with a low iteration number and swarm size. This metaheuristic is called a treble search optimizer (TSO) which comes from the three searches performed in the algorithm. These searches include two directed searches and one random search. The global optimal member and one randomly selected member become the references in these two directed searches. Meanwhile, the random search focuses on finding a better member near or around the corresponding member. In TSO, each member performs all these three searches in every iteration, which means TSO does not implement segregation of roles. One best seed is selected in every search, so three seeds from three searches are generated by a corresponding member in every search. Then, the best seed becomes the final seed as a substitute for the current corresponding member.

Based on this explanation, below are the novelties and contributions of this work.

1) This work promotes a novel swarm-based metaheuristic that is free from using metaphors, named treble search optimizer (TSO).

2) TSO was designed to overcome the optimization problem in the low iteration and swarm circumstance.

3) TSO performs three searches (two directed searches and one random search) where several seeds are generated from each search.

4) The performance of TSO is investigated using 23 classic functions to overcome.

2. RELATED WORKS

Investigating the existing metaheuristics is the first and main critical step in proposing a new metaheuristic. Within this investigation, it is very important to highlight the distinction, novelty, or uniqueness of the metaheuristic. This step is also important because there are hundreds of metaheuristics already in existence. Proposing a new metaheuristic without investigating the existing ones, especially the shortcoming ones, may end with proposing a metaheuristic like the existing ones. This investigation is also important because a new metaheuristic can be developed by modifying or hybridizing several existing metaheuristics.

Investigating a metaheuristic can be performed by classifying the metaheuristic based on several parameters. First, a metaheuristic should be classified as whether it uses metaphors. As mentioned, a metaphor-based metaheuristic should be rigorously investigated to address its distinct approach by abstracting the metaphor. Second, the algorithm and mathematical model following the algorithm should also be reviewed. After that, several parameters can be used to classify the metaheuristic, such as the number of searches, segregation of roles, and so on.

In many shortcoming metaheuristics, implementing multiple search strategies has become more popular rather than performing a single search strategy, as shown in Table 1. The main reason is that there is not any single search that can guarantee finding the optimal member. Using the global best member or the best member among the swarm becomes the most popular option so that this member is used in many swarm-based metaheuristics, such as KMA [6], HPKA [9], ASBO [18], and so on. Some other metaheuristics use local best member for their reference, such as in MPA [8], GSO [17], and so on. Meanwhile, there are also metaheuristics using the other members within the swarm as their reference, such as in TIA [16], NGO [7], and so on. Even moving toward

the best member may guide the entire swarm toward the local optimal entrapment because the global best member lies somewhere else in the search space.

Table 1: List of shortcoming metaheuristics

No	Metaheuristic	Metaphor	Segregation of Roles	Number of Searches	Maximum Iteration	Swarm Size
1	KMA [6]	komodo dragon	yes	4	5,000 evaluations	5, 20-200
2	NGO [7]	northern goshawk	no	2	1,000	20-80
3	HPKA [9]	pelican and komodo dragon	yes	4	200	20
4	MPA [8]	marine predator	yes	5	500	50
5	CO [11]	cheetah	yes	3	12×10^5 evaluations	6
6	COA [10]	coati	no	3	200, 1,000	n/a
7	HLBO [15]	leader	no	2	1,000	n/a
8	MLBO [13]	leader	no	1	n/a	n/a
9	RSLBO [14]	leader	no	1	n/a	n/a
10	ASBO [18]	-	no	3	1,000	20-80
11	GSO [17]	-	no	1	1,000	30
12	TIA [16]	-	no	1	50	10
13	this work	-	no	3	40	5

The list of swarm-based shortcoming metaheuristics is displayed in Table 1. Table 1 consists of five pieces of information related to the corresponding metaheuristics. The first information is the metaphor used in the metaheuristic. The second information is whether the corresponding metaheuristic implements the segregation of roles. The third information is the number of searches implemented in the corresponding metaheuristic. The fourth information is the maximum iteration set in the first appearance of the corresponding metaheuristic. The fifth information is the swarm size set in the first appearance of the corresponding metaheuristic. There are 14 things that could be improved in metaheuristics in Table 1. The last row of Table 1 presents the attributes of TSO to give a clear view regarding the novelty and position of this work.

Table 1 presents many areas for improvement in the metaheuristic's use of metaphors, especially animals. Some metaheuristics use the term leader while others do not use a metaphor. Most metaheuristics do not perform segregation of roles so that in these metaheuristics, all members perform all searches adopted in the corresponding metaheuristic. Besides, most shortcoming metaheuristics perform multiple searches rather than a single search. In their first appearance, many metaheuristics are challenged to overcome optimization problems in the high maximum iteration or swarm size. Based on this explanation, the opportunity to propose a new swarm-based metaheuristic that is metaphor-free implements multiple searches and is challenged to overcome problems in the low maximum iteration and low swarm size is still open.

3. MODEL

TSO is built based on two approaches. First, TSO performs three searches where each output will be compared to find the best output. Second, there are multiple seeds generated in every search so that the best seed among these seeds will be chosen to compete with other selected seeds from other searches mentioned in the first approach.

algorithm 1: treble search optimizer (TSO)	
1	**output**: s_b
2	**for all** s **in** S
3	initialize s using Eq. (1)
4	update s_b using Eq. (2)
5	**end for**
6	**for** t=1 **to** t_{max}
7	**for all** s **in** S
8	select s_s using Eq. (3)
9	**for** j=1 **to** n_c
10	generate c_1, c_2, c_3 using Eq. (4), Eq. (5), Eq. (6)
11	**end for**
12	select c_{s1}, c_{s2}, c_{s3} using Eq. (7), Eq. (8), Eq. (9)
13	select c_f using Eq. (10)
14	update s using Eq. (11)
15	update s_b using Eq. (2)
16	**end for**
17	**end for**

c_1	first search seed
c_2	second search seed
c_3	third search seed
C_1	set of first search seed
C_2	set of second search seed
C_3	set of third search seed
c_{s1}	selected seed among first search seeds
c_{s1}	selected seed among second search seeds
c_{s3}	selected seed among third search seeds
c_f	final seed
f	objective function
n_c	number of seeds
s	member
S	set of members
s_b	global best member
s_s	selected member
s_u	upper boundary
s_l	lower boundary
t	iteration
t_{max}	maximum iteration
U	uniform random
U_r	real uniform random number
U_i	integer random number

In TSO, each corresponding member performs three searches which are two directed searches and one random search. The first directed search generates several seeds along the way, the corresponding member toward the global best member. The second directed search generates seeds relative to a selected random member within the swarm. In the second directed search, these seeds may be in the direction of the corresponding member toward the selected member or away from the selected member. This choice depends on the quality of the corresponding member and the randomly selected member. The first direction occurs if this randomly selected member is better than the corresponding member. Otherwise, the second direction takes place. In the third search, several seeds are generated around the corresponding member.

As a metaheuristic, TSO consists of two phases: initialization and iteration. In the initialization, all members are uniformly randomized within the search space. Meanwhile, the iteration phase represents the improvement where three searches are performed. At the end of the process, the global best member becomes the final member, i.e., the algorithm

output.

The best seed is then chosen in every search. It means that there are now three selected seeds from three searches. Then, these seeds will compete among each other so that the best seed among these three seeds becomes the final seed. This final seed is then compared with the corresponding member. This final seed substitutes the corresponding member only if this final seed is better than the current corresponding member. Otherwise, the corresponding member remains static at the end of this iteration. This concept is then transformed into an algorithm and mathematical model. The mathematical model is displayed in Eq. (1) to Eq. (11). The formalization of TSO is displayed in algorithm 1. Below are annotations used in algorithm 1 and the following mathematical model. The visualization of TSO is presented in Fig. 1.
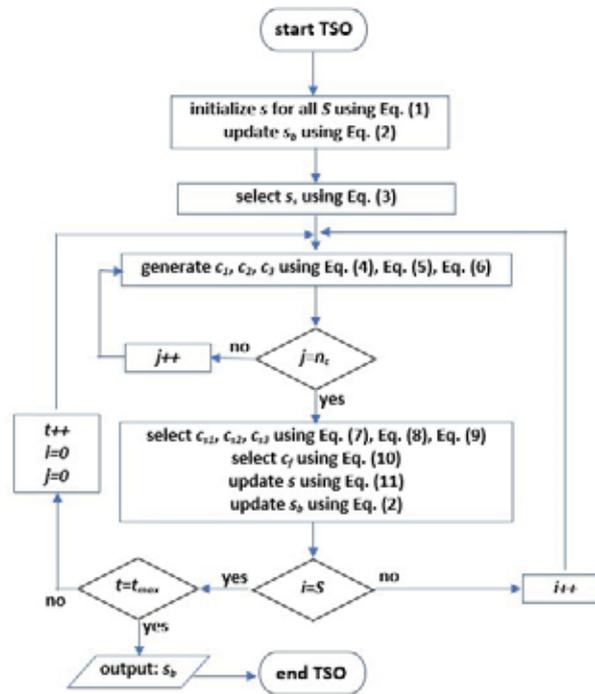
Fig. 1: Flowchart of treble search optimizer.

$$s = U_r(s_l, s_u) \tag{1}$$

$$s_b' = \begin{cases} s, f(s) < f(s_b) \\ s_b, otherwise \end{cases} \tag{2}$$

$$s_s = U(S) \tag{3}$$

$$c_1 = s + U_r(0,1)(s_b - U_i(1,2)s) \tag{4}$$

$$c_2 = \begin{cases} s + U_r(0,1)(s_s - U_i(1,2)s) \\ s + U_r(0,1)(s - U_i(1,2)s_s) \end{cases} \tag{5}$$

$$c_3 = s + 0.1U_r(-1,1)(s_u - s_l) \tag{6}$$

$$c_{s1} = c_1 \in C_1 \wedge min(f(c_1)) \tag{7}$$

$$c_{s2} = c_2 \in C_2 \wedge min(f(c_2)) \tag{8}$$

$$c_{s3} = c_3 \in C_3 \wedge min(f(c_3)) \tag{9}$$

$$c_f = c, min(c_1, c_2, c_3) \tag{10}$$

$$s' = \begin{cases} c_f, f(c_f) < f(s) \\ s, otherwise \end{cases} \tag{11}$$

Below is the detailed explanation of Eq. (1) to Eq. (11). The global best member becomes the final solution. Lines 2 to 5 represent the initialization phase. Lines 6 to 17 represent the iteration phase. Equation (1) describes that the initial member is uniformly randomized between the lower and upper boundary, i.e., search space. Equation (2) describes that the corresponding member substitutes the current global best member if this corresponding member is better than the current global best member. Equation (3) describes randomly selecting a member among the set of members. Equation (4) describes that the seed of the first search is generated along the way from the corresponding member toward the global best member. Equation (5) describes that the seed of the second search is generated based on the relation between the corresponding member and the randomly selected member. Equation (6) describes that the seed of the third search is generated near the corresponding member. Equations (7) to Eq. (9) describes that the best seed is selected from among seeds in every search. Equation (10) describes that the best seed among these three selected seeds becomes the final seed. Equation (11) describes that the final seed substitutes the current corresponding member if this final seed is better than the current corresponding member.

4. RESULTS

This section presents the experiment performed to evaluate the performance of TSO and its result. There are two experiments regarding this work. The first experiment is performed to evaluate the performance of TSO in overcoming a set of benchmark functions and the performance comparison between TSO and the sparing metaheuristics. The second experiment is performed to evaluate the hyperparameter of TSO. The 23 classic functions are chosen as the benchmark functions.

These 23 classic functions are chosen based on several reasons. The first reason is that these functions represent various problems with specific circumstances and challenges. The second reason is that these functions are very popular, so they are chosen in many studies proposing a new metaheuristic. These functions can be categorized into three groups: seven high-dimension unimodal functions, six high-dimension multimodal functions, and ten fixed-dimension multimodal functions. These functions also represent problems with various search spaces, from narrow to large ones.

In the first experiment, TSO is compared with five shortcoming metaheuristics: SMA, HPKA, MLBO, GSO, and TIA. These metaheuristics are chosen mainly because they are new. In their first appearance, these metaheuristics outperformed many previous metaheuristics. SMA outperformed many metaheuristics, such as the whale optimization algorithm (WOA), moth-flame optimizer (MFO), grey wolf optimizer (GWO), bat algorithm (BA), sine cosine algorithm (SCA), particle swarm optimization (PSO), firefly

algorithm (FA), multi-verse optimizer (MVO), salp swarm algorithm (SSA), ant lion optimizer (ALO), and differential evolution (DE) [20]. HPKA outperformed four metaheuristics: GWO, MPA, KMA, and POA [9]. MLBO outperforms several metaheuristics, such as PSO, genetic algorithm (GA), teaching-learning based optimizer (TLBO), GWO, emperor penguin optimizer (EPOA), and so on [13]. GSO outperformed four metaheuristics: gravitational search algorithm (GSA), SCA, tunicate swarm algorithm (TSA), and GWO [17]. TIA outperformed five sparing metaheuristics: PSO, marine predator algorithm (MPA), GSO, directed pelican algorithm (GPA), and driving training-based optimizer (DTBO) [16].

Table 2: The simulation result of the first experiment

Function	Average Fitness Score					
	SMA	**HPKA**	**MLBO**	**GSO**	**TIA**	**TSO**
1	6.6663×10^4	6.2570×10^4	2.2305×10^4	6.0029×10^4	**0.0000**	**0.0000**
2	**0.0000**	**0.0000**	1.1222×10^{51}	7.5089×10^{69}	1.7930×10^{55}	**0.0000**
3	2.1289×10^5	2.0975×10^5	5.7327×10^4	1.7531×10^5	0.0316	**0.0000**
4	8.2362×10^1	7.9015×10^1	4.7480×10^1	5.6846×10^1	**0.0000**	**0.0000**
5	1.7908×10^8	1.9984×10^8	1.9085×10^7	1.2710×10^8	4.8879×10^1	**4.8803×10^1**
6	6.9099×10^4	6.9444×10^4	2.1638×10^4	5.9789×10^4	9.4253	**8.2081**
7	1.6824×10^2	1.3522×10^2	1.6547×10^1	9.2918×10^1	0.0122	**0.0021**
8	**-6.2418×10^3**	-5.9791×10^3	-3.9934×10^3	-4.3938×10^3	-2.3210×10^3	-4.6397×10^3
9	5.4539×10^2	5.6253×10^2	4.4459×10^2	5.1443×10^2	**0.0000**	**0.0000**
10	1.9750×10^1	1.9947×10^1	1.6483×10^1	1.9463×10^1	**0.0000**	**0.0000**
11	6.2192×10^2	7.1348×10^2	1.9505×10^2	5.3193×10^2	0.0087	**0.0000**
12	4.6092×10^8	3.3887×10^8	8.1216×10^6	2.2263×10^8	0.8217	**0.6182**
13	1.1149×10^9	8.5180×10^8	5.0042×10^7	4.7534×10^8	3.0902	**2.8085**
14	1.1963×10^1	1.3565×10^1	7.6193	1.1118×10^1	7.5131	**3.7733**
15	0.0106	0.0111	0.0141	0.0277	0.0046	**0.0006**
16	-0.8973	-0.8722	-0.9689	-0.9862	-1.0219	**-1.0314**
17	0.9389	0.7064	0.4966	0.8170	1.7409	**0.3986**
18	8.0109×10^1	2.2714×10^1	9.8071	2.6717×10^1	1.7821×10^1	**3.0137**
19	-0.0455	-0.0387	-0.0486	-0.0193	**-0.0495**	**-0.0495**
20	-2.7549	-2.7309	-2.8287	-2.5539	-2.3823	**-3.2980**
21	-2.9231	-4.0803	-2.8597	-2.7178	-4.1103	**-8.2464**
22	-3.7697	-3.0703	-2.8335	-3.3707	-3.4117	**-7.6623**
23	-3.5599	-2.8837	-2.3017	-3.1063	-2.5551	**-8.8215**

In the first experiment, several parameters are set based on a certain value. The swarm size is set to 5, which represents the low swarm. The maximum iteration is set to 40, which represents low iteration. The dimension is set to 50, which represents a high-dimension problem. In HPKA, all searches have equal opportunity. The result is displayed in Table 2, while the superiority of the TSO compared with the other metaheuristics based on the group of functions is displayed in Table 3. In Table 2, the best score is written in bold font. Meanwhile, the floating-point accuracy is set to 10^{-4} so that a score less than 10^{-4} is rounded to 0.

Table 2 indicates the excellent performance of TSO in terms of finding the optimal global member and producing the best scores among the metaheuristics. TSO could find the optimal global member of seven functions: Sphere, Schwefel 2.22, Schwefel 1.2, Schwefel 2.21, Rastrigin, Ackley, and Griewank. Meanwhile, TSO could also find the

member near the global optimal member of three functions: Kowalik, Six Hump Camel, and Branin. TSO also performed the best of 22 functions out of 23 functions. However, several metaheuristics also performed the same value in six functions. These six functions are Sphere, Schwefel 2.22, Schwefel 2.21, Rastrigin, Ackley, and Hartman 3. TIA performed the same value in five functions (Sphere, Schwefel 2.21, Rastrigin, Ackley, and Hartman 3), while SMA and HPKA performed the same value in Schwefel 2.22.

Table 3 indicates the superiority of TSO among other competing metaheuristics in all groups of functions. TSO was better than SMA, HPKA, MLBO, GSO, and TIA in overcoming 21, 21, 23, 23, and 17 functions, respectively. It means that TSO was superior to MLBO and GSO. TSO was superior to SMA and HPKA. Meanwhile, TSO was still superior to TIA, although TIA is the most challenging metaheuristic to beat. This result indicates the superiority of TSO in overcoming all three kinds of problems: high-dimension unimodal problems, high-dimension multimodal problems, and fixed-dimension multimodal problems.

Table 3: TSO superiority among other metaheuristics based on group of functions

Group	Number of Functions				
	SMA	HPKA	MLBO	GSO	TIA
1	6	6	7	7	4
2	5	5	6	6	4
3	10	10	10	10	9
Total	21	21	23	23	17

The second experiment was performed to evaluate the sensitivity or hyperparameter. There were three parameters evaluated in this experiment: maximum iteration, swarm size, and the number of seeds. This experiment was performed by implementing TSO to overcome 23 classic functions with several values of these parameters. The maximum iteration in the first sub-experiment was set at 10, 20, and 30. The result is displayed in Table 4. In the second sub-experiment, the swarm size was set at 10, 15, and 20. The result is displayed in Table 5. In the third experiment, the number of seeds was set to 3, 6, and 9. The result is displayed in Table 6.

Table 4 indicates that all functions have achieved an acceptable member in the low iteration. In almost all functions, the result produced by TSO in the low iteration circumstance was still competitive compared with the result produced by other metaheuristics, as seen in Table 2. Moreover, convergence was achieved in the early iteration in nine functions (Schwefel 2.22, Schwefel, Penalized, Penalized 2, Shekel Foxholes, Six Hump Camel, Branin, Hartman 3, and Hatman 6). Among these nine functions, one function was a high dimension unimodal function, two functions were high-dimension multimodal functions, and six functions were fixed-dimension multimodal functions.

Table 5 indicates that the increase in swarm size after five members is insignificant in improving the member quality in almost all functions. Stagnancy occurred in nine functions because the optimal global member was achieved. Stagnancy also occurred in five functions, although a globally optimal member had yet to be achieved. Less significant improvement occurred in nine functions.

Kusuma and Dinimaharawati

Table 4: Performance of TSO with several values of maximum iteration

Function	Average Fitness Score		
	$t_{max} = 10$	$t_{max} = 20$	$t_{max} = 30$
1	2.6467	0.0001	0.0000
2	0.0000	0.0000	0.0000
3	7.7478×10^2	1.3350	0.0034
4	1.8470	0.0259	0.0003
5	1.3821×10^2	4.8819×10^1	4.8819×10^1
6	1.1822×10^1	8.0214	8.1741
7	0.0193	0.0043	0.0019
8	-4.0668×10^3	-4.2352×10^3	-4.6166×10^3
9	3.5869×10^1	0.0024	0.0000
10	0.7767	0.0020	0.0000
11	0.5366	0.0060	0.0035
12	0.7894	0.6502	0.6806
13	3.2482	2.9358	2.8653
14	4.4341	4.9429	3.6936
15	0.0017	0.0006	0.0005
16	-1.0304	-1.0309	-1.0308
17	0.3989	0.3984	0.3984
18	6.6955	3.0042	3.0079
19	-0.0495	-0.0495	-0.0495
20	-3.2050	-3.239	-3.2961
21	-5.7886	-7.2599	-7.9212
22	-5.8045	-7.1234	-7.5447
23	-6.1753	-6.9030	-7.8198

Table 5: Performance of TSO with several values of swarm size

Function	Average Fitness Score		
	$N(X) = 10$	$N(X) = 15$	$N(X) = 20$
1	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000
3	0.0000	0.0000	0.0000
4	0.0000	0.0000	0.0000
5	4.8717×10^1	4.8717×10^1	4.8713×10^1
6	6.8744	6.4879	6.2007
7	0.0007	0.0005	0.0004
8	-5.7240×10^3	-5.2949×10^3	-5.3398×10^3
9	0.0000	0.0000	0.0000
10	0.0000	0.0000	0.0000
11	0.0000	0.0000	0.0000
12	0.4685	0.4531	0.3879
13	2.6890	2.5259	2.4414
14	2.4780	1.1478	1.0927
15	0.0014	0.0004	0.0004
16	-1.0316	-1.0316	-1.0316
17	0.3981	0.3981	0.3981
18	3.0015	3.0000	3.0000
19	-0.0495	-0.0495	-0.0495
20	-3.3071	-3.3038	-3.3137
21	-9.2349	-9.4629	-9.6812
22	-9.6167	-1.0084×10^1	-1.0193×10^1
23	-9.1412	-1.0034×10^1	-1.0293×10^1

Table 6: Performance of TSO with several values of number of seeds

Function	Average Fitness Score		
	n(C) = 3	n(C) = 6	n(C) = 9
1	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000
3	0.0005	0.0000	0.0000
4	0.0000	0.0000	0.0000
5	4.8850×10^1	4.8775×10^1	4.8775×10^1
6	8.5787	7.7692	7.4645
7	0.0033	0.0012	0.0005
8	-4.4692×10^3	-4.6142×10^3	-4.8921×10^3
9	0.0000	0.0000	0.0000
10	0.0000	0.0000	0.0000
11	0.0000	0.0000	0.0000
12	0.7405	0.6131	0.5368
13	2.9124	2.7398	2.7218
14	4.7629	5.9951	3.5355
15	0.0023	0.0015	0.0013
16	-1.0310	-1.0310	-1.0316
17	0.3987	0.3986	0.3982
18	3.0053	3.0007	3.0009
19	-0.0495	-0.0495	-0.0495
20	-3.2625	-3.2822	-3.2983
21	-8.0977	-6.3341	-8.3292
22	-7.0566	-8.3640	-7.9018
23	-6.5042	-8.6724	-8.5935

Table 6 indicates that the increase in the number of seeds was less sensitive to the improvement of the performance of TSO. The average fitness score tended to fluctuate or remain static in almost all problems. Meanwhile, the less significant improvement occurred in six functions (Penalized, Shekel Foxholes, Branin, Quartic, Shekel 5, Shekel 7, and Shekel 10).

5. DISCUSSION

This section presents an in-depth evaluation of the relation between the result and the findings. This discourse is divided into four parts. The first part is a discourse related to the performance of the TSO and the linkage with the chosen exploitation-exploitation strategy. The second part is a discourse regarding the hyperparameter evaluation. The third part is a discourse related to the algorithm complexity of TSO. The fourth part is a discourse related to the limitation of this work, especially the metaheuristic.

The first discourse is related to the evaluation of the experiment result. TSO performed effectively in overcoming the 23 classic functions. Its performance was superior in all groups of these functions. Based on the superior result in overcoming unimodal functions, TSO performed exploitation effectively. Moreover, TSO was also good at performing exploration based on the superior result in overcoming multimodal functions, whether they were high-dimension multimodal functions or fixed-dimension multimodal functions. The performance gap between TSO and the sparing metaheuristics was also broad, especially in overcoming high-dimension functions. This gap needs to be more comprehensive in overcoming fixed-dimension functions.

The superiority of TSO proves that the strategy implemented in TSO is better than the strategy implemented in the sparing metaheuristics. First, implementing multiple searches is proven better than a single search because each has its strengths and weaknesses. Second, each member needs to perform multiple searches in every iteration. Third, the tournament-based approach was better than the sequential-based approach.

The second discourse is related to the sensitivity analysis of the hyperparameter. This work evaluated three parameters: maximum iteration, swarm size, and the number of seeds. In general, the increase in maximum iteration improves the quality of the member. In the low maximum iteration, increasing maximum iteration improves the member mostly in overcoming the unimodal functions. On the other hand, the increase in the maximum iteration is less significant in improving the quality of the member. On the contrary, the swarm size does not improve the member quality in overcoming unimodal functions. The increase in swarm size improves the member quality, mostly in overcoming the multimodal functions. However, this improvement is also insignificant because the near-optimal or optimal global member has been found in the low maximum iteration and low swarm size circumstances. Increasing the number of seeds improves the quality of members in overcoming several functions. These functions can be found in both unimodal and multimodal functions. However, the improvement could be more significant.

The third discourse is related to algorithm complexity. The complexity of TSO can be displayed as $O(3t_{max}.n(X).n(C))$. Based on this presentation, the complexity is linear to one of three parameters: the maximum iteration, swarm size, or the number of seeds. Fortunately, the computational process of TSO is still competitive because the acceptable member has reached a low maximum iteration, low swarm size, and a low number of seeds.

The fourth discourse is related to the limitation of the algorithm and this work. There are several limitations regarding this algorithm and this work. First, TSO still needs to find the final member, the optimal global member or near-optimal member in overcoming six functions. These six functions are two high-dimension unimodal functions (Rosenbrock and Step), three high-dimension multimodal functions (Schwefel, Penalized, and Penalized 2), and one fixed-dimension multimodal function (Hartman 3). This fact strengthens the no-free-lunch theory. Although TSO is superior among the sparing metaheuristics, there are still problems where TSO needs to find the optimal global member. Meanwhile, there is an opportunity to find the optimal global member for these functions by setting the system to high maximum iteration and swarm size. However, this scenario was not performed in this work as its focus is on the low maximum iteration and low swarm size. This limitation can be used as a baseline to improve the current form of TSO so that the improved version can overcome these six functions in future studies. This work has a limitation on implementing TSO in overcoming theoretical optimization problems only. In some studies, the proposed metaheuristic has been challenged to overcome the theoretical problems only, while in other studies, the metaheuristic has also been challenged to overcome practical problems.

6. CONCLUSION

The development and evaluation of a new swarm-based metaheuristic, treble search optimizer (TSO) has been introduced in this paper. Referring to its name, the central concept of TSO is performing three searches: two directed searches and one random search. Each member in every iteration performs these three searches. Several seeds are generated from each search. The experiment result presents that TSO performed

effectively in overcoming 23 classical functions. TSO outperformed five shortcoming metaheuristics chosen as the sparing metaheuristics in this work. TSO was better than SMA, HPKA, MLBO, GSO, and TIA in overcoming 21, 21, 23, 23, and 17 functions, respectively. TSO could find the optimal global member of seven functions in low maximum iteration and low swarm size circumstances. When the swarm size was set to moderate, there were three more functions where TSO can find their optimal global member. TSO also performed the best member of 22 functions out of 23 functions.

Future studies can be conducted in several ways. Improvement is still open for TSO, especially in finding the optimal global member of the six functions that TSO still needs to find in this work. More theoretical experiments should be performed to enrich the investigation of the strengths and weaknesses of TSO. Moreover, future studies can also be performed by implementing TSO to overcome many kinds of practical optimization problems.

ACKNOWLEDGMENT

REFERENCES

[1] Ramezanpour MR, Farajpour M. (2022) Application of artificial neural networks and genetic algorithm to predict and optimize greenhouse banana fruit yield through nitrogen, potassium and magnesium. PLoS ONE, 17(2): e0264040. https://doi.org/10.1371/journal.pone.0264040.

[2] Thammachantuek I, Ketcham M. (2022) Path planning for autonomous mobile robots using multi-objective evolutionary particle swarm optimization. PLoS ONE, 17(8): e0271924. https://doi.org/10.1371/journal.pone.0271924.

[3] Sivakumar R, Angayarkanni SA, Ramana RYV, Sadiq AS. (2022) Traffic flow forecasting using natural selection based hybrid bald eagle search-grey wolf optimization algorithm. PLoS ONE, 17(9): e0275104. https://doi.org/10.1371/journal.pone.0275104.

[4] Hoballah A, Azmy AM. (2023) Constrained economic dispatch following generation outage for hot spinning reserve allocation using hybrid grey wolf optimizer. Alexandria Engineering Journal, 62: 169-180. https://doi.org/10.1016/j.aej.2022.07.033.

[5] Sowmya R, Sankaranarayanan V. (2022) Optimal scheduling of electric vehicle charging at geographically dispersed charging stations with multiple charging piles. International Journal of Intelligent Transportation Systems Research, 20: 672-695. https://doi.org/10.1007/s13177-022-00316-2.

[6] Suyanto, Ariyanto AA, Ariyanto AF. (2022) Komodo mlipir algorithm. Applied Soft Computing, 114: 108043. https://doi.org/10.1016/j.asoc.2021.108043.

[7] Dehghani M, Hubalovsky S, Trojovsky P. (2021) Nortehrn goshawk optimization: a new swarm-based algorithm for solving optimization problems. IEEE Access, 9: 162059-162080. doi: 10.1109/ACCESS.2021.3133286.

[8] Faramarzi A, Heidarinejad M, Mirjalili S, Gandomi AH. (2020) Marine predators algorithm: a nature-inspired metaheuristic. Expert Systems with Applications, 152: 113377. https://doi.org/10.1016/j.eswa.2020.113377.

[9] Kusuma PD, Dinimaharawati A. (2022) Hybrid pelican komodo algorithm. International Journal of Advanced Computer Science and Applications, 13(6): 46-55. https://dx.doi.org/10.14569/IJACSA.2022.0130607.

[10] Dehghani M, Montazeri Z, Trojovska E, Trojovsky P. (2023) Coati optimization Algorithm: A new bio-inspired metaheuristic algorithm for solving optimization problems. Knowledge-Based Systems, 259: 110011. https://doi.org/10.1016/j.knosys.2022.110011.

[11] Akbari MA, Zare M, Azizipanah-abarghooee R, Mirjalili S, Deriche M. (2022) The cheetah optimizer: a nature-inspired metaheuristic algorithm for large-scale optimization problems. Scientific Reports, 12: 10953. https://doi.org/10.1038/s41598-022-14338-z.

[12] Braik MS. (2021) Chameleon swarm algorithm: a bio-inspired optimizer for solving engineering design problems. Expert Systems with Applications, 174: 114685. https://doi.org/10.1016/j.eswa.2021.114685.

[13] Zeidabadi FA, Doumari SA, Dehghani M, Malik OP. (2021) MLBO: mixed leader based optimizer for solving optimization problems. International Journal of Intelligent Engineering and Systems, 14(4): 472-479. doi: 10.22266/ijies2021.0831.41.

[14] Zeidabadi FA, Dehghani M, Malik OP. (2021) RSLBO: random selected leader based optimizer. International Journal of Intelligent Engineering and Systems, 14(5): 529-538. doi: 10.22266/ijies2021.1031.46.

[15] Dehghani M, Trojovsky P. (2022) Hybrid leader based optimization: a new stochastic optimization algorithm for overcoming optimization applications. Scientific Reports, 12: 5549. https://doi.org/10.1038/s41598-022-09514-0.

[16] Kusuma PD, Novianty A. (2023) Total interaction algorithm: a metaheuristic in which each agent interacts with all other agents. International Journal of Intelligent Engineering and Systems, 16(1): 224-234. doi: 10.22266/ijies2023.0228.20.

[17] Noroozi M, Mohammadi H, Efatinasab E, Lashgari A, Eslami M, Khan B. (2022) Golden search optimization algorithm. IEEE Access, 10: 37515-37532. https://doi.org/10.1109/ACCESS.2022.3162853.

[18] Dehghani M, Hubalovsky S, Trojovsky P. (2022) A new optimization algorithm based on average and subtraction of the best and worst members of the population for solving various optimization problems. PeerJ Computer Science, 8: e910. https://doi.org/10.7717/peerj-cs.910.

[19] Swan J, Adriaensen S, Brownlee AEI, Hammond K, Johnson CG, Kheiri A, Krawiec F, Merelo JJ, Minku LL, Ozcan E, Pappa GL, Garcia-Sanchez P, Sorensen K, Vob S, Wagner M, White DR. (2022) Metaheuristics in the large. European Journal of Operational Research, 297(2): 393-406. https://doi.org/10.1016/j.ejor.2021.05.042.

[20] Li S, Chen H, Wang M, Heidari AA, Mirjalili S. (2020) Slime mould algorithm: A new method for stochastic optimization. Future Generation Computer Systems, 111: 300-323. https://doi.org/10.1016/j.future.2020.03.055.

THE EFFECTIVENESS OF 6T BEAMFORMER ALGORITHM IN SMART ANTENNA SYSTEMS FOR CONVERGENCE ANALYSIS

SADIYA THAZEEN* AND MALLIKARJUNASWAMY SRIKANTASWAMY

*Department of Electronics and Communication Engineering,
JSS Academy of Technical Education, Bengaluru, India*

*Corresponding author: thazeen.sadiya@gmail.com

ABSTRACT: Recently, the need for more capacity in wireless networks has motivated this current research towards the creation of standards and algorithms that selectively take advantage of space. The development of smart antenna arrays and related beamforming algorithms has received a lot of attention. Cell Planning is an important process in networking, which is used to ensure coverage and avoid interference. Also, cell planning plays a vital role in the placement of base stations in a network. The communication between the base station (BS) and mobile station can happen either using a single antenna or an array of antenna elements. In the case of using a single antenna, if the Electromagnetic (EM) wave has low SNR, then BS cannot decode the data and drops the signal. Conversely, when an array is used, a signal with low SNR also falls on the base station, and due to multiple delayed copies of the same signal, the data gets decoded successfully. With Space Division Multiple Access (SDMA), the frequencies allotted for mobile communication are reused to provide channel access to multiple users at the same time preserving the allowable reuse distance in network architecture, thus increasing the channel capacity and facilitating multiple users separated by a distance at the same time with frequency reuse. The smart antenna system at the base station performs the transceiver function. The transmission phase uses the output from the reception i.e., the detected user direction radiates a beam towards the desired user for communication to narrow the beam. The proposed 6T Beamformer method is a six-tap-based system with three taps having fixed step sizes and the other three having variable step sizes. With the execution of each tap or module, better convergence and quality of service are achieved. In the result analysis, the proposed method is compared with existing high-performing algorithms like LMS, Griffiths, and VSSLMS against Mean Square Error (MSE) to show that it converges faster at the 9th iteration which is better than others in all the probabilities.

ABSTRAK: Dewasa ini, keperluan terhadap lebih kapasiti dalam rangkaian tanpa wayar menjadi motivasi kepada kajian terkini dalam membentu piawai dan algoritma yang menjimatkan ruang. Pembangunan tata susun antena pintar dan algoritma pembentukan pancaran telah mendapat perhatian ramai. Merancang sel adalah proses penting dalam jaringan, bagi memastikan liputan terhasil dan mengelak dari gangguan. Juga, merancang sel memainkan peranan penting dalam menempatkan tapak stesen dalam rangkaian. Komunikasi antara stesen pusat (BS) dan stesen bergerak dapat berlaku samada menggunakan antena tunggal atau elemen tata susunan antena. Dalam kes antena tunggal, jika gelombang Elektromagnetik (EM) mempunyai SNR rendah, BS tidak dapat menafsirkan kod data dan signal akan terabai. Sebaliknya, apabila susun atur digunakan, signal dengan gelombang SNR rendah akan terus ke stesen pusat dan disebabkan beberapa gelombang sama yang tertunda, data dapat ditafsir dengan sempurna. Melalui Capaian Pelbagai Pembahagi Ruang (SDMA), frekuensi yang ditimbulkan bagi komunikasi bergerak telah diguna balik bagi menyediakan kemasukan saluran kepada pelbagai

pengguna pada waktu sama memelihara jarak guna balik yang dibenarkan dalam binaan rangkaian, oleh itu menambah kapasiti saluran dan membantu gandaan pengguna yang dipisahkan oleh jarak dengan kekerapan guna balik pada masa sama. Sistem antena pintar di stesen pusat pula menjalankan fungsi pemancar. Fasa pemancaran ini menggunakan pengeluaran dari penerima iaitu, pengguna yang dikesan dari arah pancaran, akan memancarkan gelombang kepada pengguna yang memerlukan komunikasi, ini dapat mengecilkan jarak pancaran. Kaedah yang dicadangkan ini menghasil pancaran 6T iaitu sistem berdasarkan-enam-tap di mana tiga tap mempunyai saiz langkah yang tetap dan tiga lagi mempunyai saiz langkah berubah. Dengan pelaksanaan ini setiap tap atau modul mempunyai penumpuan yang lebih baik dan servis yang berkualiti terhasil. Dapatan kajian menunjukkan, kaedah yang dicadangkan dapat dibandingkan dengan algoritma berprestasi tinggi sedia ada seperti LMS, Griffiths, dan VSSLMS berbanding min kuasa dua ralat (MSE) bagi menunjukkan ia tertumpu lebih laju pada iterasi ke 9, iaitu lebih baik daripada ke semua kebarangkalian.

1. INTRODUCTION

Communication systems involving 5G not only generate higher performance in convergence, capacity, and data rate but also reduce the latency, computational complexity, and cost. This research proposes a 2-stage suboptimal beamforming method to provide a solution to the problems encountered in the physical layer of the 5G system [1,2]. The role of adaptive algorithms in digital signal processing is pivotal. This study considers the application of the three existing high-performing adaptive algorithms in the quality of voice-over-IP with modifications in the step size, as step size contributes greatly to the convergence speed of the algorithms [3]. The grade of the antenna array will be better if the radiation is sent to the desired user [4]. The grade is directly dependent upon frequency or transmitter usage [5]. All the mobile users in the hexagonal cell will receive the signal with the same frequency at the same time but the angle of variation will distinguish the sources. The new arrays send the narrow beam toward actual mobile users to save an amount of energy [6]. The phase shifts are used to calculate the amount of energy reduced from the initial stage [7]. The concept of scheduler-based transmission will make the antenna array elements rest/constant at regular intervals and assign phase shifts to array elements to form the main beam [8]. The capacity for communication can be increased with the help of intelligent systems [9]. Digital signal processing (DSP) is that crucial determining factor in advancing wireless communications technology and increasing the capacity of limited RF spectrum [10].

Multiple techniques are used to decode the signal based on either summation of signals or by using maximum Signal to Noise ratio (SNR) so that the signal formation is better at the receiving end [11]. During the reception phase, the electromagnetic (EM) waves are obtained as the original signal and delayed versions of the original signal and are added up to get the main signal data. The maximum SNR method will consider each signal, then find the SNR and choose the signal which has the maximum SNR [12]. The phase shifts are computed by taking the actual signal, noisy signal, and jammed signals to form the total signal. The total signal phase shifts are multiplied by the initial set of phase shifts to form the array output. In a legacy system, an array of dipole elements is arranged linearly. The energy wastage is greater whenever the radiation falls in other directions. To overcome that, the following contributions are made:

- The receiving phase's result or the user direction was detected and is used in the transmission phase to guide a beam in the direction of the intended user.

- Then, the 6T Beamformer, a six-tap device used in the suggested technique, has three taps with fixed step sizes and the remaining three with variable step sizes.

- Better convergence is achieved with each tap or module that is executed in the result analysis.

2. LITERATURE REVIEW

Energy Efficient Multicasting via Smart Antennas in Multipath Environments has been proven by Tong and Ramanathan [13]. Nodes can direct their transmission energy in the appropriate directions by utilizing an adaptive antenna array, which allows them to conserve energy. The fact that this work functions in rich scattering surroundings is a noteworthy aspect. The issue was framed as a non-linear issue. Furthermore, it suggested two heuristics that are effective in computing. According to the numerical simulations, the two heuristics can significantly reduce power consumption when contrasted to the single antenna scenario. To fix the design problem and test the performance of the developed approaches, researchers likewise utilized simulation annealing.

Ali and Hassan [14] used a microstrip antenna array to show a hybrid technique. This study contains the theoretical hybrid method, which integrates two methods to boost their effectiveness. This is accomplished by making use of two algorithms' strengths while avoiding their weaknesses. The least mean square and sample matrix inversion methods are those two techniques. The resonant frequency of the microstrip antenna, which operates between 1700 and 1950 MHz, is 1850 MHz.

The influence of different parameters on minimal mean square error has been established by Anjaneyulu et al. [15]. This evaluation considers the Least Mean Square technique settings as well as antenna variables including step-size fluctuation and signal-to-noise variation. The major targets of all parameter changes are mean square error, null depth, convergence speed, and beamforming efficacy. The LMS with suitable antenna input variables resulted in successfully displaying information at the location of the interferers and minimizing the error, as seen in all calculated values.

Mayyas and Aboulnasr [16] presented the sensor which was placed at a distance, d, from each other. The weights are computed by making use of directivity, white noise, and interference ratio. Once the phase shifts were computed using the previous version of phase shifts, step size, reference signal, and error signal, the trace of the autocorrelation was found. From the weights, the maximum value has been found and then normalized phase shifts are obtained.

Saqhib et al. [17]. describes the path establishment process in a wireless sensor network (WSN) for sending data packets and highlights how repeated participation in multiple paths can lead to reduced network lifetime. The paper seeks to provide recapitulation of WSN, the effects of lifetime ratio, and an energy-efficient routing protocol through numerical survey.

Mallaiah et al. [18] demonstrated the Leaky LMS algorithm, which was responsible for finding the phase shifts and then applying the components to create the beam. Leaky LMS was developed on top of the LMS method with the variation in the computation of step size. The step size depends upon the Eigenvalue and then the maximum value of Eigen is found. The new phase shifts were calculated through earlier step size and phase shift.

Beamforming by Small-Spacing Microphone Arrays through Constrained LASSO has been demonstrated by Wang et al. [19]. The Capon approach also determines the cross-correlation between the BS signal and the real reference signal. Toeplitz execution will be used on the phase shifts to discover new phase shifts. Every component of the array was therefore subjected to the phase shifts to create the main beam.

3. SYSTEM MODEL

Consider a uniform linear array that has all dipole elements with each element separated from the other at an equal distance.
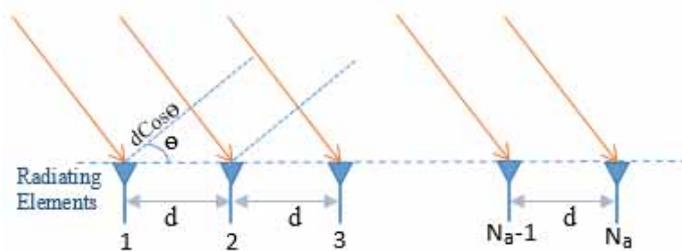

Fig. 1: Uniform Linear Array elements.

Figure 1 shows radiating or antenna elements arranged linearly. Na is the number of antenna elements separated by an equal distance d. The electromagnetic (EM) wave along with the delayed versions falls on the array elements. The representation of the electromagnetic wave will create the delay vector having Na* 1 with Na as the number of elements. The disturbance can be suppressed with the help of multiple wavelengths. The delayed electromagnetic wave vector can be summarized as Eq. (1) below,

$$EM(\theta d) = \begin{bmatrix} 1 \\ e^{-j\pi \sin(\theta d)} \\ . \\ . \\ . \\ e^{-j\pi(N_a-1)\sin(\theta d)} \end{bmatrix} \tag{1}$$

where, Na is the number of array elements, and theta is the direction of impinging electromagnetic wave.

The reference signal is generated for the mobile station which can be a sine or cosine reference signal with Ns number of samples and is provided by the values as Eq. (2)

$$r(i) = cos(pi * i / Ts) \tag{2}$$

where,

i = sample index

Ts = number of samples

The reference signal would be better if the sampling is done using Nyquist criteria in which the sampling period should have a certain minimum value. The sampling frequency should be greater or equal to the minimum value of $1/2*Ts$. The index will be varied between $0 <= I <= Ns$ where Ns is the maximum number of samples.

The array manifold vector is computed for the set of N users and the combination of direction vectors across interference user's Njammer directions $\{\theta_1, \theta_2, \ldots \ldots \ldots \ldots \ldots \theta_{Njammer}\}$ which is stated in Eq. (3),

$$J_M = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ e^{-j\pi\sin(\theta_1)} & e^{-j\pi\sin(\theta_2)} & & e^{-j\pi\sin(\theta_{Nj})} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ e^{-j\pi(N_a-1)\sin(\theta_1)} & e^{-j\pi(N_a-1)\sin(\theta_2)} & & e^{-j\pi(N_a-1)\sin(\theta_{Nj})} \end{bmatrix} \quad (3)$$

The jammer and noise signals are random and can be generated as Nj number of samples. The electromagnetic wave received at the base station will have pure directional vector and reference signal combinations. The jammer signal, along with directional vectors across all the directions and noise signal, can be provided as Eq. (4).

$$BSRS_{signal} = D_v(\theta_0)RS + JS * Jm + NS$$

Where,

$D_v(\theta_0) = direction\ vector\ for\ an\ angle\ \theta_o$

$RS = Total\ reference\ signal\ in\ matrix\ format\ generated$

$at\ the\ mobile\ station \qquad (4)$

$JS = Total\ interference\ signal\ samples\ which\ are$

$random\ in\ nature$

$Jm = Jammer\ Signal\ for\ all\ jammer\ directional\ vectors$

$NS = Noise\ Signal\ Vector\ Matrix$

The BSRS matrix along with the Hermitian BSRS matrix can be multiplied to have an auto-correlation matrix which is represented in Eq. (5)

$$R_{BSRS_{signal}} = BSRS_{signal} * BSRS_{signal}^H$$

Where,

$BSRS_{signal} = total\ recived\ signal\ at\ BS \qquad (5)$

$BSRS_{signal}^H = hermitian\ transpose\ of\ BSRS$

The error signal at the base station has been found through the combination of all electromagnetic waves as an array total which is referred to in Eq. (6)

$$AT(n) = Ps^H(n)\ BSRS(n)$$

Where

$Ps^H(n) = hermitian\ transpose\ of\ array\ phase$

$\qquad shifts\ which\ are\ applied\ to\ individual \qquad (6)$

$\qquad\qquad phase\ shifters$

$BSRS(n) = BSRS\ for\ n^{th}\ sample$

The error vector can be obtained using the difference among the reference signal which can be computed as Eq. (7)

$$ev(n) = \left| BSRS(n) - AT(n) \right|$$

(7)

Each of the methods will have to compute the phase shifts in such a way that the error vector value has to be minimized while considering some of the existing algorithms, such as Least Mean Square (LMS) [16], which is the robust approach to computing the weights. Then, the Griffiths algorithm [20] extracts a particular signal among many signals' incident at BS on the antenna array elements. Finally, the Variable step size Griffiths method [21] is formed by combining LMS along with the VSSLMS method to improve the performance. The step size is varied between 0 and the upper limit so that better and faster convergence can be achieved. The simulation results and their analysis are described in section 5.

4. PROPOSED 6T BEAMFORMER ALGORITHM

Figure 2 shows the functional block diagram of the Smart Antenna system at the BS where the antenna array picks up the signal from the user trying to access the communication channel to establish communication. The user direction is estimated using actual and delayed versions of the signal and the output is passed onto the beamforming block of the system where the advanced signal processing computes the weights for each iteration considering the output feedback and radiates a beam that is maintained sharp and noise-free to the best possible extent for the communication link to be of high quality.
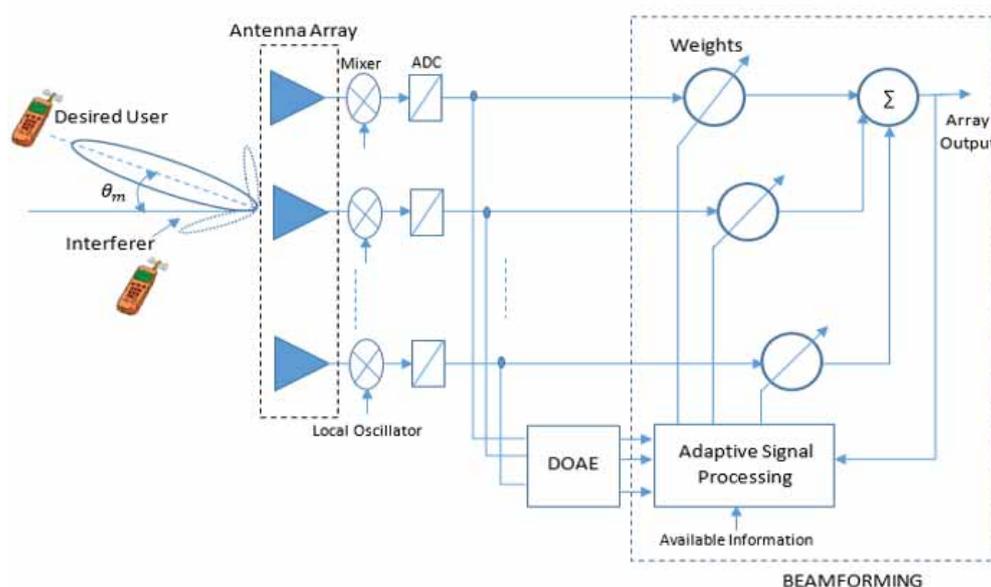
Fig. 2: Overall diagram for Smart Antenna system.

The proposed 6T Beamformer [22] is a six-tap phase shifts method with the combination of leaky Least Mean Square (LLMS) and Variable Step Size Least Mean Square (VSSLMS). Convergence is an important performance parameter wherein the convergence factor is the ith number of iterations at which the mean square error reaches zero. The performance of the algorithm is improved by lowering the iteration count and increasing the convergence factor.

Fig. 3: Six-Tap Phase Shifts Applier.

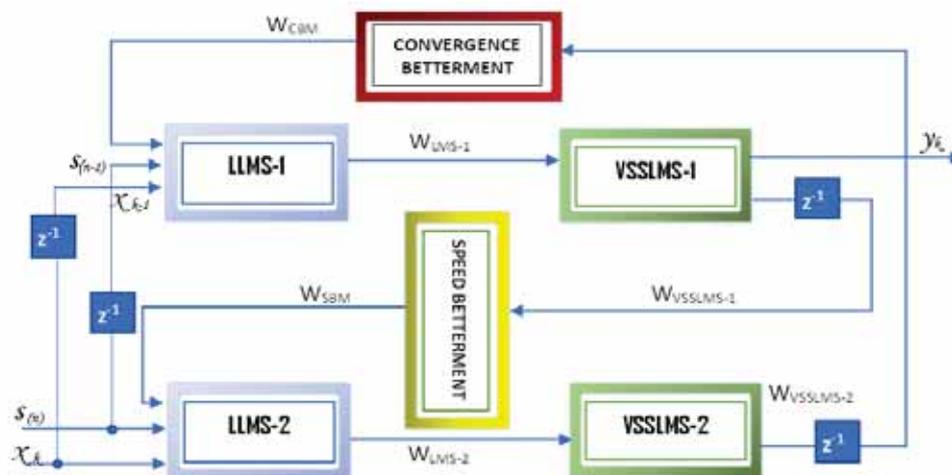Figure 3 shows the arrangement of six tap phase shifts applier which can be divided into 7 different blocks. In the first block, the phase shifts are computed using a leaky Least Mean Square (LLMS) method i.e. LLMS (1) to obtain k^{th} phase shifts. A Speed Better module will provide the computation of weights in a better fashion which computes k+1 phase shifts and then it is continued to Variable Step Size Least Mean Square (VSSLMS-1) to compute k+3. It is followed by Convergence Betterment (CB) method which will compute k+4 phase shifts which will then be sent to LLMS (2) and VSSLMS (2) to have k+5, k+6. Due to the execution of the modules, the convergence will be improved and then the error will be reduced to a lower value.

4.1 Leaky Least Mean Square

This method is one of the versatile beamforming methods which makes use of phase shifts that are applied to specific antenna elements followed by the formation of an array factor. But LMS faces certain disadvantages which are a) the rate at which convergence is achieved is quite low and b) drifting of weight causes more error. To overcome these disadvantages, the leaky factor is computed and used in the phase shifts (Eq.(8-9)). The phase shifts of the Leaky Least Mean Square (LLMS) magnitude will be kept within bounded limits so that the drifting issue can be resolved.

The step size for the LLMS module is computed as Eq. (8)

$$SS(n)_{LLMS} = \alpha * SS(n+1)_{LLMS} + \gamma * e(n)^2 \tag{8}$$

where,

α = initial step magnitude between 0 to 1

γ = average square of the error signal

The phase shifts vector used by LLMS can be defined as Eq. (9)

$$ps(n+1) = [1 - 2 * SS(n)] ps(n) + e * (n) \ BSRS(n) \tag{9}$$

4.2 Speed Betterment

The convergence speed can be improved using speed betterment-based phase shift computation which will take into consideration cross-correlation, step size, and previous phase shifts, as stated in Eq. (10).

$$P_{SB}(n+1) = P_{SB}(n) + ([I - 2\mu_{SI}BSRS_{xx}]^{-1}[P(n) - 2\mu_{SI}BSRS_{xs}]) \tag{10}$$

The speed betterment phase shifts are computed using previous phase shifts, the base station received signal and then step size can be computed using Eq. (11).

$$\mu_{SB} = \frac{1}{3tr(R_{SB})} \tag{11}$$

where R_{SB} is the autocorrelation of Base Station Received Signal (BSRS) and $tr(R_{SB})$ is the trace of autocorrelation matrix, which is obtained by performing the square root of the sum of square values of each principal element.

4.3 Variable Step Size Least Mean Square

When the SNR of the signals are varying, the constant value of step size is not helpful. The step size can be obtained using a fast and robust variation of the VSSLMS method by making use of the normalized sigmoid method. It overcomes the convergence by adjusting the step size value in an adaptive fashion which is used in the computation of phase shifts.

The VSSLMS step size is varied to have better convergence which is expressed in Eq. (12).

$$\begin{aligned} SS_s(n+1) &= SS_{supper} \quad if \ SS_s(n+1) > SS_{supper} \\ &= 0 \qquad\quad if \ SS_s(n+1) < 0 \\ &= SS_s(n+1) \ otherwise \end{aligned} \tag{12}$$

where,

SS_{supper} is the upper limit of step size.

The upper limit size is computed by making use of the Eigenvalue computation of auto-correlation which is stated in Eq. (13).

$$SS_{upper} = \frac{2 + ev_m}{AV} \tag{13}$$

where,

$ev_m = \max \textit{value of eigen values}$

$AV = \textit{array value of mean square error} (MSE)$

The phase shifts of the VSSLMS method can be computed using the following Eq. (14).

$$ps(n+1) = ps(n) + SS(n)\,e*(n)\ BSRS(n) \,_{\cdots\cdots} \tag{14}$$

4.4 Convergence Betterment

This module helps in increasing the speed at which the error can be minimized by making use of a reference signal to the array output.

The phase shifts are computed using the convergence betterment module by making use of the Aitken Process, which is referred to in Eq. (15).

$$w_{CB1} = L - \frac{Nu_{LLMS}}{De_{LLMS}} \tag{15}$$

where, $Nu_{VSSGMk} = M_k M_k w_{VSSG1} + M_k N_k - M_k\, w_{VSSG1}$

$$De_{LLMSMk} = M_k M_k w_{LLMS1} + M_k N_k$$
$$- 2M_k w_{LLMS1} - N_k + w_{LLMS1}$$

$$L = ACS_k SCS_k w_{LLMS1} + ACS_k CCS_k + CCS_k \text{ where,}$$

$$ACS_k = I - 2\mu_{CB} R_{xx}$$

$$CCS_k = 2\mu_{CB} R_{xs}.$$

5. SIMULATION RESULTS

In this section, simulation results of the proposed 6T Beamformer are discussed in detail along with the comparison of the algorithm against the existing ones to mean square error which is also a measure of the converging speed of the algorithms.

This paper considers a few existing high-performance algorithms such as Least Mean Square (LMS) [15], Griffiths algorithm [19] and Variable step size Griffiths method [20] for comparison of mean square error. The radiation pattern, real phase shifts and imaginary phase shifts are simulated only for the proposed 6T Beamformer as the comparison graph will have the results overlapped. The results are simulated by taking into consideration the probability of a number of jammers from one to many along with varying the number of antenna elements from few to many numbers.

Table 1 shows the probability-1 setup considering fewer array elements and one interference at 30 degrees. Figures 4 and 5 display the real and imaginary phase shifts for the experimental setup considered.

Table 1: Experimental setup for probability-1

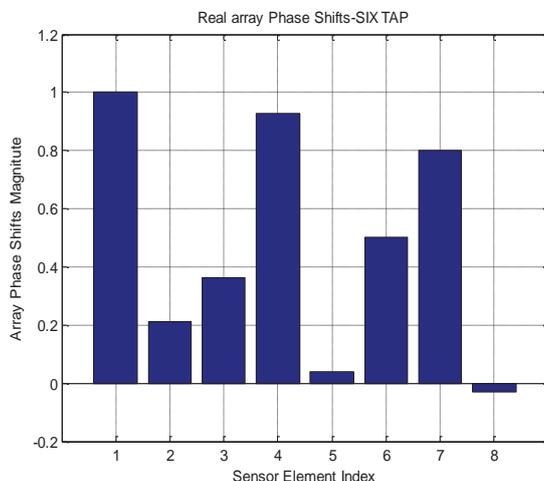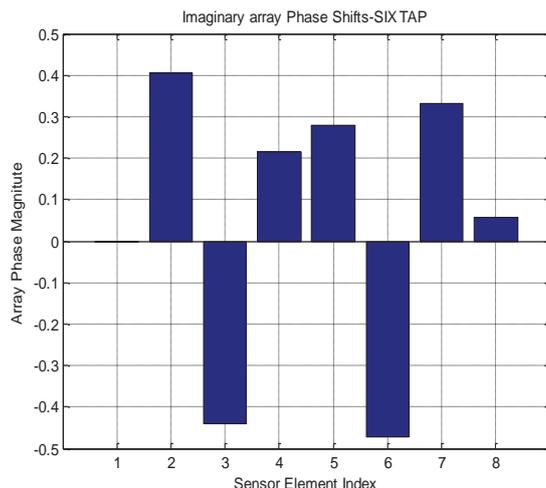| Few Antenna Elements- Single Jammer	Values
Number of Array Elements	8
Desired Direction	45
Interference Count	1
Interference Directions	30

Fig. 4: Real Array Phase Shifts.

Fig. 5: Imaginary Array Phase Shifts.

Figure 6 Reveals the radiation array pattern for the setup as mentioned in probability-1 of 45 degrees. Figure 7 shows a mean square error comparison plot for probability-1 where the proposed 6T Beamformer algorithm successfully converges faster at run count or some iterations of 11 followed by VSSG, with a run count of 45, Griffiths along with LMS with a run count of 70. Hence the proposed 6T Beamformer performs better than other methods for considering probability-1. Table 2 shows the experimental setup considering a high number of antenna elements and a single interferer at 5 degrees with the desired source at 30 degrees for probability-2.
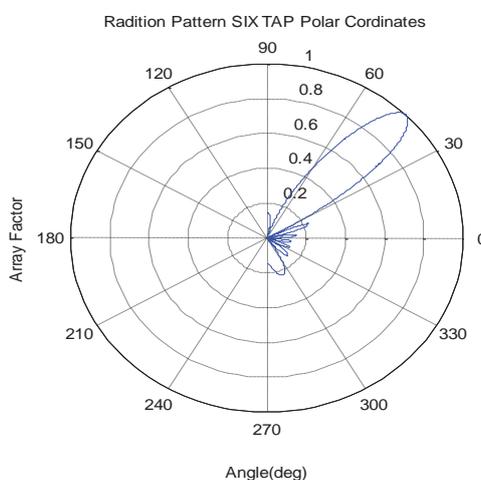
Fig. 6: Radiation pattern for 6T Beamformer at 45 degrees.

Table 2: Experimental setup for probability-2

Many Antenna Elements- Single Jammer	Values
Number of Array Elements	100
Desired Direction	30
Interference Count	1
Interference Directions	5

Fig. 7: Comparative analysis of MSE for probability-1.

Fig. 8: Real Phase Shifts for probability-2.

Fig. 9: Imaginary Phase Shifts for probability-2.

Figures 8 and 9 display the real and imaginary phase shifts for the experimental setup considered in probability-2 of 30 degrees. Figure 10 represents the antenna radiation pattern with the main beam at the desired angle of 30 degrees and interferer given no signal strength.

Fig. 10: Radiation Pattern for 6T Beamformer at 30 degrees.

Fig. 11: Comparative analysis of MSE for probability-2.

Figure 11 shows the mean square error comparison graph for probability-2 where the 6T Beamformer converges faster at just 10 iterations which means low MSE followed by VSSG converging in 34 iterations and Griffiths, LMS at 75 each. Table 2 clearly shows that the 6T Beamformer proves to perform better in the desired direction (30 degrees).
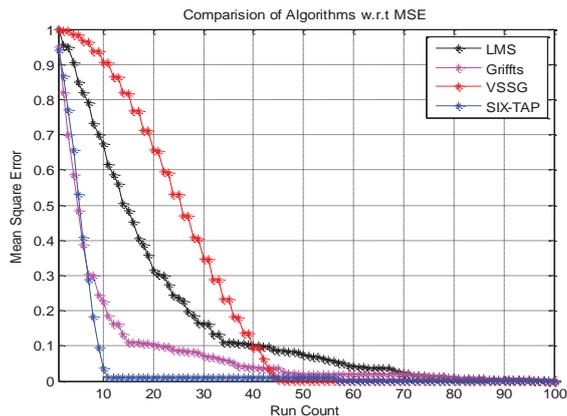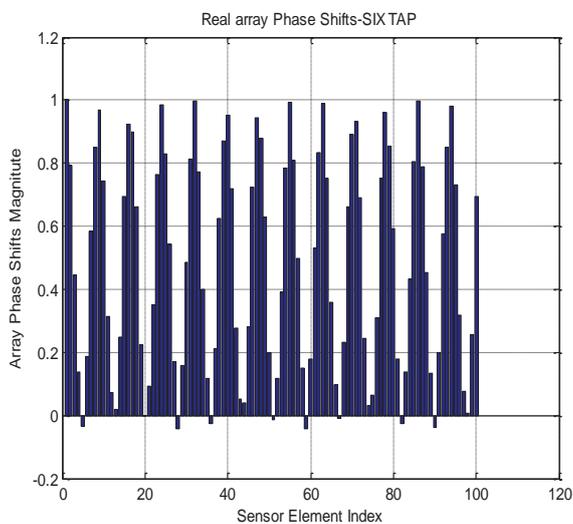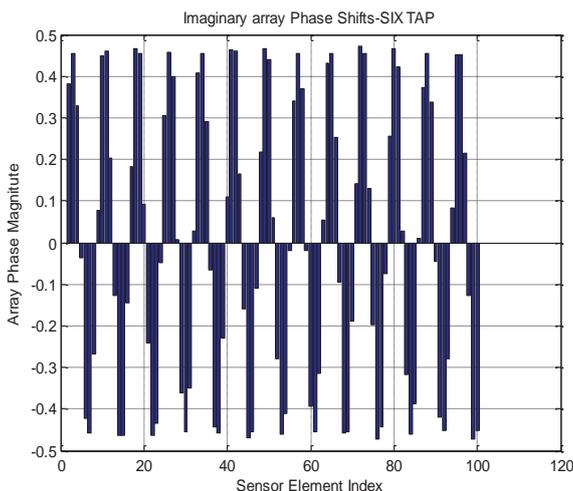
Table 3: Experimental setup for probability-3

Few Antenna Elements- Multiple Jammer	Values
Number of Array Elements	8
Desired Direction	60 Degrees
Interference Count	3
Interference Directions	[5 10 15]

Table 3 displays the experimental setup of lesser array elements considering multiple interferers. Here, consider 3 interferers at 5, 10, and 15 degrees each.

Figures 12 and 13 exhibit real and imaginary phase shifts for the experimental setup considered in probability-3. These individual phase shifts are exploited to each element of the array for the formation of the main beam in the desired direction of 60 degrees with multiple

interferers. Figure 14 shows the radiation pattern for 6T Beamformer forming a mean beam towards the desired user direction of 60 degrees while interferers are treated with nulls.
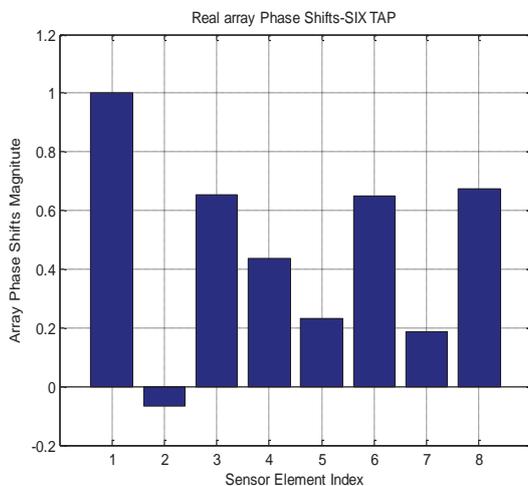
Fig. 12: Real Phase Shifts for probability-3.
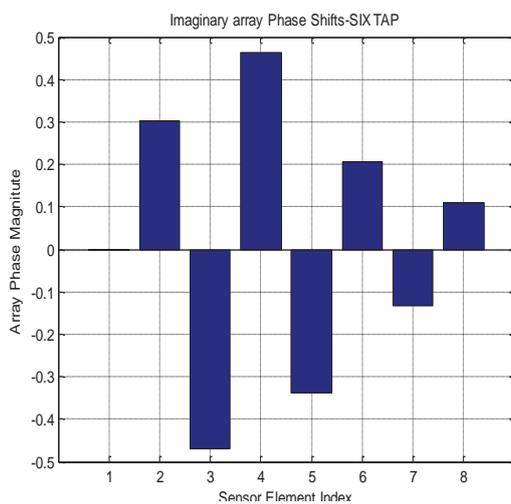
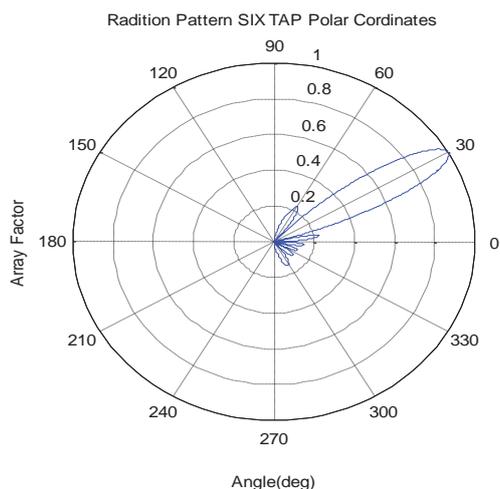Fig. 13: Imaginary Phase Shifts for probability-3.

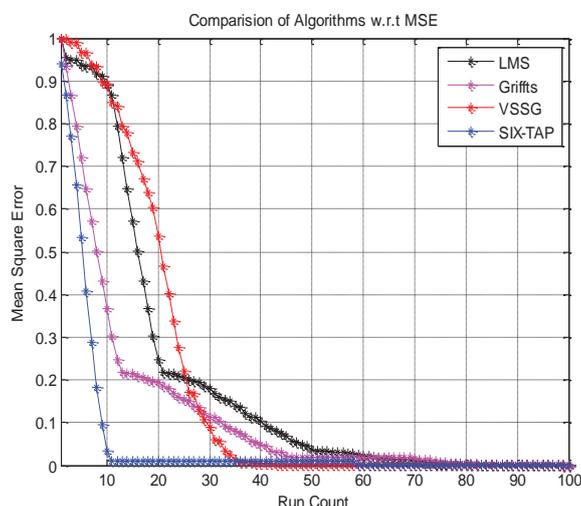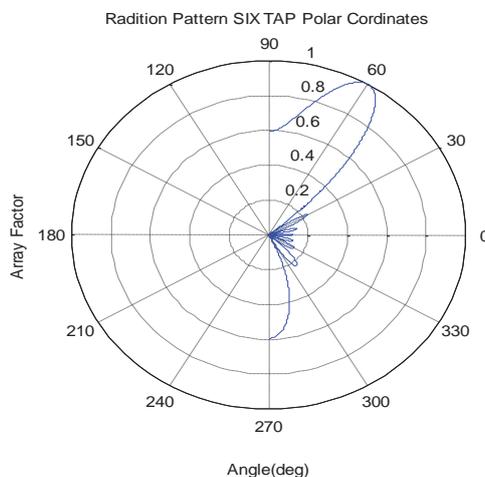Fig. 14: Radiation pattern for 6T Beamformer - probability-3.
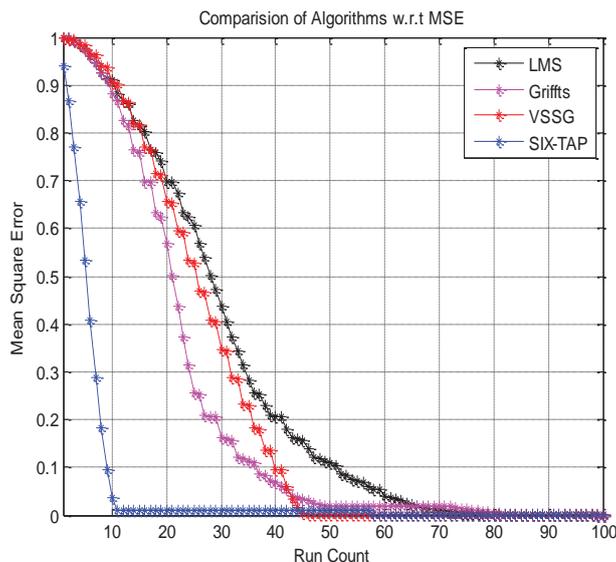
Fig. 15: Comparative Analysis of MSE for Probability-3

Figure 15 shows the mean square error comparison graph for probability-3 where the 6T Beamformer is seen converging faster, at just 10 iterations, giving low MSE followed by VSSG with a run count of 48, Griffiths and LMS converge at a run count of 80 and 78 respectively. Hence, the novel 6T Beamformer performs better than other methods for the considered probability-3. Table 4 displays the experimental setup description wherein a large number of antenna elements and multiple interferers are considered as a probability at the base station.

Table 4: Experimental setup for probability-4

Many Antenna Elements- Multiple Jammer	Values
Number of Array Elements	100
Desired Direction	15 Degrees
Interference Count	3
Interference Directions	[45 50 55]
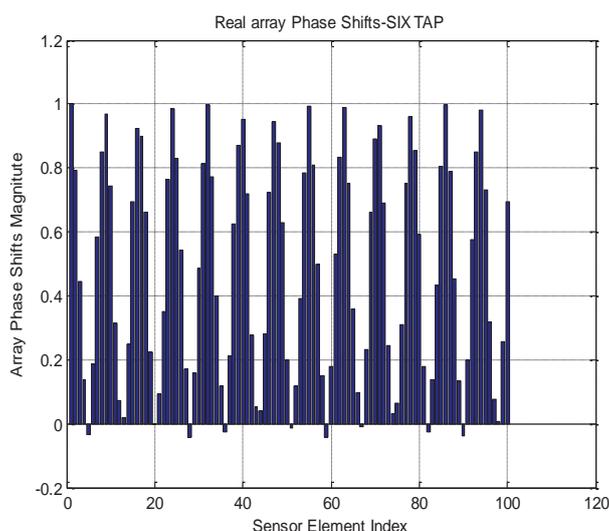
Fig. 16: Real Phase Shifts for probability-4.
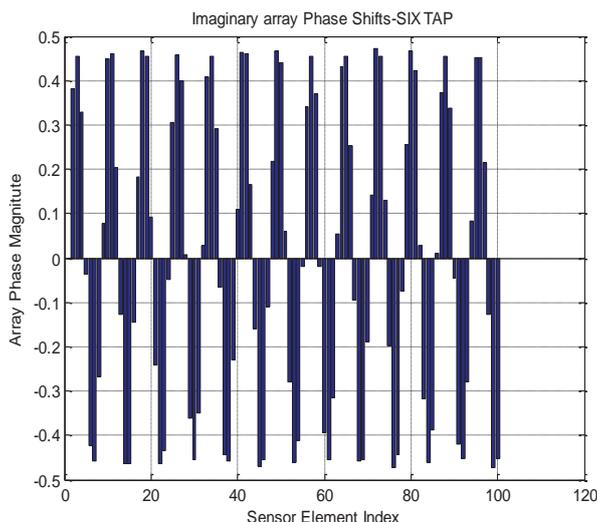
Fig. 17: Imaginary Phase Shifts for probability-4.

Figures 16 and 17 reveal real and imaginary phase shifts for the experimental setup considered in probability-4. Each phase shift is exploited to all sensor elements of the array for the formation of the main beam in the desired direction of 15 degrees with multiple interferers.
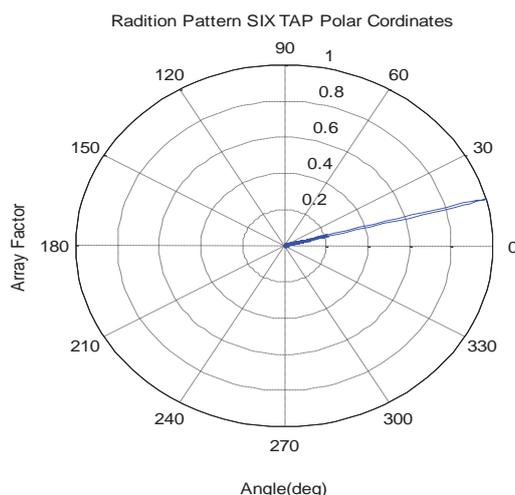
Fig. 18: Radiation pattern for 6T Beamformer - probability-4.

Figure 18 exhibits the radiation pattern for the novel 6T Beamformer Algorithm which excellently forms a narrow main beam towards the desired user direction of 15 degrees. The algorithm works very well with more antenna elements.

Figure 19 shows the mean square error comparison graph for probability-4 where the 6T Beamformer converges at just 9 iterations, implying lower MSE, followed by VSSG with a run count of 40, Griffiths and LMS converge at a run count of 70 and 62 respectively. Hence, the novel 6T Beamformer performs better than other methods for the considered probability-4. Table 5 summarizes the point of convergence obtained from the simulation results. It can be seen that the proposed 6T Beamformer converges faster than other algorithms with the least number of iterations implying low mean square error and high-speed performance of the algorithm.
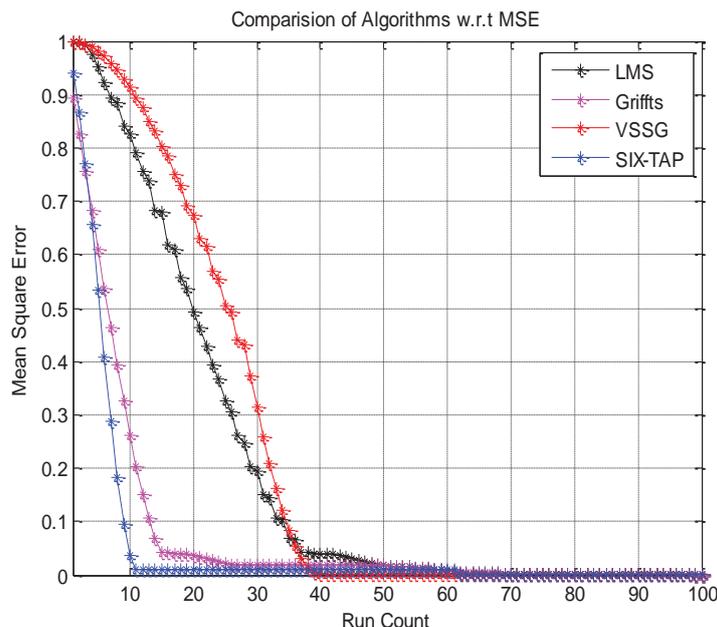
Fig. 19: Comparative analysis of MSE for probability-4.

Table 5: Summarizing the convergence point of all the algorithms.

Array Count	Jammer Count	LMS	Griffiths	VSSG	6T Beamformer
8	1	70	70	45	11
100	1	75	75	34	10
8	3	70	80	48	10
100	3	62	70	40	9

6. CONCLUSION

This research provides a novel beamforming method called 6T Beamformer to improve the converging speed by reducing the mean square error. The novel algorithm makes use of both fixed and variable step size methods along with Speediness Betterment and Convergence Betterment feedback. The feedback of weights is shared between the modules to calculate the phase shifts that are exploited to specific parts to create the main beam for actual mobile users. LMS, Griffiths, and VSSG are described in brief and these methods are then compared to the proposed Six Tap method which outperforms the existing ones for different scenarios of variation in array elements and jammer count. The convergence of SIX-TAP is high compared to other methods. The proposed technique is contrasted with other, high-performing methods in the analysis of the results to demonstrate that it converges more quickly at the 9th iteration and performs better than others in all probabilities. This also implies reduced multipath fading. The algorithm is also efficient in interference rejection and overall increasing the capacity of the RF spectrum. In the future, this research will be further extended by analyzing novel methods with various performance metrics.

REFERENCES

[1] Kabaoğlu N. (2018) Suboptimal Frequency-Selective Transceiver Design for Multicarrier Millimeter Wave MIMO Systems. Electrica, 18(1): 83-89. doi:10.5152/iujeee.2018.1813
[2] Tokan NT. (2019) Array Antenna Feeding Network Design for 5G MIMO Applications. Electrica, 19(2): 120-127. doi: 10.26650/electrica.2019.19004

[3]   Hameed AA, Ajlouni N, Orman Z, Özyavaş A. (2020) Investigating the Effectiveness of Adaptive Step Size LMS Algorithms for the Use with VOIP Applications. Electrica, 20(2): 116-123. doi: 10.5152/electrica.2020.19080

[4]   Li X, Wang J, Li Z, Li Y, Geng Y, Chen M, Zhang Z. (2022) Leaky-Wave Antenna Array with Bilateral Beamforming Radiation Pattern and Capability of Flexible Beam Switching. IEEE Trans. Anntenas Propag., 70(2): 1535-1540. doi: 10.1109/TAP.2021.3111157

[5]   Ghaffar A, Li XJ, Awan WA, Hussain N. (2021) Reconfigurable Antenna: Analysis and Applications. In Matin, MA (eds) Wideband, Multiband, and Smart Antenna Systems. Signals and Communication Technology. Springer, Cham:269-323. https://doi.org/10.1007/978-3-030-74311-6_9.

[6]   Kodgirwar VP, Deosarkar SB, Joshi KR. (2020) Design of Dual-Band Beam Switching Array for Adaptive Antenna Applications Using Hybrid Directional Coupler and E-Shape Slot Radiator. Wireless Personal Communications, 113 :423-437. https://doi.org/10.1007/s11277-020-07207-3

[7]   Nawaz H, Tekin I. (2020) Ten switched-beams with 2×2 series-fed 2.4 GHz array antenna and a simple beam-switching network. Int. J. RF Microwave Comput. Aided Eng., 30(6): e22184. https://doi.org/10.1002/mmce.22184

[8]   Baba AA, Hashmi RM, Attygalle M, Esselle KP, Borg D. (2021) Ultrawideband beam steering at mm-wave frequency with planar dielectric phase transformers. IEEE Trans. Antennas Propag., 70(3): 1719-1728.

[9]   Simone M, Fanti A, Lodi MB, Pisanu T, Mazzarella G. (2021) An in-line coaxial-to-waveguide transition for Q-band single-feed-per-beam antenna systems. Appl. Sci., 11(6): 2524. https://doi.org/10.3390/app11062524

[10] Thazeen S, Mallikarjunaswamy S, Saqhib MN. (2022) Septennial Adaptive Beamforming Algorithm. In International Conference on Smart Information Systems and Technologies (SIST): 28-30 April 2022; Nur-Sultan, Kazakhstan. pp 1-4.

[11] Cai Z, Zheng X. (2018) A Private and Efficient Mechanism for Data Uploading in SmartCyber-Physical Systems. IEEE Trans. Network Sci. Eng., 7(2): 766-775. doi: 10.1109/TNSE.2018.2830307

[12] Dakulagi V, Bakhar M. (2020) Advances in smart antenna systems for wireless communication. Wireless Personal Communications, 110(2): 931-957.

[13] Tong L, Ramanathan P. (2004) Energy efficient multicasting using smart antennas for wireless ad hoc networks in multipath environments. In IEEE Global Telecommunications Conference, GLOBECOM '04: 29 November 2004 - 03 December 2004; Dallas, TX, USA. 6: 4109-4113. doi: 10.1109/GLOCOM.2004.1379138

[14] Ali WAE, Hassan AHG. (2014) A hybrid least mean square/sample matrix inversion algorithm using microstrip antenna array. In Science and Information Conference: 27-29 August 2014; London, UK. pp 871-876. doi: 10.1109/SAI.2014.6918288.

[15] Anjaneyulu P, Rao PVDS, Sunehra D. (2021) Effect of Various Parameters on Minimum Mean Square Error and Adaptive Antenna Beamforming using LMS Algorithm. In 6th International Conference for Convergence in Technology (I2CT): 02-04 April 2021; Maharashtra, India. pp 1-5. doi: 10.1109/I2CT51068.2021.9418179.

[16] Mayyas K, Aboulnasr T. (1997) Leaky LMS algorithm: MSE analysis for Gaussian data. IEEE Trans. Signal Process., 45(4):927-934. doi: 10.1109/78.564181.

[17] Saqhib MN, Lakshmikanth S. (2020) Performance Anatomization of Routing Protocols in Wireless Sensor Network. International Journal of Advanced Trends in Computer Science and Engineering, 9(4): 5691-5699.

[18] Mallaiah N, Rao NVK, Ramakrishna D. (2019) RLS Adaptive Beamforming Algorithm in terms of Energy Efficiency For Smart Antenna System. In IEEE Indian Conference on Antennas and Propagation (InCAP): 19-22 December 2019; Ahmedabad, India. pp 1-4. doi: 10.1109/InCAP47789.2019.9134501.

[19] Wang X, Benesty J, Chen J, Cohen I. (2020) Beamforming with Small-Spacing Microphone Arrays Using Constrained/Generalized LASSO. IEEE Signal Process Lett., 27: 356-360. doi: 10.1109/LSP.2020.2971790.

[20] Roopa S, Narasimhan SV. (2021) Improved stability performance of a feedback active noise control using a Steiglitz-McBride adaptive notch filter and robust secondary path identification based on variable step size Griffiths LMS algorithm. Noise Control Engineering Journal, 69(2): 136-145. https://doi.org/10.3397/1/376914

[21] Shi L, Zhao H, Zeng X, Yu Y. (2019) Variable step-size widely linear complex-valued NLMS algorithm and its performance analysis. Signal Process., 165: 1-6. https://doi.org/10.1016/j.sigpro.2019.06.029

[22] Tuladhar SR, Buck JR. (2020) Unit Circle Rectification of the Minimum Variance Distortionless Response Beamformer. IEEE J. Oceanic Eng., 45(2): 500-510. doi: 10.1109/JOE.2018.2876584.

ENHANCEMENT OF MPPT MODULE FOR PARTIAL SHADING PHOTOVOLTAIC SYSTEM UNDER UNIFORM IRRADIANCE CONDITIONS

HAMEED ALI MOHAMMED[1], ROSMIWATI MOHD-MOKHTAR[1*]
AND HAZEM IBRAHIM ALI[2]

[1]*School of Electrical and Electronic Engineering, University Sains Malaysia,
Nibong Tebal, Pulau Pinang, Malaysia*
[2]*Control and Systems Engineering Department, University of Technology-Iraq,
Baghdad, Iraq*

[*]*Corresponding author: eerosmiwati@usm.my*

ABSTRACT: Maximum Power Point Tracking (MPPT) algorithms play a critical role in maximizing the output power of solar panels. Different MPPT techniques are evaluated based on several criteria, such as tracking speed, simplicity, and accuracy with changes in solar irradiance and ambient temperature. Under partial shading conditions (PSCs), conventional techniques fail to track global maximum power points (GMPP). This paper aims to present an automatic and accurate method to fix the complexity of determining the accurate lookup table data in an automatic and fast process under uniform irradiance conditions (UICs) and PSCs. The proposed method runs the photovoltaic (PV) module with all potential irradiance and temperature. It automatically calculates the perfect voltage reference (*Vref*) for all potential PV system cases. The *Vref* is collected in an array, sent into a two-dimensional lookup table, and used for controlling the boost converter. Simulation results verify the effectiveness of the proposed method. In addition, a comparison was also made with the conventional perturb and observe (P&O) method. Under UICs, the proposed method takes less time than the conventional P&O algorithm to reach the MPP. The time difference between them is $\Delta t = 0.133$ sec and $\Delta t = 0.04$ sec for the first scanning process at $t = 0$ sec and sudden change irradiance at $t = 1.5$ sec, respectively. As for PSCs, the proposed method reached the GMPP during pattern 104 (first peak) without any power loss, while the P&O MPPT was able to track the GMPP but with power losses of 2729.97 watts.

ABSTRAK: Algoritma Penjejakan Titik Kuasa Maksimum (MPPT) memainkan peranan penting dalam memaksimumkan kuasa keluaran panel solar. Teknik MPPT yang berbeza dinilai berdasarkan beberapa kriteria seperti kelajuan pengesanan, kesederhanaan, dan ketepatan dengan perubahan dalam sinaran suria dan suhu ambien. Di bawah keadaan teduhan separa (PSC), teknik konvensional gagal menjejak titik kuasa maksimum global (GMPP). Kajian ini bertujuan bagi membentangkan kaedah automatik dan tepat bagi membetulkan kesusahan dalam menentukan carian data berjadual secara tepat, automatik dan pantas di bawah keadaan sinaran seragam (UIC) dan PSC. Kaedah yang dicadangkan menjalankan modul fotovoltaik (PV) dengan semua potensi sinaran dan suhu dan mengira rujukan voltan sempurna (Vref) secara automatik bagi semua kes yang berpotensi dalam sebarang jenis sistem PV. Vref dikumpul dalam tata susunan, dihantar ke dalam jadual carian dua dimensi, dan digunakan bagi mengawal penukar rangsangan. Keputusan simulasi mengesahkan keberkesanan kaedah yang dicadangkan. Perbandingan juga dibuat dengan kaedah konvensional perhati dan ganggu (P&O). Di bawah UIC, kaedah yang dicadangkan mengambil masa yang lebih singkat berbanding algoritma konvensional P&O bagi mencapai MPP. Perbezaan masa antara keduanya adalah masing-masing, $\Delta t =$

0.133 saat dan Δt = 0.04 saat bagi proses pengimbasan pertama iaitu pada t = 0 saat dan sinaran perubahan mendadak pada t = 1.5 saat. Bagi PSC, kaedah yang dicadangkan mencapai GMPP semasa corak 104 (puncak pertama) tanpa kehilangan kuasa manakala MPPT P&O dapat mengesan GMPP tetapi dengan pengurangan kuasa sebanyak 2729.97 watt.

---

1. INTRODUCTION

The MPPT is a critical technique for getting maximum power from a PV panel system. The electrical energy is obtained directly from a group of connected PV cells. Solar irradiance and ambient temperature conditions affect the amount of extracted power from the PV module. The MPPT technique aims to dynamically obtain the maximum power by forcing the PV panel to work at an efficient voltage operating point (VMPP) [1,2]. The MPPT technique can be classified into two methods, under UICs and PSCs. In the UICs case, classical MPPT techniques include constant perturb and observe (P&O), incremental conductance (InC), open-circuit voltage (OCV), short circuit current (SCC), lookup table, and hill-climbing (HC) methods [3]. The nonlinearity of the P-V curve for the PV system means it only has one optimum operating point called the maximum power point (MPP). The location of this point on the P-V curve significantly depends on atmospheric conditions, such as irradiance and ambient temperature. During PSCs, the P-V curve will have several local maximum power points (LMPPs) and only one global maximum power point (GMPP) [4,5]. The aforementioned techniques are sometimes combined with other methods to find the GMPP.

Conventional MPPT techniques suffer from three major drawbacks; 1) rapid oscillations around the MPP even at the steady-state response, 2) the inability to effectively track during sudden changes in irradiance, and 3) the inability to identify the GMPP during PSCs because they may be busy in one of the LMPPs. This situation creates a complicated process to find the system's GMPP [6,7]. Research has been done in the literature to overcome the drawbacks of conventional MPPT techniques. A two-phased tracking method was presented to improve the tracking behavior of the conventional incremental conductance MPPT [8]. The method improved the performance of the conventional InC MPPT, but it was still unable to track the GMPP during PSCs.

Research by Basoglu used the concept of the OCV method to first estimate the locations of GMPP by assuming that the maximum power point was located at 0.85 and 0.8 of open-circuit voltage, respectively [9]. Then, the P&O MPPT algorithm was used to track the maximum point. The method has a good tracking performance, but it is ineffective when the peaks are not always in ascending order toward the GMPP. An optimization method called artificial bee colony (ABC) was presented in [10]. The drawbacks of this algorithm are slow tracking, and even though the technique can track under PSCs with high tracking speed, the efficiency declines when the shading patterns are changed immediately.

Also, the general drawbacks of the method were the implementation complexity, and the high financial cost. The method presented in [11] estimated the MPP by calculating the I-V curve's maximum power point voltage (VMPP). Regrettably, the method resulted in lower output power than the P&O technique. Also, the method was unable to find the GMPP under PSCs. In [12], adaptive neuro-fuzzy inference system (ANFIS) MPPT techniques were trained for the non-linear characteristic of the PV system. These algorithms had a high

accuracy under rapid changes in irradiance conditions. However, they had a high complexity in selecting the output target data and could not be implemented quickly. Authors in [13] and [14] used the Lookup Table technique and compared it with P&O MPPT.

This technique's problem was the complexity of calculating the lookup table data, especially under PSCs. Moreover, the lookup table data must be increased to get highly accurate tracking and cover all operating points, thus increasing the calculation difficulties. To solve the complexity problem, the authors in [15] and [16] combined P&O MPPT and Lookup Table MPPT. The method recorded the irradiance and duty cycle values from the P&O MPPT and sent the data into a table to use the next time. The method could not enhance the performance of P&O MPPT because it used the same duty cycle used in the P&O method, which limited this method with the P&O algorithm limitation.

For designing lookup table systems, the researchers faced a big problem with determining the table data accurately. The design process needed a large amount of accurate data to ensure accurate results. Moreover, MPPT based on the lookup table method was inefficient under PSCs. This paper proposes an automatic and accurate method to solve the problem of the calculation complexity associated with the lookup table method and makes it capable of tracking the GMPP and working efficiently under PSCs. This method can calculate and track the MPP for any PV system under UICs and PSCs and builds a complete data array that contains PMPP and corresponding voltage VMPP for all potential cases in just a few minutes.

2. SOLAR ARRAY MATHEMATICAL MODEL

The single-diode PV array equivalent circuit is shown in Fig. 1. The current source (I_{ph}) represents the cell photocurrent connected in parallel with a diode. N_s is the number of series cells, N_p is the number of parallel cells, R_{sh} and R_s are the shunt and series resistances of the cell, respectively [15,16]. The current source (I_{ph}) is described by [17]

$$I_{ph} = I_{sc} + K_i(T - 298) \cdot G/1000 \tag{1}$$
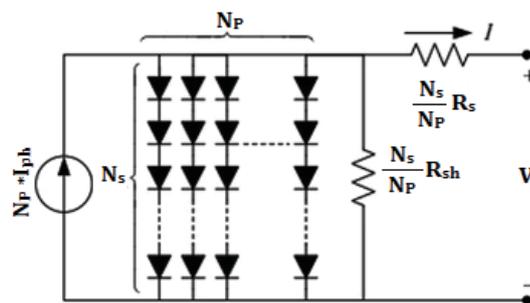

Fig. 1: Equivalent circuit of solar array [17].

where I_{SC} is the short-circuit current (A), K_i is the short-circuit current of the cell at 25 °C and 1000 W/m², T is the operating temperature (K), and G is the solar irradiation (W/m²). The reverse saturation current (I_{rs}) is

$$I_{rs} = I_{sc}/[exp(q \cdot V_{OC}/N_S \cdot k \cdot n \cdot T) - 1] \tag{2}$$

where q is the electron charge (1.6×10^{-19}), V_{OC} is the open-circuit voltage (V), n is the ideality factor of the diode, and k is the Boltzmann's constant $(1.3805 \times 10^{-23}$ J/K). The saturation current (I_0) is given by

$$I_0 = I_{rs}(T/T_r)^3 exp\left[(q \cdot E_{g0}/n \cdot k)(1/T_r - 1/T)\right] \tag{3}$$

where T_r is the nominal temperature (298 K), E_{g0} is the bandgap energy of the semiconductor (1.1 eV). The PV module output current (I) is [15]

$$I = I_{ph}N_P - I_0 N_S \left[exp\left(((V/N)s + (I \cdot R_S/N_P)/n \cdot v_t)\right) - 1\right] - I_{sh} \tag{4}$$

with

$$v_t = k \cdot T/q \tag{5}$$

$$I_{sh} = (V(N_P/N_S) + I \cdot R_s)/R_{sh} \tag{6}$$

where v_t is the diode thermal voltage (V), and I_{sh} is the current through the shunt resistance. This paper uses ERA Solar ESPMC-215 PV Model as a reference module. The PV model parameters are given in Table 1.

Table 1: Parameters of ESPMC-215 PV module

Name	Symbol	Value
Maximum power	P_{MPP}	213.15 W
Voltage at the maximum power point	V_{MPP}	29.00 V
Current at the maximum power point	I_{MPP}	7.350 A
Open circuit voltage	V_{oc}	36.30 V
Short circuit current	I_{sc}	7.840 A
Number of cells connected in series	N_s	10
Number of cells connected in parallel	N_P	40

3. PROPOSED METHOD

The P&O MPPT algorithm is the most well-known technique because of its ease of design and implementation. From the flowchart shown in Fig. 2, notice that the P&O MPPT determinations compare the change in PV power, ΔP, and change in PV voltage, ΔV, for two points, K and K-1, on a P-V curve to identify the maximum power point. The main disadvantages of this technique are inability to deal with rapid weather changes, high oscillations around the MPP, the slow response, and failure to track GMPP under PSC [18], [19].

To avoid the disadvantages of the P&O MPPT method and solve the complexity problem associated with the lookup table method and make it work efficiently under PSCs, an automatic and accurate method is proposed in this paper. In the first step, the method starts to ask the user to enter the specifications of the solar array, the maximum and minimum limits of solar radiation (G_{MAX} and G_{MIN}), and the maximum and minimum limits of the ambient temperature (T_{MAX} and T_{MIN}) respectively, as well as the number of connected cells in series (N_S) and parallel (N_P). The flowchart in Fig. 3 shows the whole tracking process of the proposed method.

Then, it runs the PV module and measures the PV voltage (V_{PV}) and current (I_{PV}) and calculates the PV power (P_{PV}). From the V_{PV} and P_{PV} that have been calculated, it generates the PV curves for all operating point cases (for all potential irradiance and temperatures). It calculates the MPP for each operating point case. After calculating MPPs, it automatically calculates the corresponding voltage for every MPP without manual calculation and considers it a reference voltage ($Vref$). After finishing all offline testing cases, it collects all the calculated reference voltages in an array and sends them with input vectors (irradiance and temperature) into a lookup table. The lookup table makes the $Vref$ array more flexible because it uses this array to map input values to output values, allowing the system to estimate the missing data. The system has been divided into three groups of PV modules connected in series to simulate the PSCs. Group 1 is set as no shading PV array, while Groups 2 and 3 are shading PV arrays with 20 shading patterns for each. The details are explained in the next section.
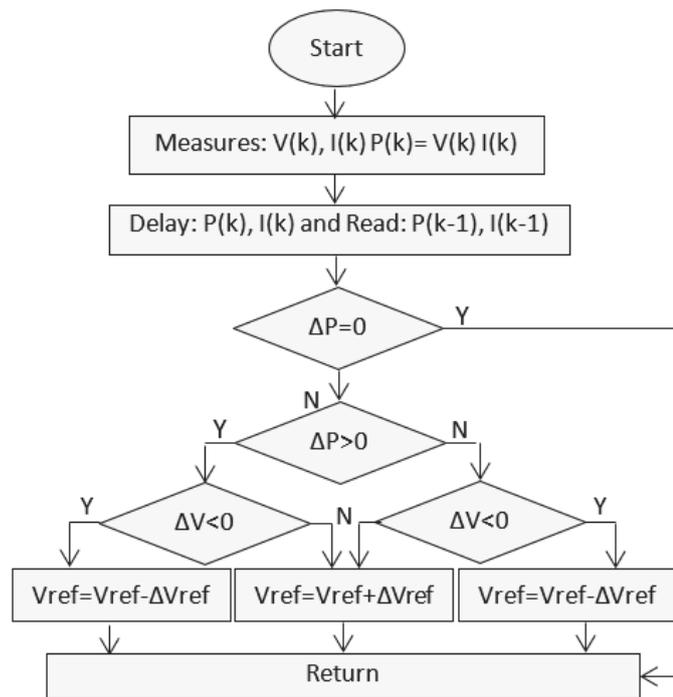
Fig. 2: Flowchart of P&O algorithm.

The proposed method is evaluated for 400 patterns (20x20) and calculates the GMPP for each pattern. It calculates the MPP irrespective of the number of peaks on the PV curve, which allows this method to work effectively under PSCs. This capability is considered one of the most vital points of the proposed method.

In this paper, a boost converter is designed and used as a power driver between the PV array and the load to adjust the solar module voltage to extract the maximum power from the PV array at the converter input side. The system uses the $Vref$ generated from the proposed MPPT method and compares it with the actual V_{PV} to generate an error signal. The Proportional Integral (PI) controller has been designed to manage the error signal and generate the duty ratio to control the boost converter. On the other hand, the parameters of the boost converter, as given in Table 2, have been calculated carefully to achieve minimum fluctuation around the MPP. Figure 4 shows the overall system, PV array, MPP method, boost converter, and PI controller.

Table 2: Parameters of a boost converter

Name	Symbol	Value
Input capacitor	C_{in}	3227 µF
Inductor	L	1.45 mH
Output capacitor	C_{out}	1000 µF
Output resistance	R_{OUT}	20 Ω
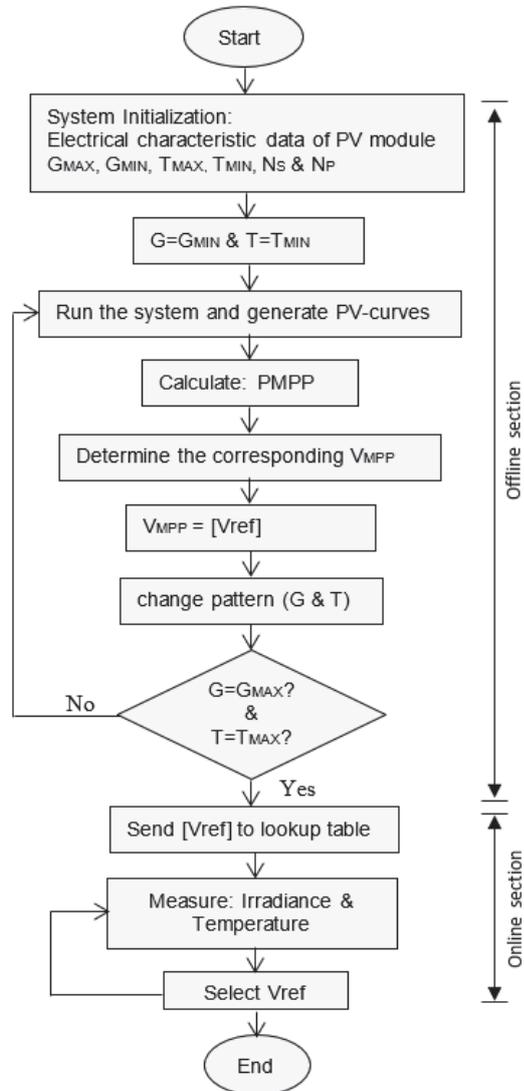
Fig. 3: The flowchart of the proposed algorithm.
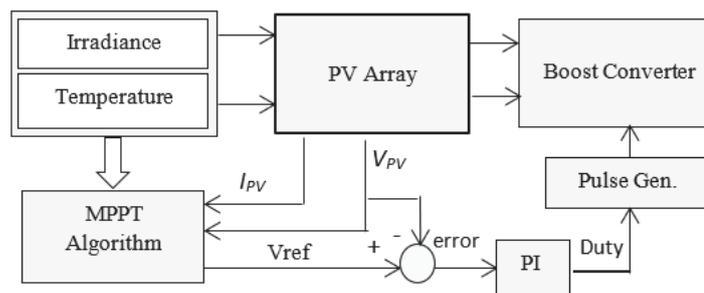
Fig. 4: The overall system with PI controller.

3.1 PV System Modeling under UICs and PSCs

Under UICs, the proposed method tests 400 patterns; 20 for changing solar irradiance (from 50 to 1000 W/m² by step change of 50 W/m²) x 20 for changing cell temperature (from 5 to 55 °C by step change of 5 °C). It runs the PV module for each pattern, measures the (V_{PV}, I_{PV}, and P_{PV}) and generates the PV curve for all these patterns, as shown in Fig. 5.

This method obtained the MPP for each pattern from the PV curves generated during the previous step, then calculates the corresponding reference voltage and sends it into an array. Figure 6 shows a 3-dimensional view of the array data, which contains a different irradiance, G and temperature, T as an input vector, and reference voltage ($Vref$) as the output array.
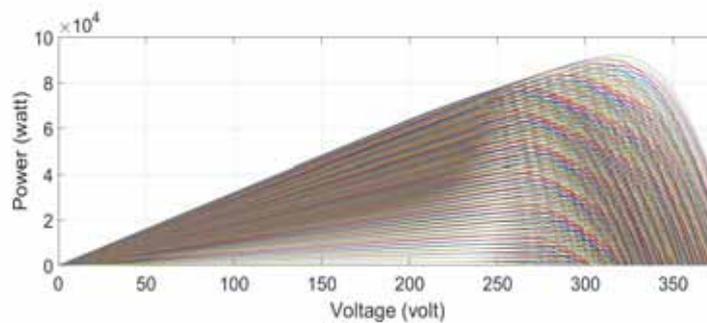
Fig. 5: The P-V curves for changing solar irradiance (50 to 1000 W/m²) and temperature (5 to 55 °C).
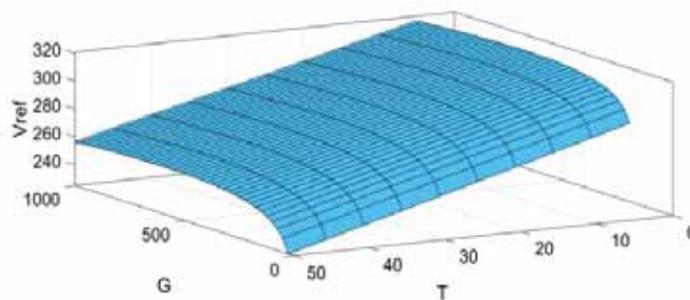
Fig. 6: The 3-dimensional view for irradiance (G), temperature (T), and reference voltage ($Vref$) under UICs.
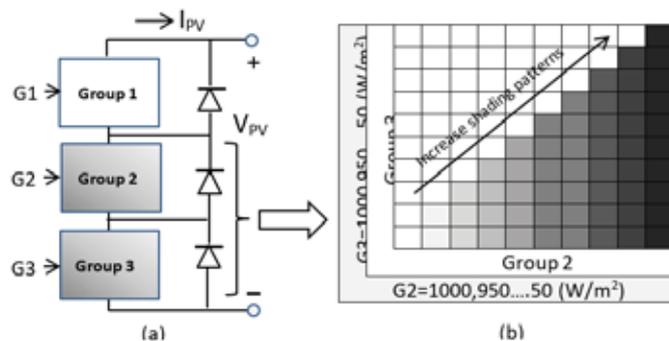
Fig. 7: Partially shaded PV array (a) three PV partial shading groups connected in series, (b) different shading patterns for two PV arrays.

Under PSCs, a PV module is shaded if three or more cells receive lower than the normal irradiance level [8,20]. Figure 7(a) shows the three groups of the PV module connected in series with 20 levels of different shading conditions for Groups 2 and 3. Under the partial

shading test, Group 1 is set as no shading PV array with 160 modules ((4–series) x (40-parallel)) and Groups 2 and 3 are shading PV arrays with 120 modules ((3–series) x (40-parallel)) for each. Figure 7(b) shows the different shading patterns (irradiance changed from 1000 to 50 W/m^2 for Group 2 and Group 3). The proposed algorithm was evaluated on 400 patterns (20x20), and the global maximum power point was calculated for each.

The P–V curves test obtained using the proposed method for a PV array with 400 modules (10×40) at a different shading pattern (one normal group and two shaded groups) is shown in Fig. 8. These curves are used to calculate the GMPP. From Fig. 8, notice that each PV-curve has one GMPP and one or more LMPP. It can also be seen that the only case in which the P-V curves contain one MPP is when all groups (Groups 1, 2, and 3) receive the same amount of irradiance.

From the obtained MPP, the proposed method calculates the voltage corresponding to the MPP for each pattern and considers it as a reference voltage (*Vref*). To evaluate the system, we assumed that the system works under three different patterns; 1) pattern 104 (G1=1000, G2=100 and G3=300 W/m^2), 2) pattern 302 (G1=1000, G2=100 and G3=800 W/m^2) and 3) pattern 320 (G1=1000, G2=1000, and G3=800 W/m^2) where G1, G2, and G3 are the amount of solar radiation falling on the groups (1, 2, and 3), respectively. These three patterns were chosen because each one gives a GMPP in a different location from the others. Fig. 9 shows a 3-dimensional view of different shading conditions for Groups 2 and 3, irradiance, and reference voltage for the PSCs test.
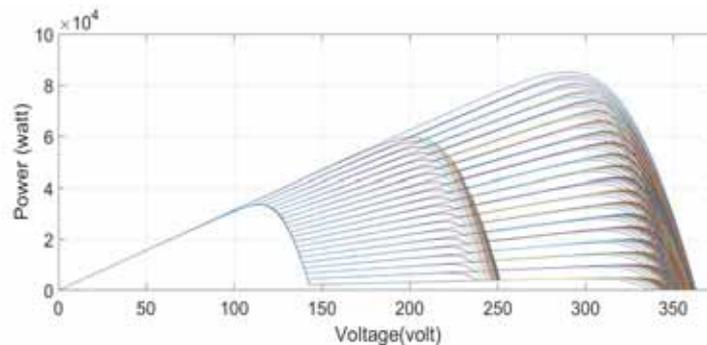
Fig. 8: PV curves testing of patterns 1 to 400 under PSCs.
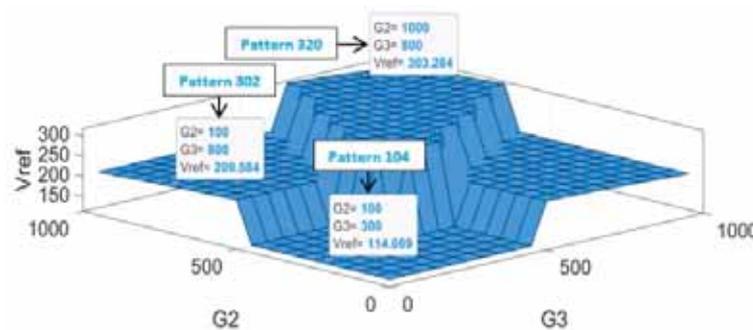
Fig. 9: The 3-dimensional view for irradiance and reference voltage under PSCs.

4. RESULTS AND DISCUSSION

Simulation analysis was conducted using MATLAB / SIMULINK and M-file code to verify the efficacy of the proposed method. This section analyses the proposed method's performance under uniform irradiance and partial shading conditions.

4.1 Uniform Irradiance Sudden Change

Figure 10 compares the dynamic response of the PV array output power (P_{PV}) based on the proposed and P&O methods under five-scenario uniform irradiance sudden changes (step increase on the left side and step decrease on the right side). Then, a comparison is made with the calculated MPPs as a reference signal calculated directly from the P-V curves. Fig. 10 shows that each MPPT technique (proposed and P&O) tracked the MPP for each step change. However, the proposed method reached the MPP with a faster speed response, greater accuracy, and lower steady-state fluctuations around the MPP compared with P&O MPPT at the same sudden changes in irradiance. The simulation results shown in Fig. 10 prove that the proposed method's effectiveness has been superior to the conventional P&O MPPT algorithm in terms of accuracy, tracking speed, and the fluctuations around the MPP. The reason is that it directly determines the optimum value of *Vref*, taking it from the lookup table, which makes the proposed method very fast and accurate in reaching the MPP without a need to scan the voltage range on the P-V curve. Also, selecting one value of the *Vref* caused the elimination of all fluctuations around the MPP.
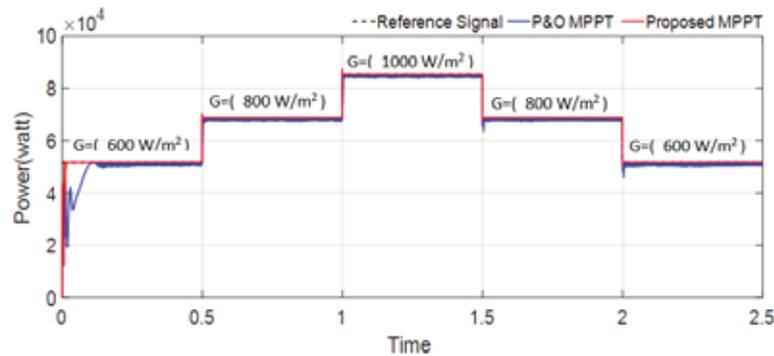
Fig. 10: Comparison of the output power of P&O and proposed MPPT techniques with the reference signal under sudden changes of irradiation.

Figure 11 shows the zoomed view of the output power during the first scanning process (*t* = 0 to 0.15 sec) and the sudden change in irradiance at *t* = 1.5 sec. From Fig. 11(a) and (b), the proposed method takes less time to reach the MPP than the conventional P&O algorithm. The time difference between them is Δt = 0.133 sec and Δt = 0.04 sec for the first scanning process at *t* = 0 sec and sudden change irradiance at *t* = 1.5 sec, respectively. This problem can be solved for the P&O algorithm by increasing the *Vref* step size ($\Delta Vref$). Unfortunately, that would cause another problem, a more significant fluctuation around the MPP. As shown in Fig. 11, the proposed method reached the MPP with no power loss, while the P&O MPPT causes power losses ($\Delta P \approx$ 51800 - 51000 = 800 watts) and ($\Delta P \approx$ 71817 - 68070 = 3737 watts) for (a) and (b), respectively.

As mentioned before, the proposed method does not need a scanning process, making it very fast to reach the MPP. In the first scanning process shown in Fig. 11(a), the conventional P&O MPPT algorithm took a very long time to reach the MPP because it depends on the initial *Vref* (in this case, *Vref* initial equal to 250 volts), while the V_{MPP} equal to 293 volts at 600 W/m². Another problem is that the P&O MPPT has very high fluctuations around the MPP compared to the proposed method, which has no fluctuations around the MPP. The fluctuations problem can be solved by decreasing the $\Delta Vref$. However, decreasing the $\Delta Vref$ would slow down the output response. Also, it is evident from Fig. 11(b) that the maximum overshot due to sudden change in irradiance is very small in the output power based on the proposed method. At the same time, it is higher for the conventional P&O

MPPT algorithm. The fluctuation and slow-down situations associated with the conventional algorithms cause additional power losses in the PV systems. These problems cannot be solved because both situations are related, so one needs to find an acceptable case between them.
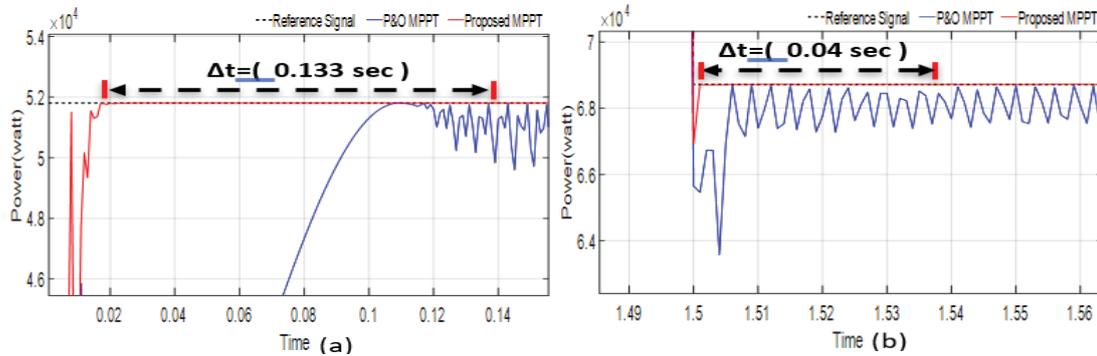
Fig. 11: Zoomed view of Fig. 10(a) due to the first scanning process (at $t = 0$ to 0.15 sec), (b) due to the sudden irradiance changes at ($t = 1.5$ sec).

4.2 Load Changes

This analysis is meant to verify the proposed method under hard conditions. The output resistance (R_{OUT}) of the boost converter is changed from 5 Ω to 20 Ω at ($t = 0.3$ sec) and then returned to 5 Ω at ($t = 0.6$ sec) as shown in Fig. 12. It can be seen from Fig. 12 that the proposed method is superior by maintaining the dynamic response without fluctuations after the R_{OUT} has been changed. Further, when the load resistance disturbance occurred at $t = 0.3$ sec, the power losses reached approximately ($\Delta P \approx 85180 – 82600 = 2580$ watts) using the conventional P&O MPTT algorithm compared to the proposed method. At the same time, the P&O MPPT generates high fluctuations around the MPP at the steady-state response. Also, when the R_{OUT} return to 5 Ω at $t = 0.6$ sec, the P&O MPPT takes about 0.11 sec to reduce the fluctuations and return the dynamic response to the first situation with power losses ($\Delta P \approx 85180 – 84200 = 980$ watts). The reason P&O MPPT was affected by changing the R_{OUT} is that the input resistance has been affected, which caused offsetting of the position of the operating point on the PV curve.
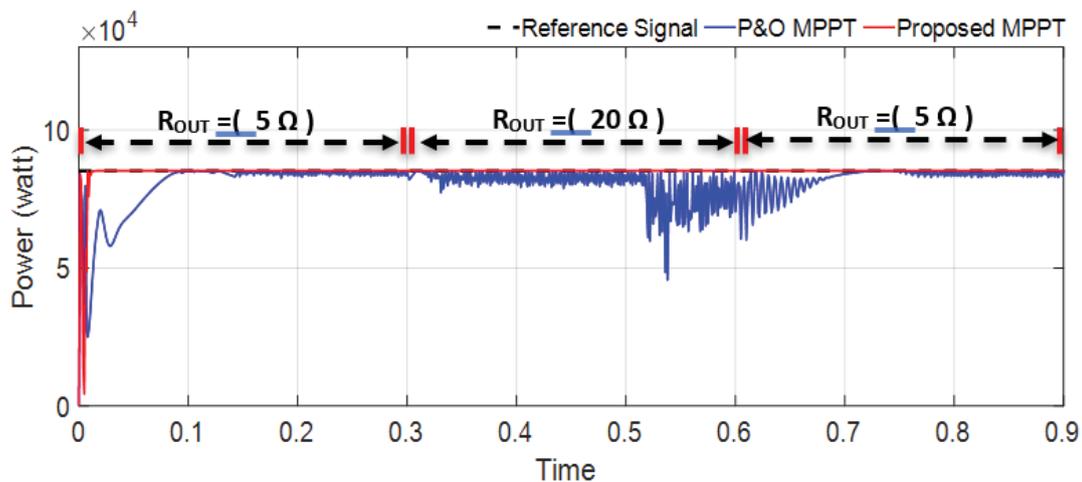
Fig. 12: Comparison of the output power of P&O and proposed MPPT techniques under load changes.

4.3 Partial Shading Conditions

This analysis aims to verify the proposed method's ability to track GMPP under PSC. Three shading patterns have been generated where three PV module groups have been connected in series to simulate the partial shading conditions. The three patterns have three different shapes based on the location of the GMPP on the P-V curve, where each one has one GMPP. The GMPP for pattern 104 (point A) is located at the first peak, making it easy to track if the voltage range scan starts from zero, while for pattern 302 and pattern 320, the GMPPs are located at the second (point B) and third (point C) peak, respectively, which needs a unique tracking technique. Figure 13 shows the three different partial shading P–V curves for 1) pattern 104 (G1=1000, G2=100 and G3=300 W/m^2), 2) pattern 302 (G1=1000, G2=100 and G3=800 W/m^2) and 3) pattern 320 (G1=1000, G2=1000 and G3=800 W/m^2).

It is observed that each P-V curve has 3 peaks, and the MPPT algorithm needs to scan all these peaks to decide which one represents the GMPP. Figure 13 shows the P-V curves of the sudden change scenario, assuming that a series of partial shading had just happened, which shifts the operating points from pattern 104 to pattern 302 to pattern 320 under PSCs at t = 1 and 1.5 sec, respectively. When partial shading scenarios occur, the P-V curves give multi LMPs and single GMPP for each pattern, where the goal here is to track the GMPP, which is point (A, B and C) in the shortest possible time.
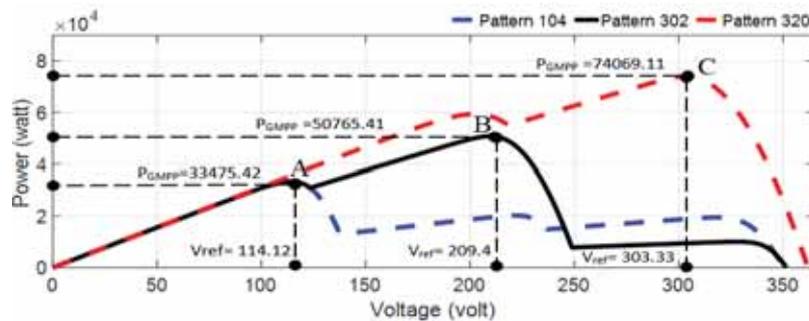
Fig. 13: Three different partial shading P–V curves for patterns 104, 302 and 320.

Figure 14 shows the PV output power response of the proposed method compared with the conventional P&O MPPT algorithm under PSCs, where the scenarios shown in Fig. 13 have been used. As seen in Fig. 14, both proposed, and conventional P&O MPPT has reached the GMPP at the first pattern (point A) at the first zone on the P-V curve. The P&O MPPT algorithm reached the GMPP because it is located in the first zone, making it easy to track based on the initial scanning voltage, which has been set in zone one. Even though the conventional MPPT algorithm tracked the GMPP, it took more time with very high fluctuations, while the proposed method reached the GMPP faster with fewer fluctuations around the GMPP ($\Delta t \approx$ 8 msec).

At t = 1 sec, the operational point jumped from pattern 104 to pattern 302, which means that the GMPP has been moved to zone two on the P-V curve, point B. At this operational point, the conventional P&O MPPT algorithm could not track the GMPP because it is located on the second peak on the PV curve, while the start scanning point located within peak one depends on the first GMPP (point A). The conventional P&O MPPT algorithm will remain stuck in zone one when the scanning point crosses point A. The P&O MPPT algorithm would return due to low power indication. At this period, the P&O MPPT algorithm failed to track the GMPP, while the proposed method tracked the GMPP

accurately as fast as possible. The same process occurred when the operational point jumped to pattern 320.
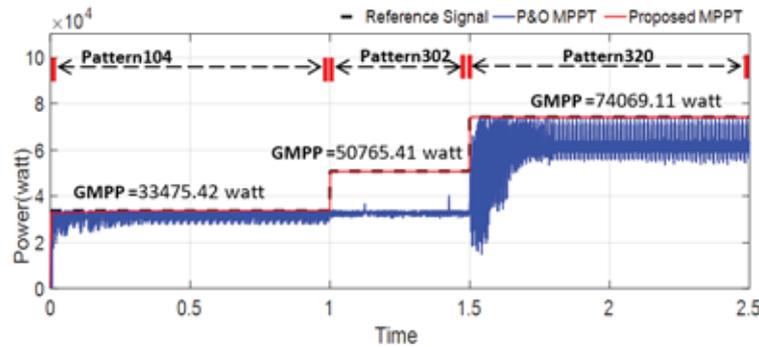
Fig. 14: Comparison of the output power of P&O and proposed MPPT techniques
with the reference signal under PSCs with three different patterns.

In contrast, the proposed method reached the GMPPs of all three patterns in the shortest time without any fluctuations around the GMPP. The conventional MPPT algorithms fail to track the GMPPs under PSCs because they cannot jump from one zone to another on the P-V curve. At the same time, the proposed method was not affected by changing the operational point, which gives enough flexibility to move on the P-V curve.

Figure 15 shows the zoomed view of Fig. 14 for the three different patterns during the time intervals (0.1 to 0.16 sec), (1.1 to 1.25 sec), and (1.75 to 1.85 sec) for patterns 104, 302, and 320, respectively. It shows how the proposed method and the P&O MPPT algorithm tracked the GMPP when partial shading occurs. As shown in Fig. 15(a), the proposed method reached the GMPP during pattern 104 without any power loss, compared to the reference power signal, while the P&O MPPT tracked the GMPP but with power losses (ΔP = 33475.42 - 30745.45 = 2729.97 watts). Also, the fluctuation around the GMPP is very high compared to the proposed method. The P&O MPPT algorithm can track the GMPP at this period because GMPP is located within the first peak. In the patterns 302 and 320 shown in Fig. 15(b) and Fig. 15(c), respectively, the P&O MPPT failed to track the GMPPs (50765.41 and 74069.11 watts) and instead tracked the first LMPP = 33493 watts for patterns 302 and second LMPP = 61499 watts for patterns 320, and missed tracking the GMPP because it did not locate in the second peak (pattern 302), or third peak (pattern 320). The overall simulation results substantiate that the proposed method's effectiveness has been superior to the conventional P&O MPPT algorithm in terms of accuracy, tracking speed, fluctuations around the MPP, and tracking the GMPP under PSCs.
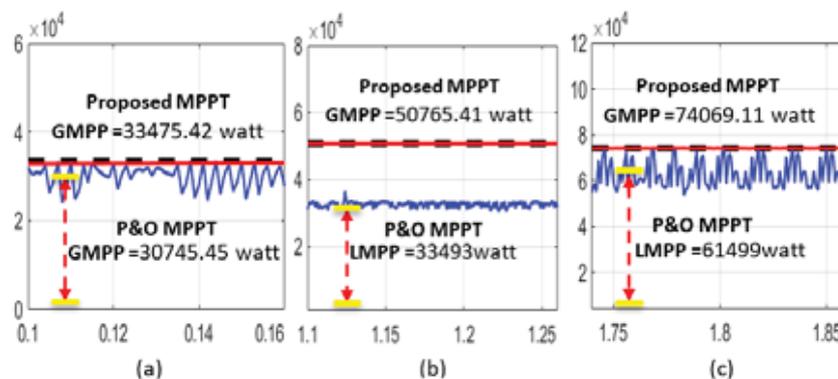
Fig. 15: Zoomed view of Fig. 14 (a) due to pattern 104, (b) due to pattern 302,
(c) due to pattern 320.

5. CONCLUSION

This paper proposed an automatic and accurate method for tracking the MPP for any PV system under UICs and PSCs. It has fixed the complexity of determining the accurate lookup table data in an automatic and fast process under UICs and PSCs. The proposed method has been running the PV module with all potential irradiance and ambient temperatures to get the PV curve and then calculating the corresponding V_{MPP} automatically for each case. Then, the calculated array voltage was set as the PV array reference operating voltage [*Vref*]. Five scenarios of sudden irradiance changes under UICs and three-pattern scenarios under PSCs were done to verify the proposed method. Compared with the conventional P&O MPPT technique, the proposed method harvests higher power from the PV system because it has a fast-tracking response, robust stability, and fewer fluctuations around the MPP. From the study, the lookup table data generated automatically using the proposed method worked ideally compared to conventional data collection methods. This method gives the lookup table control system the ability to track the GMPP under PSCs, which solves the weakness point associated with this kind of control system. A future recommendation to validate this method under all potential challenges is to test it with more extended string PV modules. This method will be the subsequent research investigation to be considered.

ACKNOWLEDGMENT

REFERENCES

[1] Heelan MY, Al-Qrimli FAM. (2020) Design and simulation of neuro-fuzzy-based MPPT controller for PV power system. Proceedings of the Int. Conf. on Electrical, Communication, and Computer Engineering (ICECCE 2020), 12-13 June 2020, Istanbul, Turkey, pp. 1-6. doi: 10.1109/ICECCE49384.2020.9179287.

[2] Zhang X, Gamage D, Wang B, Ukil A. (2021) Hybrid maximum power point tracking method based on iterative learning control and perturb & observe method. IEEE Trans. on Sustainable Energy, 12(1): 659-670. doi: 10.1109/TSTE.2020.3015255.

[3] Ostadrahimi A, Mahmoud Y. (2021) Novel spline-MPPT technique for photovoltaic systems under uniform irradiance and partial shading conditions. IEEE Trans. on Sustainable Energy, 12(1): 524-532. doi: 10.1109/TSTE.2020.3009054.

[4] Xu S, Gao Y, Zhou G, Mao G. (2021) A global maximum power point tracking algorithm for photovoltaic systems under partially shaded conditions using modified maximum power trapezium method. IEEE Trans. on Industrial Electronics, 68(1): 370-380. doi: 10.1109/TIE.2020.2965498.

[5] Kumar N, Hussain I, Singh B, Panigrahi BK. (2018) Framework of maximum power extraction from solar PV panel using self-predictive perturb and observe algorithm. IEEE Trans. on Sustainable Energy, 9(2): 895-903. doi: 10.1109/TSTE.2017.2764266.

[6] Bhattacharyya S, Kumar DSP, Samanta S, Mishra S. (2021) Steady output and fast tracking MPPT (SOFT-MPPT) for P&O and InC algorithms. IEEE Trans. on Sustainable Energy, 12(1): 293-302. doi: 10.1109/TSTE.2020.2991768.

[7] Ahmed RM, Zakzouk NE, Abdelkader MI, Abdelsalam AK. (2021) Modified partial-shading-tolerant multi-input-single-output photovoltaic string converter. IEEE Access, 9: 30663-30676. doi: 10.1109/ACCESS.2021.3058695.

[8] Hsieh G, Hsieh H, Tsai C, Wang C. (2013) Photovoltaic power-increment-aided incremental - conductance MPPT with two-phased tracking. IEEE Trans. on Power Electronics, 28(6): 2895-2911. doi: 10.1109/TPEL.2012.2227279.

[9] Basoglu ME. (2019) Module level global maximum power point tracking strategy. Proceedings of the 2019 4th Int. Conf. on Power Electronics and their Applications (ICPEA), 25-27 Sept 2019, Elazig, Turkey, pp. 1-5. doi: 10.1109/ICPEA1.2019.8911173.

[10] Gonzalez-Castano, Restrepo C, Kouro S, Rodriguez J. (2021) MPPT algorithm based on artificial bee colony for PV system. IEEE Access, 9: 43121-43133.
doi: 10.1109/ACCESS.2021.3066281.

[11] Zadeh MJZ, Fathi SH. (2017) A new approach for photovoltaic arrays modeling and maximum power point estimation in real operating conditions. IEEE Trans. on Industrial Electronics, 64(12): 9334-9343. doi: 10.1109/TIE.2017.2711571.

[12] Reddy D, Ramasamy S. (2019) An artificial intelligent MPPT controller based three level SEPIC topology for 1.2kW solar PV system. Proceedings of the 2019 Innovations in Power and Advanced Computing Technology (i-PACT), 22-23 Mar 2019, Vellore, India, pp. 1-6.
doi: 10.1109/i-PACT44901.2019.8960135.

[13] Chandran A, Mathew BK. (2020) Comparative analysis of P&O and fuzzy logic controller based MPPT in a solar cell. Proceedings of the 2020 Third Int. Conf. on Smart Systems and Inventive Technology (ICSSIT), 20-22 Aug 2020, Tirunelveli, India, pp. 622-626.
doi: 10.1109/ICSSIT48917.2020.9214187.

[14] Udayalakshmi JK, Sheik MS. (2018) Comparative study of perturb & observe and lookup table maximum power point tracking techniques using MATLAB/Simulink. Proceedings of the 2018 Int. Conf. on Current Trends towards Converging Technologies (ICCTCT), 1-3 Mar 2018, Coimbatore, India, pp. 1-5. doi: 10.1109/ICCTCT.2018.8550835.

[15] Rolevski M, Zecevic Z. (2020) MPPT controller based on the neural network model of the photovoltaic panel. Proceedings of the 24th Int. Conf. on Information Technology (IT), 18-22 Fen 2020, Zabljak, Montenegro, pp. 1-4. doi: 10.1109/IT48810.2020.9070299.

[16] Bouazza F, Mohamed M, Bouziane B, Abdelhak D. (2019) Photovoltaic generator modeling based on lookup table approach and implementation on STM32F407 board of the perturb & observe algorithm based MPPT. Proceedings of the 2018 Int. Conf. on Applied Smart Systems (ICASS), 24-25 Nov 2018, Medea, Algeria, pp. 1-7.
doi: 10.1109/ICASS.2018.8652000.

[17] Ngo S, Chiu C. (2020) Simulation implementation of MPPT design under partial shading effect of PV panels. Proceedings of the 2020 Int. Conf. on System Science and Engineering (ICSSE), 31 Aug – 3 Sept 2020, Kagawa, Japan, pp. 1-6. doi: 10.1109/ICSSE50014.2020.9219306.

[18] Alhussain HMA, Yasin N. (2020) Modeling and simulation of solar PV module for comparison of two MPPT algorithms (P&O & INC) in MATLAB/Simulink. Indonesian Journal of Electrical Engineering and Computer Science, 18(2): 666-677.
doi: 10.11591/ijeecs.v18.i2.pp666-677.

[19] Moreira HS, Gomes dos Reis MV, de Araujo LS, Perpetuo e Oliveira T, Villalva MG. (2017) An experimental comparative study of perturb and observe and incremental conductance MPPT techniques for two-stage photovoltaic inverter. Proceedings of the 2017 Brazilian Power Electronics Conference (COBEP), 19-22 Nov 2017, Juiz de Fora, Brazil, pp. 1-6.
doi: 10.1109/COBEP.2017.8257370.

[20] Ghasemi MA, Foroushani HM, Parniani M. (2016) Partial shading detection and smooth maximum power point tracking of PV arrays under PSC. IEEE Trans. on Power Electronics, 31(9): 6281-6292. doi: 10.1109/TPEL.2015.2504515.

SYSTEM IDENTIFICATION (SI) MODELLING, CONTROLLER DESIGN AND HARDWARE TESTING FOR VERTICAL TRAJECTORY OF UNDERWATER REMOTELY OPERATED VEHICLE (ROV)

FAUZAL NAIM ZOHEDI[1]*, MOHD SHAHRIEEL MOHD ARAS[1], HYREIL ANUAR KASDIRIN[1], MOHD BAZLI BAHAR[1] AND LOKMAN ABDULLAH[3]

[1]Faculty of Electrical Engineering, Universiti Teknikal Malaysia Melaka,
[2]Center for Robotics and Industrial Automation (CeRIA),
[3]Faculty of Manufacturing Engineering, Universiti Teknikal Malaysia Melaka,
76100 Durian Tunggal, Melaka, Malaysia

*Corresponding author: fauzal@utem.edu.my

ABSTRACT: Underwater remotely operated vehicles (ROV) are important in marine industries to accomplish underwater exploration and surveying. The underwater environment makes it hard for ROV operators to control the manipulator while holding position simultaneously. This led to modelling and controller design for the vertical trajectory of ROV. In this paper, the System Identification (SI) modeling technique was used to model the vertical trajectory of the ROV. Then, the Proportional, Integral, and Derivative (PID) controller was implemented to control the trajectory. The SI modelling technique was used as it estimates the model based on the input and output relationship. MATLAB SI toolbox was used as the analytical software. Step and multiple step inputs were given to the system and the responses were recorded. The model with the best fit of 84.7% was selected and verified by comparing with actual output. The model response was then analyzed and the PID controller was implemented. The actual model had high percent overshoot (%OS) and steady state error (SSE). The PID implementation successfully reduced the %OS and eliminated the SSE.

ABSTRAK: Kenderaan bawah air kendalian jauh (ROV) adalah penting dalam industri marin bagi melaksanakan penerokaan dan pemerhatian bawah air. Persekitaran dalam air menyukarkan pengendali ROV bagi memanipulasi manipulator sambil memastikan kedudukan ROV secara serentak. Ini membawa kepada pemodelan dan mereka bentuk kawalan pergerakan menegak bagi ROV. Kajian ini menggunakan teknik pemodelan Sistem Pengenalan (SI) bagi memodelkan pergerakan menegak ROV. Kemudian, kawalan seimbang, menyeluruh dan terbitan (PID) dilaksanakan bagi mengawal trajektori. Teknik pemodelan SI digunakan kerana ia menganggarkan model berdasarkan kemasukan dan keluaran. Aplikasi MATLAB SI digunakan sebagai perisian analisis. Masukan satu langkah dan berbilang kali masukan telah dijalankan dan respons sistem direkodkan. Model yang paling sesuai mencapai 84.7% dipilih dan disahkan dengan perbandingan nilai keluaran sebenar. Respons model kemudiannya dianalisis dan kawalan PID dilaksanakan. Model sebenar mempunyai peratusan tinggi melampaui (%OS) dan ralat keadaan stabil (SSE). Pelaksanaan PID telah berjaya mengurangkan %OS dan menghapuskan SSE.

1. INTRODUCTION

Underwater remotely operated vehicles (ROV) are robots that substitute humans for underwater investigation and exploration. Underwater surveillance is important for sustainability of resources and underwater research [1]. Underwater exploration and surveillance, oil and gas pipeline monitoring, ship hull cleaning and other [2] complicated tasks require a highly skilled ROV operator to operate the ROV. The ability to manipulate the manipulator and maneuver the ROV at the same time is a must [3,4] for the operator. To ease the operator, automatic trajectory and position holding is necessary to ensure accurate control of the manipulator is achieved [4]. Other than that, there are also disturbances made by currents and waves [5,6] in the underwater environment. It increases difficulties for ROV operators to maintain the ROV at a certain depth.
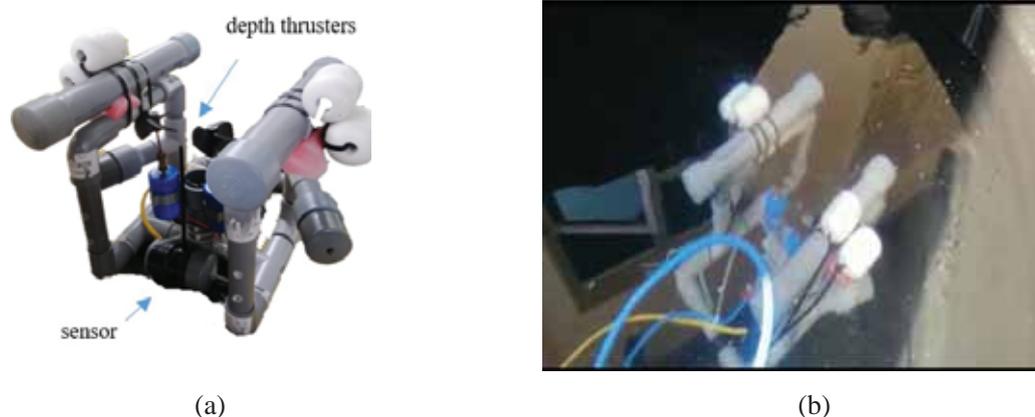
Fig 1: (a) ROV used in the research project; (b) ROV tested in controlled environment.

Underwater remotely operated vehicles (ROV) have six (6) degrees of freedom (DOF) which are surge (forward-reverse), sway, heave (vertical), roll, yaw (left and right) and pitch [7,8]. All these six degrees of freedom are coupled together. This coupling makes it difficult to model the ROV using mathematical modelling as many assumptions need to be made. Another alternative for modelling the ROV is using the System Identification (SI) technique. The SI technique is based on experimental data of the actual system. Input data is given to the system and the output response is recorded. These data are then run through the SI MATLAB toolbox to estimate the model based on the input-output relationship.

In this paper, a small ROV prototype was used. It was driven by 4 thrusters where 2 thrusters for vertical trajectory while another 2 thrusters were for forward-reverse and left-right turns. The vertical thrusters were put at the center of the ROV and placed side-by-side to provide stability to the ROV. The forward and reverse thrusters were placed at both sides of the ROV with 45 degrees angle that eased the turning process of the ROV. The ROV did not need to move fast as its solely used for investigation. The ROV had slightly positive buoyancy to ensure its ability to float if anything happened to the system. This position was measured as its initial position. The small ROV prototype used in the research project is shown in Fig. 1.

The ROV was tethered using a local area network cable (LAN). This cable provided 12V direct current (DC) to all 4 thrusters. The tethered cable also supplied 5V direct current (DC) for the sensor and was also used for data transfer between sensor and controller. The depth of the ROV was sensed using an MPC4250 pressure sensor. This sensor produced an output of 1.825 V at 1 atmosphere (ATM) and 1.8624 V at 22 cm in the water. The 22 cm is the location of sensor at the initial ROV condition floating in the water. This paper focuses on SI modelling

and the Proportional, Integral, and Derivative (PID) controller was designed for the vertical trajectory of a newly developed small ROV. PID was selected as the controller due to its simple and easy implementation.

There are several basic controllers that have been implemented to perform ROV depth control. The controllers are the PID, a fuzzy logic controller (FLC), a neural network controller (ANN), and a sliding mode control (SMC) [4,9-13]. The easiest controller to implement is the PID controller as it can be tuned using auto-tuning in MATLAB Simulink[14]. The PID is good at controlling linear systems. The FLC controller is good at adapting with disturbance [15,16] but difficult to tune. ANN tunes its system by using historical data or online data [17,18]. Due to this, the ANN controller has a slow reaction time. The SMC controller works based on limitations set for the system to follow a slide path [19]. Its disadvantage is energy waste as it bounces through the path. Table 1 shows a summary of advantages and disadvantages of the basic controllers discussed [20].

Table 1: Comparison of the controller design

Type of controller	Advantages	Disadvantages
PID	Easy to execute and maintain	Only for linear systems
FLC	No mathematical modelling required and precise order	Complicated in the tuning process
ANN	Convergence to a precise model	Slower response
SMC	Non-linear system	Energy waste occurs

2. SYSTEM IDENTIFICATION MODELLING

The System Identification (SI) modelling approach starts with data gathering of the inputs and outputs of the system by experimentation [21-24]. In this project, step input and multiple stairs input were given to the system. The step input was used to model the system while the multiple stairs input was used to verify the model with the actual system. SI will estimate the relationship between the input and the output and generate a model with percentage best fit value. The best fit value will indicate how close the model is to the actual system. The SI approach consists of five (5) steps [20], [25] which are observation and data gathering, model structure selection, model estimation, model validation, and application. Fig. 2 shows the flow chart of the SI approach.

From Fig. 2, the first step is observation and data gathering. In this process, the system is observed, and input-output data are gathered. The next 3 flows, model structure selection, model estimation, and model validation are selected automatically by the MATLAB SI toolbox. Model structure selection is where the estimation method is selected. Model estimation is where the system will create a model, and model validation is where the generated model is validated. The last step is model application where the generated model is applied with any controller to control the system. The generated model must at least obtain an 80% best fit value to be considered as an acceptable model. The generated model can be in transfer function or state space before implementing to any system, the model must also be checked for controllability. The state space of the model is substituted into Eq. 1. If the Mc rank is equal to matrix A in the state space, the system is controllable.

$$M_C = [B \ AB \ A^2B \ A^{n-1}B] \qquad (1)$$
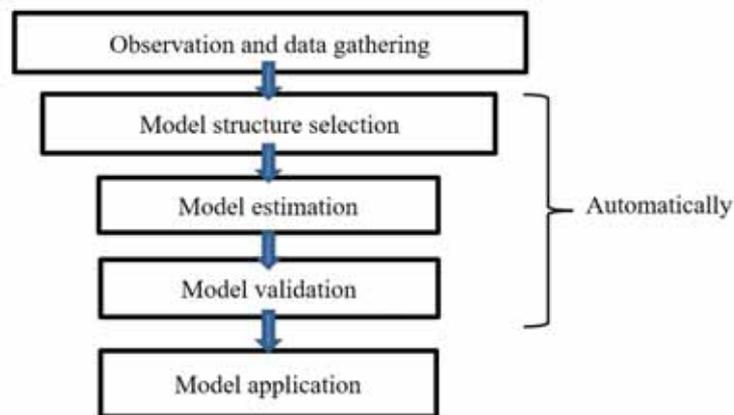
Fig. 2: Flow chart of SI approach.

2.1 Experimental Setup

In the experimental setup, there are 5 components as shown in Fig. 3. The components are a computer for user interface, a microcontroller Arduino Mega 2560 as interface between computer and ROV, a motor driver L298N to drive the thruster motor, a pressure sensor MPX425 to measure water pressure for depth, and the ROV.
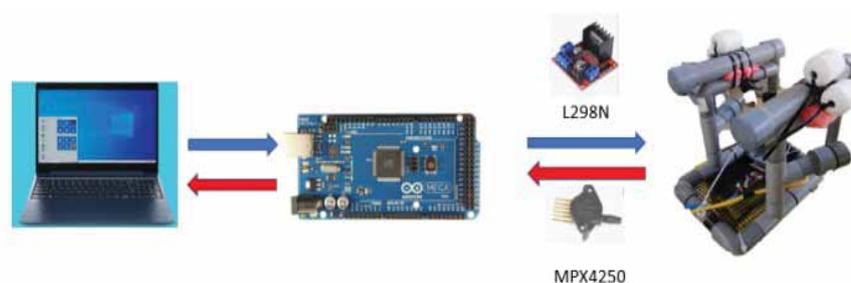
Fig. 3: Experimental setup.

The vertical trajectory of the ROV was observed by giving a step input and multiple stairs input to the system. The setups are listed below:

- Input step signal = 57 cm depth (limitation of tank)
- Repetitive stairs input = 57 cm and 40 cm
- Sampling time = 0.001 s
- Initial depth = 22 cm
- Initial volt = 1.825 V
- Triggered = 0.5 s for step and 0 s multiple input
- Experimental time = 10 s (step input) and 50 s (multiple input)

As the step given was 57 cm, the ROV submerged until 57 cm and maintained that position for 3 s. For multiple stairs input, the ROV submerged between 57 cm and 40 cm. From the sampling time, the step input produced 30000 data points while the multiple stairs input produced 40000 data points. The Simulink block diagram for data gathering is shown Fig. 4. The highlighted block in Fig. 4 shows the conversion block from volt to depth. The measurement from the sensor was divided with 1023 and multiplied with 5 V. This is due to the analog to digital conversion (ADC) of the Arduino that has a 10-bit configuration. The measured voltage was then converted to depth in centimeters.

Fig. 4: Simulink block diagram for data gathering.

3. CONTROLLER DESIGN

For this research project, the PID controller was implemented by simulation and experimentation. It was first developed by Elmer Sperry in 1911. Proportional (P) is used to reduce the time rise (Tr) but affects the percentage of Overshoot (%OS). Derivative (D) is used to decrease the %OS and increase the stability, while Integral (I) is used to eliminate the steady state error (SSE) but may affect the performance of the system [26]. Equation 2 shows the equation of the PID controller where Kp is the gain of P, Kd is the gain of D, and Ki is the gain for I.

$$\frac{output\ u(t)}{err\ (e)} = Kp + \frac{Ki}{s} + Kds \qquad (2)$$

Figure 5 shows the basic block diagram used in this research project for PID controller implementation. The P, I, and D were connected in parallel with each other as highlighted in Fig. 5. Table 2 shows the effect of varying Kp, Ki, and Kd (Fig. 5) to the transient parameters; [26] time rise (Tr), percentage overshoot (%OS), Settling Time (Ts), and Steady State Error (SSE). For this research project, the PID was tuned manually to get the result. A step input of 57 cm depth was given to the system and the result was discussed and analyzed.

Fig. 5: Basic PID block diagram.

Table 2: The effect of varying Kp, Ki, and Kd (Fig.5) to the transient parameters

Parameters	Time Rise (Tr)	Overshoot (%)	Settling Time (Ts)	Steady State Error (SSE)
Kp	Decrease	Increase	Small change	Decrease
Ki	Decrease	Increase	Increase	Eliminate
Kd	Small change	Decrease	Decrease	Small change

4. RESULTS AND DISCUSSION

This section displays results and discussion for SI modelling and controller implementation.

4.1 SI Modelling

The experimental data for SI modelling is shown Fig. 6(a) and Fig. 6(b). The dash in the figure shows input given while the line given is the output response. The curves and spikes in the result occur because of the vibrations of the thruster motor and tolerance of the sensor.

The SI approach was implemented using the step data in MATLAB SI toolbox. The result shows 84.7% best fit and is considered acceptable. The nonlinear squares method was automatically selected by the SI system. The transfer function generated is shown in equation 3 below.

$$H(s) = \frac{0.3808s^2 + 0.5898s + 1.87}{s^2 + 1.554s + 2.32} \tag{3}$$

(a)   (a)
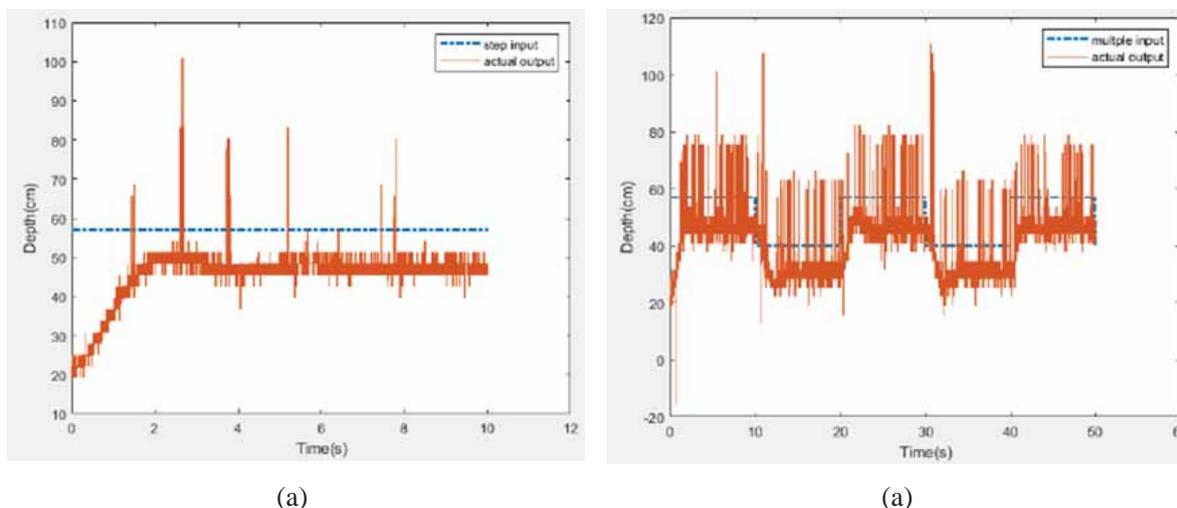
Fig. 6: Step input (a) and multiple stairs input (b) result.

The transfer function was then converted to state space and checked for controllability using Eq. 1. The result showed that the transfer function is controllable as the ranks of Mc and matrix A are similar. The transfer function was then verified by comparing it with the actual step output and multiple stairs output shown in Fig. 7 and Fig. 8.

The dashes in the figures were the model output while the wavy and spiked lines were the actual output. From the comparison in Fig. 7 and Fig 8, the model output was almost identical to the actual output. A bit overshoot and steady state error can be seen clearly in the model output and the actual output. The overshoot is 8.18%, the steady state error is about 11.4cm, and the settling time is 5.23 s. The overshoot and the error need were reduced by using the designed controllers.
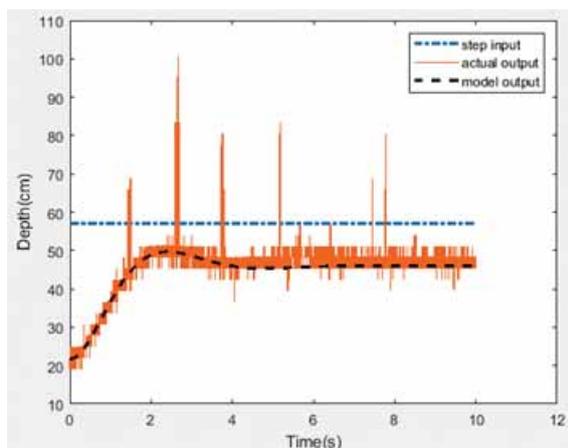
Fig. 7: step model output vs actual step output.
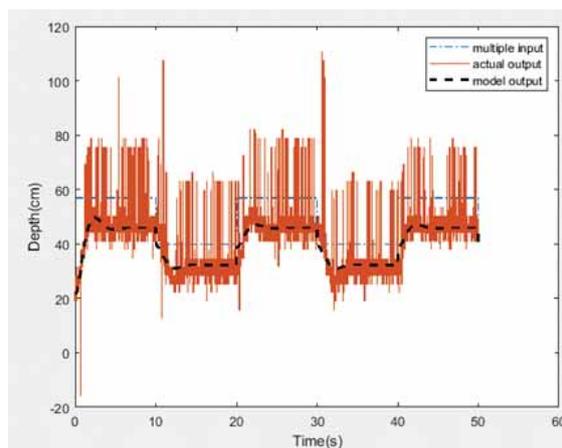
Fig. 8: Multiple stairs model output vs actual multiple stairs output.

4.2 Controller Implementation

The controller applied to the ROV model was PID. Figure 9 shows the result of the controller implementation.
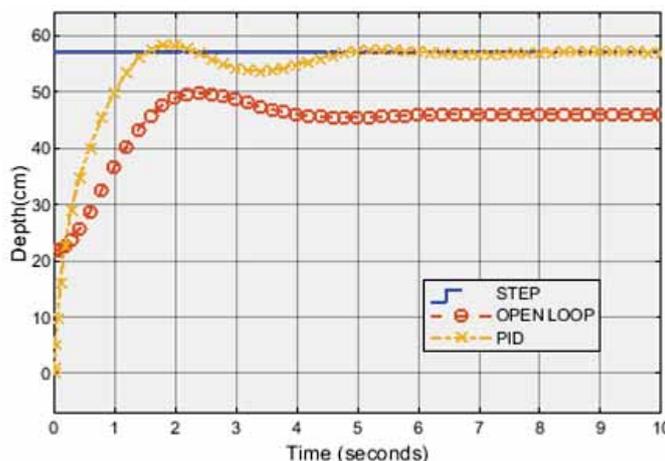
Fig. 9: PID controller implementation result.

The graph result in Fig. 9 was analyzed in terms of percentage overshoot (%OS), settling time (Ts), time rise (Tr) and steady state error (SSE). The result was tabulated in Table 3.

Table 3: Original system versus PID controller result

	Original System	**PID**
OS (%)	8.18	0.8
Ts (s)	5.23	4.37
Tr (s)	1.09	2.32
SSE (cm)	11.4	0

From the table, the PID shows remarkable reduction in the percentage overshoot (%OS) and steady state error (SSE). %OS went from 8.18% to 0.8% while SSE went from 11.4 to 0c

m. The settling time (Ts) decreased about 0.86 s from 5.23 s to 4.37 s. The drawback of the PID was the time rise (Tr) became slower by about 1s from 1.09 to 2.32s.

From this simulation result, the PID controller was then implemented experimentally to the ROV vertical trajectory. The result is shown in Fig. 10. From this figure, the experimentation for the PID controller looked almost identical to the simulation result. It shows that the PID controller can be implemented to the real ROV trajectory system.
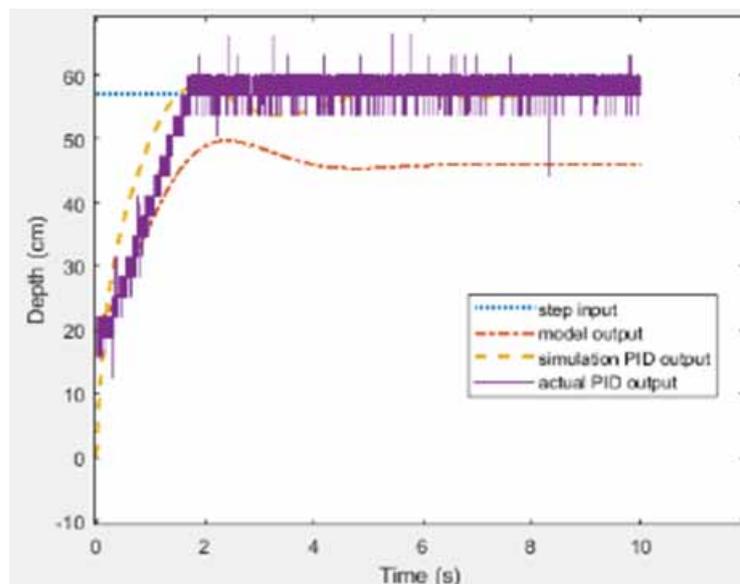
Fig. 10: Actual output result with PID controller implementation.

5. CONCLUSION

The modelling of vertical trajectory of an ROV is important to design a depth controller for the ROV. For the depth controller to maintain the ROV at certain position may offer more flexibility to ROV operators to manipulate the ROV manipulator or focus on any surveillance task. Modelling using the SI approach has proven to be successful as it manages to replicate the trajectory movement of the ROV at 84.7% best fit. The comparison between actual output and model output also shows near-identical results. In the controller design for the vertical trajectory, PID shows acceptable results in simulation and experimentally. Identical simulation and experimental results indicate that the PID controller can definitely be implemented in real life ROV trajectory control. The promising result shown by PID in all performance parameters used improve the original system. This paper is hopefully beneficial to engineers or controller designers to implement the SI model approach and design a PID controller for any real-life system.

ACKNOWLEDGMENT

REFERENCES

[1] Li G, Wu J, Tang T, Chen Z, Chen J, Liu H. (2019) Underwater acoustic time delay estimation based on envelope differences of correlation functions. Sensors, 19(5): 1-11. doi: 10.3390/s19051259.

[2] Binugroho EH, Wafiqqurochman, Mas'Udi MI, Setyawan B, Dewanto RS, and D. Pramadihanto D. (2019) EROV: Depth and Balance Control for ROV Motion using Fuzzy PID Method. IES 2019 - Int. Electron. Symp. Role Techno-Intelligence Creat. an Open Energy Syst. Towar. Energy Democr. Proc., pp. 637-643. doi: 10.1109/ELECSYM.2019.8901673.

[3] Goheen KR, Jefferys ER. (1990) Multivariable self-tuning autopilots for autonomous and remotely operated underwater vehicles. IEEE J. Ocean. Eng., 15(3): 144-151. doi: 10.1109/48.107142.

[4] Aras MSM, Abdullah SS. (2015) Adaptive simplified fuzzy logic controller for depth control of underwater remotely operated vehicle. Indian J. Geo-Marine Sci., 44(12): 1995-2007.

[5] B. Huang and Q. Yang, "Double-loop sliding mode controller with a novel switching term for the trajectory tracking of work-class ROVs," Ocean Eng., vol. 178, no. March, pp. 80–94, 2019, doi: 10.1016/j.oceaneng.2019.02.043.

[6] N. Kumar and M. Rani, "An efficient hybrid approach for trajectory tracking control of autonomous underwater vehicles," Appl. Ocean Res., vol. 95, no. October 2019, p. 102053, 2020, doi: 10.1016/j.apor.2020.102053.

[7] G. Antonelli, S. Chiaverini, N. Sarkar, and M. West, "Adaptive control of an autonomous underwater vehicle: Experimental results on ODIN," IEEE Trans. Control Syst. Technol., vol. 9, no. 5, pp. 756–765, 2001, doi: 10.1109/87.944470.

[8] M. C. Nielsen, M. Blanke, and I. Schjølberg, "Efficient Modelling Methodology for Reconfigurable Underwater Robots," IFAC-PapersOnLine, vol. 49, no. 23, pp. 74–80, 2016, doi: 10.1016/j.ifacol.2016.10.324.

[9] M. S. M. Aras, S. N. B. S. Salim, E. C. S. Hoo, I. A. B. W. A. Razak, and M. H. Bin Hairi, "Comparison of fuzzy control rules using MATLAB toolbox and simulink for DC induction motor-speed control," SoCPaR 2009 - Soft Comput. Pattern Recognitpp. 711–715, 2009, doi: 10.1109/SoCPaR.2009.143.

[10] Z. Tang, Luojun, and Q. He, "A fuzzy-PID depth control method with ., overshoot suppression for underwater vehicle," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 6329 LNCS, no. PART 2, pp. 218–224, 2010, doi: 10.1007/978-3-642-15597-0_24.

[11] M. S. Bin Mohd Aras, S. M. S. B. Syed Abdul Hamid, F. B. A. Azis, F. A. B. Ali, and S. Shah B Abdullah, "Study of the effect in the output membership function when tuning a Fuzzy logic controller," Proc. - 2011 IEEE Int. Conf. Control Syst. Comput. Eng. ICCSCE 2011, pp. 1–6, 2011, doi: 10.1109/ICCSCE.2011.6190485.

[12] M. L. Corradini and G. Orlando, "Brief Papers," Brain Cogn., vol. 32, no. 2, pp. 273–344, 1996, doi: 10.1006/brcg.1996.0066.

[13] M. . Aras, M.S.M., Abdullah, S.S. , Baharin, K.A. Nor, A.S.B.M. ,Mohd Zambri, "Model identification of an underwater remotely operated vehicle using system identification approach based on NNPC," Int. Rev. Autom. Control, vol. 8, no. 2, pp. 149–154, 2015.

[14] J. Ko, N. Takata, K. Thu, and T. Miyazaki, "Dynamic modeling and validation of a carbon dioxide heat pump system," Evergreen, vol. 7, no. 2, pp. 172–194, 2020, doi: 10.5109/4055215.

[15] K. Ishaque, S. S. Abdullah, S. M. Ayob, and Z. Salam, "Single input fuzzy logic controller for unmanned underwater vehicle," J. Intell. Robot. Syst. Theory Appl., vol. 59, no. 1, pp. 87–100, 2010, doi: 10.1007/s10846-010-9395-x.

[16] M. W. N. Azmi et al., "Comparison of controllers design performance for underwater remotely operated vehicle (ROV) depth control," Int. J. Eng. Technol., vol. 7, no. 3.14 Special Issue 14, pp. 419–423, 2018.

[17] N. Weake, M. Pant, A. Sheoran, A. Haleem, and H. Kumar, "Optimising parameters of fused filament fabrication process to achieve optimum tensile strength using artificial neural network," Evergreen, vol. 7, no. 3, pp. 373–381, 2020.

[18] H. Han, M. Hatta, and H. Rahman, "Smart ventilation for energy conservation in buildings," Evergreen, vol. 6, no. 1, pp. 44–51, 2019, doi: 10.5109/2321005.

[19] H. Yu, C. Guo, and Z. Yan, "Globally finite-time stable three-dimensional trajectory-tracking control of underactuated UUVs," Ocean Eng., vol. 189, no. March, p. 106329, 2019, doi: 10.1016/j.oceaneng.2019.106329.

[20] F. N. Zohedi, M. S. Mohd Aras, H. A. Kasdirin, and M. B. Bahar, "A new tuning approach of Single Input Fuzzy Logic Controller (SIFLC) for Remotely Operated Vehicle (ROV) depth control," Evergreen, vol. 8, no. 3, pp. 651–657, 2021, doi: 10.5109/4491657.

[21] T. N. Dief and S. Yoshida, "System identification for Quad-rotor parameters using neural network," Evergreen, vol. 3, no. 1, pp. 6–11, 2016, doi: 10.5109/1657380.

[22] T. N. Dief and S. Yoshida, "System identification and adaptive control of mass-varying quad-rotor," Evergreen, vol. 4, no. 1, pp. 58–66, 2017, doi: 10.5109/1808454.

[23] M. Sanap, S. Chaudhari, C. Vartak, and P. Chimurkar, "HYDROBOT: An underwater surveillance swimming robot," Proc. - 2018 Int. Conf. Commun. Inf. Comput. Technol. ICCICT 2018, vol. 2018-Janua, pp. 1–7, 2018, doi: 10.1109/ICCICT.2018.8325872.

[24] M. S. M. Aras, S. S. Abdullah, M. Z. A. Rashid, A. A. Rahman, and M. A. A. Aziz, "Development and modeling of unmanned underwater remotely operated vehicle using system identification for depth control," J. Theor. Appl. Inf. Technol., vol. 56, no. 1, pp. 136–145, 2013.

[25] M. H. F. Taib, M. N., Adnan, R. and Rahiman, Practical System Identification. Penerbit UiTM, 2007.

[26] M. A. Salim, A. Noordin, and A. N. Jahari, "A robust of fuzzy logic and proportional derivative control system for monitoring underwater vehicles," 2nd Int. Conf. Comput. Res. Dev. ICCRD 2010, pp. 849–853, 2010, doi: 10.1109/ICCRD.2010.187.

EYE BLINK IDENTIFICATION AND REMOVAL FROM SINGLE-CHANNEL EEG USING EMD WITH ENERGY THRESHOLD AND ADAPTIVE FILTER

LUTFI AIZAT MOHAMED JEFRI, FARIDAH ABD RAHMAN[*],
NOREHA ABDUL MALIK AND FARAH NADIA MOHD ISA

*Dept. of Electrical and Computer Engineering,
International Islamic University of Malaysia, Kuala Lumpur, Malaysia*

*Corresponding author: faridah_ar@iium.edu.my

ABSTRACT: Electroencephalography (EEG) is a non-invasive method for measuring electrical activity in the brain, which reflects the underlying neural activity of the brain. In recent years, portable EEG devices become more ubiquitous in domestic uses, research and clinical applications due to their compact design and ease of use in various settings. Like many other bio signal modalities, EEG devices are prone to the interference of physiological artifacts, mainly from eye blinking. However, since portable EEGs are equipped with only a few channels at most or sometimes just contain a single channel, removing the eye blink artifact from the EEG data is a challenge. The conventional artifact removal method using source separation cannot be applied to a single-channel EEG signal. Eye blink artifact removal is important because its spectrum overlaps with the EEG's theta and delta frequency bands, which can be confused with brain activity. Univariate-based removal method is compatible with EEG data with few channels. This paper presents a method to remove eye blink artifact based on single-channel EEG processing using Empirical Mode Decomposition (EMD) and Adaptive Noise Cancellation (ANC) system. By applying energy thresholds in EMD, there is no need to incorporate EMD with other methods to extract eye blink component accurately. ANC is used to converge the extracted eye blink component for effective eye blink artifact removal with very minimal changes to affected EEG data. The proposed method was tested on simulated EEG signals, and the result showed a good Root Mean-Square Error (RMSE) average value of the cleaned EEG (0.3211 ± 0.2738) and a high Correlation Coefficient (CC) average value of the cleaned EEG (0.9430 ± 0.0839).

ABSTRAK: Electroensefalografi (EEG) adalah kaedah bukan invasif untuk mengukur aktiviti elektrik di dalam otak, yang mencerminkan aktiviti saraf dalam otak. Kebelakangan ini, peranti EEG mudah alih menjadi lebih meluas dalam kegunaan domestik, penyelidikan dan aplikasi klinikal kerana reka bentuknya yang padat dan kemudahan penggunaan dalam pelbagai tetapan. Seperti kebanyakan modaliti biosignal yang lain, peranti EEG terdedah kepada gangguan artifak fisiologi, terutamanya daripada kerdipan mata. Walau bagaimanapun, memandangkan EEG mudah alih dilengkapi dengan paling banyak pun hanya beberapa saluran, atau kadangkala hanya satu saluran, mengalih keluar artifak kerdipan mata daripada data EEG adalah satu cabaran. Kaedah penyingkiran artifak konvensional menggunakan pemisahan sumber tidak dapat digunakan pada alat EEG satu saluran. Penyingkiran artifak kerdipan mata adalah penting kerana spektrumnya bertindih dengan jalur frekuensi teta dan delta EEG, maka boleh dikelirukan dengan aktiviti otak. Kaedah penyingkiran berasaskan univariat adalah serasi untuk data EEG dengan saluran yang sedikit. Kertas kerja ini membentangkan kaedah untuk membuang artifak kelipan mata berdasarkan pemprosesan EEG saluran tunggal menggunakan Penguraian Mod Empirikal (EMD) dan Pembatalan Bunyi Adaptif (ANC). Dengan menggunakan ambang tenaga dalam EMD, tiada

keperluan untuk menggabungkan EMD dengan kaedah lain bagi mengekstrak komponen kerdipan mata dengan tepat. ANC digunakan untuk menumpu komponen kerdipan mata yang diekstrak bagi penyingkiran artifak kerdipan mata yang berkesan dengan perubahan yang sangat minimum pada data EEG yang terjejas. Kaedah yang dicadangkan telah diuji pada signal EEG yang disimulasi, serta hasilnya menunjukkan nilai purata Ralat Min Kuasa Dua Purata (RMSE) yang baik bagi EEG yang dibersihkan (0.3211±0.2738), dan nilai purata Pekali Korelasi (CC) yang baik bagi EEG yang dibersihkan (0.9430±0.0839).

KEY WORDS: *Adaptive Filter, Eye Blink Extraction, EMD, Energy Thresholding, Univariate Processing*

1. INTRODUCTION

Electroencephalography (EEG) is one of the brain monitoring modalities that records electrical signals that originate from the neuronal activities across the cerebral cortex of the brain. These signals correspond to the higher level of the brain functions such as thinking, solving problems, memory, learning, feeling and processing emotions, intelligence, behavior regulation and body movement to some extent. As such, coupled with EEG equipment that is mainly easy to be used, there are many types of research on the matter related to the brain, utilizing EEG as a medium or small input data to analyze and explore the exactitude of a hypothesis and to synthesize solution for the objective of the research. For example, Jacobsen et al. use EEG to study EEG's beta power variability that corresponds to gait movement variability down the terrain [1]. The design of the EEG equipment is relatively of simpler implementation, making it possible to be miniaturized into a compact and portable EEG device [2]. Likewise, the nature of portable EEG that is simpler and less cumbersome to be set up than medical-grade EEG, making it stand out for use. As such, the interest in incorporating the portable EEG device into brain related research and application development is becoming increasingly ubiquitous [3]–[5]. A review of different brands of portable EEG is also made of their respective efficacy on different domains of brain related study [6]. The application of EEG in research and domestic use is expected to remain ubiquitous.

EEG signals propagate from the cerebral cortex, an outer layer of the brain, then across the meninges, skull, and scalp of the subject to the EEG electrodes, making them prone to contamination with other signals that are not of brain origin, which are called artifacts. These artifacts may be external, meaning they originate from the EEG equipment like electrical interference and alternating current artifact, or internal, meaning signals that originate from a physiological body such as electrooculogram (EOG), electromyogram (EMG) and electrocardiogram (ECG). Although the internal artifacts are of physiological origin, these signals may not be desirable for brain-related analysis because they are not implicative to brain activity during the EEG recording session. Hence, in most cases, artifacts need to be discarded in order to obtain clean, pure EEG data. Unlike internal artifact, external artifacts contamination can be avoided by correct equipment set up. While internal artifacts prevention is accomplished by limiting the subject movements and making sure the subject is calm during EEG recording. Likewise, there have been many techniques being proposed on how to remove these internal artifacts. Nevertheless, eye blinks are considered one of the most pervasive artifacts contaminating the EEG signals. Eye blink artifact is persistent in its presence during EEG recording even with restrictive and controlled experimental design. Ultimately, there would be many scenarios where eye blink artifact removal comes in handy when it comes to EEG recording.

Therefore, with relation to portable EEG, the univariate-based eye blink removal method is preferred compared to the multivariate-based method. This is because the portable

EEG is only equipped with a few EEG channels. Univariate refers to a single EEG channel use, while multivariate refers to the use of multiple EEG channels. The multivariate-based method generally needs a larger number of inputs as its performance is linearly proportional to the number of inputs. Therefore, the univariate method is more effective than the multivariate method in processing EEG data retrieved from portable devices as it only needs one input to perform effectively. The univariate method also needs to be able to remove the eye blink without affecting the EEG component that is superimposed with the eye blink artifact while maintaining its quality of being straightforward in terms of implementation.

In this paper, we propose a method of eye blink region identification using five samples window-mean and an energy thresholded-Empirical Mode Decomposition (EMD) to extract the eye blink component, coupled with Recursive Least-Mean Square (RLS) adaptive filtering to remove the eye blinks from single-channel EEG. The proposed method aims to approximate the eye blinks artifact within the identified eye blink region with high similarity as a reference in the adaptive noise cancellation system.

2. RELATED WORKS

In this section, a few univariate eye blink removal methods are briefly reviewed, along with the basis of the proposed technique is elaborated. These includes the eye blink detection method, eye blink component extraction and eye blink artifact filtering.

2.1 EEG Eye Blink Artifact Detection

Eye blink artifact is constituted of the overlapped frequency spectrum with EEG theta and delta frequency band (0.3 Hz to 7 Hz) that is indicated by 10 to 100 times bigger amplitude-wise deflections within the EEG data [7]. Occurrences of eye blink artifact contamination in EEG data can be traced by abrupt prominent EEG magnitude change. Therefore is a foolproof way of detecting the eye blink artifact. This is because the propagation of potential from corneal-retinal dipole by the eye lid during eye blinking activity exhibits abrupt deflective spikes in EEG recording [8]. The specific EEG recording brand may influence the temporal characteristic of the deflections. For example, a voluntary eye blink artifact captured using an OpenBCI amplifier, as shown in Fig. 1. (a) has different morphology when compared to an eye blink artifact that is captured using g.Tec amplifier in Fig. 1. (b), and Emotiv device in Fig. 1. (c).
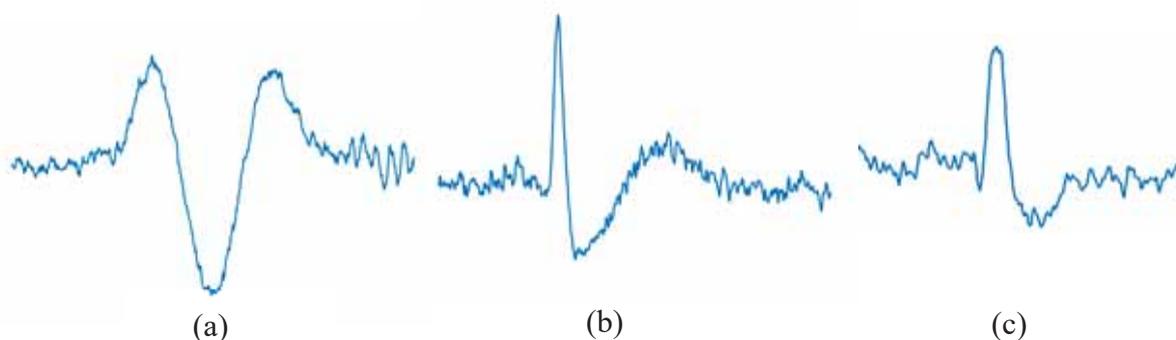

(a)　　　　　　　　(b)　　　　　　　　(c)

Fig. 1. Example of voluntary eye blink captured in frontal channel of three different
EEG devices (a) OpenBCI, (b) g.Tec cap-type device, (c) Emotiv

Nevertheless, the consistency of the recorded eye blink artifact characteristic across any EEG equipment brand is shown temporally by the huge deflective spike in any EEG data that

contains the eye blink artifact. As such, most eye blink detection methods use amplitude-based quantification as means for non-eye blink and eye blink region segregation. This may be in terms of windowed amplitude mean as proposed by authors of [9], highest amplitude as reference for eye blink temporal span approximation as posited in [7], amplitude power, as demonstrated in [10], and amplitude energy as being presented in [11].

Therefore, amplitude-based quantification is suitable to be incorporated in detecting the EEG segment of interest. For the proposed method in this paper, amplitude-based thresholding based on five samples window-mean is used to gauge the EEG data amplitude differences to dismiss the samples with amplitude lower than the threshold but retain the samples that are higher than the threshold to mark the eye blink region, as being demonstrated in [12]. This approach is independent of any assumption on the duration of the eye blinking activity to avoid including clean EEG region with the identified eye blink region or excluding any portion of the eye blink region.

2.2 Eye Blink Component Extraction

Decomposition-based univariate signal processing is a popular class of methods for eye blink removal. Lately, Singular Spectrum Analysis (SSA), with various implementation designs, has been proposed to remove eye blink artifact from a single EEG channel. However, the number of components or subspaces that an input can be composed into must be preset by mapping the single-channel EEG into a multivariate data matrix before decomposition by SSA is executed. Subsequently, the appropriate workaround is introduced to ensure satisfactory eye blink component reconstruction. For example, Maddirala et al. use Hjorth mobility as a threshold to choose the SSA subspaces for the eye blink reconstruction [13]. While in [14], the eigenvalue ratio is used as the threshold to select the subspaces that correspond to the eye blink component for eye blink reconstruction. Nevertheless, different temporal lengths of the recorded EEG and also possibly inconsistent morphology of the eye blink artifact due to different use of EEG devices may cause the variability on the suitable number of subspaces (M) to be preset for the signal decomposition that yields good component separation, which ultimately, causes variability to the threshold value for SSA subspaces selection. The ambiguity of the value that needs to be used in the parameters setting requires preparatory testing, which may complicate the implementation of the proposed methods. In contrast, EMD may bypass this problem for its empirical way of decomposing a signal. Therefore, EMD is adaptable to variable eye blink morphology and EEG temporal length without any respective parameters' adjustment.

EMD is a data-driven method that decomposes signals in the time domain by a sifting process. This process identifies all local extrema in the input signal, starting with those corresponding to the smallest oscillation period. Then, the identified oscillatory mode is extracted, which also yields the intermediate residual, as shown in Fig. 2. This oscillatory mode is called the intrinsic mode function (IMF). The identified extrema are used to construct upper and lower envelopes before being averaged to find the mean envelope. The mean envelope is then subtracted from the signal to acquire the oscillatory mode that consists of a spectral that is higher than the spectral of the derived mean envelope. A thorough explanation of how EMD operates is well presented by Zeller et al. in [15]. Generally, EMD segregates signal components locally and separates the data into locally non-overlapping time scale components.

From the decomposing eye blink region perspective, the first IMF is extracted from the signal, the intermediate residual is produced, and the sifting process is iterated again using the intermediate residual to extract other IMFs until it becomes a monotonic component. Consequently, the later extracted IMF has slower oscillatory characteristics than the former.

This shows that EMD yields IMF components that start with the highest oscillatory rate to the final monotonic residual. Moreover, using the local extrema to derive the IMF allows the derivation to be automatic and adaptive time-variant filtering, which results from IMFs that exhibit non-stationarity nature with amplitude-frequency modulated (AM/FM) characteristics. As such, the nature of EMD sequential operations in filtering out IMFs can be approximated as a dyadic filter. This is because, on average, the amount of the extrema is reduced by one-half from an IMF to the next IMF.
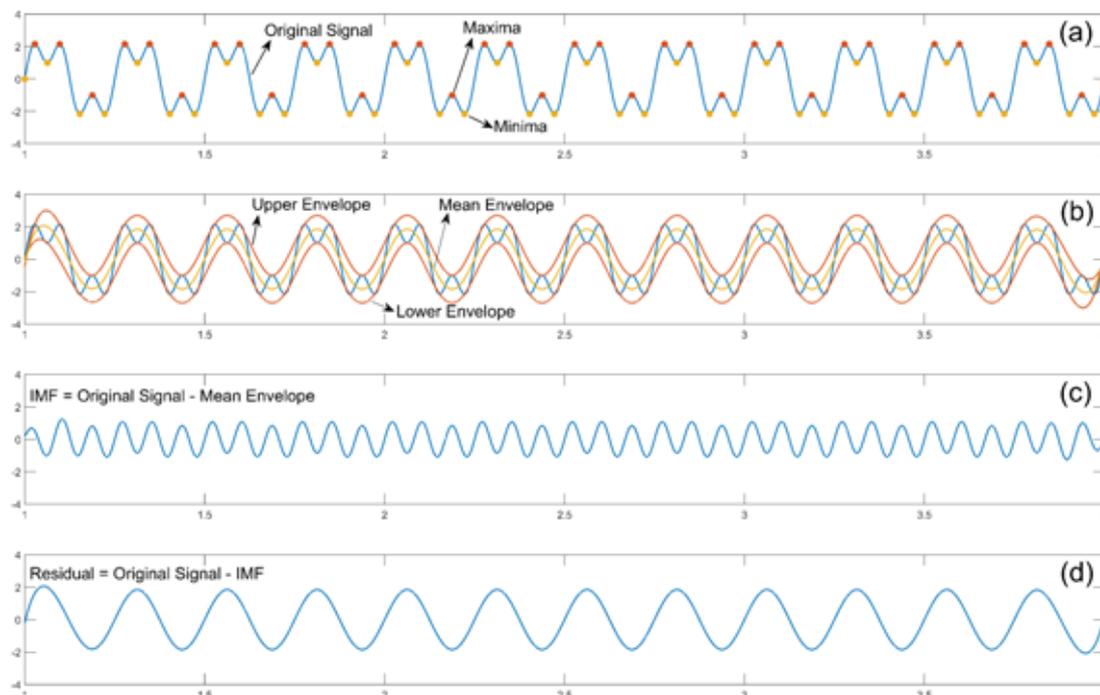
Fig. 2. The simplified sifting process visualization. (a) Extrema identification. (b) Mean envelope construction. (c) First IMF. (d) The residual, or the last IMF.

Since the advent of EMD, it has been integrated into the process of EOG artifact removal and has subsequently been widely disseminated in the academic literature. In [16], the correlation coefficient (CC) between the summed up IMFs with the true EEG component is used as an indicator for eye blink component presence, which is calculated after every time an IMF is added up from the first IMF. When the CC decreases to lower than 0.9, the latest IMF that is added up is deemed to have eye blink component. Hence, starting from that latest IMF to the residual, the IMFs are added up to reconstruct the eye blink component that can be used as a reference for adaptive noise cancellation (ANC). However, this technique needs the ground truth pure EEG that is similar in its temporal features to the EEG within the eye blink-contaminated data in order to calculate the CC value, which is counterproductive. While in [17], CC is also used between the artifactual EEG segment and the IMFs. Due to the scale of the eye blink component being larger than the EEG signal, the CC value obtained with IMFs containing eye blink component is larger. The CC threshold of 0.5 is used to discriminate between IMFs representing EEG and eye blink component, of which the IMFs with CC that is larger than 0.5 will be discarded before reconstruction. This method is vulnerable to the loss of some EEG components, especially the one that lies within the delta band. In [18], the 95% confidence interval of log energy of fractional Gaussian noise (fGn) is used as a threshold to select IMFs representative of EOG artifact. However, there is no workaround to deal with the

potential of EMD's mode-mixing effect which would possibly cause some EEG component removal along with the reconstructed EOG.

The ability of EMD to decompose a signal into their respective oscillatory modes also makes it possible to employ multivariate data processing for a single EEG channel. For example, EMD is incorporated with canonical correlation analysis (CCA) in a way that the EMD is made to produce partial separation of eye blink artifact from the EEG data. The CCA is used to find the canonical weights that can be used to approximate the eye blink component that is free from EEG mixing [19]. Nevertheless, EMD can be utilized solely to extract the eye blink component without combining it with other techniques for its empirical method in decomposing the input signal.

We proposed using five samples window-mean thresholding to identify the eye blink regions within the single-channel EEG. Each identified region is then decomposed using EMD to extract the eye blink component. Thus, EMD is used only upon the isolated region as decomposing a time series with the transient component, which in this case is the eye blink component, causes mode-mixing due to over-sifting that may be induced during IMF derivation, as discussed by Zeller et al. [15]. Also, the EMD used for the proposed method is as proposed by Rato et al. [20], as its implementation addresses the problem of the EMD in terms of extrema computation, extrema interpolation, boundary conditions and sifting criterion.

2.3 Eye Blink Artifact Removal

Adaptive filtering is a filter which receives reference input $r_1(n)$ as its input impulse to find the least resultant error e(n) between the produced impulse response y(n) and the targeted component $r_2(n)$. An adaptive filter can be represented as an impulse filter as the following Eq. 1:

$$y(n) = \sum_k^L w_k \cdot r_1(n-k) \tag{1}$$

The parameter L is the order of the filter, w_L is the filter coefficient with span of L, $r_1(n)$ is the reference input and y(n) is the impulse response.

The $r_1(n)$ is useful as it is the essential factor in producing the impulse response that is converged to the desired component $r_2(n)$ that is mixed with the input signal s(n). This is because the source of $r_1(n)$ comes from the same source as $r_2(n)$. In the context of the interest, $r_1(n)$ and $r_2(n)$ represent eye blink component. This means the eye blink component must be available to remove the respective eye blink artifact from the affected EEG segment.

In this paper, the eye blink component is extracted from the identified eye blink region in noisy EEG signals, as opposed to relying on the EOG channel which records the eye activities such as blinking and movement. The eye blink component extraction method is incorporated because of consideration for the ubiquity use of portable EEG in research and domestic uses. For the univariate method, eye blink removal is more plausible in its use of EEG data recorded using portable EEG devices as it only consists of a few EEG channels and many have no EOG electrode. Therefore, the eye blink removal method is designed to operate without the requirement for an EOG channel by extracting the eye blink component from noisy EEG. This extracted eye blink component is used as the reference r_1 in adaptive filtering.

In adaptive filtering, $r_2(n)$ removal is done by producing y(n) that is equivalent to the eye blink component within the EEG segment to not affect the EEG component during the removal. In order to find y(n), the error e(n) is used as feedback to assist the convergence of y(n). As shown in Eq. 2, the e(n) is the difference between the y(n) and $r_2(n)$ that is desired to

be as small as possible, as the smaller the e(n) results better converged y(n), as depicted in Eq. 3.

$$e(n) = r_2(n) - y(n) \tag{2}$$

$$e(n) \approx 0, y(n+1) \approx r_2(n+1) \tag{3}$$

In order to cater for the next y(n) computation, the current e(n) is used as feedback into the adaptive algorithm. The adaptive algorithm optimizes the weight to converge the y(n+1) to $r_2(n + 1)$, as being described in the following:

$$w_k(n+1) = w_k(n) - 2 \cdot \mu \cdot e(n) \cdot r_1(n - k) \tag{4}$$

The value of μ is between 0 and 1.

$$y(n+1) = \sum_k^L w_k \cdot r_1(n + 1 - k) \tag{5}$$

The weight optimization in Eq. 4 is based on the least mean square error (LMS) algorithm. The μ parameter is a constant to ensure fast adaptation for $w_k(n+1)$ that can converge y(n+1) to $x_2(n)$ effectively.

In this paper, recursive least mean square error (RLS) is used for weight optimization because RLS needs lesser n iteration counts or steps number to converge the y(n) to its respective $r_2(n)$, hence a faster convergence rate. Faster convergence impacts its output, especially when the input is of a short segment. This applies to the marked eye blink region as its temporal length may span at most by only 2 seconds. The RLS algorithm is as the following:

$$w_k(n+1) = w_k(n) - 2 \cdot \mu \cdot e(n) \cdot K(n) \tag{6}$$

K(n) is the filter coefficient vector given by Eq. 7 as the following.

$$K(n) = \frac{P(n) \cdot r_1(n)}{\lambda + r_1^T(n) \cdot P(n) \cdot r_1(n)} \tag{7}$$

P(k) is the inverse correlation matrix with its initialization as the following:

$$P(0) = \begin{bmatrix} \delta^{-1} & 0 & \cdots & 0 \\ 0 & \delta^{-1} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \delta^{-1} \end{bmatrix} \tag{8}$$

The parameter δ is the regulatory factor and the constant k is the length of the tap input vector. The P(k) adaptation is done as depicted by the following equation.

$$P(n+1) = \lambda^{-1} \cdot P(n) - \lambda^{n-1} \cdot K(n) \cdot u^T(n) \cdot P(n) \tag{9}$$

The fact that P(n) and K(n) need each other during the weight optimization is why it is called a recursive least mean square adaptive algorithm.

3. METHODOLOGY

The implementation of the proposed method is described in this section. Fig. 3 illustrates the outline of the methodology in terms of the block diagram. Firstly, the procedure for EEG dataset preparation is described. Then, the eye blink region detection, eye blink component extraction and eye blink artifact removal are explained. Finally, the performance metric is calculated upon the removal output accordingly with the original simulated EEG data that is juxtaposed with the performance metric from Egambaram et al. [19] to prove the effectiveness of our method.

```
┌─────────────────┐      ┌─────────────────┐      ┌─────────────────────┐
│   EEG data      │ ───▶ │  Eye blink region│ ───▶ │  Eye blink component │
│   simulation    │      │  identification  │      │     extraction       │
└─────────────────┘      └─────────────────┘      └─────────────────────┘
                                                             │
                                                             ▼
┌─────────────────┐      ┌─────────────────┐      
│   Performance   │ ◀─── │   Eye blink     │ ◀────
│ metric calculation│    │ artifact removal │      
└─────────────────┘      └─────────────────┘      
```
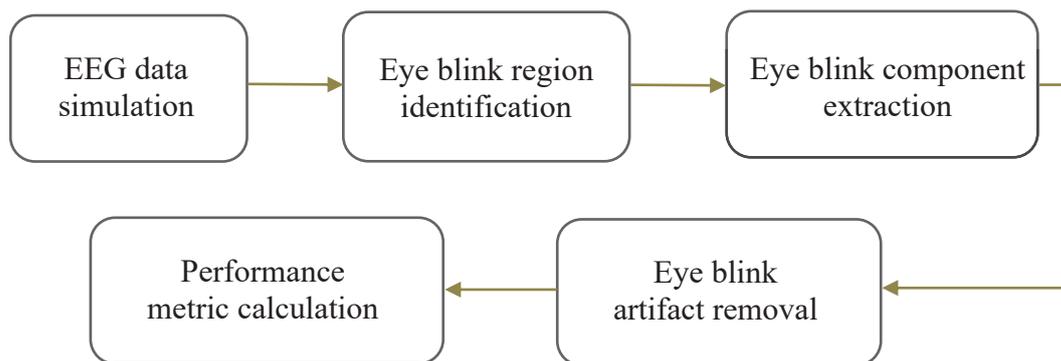
Fig. 3. The block diagram of a methodology for our eye blink removal

3.1 EEG Dataset Preparation

The EEG dataset used has been schemed to have an EEG component that can be ground truth data to assess the performance of the proposed method in preserving the brain-originated signals in the noisy EEG after eye blink artifact removal. Thus, EEG data is simulated by superimposing the pure EEG data with eye blink artifact. The superimposition of eye blink artifact can be understood as a linear mix between two signals:

$$X(t) = Y(t) + Z(t) \tag{10}$$

$X(t)$ is the eye blink contaminated-EEG, while $Y(t)$ is the pure EEG component and $Z(t)$ is the eye blink artifact. The EEG data used in the experiment to assess the proposed technique is simulated in a way that it is possible to compare the performance metrics values, which is documented in [19]. Exponential functions represent the eye blink components $Z(t)$ with different peak magnitudes that can be expressed mathematically as follows:

$$Z(t) = \sum A e^{-(10t-B)^2} \tag{11}$$

A is the peak height ranges from 7 to 19 in magnitude and B is the temporal location of the peaks.

Fig. 4 (a) shows an example of the synthetic eye blink artifact. As for the EEG component, it is synthesized by using the pink noise for which its amplitude range between 0.0±3. Fig. 4 (b) illustrates the synthetic EEG, $Y(t)$. In order to obtain a synthetic eye blink artifact contaminated-EEG, $X(t)$, the synthetic eye blink component and synthetic EEG are added together as demonstrated by Eq. 10. Fig. 4 (c) depicts the synthetic $X(t)$. This dataset is used to evaluate the equivalent performance metrics that later are used to compare with other EMD-based technique proposed by Egambaram et al. Every unit of $Y(t)$, $Z(t)$ and the resultant $X(t)$ are simulated with the length of 10 seconds at 256 sampling rate or 2560 samples.

3.2 Eye Blink Region Detection

In order to extract the eye blink component, the eye blink regions and artifact-free regions within the EEG data need to be identified. The targeted processes only on the specified EEG regions not only reduce the computational requirement and time but also prevent the mode-mixing in the decomposed IMFs during eye blink component extraction by EMD. To be able to mark the eye blink region, amplitude thresholding is used to exploit the conspicuous difference of amplitude between pure EEG and eye blink artifact. In order to do so, the pure EEG maximum amplitude of the eye blink artifact-contaminated EEG needs to be ascertained beforehand. Accordingly, EEG data that is recorded or prepared that corresponds to the pure EEG component in the EEG of interest is computed as its maximum amplitude. This maximum

amplitude value is a discriminatory factor in the eye blink artifact-contaminated EEG data. The following is a pseudocode to describe the steps in setting the amplitude threshold value:

Algorithm: Finding the threshold for the deflection detection algorithm for each EEG data channel.

Input: Array of pre-processed baseline EEG data, [data]

Output: Threshold for deflection detection

$[data_{mean}] \leftarrow$ mean([EEG_data(1:5:end)])

$[data]min \leftarrow$ min($[data_{mean}]$)

$[data]max \leftarrow$ max($[data_{mean}]$)

$[value]max \leftarrow$ max(abs([data]max), abs([data]min))

threshold $\leftarrow [value]max +1$

From this pseudocode, the mean of every five-sample window in the EEG data is used to determine the threshold instead of the individual sample because this window-mean can act like smoothing against the outlier amplitudes from EEG high frequency-component that may wrongly influence the determination for the suitable threshold value. If the value of the threshold is too high, it may bypass the EEG segment that contains an eye blink artifact (false negative), but if it is too low, it may include a clean EEG segment (false positive). Fig. 5 (a) illustrates the threshold (red color) that discriminates the eye blink region and the clean EEG. In the case of EEG data used to test the proposed method, the synthetic EEG data Y(t) is used to compute the amplitude threshold.
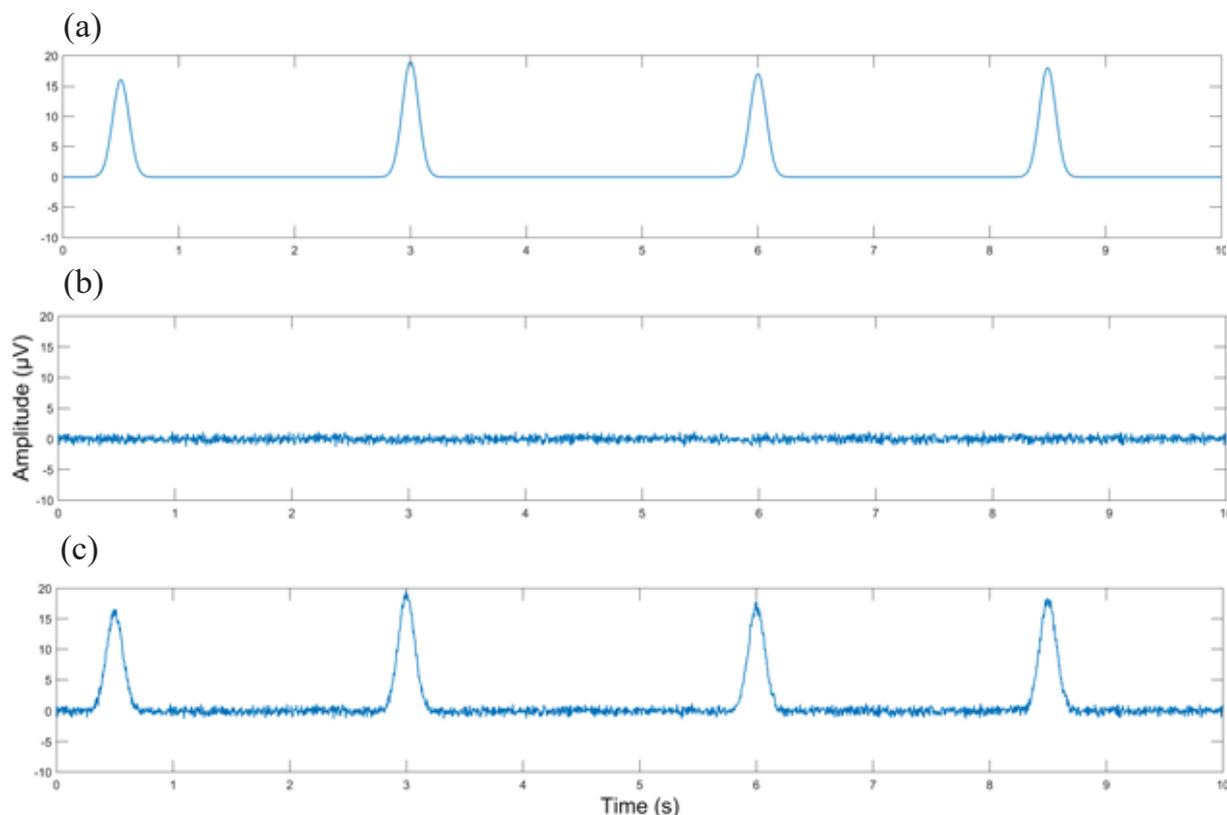

Fig. 4. (a) The simulated eye blink component Z. (b) The simulated EEG component Y. (c) The simulated eye blink contaminated-EEG X

As shown in the last line of the pseudocode, the threshold value is concluded as the maximum value added by 1 to dismiss clean EEG data samples effectively during eye blink artifact detection. During the detection, as visualized by Fig. 5 (b), the mean of every five samples is again computed and compared against the threshold to deem it as part of the eye blink region, as represented by an orange color or clean EEG samples (blue color) in a temporally successive manner. The detection by five-sample window-mean also helps to differentiate the five-sample segment whether it is a clean EEG sample or a contaminated EEG sample, more discretely.

The whole eye blink region is considered to be detected completely when no more 5-sample mean value is detected as an eye blink contaminated EEG segment for another 150 samples consecutively.

Before the detected eye blink-contaminated EEG samples are marked as a complete eye blink segment, the neighbouring samples at the beginning and end of the detected segment are included as part of the detected segment until the nearest zero-crossing is reached at both ends. At this point, an eye blink-region detection is complete. This is visualized in Fig. 5 (c), which shows that, in orange color, a complete eye blink region is marked. Then, the detection moves to the next five samples to detect another eye blink region.

(a) (b) (c)
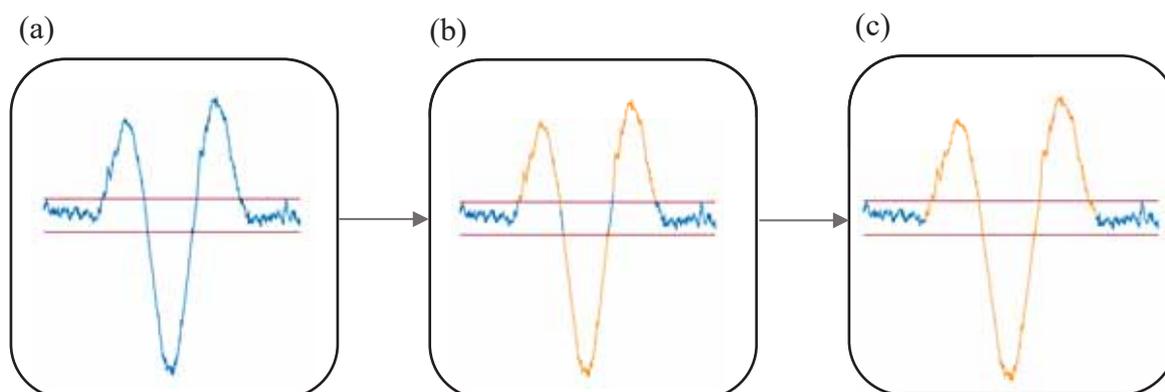
Fig. 5. (a) The window-mean threshold (red line) is used to discriminate the eye blink region. (b) Eye blink samples are detected (orange color). (c) The whole eye blink region is approximated by the nearest zero-crossing

3.3 Eye Blink Component Extraction

After the eye blink region is detected, EMD decomposes the segment. Before the decomposition, the segment is segregated into its individual deflection, as shown in Fig. 6 (b). To ensure smooth approximation of the envelope during sifting process, a constant α is used as a factor for the mean envelope e(t) whenever the e(t) is subtracted from the signal x(t) or the intermediate residual $r_n(t)$ as the following:

$$x(t) = x(t) - \alpha \cdot e(t) \tag{12}$$

The constant α value is between from 0 to 1.

Energy thresholding is deployed during IMF derivation and as a stopping criterion. For IMF derivation, qResol is used as an additional condition to derive an IMF. The threshold qResol is as an energy ratio of which when the energy of the e(t) is qResol times lower than the energy of x(t) or $r_n(t)$, the IMF is derived from the following:

$$\text{If qResol} > \frac{\text{energy of x(t) or } r_n(t)}{\text{energy of e(t)}}, \text{ IMF is derived.} \tag{13}$$

Threshold qResol aims to avoid constructing envelopes that introduce new components during IMF derivation. As for the energy threshold for the stopping criterion, the ratio of x(t) or $r_n(t)$ to the new residual produced ($r_{n+1}(t)$), which is produced after an IMF is derived, called threshold qResid. The decomposition stops when the value of qResid is lower than its set threshold:

$$\text{If qResid} > \frac{\text{energy of x(t) or } r_n(t)}{\text{energy of } r_{n+1}(t)}, \text{ the decomposition is stopped.} \tag{14}$$

The incorporation of threshold qResid influences the decomposition by EMD so that it does not produce a monotonic signal as its final residuum in a definitive manner like the original EMD would. This behavior is useful when the objective of the decomposition is to extract low oscillatory components with specific temporal features.

Before the decomposition, the deflection segregation is done according to the zero-crossings between each deflection. Deflection segregation allows the eye blink component to be extracted by EMD without being fragmented across IMFs due to over sifting. Eye blink artifact morphology is independent of any mathematical basis. However, the individual deflection generally resembles a peak or valley with one extremum, which is consistent with the EMD final residuum of which, after being generated, the decomposition stopped. Hence, as illustrated in Fig. 6 (c), the final residuum is a representative of the respective eye blink deflection instead of the monotonic signal, which ultimately can be used to reconstruct the eye blink component directly without the need to utilize other IMFs for the process.
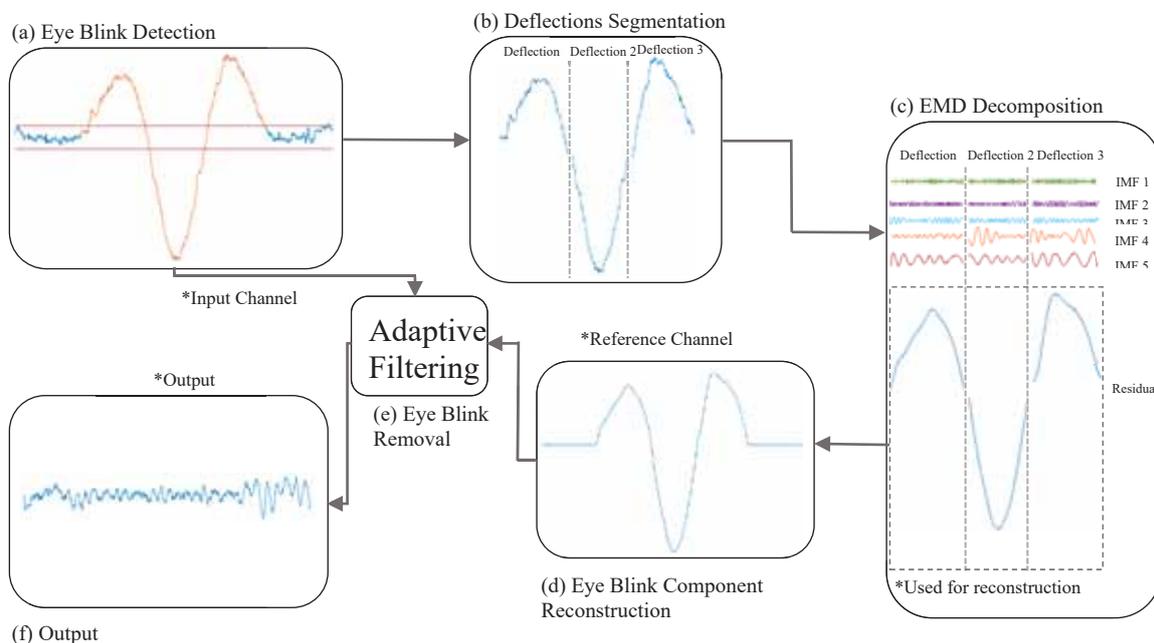
Fig. 6. (a) The identified eye blink region (orange line) is fed into EMD. (b) The eye blink is segregated into its individual deflection. (c) EMD decomposition. (d) Eye blink component extraction. (e) Eye blink artifact removal by ANC. (f) The ANC output is a cleaned signal.

The consistency of this occurrence is important so that only residuum is needed for the eye blink component reconstruction instead of the inclusion of a deliberate selection of suitable IMFs so that automatic eye blink component reconstruction is possible. In the proposed method, the residua are concatenated at their ends to form the respective eye blink component by means of spline interpolation, as visualized by Fig 6 (d). This reconstructed eye blink component has a similar deflection shape as the respective eye blink artifact contained in the EEG segment and thus can be utilized as a reference channel for adaptive filtering to remove the eye blink artifact from the EEG segment.

3.4 Eye Blink Artifact Removal

After the eye blink component is reconstructed, it is used as reference data $r_1(n)$, as shown in Fig. 7. The block diagram of the proposed adaptive filter that is used as adaptive noise cancellation (ANC) system is shown. The identified eye blink region is fed into the ANC as primary input $d(n)$. The RLS-based algorithm is used in the adaptive filter to remove the eye blink artifact from the marked EEG segment. For the respective EEG segments that correspond to the eye blink affected EEG region, their length span ranges from 0.25 to 1.5 seconds. Hence, δ value is set to be 0.99 and λ value to be 10 with a coefficient order of 2 to ensure efficient and stable impulse response convergence.
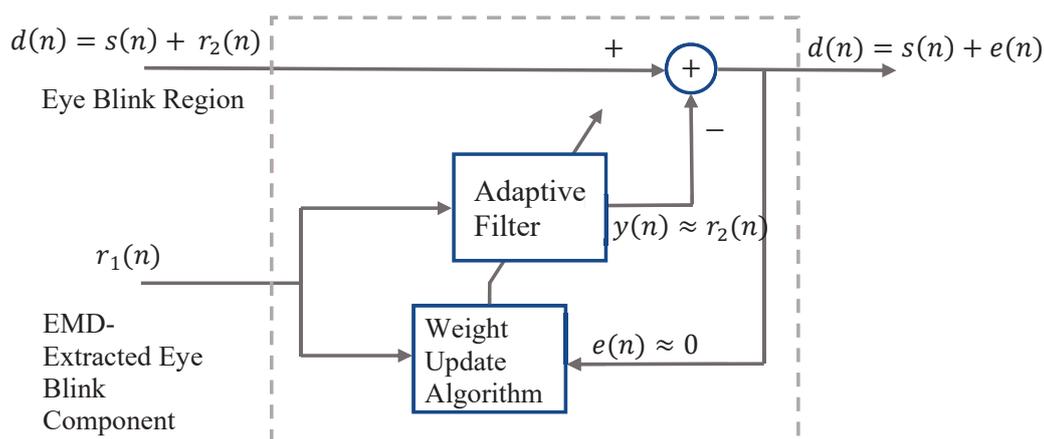
Fig. 7. ANC structure

3.5 Performance Metric

In order to measure the performance of the proposed method in removing the eye blink artifact, two data-centric measures are used on the processed EEG segments accordingly. The performance metrics are correlation coefficient (CC) and relative root-mean-square error (RMSE). The values generated from these performance metrics indicate how well the processed EEG data is preserved. The CC and RMSE are measured for the EEG component and eye blink component, respectively; between before and after being processed for eye blink artifact removal. The following are the mathematical descriptions of the performance metrics.

The CC is measured as the following:

$$CC_{EEG} = \frac{cov(Y_{pinknoise}, Y_{output})}{std(Y_{pinknoise}) * std(Y_{output})} \tag{15}$$

$$CC_{eye\,blink} = \frac{cov(Z_{simulated}, Z_{reconstructed})}{std(Z_{simulated}) * std(Z_{reconstructed})} \tag{16}$$

The RMSE is measured as the following:

$$RMSE_{EEG} = \sqrt{\frac{\sum_{t=1}^{n}(Y_{pinknoise}(t) - Y_{output}(t))^2}{n}} \tag{17}$$

$$RMSE_{eye\ blink} = \sqrt{\frac{\sum_{t=1}^{n}(Z_{simulated}(t) - Z_{reconstructed}(t))^2}{n}} \tag{18}$$

The variable $X(t)_{simulated}$ is the simulated eye blink contaminated-EEG, $Y_{pinknoise}$ is the simulated EEG component, Y_{output} is the cleaned simulated EEG component, $Z_{simulated}$ is the simulated eye blink component and $Z_{reconstructed}$ is reconstructed eye blink component.

4. RESULTS AND DISCUSSION

4.1 EEG Dataset Structure

The EEG dataset used to measure the performance metric is fully-simulated EEG data as described in Subsection 3.1. This is done to compare the performance metrics with the eye blink removal technique proposed by Egambaram et al. [19]. The generation of fully-simulated EEG for component Y(t), Z(t) and X(t) are iterated until 100 units are obtained. This total unit of EEG data is expedient to gauge the reliability of the performance metrics values in consistently proving the eye blink removal method of its level of performance.

4.2 Proposed Eye Blink Component Extraction

The extracted eye blink component is desired to be identical to its source. In this regard, a good eye blink component extraction is an eye blink component extraction with good precision of its features with the features of the original eye blink artifact. This is to ensure only the eye blink artifact is subtracted from the identified EEG segment during adaptive filtering. Generally, correct eye blink component extraction is also important for any rejection-based eye blink removal method, which includes the ICA-based method. Fig. 8 shows a sample of the reconstructed eye blink component (line in yellow colour) that being plotted in such a way that it can be compared with its respective eye blink reconstruction (line in blue colour). Two windows each magnifies a simulated eye blink region that originates from the main plot.
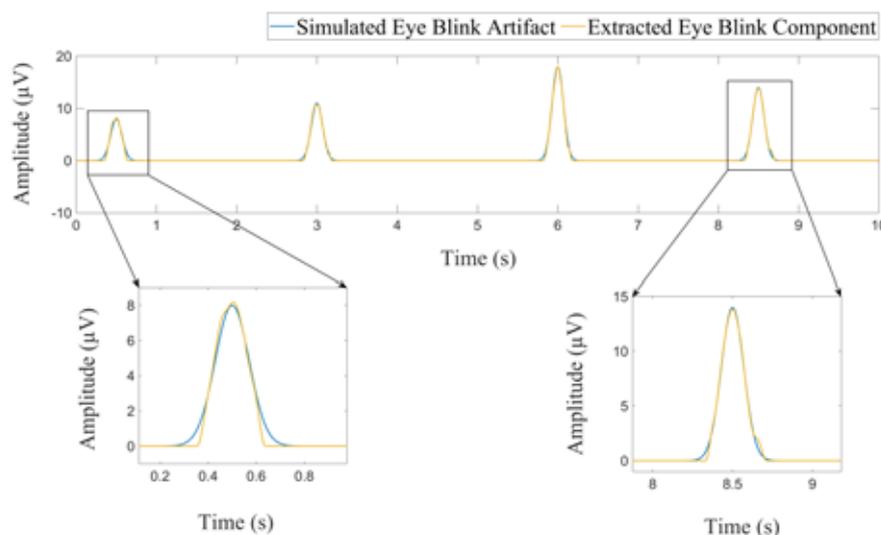

Fig. 8. The simulated eye blink component and its eye blink reconstruction.

The result shown in Table 1 contains the $CC_{eye\ blink}$ and $RMSE_{eye\ blink}$ values computed using the proposed eye blink component extraction method being juxtaposed with the respective performance metrics from the proposed FastEMD-CCA by Egambaram et al. The better value in comparison is bolded for every measurement.

Table 1: Performance metrics for eye blink component extraction between the proposed method and FastEMD-CCA.

Performance Metrics	EMD-AF (Proposed Method)	FastEMD-CCA
$CC_{eye\ blink}(\mu \pm \sigma)$	**0.9915 ± 0.0358**	0.9754 ± 0.0055
$CC_{eye\ blink}(95\%\ CI)$	**0.9880-0.9950**	0.9743-0.9765
$RMSE_{eye\ blink}(\mu \pm \sigma)$	**0.2960 ± 0.2998**	0.6580 ± 0.0776
$RMSE_{eye\ blink}(95\%\ CI)$	**0.2666-0.3253**	0.6426-0.6734

The result above shows that the eye blink reconstruction using EMD with spline interpolation resembles the original eye blink artifact better than the combination of EMD with CCA. Its CC value evidences this is statistically higher than FastEMD-CCA, while the proposed method's RMSE is statistically lower than that resulting from FastEMD-CCA. Hence, with appropriate energy thresholding for IMF derivation (qResol) and as decomposition stopping criterion (qResid), EMD is already substantial for excellent eye blink reconstruction. In the experimentation of implementing the EMD based on Rato et al. revision, α value is set to 0.5 for smooth but relatively efficient e(t) formation regulation, qResid is set to 55 for satisfactory IMF derivation and qResid is set to 57 for sufficient eye blink component extraction. Fig. 6 shows a sample of the simulated eye blink contaminated-EEG (line in orange colour) with its respective eye blink reconstruction (line in blue colour).

4.3 Proposed Eye Blink Removal

The extracted eye blink component is used to remove its eye blink source from its respective region. This can be done by means of component cancellation. In our proposed method, the extracted eye blink component is converged first by means of producing an appropriate impulse response using the RLS weight adaptation algorithm. The impulse response is then used to cancel out the eye blink artifact from the marked region by ANC. Fig. 9 displays the original simulated EEG component (line in red colour) being overlapped with its cleaned counterpart (line with yellow colour) to visualize any differences between them. The cleaned EEG is the output of the proposed method. Also, the eye blink-contaminated EEG (line with blue colour) is also plotted into the same graph to visualize the difference between it and the cleaned EEG.

In the following, the results on the performance metrics indicating the degree of possible change that may occur on the ground truth EEG are tabulated in Table 2. This degree of change is measured between the original simulated EEG data and the EEG data extracted after eye blink artifact removal performed by the proposed method. As for the metric, CC and RMSE are also used. The better value in comparison is bolded for every measurement. A higher CC value, closer to 1, or if it is 1 itself reveals that the cleaned segment is nearly or perfectly similar with its original counterparts. The lower value of RMSE points to a lower degree of difference between the cleaned segment and its original EEG component.
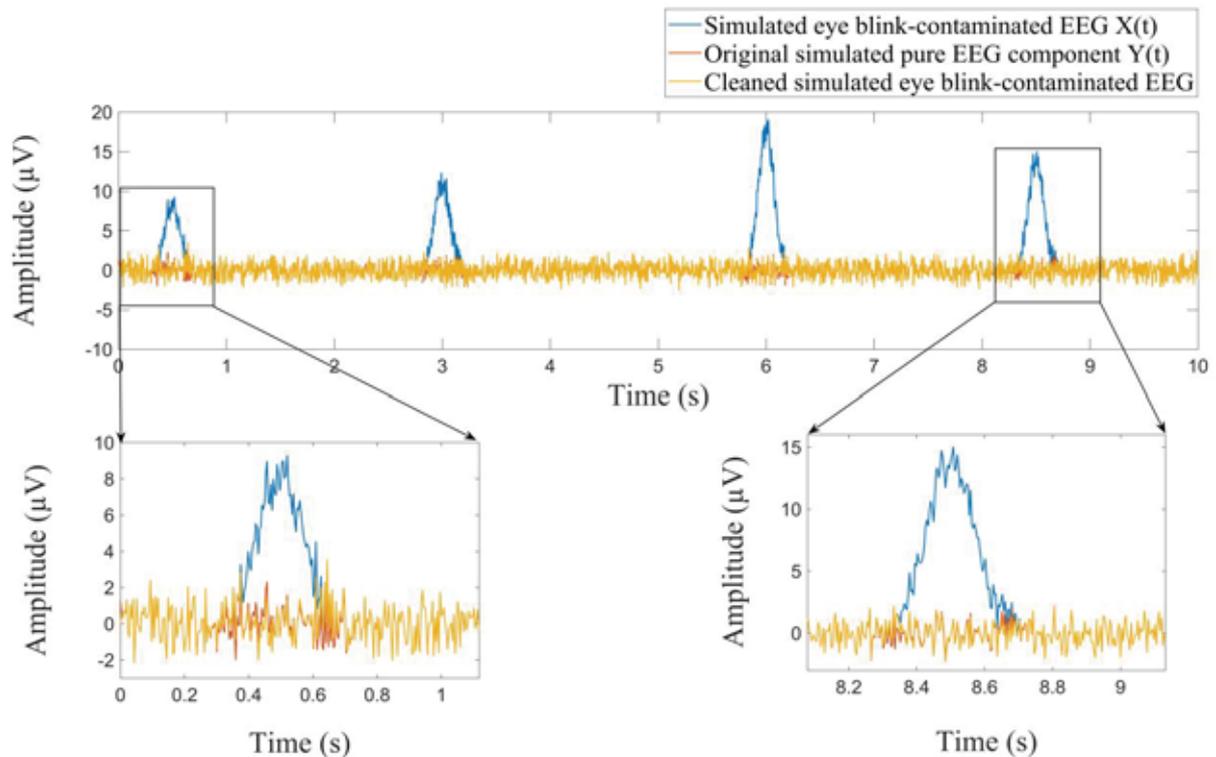
Fig. 9. The original simulated EEG data (line with blue colour) and its respective component retrieved after eye blink removal (line with yellow colour).

Table 2: Performance metrics for eye blink component removal between proposed method (EMD-AF) and FastEMD-CCA.

Performance Metrics	EMD-AF (Proposed Method)	FastEMD-CCA
$CC_{EEG}(\mu \pm \sigma)$	0.9430 ± 0.0839	0.7478 ± 0.0687
CC_{EEG}(95% CI)	0.9348-0.9513	0.7341-0.7614
$RMSE_{EEG}(\mu \pm \sigma)$	0.3211 ± 0.2738	0.6580 ± 0.0776
$RMSE_{EEG}$(95% CI)	0.2942-0.3479	0.6426-0.6734

The CC obtained from Egambaram et al. experimentation for the EEG component shows that its CC value dropped from the average of 0.9754, which is the CC obtained for the eye blink component, to 0.7478, which is the value that corresponds to the obtainment of ground truth EEG component after eye blink removal. This signifies that direct rejection of the eye blink artifact using its extracted eye blink component derived may cause distortion to the cleaned EEG segment. Therefore, it is substantial to introduce a mechanism that ensures the extracted eye blink component is converged to its eye blink artifact within the respective EEG segment in order to minimize the distortion that may be caused after the removal process. Adaptive filtering is serviceable on the matter of converging the extracted eye blink with the contained eye blink artifact. This is proven by the higher CC and lower RMSE value for the cleaned EEG segment by the adaptive filtering utilization in the proposed method.

FastEMD is meant not completely to decompose the EEG segment. Instead, it decomposes the respective EEG into a finite number of IMFs set beforehand, assuming that the first three IMFs are of EEG component and the remaining contains eye blink artifact. Subsequently, after the IMFs are divided into the non-artifactual and artifactual components,

they are added up to reconstruct into two components: non-artifactual and artifactual. The CCA is later used to find the correlation between the two components to separate the EEG component that remains within the artifactual components.

This paper intended to test the capability of EMD in extracting the eye blink component that is similar in its temporal feature without being complemented by other data processing techniques. Its heuristic approach in signal decomposition is the desirable quality that is expected to be harnessed in retrieving specific components. Therefore, with an appropriate workaround to the EMD, like energy thresholding, EMD alone is effective for eye blink component reconstruction that is handy for its subsequent removal.

5. CONCLUSION

In this paper, we proposed the five samples-window mean as the basis for amplitude thresholding that was used to identify the eye blink region in the eye blinks-contaminated single-channel EEG. We demonstrated that incorporation of energy thresholds with EMD decomposition to extract the eye blink deflections makes it possible to automate the eye blink reconstruction process. The identified eye blink region and extracted components are used as reference in RLS-ANC system to remove the eye blinks artifact in eye blink region while preserving the original EEG in eye blink-free region. The notable contribution of the proposed method is the utilization of the data-driven decomposition nature of EMD to extract the eye blink deflections that is monumental in the eye blink component reconstruction. The proposed method is evaluated using synthetic EEG datasets, and results show better performance compared to existing fastEMD-CCA technique. The results show that the proposed method was successful in extracting eye blink components with high similarity to its simulated original waveforms and the removal of eye blink using ANC was successful without altering or distorting the original EEG. The results also demonstrated that EEG components decomposition and separation using EMD is sufficient for eye blink reconstruction without relying on the multivariate data processing methods that is significantly higher in complexity. In this study, eye blink identification and removal process were performed offline. For real time implementation, the routine for accommodating delay must be designed and incorporated adequately in distributive manner across the EEG recording and the eye blink removal process because the EMD decomposition process is intrinsically sequential. Nevertheless, the proposed method is a completely univariate eye blink removal method that extracts the eye blink component in an empirical manner which complements the adaptive denoising process for effective eye blink removal in a single EEG channel.

ACKNOWLEDGEMENT

REFERENCES

[1] Artoni F, Fanciullacci C, Bertolucci F, Panarese A, Makeig S, Micera S, Chisari C. (2017). Unidirectional brain to muscle connectivity reveals motor cortex control of leg muscles during stereotyped walking. Neuroimage, 159:403–416. doi: 10.1016/j.neuroimage.2017.07.013.
[2] Mandekar S, Holland A, Thielen M, Behbahani M, Melnykowycz M (2022). Advancing towards Ubiquitous EEG, Correlation of In-Ear EEG with Forehead EEG. Sensors,

22(4):1568. https://doi.org/10.3390/s22041568

[3]   Lau-Zhu A, Lau MPH, McLoughlin G (2019). Mobile EEG in research on neurodevelopmental disorders: Opportunities and challenges. Dev Cogn Neurosci., 36:100635. doi: 10.1016/j.dcn.2019.100635.

[4]   Nagar P and Sethia D (2019). Brain Mapping Based Stress Identification Using Portable EEG Based Device. 11th International Conference on Communication Systems and Networks (COMSNETS), 601–606. doi: 10.1109/COMSNETS.2019.8711009.

[5]   Bhavnani S, Parameshwaran D, Sharma KK, Mukherjee D, Divan G, Patel V and Thiagarajan TC. (2022). The Acceptability, Feasibility, and Utility of Portable Electroencephalography to Study Resting-State Neurophysiology in Rural Communities. Front. Hum. Neurosci, 16:802764. doi: 10.3389/fnhum.2022.802764.

[6]   Sawangjai P, Hompoonsup S, Leelaarporn P, Kongwudhikunakorn S, and Wilaiprasitporn T. (2020). Consumer Grade EEG Measuring Sensors as Research Tools: A Review. IEEE Sensors Journal, 20(8):3996–4024. doi: 10.1109/JSEN.2019.2962874.

[7]   Shahbakhti M et al. (2021). VME-DWT: An Efficient Algorithm for Detection and Elimination of Eye Blink from Short Segments of Single EEG Channel. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 29:408–417. doi: 10.1109/TNSRE.2021.3054733.

[8]   Iwasaki M, Kellinghaus C, Alexopoulos A V, Burgess R C, Kumar A N, Han Y H, Lüders H O, and Leigh R J. (2005). Effects of eyelid closure, blinks, and eye movements on the electroencephalogram. Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology, 116(4), 878–885. https://doi.org/10.1016/j.clinph.2004.11.001.

[9]   Kanoga S, Kanemura A, and Asoh H. (2019). Multi-scale dictionary learning for ocular artifact reduction from single-channel electroencephalograms. Neurocomputing, 347:240–250. doi: 10.1016/j.neucom.2019.02.060.

[10]   Agarwal M and Sivakumar R. (2019). Blink: A Fully Automated Unsupervised Algorithm for Eye-Blink Detection in EEG Signals. 57th Annual Allerton Conference on Communication, Control, and Computing, Allerton, 1113–1121. doi: 10.1109/ALLERTON.2019.8919795.

[11]   G Bhaskar N Rao, Anumala V S, Peri P P and Aneesh S. (2020). Automatic Detection and Correction of Blink Artifacts in Single Channel EEG Signals. International Journal of Advanced Computer Science and Applications (IJACSA), 11(1). http://dx.doi.org/10.14569/IJACSA.2020.0110144

[12]   Jefri L A M, Rahman F A, Malik N A, and Isa F N M. (2022). Eye blink region detection in EEG using amplitude thresholding. 8th International Conference on Mechatronics Engineering (ICOM 2022), 77–83. doi: 10.1049/icp.2022.2269.

[13]   Maddirala A K and Veluvolu K C. (2022). SSA with CWT and k-Means for Eye-Blink Artifact Removal from Single-Channel EEG Signals. Sensors, 22(3). doi: 10.3390/s22030931.

[14]   Maddirala A K and Veluvolu K C. (2021). Eye-blink artifact removal from single channel EEG with k-means and SSA. Sci Rep., 11(1). doi: 10.1038/s41598-021-90437-7.

[15]   Zeiler A, Faltermeier R, Keck I R, Tomé A M, Puntonet C G, and Lang E W. (2010). Empirical mode decomposition - An introduction. Proceedings of the International Joint Conference on Neural Networks, 2010. doi: 10.1109/IJCNN.2010.5596829.

[16]   Shahbakhti M, Bavi M and Eslamizadeh M. (2013). Elimination of Blink from EEG By Adaptive Filtering without Using Artifact Reference. 2013 4th International Conference on Intelligent Systems, Modelling and Simulation, 190-194, doi: 10.1109/ISMS.2013.106.

[17]   Patel R, Janawadkar M P R, Sengottuvel S, Gireesan K, and Radhakrishnan T S. (2016). Suppression of Eye-Blink Associated Artifact Using Single Channel EEG Data by Combining Cross-Correlation with Empirical Mode Decomposition. IEEE Sens J., 16(18):6947–6954. doi: 10.1109/JSEN.2016.2591580.

[18]   Molla M K I, Rabiul Islam M, Tanaka T, and Rutkowski T M. (2012). Artifact suppression from EEG signals using data adaptive time domain filtering. Neurocomputing, 97:297–308.

doi: 10.1016/j.neucom.2012.05.009.

[19] Egambaram A, Badruddin N, Asirvadam V S, Begum T, Fauvet E, and Stolz C. (2020). FastEMD–CCA algorithm for unsupervised and fast removal of eyeblink artifacts from electroencephalogram. Biomed Signal Process Control, 57:2020. doi: 10.1016/j.bspc.2019.101692.

[20] Rato R T, Ortigueira M D, and Batista A G. (2008). On the HHT, its problems, and some solutions. Mech Syst Signal Process, 22(6):1374–1394. doi: 10.1016/j.ymssp.2007.11.028.

MODIFIED SEIRD MODEL: A NOVEL SYSTEM DYNAMICS APPROACH IN MODELLING THE SPREAD OF COVID-19 IN MALAYSIA DURING THE PRE-VACCINATION PERIOD

**Norsyahidah Zulkarnain, Nurul Farahain Mohammad
and Ibrahim Shogar***

*Department of Computational and Theoretical Sciences, Kulliyyah of Science,
International Islamic University Malaysia,
Jalan Sultan Ahmad Shah, 25200 Kuantan, Pahang, Malaysia*

*Corresponding author: shogar@iium.edu.my

ABSTRACT: Mathematical modelling is an effective tool for understanding the complex structures and behaviors of natural phenomena, such as coronavirus disease 2019 (COVID-19), which is an infectious disease caused by a life-threatening virus called SARS-CoV-2. It has rapidly spread across the world in the last three years, including Malaysia. Adopting a novel system dynamics approach, this paper aims to explain how mathematics can play a significant role in modelling the COVID-19 spread and suggests practical methods for controlling it. It forecasts the data of infected (I), recovered (R) and death (D) cases for decision-making. This paper proposes a modified Susceptible-Exposed-Infected-Recovered-Death (SEIRD) model with time-varying parameters considering the sporadic cases, the reinfection cases, the implementation of a movement control order, and the percentage of humans abiding by the rules to forecast future growth patterns of COVID-19 in Malaysia and to study the effects of the consideration on the number of forecasted COVID-19 cases, during the pre-vaccination period. This study implemented the preliminary stage of forecasting the COVID-19 data using the proposed SEIRD model and highlighted the importance of parameter optimization. The mathematical model is solved numerically using built-in Python function *'odeint'* from the Scipy library, which by default uses LSODA algorithm from the Fortran library Odepack that adopts the integration method of non-stiff Adams and stiff Backward Differentiation (BDF) with automatic stiffness detection and switching. This paper suggests that the effects of factors of sporadic cases, reinfection cases, government intervention of movement control order and population behavior are important to be studied through mathematical modelling as it helps in understanding the more complex behavior of COVID-19 transmission dynamics in Malaysia and further helps in decision-making.

ABSTRAK: Pemodelan matematik adalah alat berkesan bagi memahami struktur kompleks dan tingkah laku fenomena semula jadi, seperti penyakit coronavirus 2019 (COVID-19), iaitu penyakit berjangkit yang disebabkan oleh virus pengancam nyawa yang dipanggil SARS-CoV-2. Ia telah merebak dengan pantas ke seluruh dunia sejak tiga tahun lepas, termasuk Malaysia. Mengguna pakai pendekatan baharu sistem dinamik, kajian ini bertujuan bagi menerangkan bagaimana matematik boleh memainkan peranan penting dalam membentuk model penyebaran COVID-19, dan mencadangkan kaedah praktikal bagi mengawalnya. Model ini dapat meramalkan data sebenar kes yang dijangkiti, pulih dan kematian bagi membuat keputusan. Kajian ini mencadangkan model populasi Rentan-Terdedah-Terjangkiti-Pulih-Mati (SEIRD) yang diubah suai bersama parameter masa berbeza seperti kes sporadis, kes jangkitan semula, pelaksanaan perintah

kawalan pergerakan, dan peratusan manusia patuh peraturan bagi meramal pertumbuhan corak kes COVID-19 di Malaysia pada masa hadapan dan mengkaji kesan–kesan pertimbangan parameter tersebut ke atas bilangan kes COVID-19 yang diramalkan ketika tempoh sebelum vaksinasi. Kajian ini melaksanakan peringkat awal ramalan data COVID-19 menggunakan model SEIRD yang dicadangkan dan menekankan kepentingan pengoptimuman parameter. Model matematik ini diselesaikan secara berangka menggunakan fungsi terbina Python *'odeint'* daripada perpustakaan Scipy, yang menggunakan algoritma LSODA daripada perpustakaan Fortran Odepack menerusi kaedah penyepaduan Adams tidak kaku dan Pembezaan Belakang (BDF) kaku dengan pengesanan dan pertukaran kekakuan automatik. Kajian ini mencadangkan kesan faktor kes sporadis, kes jangkitan semula, campur tangan kerajaan terhadap perintah kawalan pergerakan dan tingkah laku penduduk adalah penting untuk dikaji melalui pemodelan matematik kerana ia membantu dalam memahami tingkah laku yang lebih kompleks dalam dinamik penularan COVID-19 di Malaysia dan seterusnya membantu dalam membuat keputusan.

KEY WORDS: *SEIRD model; system dynamics; systems thinking approach; COVID-19; simulation; Malaysia*

1. INTRODUCTION

Coronavirus disease 2019 is an infectious disease caused by a newly discovered coronavirus called SARS-CoV-2, which is a type of virus known to cause respiratory infections in humans [1]. SARS-CoV-2 has been identified and characterized as 75 to 80% identical to the SARS-CoV which refers to the virus that emerged during the SARS outbreak of 2002 and 2003 that was first recognized in Guangdong Province, China [2][3]. Furthermore, it is found to be categorized within the subgenus *Sarbecovirus* of the genus *Betacoronavirus* that is even more closely related to two bat-derived SARS-CoV which are bat-SL-CoVZC45 and bat-SL-CoVZXC21 [4].

Besides the general characteristics of SARS-CoV-2, some common clinical symptoms of the patients are fever, dry cough, fatigue, and gradual dyspnea (which refers to shortness of breath) [5]. Further development of the disease can lead to complications including pneumonia, acute respiratory distress syndrome (ARDS), septic shock, and kidney failure that eventually may cause death [6]. Although, COVID-19 is believed to be a disease that primarily affects the lungs, it can also cause damage to many other organs, including the heart, kidneys, and the brain as long-term effects when the disease is not treated well [7]. COVID-19 not only affects human health, but also the economy, which decreases the performance and profit of many industries. For instance, in Malaysia, the aviation and tourism industry were affected as 67.8% stated that they had no sales and revenue within the period of the movement control order, MCO [8].

The Malaysian Government, through the Ministry of Health, Malaysia (MOH) responded to this outbreak. Several phases of movement restriction order and standard operating procedure (SOP) were introduced to break the chain of COVID-19 spread among citizens, resulting in a significant drop in the daily data of COVID-19 cases which ended the second wave of the pandemic on 8th July 2020 [9]. However, the third wave which was more critical and challenging with the emergence of new variants of SARS-COV-2 caused sporadic cases [10] and reinfection cases [11].

According to World Health Organization (WHO) [12], dynamics of the SARS-CoV-2 are very complex, with the virus changing itself over time as it replicated through random "copying error" in a process called mutation. The variants were labelled in a simple and

easy way, using letters of the Greek alphabet which were categorized into two types of variants. The first type is the Variants of Interest (VOIs) which included Lambda and Mu, and the second type was Variants of Concern (VOCs) which included Alpha, Beta, Gamma, and Delta. Different variants may have their own properties with respect to ease of spread, the associated disease severity, or the performance of vaccines, as well as the possibility of reinfection [12]. This highlights the complexity of the transmission dynamics of the virus.

This paper explains the transmission dynamics of COVID-19 in Malaysia using compartmental models from the basic to the more complex as it helps to divide the population in Malaysia with regards to the daily data of COVID-19 cases based on each compartment. This type of mathematical model may help to reliably forecast the number of infected, recovered, and death cases in Malaysia. Thus, a System Dynamics (SD) approach through mathematical modelling can play an important role in understanding the behavior of this complex health phenomena in real life, which refers to the COVID-19 pandemic as evolving epidemiology, which may guide informed prevention and control policies.

The main purpose of this study is to modify the classical SEIRD model by considering sporadic cases, reinfection cases, government interventions of movement control order and population behavior. In this paper, we firstly (1) present the classical compartmental models of infectious disease: SI, SIR, SEIR and SEIRD and highlight its limitation, then we show how we (2) construct a new modified SEIRD model considering the aforementioned factors based on a System Dynamics (SD) approach to deal with all those limitations. This study provides a detailed explanation of how mathematics plays a significant role in modelling the COVID-19 spread and surely can contribute to understanding the transmission dynamics of COVID-19 in Malaysia and create awareness among society to apply future intervention measures in curbing COVID-19 spread.

2. MATHEMATICAL MODELLING

Mathematical modelling is an effective tool to understand the complex systems of the real world. According to Lich [13], "systems science" is a broad term referring to a family of analytical approaches that aims to explain the behavior of complex systems of the real phenomenon that occurs. Its three important methodologies are Social Network Analysis (SNA), Systems Dynamics (SD) and Agent-based Modeling (ABM). In Systems Dynamics (SD) approach, is called as 'Hard' systems approach which refers to the quantitative systems dynamic modelling that enables scientists and decision makers to examine real life system components, and the dynamic relationships between them. Generally, SEIRD model is one of the SD approaches and it is one of such methods of modelling for infectious disease. This model consists of susceptible (S), exposed (E), infected (I), recovered (R) and death (D) and is called a compartmental model since the model can be broken down into distinct compartments, and may describe the interactions between those compartments. Basically, this type of model can be constructed in terms of ordinary differential equations (ODEs) that can be interpreted as the rate of change in the amount of the substance in the compartment referring to the input rate minus output rate for each dynamic or process as mentioned in Eq. (1) and the schematic representation of a one-compartment system as in Fig. 1.

$$\frac{dx}{dt} = \text{input rate} - \text{output rate} \qquad (1)$$
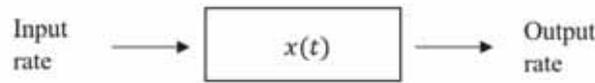
Fig. 1: Schematic diagram of one-compartment system.

This section consists of two subsections which are the classical compartmental model: SI, SIR, SEIR, SEIRD and the modified SEIRD model that we proposed.

2.1 Classical Compartmental Model: SI, SIR, SEIR, SEIRD

The origin of this deterministic epidemiological model was started from basic classical compartmental models of infectious disease which are the SI model, SIR model, and SEIR model. It was proposed by Kermack & McKendrick [14] in 1927 to study the number and distribution of cases of an infectious disease as it is transmitted through a population over time. The SI model is a basic model, and other stated models are derived models that were built according to research needs. SEIRD is the extended model of the SIR and SEIR models. The SIR model divides the population into three groups: susceptible (S), infected (I) and recovered (R). In understanding the transmission dynamic of infectious disease or how the real phenomenon, which refers to the disease spread, the schematic diagram of SIR model is framed as Fig. 2.

Fig. 2: Schematic diagram of SIR model

The SIR model is constructed in terms of non-linear differential equations as shown in Eq. (2) to Eq. (4) where $N=S+I+R$ is the total population, β is the infection rate, a coefficient accounting for the susceptible people who get infected by infectious people and γ is the parameter of infectious people which become resistant per unit of time (immune to COVID-19).

$$\frac{dS}{dt} = \frac{\beta}{N} SI \tag{2}$$

$$\frac{dI}{dt} = \frac{\beta}{N} SI - \gamma I \tag{3}$$

$$\frac{dI}{dt} = \frac{\beta}{N} SI - \gamma I \tag{4}$$

Figure 3 shows a more extended model of SIR, which is the SEIR model, where it considered a latent period, a new compartment E representing the exposed individuals that are in the incubation period. It is added between the compartment of susceptible (S) and infected (I). Many diseases, especially infectious disease, have a latent phase that refers to the very beginning part of the disease progression where the individual has been exposed or had contact with an infected (I) person yet remain asymptomatic and are typically regarded as not infectious.
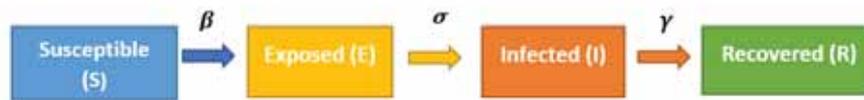
Fig. 3: Schematic diagram of SEIR model.

This delay between the acquisition of infection and the infectious state can be incorporated within the SIR model by adding an exposed (*E*) population, and letting individuals who have contact with infected (*I*) but are not yet infectious move from *S* to *E* and from *E* to *I* once they have tested positive accompanied by illness symptoms. In a closed population, assuming no births or natural deaths, the SEIR model can be constructed in terms of ordinary differential equations (ODE) as per Eq. (5) to Eq. (8) where is the total population, α represents the incubation rate. The difference between the exposed (*E*) and infected (*I*) is that the former has contacted the infected (*I*) person and are asymptomatic but not infectious, while the latter is symptomatic and infectious. SEIR has been used to model breakouts in Malaysia as a preliminary study.

$$\frac{dS}{dt} = \frac{\beta}{N} SI \tag{5}$$

$$\frac{dE}{dt} = \frac{\beta}{N} SI - \sigma E \tag{6}$$

$$\frac{dI}{dt} = \sigma E - \gamma I \tag{7}$$

$$\frac{dR}{dt} = \gamma I \tag{8}$$

A further model, the SEIRD, considers the group of Dead (*D*) for the forecast of the spread of COVID-19 has been proposed and adopted by many other epidemiological researchers [1]. The SEIRD model accounts for five different groups, namely, susceptible (*S*), exposed (*E*), infectious (*I*), recovered (*R*), and dead (*D*). Figure 4 shows how the population in each compartment during the pandemic progress in sequence.
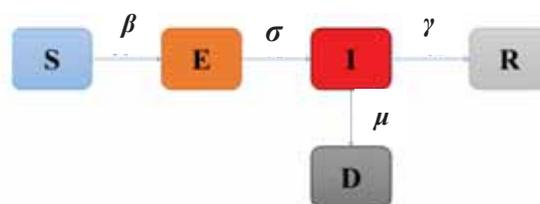
Fig. 4: Schematic Diagram of Basic SEIRD Model [18].

In their study, Shao & Shan [18] highlighted four conversions or transmission dynamics in the basic SEIRD model:

- **First, *S* → *E***. Relevant evidence shows that during the COVID-19 infection process, not only confirmed patients have infectious capacity, but also those who are asymptomatic. They have the ability to infect others as well without anyone knowing it. It can be considered as a sporadic case. Therefore, in the SEIRD model, the

individual with susceptible (S) status will change to exposed (E) status with a certain probability after contacting the infected (I) individual or the exposed (E) individual.

- **Second, $E \rightarrow I$**. Relevant evidence shows that the longest incubation period for COVID-19 is 14 days and the shortest is 1 day. An individual with exposed (E) status may be transformed into an infected (I) after the incubation period ends.

- **Third, $I \rightarrow R$**. Once the infected (I) individual has been confirmed, they will be isolated and treated with a certain probability in hospital and will change to the recovered (R) or death (D) status.

- **Fourth, $I \rightarrow D$**. Relevant evidence also shows that infected patients commonly die after 15 days without effective treatment.

It can be written in terms of systems of non-linear ordinary differential equations as per Eq. (9) to Eq. (13).

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \tag{9}$$

$$\frac{dE}{dt} = \frac{\beta SI}{N} - \sigma E \tag{10}$$

$$\frac{dI}{dt} = \sigma E - \gamma I - \mu I \tag{11}$$

$$\frac{dR}{dt} = \gamma I \tag{12}$$

$$\frac{dD}{dt} = \mu I \tag{13}$$

All variables related to the basic SEIRD model are defined as Table 1 while epidemiological parameters involved are as stated in Table 2.

Table 1: Variables of SEIRD model

Variables	Descriptions
Susceptible (S)	Population who can have the disease/ are vulnerable to COVID-19
Exposed (E)	Population who are in incubation period, asymptomatic but have the ability to spread the virus and are about to get infected by the disease (asymptomatic infective)
Infectious (I)	Population who already have the disease and can spread it/ symptomatic infected
Recovered (R)	Population who has been cured and immunized from the disease for a certain time
Death (D)	Population who has died because of the disease

Table 2: Epidemiological Parameters of SEIRD model

Symbols	Parameters	Descriptions
β	Infection rate	Probability accounting for the Susceptible get infected by the Infected
σ	Incubation rate	Rate of latent individuals becoming infectious (probability of susceptible people becoming exposed), 1/average latency or incubation.
γ	Recovery rate	Probability infectious people who become resistant per unit time, 1/average recovery time or the period between onset of symptom and recovered
μ	Mortality rate	Probability infectious people who died due to COVID-19 per unit time

Sporadic cases occur due to contact between susceptible (S) and exposed (E), Muka & Sannyal [10] later consider two types of infectious rate in their proposed SEIRD model, as shown in Fig. 5 below.
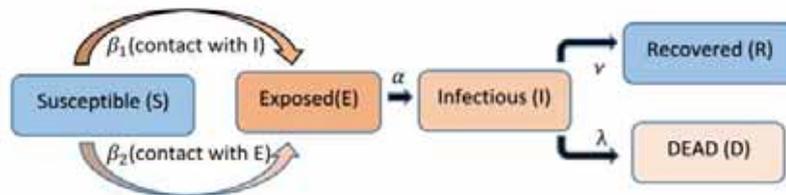
Fig. 5: Schematic diagram of basic SEIRD model considering two types of infectious rate that are constant [9].

The constructed SEIRD model may be expressed in terms of the following non-linear ordinary differential equations, as in Eq. (14) to Eq. (18), where the total population is $N = S + E + I + R + D$. The infectious rates, β_1 and β_2, control the rate of tranmission for contact with infected (I) and exposed (E) respectively. In this model, β_1 represents the probability of infection per exposure when a susceptible (S) individual has contact with an infected (I) patient and becomes a latent exposed (E) individual. While β_2 represents the potential rate per exposure when a susceptible (S) individual has mutual contact with an exposed (E) individual and becomes a latent exposed (E) individual. A detailed diagram is shown in Fig. 5. Since the probable contact between susceptible (S) and exposed (E) individuals are larger than that of between susceptible and infected individuals, Muka & Sannyal [10] assumed that $\beta_2 = 5\beta_1$. The incubation rate, α, is the rate of latent individuals becoming infectious.

$$\frac{dS}{dt} = -\frac{1}{N}(\beta_1 I S) - \frac{1}{N}(\beta_2 E S) \tag{14}$$

$$\frac{dE}{dt} = \frac{1}{N}(\beta_1 I S) + \frac{1}{N}(\beta_2 E S) - \sigma E \tag{15}$$

$$\frac{dI}{dt} = \sigma E - \gamma I - \lambda I \tag{16}$$

$$\frac{dR}{dt} = \gamma I \tag{17}$$

$$\frac{dD}{dt} = \lambda I \tag{18}$$

Moreover, Piccolomiini & Zama [15] have also proposed an adaptive SEIRD model which is a SEIRD rational model or can be simply named as SEIRD (rm) that considers several restricting measures imposed in Italy. Since there are such considerations, the whole-time interval [0, T] has been partitioned into two sub-intervals: [0, t_0] and [t_0, T] where t_0 corresponds to time when the restrictions start to produce a valuable change. Applied restrictions should decrease the number of contacts between I and S, consequently, decrease the value of transmission rate β. Thus, they model coefficients of β as a decreasing time-varying function $\beta(t)$ as follows:

$$\beta(t) = \begin{cases} \beta_0 & if\ t < t_0 \\ \beta_0(1 - \rho\ \frac{t - t_0}{t})) & otherwise \end{cases}, \quad \rho \in (0,1) \tag{19}$$

In their present, the constant value of $\rho = 0.75$ has been used but might be calibrated in the future. By substituting $\beta(t)$ in the S and E equations in basic SEIRD model, the SEIRD rational model, SEIRD (rm), is obtained. There is much limitation that can be considered and assumptions that can be reduced to model the real phenomenon of the systems specifically the systems that refer to the transmission dynamics of COVID-19.

2.2 Modified SEIRD model (Proposed Model)

In this study, the basic epidemiological model, Susceptible-Exposed-Infected-Recovered-Death model (SEIRD), will be modified considering four main factors: sporadic cases, reinfection cases, government intervention of movement control order, and population behavior. These will be implemented as time-varying parameters in this model, since a few of the epidemiological parameters, especially transmission rate that refers to the infection rate β, are believed to naturally change over time and are not constant. This is due to the government interventions taken to curb the spread of COVID-19 in Malaysia. Generally, once government interventions are taken to restrict the movement of the population, contact between them can be reduced, which may consequently reduce the value of the infection rate β.

The SEIRD model is a deterministic compartmental model that consists of five different compartments [15]. It is represented by a set of non-linear ordinary differential equations (ODE) that describe how the system evolves in time. The total population of Malaysia in 2020, which was $N \sim 32657300$ individuals as reported by Department of Statistic Malaysia [19], are divided into five different compartments: susceptible (S), exposed (E), infected (I), recovered (R) and death (D). The word 'compartments' signifies the division of the population into mutually exclusive groups.

Susceptible (S) refers to the population who can be infected by the disease. Exposed (E) refers to the population who are in the incubation period where they are not yet showing any symptoms, however they are silently infectious and called asymptomatic patients. Contact with exposed (E) may be addressed as sporadic cases. Infected (I) refers to the population who already have the disease by showing symptoms and can infect others who have close contact with them called symptomatic patients. Recovered (R) refers to the

population who has been cured of the disease, while death (D) refers to the population who died because of COVID-19.

Several assumptions have been made to model the reliable transmission dynamics of COVID-19 during the pre-vaccination period due to the limited data. The assumptions are as follows:

1. Malaysia population was a closed population due to the implementation of international travel restrictions enforced on 25th January 2020 which limits the movement of travelling foreigners into Malaysia, $N = S + E + I + R + D$.

2. There has yet to be an immunization for SARS-CoV-2 since the period of this study is during pre-vaccination. Thus, all of the Malaysian population regardless of gender and age is susceptible to COVID-19.

3. The population in Malaysia was assumed to be unchanged due to the short time of the model development and projection. Newborn and natural death were not considered. These were negligible because the outstanding period of the disease is shorter than the lifetime of a human.

4. There was a chance for the recovered (R) population to be reinfected again.

5. The sporadic case is now considered where we assume the disease can also probably be spread by exposed (E) individuals.

Assumption 4 and assumption 5 in this study will extend the model-making assumption that fits well with the real phenomenon of the spread of COVID-19. Considering different transmission dynamics in the community and the assumptions made above, Fig. 6 shows the schematic diagram that portrays how the population in each compartment progresses in sequence with a certain transition rate that considers reinfection cases. Referring to Fig. 6, the *black line arrow* and *black arrow curved down* indicate how the population progresses in sequence from one compartment to another subsequent compartment with certain probabilities, while the *yellow arrow curved down* shows the interaction or contact between everyone in stated compartments.
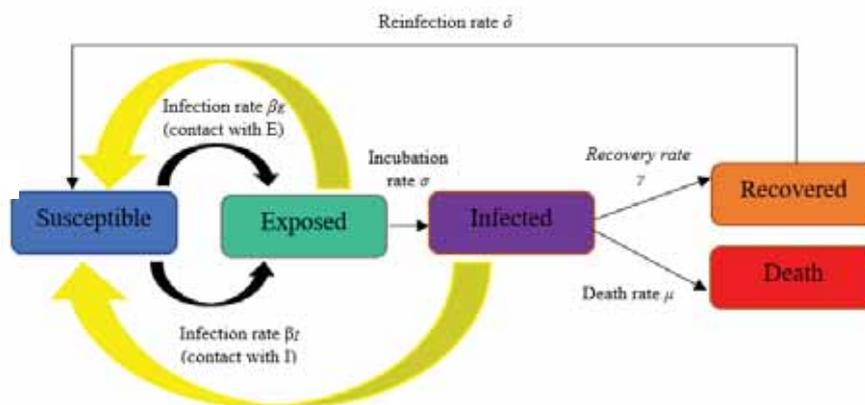
Fig. 6: Schematic diagram of modified SEIRD model.

In addition, the recovery rate γ is the probability of the infected population to become resistant, recovered, and immunized from COVID-19, which will progress the population to the recovered population, where $\gamma = \frac{1}{T_i}$ with T_i as the average of recovery duration [6] and can be called the infectious period as well [22]. Mortality rate μ may also be determined

using this formula ($\mu = \frac{D}{N}$) over a certain period where D refers to the cumulative death up to a certain date and N refers to the total population in the country. Moreover, the reinfection rate δ can easily be calculated using the simple formula of $\delta = \frac{Re}{I}$ where R_e refers to number of reinfected while I refers to number of infected persons involved in the study. Malhotra [11] highlighted that fully vaccinated health care workers in India had lower risk of reinfection compared to unvaccinated and partially vaccinated with the percentage of 1.6%, 12.7% and 11% respectively. However, this study will neglect the vaccination status of individuals and will use the aforementioned formula for the reinfection rate δ.

By following the assumptions, the modified SEIRD model that has been proposed by Jamil & Muhammad [23] that considers reinfection cases and modified SEIRD model that has been proposed by Muka & Sannyal [10] that consider sporadic cases which highlighted on two types of infection rates can be used for constructing a new form of modified model in this study. This modified model is based on five compartments in terms of a system of the non-linear ordinary differential equations (ODEs) as shown in Eq. (20) to Eq. (25) with the initial conditions of $E_0 = 3375$, $I_0 = 1$, $R_0 = 22$, $D_0 = 0$.

$$\frac{dS(t)}{dt} = -\frac{\beta_I(t)\,S(t)I(t)}{N} - \frac{\beta_E(t)\,S(t)E(t)}{N} + \delta R(t) \tag{20}$$

$$\frac{dE(t)}{dt} = \frac{\beta_I(t)\,S(t)I(t)}{N} + \frac{\beta_E(t)\,S(t)E(t)}{N} - \sigma E(t) \tag{21}$$

$$\frac{dI(t)}{dt} = \sigma E(t) - \gamma I(t) - \mu I(t) \tag{22}$$

$$\frac{dR(t)}{dt} = \gamma I(t) - \delta R(t) \tag{23}$$

$$\frac{dD(t)}{dt} = \mu I(t) \tag{24}$$

$$N = S(t) + E(t) + I(t) + R(t) + D(t) \tag{25}$$

Equations (20) to (25) are the rate of change in the amount of population in each compartment. Firstly, $\frac{dS}{dt}$ refers to the rate of change of population in the susceptible compartment. The input rate refers to δR as recovered (R) will be leaving recovered (R) compartment and move into the susceptible (S) compartment again once reinfected with SARS-CoV-2. Meanwhile, $-\frac{\beta(t)\,S(t)I(t)}{N} - \frac{\beta(t)\,S(t)E(t)}{N}$ is the output rate that refers to susceptible (S) individuals that will be leaving susceptible (S) compartment. It is formulated as aforementioned to represents the fraction of population that can be infected by SARS-COV-2 and move to exposed (E) compartment. In Eq. (20) as well, the susceptible (S) population is proportional to the infected (I) population since, in reality, they will have contact with each other, especially close contact within 1 meter. It is inversely proportional. The negative sign in the formula of the total population of susceptible (S) individuals refers to the total susceptible population that will be decreasing when the susceptible (S) move to exposed (E). In addition, when reinfection cases have been considered, the number of recovered (R) individuals will be added to the susceptible (S) compartment with a certain value of reinfection rate.

Secondly, $\frac{dE}{dt}$ refers to the rate of change of the population in exposed (E) compartment where exposed move to infected. The number of individuals who are leaving the susceptible (S) compartment is added into the Eq. (21). As time goes by, the exposed (E) population will later be leaving the exposed (E) compartments and become infected (I). That is why there is subtraction of σE. Thirdly, $\frac{dI}{dt}$ refers to the rate of change in infected (I) compartment where infected move to recovered. The number of individuals who are leaving exposed (E) compartment is added into the Eq. (22). Infected will be recovered or die as time goes by so they will be leaving infected (I) compartment. That is the meaning of $-\gamma I(t) - \mu I(t)$ in the Eq. (22).

Fourthly, $\frac{dR}{dt}$ refers to the rate of change of the population in the recovered (R) compartment when the patients have been isolated and treated in the hospital. The number of individuals who leave the infected (I) compartment is added into the Eq. (23) which is represented by $\gamma I(t)$ and reinfected population that will be leaving the recovered (R) compartment will be added as well into the Eq. (23) with $-\delta R(t)$. Lastly, $\frac{dD}{dt}$ refers to the rate of change of the population in death (D) compartment where infected (I) move to death (D) when there is no effective treatment taken to the COVID-19 patients. Similarly, the number of individuals who are leaving infected (I) compartment is added into the Eqn. (24) which is represented by $\mu I(t)$.

In this study, the time-varying parameters for infection rate β_I and infection rate β_E will be implemented and highlighted since transmission dynamics of COVID-19 especially infection rate change over time. Thus, time-varying infection rate β_I and infection rate β_E were formulated as piecewise functions, while recovery rate, mortality rate, incubation rate and reinfection rate were inferred from literature review. The formulated time-varying infection rate as proposed by Jamil & Muhammad [23] was shown in Eq. (26) and in this study it was referred to as the infection rate β_I. While for infection rate β_E, this study adopts a highlight by Muka & Sannyal [10] where they assumed that $\beta_2 = 5\beta_1$, in which it can be restated as $\beta_E = 5\beta_I$ and the time-varying infection rate is formulated as Eq. (27). Equation (26) considers government intervention such as the movement control order that was implemented from 18th March 2020 until 3rd May 2020 where the time interval of this piecewise function was divided by three phases as follows:

Phase I　　: Before Movement Control Order (MCO), t < $t_{lockdown}$
(27/2/20-17/3/20)

Phase II　: During Movement Control Order (MCO) and Conditional Movement Control Order (CMCO), $t_{lockdown} \leq$ t < t_{lift} **(18/3/20-9/6/20)**

Phase III　: During Recovery Movement Control Order (RMCO), t $\geq t_{lift}$ **(10/6/20-23/2/21)**

$$\beta_I(t) = \begin{cases} \beta_1 t + \beta_2, & t < t_{lockdown} \\ \beta_0 e^{-((t-t_{lockdown})/(\tau_\beta))}, & t_{lockdown} \leq t < t_{lift} \\ (1-r)(\beta_1(t - t_{lift}) + \beta_2), & t \geq t_{lift} \end{cases} \quad (26)$$

$$\beta_E(t) = 5 \begin{cases} \beta_1 t + \beta_2, & t < t_{lockdown} \\ \beta_0 e^{-((t-t_{lockdown})/(\tau_\beta))}, & t_{lockdown} \leq t < t_{lift} \\ (1-r)(\beta_1(t - t_{lift}) + \beta_2), & t \geq t_{lift} \end{cases} \quad (27)$$

There were 9 unknown parameters involved in this modified SEIRD model proposed by Jamil & Muhammad [23] which were the percentage of Malaysians who followed the SOPs r, infection rate (β_0, β_1, β_2), characteristic time of transmission τ_β, incubation rate σ, recovery rate γ, death rate μ, and reinfection rate δ. All the initial parameter values will be inferred from several items in the literature, then will be implemented in the simulation of modified SEIRD model in order to study the transmission dynamics of COVID-19 in Malaysia.

For understanding the formulated time-varying infection rate β, firstly, at the beginning of the outbreak which is Phase I: Before Movement Control Order (MCO), people had high mobility and were free to move anywhere, thus the infection rate $\beta(t)$ was assumed to be a linear function, $\beta_1 t + \beta_2$ as it increased with β_2 as its initial value, when $t=0$, $\beta(t) = \beta_2$. For Phase II, when the Movement Control Order (MCO) was introduced in Malaysia, the infection rate decayed due to physical distancing and intentional isolation, and this behavior was described by an exponential function, $\beta_0 e^{-((t-t_{lockdown})/(\tau_\beta))}$ with β_0 as the initial value of the infection rate during that phase and $1/(\tau_\beta)$ refers to the decay rate. The Conditional Movement Control Order (CMCO) will be regarded as Phase II, which is lockdown or restricted movement even though more economic and social activities were allowed but operating hours were limited and occurred under strict standard operating procedures (SOP).

Lastly, for Phase III, when the lockdown was lifted and more economic and social activities were allowed, for instance interstate travel, tourism business, unessential premises or even schools, the infection rate was assumed to follow the trend at the beginning of the outbreak which increased when no lockdown or stringent movement order was implemented. As highlighted by Jamil & Muhammad [23], a fraction of compliance to the SOP was included in the infection rate as it affected or contributed to the new values of infection rate and was based on the experience in facing the pandemic, $(1 - r)(\beta_1(t - t_{lift}) + \beta_2)$ where r was the percentage of Malaysians who followed the SOPs and practiced the 3Ws, even after the government had lifted the lockdown. The numeric value of r was between 0 and 1 as it will be regarded as percentage values. $1 - r$ here refers to the percentage of Malaysians who did not follow the SOPs, which led to the increment in the infection rate β. Meanwhile, the rest of the function is linear.

Different researchers used different settings for epidemiological models in their studies, including different methods of parameter estimation or even different functions of infection rate. This highlights the varied insights that researchers do have in modelling the transmission dynamics of COVID-19. This also becomes a reason behind the existence of new extended and modified models in modeling any behavior, process, or system in a real phenomenon. The simulation of the above proposed modified SEIRD model will be presented in section 4.

3. MATERIALS AND METHODS

In this section, we present the source of our data and details of parameters and variables that we used for simulation of classical and modified SEIRD model.

3.1 Data Sources

This study is a simulation for general cases in Malaysia during the pre-vaccination period; thus, daily data of confirmed new infected (I), recovered (R) and death (D) cases of the COVID-19 during the second wave and early third wave, which are from 27th February 2020 until 23rd February 2021, are collected from the official data of the Ministry of Health Malaysia (MOH) [24]. The data of COVID-19 were publicly provided in the GitHub of the Ministry. The data was collected and sorted in an Excel file before being imported into the Pandas Data Frame for the purpose of Python simulation. For the simulation of the SEIRD model, the data of daily total active cases of infected (I), total recovered (R) and total death (D) were used.

3.2 Data Analysis

In this section, the initial values of variables and parameters used for analysis in this study are presented and summarized in Table 3 and Table 4.

Table 3: Initial values of variables used modified SEIRD model simulation

Variables	Values used for analysis	Sources
Susceptible (S)	$S(0) = 32{,}653902$	[19]
Exposed (E)	$E(0) = 3375$	[9]
Infected (I)	$I(0) = 1$	[23]
Recovered (R)	$R(0) = 22$	[23]
Death (D)	$D(0) = 0$	[23]

Table 4: Initial values of parameters from recent findings for modified SEIRD model simulation

Parameters	Values used for analysis	Sources
β = infection rate, $R_t\gamma$	$R_0 = 2.52$ (R_t at 18/3/2020 or MCO), $\beta_0 = 0.70308$	[25]
	$R_1 = 0.1$ (expected increment daily or gradient), $\beta_1 = 0.0279$	
	$R_2 = 3.91$ (R_t at 27/2/2020 or pre-MCO, $\beta_2 = 1.09$	
σ = incubation rate, $\frac{1}{T_c}$	$T_c = 5.2$ days, $\sigma = 0.19$	[21]
γ = recovery rate, $\frac{1}{T_i}$	$T_i = 3.95$ days	[26][27]
μ = mortality rate, $\frac{D}{N}$	$\gamma = 0.279$ $\mu = 0.1$	[24]
δ = reinfection rate, $\frac{R_E}{I}$	$\delta = 0.02$	[28]
r = percentage obey SOP	95%	Assumption made in this study
τ_β = characteristics time of transmission	1	Assumption made in this study

As reported by the Department of Statistics Malaysia [19], the total population of Malaysia as of early April 2020 was approximately 32.7 million, $N \sim 32{,}657300$. Considering the assumption (2) and Eq. (25), the initial value of susceptible S was 32,653902. Due to the limited data of daily exposed, we assume the initial exposed (E) is

3,375 which is equal to the total number of positive COVID-19 cases in the Sri Petaling Mosque gathering cluster well known as the Cluster Tabligh. This huge cluster of COVID-19 in the second wave period was announced by the Director-General of Ministry of Health Malaysia (MOH), Datuk Dr Noor Hisham Abdullah on 8th July 2020. The gathering among tabligh was held from 27th February 2020 until 1st March 2020. It is possible to assume the total positive cases in this cluster was equal to the initial value of exposed (E) since the gathering was first held on 27th February 2020, with the understanding that the exposed population would progress to the infected population. The initial value of infected (I), recovered (R) and death (D) are 1, 22 and 0 respectively.

4. RESULTS AND DISCUSSIONS

In this section, we firstly present the historical data of Malaysian COVID-19 cases during the pre-vaccination period which is from the beginning of second wave up to the early third wave as an overview before we show the results of classical SEIRD model. Simulations of classical and modified SEIRD model in this study have been done using Python programming language via Jupyter Notebook and using the values of parameters and variables as summarized in Table 3 and Table 4. The SciPy *odeint* function is used to solved the ODEs of both classical and modified SEIRD model. Before proposing a new modified SEIRD model, the output of classical SEIRD model that we adopted in this study will be compared with other published studies that implemented classical SEIR and SEIRD models as this model's validation process. Then, the results of the simulation of our new proposed model of SEIRD is presented and the four effects that are taken into consideration in modelling the SEIRD model are highlighted.

4.1 Historical Data of Malaysian COVID-19 Cases
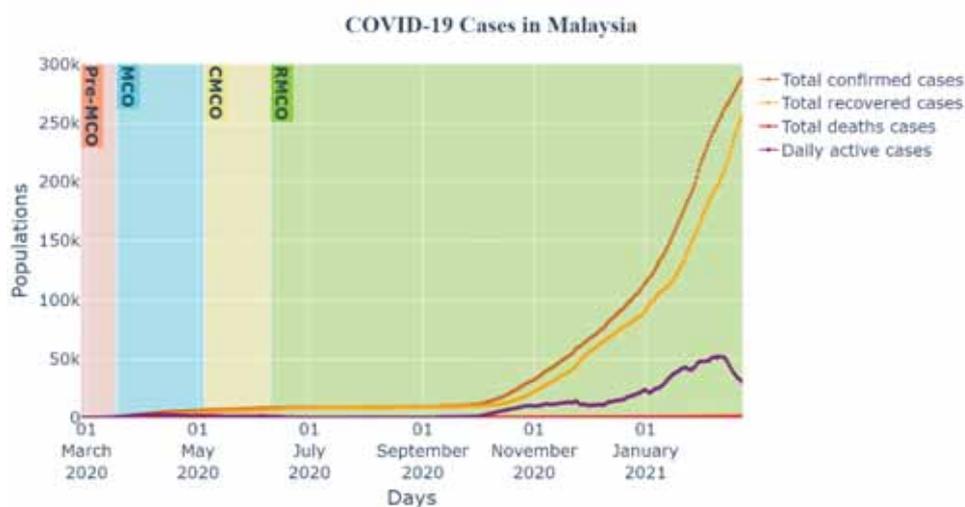
Fig. 7: Total confirmed cases, total recovered cases, total deaths cases and daily active cases of COVID-19 in Malaysia during pre-vaccination period from 27th February 2020 until 23rd February 2021.

Figure 7 exhibits the historical data of Malaysian COVID-19 cases starting from pre-MCO (27/2/2020), which is the beginning of second wave of COVID-19 in Malaysia, until the period of RMCO (23/2/2021) which is at the middle of third wave. They are divided into four different shaded regions according to their respective phases which are pre-MCO phase (non-intervention phase), Movement Control Order (MCO), Conditional Movement Control Order (CMCO), and Recovery Movement Control Order (RMCO). Based on the reported

data [3], the peak infection that is represented by the daily active cases during the second wave was depicted on 5th April 2020 with 2587 infected (*I*) cases before following with a steady decrement in the curve. Meanwhile, the peak infection during the third wave was depicted 6 months after the date of peak infection of the second wave which is 10th February 2021 with 52,187 infected (*I*) cases. Figure 7 was generated using Python programming language via Jupyter Notebook.

4.2 Classical SEIRD Model Simulation

Using the variables and parameters as detailed in Table 3 and Table 4, the classical SEIRD model simulation based on Eq. (9) to Eq. (13) was produced, as seen in Fig. 8. The figure portrays the simulation of the classical SEIRD model when no consideration of sporadic cases, reinfection cases, government interventions of movement control order and population behavior. This classical SEIRD model uses a constant value of infection rate β because the model assumes that the dynamics of COVID-19 spread, which were represented by epidemiological parameters, were not changing over time. This assumption was made because the government interventions of movement control order was not considered. Thus, the value of the infection rate used is constant. Based on the forecasted data in Fig. 8, the peak infection of COVID-19 is estimated to reach 3.05 million infected (*I*) cases, or about 9.34% of the Malaysian population, on 25 days of simulation from 27th February 2020 which is on 23rd April 2020. The classical SEIRD model forecasts the peak infection as high as the stated numbers, which was larger than the actual data (2587 infected (*I*) cases on 5th April 2020). This huge predicted peak infection is consistent with the higher values of infection rate β before the implementation of MCO that has been used in the classical SEIRD model simulation. Besides, according to the graph shown in Fig. 8, the number of infected (*I*) cases would gradually subside and plateau on 24th May 2020 before finally ending by day 23rd August 2020. Meanwhile, the cumulative number of recovered (*R*) and death (*D*) cases in Malaysia are forecasted to reach 22.39 million and 8.02 million respectively by day 200. This classical SEIRD model uses the basic systems of non-linear ordinary differential equations that can be improved more according to new assumptions or more factors that have been considered. Thus, before proposing a new modified SEIRD model, model reliability validation has been done and will be discussed in the next subsection.

Fig. 8: Forecast of COVID-19 pandemic in Malaysia using classical SEIRD model for 200 days during pre-vaccination period with constant parameters, $\beta = 1.09$.

4.3 Model Validation

In order for this study to propose a new modified SEIRD model, model reliability validation has been done by comparing the simulation's output of published literature in Malaysia which adopted similar and almost-similar classical SEIRD models to the one in

this study. The similar classical SEIRD model refers to the same systems of ordinary differential equations implemented in the study. Since there are few studies that adopted the classical SEIRD model 100% in terms of Eq. (9) to Eq. (13), we found that it was reasonable to also include another version of the classical SEIRD model (a bit modified) and the classical SEIR model that excludes the death (D) compartment, which is almost similar to our classical SEIRD model in the validation. It was reasonable to use the output to calculate the Mean Absolute Percentage Error (MAPE) to bring insight to determine the model performance. The lower the MAPE, the higher the accuracy of the model. The three studies included in the comparison consist of two studies that adopted the classical SEIR model and one study which adopted another version of the classical SEIRD model.

The percentage of peak infection is one of the crucial outputs in infectious disease modelling as it is the possible maximum total number of populations that can be infected during the simulation period. Thus, the output has been compared in terms of peak infection (%) as summarized in Table 5. In making the comparison with published works, firstly our classical SEIRD model is used to simulate the peak infection using the similar values of parameters and initial conditions, as the said literature shows. Then, the absolute percentage error (APE) between simulated peak infected cases by literature and simulated peak infected cases by this study's model is calculated, as seen in Table 6. Mean Absolute Error (MAE) is calculated to be 7.49% while the accuracy of the classical SEIRD model performance, regarded as the benchmark before proposing this new modified model, is 92.51%. This study will only simulate the model without performing any parameter fitting; thus, this kind of validation method helps in determining if either the existing classical SEIRD model, that acts as a benchmark, or the proposed modified model are sufficiently reliable to proceed.

Table 5: Comparison of the simulated peak infection by using different settings of epidemiological parameters for classical SEIR and SEIRD model by published literature in Malaysia

Author	Country	Model	Period of simulation (Days)	Simulated peak infection date (From t_0)	Simulated peak number of infected cases	Peak infection (%)
Aidalina & Lim [29]	Malaysia	Basic SEIR *Not consider Infected becoming dead and not consider Death cases (no D compartment)*	100	$t_0=$ 18/3/2020. 5/4/2020 (18 days)	13.76×10^6	43
Labadin & Hong [30]	Malaysia	Basic SEIR *Consider Infected becoming dead but not consider Death cases (no D compartment)*	150	$t_0=$ 27/2/2020. 15/5//2020 (78 days)	4×10^6	12.27
Alsayed et al. [31]	Malaysia	Basic SEIRD *same with classical SEIRD model in this study but a bit different in its equations (dE/dt and dS/dt)*	365	$t_0=$ 24/1/2020. 26/7/2020 (184 days)	2.5×10^6	7.668

Table 6: Comparison of the simulated peak infection between published literature
and the simulated peak infection by our classical SEIRD model

Author	Peak infection simulated by published literature (%)	Peak infection simulated by the classical SEIRD model (%)	Absolute percentage error (%)
Aidalina & Lim [29]	43	43.75	1.74
Labadin & Hong [30]	12.27	12.27	0
Alsayed et al. [31]	7.67	9.26	20.73

Table 7: Mean absolute percentage error (MAPE) (%) and accuracy (%)

Mean Absolute Percentage Error (MAPE) (%)	Accuracy (%)
7.49	92.51

4.4 Modified SEIRD Model Simulation

In this study, we propose a new SEIRD model that has been constructed in terms of systems of non-linear ordinary differential equations (ODEs) as in Eq. (20) to Eq. (24). This proposed model has been considered novel for four basic factors in modelling the behavior of COVID-19 spread in Malaysia, which are consideration of 1) sporadic cases, 2) reinfection cases, 3) government interventions of movement control order, and 4) population behavior. It is believed that this new model would provide a better understanding of the complex structures and behaviors of COVID-19, and finally assist decision makers in planning a strategy for ending the pandemic. Using the variables and parameters as detailed in Table 3 and Table 4, the modified SEIRD model simulation based on Eq. (20) to Eq. (24) was produced. This figure portrays the simulation of modified SEIRD model when the above factors are considered. These factors highlight the complexity of virus behavior and human behavior as discussed in section 2. This modified SEIRD model uses time-varying value of infection rate as per Eq. (25) to Eq. (26) because the model assumes that the dynamics of COVID-19 spread, which is represented by epidemiological parameters, are changing over time. This assumption was made because the government interventions of movement control order were considered. Thus, the value of infection rate varied over time. That was the reason for the implementation of time-varying parameters as a recent simulation of modified SEIRD model should use parameters that are dependent on time.

Based on the forecasted data in Fig. 9, the peak infection of COVID-19 is estimated to reach 8 million infected cases, which is about 24.5% of the Malaysian population, on 6 days of simulation from 27th February 2020 which is 3rd March 2020. The proposed SEIRD model forecast the peak infection earlier than the classical model as it also considers the sporadic cases, which is the spread of COVID-19 by asymptomatic patients, which now involves two types of infection rate β_I and β_E in the modelling. This early prediction of peak infection is consistent with the understanding of the effects of sporadic cases which may contribute to higher infected daily cases. Simply, it is due to the implementation of time-varying parameters of infection rates and modification on the modified mathematical model. Besides that, according to the Fig. 9, the number of infected (I) cases would gradually subside and becomes plateaued on 5th June 2020. Meanwhile, the total active recovered cases and the number of death cases of Malaysian was forecasted to reach 8 million and 13 million

respectively by 13th September 2020. In this modified SEIRD model, the total active recovered cases, represented by recovered (R), must be determined when simulation of modelling with parameter fitting has been done. It is in contrast with the classical SEIRD model, as it will take the daily cumulative number of recovered (R) cases. Table 8 shows the summary of peak infection, the plateau, and the subsiding date from both epidemic models.

Fig. 9: Forecast of COVID-19 pandemic in Malaysia for 200 days during pre-vaccination period using modified SEIRD model with time-varying parameters of β_I and β_E .

Table 8: Comparison of predicted date of infected cases between classical and modified SEIRD model

Infected cases (I)	Actual data	Classical SEIRD model	Modified SEIRD model
Peak infection	5th April 2020	23rd April 2020	3rd March 2020
Subside	21th June 2020	24th May 2020	1st May 2020
Plateau	8th July 2020	23rd August 2020	5th June 2020

4.5 Effects of Considered Factors to Modified SEIRD Model

The four factors that have been considered in this study bring effects to the simulation of SEIRD model: either changing the number of cases, the pattern of the simulation, or the date of peak infection. The effects of these factors are crucial to highlight in this study as the simulation of modified SEIRD model will forecast the future COVID-19 cases and mimic the real phenomenon when the effects have been taken into consideration. The effects of each factor can be seen in Fig. 10 – Fig. 13. From these simulations, we may observe how the factors may have influenced the number of infected (I), recovered (R) and death (D) cases of COVID-19 pandemic in Malaysia. Figure 10 depicts the simulation of modified SEIRD when considering the factor of sporadic cases. This factor considers the transmission of COVID-19 when S having contact with E which completely introduce new type of infection rate β_E. Thus, in simulating the modified SEIRD model, there are two types of infection rate, which are infection rate β_I (S contact with I becoming E) and infection rate β_E (S contact with E becoming E). When simulating the modified SEIRD model, only changes of β_E are taken. Meanwhile, other parameters and variables remain the same. This is similar to the other factors when simulation has been done. Based on Fig. 10, it clearly shows that as β_E increases, the peak infected, peak recovered, and peak death increase. The

peak infected will become higher and faster when there are sporadic cases reported in Malaysia, as indicated by Fig. 10. These results highlight the crucial impact that this factor brings to the output of the SEIRD model simulation. In the real scenario of the COVID-19 pandemic, when sporadic cases started to be reported in Malaysia and when there was no quick government intervention implemented, a higher than usual daily total active cases were reported, which also causes the time of peak infection to be reached earlier. This modified SEIRD model simulation brings important insights to the government authorities and other decision makers to take action earlier if such cases happen, either to implement a movement control order or introduce a vaccination program to the community in order to ensure that the Malaysian healthcare system will not collapse. The simulated number of peak infected, peak recovered, and peak death are the crucial indications for the government to decide which intervention should be implemented and when to implement it.

Fig. 10: Forecast of COVID-19 pandemic in Malaysia for 50 days during pre-vaccination period as β_E increases (β_E refers to factor 1: sporadic cases).

Furthermore, Fig. 11 depicts the simulations when considering the factor of reinfection cases, Fig. 12 depicts consideration of the factor of government interventions of movement control order, and Fig. 13 depicts consideration of population behavior. When the factor of reinfection cases has been considered, δR is a part of the mathematical model. Meanwhile, when the factor of government interventions of movement control order has been considered, decay exponential function of infection rate β_I and β_E need to be implemented in the simulation of the SEIRD model. r in the function of time-varying parameters of infection rate represents the Malaysian population behavior in abiding standard operating procedures (SOP) while ($1-r$) refers to the population who neglect the rules.
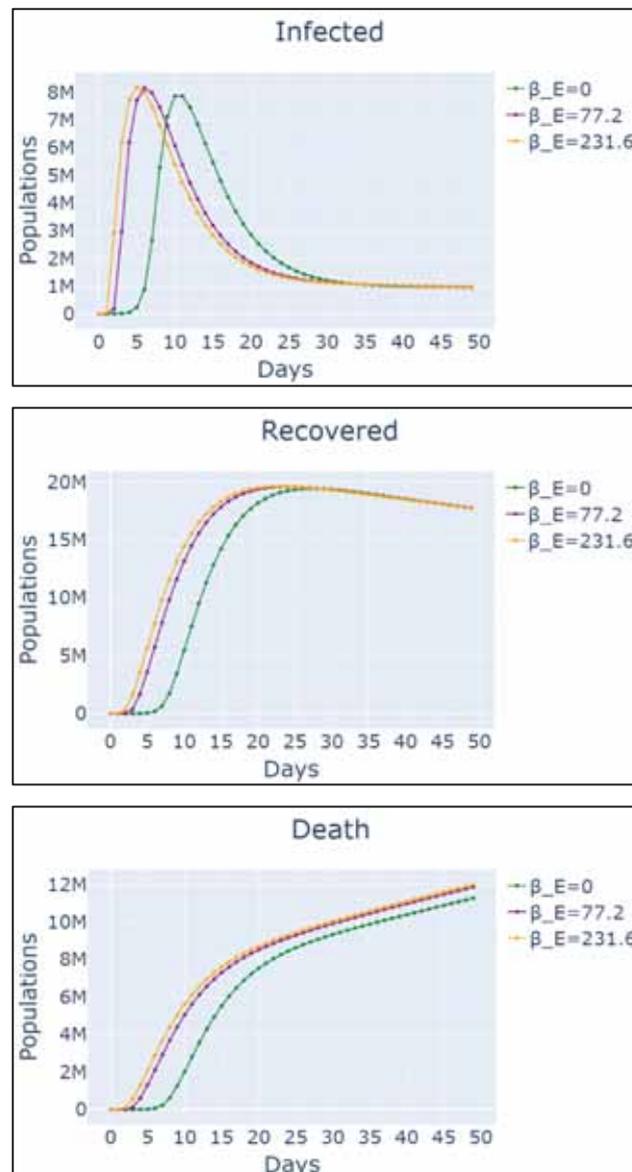

Fig. 11: Forecast of COVID-19 pandemic in Malaysia for 50 days during pre-vaccination period as δ increases (δ refers to factor 2: reinfection cases).

Based on Fig. 11, it clearly portrays the effects of reinfection cases as the value of δ increases. When δ increases, the duration of infected cases subsides and reaches a plateau. This means that the pandemic of COVID-19 in Malaysia will require a longer time to resolve when there are reinfection cases reported in Malaysia and will cause new higher peak infected if there is no government intervention taken. Figures 12 and 13 depict the effects of human behaviors towards COVID-19 cases, which are the action taken by government in implementing MCO and action taken by population in abiding by the rules and regulations or Standard Operating Procedures (SOP). Figure 12 shows significant effects on the reduction of the number of peak infected cases, peak recovered cases, and peak death cases when there is implementation of Movement Control Order (MCO) by government.

Fig. 12: Forecast of COVID-19 pandemic in Malaysia for 50 days during pre-vaccination period as β_I and β_E decrease (β_I and β_E decrease refer to factor 3: government intervention of MCO).

Figure 13 shows the effects of population behavior towards the COVID-19 reported cases which are infected (I), recovered (R), and death (D), as the value of r decreases. The figure highlights that the higher the percentage of population who abide by the rules (higher r), the lesser the number of peak infected, peak recovered, and peak death cases. This will surely bring a positive impact if compared with lower percentage of the population who abide by the rules. We set the value of r to be 99%, 60% and 20% just to show the percentage of their compliance. 99% means strong compliance, 60% means moderate compliance and 20% means weak compliance. The figure shows the importance of strong compliance by the Malaysian population after lifting the MCO, as it may delay the increase of infected and thereby decrease the peak case level. This will surely help Malaysia to recover slowly either in economical or health aspects.
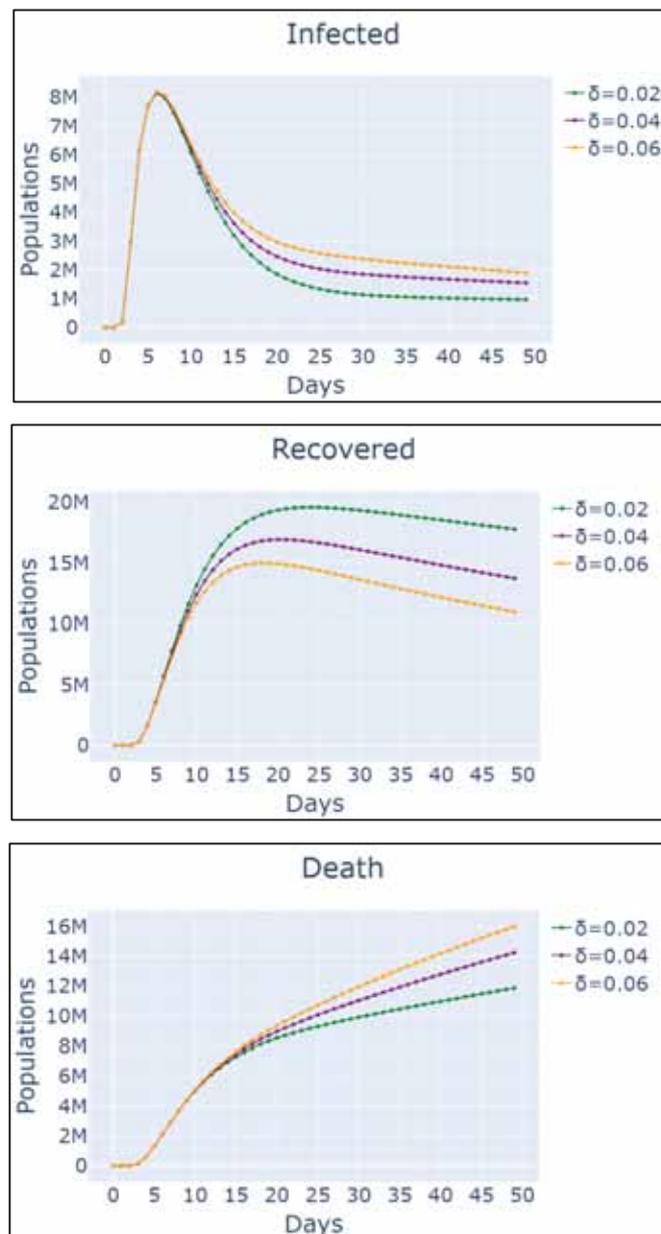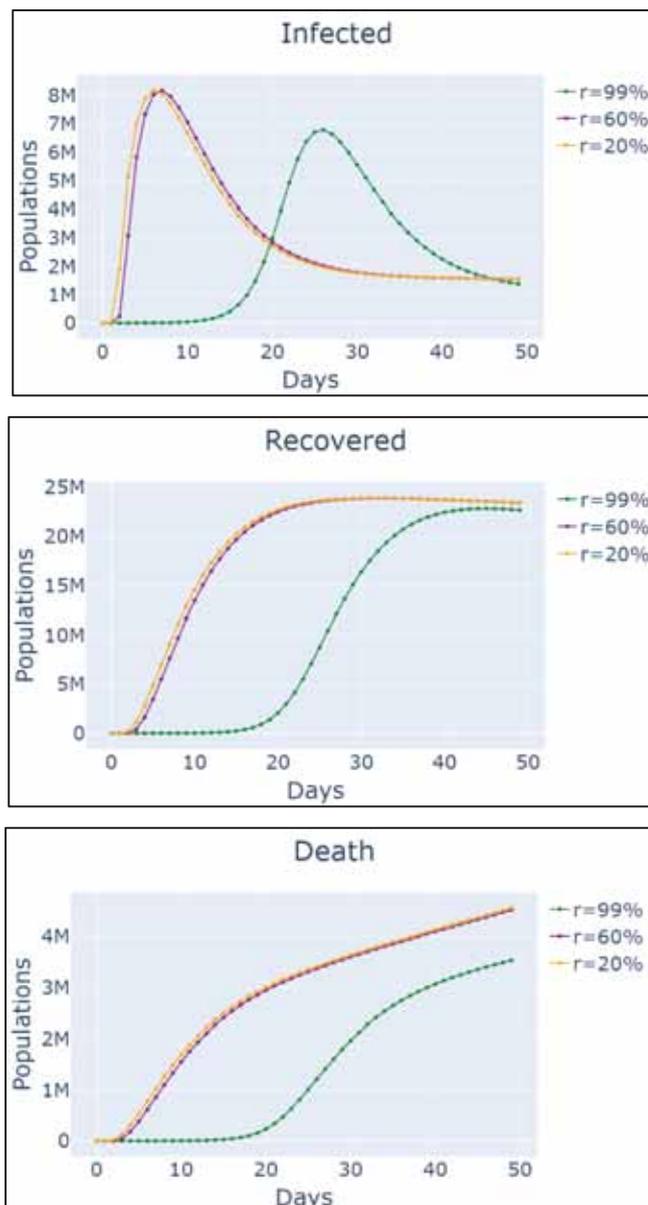

Fig. 13: Forecast of COVID-19 pandemic in Malaysia for 50 days during pre-vaccination period as r increases (r refers to factor 4: population behavior).

5. CONCLUSION

This study explained a system dynamics approach in modelling a modified SEIRD model considering sporadic cases, reinfection cases, government interventions of movement control order and population behavior in Malaysia, as these are found to be important to be studied on the effects of the factors towards the simulation of COVID-19 spread in Malaysia during pre-vaccination period. This study implemented the preliminary stage of forecasting the COVID-19 data as this study has only simulated the SEIRD model without parameter fitting (optimization) which needed for a reliable forecasting of future COVID-19 cases with a minimized error of simulation. Based on subsection 4.4 of this study, the number of infected cases at peak infection in our proposed modified SEIRD model are overestimated to the million cases as parameters optimization is not done in this study. This highlights the importance of parameter optimization of every model simulation done and included in our proposed SEIRD model. Despite the limitation of parameter optimization in this study, the findings of this study highlight the effects of the complexity of virus behavior and human behavior. This study suggests that the factors of sporadic cases, reinfection cases, government intervention of movement control order, and population behavior are important to consider in modelling the more complex behavior of COVID-19 transmission dynamics as it will helps mimic the real phenomenon and the simulation can give insights for government or researchers to take a decision in the best time to implement the rules and regulations effectively in curbing the COVID-19 spread in Malaysia. In combating the COVID-19 pandemic, two things that must be considered: virus behavior and human behavior. SARS-CoV-2 behavior is very complex as it mutates over time which creates new variants of SARS-CoV-2 that led to the emergence of new COVID-19 waves. Creating awareness among community, rules and regulations, movement control order (MCO), play a crucial role to curb the pandemic and reduce the transmission of COVID-19 since rules will be abided and interaction between humans will be limited.

ACKNOWLEDGMENT

REFERENCES

[1] Wenbao W, Yiqin C, Qi W, Ping C, Ye H, Shanwen H, Yan W, Zuxiong H, Wenxiang W. (2020) Transmission dynamics of SARS-COV-2 in China: Impact of public health interventions. MedRxiv. https://doi.org/10.1101/2020.03.24.20036285

[2] Zheng J. (2020) SARS-CoV-2: An Emerging Coronavirus that Causes a Global Threat. International Journal of Biological Science, 16(10): 1678-1685. https://doi.org/10.7150/ijbs.45053

[3] Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, Tan W. (2020) A Novel Coronavirus from Patients with Pneumonia in China, 2019. The New England Journal of Medicine, 382(8): 727-733. https://doi.org/10.1056/NEJMoa2001017

[4] Xu X, Chen P, Wang J, Feng J, Zhou H, Li X, Zhong W, Hao P. (2020) Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. Science China Life Sciences, 63(3): 457-460. https://doi.org/10.1007/s11427-020-1637-5

[5] Huang C, Wang Y, Li X, Rem L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B. (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet, 395(10223): 497-506. https://doi.org/10.1016/S0140-6736(20)30183-5

[6] Kraemer MU, Yang CH, Gutierrez B, Wu CH, Klein B, Pigott DM, Plessis L, Faria NR, Li R, Hanage WP, Brownstein JS, Layan M, Vespignani A, Tian H, Dye C, Pybus OG, Scarpino, SV. (2020) The effect of human mobility and control measures on the COVID-19 epidemic in China. Science, 368(6490): 493-497. https://doi.org/10.1126/science.abb4218

[7] Staff MC. (2022, June 28). COVID-19 (coronavirus): Long-term effects. Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/coronavirus/in-depth/coronavirus-long-term-effects/art-20490351

[8] Report of Special Survey on Effects of COVID-19 on Companies and Business Firms (Round 1). [https://www.dosm.gov.my/v1/uploads/files/covid-19/Report_of_Special_Survey_COVID-19_Company-Round-1.pdf]

[9] BERNAMA. (2020, July 8). COVID-19: Health Ministry declares end of Malaysia's biggest cluster. BERNAMA.com. https://www.bernama.com/en/general/news_covid-19.php?id=1858578

[10] Sannyal M, Mukaddes AMM. (2022) Transmission parameters of coronavirus disease 2019 in South Asian countries. Osong Public Health Res Perspect, 13(3): 191-202. https://doi.org/10.24171/j.phrp.2021.0234

[11] Malhotra S, Mani K, Lodha R, Bakhshi S, Mathur VP, Gupta P, Kedia S, Sankar J, Kumar P, Kumar A, Ahuja V, Sinha S, Guleria R, COVID Reinfection AIIMS Consortium. (2022) SARS-CoV-2 Reinfection Rate and Estimated Effectiveness of the Inactivated Whole Virion Vaccine BBV152 Against Reinfection Among Health Care Workers in New Delhi, India. JAMA Network Open, 5(1): e2142210. https://doi.org/10.1001/jamanetworkopen.2021.42210

[12] W.H.O. (2021, May 31). Tracking SARS-CoV-2 variants. World Health Organization. https://www.who.int/activities/tracking-SARS-CoV-2-variants.

[13] Lich KH, Ginexi EM, Osgood ND, Mabry PL. (2013) A Call to Address Complexity in Prevention Science Research. Prevention Science, 14(3): 279-289. https://doi.org/10.1007/s11121-012-0285-2

[14] Kermack WO, McKendrick AG. (1927) A Contribution to the Mathematical Theory of Epidemics. Proceedings of the Royal Society of London, 115(772): 700-721. https://doi.org/10.1098/rspa.1927.0118

[15] Piccolomiini EL, Zama F. (2020) Monitoring Italian COVID-19 spread by an adaptive SEIRD model. Medrxiv. https://doi.org/10.1101/2020.04.03.20049734

[16] Korolev I. (2020) Identification and estimation of the SEIRD Epidemic Model for COVID-19. Journal of Econometrics, 220(1): 63-85. https://doi.org/10.2139/ssrn.356936710.1016/j.jeconom.2020.07.038

[17] Maugeri A, Barchitta M, Battiato S, Agodi A. (2020) Modeling the Novel Coronavirus (SARS-CoV-2) Outbreak in Sicily, Italy. International journal of environmental research and public health, 17(14): 4964. https://doi.org/ 10.3390/ijerph17144964

[18] Shao P, Shan Y. (2020) Beware of asymptomatic transmission: Study on 2019-nCoV prevention and control measures based on extended SEIR model. BioRxiv. https://doi.org/10.1101/2020.01.28.923169.

[19] Department of Statistics Malaysia (2020, July 15). Current Population Estimates, Malaysia, 2020. Department of Statistics Malaysia Official Portal. https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=155&bul_id=OVByWjg5YkQ3MWFZRTN5bDJiaEVhZz09&menu_id=L0pheU43NWJwRWVSZklWdzQ4TlhUUT09

[20] Gotz T, Heidrich P. (2020) Early stage COVID-19 disease dynamics in Germany: models and parameter identification. Journal of Mathematics in Industry, 10(1): 1-13. https://doi.org/10.1186/s13362-020-00088-y

[21] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KSM, Lau EHY, Wong JY, Xing X, Xiang N, Wu Y, Li C, Chen Q, Fheng Z. (2020) Early transmission dynamics in Wuhan, China of novel coronavirus–infected pneumonia. New England Journal of Medicine, 382(13): 1199-1207. https://doi.org/10.1056/NEJMoa2001316

[22] Read JM, Bridgen JR, Cummings DA, Ho A, Jewell CP. (2021) Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 376(1829): 20200265. https://doi.org/10.1098/rstb.2020.0265

[23] Jamil NM, Rosli N, Muhammad N. (2021) Simulation of COVID-19 outbreaks via Graphical User Interface (GUI). Journal of Public Health Research. https://doi.org/10.4081/jphr.2021.2130

[24] COVID-19 Database in Github MOH [https://github.com/MoH-Malaysia/covid19-public/blob/main/epidemic/cases_malaysia.csv]

[25] Jayaraj VJ, Rampal SN, Chiu WC, Diane WQ. (2021) The Epidemiology of COVID-19 in Malaysia. The Lancet Regional Health - Western Pacific, 17(2021): 100295. https://doi.org/10.1016/J.LANWPC.2021.100295

[26] Ganasegaran K, Ch'ng ASH, Looi, I. (2021) What Is the Estimated COVID-19 Reproduction Number and the Proportion of the Population That Needs to Be Immunized to Achieve Herd Immunity in Malaysia? A Mathematical Epidemiology Synthesis. COVID 2021, 1(1): 13-19. https://doi.org/10.3390/COVID1010003

[27] Zamri ASS, Singh S, Ghazali SM, Herng LC, Dass SC, Aris T, Ibrahim HM, Gill BS. (2021) Effectiveness of the movement control measures during the third wave of COVID-19 in Malaysia. Epidemiology and Health: September 2021; 43:e2021073. https://doi.org/10.4178/epih.e2021073

[28] Chin ESM. (2021, June 17) Report: Health minister says 113 patients reinfected with Covid-19 after release. Malay Mail. https://www.malaymail.com/news/malaysia/2021/06/17/report-health-minister-says-113-patients-reinfected-with-covid-19-after-rel/1982819

[29] Mahmud A, Lim YP. (2020) Applying the SEIR Model in Forecasting The COVID-19 Trend in Malaysia: A Preliminary Study. Medrxiv. https://doi.org/10.1101/2020.04.14.20065607

[30] Labadin J, Hong BH. (2020) Transmission Dynamics of COVID-19 in Malaysia Prior to the Movement Control Order. Medrxiv. https://doi.org/10.1101/2020.02.07.20021188

[31] Alsayed A, Sadir H, Kamil R, Sari H. (2020) Prediction of Epidemic Peak and Infected Cases for COVID-19 Disease in Malaysia. International Journal of Environmental Research and Public Health, 17(11): 4076. https://doi.org/10.3390/ijerph17114076

NON-NEWTONIAN VISCOSITY BEHAVIOR INVESTIGATION FOR MALAYSIAN WAXY CRUDE OILS AND IMPACT TO WAX DEPOSITION MODELLING

**ASTRIYANA ANUAR[1,2]*, ISMAIL MOHD SAAID[2],
PETRUS TRI BHASKORO[3] AND ROHAIZAD MOHD NORPIAH [2]**

*[1]Department of Petroleum Engineering, Universiti Teknologi PETRONAS, Perak, Malaysia
[2]Project Delivery and Technology, PETRONAS, Kuala Lumpur, Malaysia
[3]Department of Advanced Material and Subsea Technology,
PETRONAS Research Sdn. Bhd., Selangor, Malaysia*

*Corresponding author: astriyana_21000006@utp.edu.my

ABSTRACT: Wax deposition is one of the major risks that causes a serious threat to pipeline transportation during operation, if not prevented. The remediation actions are usually costly; hence mitigation methods are in place to completely avoid the issues from happening. The wax deposition modelling technique has been accepted as a tool to design and continuously optimize the wax management strategy. Non-Newtonian oil-wax viscosity is an important parameter affecting wax deposition in pipelines. The present and widely used viscosity model assumes exponential behavior as observed in the emulsion system. In this paper, it is demonstrated that this assumption may not be suitable for Malaysian waxy crude oil applications due to instantaneous change of viscosity below WAT and PPT. This paper focuses on the application of the Pedersen and Ronningsen viscosity model available in the commercial fluid and flow simulators namely PVTsim ®, Multiflash ® and OLGA ® which are widely used by the flow assurance fraternities, and how it will impact wax deposition prediction accuracy specifically when applied to Malaysian waxy crude oils.

ABSTRAK: Pemendapan lilin adalah salah satu risiko utama yang menyebabkan ancaman serius kepada pengangkutan saluran paip semasa operasi jika tidak dicegah. Proses membaiki biasanya memerlukan kos yang tinggi; oleh itu kaedah mitigasi disediakan bagi mengelakkan isu ini daripada berlaku. Teknik model pemendapan lilin telah diterima sebagai alat bagi mereka bentuk dan merupakan strategi optimum pengurusan lilin secara berterusan. Kelikatan minyak-lilin bukan Newton adalah salah satu parameter penting yang mempengaruhi pemendapan lilin dalam saluran paip. Anggaran model kelikatan semasa yang digunakan secara meluas menjangkakan tingkah laku eksponen seperti yang diperhatikan dalam sistem emulsi. Kajian ini menunjukkan bahawa kaedah anggaran mungkin tidak sesuai bagi aplikasi minyak mentah berlilin Malaysia disebabkan oleh perubahan kelikatan serta-merta di bawah WAT dan PPT. Kertas kerja ini memberi tumpuan kepada aplikasi model kelikatan Pedersen dan Ronningsen yang terdapat dalam cecair komersial dan simulator aliran iaitu PVTsim ®, Multiflash ® dan OLGA ® yang digunakan secara meluas oleh persatuan jaminan aliran, dan keberkesanan pada ketepatan anggaran pemendapan lilin khususnya apabila digunakan pada minyak mentah berlilin Malaysia.

KEYWORDS: *waxy crude oil; viscosity; non-Newtonian viscosity; wax deposition; flow modelling; rheology*

1. INTRODUCTION

Flow assurance (FA) brings the most significant financial impact especially for deepwater development and FA-critical fields (i.e., fields with severe FA issues). The risks and potential issues need to be defined as early as possible prior to development, and the types of flow assurance issues and their management plans are usually included in the field's overall operating philosophy to ensure safe and unrestricted transportation of hydrocarbon from reservoir to customer. Wax deposition is one of the major risks that causes a serious threat to pipeline transportation during operation, if not prevented. When the surrounding temperature drops below Wax Appearance Temperature (WAT), wax crystallization happens followed by deposition of the precipitated wax solids with increasing thermal gradient (due to low surrounding seabed temperature). Wax deposition of crude oil in pipelines reduces the flow area with time which affects the crude production and transportation. Wax deposits are usually removed mechanically using a scraper pig and is done periodically as the wax builds up in the pipeline. Additionally, to mitigate wax deposition in the pipelines, chemicals such as wax inhibitors, dispersants or wax crystal modifiers are commonly used in industry [1].

For a new field development, a reservoir fluid sampling and analysis program is executed as part of lab characterization activities to quantitatively evaluate the wax's properties and its tendency to deposit. Flow assurance engineers use this information as input into the fluid and flow software and heavily rely on the results of these modelling activities in defining the wax management strategy philosophy that should cover the entire field life cycle. Additionally, during the production stage, wax deposition modelling has been accepted as a tool to effectively optimize the pipeline operating envelope, specifically in determining pigging frequencies and reviewing chemical injection effectiveness in inhibiting wax deposition, however, several published papers have reported on the over and under-estimation of wax deposition rates. Consequently, modelling techniques have been observed to be unreliable without representative input and further model benchmarking or validation with experimental or field data [2,3]. Non-Newtonian oil-wax viscosity is identified as an important parameter affecting wax deposition in the pipeline [2].

This paper investigates the non-Newtonian behaviors of Malaysian waxy crude oils and reviews the suitability of the present wax viscosity model by Pedersen and Ronningsen [4] embedded in the industry standard fluid and flow simulator for wax deposition modelling namely PVTsim ® by CALSEP, Multiflash ® by KBC and OLGA ® by Schlumberger which are widely used by the FA fraternities, and additionally evaluate its impact to wax deposition modelling specifically for Malaysian waxy crude oil applications. Rheological experiments were run for six Malaysian waxy crude oil samples and used for this evaluation. The application of the Pedersen and Ronningsen model for wax deposition modelling is demonstrated using the widely accepted Matzain Model [5].

As highlighted in various published papers, methods to fit experimental viscosity data into the viscosity model for wax deposition modelling were not explicitly discussed. Rather, wax deposition modelling was performed by tuning or fitting of empirical constants in matching experimental and field wax deposition data, assuming accurate viscosity prediction using the viscosity model [2,3,6,7]. Viscosity prediction inaccuracy could lead to wax deposition prediction errors that would contribute to substantial risk in wax management especially when flow simulators are used to troubleshoot issues at the field [10]. Soedarmo et al. [2] in their paper, validated a few wax deposition models using 70 wax deposition data points from 19 different conditions and concluded that to improve modelling prediction uncertainties, better measurements of wax precipitation and viscosity are required.

It is generally summarized that current deposition models may not be able to provide reliable predictions on long-term wax deposition. Design optimization and routine maintenance operations could then be a challenge, hence the need to improve the accuracy of deposition predictions in order to strike a balance between optimal design and economical operations.

2. METHODOLOGY

The investigation is conducted by evaluating the present model performance by comparing the predicted non-Newtonian viscosity profiles which were simulated using Pedersen and Ronningsen model in the fluid simulator, PVTsim, against experimental data sets from six Malaysian waxy crude oil samples. The results are plotted to demonstrate the prediction of wax deposition using various viscosity modelling methods in matching experimental data points. The impact of viscosity predictions on wax deposition modelling are demonstrated in OLGA, a transient multiphase flow simulator.

In general, the major components of wax deposition modelling are thermodynamics, hydrodynamics, heat and mass transfer, and wax deposition mechanism. The temperature and wax concentration profile can be obtained by simultaneously solving the heat and mass transfer equations. Typically, crude oil viscosity is treated as a Newtonian fluid for multiphase flow modelling, however, for wax deposition modelling, when dealing with fluid temperatures below WAT, the effect of the resulting non-Newtonian fluid behaviors has to be considered. This is because, viscosity has considerable impact on wax deposition in the pipeline which is mainly governed by diffusion and shear stripping mechanism.

There are several rheological models published in describing the non-Newtonian waxy crude oil behaviors namely the Bingham model, also called pseudo-plastic [11], the Herschel-Buckley (HB) model, and the Casson model [4], as shown in Eq. (1) to Eq. (3), respectively.

$$\tau = \tau_B + \eta\gamma \ (Bingham) \tag{1}$$

$$\tau = \tau_{HB} + m\gamma^n \ (Herschel - Buckley) \tag{2}$$

$$\sqrt{\tau} = \sqrt{\tau_C} + \sqrt{\eta}\sqrt{\gamma} \ (Casson) \tag{3}$$

where τ is the shear stress (Pa), η is the viscosity (Pa. s) and γ is the applied shear rate (s^{-1}).

Bhaskoro [12] in his work, reviewed the applications of these rheological models for four Malaysian waxy crude oils and concluded that both the Casson model and the HB model can be used to describe the rheological behaviors below WAT of all the crude samples with the latter model generally providing a better fit with $R^2=0.9998$.The general viscosity model was then developed based on Malaysian crude oil's rheological and wax data based on the HB model and the results presented showcased better performance when compared with viscosity models previously available, as presented in Bhaskoro et al. [13]. However, the model requires additional fluid information i.e., activation energy, which is not readily available from standard testing and applied using current engineering software. It is worth noting that, the Pedersen and Ronningsen model has been widely used by the FA engineers as it is the only model that is readily available in most commercial applications such as PVTsim, Multiflash and OLGA and thus will be the context of this paper.

Pedersen and Ronningsen [4] stated that the apparent viscosity of the crude oil would increase when wax crystalized and dispersed in bulk oil, which eventually led to an increase in the pressure drop in the pipeline. When the wax particles are available at sufficiently high concentrations, it will change the crude oil flow behavior from Newtonian to non-Newtonian behavior. The transition typically occurs at roughly 10 to 15 °C below WAT and at solid wax fraction of 1 to 2wt%, but this usually depends on the waxiness/characteristics of the oil. Their first viscosity model was presented based on 713 measured viscosity data points from 18 different gas-free North Sea oils and leveraging the Casson-type rheological fluid model to describe the non-Newtonian behaviors [4]. Review methodology is presented in Fig. 1.
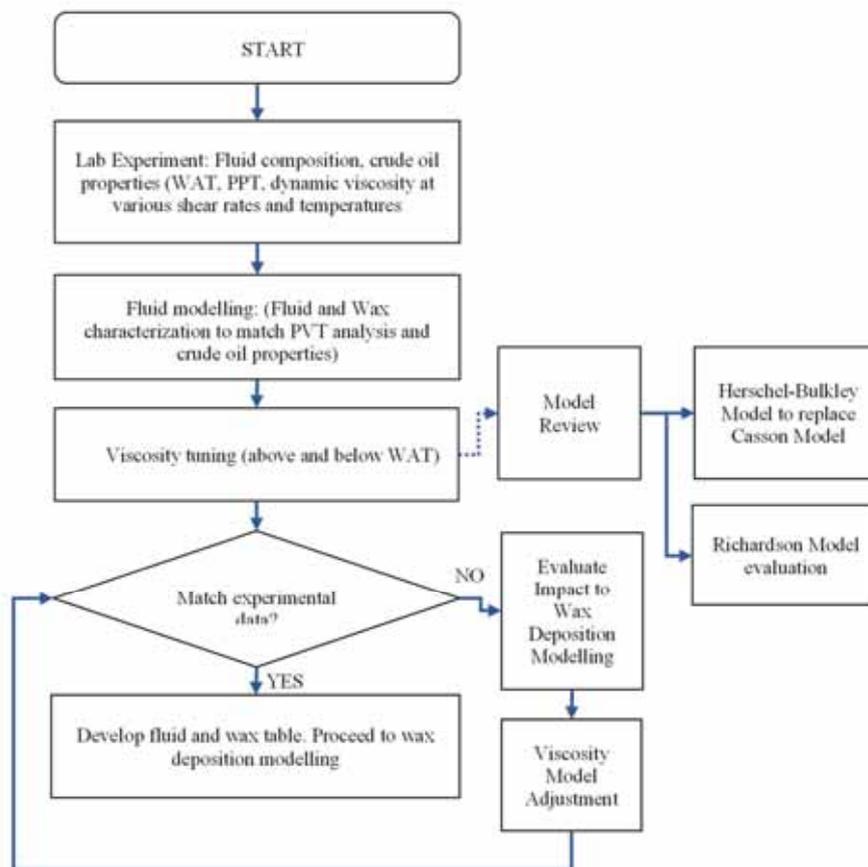
Fig. 1: Review methodology.

2.1 Non-Newtonian Viscosity Measurement

An AntonPaar MCR-302 controlled stress rheometer was used to measure crude oil viscosity covering both Newtonian and non-Newtonian ranges. Cross-hatched parallel plate geometry made of stainless steel with a diameter of 40 mm was used to avoid slippage during measurement. The groove roughness is set higher than 10 μm as suggested by Barnes [14] for a condition where complete adhesion is violated. Similar procedure proposed by Yoshimura and Prud Homme [15] was adopted by which minimal slip effect was achieved.

The crude oil sample, and both lower and upper roughened geometry (the roughened parallel plate), were pre-heated around 10 °C above the WAT but below 60 °C to avoid

evaporation of light components. Approximately 2 mL of (liquid) crude oil sample was loaded onto the measurement gap set at 500 microns using a pipette. Japper-Jaafar et al., [16] explained that this gap size aims to provide some degree of freedom for the wax crystals' movement and to minimize wall effects. The sample was maintained at the temperature for 5 minutes to ensure complete wax dissolution. Then, viscosity of the sample was measured from approximately 10 °C above WAT down to 15 °C at a cooling rate of 0.5 °C/min and under continuous shear rates of 10 s^{-1}, 50 s^{-1}, 100 s^{-1}, 300 s^{-1}, 500 s^{-1} and 1000 s^{-1}.

The viscosity measurement procedures to ensure accuracy and repeatability in this paper is as per described in Petrus and Azuraien [9]. To address the viscosity measurement reproducibility, the same sample from previous work by Bhaskoro et al., [13] using one of the Malaysian crude oils, Fluid S06 (which is represented as Crude oil A in his paper) was used for the current study. Temperature dependent-viscosity measurements for crude oil A were conducted using a controlled stress rheometer AR-G2 and DHR-1 from TA instrument while the current study uses other equipment which is an AntonPaar MCR-302 rheometer in a different laboratory. A log scale plot is presented in Fig. 2 to demonstrate the viscosity trends based on current and previous works.
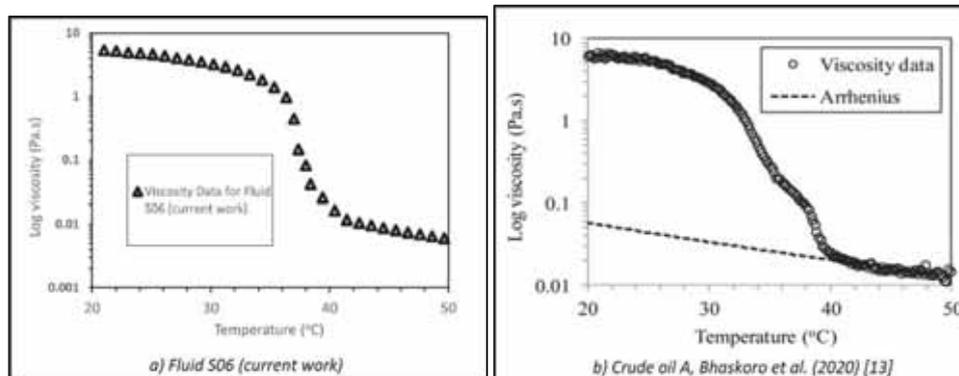
Fig. 2: Viscosity measurement reproducibility test.

Fig. 2 compares viscosity measurement from a) current work and b) previous work as presented in Bhaskoro et al., [13] that gives the same trend of instantaneous steep change of viscosity trends below WAT and PPT. Additionally, Petrus and Azuraien [9] observed similar trend when conducting rheological measurement for a waxy crude oil from the South East Asia region, as presented in Fig. 2 and Fig. 3 in their paper .

2.2 Non-Newtonian Viscosity Modelling

In general, the viscosity model by Pedersen and Ronningsen [4] has been incorporated in PVTsim, Multiflash and OLGA, however, in PVTsim, the viscosity model has been modified using proprietary data from Statoil [17].

The apparent viscosity of oil with suspended wax particles (a shear-rate-dependent viscosity model) is calculated as follows,

$$\eta = \eta_{liq}\left[exp(D.\theta_{wax}) + \frac{E.\theta_{wax}}{\sqrt{\frac{dv_x}{dy}}} + \frac{F.\theta_{wax}^4}{\frac{dv_x}{dy}}\right] \qquad (4)$$

where η_{liq} is the viscosity (Pa.s) of the oil not considering solid wax and $\varnothing_{wax}$ the volume fraction of the precipitated wax in the oil-wax suspension, dv_x/dy is the shear rate (s^{-1}), and the empirical constant/parameters D, E, F will have the following values: D = 37.82, E = 83.96, and F = 8.559 x 10^6 (shear rates in s^{-1}) [18]. Calculation of the viscosity of the wax/oil dispersion in OLGA and Multiflash simulators assumed the model directly proposed by Pedersen and Ronningsen [4].

PVTsim software, assumed a similar model by Pedersen and Ronningsen, where it has, however, been modified in 2006 using proprietary data from Statoil. The empirical constant/parameters D, E, F in Eq. (4) will have the following values: D = 18.12, E = 405.1, and F = 7.876 x10^6 (viscosities in mPa.s and shear rates in s^{-1}) [17].

Pedersen and Ronningsen further presented a new viscosity model with empirical constants D, E, F that accounts for wax inhibitors effect based on 12 different commercial wax crystal modifiers [19]. The empirical constants used in this case are D = 22.42, E = 1189, and F = 1.335 x 10^6 (viscosities in mPa.s and shear rates in s^{-1}). It should be noted that the study suggests that the inhibitors used only have marginal influence on the amount of solid wax formed but do exhibit pronounced effect on pour point and apparent viscosity, especially in the temperature interval from around 10 to 25 °C. The effect of the inhibitors on viscosity can be modelled by assigning to wax molecules in the range of C21 to C45 a lower melting temperature in the presence of wax inhibitors than that without.

In the current modelling practice, the wax-oil viscosity behavior is modelled and regressed in fluid simulators (i.e., PVTsim and Multiflash) to match the measured viscosity data for oil with suspended wax at different shear rates. The wax multipliers/correction factors are then applied to calculate the final wax viscosity multipliers using Eq. (5) to (7) which are served as input into OLGA or any other flow simulator when using the Matzain wax deposition model using input keys of VISCMULTD, VISCMULTE and VISCMULTF [17]. This correction is needed to account for different wax viscosity models used by different software (in this case, OLGA and PVTsim):

$$VISCMULTD_{OLGA} = VISCMULTD_{PVTsim} \times \frac{D_{PVTsim}}{D_{OLGA}} \qquad (5)$$

$$VISCMULTE_{OLGA} = VISCMULTE_{PVTsim} \times \frac{E_{PVTsim}}{E_{OLGA}} \qquad (6)$$

$$VISCMULTF_{OLGA} = VISCMULTF_{PVTsim} \times \frac{F_{PVTsim}}{F_{OLGA}} \qquad (7)$$

It should be additionally noted that this is only applicable when transferring multipliers from PVTsim to OLGA. The D, E, F empirical constants/parameters used by all software are tabulated in Table 1. In OLGA, shear rates are simulated and serve as input for the apparent oil viscosity calculation (4) to account for suspended wax particles.

Table 1: D, E, F empirical constants in Fluid and Flow Simulators

D,E,F Constants	OLGA/ Multiflash	PVTSIM
D	37.82	18.12
E	83.96	405.1
F	8.56 X 10^6	7.876 X 10^6

2.3 Matzain Wax Model

The Matzain wax model is a widely used wax model and it is available in most commercial multiphase flow simulators for industry applications. Even though few wax deposition models have been introduced and enhanced since the first time the Matzain Wax Model was developed more than 20 years ago, the investigation uses the Matzain Wax Model which has been reported to be effective in estimating wax deposition based on multiple validation works [2,7,18].

The Matzain Wax Model [5] is a semi-empirical model based on the Equilibrium model (EM) which considers molecular diffusion and shear stripping mechanism in predicting wax deposition. The kinetic wax deposition model is described as below:

$$\frac{d\delta}{dt} = \frac{\pi_1}{1+\pi_2} D_{ow} \left[\frac{dC_w}{dT} \frac{dT}{dr} \right] \tag{8}$$

$$D_{ow} = 7.4x10^{-8} \frac{T_w (\psi MW)^{0.5}}{\mu_{o,f} V^{0.6}} \tag{9}$$

where, δ is the thickness of wax layer deposited on the wall (m), D_{ow} is the diffusion coefficient evaluated at the wax-oil/gas interface using Wilke and Chang correlations [21] in Eq. (9), C_w is the concentration of wax in solution (weight %), r is the pipe radial distance (m) and T is the fluid temperature (°C). In Eq. (6), ψ is the oil association parameter assumed equal to 1.0 and recommended for a crude oil system. D_{ow} is in cm²/s, $\mu_{o,f}$ which is the viscosity of wax-oil in mPa.s and wax molar volume (V) is in cm³/mol.

π_1 is an empirical correlation introduced to account for the porosity effect on the rate of wax build up and for other deposition enhancement mechanisms not considered by the diffusion mechanism alone (e.g., turbulent mass diffusion mechanism) as shown in Eq. (10) below,

$$\pi_1 = \frac{C_1}{1 - C_L/100} \tag{10}$$

$$C_L = 100 \left(1 - \frac{N_{Re,f}^{0.15}}{8} \right) \tag{11}$$

where, C_L describes the amount of oil trapped in the wax layer and $C_1 = 15$.

The dimensionless parameter $N_{Re,f}$ is a function of the effective inside radius of the pipeline and is evaluated at the wax-oil/gas interface as shown in Eq. 12, where ρ_{oil} is the density of oil in kg/m³, v_{oil} is oil velocity in m/s, d_w is the effective inside pipe diameter as a result of wax build-up in m and μ_o is the viscosity of oil in mPa.s.

$$N_{Re,f} = \frac{\rho_{oil} v_{oil} d_w}{\mu_{oil}} \tag{12}$$

π_2 accounts for the wax limiting effect due to shear stripping and is defined as below:

$$\pi_2 = 1 + C_2 N_{SR}^{c_3} \tag{13}$$

where $C_2 = 0.055, C_3 = 1.4$

The flow regime dependent Reynolds number (N_{SR}) is calculated for each regime as shown below:

For single phase and stratified wavy flow:

$$N_{SR} = \frac{\rho_{oil} v_{oil} \delta}{\mu_{o,f}} \tag{14}$$

For bubble and slug flow:

$$N_{SR} = \frac{\rho_{mix} v_{oil} \delta}{\mu_{o,f}} \qquad (15)$$

For annular flow:

$$N_{SR} = \frac{\sqrt{\rho_{mix} \rho_{oil}}\, v_{oil} \delta}{\mu_{o,f}} \qquad (16)$$

The above expressions demonstrated that the wax deposition thickness profiles in pipeline according to the Matzain wax deposition model, have been modelled as dependent on the wax deposition mechanisms, flow conditions and flowing fluid properties, including oil viscosity and wax-oil viscosity. Wax-oil viscosity here refers to wax-oil suspension viscosity and this property is evaluated at the wax-oil/gas interface based on the Matzain Model [5].

Soedarmo et al., [2] presented wax deposition model's performance using Film Mass Transfer (FMT) and Equilibrium Model (EM) which can be affected by viscosity input as presented in Fig. 3, included in their paper.
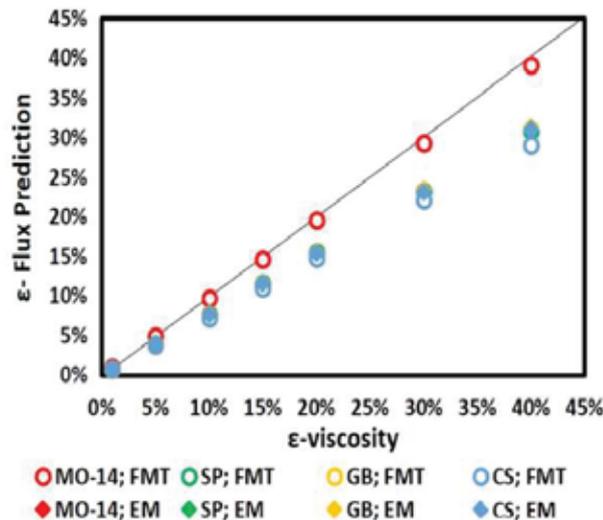
Fig. 3: Wax Mass Flux Prediction Errors: FMT and EM model sensitivity
to uncertainties in viscosity.

As viscosity plays major roles in wax deposition predictions, it is crucial to get accurate predicted values from the viscosity model, to ensure accurate wax deposition modelling for field consumption. The traditional approach in tuning the deposition model is for the simulation models to match wax deposition tests which are ideally conducted using large or smaller scale pipeline tests at laboratory, also called flow loop. However, due to the cost associated with the test, often, it is not conducted, and empirical constants in the deposition models were fitted to match field data (e.g., temperature, pressure). Without further understanding on how the waxy crude oil deposit, model fittings could introduce unphysical behavior. To reduce the wax deposition modelling fitting ranges and uncertainties, it is prudent to identify the effect of each parameters used and to evaluate methods to eliminate tuning uncertainties, which in this case is focusing on the application of present non-Newtonian viscosity model for Malaysian waxy crude oils.

3. CASE STUDY ON MALAYSIAN WAXY CRUDE OILS

3.1 Materials

Six (6) waxy crude oil samples from Malaysian fields are used for this study. The basic properties measured are tabulated in Table 2 as below. Three data sets (Fluid D02, Fluid S06, and Fluid D01a) will be used to showcase viscosity prediction via modelling (covering fluid and flow simulations) and all six data sets will be used to showcase Malaysian waxy crude oils' viscosity measurements and behaviors.

Table 2: Fluid experimental data sets

Crude Oil Samples	Wax Appearance Temperature (ºC)	Pour Point Temperature (ºC)	Wax Content (wt%)
Fluid D01a[a]	37	17	20
Fluid D02	44	33	7.6
Fluid D01b[a]	40	27	18.3
Fluid W101	43	36	15.5
Fluid W103	36.5	18	11.4
Fluid S06	49.5	39	26.7

[a] Fluid D01a and Fluid D01b are fluids taken from commingling header; Fluid D01a has been added with Pour Point Depressant while other fluids in this list are blank waxy crude oil samples directly from the wells.

For the purpose of wax deposition modelling, the viscosity impact will be evaluated mainly from non-Newtonian rheological behaviors perspective (below WAT). Experimental viscosity data at various shear rates for the above crude oils are plotted in Fig. 4, Fig. 5, Fig. 6**Error! Reference source not found.**, Fig. 7, Fig. 8, and Fig. 9 respectively. The shear rates for specific crude oil are determined based on pipeline size and production rates accordingly.
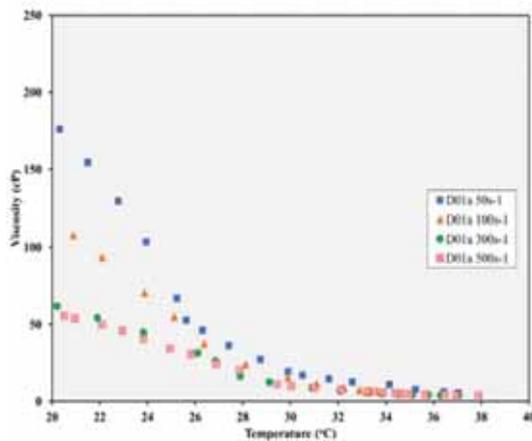
Fig. 4: Experimental viscosity data below WAT for Fluid D01a at various shear rates.
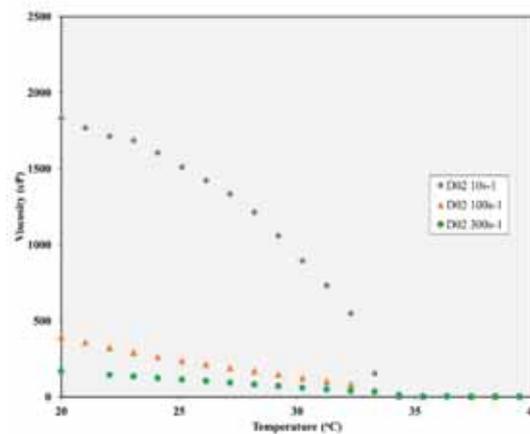
Fig. 5: Experimental viscosity data below WAT for Fluid D02 at various shear rates.
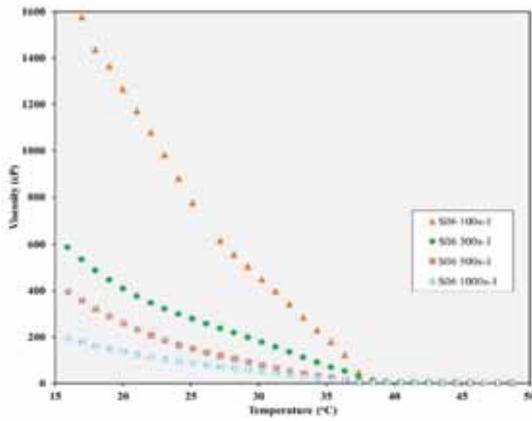
Fig. 6: Experimental viscosity data below WAT for Fluid S06 at various shear rates.
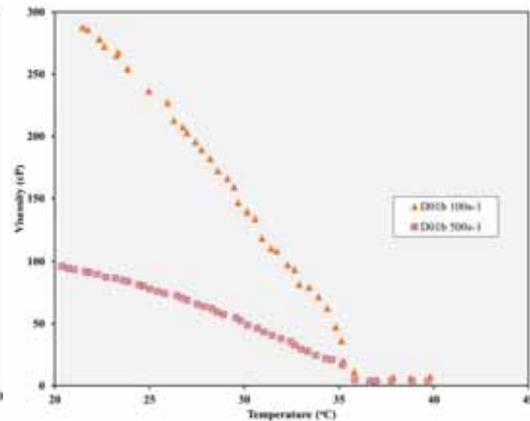
Fig. 7: Experimental viscosity data below WAT for Fluid D01b at various shear rates.
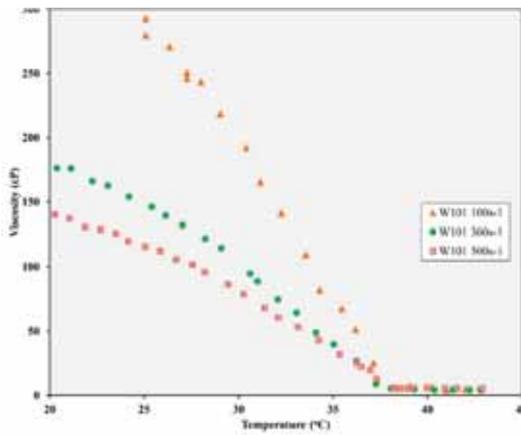
Fig. 8: Experimental viscosity data below WAT for Fluid W101 at various shear rates.
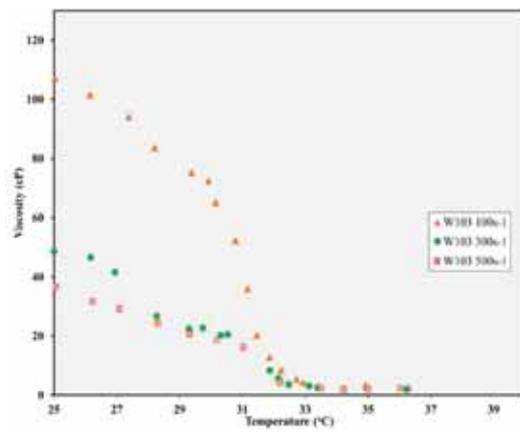
Fig. 9: Experimental viscosity data below WAT for Fluid W103 at various shear rates.

Additionally, three (3) material parameters; τ_{yo}, k and n, as a function of temperature are presented in Table 3 for Fluid S06.

Table 3: Herschel-Bulkley parameters obtained at various temperatures after dynamic cooling at 1 °C/min for Fluid S06

T (°C)	τ_{yo} (Pa)	k (Pa.s)	n
55	0	0.00305	1
52.5	0	0.00324	1
50	0	0.00371	1
47.5	0	0.00414	1
45	0	0.00478	1
42.5	0	0.00534	1
40	0.2	0.0007	0.973
37.5	15.5	0.0699	0.9794
35	16.2	0.253	0.9374
32.5	22	0.2462	0.9184

30	25	0.4012	0.9199
27.5	26.5	0.5767	0.9099
25	30	0.9118	0.906
22.5	32	1.0201	0.9158
20	40.3	1.7515	0.8733

4. RESULTS AND DISCUSSION

4.1 Regression of Viscosity Model to fit Experimental Data

Prior to viscosity tuning in PVTsim, fluid and wax characterization activities were performed on each fluid to match the experimental dataset comprising of GC (gas chromatography), PVT data (pressure volume temperature), n-paraffin distribution, wax content, WAT, and crude oil properties. In general, tuning enables the viscosity model to match the given experimental data at specific shear rates and temperature ranges. This is done in fluid simulators by adjusting wax viscosity multipliers/correction factors manually through a trial-and-error method. The tuning process was repeated until it matched the experimental data.

4.2 Impact of non-Newtonian viscosity on Wax Deposition Modelling

Equations (8) to (16) from Matzain model demonstrates how viscosity plays roles in determining the wax deposition thickness profile in pipelines. Since wax deposition is a function of wax deposition mechanisms, flow regime, and many other factors, to enable direct evaluation of the impact of non-Newtonian viscosity model to wax deposition thickness, the OLGA runs will focus on simulating the diffusion coefficient, D_{ow}, as it has a direct inverse effect to wax-oil viscosity on top of evaluating general impact to wax mass (net effect). Viscosity tuning below WAT in matching experimental data are achieved using either automatic or manual tuning and the impact to wax deposition will only be demonstrated for three (3) fluids: Fluid D01a, Fluid D02 and Fluid S06, since Fluid D01b, Fluid W101 and Fluid W103 are produced from the same field but from different reservoirs.

Flow loop models were constructed in OLGA to showcase the impact on the viscosity model predictions in flow simulators. The Matzain model is used, and wax deposition tuning factors are kept as default. The goal is to simulate wax deposition modelling and evaluate the impact of applying different viscosity multipliers. Tuning to wax deposition tests is currently not considered for this work. To demonstrate the immediate effect of viscosity multipliers to wax deposition, departing fluid temperatures are set to below WAT. Additionally, it is worth noting that VISHL, which is the effective viscosity including the non-Newtonian wax effect, is determined at fluid temperature in OLGA.

4.2.1 Fluid D01a – Viscosity Behavior Transition 4 °C below WAT

Dashed lines in Fig. 10 represent the viscosity predictions using the Pedersen and Ronningsen model based on various tuning methods. For Fluid D01a, automatic viscosity tuning at various shear rates (Set 3) returned a good matching to experimental data only at higher shear rates of 500 s^{-1} and at higher temperatures. Manual tuning on viscosity at various shear rates (Set 1) showed improvement in viscosity predictions however, they still suffer underprediction of up to 2 times, especially at low temperature ranges. D, E, F wax viscosity multipliers/correction factors are presented in Table 4. Better match to measured

viscosity data is achieved with (Set 2) which was selected as the basis for wax deposition modelling.

Fluid D01a flows from Field Y through a 10-km pipeline with seabed temperatures ranging from 26 °C to 30 °C and at shear rates of 100 s^{-1} and lower. To enable correct representation of the viscosity profile for Fluid D01a, the viscosity model is manually adjusted to match experimental shear rates at 50 s^{-1} and 100 s^{-1} (Set 2) matching the pipeline shear rates. Manual tuning enabled matching to the lowest operating temperature of 26 °C and higher.
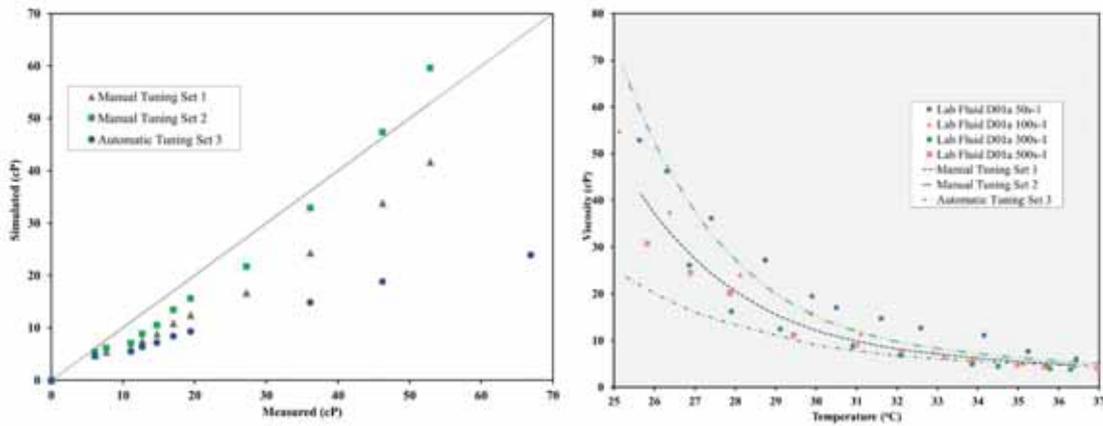
Fig. 10: Tuned viscosity below WAT at 50 s^{-1} for Fluid D01a using various tuning methods.

Table 4: Viscosity Multipliers/Correction Factors used in OLGA for Fluid D01a

Viscosity Multipliers /Correction Factors	Manual Tuning at Various Shear Rates (Set 1)	Manual Tuning at Specific Shear Rates/ Temperature (Set 2)	Automatic Tuning at Various Shear Rates (Set 3)
D	0.292	1.437	0.767
E	5.675	0	0
F	3.030	0.575	0.575

In OLGA, a 0.25-inch and 1-meter flow loop model with fluid departing temperature of 30 °C (below WAT), seabed ambient temperature of 26 °C and a mass flow of 21 kg/s have been specified as mass source. The fluid D, E, F wax viscosity multipliers used in OLGA after tuning in PVTsim are as per Table 4. Fig. 11 presents the wax mass (net effect) using three different sets of wax viscosity multipliers, and the Apparent Wax Diffusion Coefficient for oil/wax dispersion, D_{ow} is tabulated in Table 5.
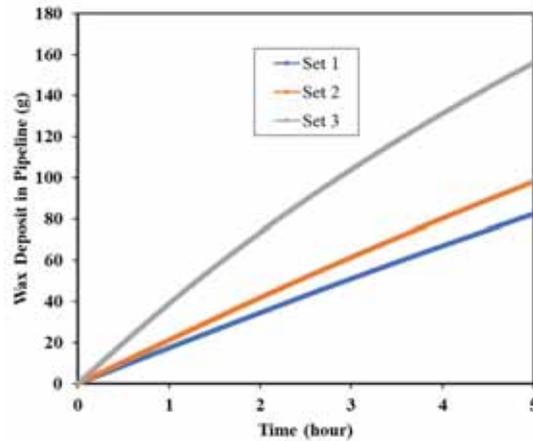
Fig. 11: Wax deposit mass in branch predicted in OLGA using various viscosity tuning methods based for Fluid D01a.

Table 5: Wax deposition modelling results using different viscosity multipliers for Fluid D01a

Simulation Run (5-hours)	Wax Mass (gram)	Apparent Wax Diffusion Coefficient (m²/s)	Predicted Viscosity	Impact to Wax Mass
Set 1	82.31	5.592×10^{-11}	Higher viscosity	-15.97% Error
Set 2	97.95	6.312×10^{-11}	Baseline (Proposed tuning)	
Set 3	155.77	9.717×10^{-11}	Lower viscosity	+59.03% Error

Table 5 summarizes the impact of various viscosity multipliers on wax deposition prediction. We can observe significant increase up to 59% in wax deposition prediction using automatic tuning (Set 3), which tends to underpredict the viscosity, and 16% reduction using manual tuning at various shear rates (Set 1) when compared to Set 2. The results showed that underprediction on viscosity value will give high wax deposition mass and vice versa. In this case, set 2 has been used as baseline since it provides the best match with experimental data at the respective pipeline shear rates and temperatures.

4.2.2 Fluid D02 – Viscosity Behavior Transition 10 °C below WAT

Fluid D02 was initially tuned in PVTsim at all shear rates (automatic tuning) and the simulated wax viscosity multipliers are VISCMULTD = 0, VISCMULTE = 0 and VISCMULTF = 0.5632. The zero values indicated that the value is not found to be sensitive to any further tuning. It can be observed from Fig. 12 that after tuning to experimental data, Fluid D02 is still unable to accurately match experimental data at all shear rates, especially at low shear rate, therefore tuning is conducted based on one interested shear rate most potentially experienced by the pipeline. The interested shear rate is calculated using pipeline flowrates and geometries.

Fig. 12: Tuned viscosity below WAT at various shear rates for Fluid D02 using various tuning methods.
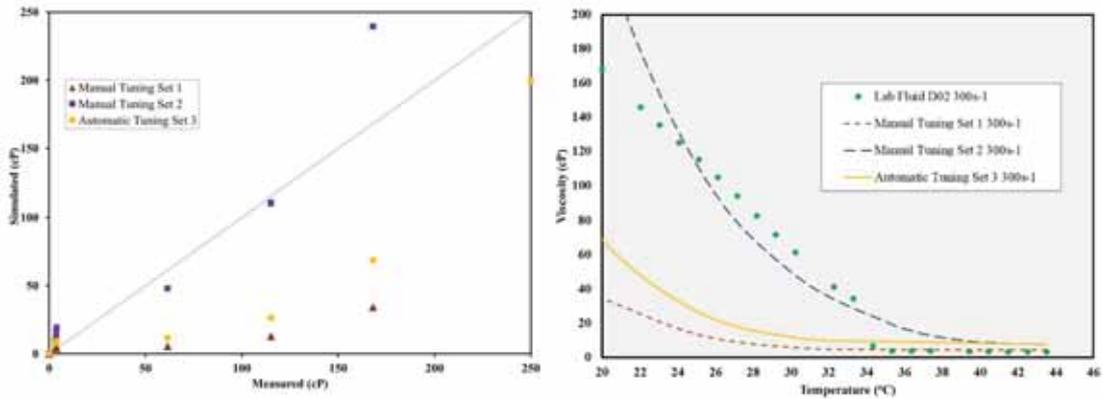
Fig. 13: Tuned Viscosity below WAT at 300 s^{-1} for Fluid D02.

After tuning, present model showcased an underprediction of viscosity up to 4 times (refer to Fig. 13), which is beyond 100% error, especially at operating fluid temperature of below pour point of 33 °C. Additionally, it can be observed that present viscosity model has not able to match the transition region between 33 °C to 40 °C. For Fluid D02, the average seabed ambient temperature experienced by the pipeline is 22 °C, hence it is critical to get the viscosity model to match experimental data, especially at the lower temperature ranges.

Table 6: Viscosity Multipliers/Correction Factors used in OLGA for Fluid D02

Viscosity Multipliers /Correction Factors	Manual Tuning at Various Shear Rates (Set 1)	Manual Tuning at Specific Shear Rates/ Temperature (Set 2)	Automatic Tuning at Various Shear Rates (Set 3)
D	0	0	0
E	0	12.062	0
F	0.5182	1.251	1.183

A 1.75-inch and 3-meter model is constructed in OLGA, with departing temperature at 40 °C (below WAT), seabed temperature of 22 °C and a mass flow of 0.513 kg/s which were specified as mass source. The D, E, F wax viscosity multipliers used in OLGA after tuning in PVTsim are as per Table 6. Field Z flows Fluid D02 at a shear rate of 300 s^{-1} and a minimum seabed temperature of 22 °C as described. Fig. 14 presents the Wax Mass (net effect) at three different sets of viscosity multipliers, and the Apparent Wax Diffusion Coefficient for oil/wax dispersion, D_{ow} is tabulated in Table 7.

As per Fig. 12 and Fig. 13, it is anticipated that both automatic (Set 3) and manual tuning of viscosity data at all shear rates (Set 1) would return a much lower viscosity predictions, which would return a high wax mass (up to 336%), as summarized in Table 7, and are in agreement with previous results. Manual tuning of Set 2 improved viscosity predictions for viscosity below 150 cP as presented in Fig. 13.
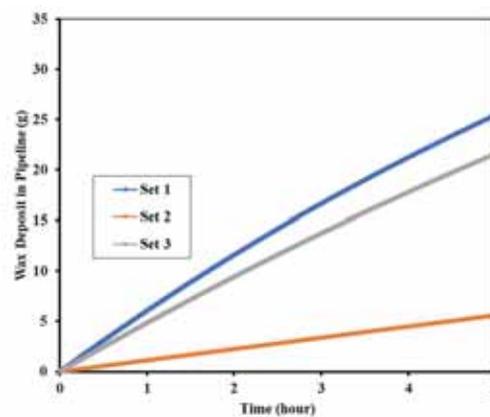
Fig. 14: Wax deposit mass in branch predicted in OLGA using
various viscosity tuning methods based for Fluid D02.

Table 7: Wax deposition modelling results using different viscosity multipliers for Fluid D02

Simulation Run (5-hours)	Wax Mass (gram)	Apparent Wax Diffusion Coefficient (m²/s)	Predicted Viscosity	Impact to Wax Mass
Set 1	29.34	6.25x10^{-11}	Lower viscosity	+336.61% Error
Set 2	6.72	1.12x10^{-11}	Baseline (Proposed tuning)	
Set 3	25.18	4.98x10^{-11}	Lower viscosity	+274.70% Error

4.2.3 Fluid S06 – Viscosity Behavior Transition 15 °C below WAT

A more comprehensive data set is presented for Fluid S06. Viscosity measurement is conducted at shear rates of 1000 s^{-1}, 500 s^{-1}, 300 s^{-1}, and 100 s^{-1}. The fluid was initially tuned in PVTsim at both shear rates (automatic tuning) and the resultant wax viscosity multipliers are VISCMULTD = 0.5867, VISCMULTE = 0.2271 and VISCMULTF = 2.6581. Generally, at higher shear rates (more than 500 s^{-1}), the Pedersen and Ronningsen model is able to match experimental data at most points, however at lower shear rates of 100 s^{-1}, 300 s^{-1}, and especially below Fluid S06 pour point of 39 °C, the model showcased an underprediction of up to 7 times. Manual tuning is conducted based on one interested shear rate (shear rate of 1000 s^{-1}) and is presented in Fig. 15.
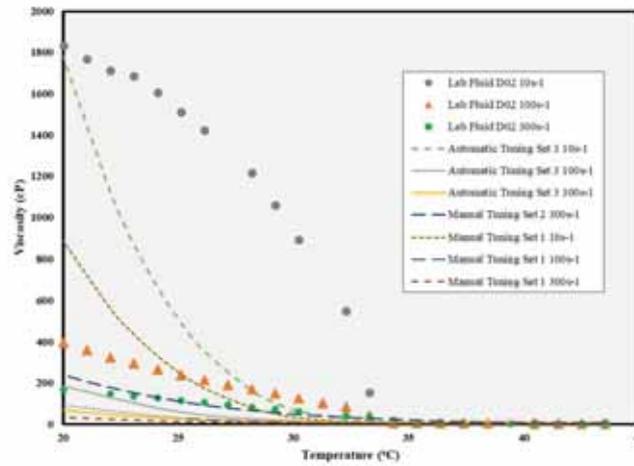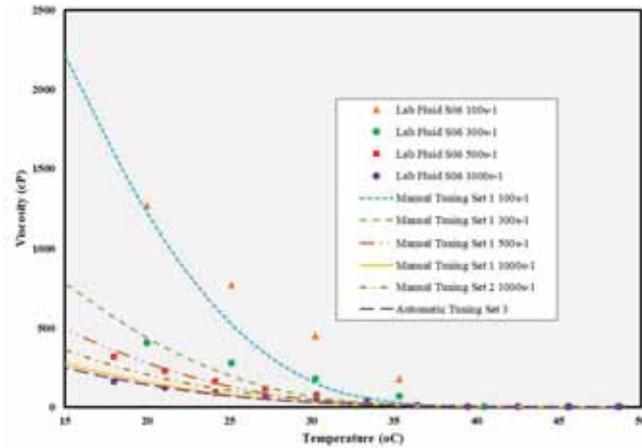
Fig. 15: Tuned viscosity below WAT for Fluid S06 using various tuning methods.
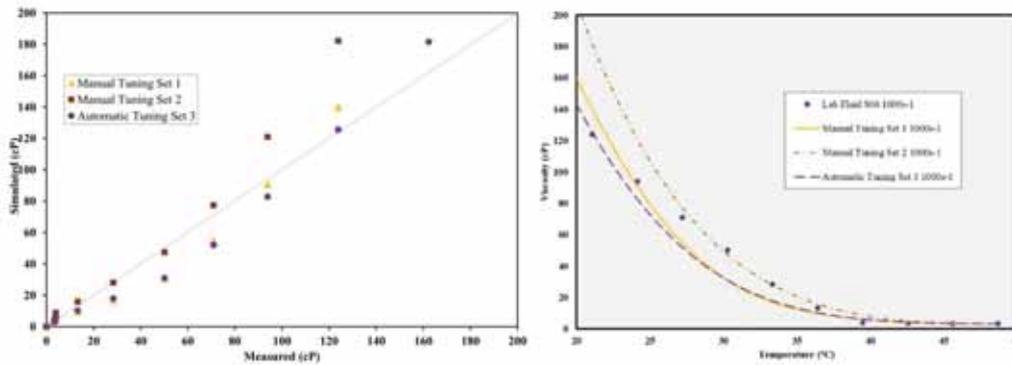
Fig. 16: Tuned Viscosity below WAT at 1000 s^{-1} for Fluid S06.

Fig. 16 showcases better predictions when fluid was tuned to specific shear rate of 1000 s^{-1} with a lower underprediction (less than 2 times) when compared to tuning at various shear rates in Set 1 and Set 3 (7 times). Manual tuning was performed to get a better match at a specific operating temperature range and shear rate. Even using manual tuning, the tuned models were not able to match experimental data at temperatures below 27 $^{\circ}$C accurately, and for immediate field application, the overprediction of viscosity could have an impact in terms of wax deposition as well as pressure drop in the pipeline, and this is deemed to be unacceptable.

Field X flows fluid S06 in a 7 km pipeline with seabed ambient temperature ranging from 22 $^{\circ}$C to 30 $^{\circ}$C. It is predicted that impact on wax deposition will be more apparent when ambient temperature drops below 27 $^{\circ}$C, but in the case of Fluid S06, due to fluid flowing at high shear rates, the impact to wax deposition is predicted to not be as significant. A 0.25-inch and 1-meter model was constructed in OLGA with departing temperature of 40 $^{\circ}$C, seabed temperature of 22 $^{\circ}$C and a mass flow of 0.02 kg/s were specified as mass source. The fluid D, E, F viscosity multipliers used in OLGA after tuning in PVTsim are as follows.

Table 8: Viscosity Multipliers/Correction Factors used in OLGA for Fluid S06

Viscosity Multipliers /Correction Factors	Manual Tuning at Various Shear Rates (Set 1)	Manual Tuning at Specific Shear Rates/	Automatic Tuning at Various

		Temperature (Set 2)	Shear Rates (Set 3)
D	0.281	0.552	0.513
E	1.337	11.160	0.914
F	2.446	0	0

Field X flows Fluid S06 at higher shear rates of 1000 s^{-1} and minimum seabed temperature ranges of 22 $^{\circ}$C. It is expected that the impact of applying different sets of viscosity multipliers would not be apparent as the Pedersen and Ronningsen model can match fairly well at high shear rates. Fig. 17 presented the Wax Mass (net effect) at three different sets of viscosity multipliers, and the Apparent Wax Diffusion Coefficient for oil/wax dispersion, D_{ow} is tabulated in Table 9.
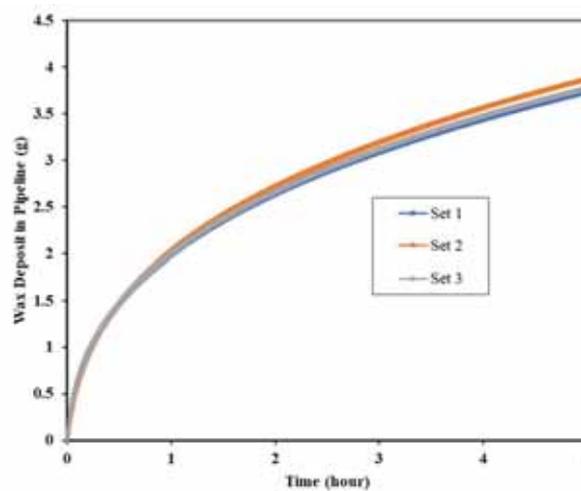
Fig. 17: Wax deposit mass in branch at different wax viscosity multipliers as predicted in OLGA based on the Matzain Wax Model at different viscosity multipliers for Fluid S06.

Table 9: Wax deposition modelling results at different viscosity multipliers for Fluid S06

Simulation Run (5-hours)	Wax Mass (gram)	Apparent Wax Diffusion Coefficient (m²/s)	Predicted Viscosity	Impact to Wax Mass
Set 1	4.210	2.09x10^{-11}	No significant difference at high shear rates	
Set 2	4.370	1.62x10^{-11}		
Set 3	4.259	1.93x10^{-11}		

It is concluded that since fluid flows at higher shear rates, the impact of different viscosity multipliers will not be as significant as compared to lower shear rates. In this case, the present model deemed to be acceptable.

4.3 Suitability of Present Model on Malaysian Waxy Crude Oil Applications

The Pedersen and Ronningsen model was developed based on viscosity data of 18 North Sea oils (API gravity ranging from 23.8 to 47.6), measured at temperatures between 40 and 0 $^{\circ}$C, and shear rates ranging from 30 s^{-1} to 500 s^{-1}. The model incorporated the Casson rheological model as per Eq. (3) which has been found to represent the equilibrium flow properties of waxy oils very well at high shear rates (up to 700 s^{-1}) and the Richardson

viscosity model to describe Newtonian behavior based on analogy from the oil/water emulsion system [4].

As per Fluid S06 review, generally at higher shear rates (more than 500 s^{-1}), the Pedersen and Ronningsen model is able to match experimental data at most points, however at lower shear rates of 100 s^{-1} and 300 s^{-1} and especially below Fluid S06 pour point of 39 °C, the model showcased an underprediction of up to 7 times. This is because high shear can continuously destroy wax particles and network leading to much smaller wax aggregates or crystals. It makes the fluid have slurry-like behavior which is found to be similar to that observed in the water/oil emulsion system, which is included in the Pedersen and Ronningsen model to describe the Newtonian behavior of the fluid [4].

As per Malaysian crude oil data sets earlier depicted in Table 2 and Fig. 4 to Fig. 9, Malaysian waxy crude oils of higher WAT and PPT tend to exhibit an instantaneous steep change of viscosity trends below WAT and PPT which does not follow the general relationship as described by the Pedersen and Ronningsen model. One hypothesis that can be acknowledged from previous study is that rheological data from Malaysian waxy crude oils were well-described using HB model [12], hence viscosity prediction using the present model could potentially suffer from inaccuracy mainly due to incorporation of the Casson rheological model instead of the HB model. To test this hypothesis, the HB model is fitted into the present model and its performance is further evaluated as below:

The first term is the viscosity which is defined as ratio of shear stress to shear rate.

$$\tau_{xy} = \eta \left(\frac{dVx}{dy} \right) \tag{17}$$

Eq. (17) is integrated into Eq. (2) giving a Herschel-Bulkley model presented as below:

$$\eta_c \frac{dV}{dy} = \tau_0 + k \left(\frac{dV}{dy} \right)^n \tag{18}$$

Assuming that $\tau_0 = \eta_{liq}\theta_{wax}{}^4$, $k = \eta_{liq}\theta_{wax}$ and n = Gθ_{wax} and by applying Richardson model, Eq. (19) may be re-written into a new equation as below,

$$\eta = \eta_{liq} exp(D\theta_{wax}) \left(\frac{F\theta_{wax}{}^4}{dV/dy} + E\theta_{wax} \left(\frac{dV}{dy} \right)^{G\theta_{wax}-1} \right) \tag{19}$$

where G is the additional viscosity multiplier. It is worth highlighting that Eq. (2) as described in Pedersen and Ronningsen [4] as the original model, is not exactly the same with what is described as Eq. (20) in Pedersen [19]. In this case study, performance from both equations will be evaluated.

Pedersen [19] presented a new Eq. (20) to describe viscosity behavior below WAT. Equation (21) is extended to include the HB rheological model.

$$\eta = \eta_{liq} exp(D.\theta_{wax}) \left[1 + \frac{E.\theta_{wax}}{\sqrt{\frac{dv_x}{dy}}} + \frac{F.\theta_{wax}^4}{\frac{dv_x}{dy}} \right] \tag{20}$$

$$\eta = \eta_{liq} \left(exp(D.\theta_{wax}) + k.\theta_{wax} \left(\frac{dV}{dy} \right)^{n-1} + \frac{\tau.\theta_{wax}^4}{\frac{dv_x}{dy}} \right) \tag{21}$$

It can be observed from Fig. 18 that even though the HB model has been incorporated to replace the Casson model in the present viscosity model, it is still unable to accurately describe Malaysian waxy crude oil behavior, especially to address the steep change. For this evaluation, the performance of all models is showcased using Fluid S06 at low shear rates of 100 s⁻¹. Additionally, with adjustment, both the Pedersen and Ronningsen model as per Eq. (2) and Eq. (20), performed comparably well as illustrated in Fig. 18 although it shows a big difference at low temperatures.

The second hypothesis that can be deduced is the suitability of the Richardson oil/water emulsion viscosity model embedded in the present model on Malaysian waxy crude oil application. According to Pedersen and Ronningsen model, the Richardson model can be described below [4].

$$\eta = \eta_c exp(D\theta) \tag{22}$$
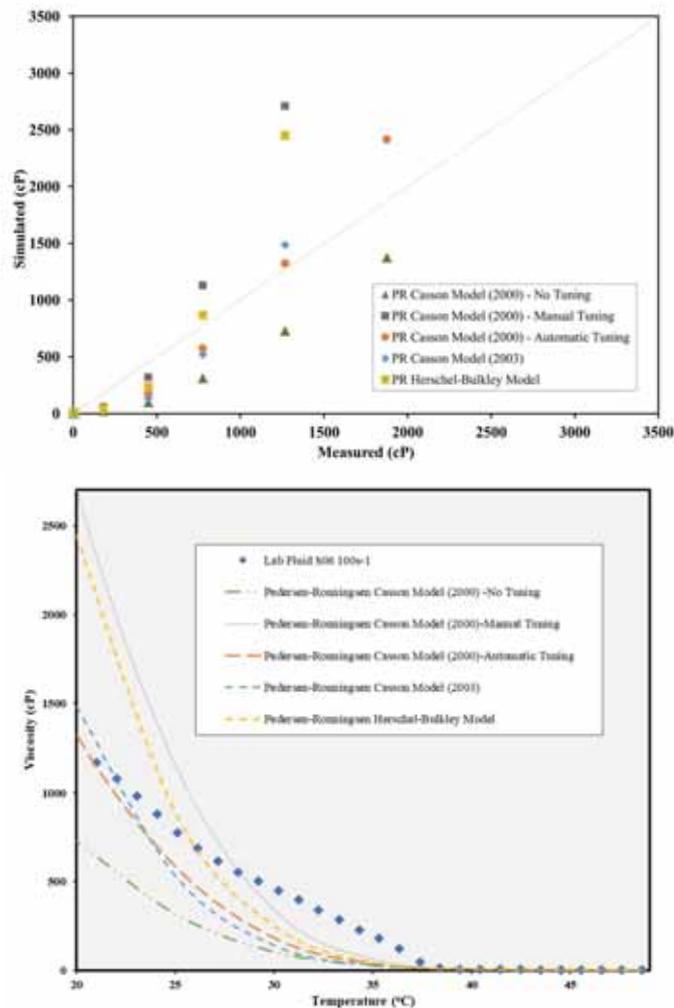
Fig. 18: Viscosity prediction at 100 s⁻¹ based on various models.

The Richardson model applies an exponential relationship to define the viscosity of an oil/emulsion, which could be main reason why Malaysian waxy crude oils suffer for over

and under-prediction when relying on present viscosity models to describe both Newtonian and non-Newtonian behavior of the fluid. Fig. 19, Fig. 20 and Fig. 21 were plotted to showcase the exponential behavior of the 18 North Sea oils with the highest WAT of 50 ℃ and PPT of 32 ℃ as listed in Pedersen and Ronningsen model. Sixteen out of 18 data sets show relationship that accurately matched exponential behavior, as described by Richardson and this regression model has been the baseline for the Pedersen and Ronningsen model.

Six Malaysian crude oil viscosity distribution data sets and exponential behaviors are plotted in Fig. 22 and Fig. 23 which generally showcases the behavior of the Malaysian waxy crude oils that do not follow the exponential relationship as described, especially at low shear rates and at temperatures below WAT and close to PPT, therefore, from this observation, it can be concluded that the present model is highly crude-specific and may not be suitable for Malaysian waxy crude oils wax deposition modelling application. Even with the HB model adjustment, modified Pedersen and Ronningsen model is still unable to showcase an acceptable viscosity trend. Due to the steep and instantaneous change of viscosity behavior below WAT and PPT, it is difficult to tune the viscosity model in fluid simulator for the purpose of wax deposition modelling. The present model has not been able to accurately describe the non-Newtonian viscosity behavior for Malaysian waxy crude oils, which in return will impact the accuracy of the wax deposition modelling.
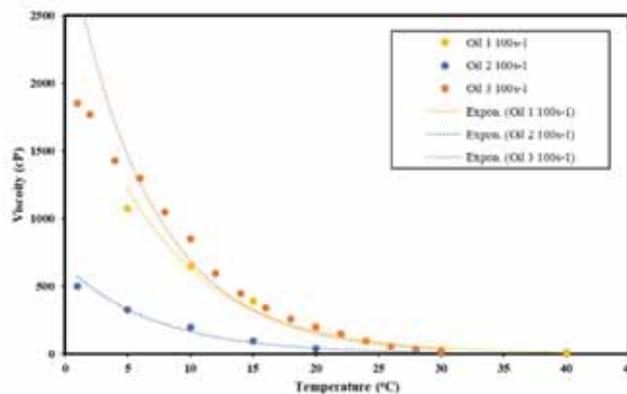
Fig. 19: Viscosity Distribution based on Pedersen and Ronningsen model (Oil 1, Oil 2, and Oil 3).
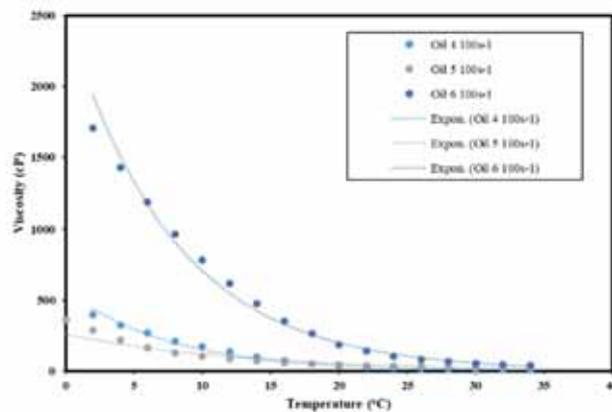
Fig. 20: Viscosity Distribution based on Pedersen and Ronningsen model (Oil 4, Oil 5, Oil 6, and Oil 7)
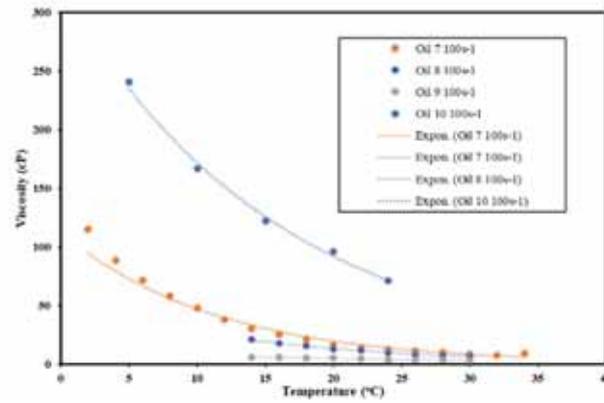
Fig. 21: Viscosity Distribution based on Pedersen and Ronningsen model (Oil 8, Oil 9, and Oil 10).
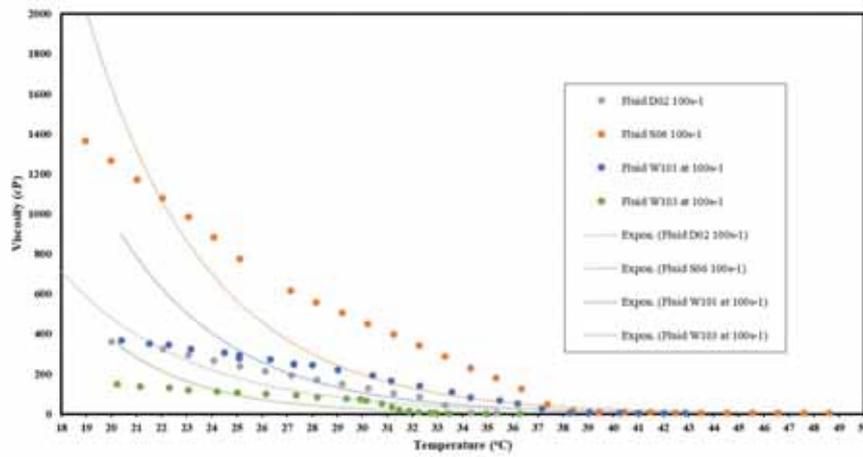
Fig. 22: Viscosity exponential relationship of Malaysian waxy crude at 100 s^{-1}.
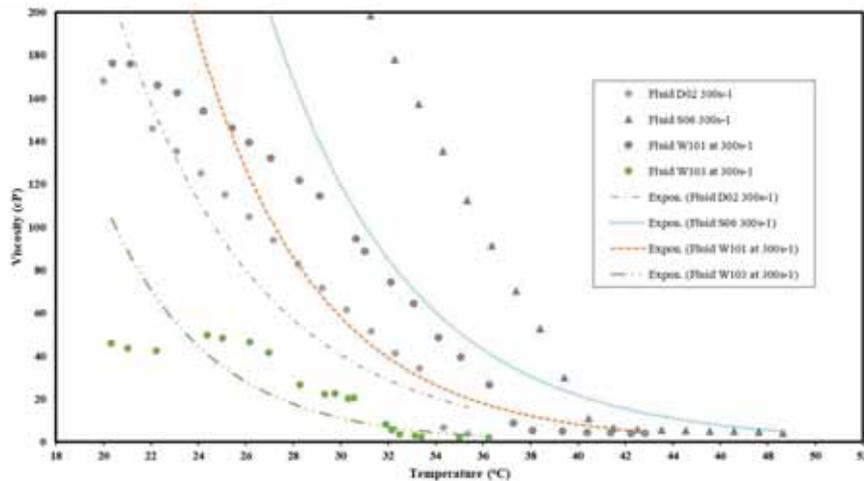
Fig. 23: Viscosity exponential relationship of Malaysian waxy crude at 300 s^{-1}.

It is worth noting that the present model demonstrates viscosity prediction dependencies to wax content and shear rates, but viscosity of crude oil can be influenced by many other factors including pressure, temperature, crude oil composition and presence of other compounds [20]. The effect of pressure on viscosity of the reservoir fluid is commonly demonstrated as part of PVT testing. With an increase of pressure, viscosity of crude oil reduces which would help in transporting fluid in the reservoir/tubing/pipeline, however it would also incur other risks (e.g., safety when operating at high pressure).

As presented in the Section 3.1, viscosity measurement increases with an increase in temperature, and more prominently as the fluid changes from Newtonian to non-Newtonian region. In fluid flow simulators, wax viscosity is calculated at fluid temperature, hence it is crucial to get a representative thermal-hydraulic flow profile in the pipelines. The limitations of the present model to match Malaysian crude oil experimental data specifically at low temperatures (e.g. up to 4 times underprediction for Fluid D02) would contribute to higher facilities cost to manage wax.

Crude oil composition as well as the other components in the crude oil play the most important roles in viscosity. Santos et al., [20] summarized a few findings from various researchers highlighting the impact of asphaltene concentration on the viscosity, which shows that viscosity can increased sharply with the presence of higher asphaltene content. Malaysian crude oil however, showcased a very low asphaltene content as reported by Sulaimon and Yusoff [8] ranging from 0.13-0.34 wt% which is hypothesized may not have impacted viscosity measurement greatly. It is however a good practice to conduct SARA analysis as part of fluid property screening to determine asphaltene content prior to conducting a more detailed analysis such as dynamic viscosity measurement.

Nevertheless, the present model is deemed to be acceptable when dealing with high shear rates and at higher temperatures as observed in the simulation cases. It is very prudent to determine pipeline specific shear rates, and when dealing with lower ambient temperatures, the prediction error and its impact needs to be accounted for in the decision making.

As a way forward, a more reliable model such as that proposed by Bhaskoro et al., [13] that were developed based on Malaysian waxy crude oil should be used in engineering applications. As described in his paper, above WAT, viscosity of waxy crude oil follows Arrhenius Law, while below WAT, the viscosity deviates from Arrhenius behavior up to several orders of magnitude due to the precipitated wax crystals. The general temperature-dependent viscosity model was developed by modifying the HB model and incorporating the critical physico-chemical properties dictating the viscosity behavior which are activation energy of viscous flow, molecular weight, total wax content, and amount of precipitated wax. The model as described has showcased its superior capability in predicting both Newtonian and non-Newtonian viscosity of the waxy crude oils without rheological data input specifically on its application for waxy crude oils from the Malay basin. The same Malaysian crude oil, Fluid S06 (which is represented as Crude oil A in his paper) has been tested, and its performance is presented below [13]. Generally, Fig. 24, as presented in Bhaskoro et al., [13] shows that the model can match experimental datasets better when compared to the present model available in commercial simulators.
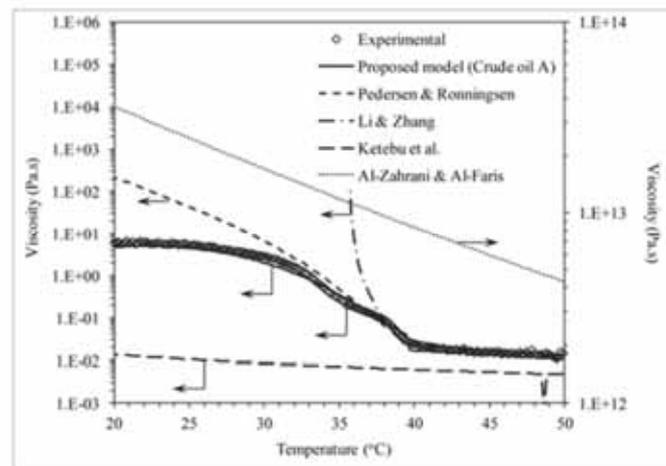
Fig. 24: Performance of Bhaskoro general viscosity model as compared to various published viscosity models.

At the time of this study, however, the said model is not available within the fluid and flow simulators being used. Additionally, although it has modified the exponential terms and developed a generally more suitable model for Malaysian waxy crude oils, this model is still unable to match the viscosity trend specifically at the transition regions (below WAT and close/below PPT).

Ideally, whenever possible, the best way is to enable direct input of measured viscosity data to be used in wax deposition simulation. This can help directly in reducing prediction errors. It should be noted that for pipelines with a high likelihood to experience rapid dynamic changes in flow as well as having wider temperature ranges, the use of the present non-Newtonian viscosity model would not be able to describe the viscosity behavior accurately matching the experimental data. Most of the crude oils shown in this example, are being produced from shallow water depth with a seabed temperature range of 20 to 26 °C and ambient temperature ranging from 30 to 36 °C, therefore as a way forward for the wax deposition modelling work, manual tuning is conducted for the viscosity model focusing on tuning at specific temperature ranges and interested shear rates. Additionally, more attention should be given when dealing with viscosity at low shear rates.

5. CONCLUSION

The Pedersen and Ronningsen model is generally unable to accurately describe the non-Newtonian viscosity behavior for the six (6) samples of Malaysian waxy crude oils, especially at low shear rates and low temperatures ranges due to immediate transition/instantaneous change of viscosity behavior below WAT and PPT. From the analysis, it is observed that viscosity prediction inaccuracy could lead to wax deposition prediction error of up to 336%. This would contribute to high risk in wax management especially when flow simulators are used to troubleshoot issues at the field.

Manual tuning is proposed to be performed focusing only on regressing the Pedersen and Ronningsen model to match experimental viscosity datasets at the specific shear rate and specific temperature ranges like what is experienced by the pipelines. Most of the crude oils shown in this paper are being produced from shallow water depth with seabed temperature range of 20 °C to 26 °C hence it is critical to match at lower temperature ranges.

Ideally, the best method is to enable direct input of experimental viscosity data sets for the wax deposition modelling hence this feature should be made within current commercial engineering software. Even though that Bhaskoro et al. [13] have tried to improve the model that suits Malaysian waxy crude oils, this model is yet to be made available in commercial applications for use by FA engineers. Additionally, it is worth noting that this paper also demonstrates that for Malaysian waxy crude oils, one cannot rely on modelling prediction alone without regression to measured viscosity data. Getting measured viscosity data sets at various conditions are still required. This test is not a tedious test, rather it can be done within hours and minimal sample volumes are required.

Further improvement on the wax viscosity model is required to describe Malaysian waxy crude oil non-Newtonian behaviors. Incorrect viscosity predictions will lead to difficulties when tuning the wax model and might lead to unphysical behaviors of the wax deposition. Additionally, to enable robust wax deposition modelling for field application, it is crucial to enable correct representation of each parameter that governs the wax deposition prediction in the pipelines.

ACKNOWLEDGMENT

REFERENCES

[1] Zhang X, Pedrosa N, Speranza A, Queimada A, Richard S. (2019) Impact of fluid flow and thermodynamic wax models on multiphase wax deposition simulation. Paper presented at the BHR 19th International Conference on Multiphase Production Technology, Cannes, France, June 2019.

[2] Soedarmo AA, Daraboina N, Sarica C. (2017) Validation of wax deposition models with recent laboratory scale flow loop experimental data. Journal of Petroleum Science and Engineering, 149: 351-366. https://doi.org/10.1016/j.petrol.2016.10.017

[3] Singh A, Lee H, Singh P, Sarica C. (2011) Flow assurance: Validation of wax deposition models using field data from a subsea pipeline. Offshore Technology Conference. https://doi.org/10.4043/21641-ms

[4] Pedersen KM, Rønningsen HP. (2000) Effect of Precipitated Wax on Viscosity: A Model for Predicting Non-Newtonian Viscosity of Crude Oils. Energy & Fuels, 14(1): 43-51. https://doi.org/10.1021/ef9901185

[5] Matzain AB. (1999) Multiphase Flow Paraffin Deposition Modeling (Doctoral dissertation). University of Tulsa, Oklahoma.

[6] Giacchetta G, Marchetti B, Leporini M, Terenzi A, Dall'acqua DSV, Capece L, Grifoni, RC. (2017) Pipeline wax deposition modeling: A sensitivity study on two commercial software. Petroleum, 5(2): 206-213. https://doi.org/10.1016/j.petlm.2017.12.007

[7] Coronado O, Kak A. (2017) Workflow to Evaluate Wax Deposition Risk Along Subsea Production Systems. Offshore Technology Conference. https://doi.org/10.4043/27855-ms

[8] Sulaimon AA, Yusoff ME. (2015) Wax and Asphaltene Deposition Tendency of Malaysian Crude Oils. Proceedings of the International Conference on Integrated Petroleum Engineering and Geosciences. https://doi.org/10.1007/978-981-287-368-2_15

[9] Petrus TB, Azuraien J. (2014) Rheological Measurement of Waxy Crude Oil under Controlled Stress Rheometer: Determination of the Setting Parameters. Journal of Applied Sciences, 14(18): 2047-2053. https://doi.org/10.3923/jas.2014.2047.2053

[10] Anuar A, Rasib AA, Norpiah RM, Ali Z. (2022) Pigging the Unpiggable: A case study of waxy crude oil pigging simulation and field implementation. SPE Asia Pacific Oil & Gas Conference and Exhibition. https://doi.org/10.2118/210724-ms

[11] Vinay G, Wachs A, Agassant JF. (2006) Numerical simulation of weakly compressible

Bingham flows: The restart of pipeline flows of waxy crude oils. Journal of Non-Newtonian Fluid Mechanics, 136: 93-105. https://doi.org/10.1016/j.jnnfm.2006.03.003

[12] Bhaskoro PT. (2017) Viscosity prediction for waxy crude oil – Effects of physico-chemical parameters (Doctoral dissertation). Universiti Teknologi PETRONAS, Perak, Malaysia

[13] Bhaskoro PT, Japper@Jaafar A, Sariman MZ, Norpiah R, Shafian SR. (2020) Viscosity Model of Waxy Crude Oil Based on the Physico-Chemical Properties. Offshore Technology Conference Asia. https://doi.org/10.4043/30084-ms

[14] Barnes HA. (1997) Thixotropy - a review. Journal of Non-Newtonian Fluid Mechanics, 70 (1–2): 1-33. https://doi.org/10.1016/s0377-0257(97)00004-9

[15] Yoshimura A, Prud Homme RK. (1988) Wall slip corrections for couette and parallel disk viscometers. Journal of Rheology, 32(1): 53-67. https://doi.org/10.1122/1.549963

[16] Japper-Jaafar A, Bhaskoro P, Sean L, Sariman M, Nugroho H. (2015) Yield stress measurement of gelled waxy crude oil: Gap size requirement. Journal of Non-Newtonian Fluid Mechanics, 218: 71-82. https://doi.org/10.1016/j.jnnfm.2015.02.001

[17] PVTsim Nova Method Documentation

[18] OLGA Method Documentation

[19] Pedersen KS. (2003) Influence of Wax Inhibitors on Wax Appearance Temperature, Pour Point, and Viscosity of Waxy Crude Oils. Energy & Fuels, 17(1): 321-328. https://doi.org/10.1021/ef020142

[20] Santos IC, Oliveira P, Mansur CRE. (2017) Factors that affect crude oil viscosity and techniques to reduce it: A review. Brazilian Journal of Petroleum and Gas, 11(2): 115-130. https://doi.org/10.5419/bjpg2017-0010

[21] Wilke, C.R. and Chang, P. (1995) Correlation of Diffusion Coefficient in Dilute Solutions. AIChE Journal, 1(2):264-270.https://doi: 10.1002/AIC.690010222

BEHAVIORS OF THERMOELASTIC PROPERTIES IN NI-TI-BASED SHAPE MEMORY ALLOYS, PROCESSED BY METAL FORMING TECHNIQUES

TARSILA TENÓRIO LUNA DA SILVA[1*], FILLIPE VIRGOLINO[2], NIÉDSON SILVA[3], CARLOS OLIVEIRA[4], CARLOS ARAÚJO[5], OSCAR OLIMPIO DE ARAUJO FILHO[4] AND CEZAR GONZALEZ[4]

[1]*Departamento de Engenharia Mecânica (DEM), Universidade Federal da Bahia (UFBA), Salvador, Brazil*
[2]*Departamento de Engenharia Mecânica (DEMEC), Universidade de Pernambuco (UPE), Recife, Brazil*
[3]*Departamento de Engenharia Mecânica (DEMEC), Instituto Federal de Pernambuco (IFPE), Caruaru, Brazil*
[4]*Departamento de Engenharia Mecânica (DEMEC), Universidade Federal de Pernambuco (UFPE), Recife, Brazil*
[5]*Departamento de Engenharia Mecânica (DEMEC), Universidade Federal de Campina Grande (UFCG), Campina Grande, Brazil*

[*]Corresponding author: tarsila.silva@ufpe.br

ABSTRACT: In this work, the thermoelastic properties of Ni-Ti shape memory alloys (SMA) processed by conventional rolling and equal channel angular extrusion (ECAE) were investigated. SMAs have two phases: Austenite (at high temperature) and Martensite (at low temperature). The samples were compared under five different thermal and processing conditions: homogenized, rolled, rolled-annealed, extruded, and extruded-annealed. The homogenized sample served as a reference. The samples were analyzed by differential scanning calorimetry (DSC) to determine the thermoelastic transformation temperatures. Images were taken using scanning electronic microscopy (SEM) in conjunction with energy dispersive spectroscopy (EDS). The dynamic area was completed for two tests: under constant load bending (simulation of the memory effect to determine the reversible thermoelastic strain) and dynamic mechanical analysis (DMA). The results showed that the plastic forming processes alter the properties, especially for samples exposed to the ECAE, which can block the martensitic phase. However, R-phase (a rhombohedral phase), that can appear at low temperatures before the martensitic phase, emerges totally when the extruded sample suffers annealing. The images of SEM, confirmed by EDS, show that any type of forming process and the presence of precipitates have a significant influence on the behavior of the elastic property. It was found that extrusion has a greater effect on the restoring properties of the alloys than rolling. This analysis is of great importance for the use of SMA in applications requiring high mechanical strength combined with the functional properties of shape recovery through martensitic phase transformations.

ABSTRAK: Kajian ini adalah berkaitan sifat-sifat bentuk aloi ingatan (SMA) termoelastik Ni-Ti yang diproses melalui penggelek konvensional dan penyemperitan sudut saluran sama (ECAE). SMA mempunyai dua peringkat: Austenit (pada suhu tinggi) dan Martensit (pada suhu rendah). Sampel dibandingkan pada lima tahap kepanasan dan proses iaitu: percampuran, penggulungan, penggulungan-rataan, perataan dan penyemperitan-rataan. Sampel campuran yang dihomogenkan dijadikan sebagai sampel rujukan. Sampel dianalisis dengan pengimbas kalorimetri pembezaan (DSC) bagi menentukan suhu

transformasi termoelastik. Imej diambil menggunakan pengimbas mikroskop elektronik (SEM) bersama spektroskopi penyebaran tenaga (EDS). Kawasan dinamik diuji dengan dua ujian: di bawah lenturan beban malar (simulasi kesan memori bagi menentukan terikan termoelastik boleh balik) dan analisis mekanik dinamik (DMA). Dapatan kajian menunjukkan bahawa proses pembentukan plastik telah mengubah sifat, terutama pada sampel yang terdedah kepada penyemperitan sudut saluran sama ECAE, yang boleh menyekat fasa martensit. Walau bagaimanapun, fasa-R (fasa rombohedral) yang boleh muncul pada suhu rendah sebelum fasa martensitik, muncul sepenuhnya apabila sampel tersemperit mengalami penyepuhlindapan (penyemperitan-rataan). Imej pengimbasan mikroskop elektron, seperti yang dibuktikan dengan spektroskopi penyebaran tenaga (EDS), menunjukkan bahawa apa-apa jenis proses pembentukan dan kehadiran mendakan mempunyai pengaruh kuat terhadap sifat elastik. Dapatan kajian juga mendapati bahawa penyemperitan mempunyai kesan yang lebih besar terhadap sifat pemulihan aloi berbanding proses penggulungan. Analisis ini sangat penting bagi penggunaan bentuk aloi ingatan (SMA) dalam aplikasi yang memerlukan kekuatan mekanikal yang tinggi bersama sifat pemulihan bentuk melalui transformasi fasa martensit.

KEY WORDS: *equal channel angular extrusion; shape memory alloy; Ni-Ti alloys*

1. INTRODUCTION

Shape memory alloys (SMA) are active materials that can restore plastic deformations through suitable thermal conditioning. The phenomenon of shape memory effect (SME) is closely related to martensitic phase transformations, which are crystallographically reversible [1]. Figure 1 summarizes the shape memory effect when comparing an ordinary plastic deformation, a superelastic deformation and a shape memory deformation.
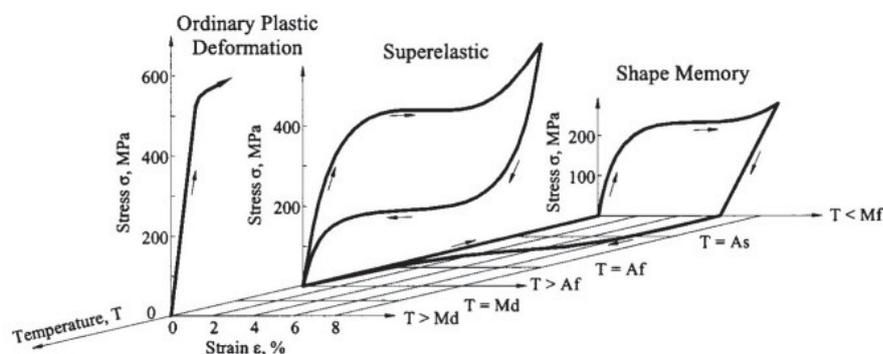
Fig. 1: Comparison between the diagram "stress x strain" in 3 forms: ordinary plastic deformation, superelastic deformation and deformation with shape memory [2] *apud* [3].

SMAs have two phases: Austenite (at high temperature) and Martensite (at low temperature). The phase transformation temperatures are austenite onset (As), austenite end (Af), martensite onset (Ms), martensite end (Mf) and martensite transformation temperature by deformation (Md). Figure 1 shows that superelastic and shape memory deformations can be as high as 8%. The SMAs can also deform with a temperature difference [2]. The fabrication of Ni-Ti family SMAs is of great importance because slight variations in chemical composition can change the critical temperatures of martensitic transformation to about 100 °C [2]. Therefore, plasma is efficiently used for the manufacturing process as it leads to negligible losses due to oxidation during melting [4]. Ni-Ti based SMAs are widely used due to their wide range of technological applications, ranging from aerospace and biomedical to petrochemical and other industries [5,6].

The SMA can be considered an electric drive if the SME is activated by the application of current pulses (Joule effect). On the other hand, if the SME is activated directly by controlled temperature changes, they are considered thermal actuators. When these alloys are subjected to plastic deformation processes, they exhibit increased mechanical strength due to their hardening (significant increase in the density of crystalline defects) and reduced grain size after deformation and heat treatment. This change in mechanical properties leads to changes in the thermoelastic properties of SMA (critical transformation temperatures, transformation enthalpies, thermoelastic deformation, thermal hysteresis and others), which are essential characteristics for the use of these materials in technological applications [7,8].

Equal channel angular extrusion (ECAE) or equal channel angular pressing (ECAP) is a metal-forming process recently developed by Segal et al. [9]. The process promotes the passage of material between two successive channels that intersect, forming angles between 90° and 120° and causing plastic deformation. These channels have identical geometric dimensions so that material does not change its dimensions after processing. In this process, the material undergoes several plastic deformations that cause high critical shear stress without changing the cross-section of the sample. Some studies show that the deformation phenomena during the ECAE process are independent of the sample size [6]. Several technological applications of these processed materials are already idealized, especially for parts that need to have a high mechanical strength obtained by plastic deformation processes without changing their dimensions during processing.

Sun et al. [10] examined the properties and hardening behavior of ECAE-processed Mg-Al binary alloys. As a result, the dynamic compressive test shows a significant strength increase attributed to the precipitates and the strength improvement from grain size refinement (Hall-Petch effect). Yang et al. [11] processed Ni-Ti-Nb SMAs by ECAP. Ni-Ti-Nb is a wide hysteresis shape memory alloy with great application potential in aerospace. The ECAP process decreased grain size (or the increase of grain boundary content), and the transformation hysteresis temperature and recovery stress of the Ni-Ti-Nb alloy increased simultaneously. Ni-Ti shape memory alloys are widely used for medical components, as they can accommodate large strains in their superelastic state. To further improve the mechanical properties of Ni-Ti, grain refinement by severe plastic deformation is applied to generate an ultrafine-grained microstructure with increased strength. Leitner et al. [12] performed experiments in ultrafine-grained Ni-Ti, processed by ECAP, and comprehensive fracture and fatigue crack growth experiments to assess its damage tolerance, which is essential for safely using this material in medical applications.

Other methods of plastic forming were studied. Jiang et al. [13] studied the mechanically induced martensite in Ni-Ti-Fe alloys, with shape memory effect, manifesting the suggested in the deformation plane, concluding that the plastic deformation necessary to produce the product of the dislocations plays an essential role in the formation of mechanically induced martensitic transformation, unlike stress-induced martensite where control of dislocations only occurs after deformation. Zhang et al. [14] subjected Ni-Ti-based superelastic alloys to severe cold rolling and observed that stress-induced martensite reorientation (B19' phase) dominates the reaction. Liang et al. [15] described the martensitic strain states of cold-rolled Ti50.8at%Ni alloys for different thickness reduction levels. Silva [16] rolled two Ni-Ti wires with a shape memory effect. Both wires showed important characteristics and high performance, such as good responses to several cold rolling, low thermal hysteresis, reasonable dimensional control, nanostructure, and excellent properties mechanics. It can be used in highly complex engineering applications requiring precision in actuation and control.

This study aims to obtain a Ni-Ti alloy with an SME and subject these materials to mechanical forming processes by conventional rolling and ECAE. In the work, the properties of Ni-Ti SMA were analyzed using DSC analysis, SEM-EDS and a constant bending load test. The latter test simulates the shape memory effect and the thermoelastic properties are determined from the curves obtained.

2. MATERIALS AND METHODS

A SMA of the equiatomic composition of Ni-50at.%Ti, corresponding to Ni-45wt.%Ti, was obtained by the plasma skull-push pull (PSPP) process with a controlled atmosphere melting chamber. A plasma flame of a tungsten electrode melts the metallic material. Plasma energy is enough for melting most metallic alloys cleanly and efficiently. The furnace chamber receives high pressure by vacuum in its mold cavity, continuously filling the mold in a push-pull sequence. The machine has a mode control that determines time and temperature according to the alloy [17,18]. It was found that to ensure good homogeneity of the SMA, the melting process should be repeated five times. Heat treatments were made in a Jung muffle furnace, model LF4212, followed by cooling in water. The ingot was homogenized for 120 min in a muffle furnace at 900 °C, followed by quenching in water at room temperature (~25 °C).

The ingot was cut on a precision cutter with a diamond disk to produce samples as prismatic bars for metal forming processes: conventional mechanical rolling process and ECAE. The sample used in the rolling process has 5×1.5 mm². Hot rolling was performed on a conventional rolling machine. After each step, the sample was heated in a muffle furnace at 900 °C for 1 min. The sample was deformed from 1.5 mm to 0.5 mm, with a final deformation rate of 66.7% (20 steps with 0.05 mm for each step). For the angular extrusion, the sample was a bar with a 5×5 mm² square cross-section. The die for the angular extrusion process has an angle of 130° between internal channels. Angular extrusion was performed on a universal mechanical testing machine (EMIC), controlled to operate in a compression test mode with 0.5 mm·min⁻¹ speed at room temperature (~25 °C).

This work obtained samples after homogenization (using a sample without plastic deformation as a reference), mechanical rolling, and ECAE. The samples obtained for each plastic deformation process were subjected to annealing at 450 °C for 15 min and cooled in water at room temperature for stress relief. Table 1 presents the nomenclature and procedures applied to each of the samples analyzed in this study.

Table 1: Nomenclature and procedures of the Ni-Ti samples

Sample	Processes applied
H	Homogenized at 900 °C for 120 min
RL	Rolled
RA	Rolled and annealed at 450 °C for 15 min
E	Extruded
EA	Extruded and annealed at 450 °C for 15 min

The calorimetric properties (temperatures and enthalpies of austenitic, martensitic, and rhombohedral phase transformations and their thermal hysteresis) were performed by a Mettler TA3000 DSC. The DSC technique is very effective for determining the record of phase transformation temperatures in LMF. The advantage of this technique is the possibility of measuring the energy involved in the process based on measuring the area of

the endothermic peak in heating and the exothermic peak in cooling. Even so, the transformation enthalpies, thermal hysteresis (difference between the transformation peak temperatures), and phase transformation temperatures (As, Af, Ms, Mf measured by crossing the tangents at the peaks) are measured. This technique also allows obtaining graphs, in which it is possible to analyze whether the phase transformation occurs in one or two steps [19,20]. The DSC heating chamber was filled with nitrogen gas to prevent the oxidation of the material. The samples were analyzed over a temperature range of -60 °C to 120 °C, with a heating/cooling rate of 10 °C/min.

The DMA experiment consisted of applying an oscillating force to a sample in uniaxial tension mode and investigating the response of the material to this applied load. The equipment allows working between 25 °C and 500 °C in bending, tensile and uniaxial cantilever (curved at one end and free at the other) modes [21,22].

In this work, DMA tests were performed on equipment installed in an anti-vibration bed to avoid interference during data acquisition. The experiment was conducted in single cantilever mode, following standard parameters: frequency of 1 Hz, a heating rate of 2 °C/min and a vibration amplitude of 5×10^{-3} mm [23]. The SMA Ni-Ti slides used had a useful length and width of approximately 17 and 5 mm, respectively. Prior to all tests, the samples were cooled to temperatures of about -30 °C to collect data on the reverse martensitic transformation upon heating. As with the DSC tests, the transformation temperatures obtained from the DMA curves were determined using the tangent crossing method. Conventional metallographic procedures (grinding and polishing with diamond paste) were performed and Kroll reagent (2 mL HF, 4 mL HNO_3, and 100 mL H_2O) was used for etching. The microstructures of the samples were recorded with a Hitachi TM3000 SEM. Together with the SEM-EDS tests are carried out to semi-quantitatively check the chemical composition of the sample.

A special apparatus was developed to determine the thermoelastic properties of SMA under constant bending load (cantilever test). The load is applied to the sample through a pulley system. A programmable silicon oil bath is used to perform heating and cooling cycles between -10 °C and 100 °C. During cooling, the plate (sample) is deformed, and when heated, the sample regains its original shape. A Linear Variable Displacement Transducer (LVDT) and a thermocouple measure the deformation and temperature, respectively [24].

In this work, the samples were individually tested under a constant bending load of 80 MPa and thermally cycled 20 times between 20 °C and 100 °C with a rate of 10 °C.min^{-1} during heating and 3 °C.min^{-1} for cooling. The samples have prismatic bar geometry with dimensions of $20 \times 5 \times 0.5$ mm³.

Figure 2 shows representative thermoelastic deformation versus temperature curves obtained from data collected in the thermomechanical tests under constant load. There are two cycles, in which thermoelastic properties are determined. The martensitic transformation temperatures (austenite and martensite phases) can be found across the tangent line technique. Thermoelastic deformations (ε_t) and previous plastic strain (x) are found in the ordinate axes as indicated in Fig. 2. The thermal hysteresis ($Ht = A_{50} - M_{50}$) can be also calculated [25].
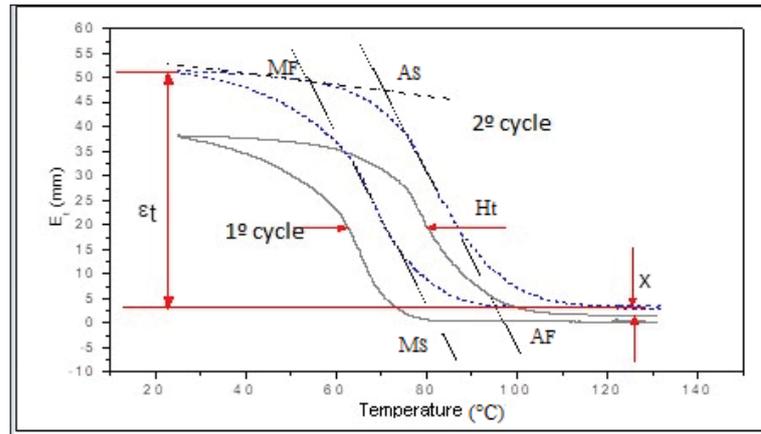
Fig. 2: Representative thermoelastic deformation versus temperature curves obtained during a typical test of thermomechanical cycles in simple bending mode. The investigated parameters are pointed out in this figure [25]. "Reproduced with permission from Gonzalez, C.H. and De Araújo, C.J, Study of Martensitic Stabilization Under Stress in Cu–Al–Be Shape Memory Alloy Single Crystal; published by Materials Science and Engineering, 2004".

3. RESULTS AND DISCUSSIONS

3.1 Calorimetric Curves (DSC)

Figures 3-6 show calorimetric curves for the samples submitted to the five conditions: homogenized (H), rolled (RL), rolled and annealed (RA), and extruded (E), respectively.

Figure 3 was used as the reference sample. In the case of the rolled sample (RL) (Fig. 4), the start temperature of the R-phase remains practically unchanged from to M_S temperature of the homogenized (H) sample (Fig. 3). The curve of the rolled sample (Fig. 4) indicates that the plastic deformation process causes a decrease in M_S temperature during cooling and A_S temperature is also reduced. In principle, it can be assumed that the initial inflection during cooling is due to the presence of the rhombohedral phase (R-phase) [1]. The plastic deformation process increases the linear defects density (dislocations) that may hinder or even inhibit direct (austenite => martensite) and reverses (martensite => austenite) martensitic transformations.
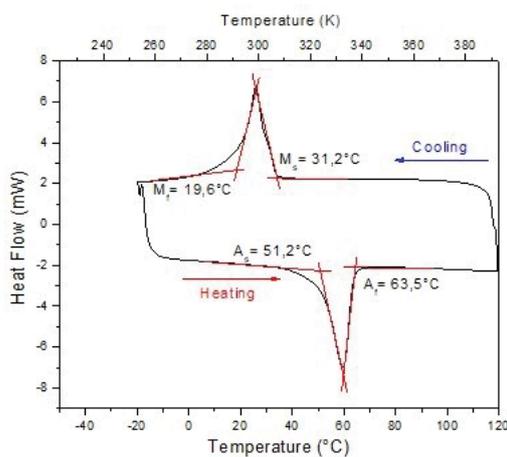
Fig. 3: Calorimetric curves with martensitic transformations of the Ni-Ti alloy homogenized (H).
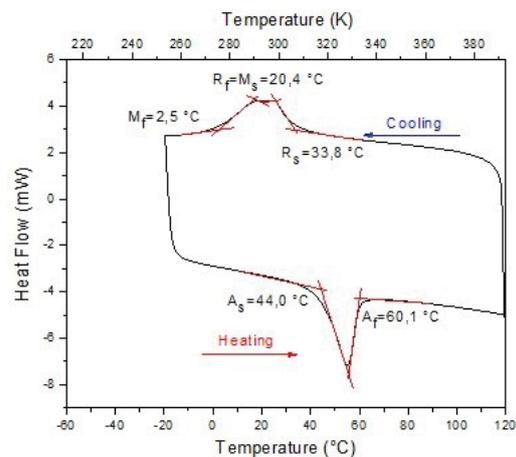
Fig. 4: Calorimetric curves with martensitic transformations of the Ni-Ti alloy rolled (RL).
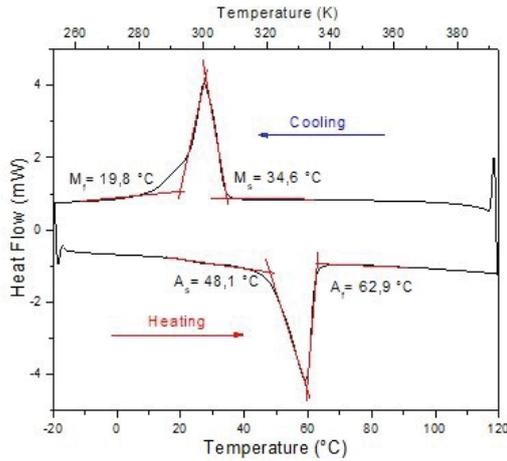
Fig. 5: Calorimetric curves with martensitic transformations of the Ni-Ti alloy rolled and annealed (RA).
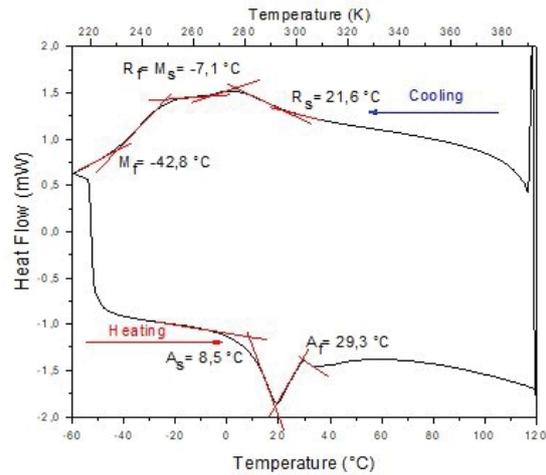
Fig. 6: Calorimetric curves with martensitic transformations of the Ni-Ti alloy extruded (E).

In the rolled and annealed (RA) sample (Fig. 5), it was observed that the calorimetric curve and its data are like the reference sample (Fig. 3), so this treatment eliminates the effects of the mechanical rolling process in this alloy.

The extruded (E) sample's (Fig. 6) DSC shows the peaks of the R-phase (rhombohedral) and monoclinic martensite phase (B19'), with lower temperatures of R_S and M_S than in the roller (RL) sample (Fig. 4). This clearly shows that the angular extrusion process introduces crystalline defects much larger than rolled process, however, it produces no reduction in the geometric dimensions of the sample.

Table 2 summarizes the main results obtained through the DSC test, the temperatures of the martensitic transformations, including the R-phase, the heating and cooling enthalpies, and the thermal hysteresis reaction.

Table 2: Martensitic transformations, enthalpies, and hysteresis of the Ni-Ti alloy samples

Sample	A_S [°C]	A_F [°C]	$\Delta H_{Heating}$ [J/g]	R_S [°C]	R_F [°C]	M_S [°C]	M_F [°C]	$\Delta H_{Cooling}$ [J/g]	H_T [°C]
H	51.2	63.5	10.3	-	-	31.2	19.6	10.9	32.3
RL	44.0	60.1	13.4	33.8	20.4	20.4	2.5	14.7	39.7
RA	48.1	62.9	15.0	-	-	34.6	19.8	13.4	28.3
E	8.5	29.3	17.2	21.6	-7.1	-7.1	-42.8	18.2	36.4
EA	26.8	42.4	17.0	31.3	14.0	-15.4	-35.5	16.0	57.8

In Fig. 7 the calorimetric curve of the extruded and annealed sample (EA) shows a complete separation between the peaks of the R-phase and the martensitic phase, with a larger temperature interval between peaks. This heat treatment was able to restore the R_S temperature of the R-phase, in this way, it was effective to eliminate (or minimize) the effects of the angular extrusion process in this crystal structure. For the martensitic phase, this treatment has become ineffective or even complicates the process to initiate the transformation.
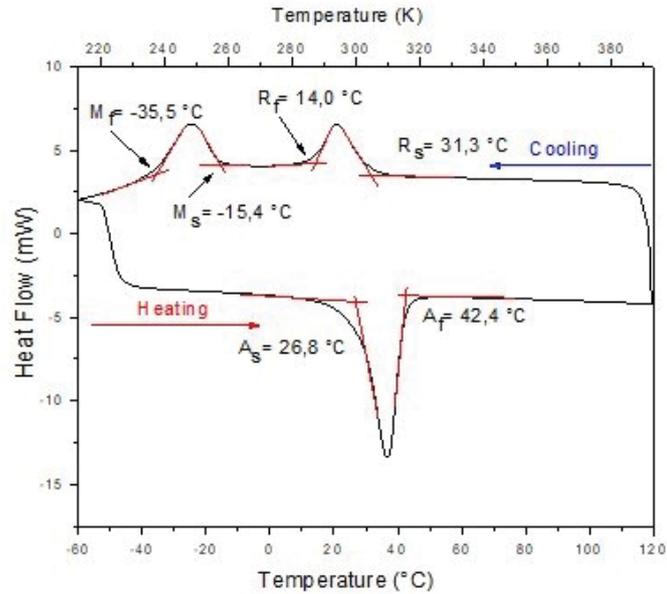
Fig. 7: Calorimetric curves with martensitic transformations of the
Ni-Ti alloy extruded and annealed (EA).

3.2 Thermomechanical Test

Figures 8-12 show the thermomechanical cycling curves of the Ni-Ti samples in the bending tests under constant load. Table 3 summarizes the results obtained in these tests.
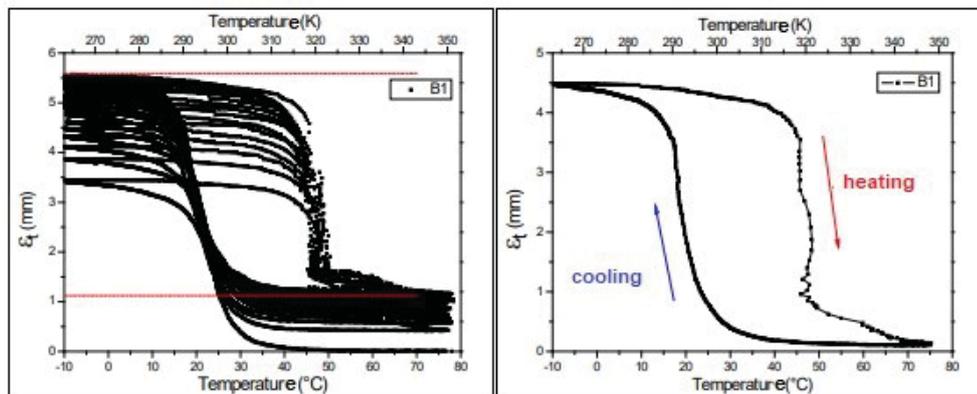
Fig. 8: Thermomechanical cycling curves of the Ni-Ti alloy homogenized (H).
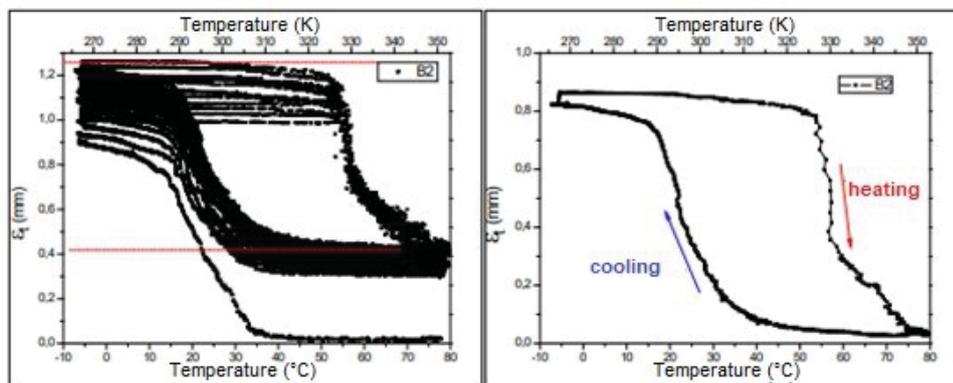
Fig. 9: Thermomechanical cycling curves of the Ni-Ti alloy rolled (RL).
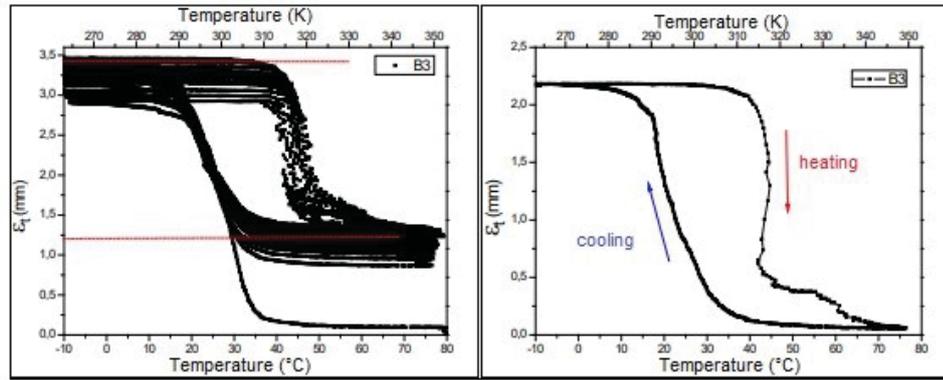
Fig. 10: Thermomechanical cycling curves of the Ni-Ti alloy rolled and annealed (RA).
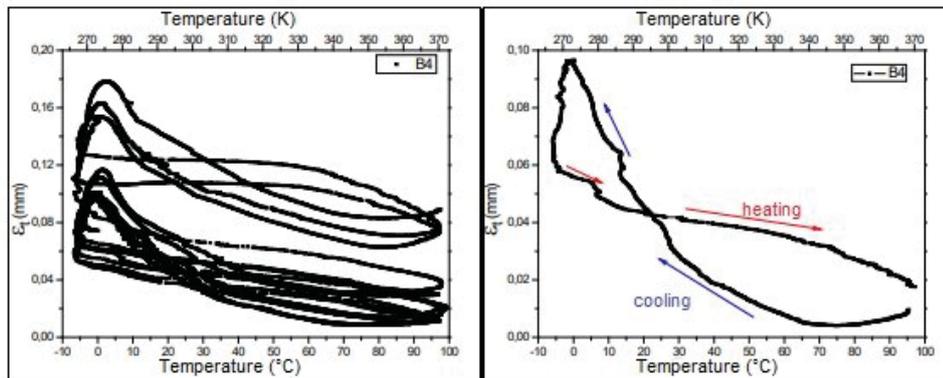
Fig. 11: Thermomechanical cycling curves of the Ni-Ti alloy extruded and annealed (EA).
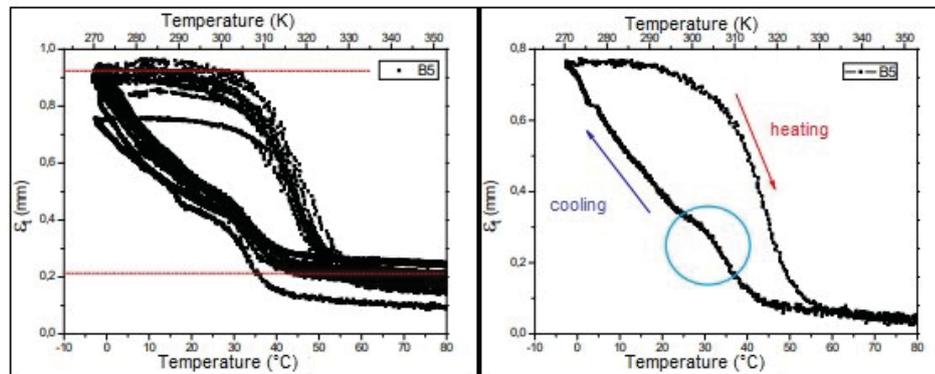
Fig. 12: Thermomechanical cycling curves of the Ni-Ti alloy in the extruded (E) condition.

Table 3: Results of the thermomechanical cycling curves of the Ni-Ti samples

Sample	A_s [°C]	A_F [°C]	R_S [°C]	$R_F = M_S$ [°C]	M_F [°C]	ε_T [mm]
H	44.9	49.8	-	26.5	15.1	4.50
RL	53.2	74.6	-	33.2	12.6	0.86
RA	42.9	50.2	-	30.9	15.1	2.21
E	-	-	-	-	-	0.09
EA	33.2	50.8	41.7	29.1	-2.6	0.74

In Fig. 8, the homogenized (H) sample had the maximum reversible thermoelastic deformation. In the calorimetric tests, the critical temperatures of phase transformation of the rolled and annealed (RA) sample (Fig. 10) returned to effectively the same temperatures

obtained for the reference sample (H), as seen in Table 2, which also shows the enthalpies of transformation. However, this treatment was not enough to restore the reversible thermoelastic deformation obtained in the reference sample, since the maximum value obtained for the annealed sample was lower than the reference sample (H). From the results of thermoelastic tests, this treatment appears to have not been able to eliminate all the effects of the rolling process.

The extruded (E) sample (Fig. 11) practically did not show any thermoelastic deformation. This was probably due to the limitation of the thermal bath, which has a minimum temperature of -10°C, but in these cases (Fig. 12) it is also possible that the energy provided by the system was not enough to complete the entire transformation due to the blocking provided by the large increase in dislocation density. When external stress is applied to a shape memory alloy it promotes the increase of the critical transformation temperatures. The load applied was not enough for these temperatures to increase sufficiently to be detected by the device. The extruded and annealed (EA) sample (Fig. 12) showed a low recovery deformation, indicating a high degree of hardening when compared with the reference sample (H).

The heat treatment and training process seriously affects the internal friction behavior in Ti-Ni alloys. These events have an importance on the reconfiguration of Ti-Ni alloy defects, inducing modification that might change the material properties such as memory effect, damping capacity, strength, hardness, and others [26-28].

3.3 DMA Analysis

Figure 13 shows the damping curves, during heating, provided by the dynamic tests. It can be noticed that the decrease in the energy absorption capacity of the shaped samples, especially in the sample that was submitted to the ECAE, sample (E). The mechanical rolling process generates elongated grains according to the deformation direction [29,30]. These elongated grains are in a preferential direction and the result of this is the improvement of the damping capacity for RL and RA samples. The random grains obtained in sample E induce more stress field dislocation blocking the phase transformation. The large damping capacity of the EA sample is due to heat treatment who is efficient to liberate blocking regions and dislocations.
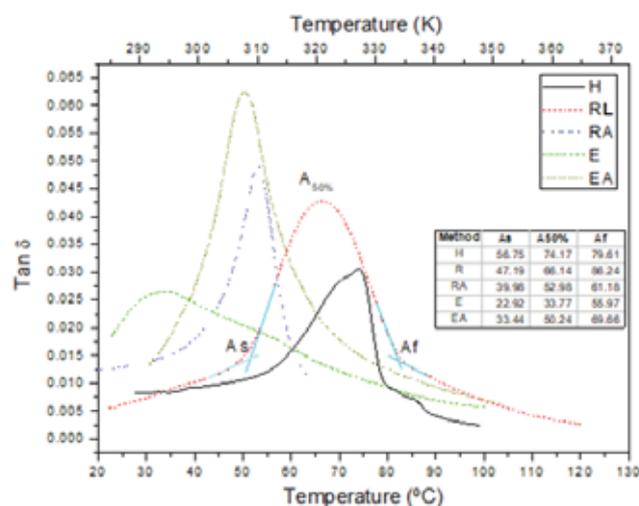
Fig. 13: DMA analysis. Temperature versus damping capacity in the conditions: homogenized (H), rolled (RL), rolled and annealed (RA), extruded (E), and extruded and annealed (EA).

According to the literature, Ni-Ti-based alloys have a high damping capacity during phase transformation and in the martensitic phase, whereas their austenitic phase shows a much lower energy dissipation [31]. The high damping capacity of the martensite phase is related to the hysteretic movement of interfaces (martensite variant interface and twin boundaries) [32,33]. It is established that the global internal friction is also controlled by dislocations and their interactions with other lattice defects. The high damping capacity of Ti-Ni alloys comes from the movement of several twin boundaries. The accommodation of twin boundaries under external stress is believed to be responsible for high damping capacity [34].

3.4 Micrography Analysis (SEM)

Figures 14-15 show micrographs of the samples. Figure 14a exhibits the microstructure of the homogenized sample and consists of equiaxial and heterogeneous grains with segregation as found in Jiang's et al. [35] work. Furthermore, as they exhibit a eutectic structure around grain boundaries, it was verified the regions with the highest percentage of Ni due to the exclusion of this element during solidification, confirmed by EDS [24].

Figure 14b and 14c show the rolled (RL) and the rolled and annealed (RA) samples. The arrows indicate this preferential orientation of the grains, which became more evident after the annealing heat treatment was carried out (Fig. 14c). Literature also suggests that with rolled process some particles of Ti_2Ni and others might be found in the grain boundaries provoking a second transformation in the grain boundaries and it is also related to R-phase transformation [36,37].
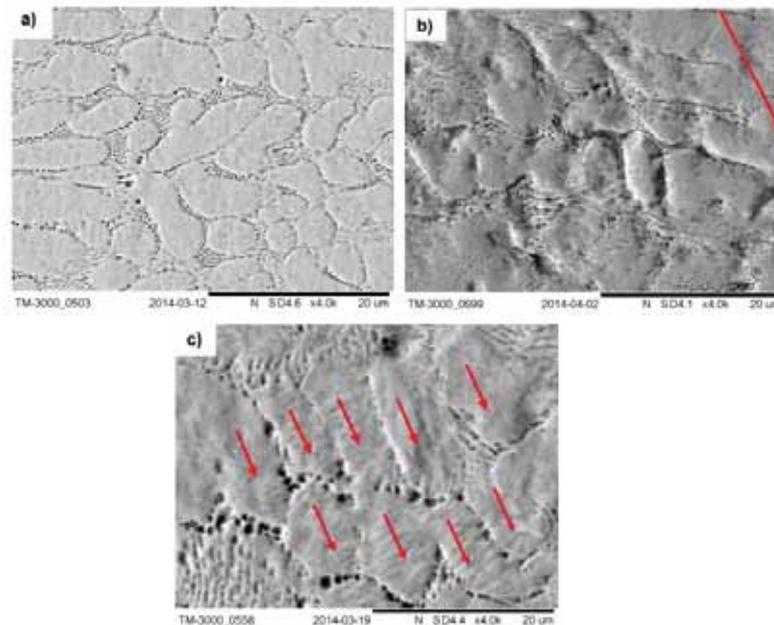
Fig. 14: SEM micrographs of samples: (a) homogenized (H), (b) rolled (RL) and (c) rolled and annealed (RA).

Figure 15 shows the microstructure for extruded (E) and extruded and annealed (EA) samples. In the images, it is possible to see the reduced grain compared with the homogenized (H) sample and particles provided by the high-stress state made by the mechanical process. These samples were subjected to severe shear deformations during the forming process, which can be confirmed by the reduction of transformation temperatures, damping capacity, and thermoelastic deformation. From the images, despite not having a

morphology with preferential orientation, it is possible to visualize the contours and decrease the grain sizes of the homogenized (H) sample. There is the presence of particles between the grains, which suggests that they are precipitated from the state of tension caused by ECAE.

Unlike samples annealed after conventional rolling, the micrographs of rolled and extruded (Fig. 15b) samples show that the latter process produces grains without preferential orientation. Probably, the configuration of the stress fields generated by the increased dislocation density must be highly disoriented. These facts seem to be responsible for the greatest loss of thermoelastic properties in the extruded samples.
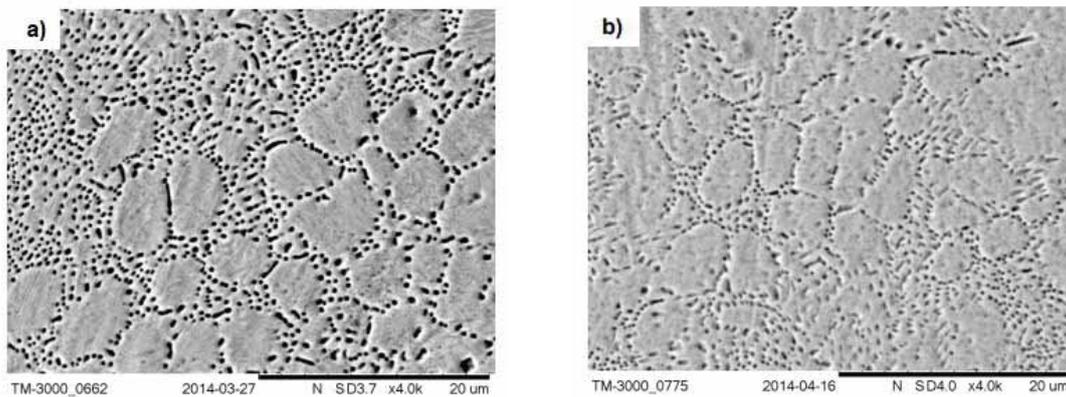
Fig. 15: SEM micrographs of samples: a) extruded (E) and b) extruded and annealed (EA).

3.5 Energy Dispersive Spectroscopy (EDS)

Figure 16 shows the analyzed regions of the sample as cast, where the percentage of elements in the alloy matrix (A), inside the grain (B), and in the region of its contour (C) was verified.
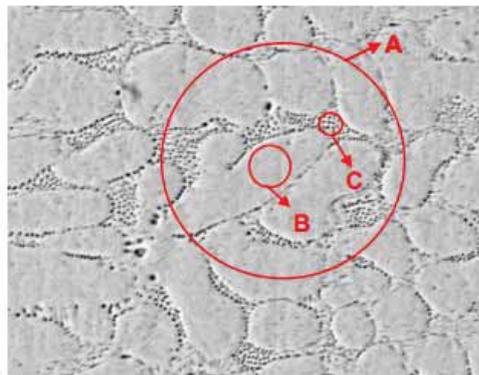
Fig. 16. EDS' analyzed regions: alloy matrix (A), inside the grain (B),
and in the region of its contour (C).

Figures 17, 18, and 19 showed results of each region introduced in Fig. 16. The percentage in weight is close to the nominal of the alloy, considering it to be a semi-quantitative analysis with a certain margin of error (around 5%). In turn, it was revealed the presence of a high percentage of nickel in the region along the grain boundaries, a probable result of the precipitation of this element during the solidification of the Ni-Ti alloys [38].
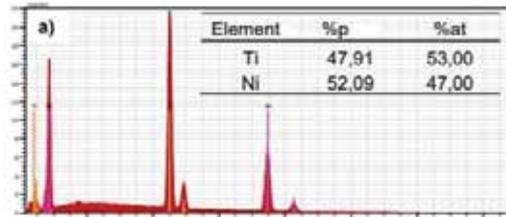
Fig. 17: EDS of area A showed in Fig. 16.
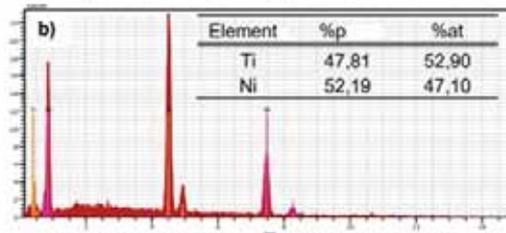
Fig. 18: EDS of area B showed in Fig. 16.
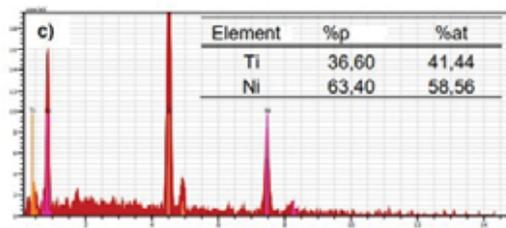
Fig. 19: EDS of area C showed in Fig. 16.

An EDS analysis for the equiatomic alloy showed a higher amount of Ni in the grain boundary region. Although the percentage of Ti is still lower, there is an increase in the detection of this element compared to the homogenized (H) sample, which suggests the presence of precipitates related to the mechanical conformation, Ti_2Ni.

4. CONCLUSION

Plastic forming processes by mechanical rolling and angular extrusion promoted increased mechanical resistance of samples due to the reduction of the grain size caused by shear stress during plastic deformation and introducing defects in the crystalline structure. Mechanical rolling and angular extrusion increased the mechanical resistance of samples by reducing the grain size and introducing defects in the crystalline structure. These processes have significantly affected the thermoelastic behavior of the shape memory alloy. Although one cannot compare directly (quantitatively) the influence of the two processes on the properties of shape memory alloys, it became evident that the process of angular extrusion seems to be much more effective in hardening the samples.

The results of the calorimetry curves when compared with data curves under constant load tests are complementary. The critical phase transformation temperatures of the samples were reduced by plastic deformation processes since there is a higher need for energy to perform the transformations.

The micrographs of rolled and extruded samples show that the latter process produces grains without preferential orientation. Probably, the configuration of the stress fields

generated by the increased dislocation density must be highly disoriented. These facts seem to be responsible for the greatest loss of thermoelastic properties in the extruded samples.

In the technological applications of the SMA, it is important to take into consideration the fact that the increased mechanical strength is achieved at the expense of reduced transformation temperatures and of the thermoelastic properties of the shape memory effect and may be recovered partly by heat treatment.

The extrusion-forming process proved to be more severe than the rolling process to the shape-recovery properties presented by the alloys. This analysis is of great importance in the use of LMF in applications that require high mechanical strength combined with the functional properties of shape recovery provided by martensitic phase transformations.

ACKNOWLEDGEMENT

REFERENCES

[1] Otsuka K, Wayman CM. (1998) Shape Memory Materials. Cambridge, Cambridge University Press.

[2] Qader IN, Kök M, Dagdelen F, Abdullah SS. (2020) The effect of different parameters on shape memory alloys. Sakarya University Journal of Science, 24(5): 881-902.
doi: 10.16984/saufenbilder.733645

[3] DesRoches R, McCormick J, Delemont M. (2004) Cyclic properties of superelastic shape memory alloy wires and bars. Journal of Structural Engineering, 130(1): 38-46.
doi:10.1061/(ASCE)0733-9445(2004)130:1(38)

[4] Antunes AS, Tosetti JPV, Otubo J. (2013) High shape recovery Ni–Ti SMA wire produced from electron beam melted ingot. Journal of Alloys and Compounds, 577(1): 265-267.
doi:10.1016/j.jallcom.2012.03.043

[5] De Araújo CJ, Gomes AAC, Silva JA, Cavalcanti AJT, Reis RPB, Gonzalez CH. (2009) Fabrication of shape memory alloys using the plasma skull push-pull process. Journal of Materials Processing Technology, 209(7): 3657-3664. doi:10.1016/j.jmatprotec.2008.08.025

[6] Silva NJ. (2018) Estudo de ligas com memória de forma Ni-Ti processadas por laminação e extrusão angular. PhD thesis. Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências.

[7] Silva KCA. (2014) Estudo de ligas com memória de forma Ni-Ti processadas por laminação e extrusão angular. PhD thesis. Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências.

[8] Figueiredo AM, Modenesi P, Buono V. (2009) Low-cycle fatigue life of superelastic NiTi wires. International Journal of Fatigue, 31(4): 751-758. doi:10.1016/j.ijfatigue.2008.03.014

[9] Segal VM. (1995) Materials processing by simple shear. Materials Science and Engineering: A, 197(2):157-164. doi:10.1016/0921-5093(95)09705-8

[10] Sun X, Wu DY, Kang M, Ramesh KT, Kecskes LJ. (2023) Properties and hardening behavior of equal channel angular extrusion processed Mg-Al binary alloys. Materials Characterization, 195: 112514. doi:10.1016/j.matchar.2022.112514

[11] Yang Z, Li H, Zhang Y, Liu X, Gu Q, Liu X. (2022) ECAP based regulation mechanism of shape memory properties of NiTiNb alloys. Journal of Alloys and Compounds, 897: 163184.
doi:10.1016/j.jallcom.2021.163184

[12] Leitner T, Sabirov I, Pippan R, Hohenwarter A. (2017) The effect of severe grain refinement on the damage tolerance of a superelastic NiTi shape memory alloy. Journal of the Mechanical Behavior of Biomedical Materials, 71:337-348. doi:10.1016/j.jmbbm.2017.03.020

[13] Jiang SY, Yu JB, Zhang YQ, Xing XD. (2020) Mechanically-induced martensite transformation of NiTiFe shape memory alloy subjected to plane strain compression. Trans. Nonferrous Met. Soc. China, 30: 1325-1334. doi:10.1016/S1003-6326(20)65299-2

[14] Zhang J, Chen T, Li W, Bednarcik J, Dippel AC. (2020) High temperature superelasticity realized in equiatomic Ti-Ni conventional shape memory alloy by severe cold rolling. Materials and Design, 193: 1008875. doi:10.1016/j.matdes.2020.108875

[15] Liang Q, Zhao S, Liang C, Zhao T, Wang D, Ding X, Li S, Wang Y, Zheng Y, Ren X, Mills M, Wang Y. (2022) Strain states and unique properties in cold-rolled TiNi shape memory alloys. Acta Materialia, 231: 117890, 2022. doi:10.1016/j.actamat.2022.117890

[16] Silva TTL. (2022) Comportamento termoelástico de fios de Ni-Ti com efeito memória de forma processados por laminação a frio. PhD thesis. Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências.

[17] Silva TTL, Virgolino FSS, Oliveira CAN, de Araújo CJ, Gonzalez CH. (2021) Fabricação de uma liga equiatômica de TiNi com efeito memória de forma pelo método plasma skul-push pull. Revista Engenharia e Tecnologia, 13(4): 270-280. Retrieved from https://revistas.uepg.br/index.php/ret/article/view/19659

[18] Simões JB. (2017) Fabricação de componentes miniaturizados de ligas com memória de forma NITI usando fundição de precisão. João Pessoa, Ideia.

[19] Virgolino FSS. (2017) Comportamento em fadiga termomecânica de fios de liga com memória de forma Ni-Ti-Cu. Dissertation. Universidade Federal de Pernambuco, Centro de Tecnologia e Geociência.

[20] Aghamiri SMS, Ahmadabadin MN, Shahmir H, Naghdi F, Raygan S. (2013) Study of thermomechanical treatment on mechanical-induced phase transformation of NiTi and TiNiCu wires. Journal of the mechanical behavior of biomedical materials, 21: 32-36. doi:10.1016/j.jmbbm.2013.01.014

[21] Karaca HE, Kaya I, Tobe H, Basaran B, Nagasako M, Kainuma R, Chumlyakov S. (2013) Shape memory behavior of high strength Ni54Ti46 alloys. Materials Science & Engineering, 580: 66-70. doi:10.1016/j.msea.2013.04.102

[22] Khaleghi F, Allafi JK, Chianeh VA, Noori S. (2013) Effect of short-time annealing treatment on the superelastic behavior of cold drawn Ni-rich NiTi shape memory wires. Journal of Alloys and Compounds, 554: 32-38. doi:10.1016/j.jallcom.2012.11.183

[23] Van Humbeeck J. (2003) Damping capacity of thermoelastic martensite in shape memory alloys. Journal of Alloys and Compounds, 335(1-2): 58-64. doi:10.1016/S0925-8388(03)00268-8

[24] Oliveira CAN, Gonzalez CH, De Araújo CJ, Araujo Filho OO, Urtiga Filho SUL. (2010) Thermoelastic Properties on Cu-Zn-Al Shape Memory Springs. Materials Research, 13(2): 219-223. doi:10.1590/S1516-14392010000200016

[25] Gonzalez CH, De Araújo CJ, Quadros NF, Guénin G, Morin M. (2004) Study of Martensitic Stabilization Under Stress in Cu–Al–Be Shape Memory Alloy Single Crystal. Materials Science and Engineering, 378(1-2): 253-256. doi:10.1016/j.msea.2003.11.069

[26] Yoshida I, Ono T, Asai M. (2000) Internal friction of Ti-Ni alloys. Journal of Alloys and Compounds, 310(1-2):339-343. doi:10.1016/S0925-8388(00)00957-9

[27] Yoshida I, Monma D, Iino K, Ono T, Otsuka K, Asai M. (2004) Internal friction of Ti-Ni-Cu ternary shape memory alloys, Mat Sci. Engineering Structures, 370(1-2): 444-448. doi:10.1016/j.msea.2003.05.003

[28] Genlian Fan YZ, Otsuka K, Ren X. (2006) Ultrahigh damping in R-phase state of Ti-Ni-Fe alloy. Applied Physics Letters, 89:3. doi:10.1063/1.2363173

[29] Ma Y, Du Z, Cui X, Cheng J, Liu G, Gong T. et al. (2018) Effect of cold rolling process on microstructure and mechanical properties of high strength β titanium alloy thin sheets. Progress in Natural Science: Materials International, 28(6): 711-717. doi:10.1016/j.pnsc.2018.10.004

[30] Sadeghpour S, Abbasi SM, Morakabati M, Karjalainen LP, Porter DA. (2018) Effect of cold rolling and subsequent annealing on grain refinement of a beta titanium alloy showing stress-induced martensitic transformation. Materials Science and Engineering: A, 731: 465-478. doi:10.1016/j.msea.2018.06.050

[31] Lu XL, Cai W, Zhao LC. (2003) Damping behavior of a Ti44Ni47Nb9 shape memory alloy. Journal of Materials Science Letters, 22: 1243-1245. doi:10.1023/A:1025445832316

[32] Van Humbeeck J, Liu Y (2000). The High Damping Capacity of Shape Memory Alloys. Springer-Verlag. https://doi.org/10.1007/978-3-642-59768-8_4

[33] Oliveira CAN, Gonzalez CH, Araujo Filho OO, Silva NJ, Guimaraes PB, Mendonza EN et al. (2015) Thermomechanical Analysis on Ti-Ni Shape Memory Helical Springs Under Cyclic Tensile Loads. Mat. Res., 18(2): 17-24. doi:10.1590/1516-1439.334514

[34] Wu SK, Lin HC. (2003) Damping characteristics of TiNi binary and ternary shape memory alloys. Journal of alloys and compounds, 355(1-2): 72-78. doi:10.1016/S0925-8388(03)00279-2

[35] Jiang SY, Zhang YQ. (2012) Microstructure evolution and deformation behavior of as-cast Ni-Ti shape memory alloy under compression. Transactions of Nonferrous Metals Society of China, 22(1): 90-96. doi:10.1016/S1003-6326(11)61145-X

[36] Bhagyaraj J, Ramaiah KV, Saikrishna CN, Bhaumik SK, Gouthama. (2013) Behavior and effect of Ti2Ni phase during processing of Ni-Ti shape memory alloy wire from cast ingot. Journal of Alloys and Compounds, 581: 344-351. doi:10.1016/j.jallcom.2013.07.046

[37] Duerig TW, Bhattacharya K. (2015) The Influence of the R-Phase on the Superelastic Behavior of NiTi. Shape Memory and Superelasticity, 1: 153-161. doi:10.1007/s40830-015-0013-4

[38] Chen Y, Jiang HC, Liu SW, Rong LJ, Zhao XQ. (2009) Damping capacity of TiNi-based shape memory alloys. Journal of Alloys and Compounds, 482(1-2): 151-154. doi:10.1016/j.jallcom.2009.03.148

EFFECT OF PRINTING PARAMETERS AND POST-CURING ON MECHANICAL PROPERTIES OF PHOTOPOLYMER PARTS FABRICATED VIA 3D STEREOLITHOGRAPHY PRINTING

RAFFLE RAED DIAB, ABASS ENZI AND OMAR HASHIM HASSOON

Department of Production Engineering and Metallurgy,
University of Technology, Baghdad, Iraq

*Corresponding author: Abass.M.Jabber@uotechnology.edu.iq

ABSTRACT: Three-dimensional printing has recently come into the spotlight due to its promising potential to create physically three-dimensional parts or structures through computer-aided design. While there are many options for 3D printing methods, photopolymerization 3D printing has garnered much attention because of its high resolution. However, the mechanical properties of photopolymerized 3D printed parts can vary widely depending on the manufacturing parameters and post-processing settings used. This research focuses on studying the effect of printing variables on the mechanical properties of samples printed using a Stereolithography machine (Formlabs, Form+3). Three variables are used: layer thickness (25 and 50 µm), part orientation (X and Z directions), and post-curing. Also, eight groups of 3D-printed photopolymer specimens for twenty-four specimens are used for the tensile test results. The results showed the printing variables affected the mechanical properties of samples, which were proven by Young's modulus, ultimate stress, and ultimate strain.

ABSTRAK: Pencetakan tiga dimensi baru-baru ini menjadi perhatian kerana potensinya yang menjanjikan bagi mencipta bahagian atau struktur tiga dimensi secara fizikal melalui reka bentuk bantuan komputer. Walaupun terdapat banyak pilihan bagi kaedah percetakan 3D, pencetakan 3D fotopolimerisasi telah mendapat banyak perhatian kerana resolusinya yang tinggi. Walau bagaimanapun, sifat mekanikal bahagian bercetak 3D fotopolimer adalah pelbagai bergantung pada parameter pembuatan dan tetapan pasca pemprosesan yang digunakan. Kajian ini memberi tumpuan kepada kesan pembolehubah cetakan terhadap sifat mekanikal sampel yang dicetak menggunakan mesin Stereolitografi (Formlabs, Form+3). Tiga pembolehubah digunakan: ketebalan lapisan (25 dan 50 µm), orientasi bahagian (arah X dan Z), dan pasca pengawetan. Juga, lapan kumpulan spesimen fotopolimer cetakan 3D untuk dua puluh empat spesimen digunakan bagi mendapatkan keputusan ujian tegangan. Dapatan kajian menunjukkan pembolehubah cetakan mempengaruhi sifat mekanikal sampel, dibuktikan oleh modulus Young, tegangan utama, dan tarikan utama.

1. INTRODUCTION

3D printing is one of the modern manufacturing methods that are gaining popularity because of the possibility of its use in various fields of engineering, medicine, and more. Compared to other manufacturing methods, simple and complex parts can be easily

manufactured in record time using 3D printing. Also, the manufacturing process' waste and its cost are very limited in the printing process, which reduces the manufacturing cost [1]. According to the annual growth rate, 3D printing industry sales in 2020 reached more than 8 billion dollars in sales, equivalent to 14%. In 2013, the worldwide demand for 3D printing materials reached about 2 tons, which is anticipated to grow due to the increased use of printed products [2]. Operating 3D printers and manufacturing is simple; anyone can efficiently handle the machines and manufacture parts. The manufacturing process begins with drawing the 3D part with one of the engineering drawing programs, such as SolidWorks. Then the drawing is saved in STL file format. After that, the file is sent to the 3D printer to start the manufacturing process layer-by-layer after the required manufacturing process parameters are determined [3-7]. As shown in Fig. 1, stereolithography is one of the most important methods of producing 3D parts with good quality. SLA system uses a laser to polymerize a liquid resin and transform it into a solid part by a process called photopolymerization [8,9]. Printing thinner layers results in more cohesion and higher mechanical properties, but it does so at the expense of increased construction time [10]. Layers in Stereolithography are kept in a semi-reacted "green state" with polymerizable groups between them because the polymerization reaction is incomplete, and that helps with layer-to-layer bonding by supplying layers for subsequent polymerization. At the same time, post-curing procedures are used to finish the reaction and covalently bind successive layers. After curing, UV is typically employed to complete polymerization by activating photoinitiators [11].
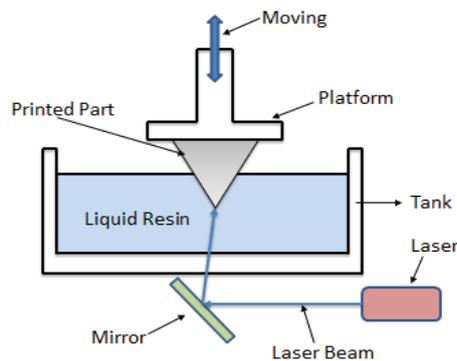
Fig. 1: Scheme of Stereolithography 3D printer machine.

Researchers used an SLA printer in previous work to produce samples with a wide range of print orientations and layer thicknesses [12]. Aznarte et al. [13] examined how 3D printing variables affect the final mechanical properties of specimens created using Digital Light Processing (DLP) 3D printing. Design, printing, and testing in this research were done on several ISO-compliant tensile test specimens. The effect of variables (layer thickness, exposure time, and part orientation) was examined for the elastic modulus, ultimate strain, ultimate tensile strength, and printing time, along with the economic impact of the researched factors regarding printing time. Results presented design guidelines for Vat Photopolymerization procedures. Khalid et al. [14] used the PR 48 photopolymer samples in this investigation using the FORM 2 SLA printer. The elastic modulus and hardness of printed specimens with 50 and 100 μm resolutions were evaluated employing a nanoindentation testing machine at 0°, 45°, and 90° directions. UV was used to cure the samples, and the impact of UV curing duration was examined. Results demonstrated that the elastic modulus and hardness for 100 μm print resolution and 0° orientation was greater if compared to 45° and 90° orientation by 35% and 390%, respectively. Elastic modulus and

hardness for 0° direction were greater than 45° direction by 106%, and greater also for 90° orientation by 92% for 50 μm print resolution. The findings demonstrated that mechanical properties strongly depended on resolution, print direction, and UV curing time. Agrawal [15] attempted to determine which orientation angle is best for which kinds of loads, specifically the fracture test and the dynamic mechanical analysis (DMA) test. Various mechanical property values were obtained during the inspection process. Results showed the orientation angle had a significant impact on the examination process. An average of three samples were taken for each test to reduce the error. After considering the stress-strain and load-extension graphs, the researcher concluded that the orientation angle should be 0°.

The parts manufactured using the SLA system change their mechanical properties according to the selected variables of the printing process. Therefore, some printed samples have poor mechanical properties due to the values of the used printing process variables, as the printing process variables have a significant impact on the mechanical properties of the produced samples. Therefore, the work aim is to determine the effect of part orientation in the X and Z axes, layer thicknesses of 25 and 50 μm, and post-curing on the mechanical properties of printed samples that are fabricated using the SLA system. The specimens' elastic modulus, ultimate stress, and ultimate strain are evaluated and analyzed depending on tensile test results to recognize the variables' values that affect printed specimens' mechanical properties.

2. MATERIAL AND EXPERIMENTAL DETAILS

A detailed description presents the manufacturing process and sample preparation, where the material used in the printing process is explained, as well as the variables of the printing process, the preparation of the number of experiments that are completed, the preparation of samples, and the examination process for samples that are produced through the printing process by a tensile test device.

2.1 Material

Clear resin is ideal for fluidics and mold making, optics, lighting, and any component needing translucency or displaying internal characteristics. It possesses several crucial characteristics, including quality. Formlabs' precisely crafted clear resin captures a model's finest features. Formlabs clear resin is excellent for quick prototyping and product development because it produces accurate, durable pieces, a glossy appearance, and the surface finish of the printed parts is smooth [15]. In the present investigation, specimens are printed with photocurable acrylic-based resin FLGPCL4 (Formlabs, MA, USA).

2.2 Process Parameters

3D printing is one of the basic methods for producing prototype parts. Still, 3D printing is not considered one of the mass production methods due to the long production time, anisotropy, etc. In addition, some critical issues face the 3D printing process, including accuracy, curvature, anisotropy, and the formation of voids inside the manufactured parts. The properties of 3D printed parts are dependent on printing parameters such as temperature, 3D printer machine resolution, layer thickness, geometries, and printing orientations. Therefore, one of the important points before the printing process is to focus on choosing the best process variables for printing to avoid defects in the manufactured parts [7,16-20]. In the current experimental study, there are three process parameters used. Two printing parameters include layer thickness and part orientation, and the third parameter is post-curing. The layer thickness is one of the most critical variables of the printing process, which affects the quality of the produced surface and the mechanical properties of the

manufactured parts. Increasing or decreasing the layer thickness affects the sample's strength. There are many directions for printing, and researchers focus on changing the direction of printing during the sample preparation process to obtain the best quality of the manufactured parts. Part orientation is one of the most studied manufacturing characteristics. Therefore, printing direction statistically impacts the mechanical properties of SLA 3D-printed parts [21-23]. Figure 2 shows the part's orientation for both the X-axis and Z-axis.
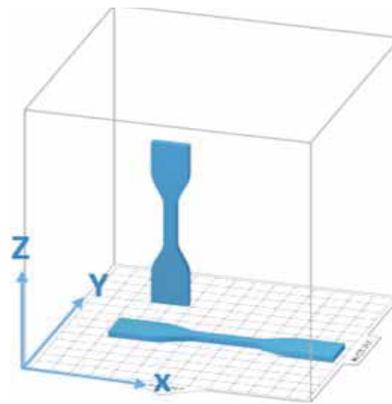
Fig. 2: Printing process directions.

Finishing, including washing and post-curing, are necessary when using SLA printing because areas of uncured resin, whether between layers or on the surface, are considered weak points and damage the material's mechanical properties. The UV post-curing of SLA printed resin can significantly improve the mechanical strength due to the complete curing of any leftover resin.

In addition, the most significant improvement in properties occurs when the UV light is at the same wavelength that the SLA printer uses to cure the resin, as each resin type has a specific wavelength for the curing process. Therefore, the appropriate wavelength must be chosen to cure the resin for the best results [24]. Immediately following the completion of the printing process, the supporting material is removed from the printed part and the part is soaked in isopropyl alcohol for 15 minutes. Alcohol liquefies any uncured resin and cleans the surface of the components. Before testing, materials were allowed to dry for 24 hours on a clean surface. Post-curing is carried out for 50 minutes in a UV chamber, previously heated to 60 °C with a light source of 405 nm and 1.25 mW/cm, see Fig. 3.

Fig. 3: Tensile test specimens during the post-curing process in Formlabs UV chamber (Formcure).

2.3 Experimental Design

SLA technology generates 3D printed parts from a liquid (photopolymer) resin by employing a UV-light source to solidify the liquid substance (resin). To construct a 3D-printed object, a build platform is submerged in a tank of photosensitive thermoset polymeric resin. Once the build platform is submerged, a UV light within the machine solidifies the material by mapping each layer of the object through the tank's bottom. After the light source has printed the layer, the platform rises to allow the swiping blade to apply a fresh coating of resin to the surface; this is continued layer-by-layer until the desired object is created [25]. Table 1 represents all experimental variables and groups that will be used to print the specimens, to ensure that the results of the studies can be reliably replicated; each sample is printed three times.

Table 1: Eight groups (A-H) of specimens with printing parameters

	Symbol	A	B	C	D	E	F	G	H
Variables	**Layer thickness (μm)**	25	25	25	25	50	50	50	50
	Part orientation	X-axis	X-axis	Z-axis	Z-axis	X-axis	X-axis	Z-axis	Z-axis
	Post-curing	Green	Post-cured	Green	Post-cured	Green	Post-cured	Green	Post-cured

2.3.1 Specimen Preparation

SolidWorks is used to make the 3D model of the specimens following ASTM D638 type IV. Fig. 4 displays the ASTM-required dimensions of the specimen. Slicing can build the model using any CAD software and export it in a 3D printable file format (STL). Each SLA printer includes software to configure printing settings and split the digital model into layers for printing. Once the part design is completed, the print preparation software transmits the instructions to the printer over a wireless or wired connection. For slicing, STL file software named Formlabs preform (Version 3.27.1) is used to slice the specimen into some layers. In addition, the factors are fed to printer software based on the process parameters used in this research.
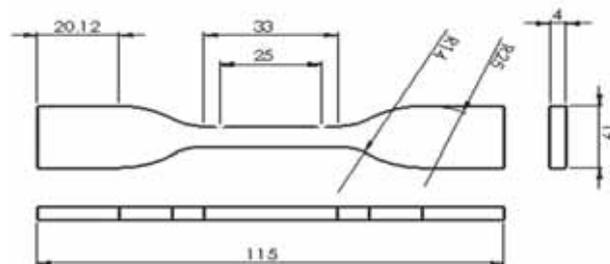
Fig. 4: Standard specimens according to ASTM D638 specimen dimensions
(All units in mm) [26].

An SLA machine (Formlabs, Form+3) is used in this research to produce the specimens. Form+3 has a 50 μm resolution in the plane parallel to the printing surface (XY resolution) and a 10 μm resolution perpendicular to the printing surface (Z resolution). Figure 5 depicts

the machine and the necessary support structure for building the specimens in all directions. It was reported that the printer's maximum build dimensions were 145 x 145 x 185 mm.
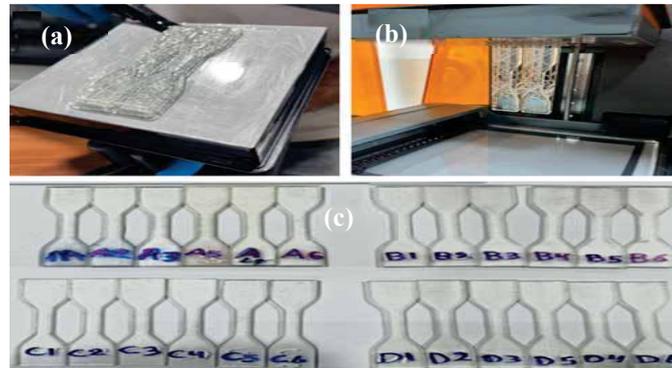
Fig. 5: Fabrication of tensile specimen (a) Longitudinal direction (b) Vertical direction (c) Twenty-four specimens.

2.3.2 Tensile Testing

A Gester universal tensile testing machine is used to test the properties of the printed specimens, with a load cell capacity of 5kN with a crosshead speed of 1mm/min; see Fig. 6.
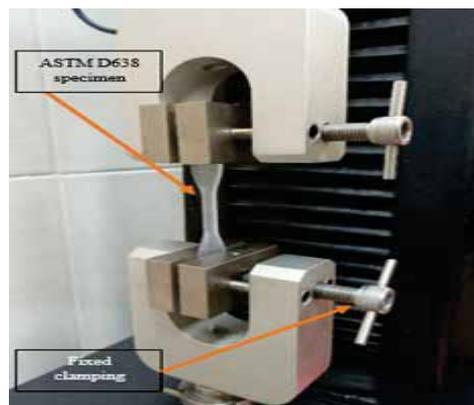
Fig. 6: Specimen placed inside the grips of the Gester universal tensile testing machine.

3. RESULTS AND DISCUSSION

In this section, the tensile test results are shown and discussed. The section is separated into three subsections that discuss the effect of several elements on Young's modulus, ultimate stress, and ultimate strain, respectively. Table 2 shows eight specimens for twenty-four groups representing a variety of process parameters.

Figure 7 shows the relationship between displacement (mm) and force (N) for green and cured samples. The curing process and layer thickness increased the tensile strength of the samples, taking into account the printing orientation. The cured samples had greater tensile strength than the green samples with decreased displacement because the material's behavior tends towards the sample's fragility. However, the green samples had more significant displacement than the cured samples. The tensile strength and displacement values change according to the printing variables, which indicate that the variables significantly impact tensile strength.

Table 2: Tensile testing results for twenty-four specimens depending on the printing parameters and post-curing

Sample	Young's modulus E (MPa)	Average of modulus E (MPa)	Ultimate stress (MPa)	Average of ultimate stress (MPa)	Strain at break (Xf)	Average of strain at break (Xf)
A1	1394.6		46.019		0.109	
A2	1135.883	1301.961	43.956	44.892	0.099	0.108
A3	1375.4		44.700		0.116	
B1	1626.75		61.8238		0.085	
B2	1659.25	1648.417	55.423	57.5566	0.0522	0.073067
B3	1659.25		55.423		0.082	
C1	1339.273		48.248		0.085	
C2	1336.182	1326.479	47.766	49.217	0.104	0.106
C3	1303.983		51.638		0.13	
D1	1818.5		62.193		0.067	
D2	1784.75	1767.274	62.112	61.707	0.066	0.067333
D3	1698.571		60.815		0.069	
E1	619.564		35.097		0.0635	
E2	916.933	796.141	35.416	33.159	0.049	0.0822
E3	851.925		28.963		0.134	
F1	1526		56.990		0.057	
F2	1347.8	1487.119	54.248	57.307	0.076	0.064
F3	1587.556		60.684		0.059	
G1	1236.167		41.537		0.076	
G2	1309.333	1228.141	44.778	42.219	0.113	0.103333
G3	1138.923		40.342		0.121	
H1	1794.5		61.731		0.0625	
H2	1449.6	1699.033	59.433	60.277	0.118	0.0815
H3	1853		59.666		0.064	
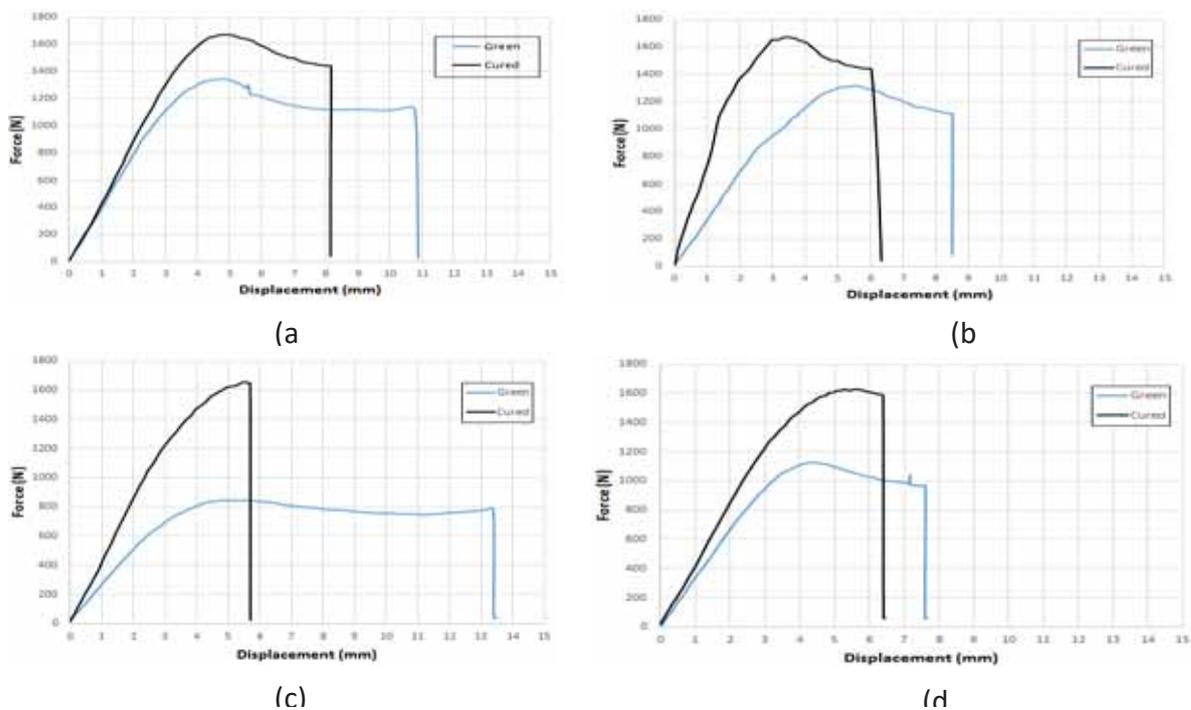
Fig. 7: Force displacement curves (a) Green and cured samples with X-axis and 25 μm (b) Green and cured samples with Z-axis and 25 μm (c) Green and cured samples with X-axis and 50 μm (d) Green and cured samples with Z-axis and 50 μm.

3.1 Effect of Layer Thickness, Part Orientation, and Post-curing on Young's Modulus

In this part, a detailed explanation, supported by values and figures, is given of the relationship between the variables of the specimens manufacturing process with Young's modulus and manufacturing time. Where it will be explained:

1. Young's modulus vs. layer thickness for samples printed with X, 0° orientation & 25 μm, 50 μm thickness.

2. Young's Modulus vs. layer thickness for samples printed with Z, 90° orientation & 25 μm, 50 μm layer thickness.

Young's modulus vs. layer thickness for samples printed with X, 0° orientation & 25 μm, 50 μm thickness: for the green specimens, the average elastic modulus produced with layer thicknesses of 25 μm and 50 μm was 1301.961 MPa, 796.141 MPa, respectively. The modulus decreased clearly by 50.582% when layer thickness was increased from 25 μm to 50 μm. For cured samples, the average elastic modulus of the printed specimens with layer thicknesses of 25 μm and 50 μm was 1648.417 MPa and 1487.119 MPa, respectively. The modulus dropped by 16.13% when layer thickness was increased from 25 μm to 50 μm). In addition, the results imply that the elastic modulus increased with a thin layer, and that happens because of the resin's exponential decay in the amount of light it transmits, increased curing speeds along the layer, and increased adhesion between layers [26]. The data also showed that curing significantly increased elastic modulus, as the elastic modulus of a green specimen that was printed with a layer thickness of 25 μm grew to 1648.417 (34.6456 %) MPa when cured. Additionally, after curing, the elastic modulus of the green specimen, which was printed with a layer thickness of 50 μm (796.141 MPa), increased to 1487.119 MPa (69.09%). The elastic modulus of 25 μm samples was higher than that of 50 μm samples because they have a lower fraction of semi-reacted resin due to superior laser beam penetration through a thinner layer, see Fig. 8 (a) & (b).

Fig. 8: Young's modulus vs. layer thickness for (a) X, 0° green samples (b) X, 0° cured samples (c) Z, 90° green samples (d) Z, 90° cured samples.

Young's Modulus vs. layer thickness for samples printed with Z, 90° orientation & 25 µm, 50 µm layer thickness: for green samples, the average elastic modulus for specimens produced with layer thicknesses of 25 µm and 50 µm was 1326.479 MPa and 1228.141 MPa, respectively. As can be seen, the modulus decreased by 9.8338 % when layer thickness was increased from 25 µm to 50 µm. For cured samples, the average elastic modulus of the specimens printed with layer thicknesses of 25 µm and 50 µm was 1767.274 MPa and 1699.033 MPa, respectively. As can be seen, the modulus dropped by 6.8241% when the layer thickness was increased from 25 µm to 50 µm. The elastic modulus of 25 µm samples was higher because they have a lower fraction of semi-reacted resin due to superior laser beam penetration through a thinner layer, see Fig. 8 (c) &(d). For green samples, the modulus of elasticity with X-direction and layer thickness varying from 50 and 25 µm increased from 796.141 to 1301.961 (increased 63.533%), and for the Z-direction increased from 1228.141 to 1326.479 MPa (increased 8%). For the cured sample, the modulus of elasticity with X-direction and layer thickness varying from 50 and 25 µm increased from 1487.119 to 1648.417 (increased 10.846%), and for Z-direction increased from 1699.033 to 1767.274 MPa (increased 4.016%). Therefore, the results showed that the modulus of elasticity with the X- direction is better than the Z-direction in both green and cured samples (see Fig. 9).

(a) (b)
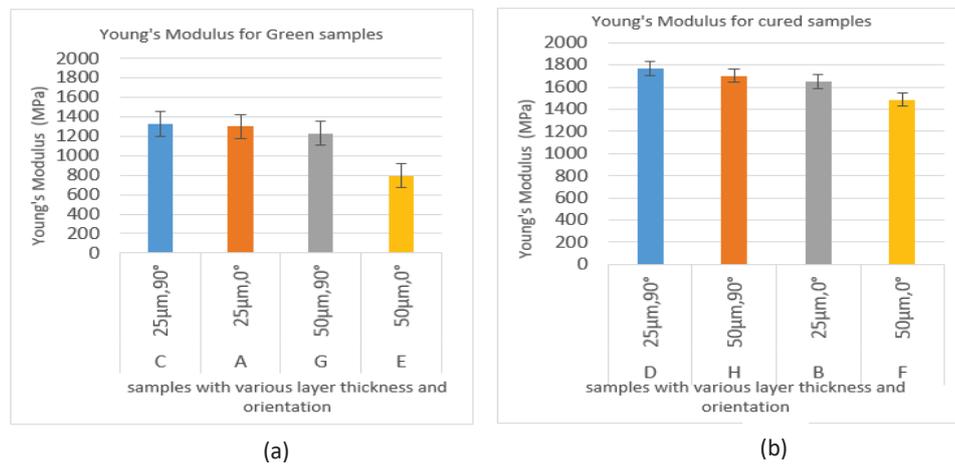
Fig. 9: Young's modulus for (a) Green samples (b) Cured samples.

Results showed that the thinner layer had higher elastic modulus, and that happened due to the exponential decay of light intensity transmission of the resin and getting higher curing rates along the layers and higher adhesion between layers.

3.2 Effect of Layer Thickness, Part Orientation, and Post-cure on Ultimate Stress

This part will explain the effect of the manufacturing process variables on the ultimate stress and printing time. Where it will be explained:

1. Ultimate stress vs. layer thickness for samples printed with X, 0° orientation & 25 µm, 50 µm thicknesses.

2. Ultimate stress vs. layer thickness for samples printed with Z, 90° orientation & 25 µm, 50 µm layer thicknesses.

Ultimate stress vs. layer thickness for samples printed with X, 0° orientation & 25 µm, 50 µm thicknesses: Figure 10 (c) & (d) demonstrates that specimens with the thinnest layer thickness can tolerate more stress than those with thicker layers. This outcome can be attributed to the resin transmittance, which allows a thinner layer to cure to a greater extent

than a thicker layer. For green samples, the average ultimate stress for specimens produced with layer thicknesses of 25 μm and 50 μm was 44.892 MPa, and 33.159 MPa, respectively. For cured samples, the average ultimate stress of the specimens printed with a layer thickness of 25 μm and 50 μm was 57.5566 MPa and 57.307 MPa, respectively.

Ultimate stress vs. layer thickness for samples printed with Z, 90° orientation, and 25 μm and 50 μm layer thickness: for green samples, the average ultimate stress for specimens produced with layer thicknesses of 25 μm and 50 μm was 49.217 MPa and 42.219 MPa, respectively. For cured samples, the average ultimate stress of the specimens printed with a layer thickness of 25 μm and 50 μm was 61.707 MPa and 60.277 MPa, respectively, see Fig. 10 (c) & (d).
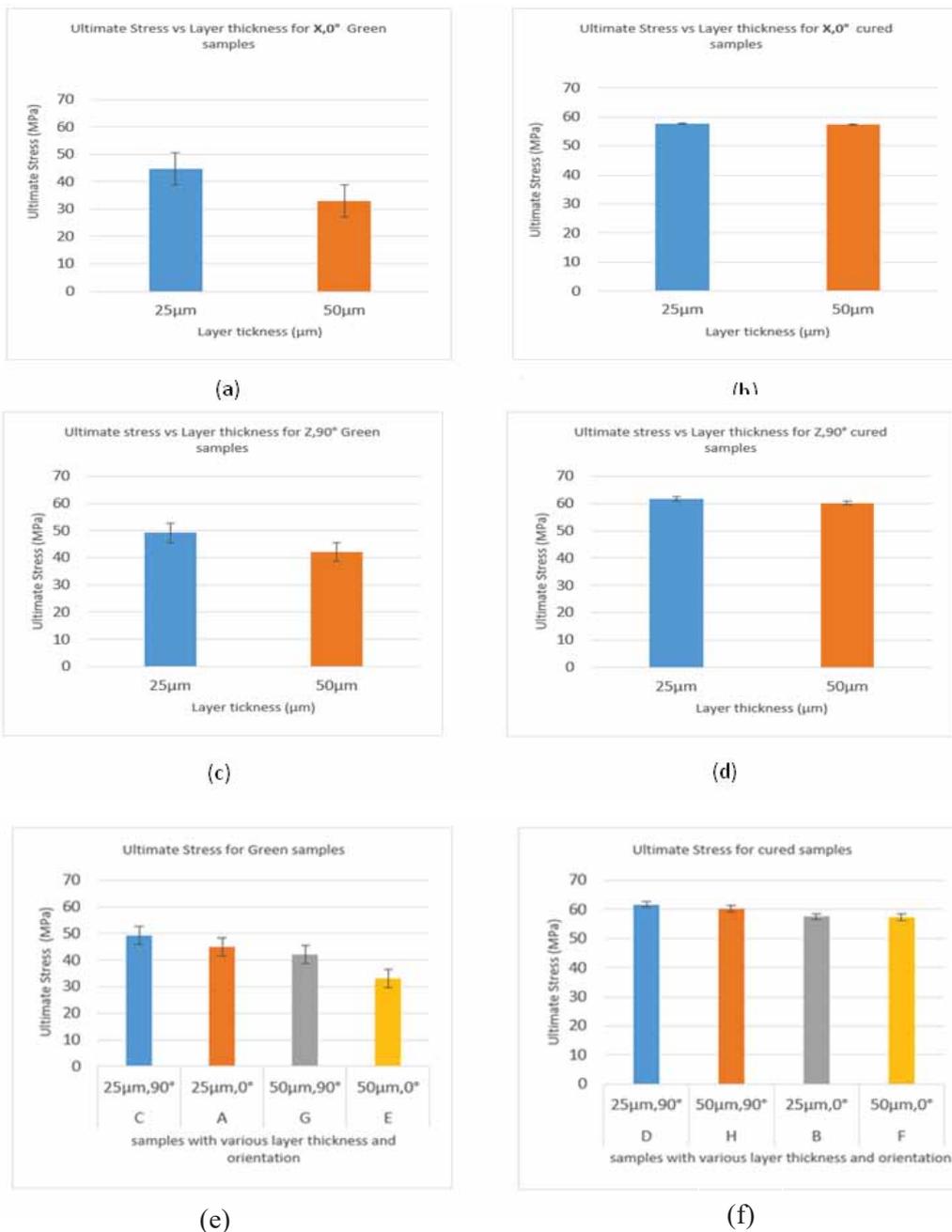
Fig. 10. Ultimate stress vs. layer thickness for (a) X, 0° green samples (b) X, 0° cured samples (c) Z, 90° green samples (d) Z, 90° cured samples (e) Green samples (f) Cured samples.

For green samples, the ultimate stress with X-direction and layer thickness varying from 50 and 25 μm increased from 33.159 to 44.892 (increased 35.384%), and for Z-direction increased from 42.219 to 49.217 MPa (increased 16.575%). For the cured sample, the ultimate stress with X-direction and layer thickness varying from 50 and 25 μm increased slightly from 57.307 to 57.556 (increased 0.004 %). For the Z-direction, it increased slightly from 60.277 to 61.707 MPa (increased by 2.372%). Therefore, the results showed that the ultimate stress with the X- direction is better than the Z-direction in both green and cured samples; see Fig. 10 (e) & (f).

The specimens with thin layers withstand greater forces than those with thicker layers, resulting from laser transmittance and providing a higher degree of curing to a thin layer than a thicker layer.

3.3 Effect of Layer Thickness, Part Orientation, and Post-cure on Ultimate Strain

The effect of manufacturing process variables on ultimate strain will be explained in this part. Where it will be explained:

1. Ultimate strain vs. layer thickness for samples printed with X, 0° orientation & 25 μm, 50 μm thickness.

2. Ultimate strain vs. layer thickness for samples printed with Z, 90° orientation & 25 μm, 50 μm layer thickness.

Ultimate strain vs. layer thickness for samples printed with X, 0° orientation & 25 μm, 50 μm thickness: for green samples, the average ultimate strain for specimens produced with layer thicknesses of 25 μm and 50 μm was 0.108 and 0.082, respectively. For cured samples, the average ultimate strain of the specimens printed with a layer thickness of 25 μm and 50 μm was 0.073 and 0.064, respectively, see Fig. 11(a) & (b).

Ultimate strain vs. layer thickness for samples printed with Z, 90° orientation & 25 μm, 50 μm layer thickness: for green samples, the average ultimate strain for specimens produced with layer thicknesses of 25 μm and 50 μm was 0.106, 0.103, respectively. For cured samples, the average ultimate strain of the specimens printed with a layer thickness of 25 μm and 50 μm was 0.0822 and 0.0815, respectively. By examining the results, the ultimate strain of the green samples was greater than that of the cured samples, see Fig. 11 (c) & (d). For green samples, the ultimate strain with X-direction and layer thickness varying from 50 and 25μm increased from 0.0822 to 0.108 (increased 31.386%), and for Z-direction, increased from 0.103 to 0.106 MPa (increased 2.912%). For the cured sample, the ultimate strain with X-direction and layer thickness varying from 50 and 25μm rose from 0.064 to 0.073 (increased 14.062%), and for Z-direction decreased from 0.081 to 0.067 MPa (fallen 20.895%). Therefore, the results showed that the ultimate strain with the X- direction is better than the Z-direction in both green and cured samples; see Fig. 11 (e) & (f).

4. CONCLUSIONS

In this study, the mechanical properties of 3D-printed photopolymers are examined and analyzed according to layer thickness, printing orientation, and post-curing. Based on the analyzed properties of elastic modulus, ultimate stress, and ultimate strain used to evaluate the printed samples, the results demonstrated that printing parameters significantly impacted mechanical properties. The results show that mechanical properties increased in X-orientation when the layer thickness varied from 50 to 25 μm in green printed samples. Therefore, the X-axis samples exhibit improvement in tensile strength and elastic modulus and have more elongation to failure when printed layers change to be thinner compared to

printed samples in the Z-axis. This could be due to the nature of the 3D printing procedure, which constructs a desired part layer-by-layer. When printing a new layer on the specimens, the additional UV-light exposure to previously printed layers will increase the polymerization of leftover unreacted monomers.
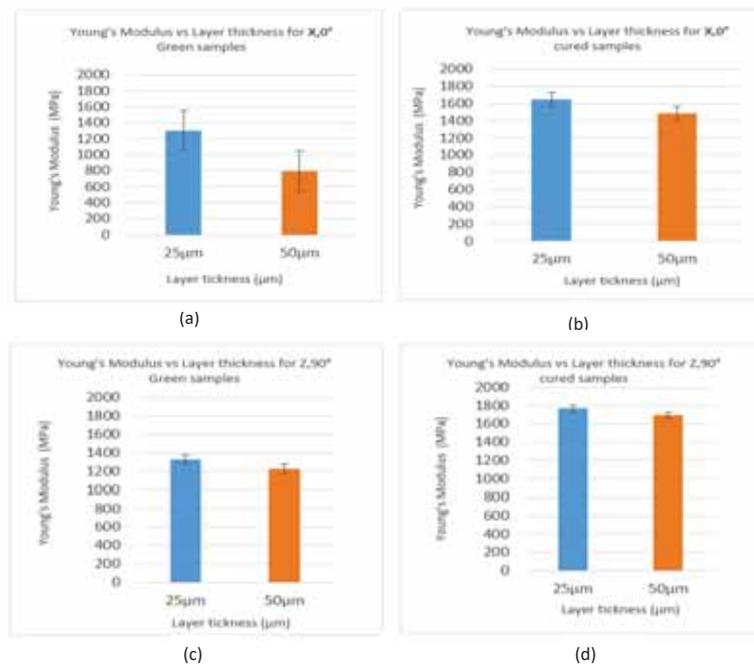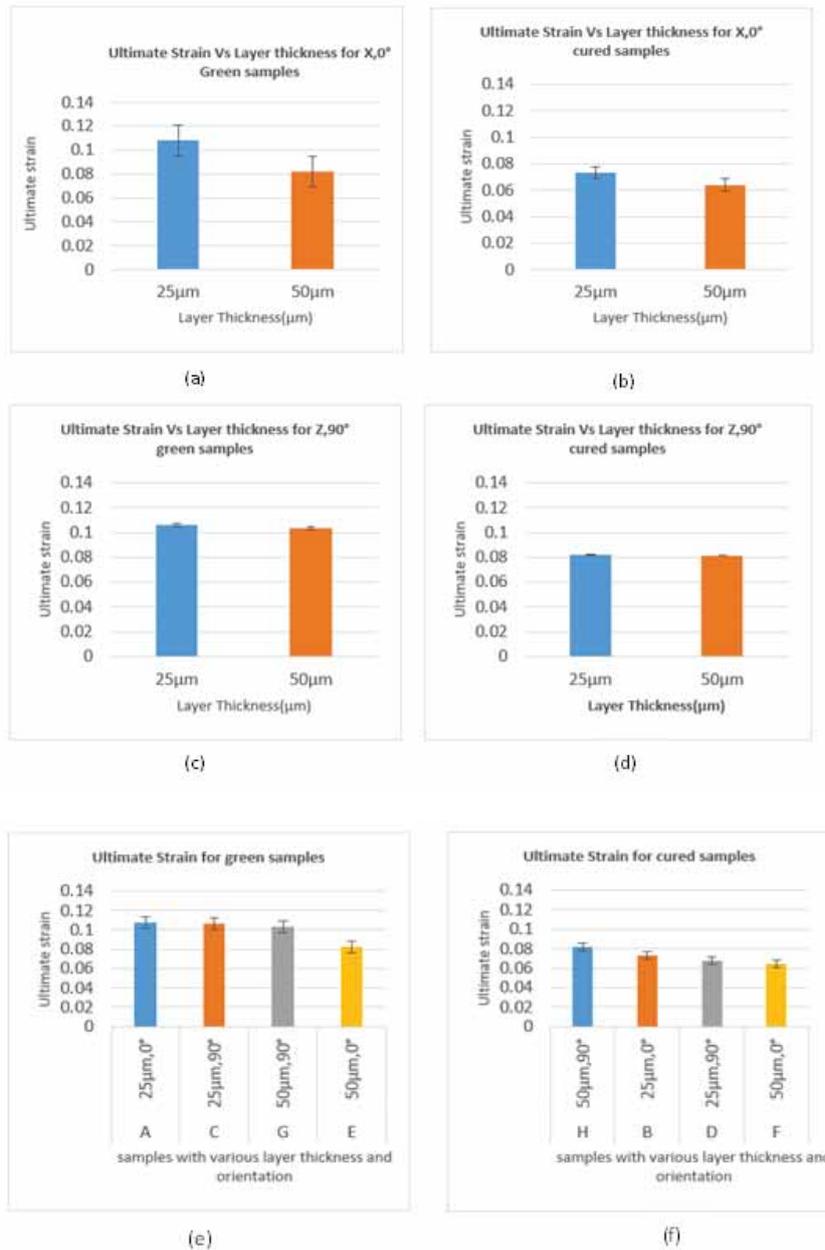

Fig. 11: Ultimate strain vs. layer thickness for (a) X, 0° green samples (b) X, 0° cured samples (c) Z, 90° green samples (d) Z, 90° cured samples (e) Green samples (f) Cured samples.

The interlayer fracture happens between the printed layers. In a thicker layer, the strength degraded faster in the specimen due to separation in the printed layers and increased interlayer stress. In contrast, the strength of specimens that are printed with a thin layer degraded slowly. Furthermore, in the case of vertical layer printing, the number of layers was large and thin. The laser-exposed surface area was large, enhancing the mechanical performance, which is distinct from horizontal printing.

The post-curing conditions had apparent effectiveness. UV curing under high temperatures and curing time improved the mechanical properties in both the X-axis and Z-axis and with various layer thicknesses. In the X- direction with 25 μm thickness, the elastic modulus increased by 26.61 % compared to the green samples. Also, the elastic modulus of the cured samples printed with 25 μm thickness and in the Z- direction increased by 33.23% compared to the green samples. The elastic modulus printed with 50μm thickness and in the X-axis increased by 86.791% compared to the green samples. Also, the elastic modulus printed with 50 μm thickness and in the vertical direction increased by 38.341% compared to the green samples. In summary, there was an increase in ultimate stress values of the samples. For the ultimate strain, the green samples were generally higher than the cured samples, as the post-curing made the material behavior more brittle. According to the results obtained, the printing orientation, layer thickness, and post-curing of the build-direction of 3D printed samples play a role in improving and controlling the anisotropy of mechanical properties of the printed samples, which is considered a challenge that is faced in the additive manufacturing process.

REFERENCES

[1] Ponce de Leon C, Hussey W, Frazao F, Jones D, Ruggeri E, Tzortzatos S, Mckerracher RD, Wills RGA, Yang S, Walsh FC. (2014) The 3D printing of a polymeric electrochemical cell body and its characterization. Chemical Engineering Transactions, 41: 1-6. https://doi.org/10.3303/CET1441001.

[2] Chong S, Chiu HL, Liao YC, Hung ST, Pan GT. (2015) Cradle to Cradle® design for 3D printing. Chemical Engineering, 45. https://doi.org/10.3303/CET1545279.

[3] Ngo TD, Kashani A, Imbalzano G, Nguyen KTQ, Hui D. (2018) Additive manufacturing (3D printing): A review of materials, methods, applications and challenges. Composites Part B: Engineering, 143: 172-196. https://doi.org/10.1016/j.compositesb.2018.02.012.

[4] Standard, A.S.T.M. (2012) Standard terminology for additive manufacturing technologies. ASTM International F2792-12a, 1-9. https://doi: 10.1520/F2792-12A.

[5] Hod L, Kurman M. (2013) Fabricated: The new world of 3D printing. John Wiley and Sons, 2013.

[6] Ivanova O, Williams C, Campbell T. (2013) Additive manufacturing (AM) and nanotechnology: promises and challenges. Rapid prototyping journal, 19 (5): 353-364. https://doi.org/10.1108/RPJ-12-2011-0127.

[7] Quan Z, Wu A, Keefe M, Qin X, Yu J, Suhr J, Byun JH, Kim BS, Chou TW. (2015) Additive manufacturing of multi-directional preforms for composites: opportunities and challenges. Materials Today, 18 (9): 503-512. https://doi.org/10.1016/j.mtadv.2019.100045.

[8] Quan Z, Larimore Z, Wu A, Yu J , Qin X, Mirotznik M, Suhr J, Byun JH, Oh Y, Chou TW. (2016) Microstructural design and additive manufacturing and characterization of 3D orthogonal short carbon fiber/acrylonitrile-butadiene-styrene preform and composite. Composites Science and Technology, 126:139-148. https://doi.org/10.1016/j.compscitech.2016.02.021.

[9] Skliutas E, Kasetaite S, Jonušauskas L, Ostrauskaite J, Malinauska M. (2018) Photosensitive naturally derived resins toward optical 3-D printing. Optical Engineering, 57(4): 041412-041412. https://doi.org/10.1117/1.OE.57.4.041412.

[10] Januszewicz R, Tumbleston JR, Quintanilla AL, Mecham SJ, DeSimone JM. (2016) Layerless fabrication with continuous liquid interface production. Proceedings of the National Academy of Sciences. 113(42): 11703-11708. https://doi.org/10.1073/pnas.1605271113.

[11] O'Neil PF. (2018) Internal void fabrication via mask projection micro-stereolithography: a rapid repeatable microfluidic prototyping technique. Ph.D Dissertation. Dublin City University, 2018.

[12]   Naik D, Kiran R. (2018). On Anisotropy, Strain Rate and Size Effects in Vat Photopolymerization Based Specimens. Additive Manufacturing, 23: 181-196. https://doi.org/10.1016/j.addma.2018.08.021.

[13] E Aznarte, Ayranci C, Qureshi AJ. (2017) Digital light processing (DLP): Anisotropic tensile considerations. 2017 International Solid Freeform Fabrication Symposium. University of Texas at Austin.

[15] Agrawal S, Ray H, Kulat A, Garhekar Y, Jibhakate R, Singh S, Bisaria H. (2023) Evaluation of Tensile Property of SLA 3D Printed NextDent Biocompatible Class I Material for Making Surgical Guides for Implant Surgery. Materials Today: Proceedings, 72: 1231-1235. https://doi.org/10.1016/j.matpr.2022.09.288.

[16] Lee CS, Kim SG, Kim HJ, Ahn SH. (2007) Measurement of Anisotropic Compressive Strength of Rapid Prototyping Parts. Journal of materials processing technology, 187: 627-630. https://doi.org/10.1016/j.jmatprotec.2006.11.095.

[17] SH. Ahn, M. Montero, D. Odell, S. Roundy, and P. Wright. (2002) Anisotropic Material Properties of Pused Deposition Modeling ABS. Rapid prototyping journal, 8: 248-257. https://doi.org/10.1108/13552540210441166.

[18] Mueller J, Shea K, and Daraio C. (2015) Mechanical Properties of Parts Fabricated with Inkjet 3D Printing Through Efficient Experimental Design. Materials and Design, 86: 902-912. https://doi.org/10.1016/j.matdes.2015.07.129.

[19]   Chockalingam K, Jawahar N, Chandrasekhar U. (2006) Influence of Layer Thickness on Mechanical Properties in Stereolithography. Rapid Prototyping Journal, 12: 106-113. https://doi.org/10.1108/13552540610652456.

[20] Rankouhi B, Javadpour S, Delfanian F, Letcher T. (2016) Failure Analysis and Mechanical Characterization of 3D Printed ABS with Respect to Layer Thickness and Orientation. Journal of Failure Analysis and Prevention, 16 :467-481.
https://doi.org/10.1007/s11668-016-0113-2.

[21] Quintana R, Choi JW, Puebla K, Wicker R. (2010) Effects of Build Orientation on Tensile Strength for Stereolithography-Manufactured ASTM D-638 Type I Specimens. The International Journal of Advanced Manufacturing Technolog, 46: 201-215. https://doi.org/10.1007/s00170-009-2066-z.

[22] Domínguez-Rodríguez G, Ku-Herrera J, Hernández-Pérez A. (2018) An Assessment of the Effect of Printing Orientation, Density, and Filler Pattern on the Compressive Performance of 3D Printed ABS Structures by Fuse Deposition. The International Journal of Advanced Manufacturing Technology, 95: 1685-1695. https://doi.org/10.1007/s00170-017-1314-x.

[23] Rohde S, Cantrell J, Jerez A, Kroese C, Damiani D, Gurnani R, DiSandro L, Anton J, Young A, Steinbach D, Ifju P. (2018) Experimental Characterization of The Shear Properties of 3D–Printed ABS and Polycarbonate Parts. Experimental Mechanics, 58: 871-884. https://doi.org/10.1007/s11340-017-0343-6.

[24] Zguris Z. (2016) How Mechanical Properties of Stereolithography 3D Prints are Affected by UV Curing. Formlabs White Paper, 1-11.

[25] Kerns J. (2015) What's the difference between stereolithography and selective laser sintering. Machine Design, 17.

[26] ASTM International (2014) Standard test method for tensile properties of plastics. ASTM international.

MULTICRITERIA DECISION MAKING ON SUPPLIER SELECTION USING SOCCER MODEL INTEGRATED WITH ANALYTICAL HIERARCHY PROCESS

MUATAZ AL HAZZA[1*], AZURA DAPIT[2], ISLAM FAISAL BOURINI[3], ZUBAIDAH MUATAZ[2] AND MOHAMMAD YEAKUB ALI[4]

[1]Mechanical and Industrial Engineering, School of Engineering,
American University of Ras Al Khaimah, Road - Ras al Khaimah, Ras Al Khaimah, UAE
[2]Department of Manufacturing and Materials Engineering,
International Islamic University Malaysia, Jalan Gombak, 53100 Selangor, Malaysia
[3]Dubai Business School, University of Dubai, Dubai, UAE
[4]Mechanical Engineering Programme Area, Faculty of Engineering,
Universiti Teknologi Brunei, Bandar Seri Begawan BE1410, Brunei Darussalam

[*]Corresponding author: muataz.alhazza@aurak.ac.ae

ABSTRACT: Supplier evaluation and selection are key components in the supply chain because supplier performance directly affects the supply chain's efficiency. Therefore, companies should think strategically when they need to select their suppliers. Thus, selecting and evaluating new suppliers is essential in shaping the supply chain's smoothness and efficiency. Selecting suppliers is a complex issue as it involves many factors and decisions to be considered and needs to be assessed with an auditing process. However, a supplier audit is an expensive way to evaluate supplier capability. This research aims to propose a supplier selection model for a corrugated carton manufacturing company. The SOCCER model, developed by Steve Rogers, was used as the supplier selection criteria. Analytical Hierarchy Process (AHP) was used in the supplier selection. A face-to-face interview method was used in collecting data. The results show that the cost structure is the ultimate concern on supplier selection which bears 44.2% of the SOCCER model, followed by operational capability (23%), customer approach (13.5%), economic performance (8.3%), strategic direction (6.9%), and lastly research and development (4.1%). The percentages inform the company how much attention they need to pay when evaluating and selecting a new supplier.

ABSTRAK: Penilaian dan pemilihan pembekal adalah kunci utama dalam rantaian bekalan kerana prestasi pembekal secara langsung melibatkan kecekapan rantaian bekalan. Oleh itu, syarikat perlu memikirkan secara strategik apabila ingin memilih pembekal. Dengan demikian, pemilihan dan penilaian pembekal baru adalah penting dalam pembentukan kelancaran rantaian bekalan dan kecekapan. Pemilihan pembekal adalah isu kompleks kerana ianya melibatkan banyak faktor dan keputusan perlu difikirkan dan perlu dinilai bersama proses audit. Namun, audit pembekal adalah mahal bagi menilai kemampuan pembekal. Kajian ini mencadangkan model pemilihan pembekal bagi syarikat pembekal kotak karton. Model SOCCER dicipta oleh Steve Rogers, telah digunakan sebagai kritia pemilihan pembekal. Proses Hirarki Analitikal (AHP) digunakan dalam pemilihan pembekal. Kaedah temuduga bersemuka digunakan dalam pengumpulan data. Dapatan kajian menunjukkan struktur harga adalah kehendak utama dalam pemilihan pembekal iaitu 44.2% daripada model SOCCER, diikuti kemampuan operasi (23%), pendekatan pelanggan (13.5%), prestasi ekonomi (8.3%), misi strategik (6.9 %), dan

akhirnya penyelidikan dan pembangunan (4.1%). Peratusan ini berguna untuk syarikat dalam memberi keutamaan dalam penilaian dan pemilihan pembekal baru.

KEYWORDS: *supplier selection; AHP; MCDM; SOCCER*

1. INTRODUCTION

Today, with a rapidly changing world and markets, companies face a vital challenge to stay competitive [1]. Markets nowadays are witnessing major changes due to the global nature of trade and rapid technological development that leads to aggressive competition between manufacturers. This rapid technological advancement changed supply chain work and made it easier to communicate and faster to deliver goods. As a result, manufacturers have realized that suppliers' performance is one of the main and important factors to survive in the market. Capable suppliers are essential to ensure the supply chain runs smoothly and efficiently. Thus, establishing a strong relationship and partnering with suppliers will result in a win-win situation where both parties would gain advantages through this relationship. However, supplier selection is complex as it should consider many factors [2]. Researchers spend most of their time finding and determining the best supplier selection criteria, resulting in many approaches and checklist assessments [3]. Each company may have different selection criteria as they have their own goals, needs, and industry types. Selecting the best supplier is a complex and challenging procedure. Therefore, choosing the best method has become one of the main success factors for manufacturers and thus, multicriteria decision-making methods (MCDM) will be useful and effective.

2. LITERATURE REVIEW

MCDM methodology is a decision-support framework that can consider multiple inconsistent criteria [4]. It is a method in which different criteria are traded off to achieve the best alternative. It includes quantitative and qualitative factors, which are considered complex decision-making tools, making it the most widely used and favorable decision methodology in many fields [5]. Different MCDM techniques employ different approaches. Throughout their analysis research, Velasquez and Hester [6] identified eleven different MCDM. However, supplier selection is a complex critical problem that must trade off various conflicting criteria such as price, quality, and delivery time. These methods have been used and applied by different researchers in supplier selection such as Multi-Attribute Utility Theory [8], Analytic Hierarchy Process (AHP) [7], Fuzzy Set Theory [9], Case-Based Reasoning [10], Data Envelopment Analysis [11], Goal Programming [12], ELECTRE method [13], Simple Additive Weighing [14], and Fuzzy TOPSIS Technique [15]. Other researchers prefer integrating two methods or techniques to have better and more effective decisions as an efficient approach [16-20].

Many researchers [21-22] claim that merging AHP with one of the intelligent methods such as Fuzzy Set Theory is favorable in decision making on the selection of suppliers due to high uncertainty in this decision-making process. However, each multicriteria technique has its advantages and disadvantages. Thus, integrating several techniques is common in multicriteria decision-making to overcome deficiencies [6].

One of the most specific multicriteria frameworks in supplier selections is the SOCCER model [3]. The acronym represents the six main criteria, S- Strategic Direction, O-Operational Capability, C- Customer Approach, C- Cost Structure, E- Economic Performance and R-Research & Development. Rogers, in his book [3], explained that the supplier assessment is needed to make sure that the supplier can handle the orders.

Therefore, the SOCCER model is simply one of the most effective checklists when auditing suppliers. Ho et al. [23] have reviewed the literature of the MCDM approaches for supplier selection in the international journals from 2000 to 2008. However, in this research, the integration between AHP as one of the effective tools in multicriteria methods with SOCCER model as one of the effective checklists in suppliers' selection auditing was implemented.

2.1 Analytical Hierarchical Process

Thomas Saaty introduced the Analytical Hierarchical Process (AHP) in 1970 [24-25]. AHP is a structured method for dealing with complex decision-making. It aids the decision-maker in setting priorities and making the best decision. It derives relative priorities on absolute scales (invariant under the identity transformation) from discrete and continuous paired comparisons in multilevel hierarchical structures [26].

AHP is a method that merges a decision's subjective and objective characteristics. The AHP considers a group of evaluation criteria and different options, among which the most effective decision will be created. First, the AHP generates a weight for every evaluation criterion according to the decision-makers pairwise comparisons of the factors. The higher Weight, the more essential is the corresponding criterion. Next, the AHP assigns a score to every possibility for a set criterion according to the decision maker's pairwise comparisons of the choices based on that criterion. The higher the score, the better the performance of choice concerning the considered criterion. Finally, the AHP combines the criteria weights and the choices' scores, determining a global score for every option and a consequent ranking. The global score for a given possibility is a weighted total of the scores it obtained concerning all the criteria.

Sipahi and Timor [27] presented a comprehensive review of applications of AHP method and ANP from 2005 to 2009. The paper additionally contains fuzzy AHP and fuzzy ANP applications. Ishizaka and Labib [28] presented a theoretically-view of the AHP articles instead of classifying them by application areas. Their paper mentioned problem modeling, pairwise comparisons, judgment scales, derivation techniques, consistency indices, incomplete matrix, synthesis of the weights, sensitivity analysis, and group decision problems.

Subramanian and Ramanathan [29] reviewed and methodologically analyzed applications of AHP in operations management from 1990 to 2009. They classified 291 application research of AHP into operations strategy, process, and product style, designing and planning resources, project management, and managing the supply chain.

AHP-primarily based techniques for supplier analysis were studied by Bruno et al. [30]. Their study underlined that the weak and strong points are rising from applying the AHP in a greater supply chain.

2.2 SOCCER Model

The SOCCER model was developed by Rogers [3] to focus on the main factors in any supplier selection and what is needed for supplier analysis. The acronym represents the six main criteria, S- Strategic Direction, O- Operational Capability, C- Customer Approach, C- Cost Structure, E- Economic Performance and R-Research & Development. Each of the criteria has five sub-criteria, as shown in Fig. 1.

Finally, after a sweeping review of the current literature in the area, it was identified that various investigators used diverse criteria and different methods. In this research, the integration between the SOCCER model developed by Steve C. Rogers and the Analytical

Hierarchy Process (AHP) will add value to the area of knowledge and will reduce the risk of selecting an inappropriate supplier. AHP aids in developing a relative weighting and prioritizing the different criteria based on the organization's objectives.

This integration considers the advantages of using a powerful qualitative method presented by the AHP method that focuses on the pairwise comparisons for every two main criteria and strengthens it by using a well-established framework that considers six main dimensions: Strategic Direction, Operational Capability, Customer Approach, Cost Structure, Economic Performance, and Research & Development. Each of those dimensions has a sub-criteria that covers all the possible factors.
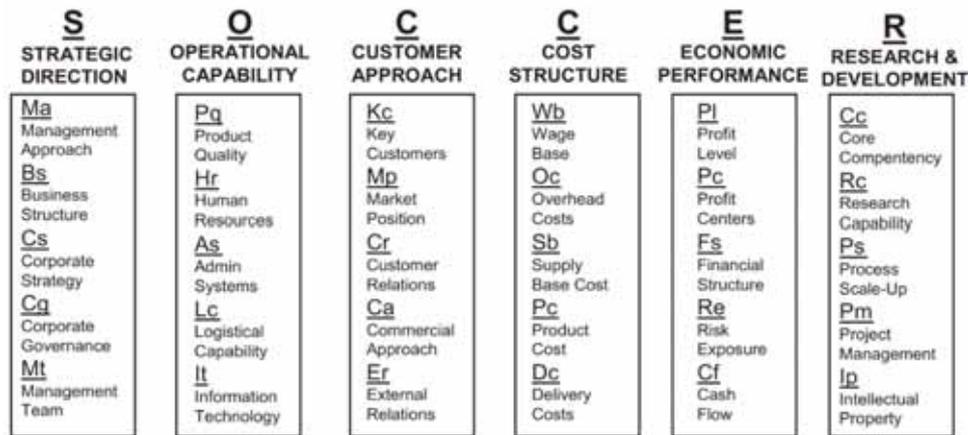
Fig. 1: Supplier analysis factors [3].

3. RESEARCH METHODOLOGY

In this research, the integration between AHP with SOCCER framework was implemented. The advantage of using the AHP method is that it focuses on pairwise comparisons for every two main criteria. To enhance the output of using the AHP method and strengthens it by using a well-established detailed framework that considers six main dimensions: Strategic Direction, Operational Capability, Customer Approach, Cost Structure, Economic Performance, and Research & Development. Each of those dimensions has a sub-criteria that covers all the possible factors. A sequential feedback loop is summarized in Fig. 2. The main steps of the research can be concluded in the following steps:

1. Analyze the suitability of using the SOCCER framework
2. Integrate the framework with the AHP method
3. Select the internal expertise in supplier selection for the company: quality assurance, customer service, and purchasing
4. Design the interview framework: Interview questions were designed to utilize the pairwise matrix.
5. Creating the pairwise matrix based on the interview results.
6. Test the consistency by avoiding bias due to the area of the expertise
7. Develop a conceptual framework

The consistency was determined using the followings steps [9,30]:

1. Calculating the consistency index CI, using Eq, (1), where n is the number of criteria in the comparison.

$$CI = (Max\ Eigenvalue - n)/(n-1) \qquad (1)$$

2. Dividing its value by the random consistency index, which is stated by Saaty depending on the value of n. The results are shown in Table 14.
3. Calculating the Consistency Ratio (CR) value using equation (2) where the value below 10% is considered consistent.

$$CR=CI/RI<0.1\sim10\% \qquad (2)$$

where CI is the consistency index, CR is the Consistency Ratio, and RI is the Random Consistency Index
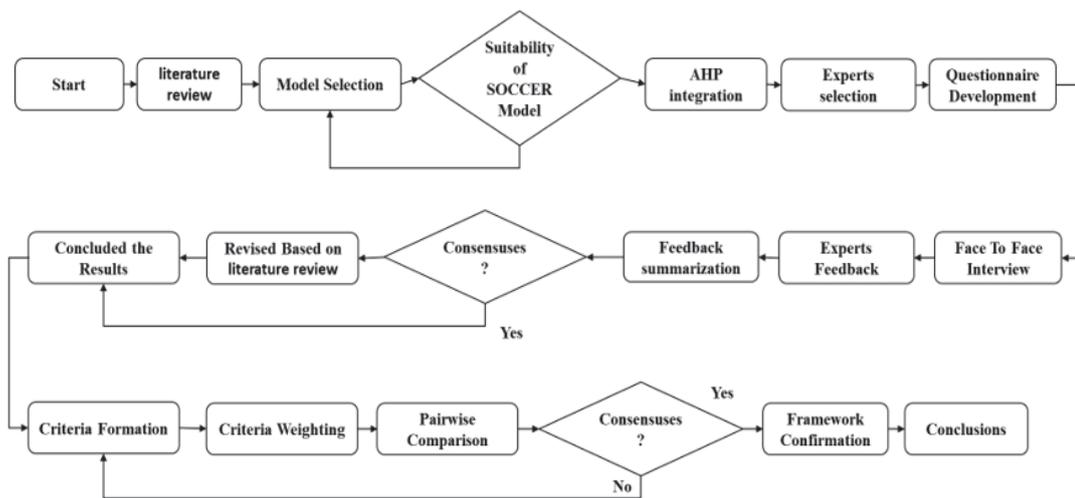
Fig. 2: Research methodology flow chart.

4. RESULTS

Supplier selection is a complex issue involving many factors and decisions due to the difficulties of trading off financial and performance evaluation. The data were gathered through interview sessions using the criteria based on the SOCCER model. The relative "priority" given to each element in the hierarchy is determined by comparing pairwise using the AHP method. The criteria ranking is decided through pairwise comparisons, and the preference scale ranking the hierarchy.

An industrial company was selected to implement SOCCER and AHP. The representatives from three different departments: quality assurance, customer service, and purchasing, were interviewed through a comprehensive individual interview session designed in order to fit for pairwise comparison. The results were concluded based on the SOCCER model. The results were analyzed based on two steps:

STEP 1: Developing a decision matrix of the SOCCER model for the main factors for each interviewee. The results are shown in Tables 1, 2, and 3.

The results show an agreement on the cost structure with the higher priority of 0.45, 0.42, and 0.44 for the quality assurance, customer service, and purchasing departments. Moreover, all the experts show agreement that the second essential criterion is the operational capability with the values of 0.27, 0.27, and 0,23.

Table 1: Decision matrix of SOCCER model by Quality Assurance

	Strategic Direction	Operational Capability	Customer Approach	Cost Structure	Economic Performance	R&D	Priority Vector
Strategic Direction	1	1/7	1/4	1/8	3	3	0.060
Operational Capability	7	1	3	1/3	6	7	0.270
Customer Approach	4	1/3	1	1/5	5	6	0.150
Cost Structure	**8**	**3**	**5**	**1**	**9**	**6**	**0.450**
Economic Performance	1/3	1/6	1/5	1/9	1	1/3	0.030
R & D	1/3	1/7	1/6	1/6	3	1	0.040

CI = 0.1116, RI = 1.24, CR = 0.09 < 0.1 OK

Table 2: Decision matrix of SOCCER model by Customer Services

	Strategic Direction	Operational Capability	Customer Approach	Cost Structure	Economic Performance	R&D	Priority Vector
Strategic Direction	1	1/6	1/5	1/8	½	2	0.050
Operational Capability	6	1	3	1/2	4	5	0.270
Customer Approach	5	1/3	1	1/4	2	4	0.140
Cost Structure	**8**	**2**	**4**	**1**	**5**	**8**	**0.420**
Economic Performance	2	1/4	1/2	1/5	1	2	0.080
R & D	½	1/5	1/4	1/8	1/2	1	0.040

CI = 0.0372, RI = 1.24, CR = 0.03 < 0.1 OK

Table 3: Decision matrix of SOCCER model by Purchasing

	Strategic Direction	Operational Capability	Customer Approach	Cost Structure	Economic Performance	R&D	Priority Vector
Strategic Direction	1	1/5	1/3	1/6	1	3	0.070
Operational Capability	5	1	3	1/4	2	5	0.230
Customer Approach	3	1/3	1	1/3	2	3	0.140
Cost Structure	**6**	**4**	**3**	**1**	**4**	**7**	**0.440**
Economic Performance	1	1/2	1/2	1/4	1	2	0.090
R & D	1/3	1/5	1/3	1/7	1/2	1	0.040

CI = 0.0868, RI = 1.24, CR = 0.07 < 0.1 OK

STEP 2: Developing a decision matrix of the SOCCER model for each interviewee for each sub criterion. Tables 4, 5, 6, 7, 8, and 9 show the decision matrix for quality assurance.

Table 4: Decision matrix of Strategic Direction by Quality Assurance

Strategic Direction	Management Approach	Business Structure	Corporate strategy	Corporate Governance	Management team	Priority Vector
Management Approach	1	1/5	1/3	5	1/6	0.070
Business Structure	5	1	4	7	1	0.360
Corporate strategy	3	1/4	1	6	1/6	0.130
Corporate Governance	1/5	1/7	1/6	1	1/8	0.030
Management team	**6**	**1**	**6**	**8**	**1**	**0.410**

CI = 0.0784, RI = 1.12, CR = 0.07 < 0.1 OK

Table 5: Decision matrix of Operational Capability by Quality Assurance

Operational Capability	Product Quality	Human Resource	Admin Systems	Logistical Capability	Information Technology	Priority Vector
Product Quality	1	1/5	1/3	5	1/6	0.070
Human Resources	5	1	4	7	1	0.360
Admin Systems	3	1/4	1	6	1/6	0.130
Logistical Capability	1/5	1/7	1/6	1	1/8	0.030
Information Technology	**6**	**1**	**6**	**8**	**1**	**0.410**

CI = 0.1008, RI = 1.12, CR = 0.09 < 0.1 OK

Table 6: Decision matrix of Customer Approach by Quality Assurance.

Customer Approach	Key Customers	Market Position	Customer Relations	Customer Approach	External Relations	Priority Vector
Key customers	1	9	1/2	2	6	0.330
Market Position	1/9	1	1/6	1/4	1/3	0.040
Customer Relation	**2**	**6**	**1**	**1**	**5**	**0.340**
Customer Approach	1/2	4	1	1	4	0.220
External Relation	1/6	3	1/5	1/4	1	0.070

CI = 0.0784, RI = 1.12, CR = 0.07 < 0.1 OK.

Table 7: Decision matrix of Cost Structure by Quality Assurance

Cost Structure	Wage Base	Overhead Costs	Supply Base Cost	Product Cost	Delivery Cost	Priority Vector
Wage Base	1	1/3	1/2	1/8	1/7	0.040
Overhead Costs	3	1	3	1/5	1/4	0.110
Supply Base Cost	2	1/3	1	1/7	1/6	0.060
Product Cost	**8**	**5**	**7**	**1**	**4**	**0.520**
Delivery Cost	7	4	6	1/4	1	0.270

CI = 0.1008, RI = 1.12, CR = 0.09 < 0.1 OK

Table 8: Decision matrix of Economic Performance by Quality Assurance

Economic Performance	Profit Level	Profit Centers	Financial Structure	Risk Exposure	Cash Flow	Priority Vector
Profit Level	1	1/2	1/6	1/3	1/4	0.050
Profit Centers	2	1	1/7	1/2	1/5	0.070
Financial Structure	**6**	**7**	**1**	**4**	**3**	**0.500**
Risk Exposure	3	2	1/4	1	1/2	0.140
Cash Flow	4	5	1/3	2	1	0.240

CI = 0.0336 RI = 1.12, CR = 0.03 < 0.1 OK.

Table 9: Decision matrix of Research & Development by Quality Assurance

Research & Development	Core Competency	Research Competency	Process Scale-Up	Project Management	Intellectual Property	Priority Vector
Core Competency	1	4	7	1/4	1/2	0.180
Research Competency	1/4	1	3	1/9	1/5	0.060
Process Scale-Up	1/7	1/3	1	1/6	1/3	0.040
Project Management	**4**	**9**	**6**	**1**	**2**	**0.470**
Intellectual Property	2	5	3	1/2	1	0.240

CI = 0.0784, RI = 1.12, CR = 0.07 < 0.1 OK

The results are concluded in Table 10. The priority vector of Strategic Direction, Operational Capability, Customer Approach, Cost Structure, Economic Performance, and R & D were management team, product quality, customer relation, Product Cost, financial structure, and project management were 0.41,0.41, 0.43, 0.52, 0.50, and 0.74, respectively. The SOCCER model overall Weight by the Quality Assurance is concluded in Table 10. The Quality Assurance preferred Cost structure (45%), followed by Operational Capability (27%), Customer Approach (15%), Strategic direction (6%), Research & Development (4%), and Economic performance (3%).

Table 10: Results of SOCCER model (Quality Assurance)

Factor	Item	Priority Vector (%)	Sub-Criteria Weight According to Main-Criteria Weight (%)
Strategic Direction (6%)	Management Approach	7.00	0.42
	Business Structure	36.00	2.16
	Corporate strategy	13.00	0.78
	Corporate Governance	3.00	0.18
	Management team	**41.00**	**2.46**
Operational Capability (27%)	**Product Quality**	**38.00**	**10.26**
	Human Resources	5.00	1.35
	Admin Systems	25.00	6.75
	Logistical Capability	26.00	7.02
	Information Technology	6.00	1.62
Customer Approach (15%)	Key customers	33.00	4.95
	Market Position	4.00	0.60
	Customer Relations	**34.00**	**5.10**
	Customer Approach	22.00	3.30
	External Relations	7.00	1.05
Cost Structure (45%)	Wage Base	4.00	1.80
	Overhead Costs	11.00	4.95
	Supply Base Cost	6.00	2.70
	Product Cost	**52.00**	**23.40**
	Delivery Cost	27.00	12.15
Economic Performance (3%)	Profit Level	5.00	0.15
	Profit Centres	7.00	0.21
	Financial Structure	**50.00**	**1.50**
	Risk Exposure	14.00	0.42
	Cash Flow	24.00	0.72
Research & Development (4%)	Core Competency	18.00	0.72
	Research Competency	6.00	0.24
	Process Scale-Up	4.00	0.16
	Project Management	**47.00**	**1.88**
	Intellectual Property	24.00	0.96

The decision matrices for each main criteria for the customer service are shown in Tables 11,12, 13, 14, 15, and 16.

Al Hazza et al.

Table 11: Decision matrix of Strategic Direction by Customer Service

Strategic Direction	Management Approach	Business Structure	Corporate Strategy	Corporate Governance	Management Team	Priority Vector
Management Approach	1	2	3	4	1/2	0.250
Business Structure	1/2	1	2	3	1/4	0.140
Corporate strategy	1/3	1/2	1	2	1/5	0.090
Corporate Governance	1/4	1/3	1/2	1	1/6	0.060
Management team	**2**	**4**	**5**	**6**	**1**	**0.460**

CI = 0.0112, RI = 1.12, CR = 0.01 < 0.1 OK

Table 12: Decision matrix of Operational Capability by Customer Service

Operational Capability	Product Quality	Human Resource	Admin Systems	Logistical Capability	Information Technology	Priority Vector
Product Quality	**1**	**8**	**6**	**4**	**3**	**0.500**
Human Resources	1/8	1	1	1/5	1/4	0.050
Admin Systems	1/6	1	1	1/5	1/3	0.060
Logistical Capability	1/4	5	5	1	3	0.250
Information Technology	1/3	4	3	1/3	1	0.150

CI = 0.0896, RI = 1.12, CR = 0.08 < 0.1 OK

Table 13: Decision matrix of Customer Approach by Customer Service

Customer Approach	Key customers	Market Position	Customer Relations	Customer Approach	External Relations	Priority Vector
Key customers	1	4	1/3	1/5	4	0.140
Market Position	1/4	1	1/6	1/8	1/2	0.040
Customer Relation	3	6	1	1/3	4	0.250
Customer Approach	**5**	**8**	**3**	**1**	**7**	**0.510**
External Relation	1/4	2	1/4	1/7	1	0.060

CI = 0.056, RI = 1.12, CR = 0.05 < 0.1 OK

Table 14: Decision matrix of Cost Structure by Customer Service

Cost Structure	Wage Base	Overhead Costs	Supply Base Cost	Product Cost	Delivery Cost	Priority Vector
Wage Base	1	1/5	1/2	1/9	1/8	0.030
Overhead Costs	5	1	3	1/6	1/5	0.110
Supply Base Cost	2	1/3	1	1/8	1/5	0.050
Product Cost	**9**	**6**	**8**	**1**	**3**	**0.520**
Delivery Cost	8	5	5	1/3	1	0.290

CI = 0.0784, RI = 1.12, CR = 0.07 < 0.1 OK

Table 15: Decision matrix of Economic Performance by Customer Service

Economic Performance	Profit Level	Profit Centers	Financial Structure	Risk Exposure	Cash Flow	Priority Vector
Profit Level	1	1/2	1/7	1/4	1/5	0.040
Profit Centers	2	1	1/6	1/3	1/4	0.070
Financial Structure	**7**	**6**	**1**	**5**	**4**	**0.520**
Risk Exposure	4	3	1/5	1	1/3	0.130
Cash Flow	5	4	1/4	3	1	0.230

CI = 0.0784 RI = 1.12, CR = 0.07 < 0.1 OK

Table 16: Decision matrix of Research & Development by Customer Service

Research & Development	Core Competency	Research Competency	Process Scale-Up	Project Management	Intellectual Property	Priority Vector
Core Competency	1	2	4	1/6	1/5	0.100
Research Competency	1/2	1	2	1/7	1/4	0.070
Process Scale-Up	1/4	1/2	1	1/8	1/5	0.040
Project Management	**6**	**7**	**8**	**1**	**3**	**0.520**
Intellectual Property	5	4	5	1/3	1	0.270

CI = 0.0672, RI = 1.12, CR = 0.06 < 0.1 OK.

The results are concluded in Table 17. The priority vector of Strategic Direction, Operational Capability, Customer Approach, Cost Structure, Economic Performance, and R & D were management team, product quality, customer relation, product cost, financial structure, and project management were 0.46, 0.50, 0.51, 0.52, 0.52, and 0.52, respectively. The Quality Assurance Preferred Cost structure (42%), followed by Operational Capability (27%), Customer Approach (14%), Economic performance (8%), Strategic direction (6%), and Research & Development (4%).

Table 17: Results of R1 based on SOCCER model (Customer Service)

Factor	Item	Priority Vector (%)	Sub-Criteria Weight According to Main-Criteria Weight (%)
Strategic Direction (5%)	Management Approach	25.00	1.25
	Business Structure	14.00	0.70
	Corporate strategy	9.00	0.45
	Corporate Governance	6.00	0.30
	Management team	**46.00**	**2.30**
Operational Capability (27%)	**Product Quality**	**50.00**	**13.50**
	Human Resources	5.00	1.35
	Admin Systems	6.00	1.62
	Logistical Capability	25.00	6.75
	Information Technology	15.00	4.05
Customer Approach (14%)	Key customers	14.00	1.96
	Market Position	4.00	0.56
	Customer Relations	25.00	3.50
	Customer Approach	**51.00**	**7.14**
	External Relations	6.00	0.84
Cost Structure (42%)	Wage Base	3.00	1.26
	Overhead Costs	11.00	4.62
	Supply Base Cost	5.00	2.10
	Product Cost	**52.00**	**24.84**
	Delivery Cost	29.00	12.18
Economic Performance (8%)	Profit Level	4.00	0.32
	Profit Centers	7.00	0.56
	Financial Structure	**52.00**	**4.16**
	Risk Exposure	13.00	1.04
	Cash Flow	23.00	1.84
Research & Development (4%)	Core Competency	10.00	0.4
	Research Competency	7.00	0.28
	Process Scale-Up	4.00	0.16
	Project Management	**52.00**	**2.08**
	Intellectual Property	27.00	1.08

The results are concluded in Table 24. The priority vector of Strategic Direction, Operational Capability, Customer Approach, Cost Structure, Economic Performance, and R&D were management team, product quality, customer relation, product cost, financial structure, and project management with 0.46, 0.51, 0.50, 0.52, 0.51, and 0.51, respectively. However, the unexpected outcome was that Logistical Capability ranked first with the priority vector of 0.517. The quality assurance preferred cost structure (44%), followed by operational Capability (23%), customer approach (14%), economic performance (9%), strategic direction (7%), and research & development (4%).

5. SUPPLIER SELECTION CRITERIA FRAMEWORK

The overall priority vector of the interviewees is concluded in Table 25, which shows that the cost structure has the highest number by 0.442, which makes it 44.2% from the SOCCER model. The second highest was an operational capability with 23%, however, other factors of customer approach, economic performance, strategic direction, and R&D, had the remaining weightage of 13.5%, 8.3%, 6.9%, and 4.1, respectively.

The decision matrix for each main criteria for the purchasing department represener are shown in tables 18, 19, 20, 21, 22, and 23.

Table 18: Decision matrix of Strategic Direction by Purchasing

Strategic Direction	Management Approach	Business Structure	Corporate strategy	Corporate Governance	Management team	Priority Vector
Management Approach	1	3	4	5	1/2	0.290
Business Structure	1/3	1	2	3	1/4	0.130
Corporate strategy	1/4	1/2	1	1	1/6	0.070
Corporate Governance	1/5	1/3	1	1	1/6	0.060
Management team	**2**	**4**	**6**	**6**	**1**	**0.460**

CI = 0.0224, RI = 1.12, CR = 0.02 < 0.1 OK

Table 19: Decision matrix of Operational Capability by Purchasing

Operational Capability	Product Quality	Human Resource	Admin Systems	Logistical Capability	Information Technology	Priority Vector
Product Quality	1	7	5	1/4	3	0.260
Human Resources	1/7	1	1/2	1/8	1/4	0.040
Admin Systems	1/5	2	1	1/6	1/3	0.060
Logistical Capability	**4**	**8**	**6**	**1**	**4**	**0.510**
Information Technology	1/3	4	3	1/4	1	0.130

CI = 0.0896, RI = 1.12, CR = 0.08 < 0.1 OK

Table 20: Decision matrix of Customer Approach by Purchasing

Customer Approach	Key customers	Market Position	Customer Relations	Customer Approach	External Relations	Priority Vector
Key customers	1	5	1/3	2	4	0.240
Market Position	1/5	1	1/8	1/5	1/2	0.040
Customer Relation	**3**	**8**	**1**	**4**	**6**	**0.500**
Customer Approach	1/2	5	1/4	1	2	0.150
External Relation	1/4	2	1/6	1/2	1	0.070

CI = 0.0336, RI = 1.12, CR = 0.03 < 0.1 OK

Table 21: Decision matrix of Cost Structure by Purchasing

Cost Structure	Wage Base	Overhead Costs	Supply Base Cost	Product Cost	Delivery Cost	Priority Vector
Wage Base	1	1/5	1/3	1/9	1/8	0.030
Overhead Costs	5	1	3	1/7	1/5	0.100
Supply Base Cost	3	1/3	1	1/7	1/5	0.060
Product Cost	**9**	**7**	**7**	**1**	**3**	**0.520**
Delivery Cost	8	5	5	1/3	1	0.290

CI = 0.0896, RI = 1.12, CR = 0.08 < 0.1 OK

Table 22: Decision matrix of Economic Performance by Purchasing

Economic Performance	Profit Level	Profit Centers	Financial Structure	Risk Exposure	Cash Flow	Priority Vector
Profit Level	1	1/2	1/9	1/6	1/7	0.030
Profit Centers	2	1	1/7	1/5	1/6	0.050
Financial Structure	**9**	**7**	**1**	**6**	**3**	**0.510**
Risk Exposure	6	5	1/6	1	1/3	0.140
Cash Flow	7	6	1/3	3	1	0.270

CI = 0.0896RI = 1.12, CR = 0.08 < 0.1 OK.

Table 23: Decision matrix of Research & Development by Purchasing

Research & Development	Core Competency	Research Competency	Process Scale-Up	Project Management	Intellectual Property	Priority Vector
Core Competency	1	3	5	1/5	1/3	0.130
Research Competency	1/3	1	2	1/7	1/4	0.060
Process Scale-Up	1/5	1/2	1	1/8	1/7	0.040
Project Management	**5**	**7**	**8**	**1**	**3**	**0.510**
Intellectual Property	3	4	7	1/3	1	0.260

CI = 0.056, RI = 1.12, CR = 0.05 < 0.1 OK.

Table 24: Results of R1 based on SOCCER model

Factor	Item	Priority Vector (%)	Sub-Criteria Weight According to Main-Criteria Weight (%)
Strategic Direction (7%)	Management Approach	29.00	2.03
	Business Structure	13.00	0.91
	Corporate strategy	7.00	0.49
	Corporate Governance	6.00	0.42
	Management team	**46.00**	**3.22**
Operational Capability (23%)	Product Quality	26.00	5.98
	Human Resources	4.00	0.92
	Admin Systems	6.00	1.38

Factor	Item	Priority Vector (%)	Sub-Criteria Weight
	Logistical Capability	**51.00**	**11.73**
	Information Technology	13.00	2.99
Customer Approach (14%)	Key customers	24.00	3.36
	Market Position	4.00	0.56
	Customer Relations	**50.00**	**7.00**
	Customer Approach	15.00	2.10
	External Relations	7.00	0.98
Cost Structure (44%)	Wage Base	3.00	1.32
	Overhead Costs	10.00	4.40
	Supply Base Cost	6.00	2.64
	Product Cost	**52.00**	**22.88**
	Delivery Cost	29.00	12.76
Economic Performance (9%)	Profit Level	3.00	0.27
	Profit Centers	5.00	0.45
	Financial Structure	**51.00**	**4.59**
	Risk Exposure	14.00	1.26
	Cash Flow	27.00	2.43
Research & Development (4%)	Core Competency	13.00	0.52
	Research Competency	6.00	0.24
	Process Scale-Up	4.00	0.16
	Project Management	**51.00**	**2.04**
	Intellectual Property	26.00	1.04

Table 25: Results of R1 based on SOCCER model

Factor	Item	Priority Vector (%)	Sub-Criteria Weight According to Main-Criteria Weight (%)
Strategic Direction (6.9%)	Management Approach	29.00	2.00
	Business Structure	12.90	0.89
	Corporate strategy	6.70	0.46
	Corporate Governance	6.00	0.41
	Management team	45.50	3.14
Operational Capability (23%)	Product Quality	25.50	5.87
	Human Resources	3.80	0.87
	Admin Systems	6.00	1.38
	Logistical Capability	**51.70**	**11.89**
	Information Technology	12.90	2.97
Customer Approach (13.5%)	Key customers	23.50	3.17
	Market Position	4.30	0.58
	Customer Relations	**50.30**	**6.79**

Cost Structure (44.2%)	Customer Approach	14.60	1.97
	External Relations	7.30	0.99
	Wage Base	3.10	1.37
	Overhead Costs	10.60	4.69
	Supply Base Cost	5.90	2.61
	Product Cost	52.10	23.03
	Delivery Cost	28.40	12.55
Economic Performance (8.3%)	Profit Level	3.30	0.27
	Profit Centres	4.90	0.41
	Financial Structure	**52.00**	**4.32**
	Risk Exposure	14.10	1.17
	Cash Flow	25.80	2.14
Research & Development (4.1%)	Core Competency	13.30	0.55
	Research Competency	6.10	0.25
	Process Scale-Up	3.70	0.15
	Project Management	51.30	2.10
	Intellectual Property	25.60	1.05

Finally, the overall weight supplier selection criteria were calculated. Figure 3 shows the percentages of the SOCCER model supplier selection criteria.

The results show that cost structures are the first criteria in supplier selection, with 44.2% of the total weight. However, out of this weight, 23% focused on product cost. These results are understandable, especially when analyzing the importance of criteria in small and medium companies. Operational Capability came next, with about 23% of the total weightage. However, it is unexpected that Logistical Capability ranked first with a priority vector of 0.517 with 11.89%. Next in the Ranks is the Customer Approach, with 13.5% of the overall weight. Customer Relations was the most focused on Customer Approach in supplier selection criteria, priority vector of 0.503. Following this were the Key Customers (0.235), Customer Approach (0.146), External Relation (0.073), and 55, and lastly Market Position (0.043). The Market Position showed a low percentage of 0.58% of overall weights due to filtering. Thus, considering market position was the last thing they would consider. The Economic Performance stood in the second last place with 8.3% of overall weight. Altogether, the interviewees agreed that Financial Structures (0.520) was the foremost factor to be considered. Financial structure is the long-term and short-term company sources of capital composition. The suppliers manage their liabilities and equity to finance their operations. Financial Structures are significant if the company wants to have a long-term relationship with the suppliers to predict the supplier's performance in the future. A well-planned financial structure means stability and will build trust in the company in the long run. Finally, the overall weight showed that cost structure, operational capability, and customer approach with a total of 80.7% were the factors that contribute the most to supplier selection.

Fig. 3: Overall weight supplier selection criteria.

6. CONCLUSIONS

This paper has illustrated a case study for the supplier selection in manufacturing corrugated cartons company in Malaysia. The Research outputs can be concluded as the followings:

1. The SOCCER model is one of the most comprehensive models covering most essential supplier selection criteria.
2. Integrating the AHP method and the SOCCER model gave a practical and valid framework.
3. Consulting the company's internal expertise will make the framework better understood.
4. Success in data gathering needs to recognize the bias due to specialty.
5. The results show that the Consistency Ratio (CR) is 10% or below to ensure data consistency. However, if the Consistency Ratio was more than 10%, the rating gathered during the In-depth interview would be revised using the recording as the reference.
6. The results show that the cost structure is the ultimate concern on supplier selection which bears 44.2% of the SOCCER model. Followed by operational capability (23%), customer approach (13.5%), economic performance (8.3%), strategic direction (6.9%), and lastly research & development (4.1%).

REFERENCES

[1] Wahid Z, Daud MRC. (2020) Study on productivity improvement of manual operations in soya sauce factory. IIUM Engineering Journal, 21(1): 202-211.
[2] Tsai CK, Phumchusri N. (2021) Fuzzy analytical hierarchy process for supplier selection: A case study in an electronic component manufacturer. Engineering Journal. 25(8): 73-86.
[3] Rogers SC. (2009) The supply-based advantage: How to link suppliers to your organization's corporate strategy. AMACOM Div American Mgmt Assn.
[4] Shukor JA, Omar MF, Kasim MM, Jamaludin MH, Naim MA. (2018) Assessment of Composting Technologies for Organic Waste Management. Int Journal of Technology. 9(8): 1579-1587.

[5] Fagundes MVC, Keler ÁC, Teles EO, de Melo SABV, Freires FGM. (2021) Multicriteria Decision-Making System for Supplier Selection Considering Risk: A Computational Fuzzy AHP-Based Approach: a computational Fuzzy AHP-based approach. IEEE Latin America Transactions, 19(9): 1564-1572.

[6] Velasquez M, Hester PT. (2013) An analysis of multicriteria decision making methods. International Journal of Operations Research, 10(2): 56-66.

[7] Al Hazza MH, Abdelwahed A, Ali MY, Sidek ABA. (2022) An Integrated Approach for Supplier Evaluation and Selection using the Delphi Method and Analytic Hierarchy Process (AHP): A New Framework. Int. J. Technol, 13(1): 16-25.

[8] Shaik M, Abdul-Kader W. (2011) Green supplier selection generic framework: a multi-attribute utility theory approach. International Journal of Sustainable Engineering, 4(01): 37-56.

[9] Chen CT, Lin CT, Huang SF. (2006) A fuzzy approach for supplier evaluation and selection in supply chain management. International journal of production economics, 102(2): 289-301.

[10] Zhao K, Yu X. (2011) A case based reasoning approach on supplier selection in petroleum enterprises. Expert Systems with Applications, 38(6): 6839-6847.

[11] Sevkli M, Lenny Koh SC, Zaim S, Demirbag M, Tatoglu E. (2007) An application of data envelopment analytic hierarchy process for supplier selection: a case study of BEKO in Turkey. International Journal of Production Research, 45(9): 1973-2003.

[12] Choudhary D, Shankar R. (2014) A goal programming model for joint decision making of inventory lot-size, supplier selection and carrier selection. Computers & Industrial Eng, 71: 1-9.

[13] Fahmi A, Kahraman C, Bilen Ü. (2016) ELECTRE I method using hesitant linguistic term sets: An application to supplier selection. Int J of Computational Intelligence Systems, 9(1): 153-167.

[14] Kaur P, Kumar S. (2013) An Intuitionistic Fuzzy Simple Additive Weighting (IFSAW) Method for Selection of Vendor. IOSR Journal of Business and Management, 15(2): 78-81.

[15] Zouggari A, Benyoucef L. (2012) Simulation based fuzzy TOPSIS approach for group multicriteria supplier selection problem. Engineering Applications of Artificial Intelligence, 25(3): 507-519.

[16] Wan SP, Xu GL, Dong JY. (2017) Supplier selection using ANP and ELECTRE II in interval 2-tuple linguistic environment. Information Sciences, 385: 9-38.

[17] Shaw K, Shankar R, Yadav SS, Thakur L. (2012) Supplier selection using fuzzy AHP and fuzzy multi-objective linear programming for developing low carbon supply chain. Expert Systems with Applications, 39(9): 8182-8192.

[18] Kuo RJ, Lin YJ. (2012) Supplier selection using analytic network process and data envelopment analysis. International Journal of Production Research, 50(11):2852-2863.

[19] Agrawal N, Kant S. (2020) Supplier selection using fuzzy-AHP: a case study. Trends in Manufacturing Processes: Select Proceedings of ICFTMM 2018, 119-127.

[20] Ramanathan R, Abdullah L, Fauadi MHFM, Mohamed MSS, Kamaludin KN. (2023) A Hybrid of Kansei Engineering (KE) And Analytical Hierarchy Process (AHP) to Develop Conceptual Designs of Portable Oil Spill Skimmer. IIUM Engineering Journal, 24(1): 183-198

[21] Bottani E, Rizzi A. (2008) An adapted multicriteria approach to suppliers and products selection - An application oriented to lead-time reduction. Int J Production Economics, 111(2): 763-781.

[22] Ulutaş A. (2019) Supplier Selection by Using a Fuzzy Integrated Model for a Textile Company. Engineering Economics, 30(5): 579-590.

[23] Ho W, Xu X, Dey PK. (2010) Multicriteria decision-making approaches for supplier evaluation and selection: A literature review. European Journal of Operational Research, 202(1): 16-24.

[24] Golden BL, Wasil EA, Harker PT. (1989) The analytic hierarchy processes. Applications and Studies, Berlin, Heidelberg.

[25] Saaty TL. (2008) Decision making with the analytic hierarchy process. International journal of Services Sciences, 1(1): 83-98.

[26] Ariff H, Salit MS, Ismail N, Nukman Y. (2008) Use of analytical hierarchy process (AHP) for selecting the best design concept. Jurnal Teknologi, 49(1): 1-18.

[27] Sipahi S, Timor M. (2010) The analytic hierarchy process and analytic network process: an overview of applications. Management Decision, 48(5): 775-808.

[28] Ishizaka A, Labib A. (2011) Review of the main developments in the analytic hierarchy process. Expert systems with applications, 38(11):14336-14345.

[29] Subramanian N, Ramanathan R. (2012) A review of applications of Analytic Hierarchy Process in operations management. International Journal of Production Economics, 138(2): 215-241.

[30] Bruno, Esposito E, Genovese A, Passaro R. (2012) AHP-based approaches for supplier evaluation: Problems and perspectives. Journal of Purchasing and Supply Management, 18(3): 159-172.

ENHANCING SUSTAINABILITY INDEX PARAMETER USING ANFIS COMPUTATIONAL INTELLIGENCE MODEL

DIAH SEPTIYANA[1,2]*, MOHAMED ABD. RAHMAN[1], TASNIM FIRDAUS MOHAMED ARIFF[1], NOR AIMAN SUKINDAR[1] AND ERRY YULIAN TRIBLAS ADESTA[3]

[1]*Department of Manufacturing and Material Engineering, Kulliyyah of Engineering (KOE), International Islamic University Malaysia (IIUM), Selangor, Malaysia.*
[2]*Department of Industrial Engineering, Faculty of Engineering, Universitas Muhammadiyah Tangerang, Banten, Indonesia.*
[3]*Department of Industrial Engineering Safety and Health, Faculty of Engineering, Universitas Indo Global Mandiri (UIGM), Palembang, Indonesia.*

*Corresponding author: dee.septie@gmail.com

ABSTRACT: The scarcity of water resource is an essential global issue in the 21st century. Therefore, one of the Sustainable Development Goals (SDG) was to ensure the availability and sustainable management of water and sanitation. To do this, it is necessary to assess whether or not the SDG has been followed using the sustainability index. However, there are a lot of sustainability indexes and many of them have the same problem, in which all sustainability index parameters have the same weightage. This problem shows us that every parameter in the sustainability index is equal, while in real life there is no equal parameter. In this paper a weightage for each parameter is proposed to enhance the sustainability index. The method to assess the sustainability index parameters was using a questionnaire by key experts in the water industry. Using ANFIS computational intelligence, the result of the assessment was then fit to the frequent parameters that exist in other sustainability indexes. This proposed method can produce a ranking and weight for each sustainability index parameter and criteria. Using this method, the weightage for each sustainability index parameter can be generated, such as environmental 0.301, engineering 0.214, economic 0.280, and social 0.205.

ABSTRAK: Kekurangan sumber air merupakan isu global yang penting dalam abad ke-21. Oleh itu, salah satu Matlamat Pembangunan Mampan (SDG) adalah bagi memastikan ketersediaan dan pengurusan air dan sanitasi yang berterusan. Bagi melaksanakan ini, adalah perlu untuk menilai sama ada SDG telah diikuti atau tidak menggunakan indeks kemampanan. Walau bagaimanapun, terdapat banyak indeks kemampanan dan kebanyakannya mempunyai masalah yang sama, di mana semua parameter indeks kemampanan mempunyai pemberat yang sama. Masalah ini menunjukkan kepada kita bahawa setiap parameter indeks kemampanan adalah sama, manakala dalam kehidupan sebenar tiada parameter yang sama. Kajian ini merupakan cadangan wajaran pemberat bagi setiap parameter bagi meningkatkan indeks kemampanan. Kaedah bagi menilai parameter indeks kemampanan adalah menggunakan soal selidik oleh pakar utama dalam industri air. Menggunakan kecerdasan pengiraan ANFIS, hasil penilaian kemudiannya diselaraskan dengan parameter kerap yang wujud dalam indeks kemampanan lain. Kaedah yang dicadangkan ini boleh menghasilkan pemeringkatan dan pemberat bagi setiap parameter dan kriteria indeks kemampanan. Menggunakan kaedah ini, wajaran pemberat bagi setiap parameter indeks kemampanan dapat dijana, seperti persekitaran 0.301, kejuruteraan 0.214, ekonomi 0.280, dan sosial 0.205.

KEYWORDS: *Sustainability Index, Computational Intelligence, ANFIS, Weightage, Water Industry.*

1. INTRODUCTION

In 2022, Indonesia's economy improved, with economic progress in both urban and rural areas. Increased employment in the agricultural, manufacturing, transportation, and logistics services sectors contributed to poverty reduction [1]. Likewise, in the Tangerang Region, Indonesia, which has seen an increase in per capita income, the total population is also growing very fast. Consequently, water needs also increase [2]. Apart from having a positive impact, it also poses new challenges in terms of water availability in the future [3, 4]. Water resources are key ecosystem components that sustain life and all social and economic processes. The scarcity of water resources is an essential global issue in the 21st century [3–6] Adequate water availability is challenged by population growth, rapid urbanization, and climate change in meeting urban water needs [7, 8]. Therefore, the second target of SDGs seeks to ensure the availability and sustainable management of water and sanitation until 2030, emphasizing the quality and sustainability of water resources in integrating and balancing the three dimensions of sustainable development such as economic, social, and environmental [9].

Although there is an increasing trend toward using sustainability index to support plans for long-term, medium-term, and short-term sustainable development, the sustainability index is not infallible on its own. This has been the cause for creation of many new sustainability indices such as the SWaM index, the Canadian Water Sustainability index, and many more [10, 11]. Even though there have been some instances of success with the use of these new and current sustainability indices, they are not entirely relevant in all applications. This was brought on by some of the sustainability indices that were created for usage in particular nations or regions [12]. One could argue that there isn't a universal gauge of water sustainability. However, in the current sustainability index, the parameters don't have weightage. The example of this can be seen in the WSSI index. This implies a problem in the sustainability index foundation, in which evaluation components are considered to have equal weight, hence making sustainability measurement inaccurate.

The aim of this paper is to develop the best sustainability index. To achieve that aim, the objective of this paper will be to combine input from various experts and computational intelligence methods [13]. There are many computational intelligence methods, such as neural network, fuzzy, and other [14–16]. However, in this paper ANFIS is prefered because it combines the fuzzy method with the neural network method [17]. This combined approach makes ANFIS preferable compared to the others. ANFIS is preferred due to its transparency in associations (reasoning to generate the membership functions and rules from a set of training data) to reach the optimal predicted value [18]. This paper proposed an ANFIS method to create differing weightage for every sustainability index parameter and create more accurate sustainability measurements.

2. METHOD AND DESIGN

In this section, the method used in this research is explained. The technique used includes questionnaires with experts and ANFIS computational intelligence.

2.1. Adaptive Neuro Fuzzy Inference System (ANFIS)

Zadeh developed the foundation of fuzzy in 1975 based on Linguistic Variables and their application to Approximate Reasoning [19]. The fuzzy rule was subsequently created to model the qualitative elements of human knowledge (reasoning based on experience) and address the issue using those as its foundations [20] Fuzzy logic is typically employed for control, including robot movement, and many others [21–24], however in reality, a fuzzy system can be applied to anything from detection to forecasting [25-26]. In 1985, Takagi and Sugeno

created Fuzzy Sugeno [27]. The fuzzy implications and reasoning system control were shown using this mathematical model [28–30]. The Fuzzy Sugeno equation can be seen in equations 1 through 6.

$$O_i^1 = \mu A_i(x) \ and \ O_i^1 = \mu B_i(x), i = 1, 2 \tag{1}$$

Where:

i = each node of ANFIS architecture.

A, B= is the linguistic label.

x = is the input to node i. (such as small, large, etc.).

Every membership function type is usable at this stage. However, to offer a maximum equal to 1 and a minimum equal to 0, generalized bell types were employed. hence:

$$\mu A i(x) = \frac{1}{1 + \left(\frac{x - c_i}{a_i}\right)^{2*b_i}} \tag{2}$$

where:

a,b,c= is the parameter set.

By multiplying the two input signals, the second step contributes to the firing strength of fuzzy inference. This is represented by each node.

$$O_i^2 = \mu A_i(x).\mu B_i(x), i = 1, 2 \tag{3}$$

The following phase involves applying normalization for each fuzzy inference firing.

$$O_i^3 = \overline{W_i} = \frac{W_i}{W1 + W2}, i = 1, 2 \tag{4}$$

where:

W= is the firing strength of the node.

$\overline{W_i}$= is the normalized firing strength of the node.

The following phase involves a calculation based on the parameters of the rule consequent in the following phase:

$$O_i^4 = \overline{W_i}.F_i = \overline{W_i}.(P_i x + Q_i x + R_i x), i = 1,2 \tag{5}$$

where:

P,Q,R = is the parameter set

The final phase computes the overall output by adding up all of the input signals:

$$O_i^5 = Overall \ Output = \sum_{k=0}^{n} \overline{W_i}.F_i = \frac{\sum_{k=1}^{n} W_i.F_i}{\sum_{k=1}^{n} W_i} \tag{6}$$

Jang Jyh Shing Roger created Adaptive Neuro Fuzzy Inference System (ANFIS) in 1993 based on fuzzy Takagi Sugeno's IF-THEN rules [31]. The fuzzy inference system can adapt organically using the ANFIS approach based on its training data. The Takagi-Sugeno fuzzy inference system serves as the foundation for the artificial neural network technique that makes up ANFIS [32]. This method can combine the benefits of fuzzy logic and neural networks into a single framework: a collection of fuzzy IF-THEN rules that can be learned to estimate nonlinear functions govern how this inference system operates. Consequently, ANFIS is regarded as a Universal Estimator (universal assessor). Equations 2 through 7 make up the core

algorithm used by ANFIS, which uses a Fuzzy Sugeno algorithm. Jang then used gradient descent and chain rule to optimize its parameter. However, to do this, it must be aware of the error rate for data training for each node output, since ANFIS learns through chain rule and gradient descent. The ANFIS algorithm can be seen from equation 7 through Equation 12:

If the data training sets have P numbers of inputs and the i-th position node outputs define 0_i, then the error function may be calculated as follows:

$$E_p = \sum_{m=1}^{\#L}(T_{mp} - 0\,_{mp}^{L})^2 \tag{7}$$

where:

E_p = is the error measure which is the sum of squared errors.

T_{mp} = is the m component from the P output target vector.

$0\,_{mp}^{L}$ = is the m component from the actual output vector that has been delivered by the P input vector.

Hence, the error rate can be calculated as:

$$\frac{\partial E_p}{\partial 0\,_{ip}^{k}} = \sum_{m=1}^{\#k+1} \frac{\partial E_p}{\partial 0_{mp}^{k+1}} \frac{\partial 0_{mp}^{k+1}}{\partial 0\,_{ip}^{k}} \tag{8}$$

Where $1 \leq k \leq$ L-1 is the error rate of an internal node. It is expressed as the linear combination error rate of nodes in the next step. Therefore, for all $1 \leq k \leq$ L and $1 \leq i \leq \#(k)$, the $\frac{\partial E_p}{\partial 0\,_{ip}^{k}}$ can be found using mathematical equations (7) and (8). Thus, a parameter of the adaptive network α can be written as:

$$\frac{\partial E}{\partial \alpha} = \sum_{0*\epsilon S} \frac{\partial E_p}{\partial 0*} \frac{\partial 0*}{\partial \alpha} \tag{9}$$

where:

S= shows the set of nodes whose output depends on α. Derivative for overall error measurement E concerning α is:

$$\frac{\partial E}{\partial \alpha} = \sum_{0*\epsilon S} \frac{\partial E_p}{\partial 0*} \frac{\partial 0*}{\partial \alpha} \tag{10}$$

Therefore, the updated mathematical equations for generic parameter α as follows:

$$\Delta \alpha = -\eta \frac{\partial E}{\partial \alpha} \tag{11}$$

where:

η = is a learning rate.

The learning rate can be written as:

$$\eta = \frac{k}{\sqrt{\sum_\alpha (\frac{\partial E}{\partial \alpha})^2}} \tag{12}$$

where:

k = is the step size or length of each gradient transition in the parametric space.

2.2. Sustainability Index Assessment Using Questionnaire with Key Experts in the Water Industry

The assessment of water supply sustainability might well be classified into three methodological methods that were established sequentially throughout time but are still in use today and are mutually supplemented, enhanced, and integrated [33]. Sustainability Development is defined by three interconnected components (environment, social, and economy), often known as aspects. Since the focus of this study is water management sustainability, the assessment of its sustainability must include extra components such as engineering components related to its intrinsic features [33–35]. The engineering component is adapted to the operating conditions of the water treatment which may affect the management of water utilities and affect environmental, social, and economic aspects. Furthermore, for this paper, the Sustainability Assessment framework is created by grouping key indications into four major components (Engineering, Environmental, Economic, and Social).

The questionnaire's primary components are in connection with the water sustainability index's theoretical model. A literature review was used to construct the framework, which consists of an initial set of components, indicators, and threshold values based on sustainability theories, water management principles, and existing water sustainability indexes from the CWSI, WSSI, WJWSI, WSI, and WPI. Criteria in designing this conceptual framework are critical for reducing researcher bias.

The questionnaire's contents include the respondents' importance ratings in the component/indicator/threshold values on a 4-point Likert scale. Where a value of 4 is 'very important', a value of 3 is 'important', a value of 2 is 'less important' and a value of 1 is 'very less important'. Questionnaires were distributed to several key experts who have backgrounds related to drinking water companies in Indonesia, especially piped drinking water in the Tangerang area. The parameter that is assessed using questionnaires by key experts can be seen in Table 1 below.

Table 1: Sustainability Index Parameter used for Questionnaire.

Component	Indicator	Sustainability Index (Existing)					Literature
		CWSI	WSSI	WJWSI	WSI	WPI	
Engineering	Water Losses	√	-	√	-	-	[36]–[42]
	Operational and Maintenance	√	√	√	√	-	[40], [43]–[46]
Environment	Quality	√	√	√	√	√	[42]–[44], [46]–[48]
	Quantity	√	√	√	√	√	[40], [43]–[46], [48]–[50]
	Continuity	√	√	√	√	√	[40], [43]–[46], [48]–[50]
Economic	Capital Investment	√	√	-	√	-	[46], [51], [52]
	Cost recovery & Operation and maintenance cost	-	√	-	-	-	[46], [51], [52]
	Ability to Pay	√	-	-	-	√	[43], [44], [48]
Social	Access for Water Supply	√	√	√	√	√	[40], [42], [43], [46], [51]
	Access for Sanitation	√	-	√	-	√	[42], [43], [46], [53], [54]

3. RESULT AND DISCUSSION

The findings of this research are presented in this section. The first step would be to start with the survey results from important water industry operation specialists. The ANFIS computational intelligence used the survey results to provide values that fit the frequently occurring sustainable index parameter. The weightage and ranking for each parameter of the sustainable index are generated in this way.

3.1 Questionnaire Result from The Experts

The assessment of the water sustainability index was carried out with a key expert in the water industry. Figure 1 shows the ranking of the sustainability index parameter as a result of using 10 parameters from different sustainability indices in Table 1.

(a)
(b)

(c)
(d)
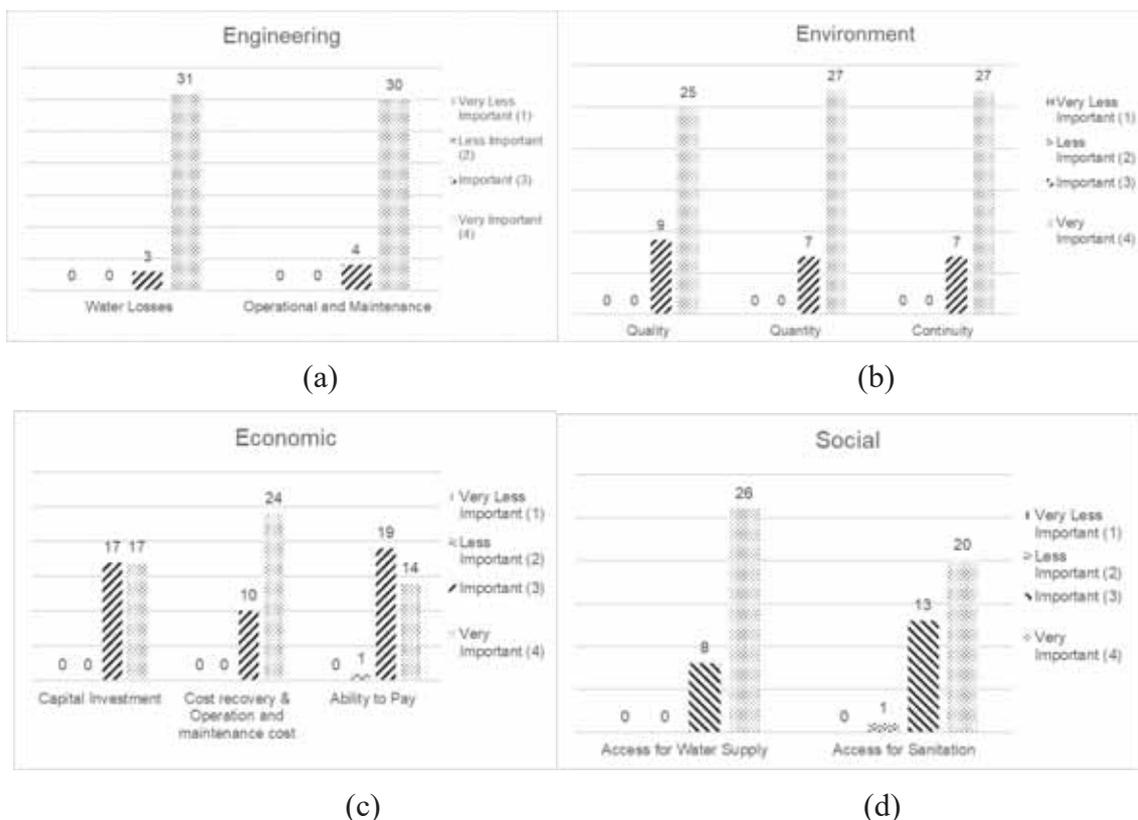
Fig. 1. Water Sustainability Index Assessment by Key Experts in the Water Industry. (a) Engineering Parameter, (b) Environment Parameter, (c) Economic Parameter, (d) Social Parameter.

3.2 ANFIS Fitting Model for Questionnaire and frequent Parameter Sustainability Index

Assisted by the ANFIS algorithm, a fitting model for the questionnaire result and frequent parameters that exist in most Sustainability Indices can be developed easily. The questioner data in Table 1 is data training for ANFIS. Thus, using Equation 2 and Equation 6, an ANFIS Fitting Model can be produced as shown in figure 2, Table 2, Table 3, Table 4, and Equation 13. In Figure 2, the generalized bell membership function is used for ANFIS fuzzification model fitting. Its membership function was formed by three parameters such as a, b and c. Using those parameters, the visualization can be built such that parameter a defines the width of the membership function input, while parameter b defines the shape of the curve on either side of the midland, and parameter c defines the center point of the membership function.

Fig. 2. ANFIS Fuzzification fitting model for the questionnaire and frequent parameter
Sustainability Index.

Table 2: ANFIS Fuzzification fitting model for the questionnaire and frequent
parameter Sustainability Index.

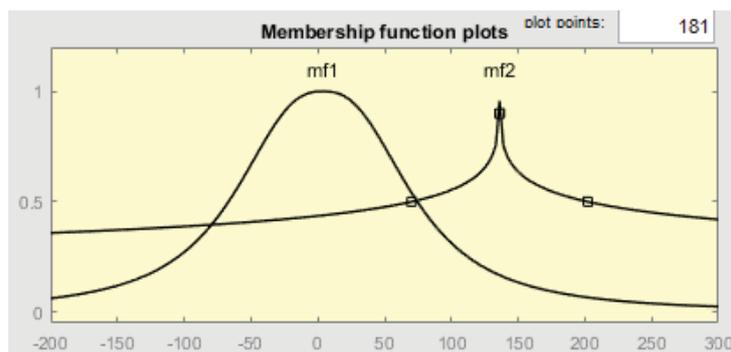| Fuzzification Input	a	b	c
mf1	70.7	1.283	3.825
mf2	65.97	0.1787	136.1

Table 3: ANFIS inference for Sustainable Index Parameter Weighting

Rule Index	Rule Firing	Rule Output
1	If (input is mf1)	Then Output is F1
2	If (input is mf2)	Then Output is F2

Table 4: ANFIS fitting model Fuzzy Constant Output

Output Constant Index (Fi)	Value
1	-2.484
2	5.234

Thus, the fitting ANFIS model for the questionnaire and frequent parameters that exist in
Sustainability Index can be seen in equation 13 below:

$$
= \frac{\sum_{k=0}^{n} \frac{1}{1+\left[\left(\frac{x-c_i}{a_i}\right)^{2*b_i}\right]} + \frac{1}{1+\left[\left(\frac{x-c_i}{a_i}\right)^{2*b_i}\right]} F_{i=1}^{n}}{\sum_{k=0}^{n} \frac{1}{1+\left[\left(\frac{x-c_i}{a_i}\right)^{2*b_i}\right]} + \frac{1}{1+\left[\left(\frac{x-c_i}{a_i}\right)^{2*b_i}\right]}}
\tag{13}
$$

Where:

x = Input variables in this research were questionnaire.

a, b, c= Defines the width of the membership function input as a, while b defines the shape
of the curve on either side of the midland, and c defines the center point of the membership
function.

$F_{i=1}^{n}$ = Defines the output variables' constant Level generated automatically by ANFIS
based on the number of fuzzy rules/inferences. In this case, it refers to the parameter that
often appears in the sustainability index.

n= Number of fuzzy rules/inferences.

Using the Fuzzy Sugeno fitting model on equation 13, the weight for each Sustainability Index parameter was determined. However, the value was normalized so that it ranged between a minimum of 0 and maximum of 1. Hence, calculating the total value of all fitting values from ANFIS is necessary to get this weight. The fitting value of each parameter will then be divided by the ANFIS's overall fitting value. This will establish the relative importance of each parameter which, when added together, will result in a value of 1. Therefore, the normalized weightage for each sustainability parameter can be seen in Table 5.

Table 5: Sustainability Index weight and ranking parameter result using ANFIS fitting.

Total Criteria Weight	Indicator	Index	Questionn aire Score	ANFIS Fitting	Ranking	Normalize Weight
Engineering (0.214)	Water Losses	E1	133	3,77	1	0,108
	Operational and Maintenance	E2	132	3,72	2	0,106
Environment (0.301)	Quality	E4	127	3,5	7	0,100
	Quantity	E6	129	3,58	3	0,102
	Continuity	S4	126	3,46	8	0,099
Economic (0.280)	Capital Investment	E7	120	3,23	9	0,092
	Cost recovery & Operation and maintenance cost	E8	115	3,04	10	0,087
	Ability to Pay	E10	128	3,54	6	0,101
Social (0.205)	Access for Water Supply	S1	129	3,58	4	0,102
	Access for Sanitation	S2	129	3,58	5	0,102

By employing this technique, it can be seen that the environmental criteria have a higher weightage (0.301) than other criteria, such as engineering (0.214), economic (0.280), and social (0.205). The fact that this criterion has a higher weightage indicates that it is more crucial in achieving a high sustainability index.

3.3 Enhanced Sustainability Index Model using new Sustainability Parameter Weightage

Each sustainability index such as CWSI, WSSI, WJWSI, WSI, and WPI can easily be enhanced using this new sustainability index parameter weightage, one such example is by enhancing the development of WSSI by Odjegba [46]. The original WSSI model can be seen on equation 14 below:

$$WSSI = \frac{S_{obt}}{MS_{obt}} \, x \, 100 \tag{14}$$

Where:

S_{obt} = Score obtained from observation.

MS_{obt} = Maximum score obtained from total parameter.

Assisted with new sustainability index parameter weightage, the enhanced WSSI sustainability index can be seen in equation 15 below:

$$WSSI_{eh} = \frac{\sum_i^n P_i \cdot W_i}{\sum_i^n P_{tot} \cdot W_i} \, x \, 100 \tag{15}$$

$WSSI_{eh}$ = Enhanced WSSI Sustainability Index

P_i = Sustainability Index parameter score obtained from observation, minimum 1 and maximum 5.

W_i = Sustainability Index parameter normalize weightage obtained from Table 3.

P_{tot} = Total Sustainability Index parameter score, maximum score for each parameter is 5.

Using this final form of enhanced WSSI sustainability index, it can reach the original WSSI score from 0 to 100 points.

4. CONCLUSION

The sustainability index has been enhanced using the new sustainability index parameter weightage in this paper. This sustainability index parameter weightage was in general form so it can be applied to the different sustainability indices. Using a leading expert in the water operations questions, the popular sustainability index parameter from CWSI, WSSI, WJWSI, WSI, and WPI was evaluated carefully. The outcome of this evaluation was then computed utilizing the computational intelligence of ANFIS to fit the frequent parameter seen in the other 5 sustainability indices before. The rank and weightage for each indicator and criterion sustainability index parameter can be proposed for engineering 0.214, economics 0.280, social 0.205, and the environment 0.301. It can be concluded that using the suggested method has successfully generated weightage for each sustainability index parameter, thus creating a better water sustainability index.

REFERENCES

[1]   World Bank Poverty & Equity Brief East Asia & Pacific – Indonesia [https://databankfiles.worldbank.org/public/ddpext_download/poverty/987B9C90-CB9F-4D93-AE8C-750588BF00QA/current/Global_POVEQ_IDN.pdf]

[2]   Badan Pusat Statistik Provinsi Banten Proyeksi Penduduk Provinsi Banten 2010-2020 [https://banten.bps.go.id/publication/2015/03/31/a723ef5e69db53cfc3eca058/proyeksi-penduduk-provinsi-banten-2010-2020.html]

[3]   E. Iglesias and M. Blanco (2008). New directions in water resources management: The role of water pricing policies. Water Resour Res, 44. doi: 10.1029/2006WR005708.

[4]   Q. Tang (2020). Global change hydrology: Terrestrial water cycle and global change. Sci China Earth Sci, 63:459–462. doi: 10.1007/s11430-019-9559-9.

[5]   N. W. Arnell (1999). Climate change and global water resources, Global Environmental Change, 9:S31-S49. doi: 10.1016/S0959-3780(99)00017-5.

[6]   E. Marris (2016). Water scarcity' affects four billion people each year, Nature. doi: 10.1038/nature.2016.19370.

[7]   Climate Change Adaptation vs Mitigation: A Fiscal Perspective [https://www.lse.ac.uk/granthaminstitute/wp-content/uploads/2015/03/Barrage_2015_Adaptation_Mitigation_Fiscal.pdf]

[8]   D. S. Brookshire and D. Whittington (1993). Water resources issues in the developing countries. Water Resour Res., 29:1883-1888, doi: 10.1029/92WR02988.

[9]   United Nations The sustainable development goals report 2019 [https://unstats.un.org/sdgs/report/2019/The-Sustainable-Development-Goals-Report-2019.pdf]

[10] M. Maiolo and D. Pantusa (2019). Sustainable water management index, swam_index. Cogent Eng., 6, doi: 10.1080/23311916.2019.1603817.

[11] The Canadian Water Sustainability Index (CWSI): Case Study Report [http://www.bibliotheque.assnat.qc.ca/01/MONO/2006/03/842674.pdf]

[12] I. Juwana, B. J. C. Perera, and N. Muttil (2010). A water sustainability index for West Java. Part 1: Developing the conceptual framework. Water Science and Technology, 62:1629–1640. doi: 10.2166/wst.2010.452.

[13] I. Juwana, B. J. C. Perera, and N. Muttil (2010). A water sustainability index for West Java. Part 1: Developing the conceptual framework. Water Science and Technology, 62: 1641–1652. doi: 10.2166/wst.2010.453.

[14] G. P. N. Hakim et al. (2023). Levenberg Marquardt artificial neural network model for self-organising networks implementation in wireless sensor network. IET Wireless Sensor Systems, 1–14. doi: 10.1049/wss2.12052.

[15] S. R. Utama, A. Firdausi, and G. P. N. Hakim (2022). Control and Monitoring Automatic Floodgate Based on NodeMCU and IOT with Fuzzy Logic Testing. Journal of Robotics and Control (JRC), 3:14-17. doi: 10.18196/jrc.v3i1.11199.

[16] A. Adriansyah, Y. Gunardi, B. Badaruddin, and E. Ihsanto (2015). Goal-seeking Behavior-based Mobile Robot Using Particle Swarm Fuzzy Controller. TELKOMNIKA, 13:528-538. doi: 10.12928/telkomnika.v13i2.1111.

[17] J. S. R. Jang (1993). ANFIS: Adaptive-Network-Based Fuzzy Inference System. IEEE Trans Syst Man Cybern., 23:665–685. doi: 10.1109/21.256541.

[18] H. A. O. Cruz, R. N. A. Nascimento, J. P. L. Araujo, E. G. Pelaes, and G. P. S. Cavalcante (2017). Methodologies for path loss prediction in LTE-1.8 GHz networks using neuro-fuzzy and ANN. In SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference: 27-30 August 2017; Brazil. Edited by Aguas de Lindoia.

[19] L. A. Zadeh (1975). The Concept of a Linguistic Variable and its Application to Approximate Reasoning. Inf Sci (N Y), 8:199–249. doi:10.1016/0020-0255(75)90036-5.

[20] N. Yusof, N. Bahiah, Mohd. Shahizan, and Y. Chun, "A Concise Fuzzy Rule Base to Reason Student Performance Based on Rough-Fuzzy Approach," in Fuzzy Inference System - Theory and Applications, 2012. doi: 10.5772/37773.

[21] G. P. N. Hakim, D. Septiyana, and S. Iswanto (2022). Survey Paper Artificial and Computational Intelligence in the Internet of Things and Wireless Sensor Network. Journal of Robotics and Control (JRC), 3:439–454. doi: 10.18196/jrc.v3i4.15539.

[22] I. R. F. Arif, A. Firdausi, and G. P. N. Hakim (2021). Nebulizer operational time control based on drug volume and droplet size using fuzzy sugeno method. Journal of Robotics and Control (JRC) 2:94-97. doi: 10.18196/jrc.2259.

[23] G. P. N. Hakim, R. Muwardi, M. Yunita, and D. Septiyana (2022). Fuzzy Mamdani performance water chiller control optimization using fuzzy adaptive neuro fuzzy inference system assisted. Indonesian Journal of Electrical Engineering and Computer Science, 28: 1388~1395. doi: 10.11591/ijeecs.v28.i3.

[24] G. P. N. Hakim, D. Septiyana, and S. Iswanto (2022). Survey Paper Artificial and Computational Intelligence in the Internet of Things and Wireless Sensor Network. Journal of Robotics and Control (JRC), 3:439–454. doi: 10.18196/jrc.v3i4.15539.

[25] D. Sahid and M. Alaydrus. (2020) Multi Sensor Fire Detection in Low Voltage Electrical Panel Using Modular Fuzzy Logic. In 2nd International Conference on Broadband Communications, Wireless Sensors and Powering on 28-30 September; Jakarta. pp 31-35.

[26] G. P. N. Hakim et al. (2022). Near Ground Pathloss Propagation Model Using Adaptive Neuro Fuzzy Inference System for Wireless Sensor Network Communication in Forest, Jungle, and Open Dirt Road Environments. Sensors, 22:3266-3288 doi: 10.3390/s22093267.

[27] T. Takagi and M. Sugeno (1985). Fuzzy Identification of Systems and Its Applications to Modeling and Control. IEEE Trans Syst Man Cybern, 15:116-132. doi: 10.1109/TSMC.1985.6313399.

[28] C. H. Chiu and Y. F. Peng (2019). Design of Takagi-Sugeno fuzzy control scheme for real world system control. Sustainability, 11. doi: 10.3390/su11143855.

[29] X. Tang, D. Ning, H. Du, W. Li, and W. Wen (2020). Takagi-Sugeno Fuzzy Model-Based Semi-Active Control for the Seat Suspension with an Electrorheological Damper. IEEE Access, 8: 98027-98037. doi: 10.1109/ACCESS.2020.2995214.

[30] X. Tang, D. Ning, H. Du, W. Li, Y. Gao, and W. Wen (2020). A takagi-sugeno fuzzy model-based control strategy for variable stiffness and variable damping suspension. IEEE Access, 8: 71628-71641. doi: 10.1109/ACCESS.2020.2983998.

[31] J. S. R. Jang (1993). ANFIS: Adaptive-Network-Based Fuzzy Inference System. IEEE Trans Syst Man Cybern, 23:665–685. doi: 10.1109/21.256541.

[32] J.-S. R. Jang, C.-T. Sun, and E. Mizutani (1997) Neuro-Fuzzy and Soft Computing-A Computational Approach to Learning and Machine Intelligence, Prentice Hall.

[33] K. Carden and N. P. Armitage (2013). Assessing urban water sustainability in South Africa - Not just performance measurement. Water SA, 39. doi: 10.4314/wsa.v39i3.1.

[34] Z. Chen, H. H. Ngo, and W. Guo (2012). A critical review on sustainability assessment of recycled water schemes. Science of the Total Environment, 426:13-31. doi: 10.1016/j.scitotenv.2012.03.055.

[35] Sustainability indicators and the assessment of Urban Water Systems: Seeking for a robust way forward [http://www.switchurbanwater.eu/outputs/pdfs/WP1-1_PAP_Sustainability_indicators_for_assessment_of_UWS.pdf]

[36] Evaluation of Sustainability index for Urban Water Management System [http://www.ipcbee.com/vol4/58-ICESD2011D30014.pdf]

[37] T. Al-Washali, S. Sharma, F. Al-Nozaily, M. Haidera, and M. Kennedy (2019). Monitoring nonrevenue water performance in intermittent supply. Water, 11. doi: 10.3390/w11061220.

[38] L. J. N. Jones, D. Kong, B. T. Tan, and P. Rassiah (2021). Non-revenue water in Malaysia: Influence of water distribution pipe types. Sustainability, 13. doi: 10.3390/su13042310.

[39] M. Farley and S. Trow (2003) Losses in water distribution networks: A practitioner's guide to assessment, London, IWA Publishing.

[40] A. Adham, S. Seeyan, R. Abed, K. Mahdi, M. Riksen, and C. Ritsema (2022). Sustainability of the Al-Abila Dam in the Western Desert of Iraq. Water, 14. doi: 10.3390/w14040586.

[41] The Canadian Water Sustainability Index (CWSI): Case Study Report [https://core.ac.uk/reader/80536600]

[42] I. Juwana, B. J. C. Perera, and N. Muttil (2009) Conceptual framework for the development of West Java water sustainability index," in 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation: Interfacing Modelling and Simulation with Mathematical and Computational Sciences on Cairns, Australia 13-17 July 2009: pp 3343-3349.

[43] Canadian Water Sustainability Index [https://publications.gc.ca/Collection/PH4-38-2007E.pdf]

[44] Linking technology choice with operation and maintenance in the context of community water supply and sanitation [https://apps.who.int/iris/handle/10665/42538].

[45] H. M. L. Chaves and S. Alipaz (2007). An integrated indicator based on basin hydrology, environment, life, and policy: The watershed sustainability index. Water Resources Management, 21:883–895. doi: 10.1007/s11269-006-9107-2.

[46] E. Odjegba, G. Oluwasanya, O. Idowu, O. Shittu, and G. Brion (2020). Sustainability indices and risk analysis of drinking water systems in Southwest Nigeria. Journal of Water Supply: Research and Technology - AQUA, 69:591–603. doi: 10.2166/aqua.2020.002.

[47] Water Safety Plans Managing drinking-water quality from catchment to consumer Water, Sanitation and Health Protection and the Human Environment World Health Organization Geneva [https://apps.who.int/iris/handle/10665/42890].

[48] C. Sullivan (2002). Calculating a Water Poverty Index. World Dev, 30:1195–1210. doi: 10.1016/S0305-750X(02)00035-9.

[49] D. Koo, K. Piratla, and C. J. Matthews (2015). Towards Sustainable Water Supply: Schematic Development of Big Data Collection Using Internet of Things (IoT). Procedia Engineering, 118:489-497. doi: 10.1016/j.proeng.2015.08.465.

[50] C. O. de Castro, Ó. C. S. Loureiro, A. V. Santos, J. Silva, and W. B. Rauen (2017). Water sustainability assessment for the region of Curitiba. International Journal of Sustainable Building Technology and Urban Development, 184-194 doi: 10.12972/susb.20170016.

[51] H. M. L. Chaves and S. Alipaz (2007). An integrated indicator based on basin hydrology, environment, life, and policy: The watershed sustainability index. Water Resources Management, 21:883–895. doi: 10.1007/s11269-006-9107-2.

[52] M. Van den Belt, T. Bowen, K. Slee, and V. Forgie (2013). Flood Protection: Highlighting an Investment Trap Between Built and Natural Capital. J Am Water Resour Assoc, 49: 681-692. doi: 10.1111/jawr.12063.

[53] C. Dong, G. Schoups, and N. Van de Giesen (2013). Scenario development for water resource planning and management: A review. Technological Forecasting and Social Change, 4: 749-761. doi: 10.1016/j.techfore.2012.09.015.

[54] D. P. LOUCKS (1997). Quantifying trends in system sustainability. Hydrological Sciences Journal, 42: 513-530. doi: 10.1080/02626669709492051.

USE OF TIO$_2$ AS A REINFORCEMENT OF CASSAVA STARCH/PVA COMPOSITES ON MOISTURE-RESISTANT PROPERTIES OF TRIBOELECTRIC NANOGENERATOR (TENG) FILM

ARIS ANSORI[1,2]*, SUDJITO SOEPARMAN[1], DENNY WIDHIYANURIYAWAN[1] AND TEGUH DWI WIDODO[1]

[1]*Department of Mechanical Engineering, Universitas Brawijaya*
Jl. Mayjen Haryono 167 Malang 65145, East Java, Indonesia
[2]*Department of Mechanical Engineering, Universitas Negeri Surabaya,*
Campus Ketintang Surabaya 60231, East Java, Indonesia

*_Corresponding author: arisansori@unesa.ac.id_

ABSTRACT: High humidity environments can accelerate the transmission, neutralization, or dissipation of frictional charges on the frictional surface of solid-solid triboelectric nanogenerator films (TENGs), which can reduce the output power. The moisture resistance properties of the TENG triboelectric film are needed to overcome these problems. Therefore, this study discusses the role of the TiO$_2$ nanofiller in cassava starch (CS) and polyvinyl alcohol (PVA) nanocomposite matrix that can increase triboelectricity through the formation of hydrogen bonds and the provision of oxygen-free electrons. The research method was to incorporate different concentrations of TiO$_2$ nanoparticles (0%, 0.5%, 1%, 5%, 10% wt, and 15% wt) into the CS-PVA nanocomposite matrix using the solvent casting method. The results showed an increase in surface polarity which was more triboelectric-positive due to the CS-PVA hydroxyl group interacting with water molecules. Increasing the concentration above 5% wt TiO$_2$ increases the density of the CS-PVA nanocomposite film which can significantly reduce water vapor permeability (WVP) and increase water resistance. The TENG performance of the CS-PVA/TiO$_2$ nanocomposite film with a concentration of 15% wt TiO$_2$ under conditions of high humidity (RH, 95%) resulted in an output voltage of 2.5-fold (~70.5 V to ~180 V), and the output current increased 2.6-fold (~5.2 µA to ~13.7 µA).

ABSTRAK: Persekitaran berkelembapan tinggi dapat mempercepatkan penghantaran, peneutralan, atau pelesapan cas geseran pada permukaan geseran filem nanopengeluaran triboelektrik pepejal (TENG), di mana mengurangkan pengeluaran tenaga. Sifat rintangan lembapan filem triboelektrik TENG diperlukan bagi mengatasi masalah ini. Oleh itu, kajian ini membincangkan peranan pengisi nano TiO2 dalam matriks nanokomposit kanji ubi kayu (CS) dan polivinil alkohol (PVA) yang dapat meningkatkan triboelektrik melalui pembentukan ikatan hidrogen dan bekalan elektron bebas oksigen. Kaedah kajian ini adalah dengan menggabungkan kepekatan nanozarah TiO2 berbeza (0%, 0.5%, 1%, 5%, 10%, dan 15%) ke dalam matriks nanokomposit CS-PVA menggunakan kaedah tuangan pelarut. Dapatan kajian menunjukkan peningkatan kekutuban permukaan yang lebih positif-triboelektrik adalah disebabkan oleh kumpulan hidroksil CS-PVA yang berinteraksi dengan molekul air. Pertambahan jisim kepekatan TiO2 melebihi 5% meningkatkan ketumpatan filem nanokomposit CS-PVA yang boleh mengurangkan kebolehtelapan wap air dan meningkatkan rintangan air dengan ketara. Prestasi TENG filem nanokomposit CS-PVA/TiO2 dengan jisim kepekatan TiO2 15% dalam keadaan berkelembapan tinggi (RH, 95%) menghasilkan voltan keluaran sebanyak 2.5 kali ganda

(~70,5 V kepada ~ 180 V), dan arus keluaran meningkat 2.6 kali ganda (~ 5,2 µA kepada ~ 13,7 µA).

KEYWORDS: *nanocomposite films; TiO₂; cassava starch/PVA; moisture resistant; triboelectric nanogenerators (TENG)*

1. INTRODUCTION

A triboelectric nanogenerator (TENG) is a micro/nanoscale energy harvesting device that utilizes static electric charges or the triboelectric effect between two materials that have different electronegativities [1]. TENG's working mechanism is a combination of contact electrification and electrostatic induction to convert mechanical energy that is abundant in the environment into electrical energy. The TENG structure consists of two electrodes made of copper or aluminum, where electrode 1 is attached to the tribo-positive layer and electrode 2 is attached to the tribo-negative layer, the TENG structure is presented in Fig. 1a. There are four working modes of TENGs, namely; contact mode, slide mode, single electrode mode, and freestanding mode [2,3]. The working modes of TENGs are presented in Fig. 1.
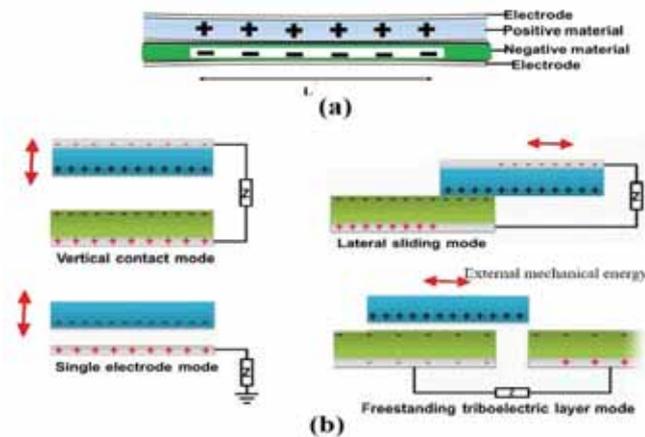

Fig. 1: The working modes of TENG (a) contact electrification phenomena between different triboelectric materials, (b) The four working modes of TENG.

The schematic diagrams of working TENG, are presented in Fig. 2. The working TENG mechanism that converts mechanical energy into electricity consists of four steps. In step I, the two layers of the triboelectric layer (tribo-positive layer and tribo-negative layer) are in contact with one another, where the contact of two surfaces of the triboelectric film layer has not resulted in charge (Fig. 2a). In step II, the application of mechanical energy to the tribo-positive layer 1 causes friction to occur, reduces the contact area between the tribo-positive layer 1 and the tribo-negative layer 2, and causes a charge difference on the surface of the triboelectric film layer. Then, electrons will flow from electrode 1 to electrode 2 through the outer circuit (resistor) to balance the surface charge on the two triboelectric layers until the contact of the two triboelectric layers is free (Fig. 2b). In step III, the sliding of the tribo-positive layer 1 optimally causes the contact of the two triboelectric layers to completely separate and there is a charge balance between the two triboelectric layers, so that there is no flow of electrons between electrode 1 and electrode 2 (Fig. 2c). In step IV, the tribo-positive layer 1 moves and creates friction and surface contact again between the tribo-positive layer 1 and the tribo-negative layer 2 in the opposite contact direction in stage II, this causes electrons to flow through an external circuit (resistor) from electrode 2 to

electrode 1 to balance the surface charges on the two triboelectric layers with opposite directions of electron flow (Fig. 2d) until the two triboelectric layers return to full contact again at the stage I position. The sliding and closing of the triboelectric layers periodically generate an alternating current [4].
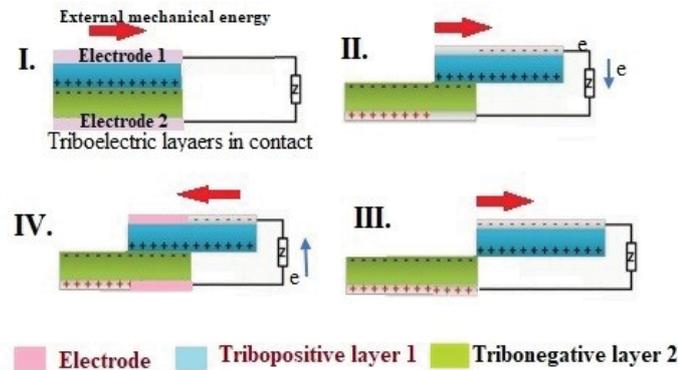

Fig. 2: Schematic diagrams working of TENG.

TENG was first developed by Wang et al. in 2013, The research was carried out utilizing the triboelectric effect of materials to harvest mechanical energy into electricity and showed high energy conversion efficiency. [5]. TENG has become a promising energy conversion device to convert mechanical energy into electrical energy due to the capacity of triboelectric materials. TENG can be a solution for new and sustainable energy harvesting, such as wind [5], ocean waves [6], human movement [7], rotational motion [8-11], flowing water [12], vibration [13], and other energy sources.

In general, the TENG device consists of two friction materials and electrodes at the top and bottom of the device wherein, physical contact between two friction materials with different levels of electronegativity can induce a triboelectric charge which can produce a potential decrease when separated by mechanical forces and increase the flow of electrons between the electrodes via an external circuit. The dominant factor that determines the performance of TENG's output is the TENG friction material, such as the use of nanocomposite friction materials on a nanoscale which has the advantages of mechanical strength, toughness, flexibility, transparency, dielectric, thermal, or electrical conductivity properties which show an effective and efficient TENG performance improvement [6,7].

In the previous research report, the use of nanocomposite films showed an improved TENG performance. One such nanocomposite material is polydimethylsiloxane (PDMS), a silicon-based organic polymer compound (organosilicon). PDMS nanocomposite films with TiO_2 deposition can cause changes in oxygen vacancies on the PDMS surface which can contribute to electron exchange and trapping and have an impact on increasing TENG output power [8]. Poly (vinylidene fluoride) (PVDF) polymer with TiO_2 NPs as filler can improve dielectric properties [9]. Polydimethylsiloxane (PDMS) with different TiO_x weight ratios as a function of dielectric constant control with 5% rutile TiO_x and 7% TiO_x anatase phase produces the highest TENG output of ~180 V/8.2 μA and 211.6 V/8.7 μA [10]. TiO_2 nanoparticles 0.2% wt in Portland cement increased the mechanical strength 1.3-fold and 3-fold the power density of 265 mW/m^2 than pure Portland cement [11]. However, TENG's output power may decrease due to high humidity environmental conditions. When the humidity is high, the water molecules in the air absorb on the surface of the friction film pair and form a conductive water film, which can increase the frictional interface conductivity and accelerate the frictional charge transfer or neutralize the charge on the next

contact. These conditions, will generate different frictional charges when the TENG solid-solid films rub against each other [12].

Several methods have been reported to solve the problem, such as; creating a hydrophobic or superhydrophobic film to reduce the adsorption of water molecules on the friction surface and loosen the texture of the structure to increase the permeability of the material [13,14]. For example, Shen et al., using a hydrophobic polyvinylidene fluoride (PVDF) film, showed the performance of TENG at 55% relative humidity with an output current of 28 μA, a voltage of 345 V, and a power density of 1.3 W/m^2 [15]. Lv et al. made a TENG single electrode using a hydrophobic ionic liquid that showed high stability performance in various weather conditions (humidity up to 80%) [16]. Zhou et al. used a TENG with superhydrophobic PDMS interlayer film that can maintain electrical output under initial conditions up to relative humidity of 80% [17]. However, modification of the hydrophobic surface when humidity is high cannot completely solve the problem of decreasing the output performance of solid-solid TENG. In addition, the hydrophobic or superhydrophobic film process requires a complicated film-making process, and its manufacture is not easy to apply to some materials.

Currently, the use of biomaterials rich in hydroxyl groups has begun to be developed as a TENG triboelectric film, several studies have shown that TENGs output performance is very good by utilizing the ability to bind water molecules by forming hydrogen bonds. In addition, biomaterials are expected to replace synthetic polymers (PDMS, PVDF, PET, and Teflon) which still dominate the TENG film. Meanwhile, several applications of TENG triboelectric films are made from biomaterials rich in hydroxyl groups, such as; pure potato starch exhibits high output performance of 60 mV to 300 mV per 4 cm^2 area [18]. The triboelectric film of tapioca starch rich in hydroxyl groups can enhance the triboelectricity effect by improving water molecules on the friction film surface through hydrogen bonding [19,20]. Potato starch biomaterials with 0.5% CaCl$_2$ showed an increase in triboelectric performance; the highest voltage output (1.2 V) was three times higher than that of pure starch [21]. In addition, the use of starch biomaterials has advantages, such as biocompatibility and biodegradability [18,19][21-23], abundant availability, easy processing, bring inexpensive, possessing hydroxyl groups, amorphous crystal structure, and rheological properties that can provide the required resistance for TENG triboelectric film applications [24].

However, starch biomaterials still have problems in the application of TENG triboelectric films, such as poor mechanical properties, ease of dissolving on prolonged contact with water, and poor tribological properties in high humidity environments and low moisture resistance when exposed to high humidity conditions as a friction film. Therefore, to increase the functionality of starch biomaterials is to modify starch biomaterials into nanocomposite films. To the best of our knowledge, the use of TiO$_2$ nanoparticles as a starch biomaterial nanofiller with moisture resistance characteristics has not been reported. Meanwhile, in this study, a type of nanocomposite film was made using a composite matrix of cassava starch (CS) and PVA with TiO$_2$ nanofiller to increase moisture resistance. Polyvinyl alcohol (PVA), rich in hydroxyl groups with the addition of TiO$_2$, will cause intra/intermolecular interactions between TiO$_2$ nanoparticles and adjacent OH groups of PVA through hydrogen bonds and form stable complex composites. PVA compatibility increases with the addition of TiO$_2$ into the film due to the formation of C-O-Ti. Meanwhile, the starch/PVA mixture changed the structure of the starch/PVA composite film by forming C-H bond groups from alkyl groups in both starch and PVA. After incorporating TiO$_2$ filler into the starch/PVA mixture, C-O bonds will form at the macromolecular level, and intramolecular hydrogen bonds are formed between two neighboring OH groups on the

same side of the plane of the carbon chain. This allows for increased miscibility and compatibility between starch/PVA and TiO_2 with the formation of hydrogen and C-O-Ti bonds which can affect the characteristics of the nanocomposite film [25-27]. Meanwhile, the basis for using TiO_2 nanofillers is that TiO_2 contains oxygen atoms that can be easily dispersed in a polar biopolymer matrix without surface modification. In addition, chemical bonds and electrostatic interactions between oxygen atoms on the metal oxide surface and functional groups in the polymer matrix make the nanoparticles disperse in the polymer matrix [28].

Meanwhile, the special emphasis in this study is to investigate the effect of TiO_2 in the CS-PVA polymer blend on the physicochemical, water resistance, and electrical properties of the nanocomposite films and the TENG characterization of the CS-PVA/TiO_2 nanocomposite films as positive friction surfaces and thin-film commercial polyimides as negative friction partners. We compared the results and analyzed the effect of TiO_2 on the moisture properties of the CS-PVA nanocomposite films. The results of the measurement of the output voltage and current of TENG CS-PVA nanocomposite film without TiO_2 produced an output voltage and current of ~25.5 V and ~3.6 µA and with 0.5% wt TiO_2 output voltage of ~50 V and an output current of ~5,9 µA (relative humidity /RH, 15%). In this case, there was a significant increase in the effect from TiO_2 in the CS-PVA polymer blend. The optimal increase of TENG output voltage and current at a weight ratio of 15% wt TiO_2 at high humidity (RH, 95%) achieved a 2.5-fold increase in output voltage (~70.5 V to ~180 V), and an output current of 2.6-fold (~5.2 µA to ~13.7 µA). The addition of a weight ratio above 5% wt TiO_2 can reduce the water vapor permeability by 65%. Therefore, these CS-PVA/TiO_2 nanocomposite films have the potential for TENG films under high humidity environmental conditions.

2. EXPERIMENTAL METHODS

2.1 Materials

The starch used in the experiment was cassava starch produced by PT. Budi Starch & Sweetener.Tbk. in Indonesia, polyvinyl alcohol (PVA) with an average molecular weight of Mw = 89,000–98,000, was procured from Sigma Aldrich. Anatase TiO_2 nanoparticles with a particle size of 20 nm were purchased from Sigma-Aldrich, glycerol (99% purity, clear, colorless, and density: 1.261 g/cm^3), aquades, and other chemicals NaOH, HCl, and NaCl were used to control humidity (RH, 15%, 55%, and 95%) from Sigma-Aldrich and Merck.

2.2 Fabrication of Cassava Starch (CS)-PVA/TiO₂ Nanocomposite Films

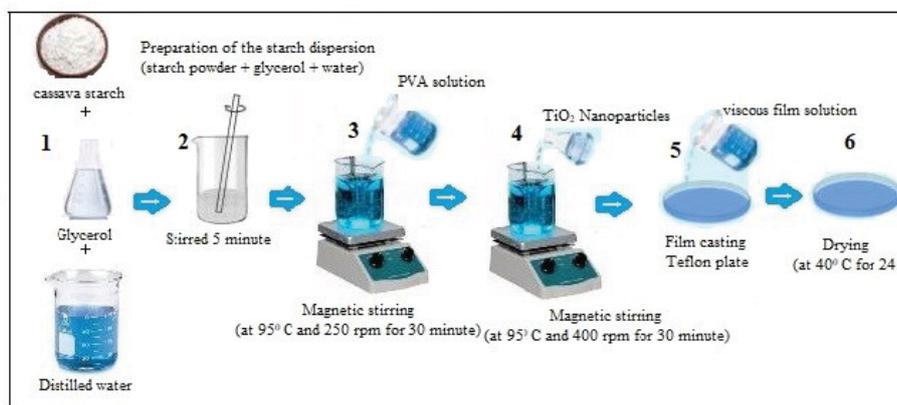Fig. 3: Preparation of CS/PVA-TiO₂ nanocomposite films by solvent casting method.

Figure 3 describes the making of cassava starch (CS)/PVA nanocomposite films with TiO_2 nanofiller using the solvent casting method. The step of making cassava starch (CS)/PVA-TiO_2 nanocomposite films consists of six steps. In step 1, the constant weight ratio of the CS/PVA mixture (70:30), cassava starch (4.2 g) was determined and dispersed with distilled water (50 mL) in borosilicate glass, and the addition of glycerol (3 g) served as a plasticizer. In step 2, the mixture of CS powder and glycerol was mechanically stirred for 5 minutes until a homogeneous solution was produced. In step 3, the mixture of CS powder and glycerol was stirred with a magnetic stirrer at 250 rpm at a temperature of 95 °C for 30 minutes to obtain a gelatinized starch suspension. Meanwhile, a solution of 1.8 g PVA with 40 mL of distilled water was prepared in the same way in another borosilicate glass, and a clear PVA solution was produced. Mixing of the two solutions (CS solution/PVA solution) was carried out at room temperature with continued stirring with a magnetic stirrer (250 rpm, 95 °C) for 30 minutes. In step 4, the TiO_2 was incorporated into the CS/PVA solution with different polymer weight ratios (0%, 0.5%, 1%, 5%, 10%, and 15% wt), and continuous stirring was carried out using a magnetic stirrer at 400 rpm, and temperature 95 °C for 30 minutes. In step 5, the production of nanocomposite films occurred, where the resulting CS/PVA-TiO_2 viscous solution was poured into Teflon molds (radius 50 mm) with the same volume of solution in each mold to produce relatively the same thickness of film samples; residual air was removed from the solution in a vacuum for 10 minutes using a vacuum machine. In step 6, the drying of the film was carried out in a hot air oven at 40 °C for 24 hours. The dry film was then carefully peeled off from the Teflon mold and produced a film with an average thickness of 100 μm. The resulting nanocomposite films were kept in a zippered bag until tests were carried out to reduce water absorption according to the Ramirez method [29].

2.3 Fabrication of CS-PVA/TiO₂ Nanocomposite Film-Based TENGs

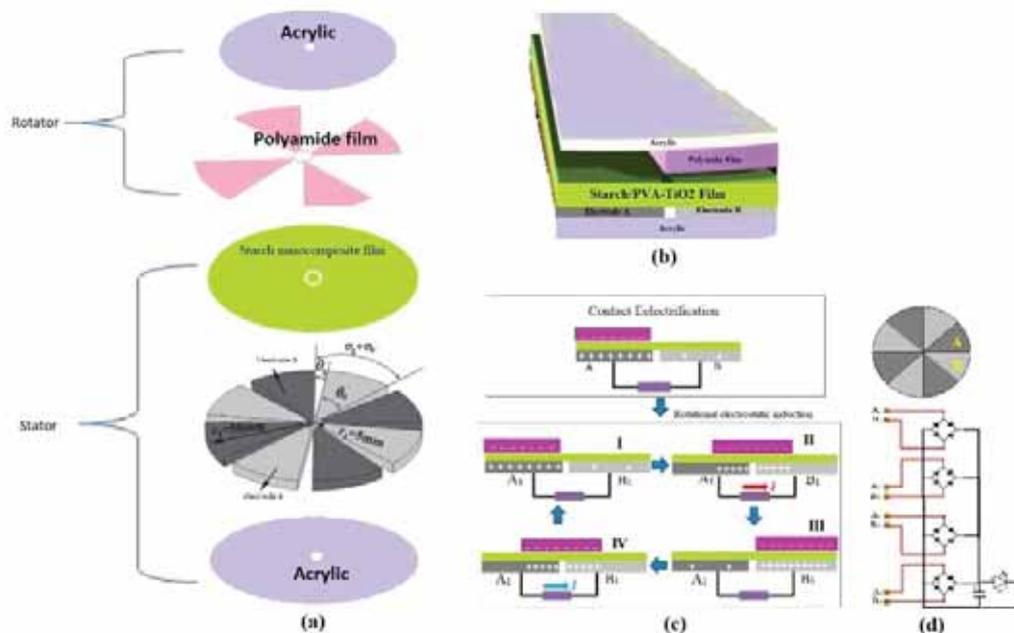Figure 4 describes the structure of the TENG mode rotary-disk freestanding (RDF-TENG).

Fig. 4: Schematic of triboelectric nanogenerator based on CS-PVA/TiO_2 nanocomposite film (a) Structure of rotor and stator (b) TENG component arrangement (c) Freestanding electret rotary (FER) working mechanism (d) TENG electrical circuit.

Rotor structure parameters with the number of film segments (n) = 4, outer radius (r_2) = 50 mm, inner radius (r_1) = 5 mm, were made of commercial polyimide thin film (0.03 mm) as a tribo-negative film glued to an acrylic surface (radius = 50 mm). Meanwhile, the stator structure parameters consisting of CS/PVA-TiO$_2$ nanocomposite films were cut to form an equal-interval electrode model with the number of film segments (n) = 8, outer radius (r_2) = 50 mm, inner radius (r_1) = 5 mm, and the nanocomposite film thickness was 100 μm. Next, the cut nanocomposite film was glued to the copper foil (electrode film A and B). Where the tip of the copper foil electrode was slightly exaggerated to connect the external circuit with the help of a conducting wire so that electrons could flow into the external circuit. External mechanical energy used a DC electric motor that was connected to the TENG rotor shaft in a freestanding rotational mode. Meanwhile, the gap between the rotor and stator was controlled with a distance of 0.10 mm to keep the friction between the surface of the polyimide film and the CS/PVA-TiO$_2$ nanocomposite film stable. The freestanding-electret rotary (FER) TENG mode has four steps. In step I, electrode A is parallel to the polyamide electret film, and there is no electrostatic induction effect of rotating rotor. In step II, the polyamide electret is parallel to the positions of electrodes A and B causing the charge at electrode A to be transferred to electrode B through an external load (resistor) and generating a current and voltage across the load resistor. In step III, the polyamide electret is parallel to electrode B, and in step IV, the polyamide electret will go to electrode segment A and when parallel to electrode A and electrode B will generate current and voltage. A schematic of triboelectric nanogenerator based on CS-PVA/TiO$_2$ nanocomposite film is presented in Fig. 4.

2.4 Physicochemical Characterizations

2.4.1 Chemical, Crystalline Structure and Morphological Characterization

The functional groups of the nanocomposite films were analyzed using Fourier transform infrared spectroscopy (FTIR) with an IRPrestige-21 spectrophotometer (Shimadzu, USA). FTIR measurements were carried out at room temperature (25 ± 1^0 C) in the wave number range of 600–4000 cm^{-1} with a resolution of 2 cm^{-1} and an average of more than 64 scans. The degree of crystallinity of the film samples was recorded using a PANalytical Xpert Pro diffractometer with Cu Kα radiation at 40 kV and 40 mA, using a 20 mm Ni filter. All sample films were scanned between $2\theta = 5^0$ to 69.99^0 at 5^0/min at a measurement temperature of 25 °C. Morphological analysis of nanocomposite film by Scanning Electron Microscope (FESEM) FEI Quanta FEG 650 to explore the effect of TiO$_2$ filler on the cell structure of the CS-PVA polymer mixture. The dry film sample was sprayed with gold-palladium under vacuum conditions to increase its conductivity. The test was carried out at a voltage of 15xKV with a magnification of 500x.

2.4.2 Water Vapor Permeability (WVP)

The combination of Fick's and Henry's laws for the diffusion of vapors or gases through the film is used as the basis of the method for calculating WVP. The standard ASTM E96 test method (ASTM, 1993) was used to test the WVP of nanocomposite films [41]. In the initial step of the test, the film sample was cut into a circle with a radius of 1.75 cm and tightly mounted on the top surface of a special permeation cup containing CaCl$_2$ anhydrous. After weighing the initial weight, the cup was stored in a desiccator containing a saturated NaCl solution with a constant relative humidity (RH) of 75% at 25 °C. The temperature in the desiccator was maintained at 25 °C throughout the experiment. Permeation cup weights were recorded at 24-hour intervals until a constant weight gain was achieved. The amount of water vapor permeating through the film was calculated from the increase in cup weight

and plotted as a function of time. Measurement of WVP ($gm^{-1}s^{-1} Pa^{-1}$) was performed three times using Eq. (1).

$$WVP = \frac{m.d}{A.t.\Delta p}\left(\frac{g}{m.P_a.s}\right) \tag{1}$$

where m = increase in cell mass over a certain period, t = time, d = average film thickness, A = effective film area, and p = difference in water vapor pressure on both sides of the film. The WVP value of the nanocomposite films in this paper represents the effective permeability, which reflects the moisture barrier properties of the films under the humidity and temperature conditions used.

2.4.3 Degradability Tests

The degradability test was carried out to analyze the resistance of the CS-PVA/TiO$_2$ nanocomposite films in water. In the initial step of testing, the sample nanocomposite film was cut to a size of (2 x 3 cm), then the sample film was immersed in distilled water for a certain time until the sample film was destroyed.

2.5 Electrical Output Measurement

The performance test of the nanocomposite film on the TENG unit was carried out in a closed cylindrical chamber with the same humidity control. The cylinder chamber was equipped with two inlets to circulate dry air and moist air, and one outlet was used to regulate air circulation in the cylinder to achieve the required relative humidity in the room accurately. The external mechanical energy of the rotor used a DC V = 24-volt, 0-500 rpm electric motor. Measurement of output voltage and current used a digital oscilloscope (DSOX6004A Digital Storage Oscilloscope) with a load resistance of 100 MΩ, before the testing film was put in a closed box for 24 hours with adjusted humidity control (15%, 55 %, and 95%) with a saturated salt solution.

3. RESULTS AND DISCUSSION

3.1 Analysis of Chemical, Structural, and Morphological of Nanocomposite Films

Structural analysis can be observed from the FTIR spectrum of the material in the range of 4000-700 cm^{-1} are presented in Fig. 5a, where the molecular bond spectra of the CS-PVA polymer matrix and the change after incorporation of TiO$_2$ filler can be seen, which has several characteristic peaks reflecting the structure of the two biopolymers. The transmittance band of about 3285 cm^{-1} was determined from the stretching vibration of the hydroxyl group (–OH) in CS and PVA polymers having a hydroxyl-rich average. The peak transmittance band of 1641.42 cm^{-1} indicates the formation of hydrogen bonds in pure starch due to the flexion of the –OH molecule and shows the hygroscopic nature of the starch polymer. Characteristics of the bending of C-OH are in the 1205 cm^{-1} transmittance band. The peak at 770-1120 cm^{-1} is associated with the stretching of the C–O bond of the macromolecular portion, where an intramolecular hydrogen bond is formed between two adjacent OH groups that are on the same side of the plane of the C–C carbon chain of the C–O–C group of the unit glucose in starch [37].

Pure PVA showed hydroxyl and acetate groups, where the intermolecular and intramolecular hydrogen –OH bonds were shown in a wider band at 3296.35 cm^{-1}. The structure of the alkaline group confirmed the peak in the transmission band 1480-1190 cm^{-1} associated with the planar buckling vibrations of C–H and O–H in both CS and PVA. The

presence of a carbonyl group (C=O) is indicated by the 1732 cm^{-1} transmission band and a hydrocarbon (alkane) group in PVA is indicated by a peak at 1373.32 cm^{-1}, as previously reported by Abdullah [30]. Meanwhile, the chemical interaction between CS-PVA polymer molecules that causes changes in the characteristic spectral peaks of the composite, the C–OH hydrocarbon group is estimated to come from the C–O vibrations, which is confirmed by the peak of 1000-1250 cm^{-1}. In addition, the vibration of the –OH group confirmed the absorption peak in the range of 3300-3000 cm^{-1}. The formation of the C–H alkene functional group due to the mixing of starch-PVA which was confirmed at a wavelength of 675-995 cm^{-1} and C–H alkane which was confirmed at a wavelength of 2850-2970 cm^{-1}.
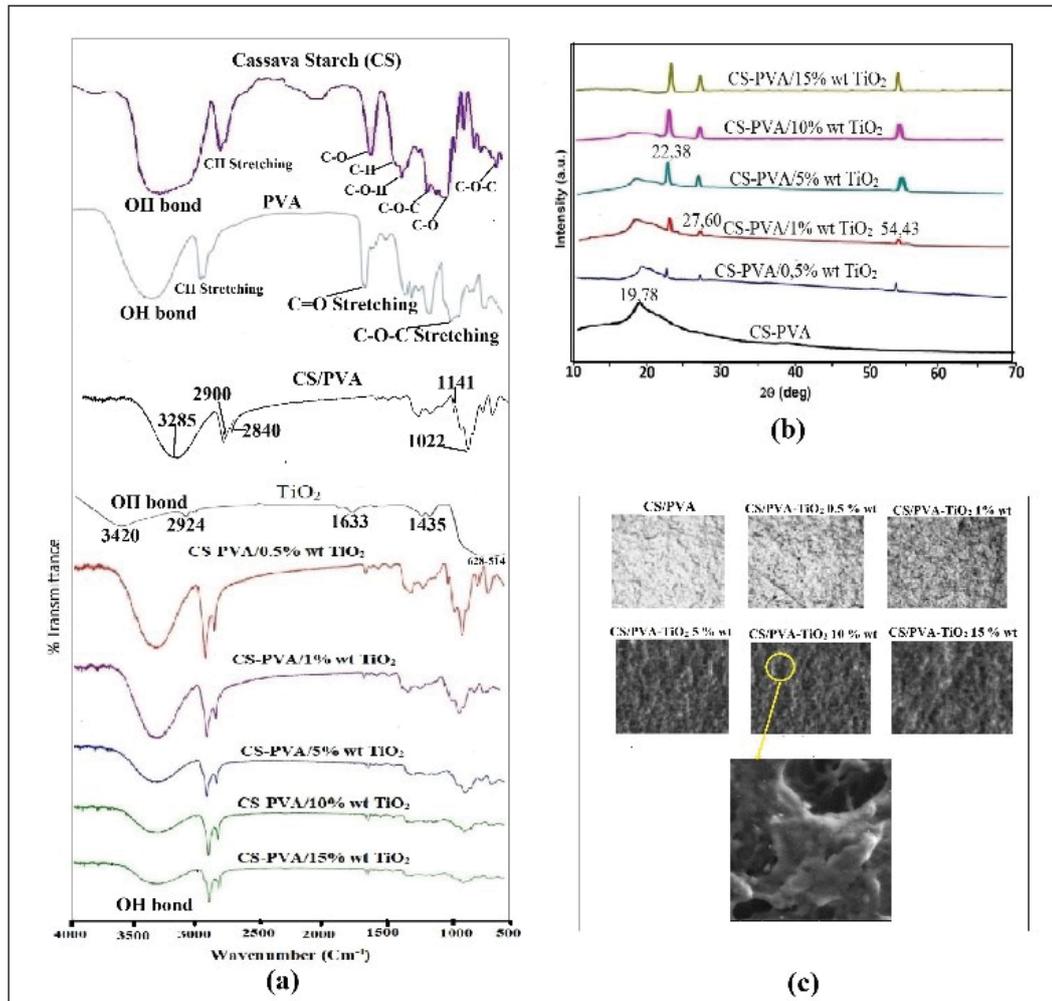
Fig. 5: Physicochemical characterizations of CS-PVA/TIO$_2$ nanocomposite films.
(a) FTIR spectra of cassava starch-PVA/TiO$_2$, (b) XRD pattern of cassava starch film,
(c) SEM of CS-PVA nanocomposite film.

In addition, the C–H functional group of the aromatic ring was confirmed at the wavelength of 690-900 cm^{-1}. The mixing of CS-PVA polymer showed flexural vibrations of the hydrogen bonding of the –OH group at 1710.86 cm^{-1} and the C–O strain at 1050-1300 cm^{-1}. The stronger interaction of CS-PVA molecules is shown by intermolecular interactions through the formation of hydrogen bonds with water vapor molecules which can increase triboelectricity. At low concentrations (below 15% TiO$_2$ wt) in the CS-PVA polymer mixture in the FTIR spectrum, the vibration peaks did not show any additional peaks. This may be due to the overlapping bands of the CS-PVA components. However, the

increase in TiO_2 concentration (above 1% TiO_2 wt) showed a slight shift in the peak positions of several bands in the range of 900-1100 cm^{-1}. It is assumed that the absorption band at 1000 cm^{-1} is associated with the TiO_2 crystal domains hydrated into the CS-PVA polymer matrix. In addition, the titanium ions of the TiO_2 particles can interact with the hydroxyl groups of CS-PVA, which make the nanocomposite films highly compatible.

Figure 5b shows the XRD pattern of the sample film, the cassava starch film produced peaks at $2\theta = 15.73^0$, 16.51^0, 17.23^0, 19.69^0, 22.14^0, and 24.37^0 whereas, an important diffraction peak was located at 19.69^0, which indicates strong intermolecular and intramolecular hydrogen bonds and is indicative of a semi-crystalline structure. This is not much different from the signal of the 20^0 peaks of cassava starch reported by Bergo et al. [31], and Adamu et al. [32]. Meanwhile, the CS-PVA composite sample formulation produced peaks at $2\theta = 13.11^0$, 17.21^0, 19.76^0, and 24.54^0, which indicated a change in peak. These peaks can reveal that polymer cassava starch (CS) is dispersed in PVA, which can change the characterization of the films [33,34]. The incorporation of TiO_2 filler into CS-PVA with variations in TiO_2 concentration below 1% wt showed peak changes that were not easily visible at low concentrations. However, increasing the concentration above 1% by weight showed peaks at $2\theta = 22.38^0$, 27.60^0, and 54.43^0 in semicrystalline structure nanocomposite films and indicated the presence of titanium dioxide embedded in the CS-PVA polymer.

Figure 5c shows the morphology of the polymer mixture of cassava starch (CS)-PVA and the morphological changes of the nanocomposite films after the addition of TiO_2 can be observed from the material micrograph. The ability of starch granules to interact with various ceramic and mineral particles can hold these particles, which causes the added TiO_2 to adhere to the surface of the starch granules well. Meanwhile, mixing CS-PVA suspension with TiO_2 particles, during the gelatinization process, the TiO_2 particles remain attached to the surface of the starch granules and interact with the continuous phase macromolecules of water-soluble amylose and PVA, this process most likely occurs through hydrogen bonding. The SEM photos show morphologically the resulting film has a heterogeneous surface relief. The TiO_2 filler particles were evenly distributed within the polymer matrix. This can be seen from the sample of the film with a low concentration of TiO_2 (0.5% wt), the filler particles are still individual. This is in contrast with the addition of TiO_2 concentration of more than 1% wt indicating the association of filler particles in the composite. In addition, the presence of TiO_2 particles with a high ratio in the film can create a winding path that prolongs the transport of water vapor molecules.

3.2 Water Vapor Permeability (WVP)

Water vapor permeability (WVP) was used as an indicator of the film's ability to absorb water vapor. WVP values of nanocomposite films from the studied samples are presented in Table 1. WVP values of CS-PVA nanocomposite films with TiO_2 are lower than those without TiO_2. The WVP value was inversely correlated with the addition of TiO_2 concentration in the CS-PVA mixture matrix. WVP value of SC-PVA/TiO_2 nanocomposite films decreased with the addition of more than 1% wt TiO_2. This is due to the presence of hydroxyl groups in the CS-PV mixture making the water vapor molecules easy to diffuse. The addition of TiO_2 nanoparticles formed a denser nanocomposite film structure compared to the unfilled film. The presence of TiO_2 in the CS-PVA matrix forms pathways in the polymer that prolong the transport of water vapor molecules. This result is not much different from the results of previous research reports, the effect of TiO_2 nanofiller in polymer films causes a decrease in moisture penetration of corn starch/PVA-TiO_2 films,

[26], Potato starch/lactucin-TiO$_2$ composite films [35], and Potato Starch/Montmorillonite-TiO$_2$ composite films [36].

Table 1: Physical properties of CS/PVA-TiO$_2$ composites

TiO$_2$ content % wt	WVP value ($\times 10^{-7}$ g/m h Pa)
0	9.46 $\pm$ 0.84
0.5	8.38 $\pm$ 0.80
1	7.54 $\pm$ 0.78
5	7.39 $\pm$ 0.75
10	7.47 $\pm$ 0.58
15	7.48 $\pm$ 0.55

3.3 Degradability Testing

The degree of damage film caused by water molecules was tested by immersing the nanocomposite film in distilled water (Fig. 6a).
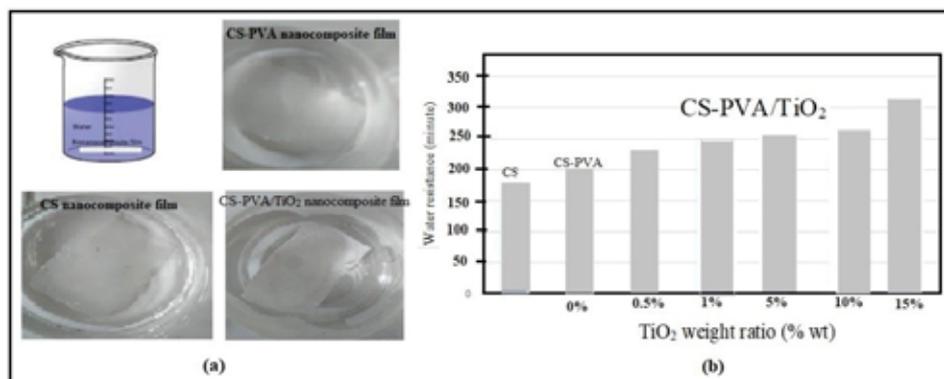
Fig. 6: (a) Test method (b) water-resistance of CS-PVA/TIO$_2$ nanocomposite films.

The damage rate of CS-PVA nanocomposite films with TiO$_2$ increased 3-fold compared to without TiO$_2$. Figure 6b shows the water resistance of CS, CS-PVA, and SC-PVA/TiO$_2$ films, the resistance of CS films is the lowest compared to CS-PVA and SC-PVA/TiO$_2$ films. The durability of CS films is affected by the number of hydroxyl groups in the polymer molecules, where the hydroxyl groups bind to water molecules by forming hydrogen bonds. The increase in the number of hydroxyl groups in the CS-PVA film can occur from the intramolecular interaction of the CS-PVA polymer mixture, which can significantly increase the binding of water molecules. The use of reinforcement TiO$_2$ in the CS-PVA composite increased the durability of the film in water for 315 minutes longer than the CS-PVA film for 200 minutes. The most optimal increase was seen at a concentration of 15% wt TiO$_2$. The presence of TiO$_2$ in the CS-PVA polymer mixture made the film denser which increased the absorption capacity of water molecules. In addition, the hydrophobic nature of TiO$_2$ can also slow down the absorption of water molecules on the film surface.

3.4 Electrical TENG

The schematic diagram of the TENG output test is shown in Fig. 7. Conditioning for humidity levels of 15%, 55%, and 95% were carried out by placing the TENG in a closed cylindrical chamber.

Fig. 7: TENG performance measurements.

The TENG operating mode rotary freestanding (RF-TENG). The stator comprises six sectors of a CS-PVA/TiO$_2$ nanocomposite film connected to a copper foil electrode, the TENG output circuit uses a rectifier diode to convert AC to DC, and a capacitor as electrical storage before being channeled to the LED load. The rotor rotation speed was controlled stable (ω = 200 rad/s), and the load was 50 MΩ. TENG 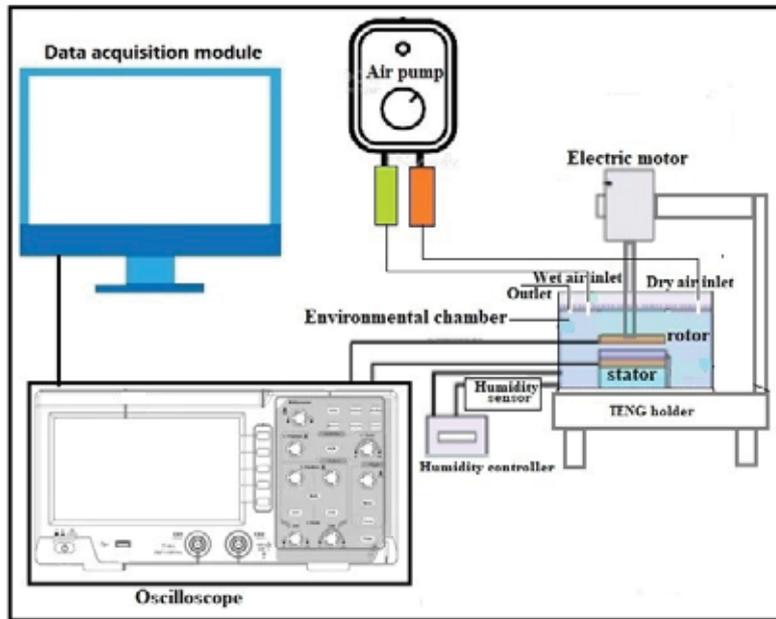output performance test with CS-PVA triboelectric film without TiO$_2$ at a low humidity level (RH, 15%) resulted in an output voltage and current of ~25.5 V and ~3.6 μA, while the effect of adding 0.5% wt TiO$_2$ in the matrix of the CS-PVA nanocomposite showed a 1.9-fold increase in output voltage from ~25.5 V to 50 V, and a 1.6-fold increase in output current from ~3.6 μA to ~5.9 μA.

Meanwhile, the addition of TiO$_2$ in the weight ratio (0% to 15% wt TiO$_2$) showed a linear increase in voltage and current, respectively, as shown in (Fig. 8b-d). Optimal output voltage and current were achieved at a concentration of 5%wt TiO$_2$ with a value of 117 V and 10.2 μA (Fig. 8b). This increase in output voltage and current is due to an increase in the value of the dielectric constant due to the presence of TiO$_2$ NPs in the CS-PVA polymer mixture which can increase the triboelectric of the film. These results show similarities with some of the results of previous research reports, such as an increase in dielectric properties of poly (vinylidene fluoride) film with TiO$_2$ NP deposition [9], Dielectric constant changes linearly as a function of weight ratio of PDMS embedded TiO$_x$ NP [10,37], TiO$_2$ doping on Portland Cement [11], natural rubber (NR)-TiO$_2$ [38].

In addition, the effect of differences in relative humidity factors (RH, 15%, 55%, and 95%) on the output performance of CS-PVA/TiO$_2$ film-based TENG at a weight ratio above 5% TiO$_2$ shows that the resulting voltage and current increase are directly correlated with increased humidity levels (Fig. 8a-c.). The output performance of TENG when the relative humidity condition increased from 15% to 55% (ratio of 15% weight of TiO$_2$) showed an increase in the output voltage of TENG 1.8 -fold from ~70.5 V to ~130 V, and the output current increased by 1.1-fold from ~5.2 μA to ~5.6 μA. Enhanced relative humidity of the cylinder chamber is carried out by continuously injecting wet air into the closed cylinder chamber to increase the humidity from 55% to 95% as shown in Fig. 8b-c. Under conditions of relatively high humidity (95%), TENG can produce a 1.4-fold increase in TENG output

voltage from ~130 V to ~180 V and a 2.4-fold increase in current from ~5.6 μA to ~13.7 μA. However, the increase in TENG output voltage and current was lower at a TiO_2 weight ratio of less than 5%, about 1.1 times. Meanwhile, the increase in TENG's optimal output voltage and current when humidity is high (RH, 95%) with a concentration of 15% wt TiO_2 can reach 2.5-fold from ~70.5 V to ~180 V, and the current increases 2.6-fold from ~5.2 μA to ~13.7 μA.
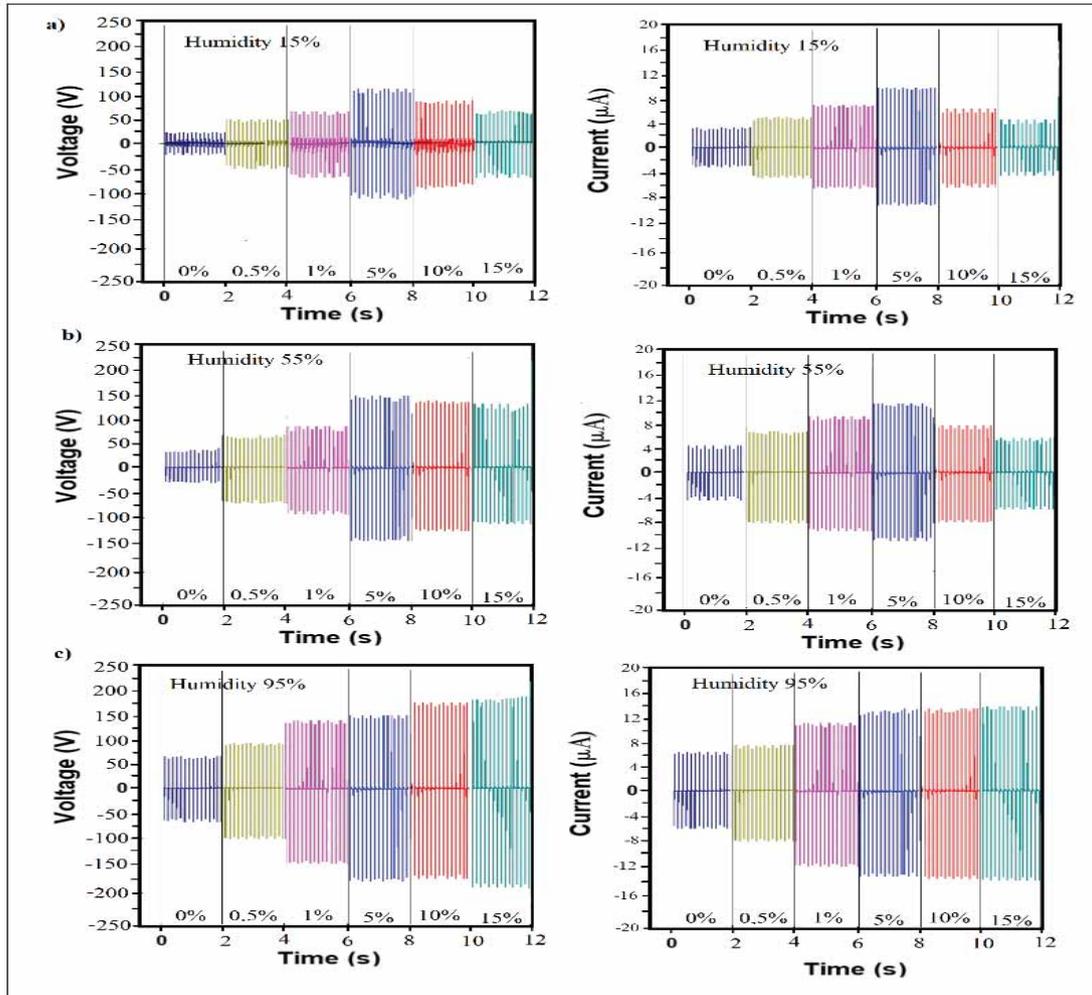
Fig. 8: Output voltage and current at film humidity level:
a) humidity 15%, c) humidity 55%, and c) humidity 95%.

The increase in TENG output voltage and current is caused by the formation of hydrogen bonds between hydroxyl groups and water molecules (when humidity is high) on the surface of the CS-PVA/TiO_2 nanocomposite film increasing the positive triboelectrification properties of the film. This is not much different from the research reported by Wang et al. using starch polymers rich in hydroxyl groups with the formation of hydrogen bonds with water molecules when humidity is high can increase the output current of TENG based on starch films [39], PVA films rich in hydroxyl chains when humidity is high. High water molecules participate in triboelectric charging [40]. Meanwhile, the effect of the presence of 15% wt TiO_2 concentration in the CS-PVA nanocomposite film showed the most optimal increase at high humidity. The output voltage and current increased with a value to 1.9-fold voltage and 1.8-fold current due to the increase in the density of the nanocomposite film. The density of the film can increase the absorption

of water molecules when humidity is high. In addition, TiO_2 in the CS-PVA matrix can form a winding path that can prolong the transport of water molecules, enhancing the triboelectric properties of the film.

4. CONCLUSIONS

This study presents a method for producing cassava starch (CS)/PVA nanocomposite films using TiO_2 reinforcement. The effects of incorporation of TiO_2 on the structure and properties of the CS/PVA composite were investigated. The FTIR spectra indicate the response of hydroxyl groups in the CS/PVA composite increased after inserting TiO_2 nanoparticles with the formation of hydrogen bonds and C-O-Ti bonds. The interaction between Ti ions and hydroxyl groups forms a complex composite mixture through intra/intermolecular hydrogen bonds. The XRD analysis reveals the TiO_2 concentration (> 1% wt) increased the crystallinity of the nanocomposite film with a semicrystalline structure and variations in crystal shape and uniform crystal size. The SEM photo indicates the TiO_2 can be evenly distributed in the film surface of the CS/PVA composite and resulting in a compatible nanocomposite film. The physical properties of the film increased significantly with the addition of TiO_2 (> 1% wt) with a decreased water vapor permeability (WVP) value $7.39\pm0.75 \times 10^{-7}$ g/m h Pa, and the degree of film damage in water decreased by 1.6-fold compared to CS/PVA nanocomposite films without TiO_2 reinforcement. Optimal performance of the CS-PVA/TiO_2 nanocomposite film as a tribo-positive friction film for TENG with a 5% wt TiO_2 concentration, while at high humidity (95% RH) with a 15% wt TiO_2 concentration, the output voltage and current of TENG increased 1.9 -fold respectively and the current is 1.8-fold due to the formation of water molecule bonds with the hydroxyl chains, and the availability of free electrons in the film due to the presence of TiO_2 bonds in the CS-PVA matrix. The findings of this study will promote starch-based nanocomposite films that have great potential for tribo-triboelectric nanogenerator (TENG) film applications for high-humidity environmental conditions.

ACKNOWLEDGMENT

REFERENCES

[1] Dzhardimalieva GI, Yadav BC, Lifintseva TV, Uflyand IE. (2021). Polymer chemistry underpinning materials for triboelectric nanogenerators (TENGs): Recent trends. Eur. Polym. J, 142: 110-163. https://doi.org/10.1016/j.eurpolymj.2020.110163
[2] Wang, J., Wu, C., Dai, Y., Zhao, Z., Wang, A., Zhang, T., & Wang, Z. L. (2017). Achieving ultrahigh triboelectric charge density for efficient energy harvesting. *Nature Communications*, *8*(1), 1–7. https://doi.org/10.1038/s41467-017-00131-4
[3] Xia K, Fu J, Xu Z. (2020). Multiple-frequency high-output triboelectric nanogenerator based on a water balloon for all-weather water wave energy harvesting. Adv. Energy Mater, 10(28): 1-9. https://doi.org/10.1002/aenm.202000426
[4] Yoon HJ, Ryu H, Kim SW. (2018). Sustainable powering triboelectric nanogenerators: Approaches and the path towards efficient use. Nano Energy, 51(June): 270-285. https://doi.org/10.1016/j.nanoen.2018.06.075

[5] Song, G., Kim, Y., Yu, S., Kim, M. O., Park, S. H., Cho, S. M., Velusamy, D. B., Cho, S. H., Kim, K. L., Kim, J., Kim, E., & Park, C. (2015). Molecularly Engineered Surface Triboelectric Nanogenerator by Self-Assembled Monolayers (METS). *Chemistry of Materials*, *27*(13), 4749–4755. https://doi.org/10.1021/acs.chemmater.5b01507

[6] Mallakpour S, Jarang N. (2016). Mechanical, thermal and optical properties of nanocomposite films prepared by solution mixing of poly(vinyl alcohol) with titania nanoparticles modified with citric acid and Vitamin C. J. Plast. Film Sheeting, 32(3): 293-316. https://doi.org/10.1177/8756087915597024

[7] Khushboo, & Azad, P. (2019). Design and implementation of conductor-to-dielectric lateral sliding TENG mode for low power electronics. In *Advances in Intelligent Systems and Computing* (Vol. 698, Issue May). Springer Singapore. https://doi.org/10.1007/978-981-13-1819-1_17

[8] Park, H. W., Huynh, N. D., Kim, W., Lee, C., Nam, Y., Lee, S., Chung, K. B., & Choi, D. (2018). Electron blocking layer-based interfacial design for highly-enhanced triboelectric nanogenerators. *Nano Energy*, *50*, 9–15. https://doi.org/10.1016/j.nanoen.2018.05.024

[9] Kum-onsa, P., Chanlek, N., Manyam, J., Thongbai, P., Harnchana, V., Phromviyo, N., & Chindaprasirt, P. (2021). Gold-nanoparticle-deposited TiO2 nanorod/poly(Vinylidene fluoride) composites with enhanced dielectric performance. *Polymers*, *13*(13), 1–14. https://doi.org/10.3390/polym13132064

[10] Huynh ND, Park H, Chung K, Choi D. (2018). Effect on TENG performance by phase control of TiO x nanoparticles. Compos. Res, 31(6): 365-370. http://dx.doi.org/10.7234/composres.2018.31.6.365

[11] Sintusiri J, Harnchana V, Amornkitbamrung V, Wongsa A, Chindaprasirt P. (2020). Portland Cement-TiO$_2$ triboelectric nanogenerator for robust large-scale mechanical energy harvesting and instantaneous motion sensor applications. Nano Energy, 74: 104802. https://doi.org/10.1016/j.nanoen.2020.104802

[12] Nguyen V, Yang R. (2013). Effect of humidity and pressure on the triboelectric nanogenerator. Nano Energy, 2(5): 604-608. https://doi.org/10.1016/j.nanoen.2013.07.012

[13] Zhang JH, Hao X. (2020). Enhancing output performances and output retention rates of triboelectric nanogenerators via a design of composite inner-layers with coupling effect and self-assembled outer-layers with superhydrophobicity. Nano Energy, 76: 105074. https://doi.org/10.1016/j.nanoen.2020.105074

[14] Ahn, J., Zhao, Z. J., Choi, J., Jeong, Y., Hwang, S., Ko, J., Gu, J., Jeon, S., Park, J., Kang, M., Del Orbe, D. V., Cho, I., Kang, H., Bok, M., Jeong, J. H., & Park, I. (2021). Morphology-controllable wrinkled hierarchical structure and its application to superhydrophobic triboelectric nanogenerator. *Nano Energy*, *85*(March), 105978. https://doi.org/10.1016/j.nanoen.2021.105978

[15] Shen J, Li Z, Yu J, Ding B. (2017). Humidity-resisting triboelectric nanogenerator for high performance biomechanical energy harvesting. Nano Energy, 40: 282-288. https://doi.org/10.1016/j.nanoen.2017.08.035

[16] Lv, P., Shi, L., Fan, C., Gao, Y., Yang, A., Wang, X., Ding, S., & Rong, M. (2020). Hydrophobic Ionic Liquid Gel-Based Triboelectric Nanogenerator: Next Generation of Ultrastable, Flexible, and Transparent Power Sources for Sustainable Electronics. *ACS Applied Materials and Interfaces*, *12*(13), 15012–15022. https://doi.org/10.1021/acsami.9b19767

[17] Zhou, Q., Lee, K., Kim, K. N., Park, J. G., Pan, J., Bae, J., Baik, J. M., & Kim, T. (2019). High humidity- and contamination-resistant triboelectric nanogenerator with superhydrophobic interface. *Nano Energy*, *57*(December 2018), 903–910. https://doi.org/10.1016/j.nanoen.2018.12.091

[18] Ccorahua R, Cordero A, Luyo C, Quintana M, Vela E. (2019). Starch-Cellulose-Based triboelectric nanogenerator obtained by a low-cost cleanroom-free processing method. MRS Adv, 4(23): 1315-1320. https://doi.org/10.1557/adv.2018.652

[19] Wang N, Zheng Y, Feng Y, Zhou F, Wang D. (2020). Biofilm material based triboelectric nanogenerator with high output performance in 95% humidity environment. Nano Energy, 77(May): 105088. https://doi.org/10.1016/j.nanoen.2020.105088

[20] Wang J, Wu H, Fu S, Li G, Shan C, He W, Hu C. (2022). Enhancement of output charge density of TENG in high humidity by water molecules induced self-polarization effect on dielectric polymers. Nano Energy, 104(Part B): 107916. https://doi.org/10.1016/j.nanoen.2022.107916

[21] Ccorahua R, Huaroto J, Luyo C, Quintana M, Vela EA. (2019). Enhanced-performance bio-triboelectric nanogenerator based on starch polymer electrolyte obtained by a cleanroom-free processing method. Nano Energy, 59(March): 610-618. https://doi.org/10.1016/j.nanoen.2019.03.018

[22] Shi, L., Dong, S., Xu, H., Huang, S., Ye, Q., Liu, S., Wu, T., Chen, J., Zhang, S., Li, S., Wang, X., Jin, H., Kim, J. M., & Luo, J. (2019). Enhanced performance triboelectric nanogenerators based on solid polymer electrolytes with different concentrations of cations. *Nano Energy*, *64*(August), 103960. https://doi.org/10.1016/j.nanoen.2019.103960

[23] Zhu Z, Xia K, Xu Z, Lou H, Zhang H. (2018). Starch Paper-Based Triboelectric Nanogenerator for Human Perspiration Sensing. Nanoscale Res. Lett, 13: 4-9. https://doi.org/10.1186/s11671-018-2786-9

[24] Torres FG, Troncoso OP, Torres C, Díaz DA, Amaya E. (2011). Biodegradability and mechanical properties of starch films from Andean crops. Int. J. Biol. Macromol, 48(4): 603-606. https://doi.org/10.1016/j.ijbiomac.2011.01.026

[25] Ben Doudou B, Vivet A, Chen J, Laachachi A, Falher T, Poilâne C. (2014). Hybrid carbon nanotube - Silica/ polyvinyl alcohol nanocomposites films: Preparation and characterisation," J. Polym. Res, 21(4). https://doi.org/10.1007/s10965-014-0420-9

[26] Kochkina NE, Butikova OA. (2019). Effect of fibrous TiO_2 filler on the structural, mechanical, barrier and optical characteristics of biodegradable maize starch/PVA composite films. Int. J. Biol. Macromol, 139: 431-439. https://doi.org/10.1016/j.ijbiomac.2019.07.213

[27] Rajesh K, Crasta V, Rithin Kumar NB, Shetty G, Rekha PD. (2019). Structural, optical, mechanical and dielectric properties of titanium dioxide doped PVA/PVP nanocomposite. J. Polym. Res, 26(4): 4-10. https://doi.org/10.1007/s10965-019-1762-0

[28] Sarkar S, Guibal E, Quignard F, SenGupta AK (2012). Polymer-supported metals and metal oxide nanoparticles: Synthesis, characterization, and applications. J. Nanoparticle Res, 14(2): 2-24. https://doi.org/10.1007/s11051-011-0715-2

[29] Alvarez-Ramirez J, Vazquez-Arenas J, García-Hernández A, Vernon-Carter EJ. (2019). Improving the mechanical performance of green starch/glycerol/Li + conductive films through cross-linking with Ca^{2+}. Solid State Ionics, 332(November): 1-9. https://doi.org/10.1016/j.ssi.2019.01.002

[30] Abdullah AM, Aziz SB, Saeed SR(2021). Structural and electrical properties of polyvinyl alcohol (PVA):Methyl cellulose (MC) based solid polymer blend electrolytes inserted with sodium iodide (NaI) salt. Arab. J. Chem, 14(11): 103388. https://doi.org/10.1016/j.arabjc.2021.103388

[31] Bergo P, Sobral PJA, Prison JM. (2011). Effect of glycerol on physical properties of cassava starch films. J. Food Process. Preserv, 34(2): 401-410. https://doi.org/10.1111/j.1745-4549.2008.00282.x

[32] Adamu AD, Jikan SS, Talip BHA, Badarulzaman NA, Yahaya S. (2017). Effect of glycerol on the properties of tapioca starch film. Mater. Sci. Forum, 888: 239-243. https://doi.org/10.4028/www.scientific.net/MSF.888.239

[33] Abral H, Hartono A, Hafizulhaq F, Handayani D, Sugiarti E, Pradipta O. (2019). Characterization of PVA/cassava starch biocomposites fabricated with and without sonication using bacterial cellulose fiber loadings. Carbohydr. Polym, 206: 593-601. https://doi.org/10.1016/j.carbpol.2018.11.054

[34] Surudžić, R., Janković, A., Bibić, N., Vukašinović-Sekulić, M., Perić-Grujić, A., Mišković-Stanković, V., Park, S. J., & Rhee, K. Y. (2016). Physico-chemical and mechanical properties and antibacterial activity of silver/poly(vinyl alcohol)/graphene nanocomposites obtained by electrochemical method. *Composites Part B: Engineering*, *85*, 102–112. https://doi.org/10.1016/j.compositesb.2015.09.029

[35] Wang Y, Zhang H, Zeng Y, Hossen Md A, Dai Y, Li S, Liu Y, Qin W. (2022). Development and characterization of potato starch/lactucin/nano-TiO$_2$ food packaging for sustained prevention of mealworms. Food Packag. Shelf Life, 33: 100-837. https://doi.org/10.1016/J.FPSL.2022.100837

[36] Oleyaei SA, Almasi H, Ghanbarzadeh B, Moayedi AA. (2016). Synergistic reinforcing effect of TiO$_2$ and montmorillonite on potato starch nanocomposite films: Thermal, mechanical and barrier properties. Carbohydrate Polymers, 152: 263-262. https://doi.org/10.1016/j.carbpol.2016.07.040

[37] Park, H. W., Huynh, N. D., Kim, W., Hwang, H. J., Hong, H., Choi, K. H., Song, A., Chung, K. B., & Choi, D. (2018). Effects of embedded TiO2-x nanoparticles on triboelectric nanogenerator performance. *Micromachines*, *9*(8). https://doi.org/10.3390/mi9080407

[38] Bunriw, W., Harnchana, V., Chanthad, C., & Huynh, V. N. (2021). Natural rubber-TiO2 nanocomposite film for triboelectric nanogenerator application. *Polymers*, *13*(13), 1–12. https://doi.org/10.3390/polym13132213

[39] Butyral, P., Anti, E., Agent, O., Benas, J., Huang, C., Yan, Z., Liang, F., & Li, P. (2022). Nanofiber-Based Odor-Free Medical Mask Fabrication Using Polyvinyl Butyral and Eucalyptus Anti Odor Agent. Polymers 2022, 14: 4447. https://doi.org/10.3390/polym14204447

[40] Wang, N., Feng, Y., Zheng, Y., Zhang, L., Feng, M., Li, X., Zhou, F., & Wang, D. (2021). New Hydrogen Bonding Enhanced Polyvinyl Alcohol Based Self-Charged Medical Mask with Superior Charge Retention and Moisture Resistance Performances. *Advanced Functional Materials*, *31*(14), 1–14. https://doi.org/10.1002/adfm.202009172

[41] Patil, S., Bharimalla, A. K., Mahapatra, A., Dhakane-Lad, J., Arputharaj, A., Kumar, M., Raja, A. S. M., & Kambli, N. (2021). Effect of polymer blending on mechanical and barrier properties of starch-polyvinyl alcohol based biodegradable composite films. *Food Bioscience*, *44*(PA), 101352. https://doi.org/10.1016/j.fbio.2021.101352

A COMPARATIVE FINITE ELEMENT ANALYSIS OF REGULAR AND TOPOLOGICALLY OPTIMISED DENTAL IMPLANTS FOR MECHANICAL AND FATIGUE RESPONSES EVALUATION

**MUHAMMAD IKMAN ISHAK[1*], RUSLIZAM DAUD[1], BAKRI BAKAR[1]
AND SITI NOOR FAZLIAH MOHD NOOR[2]**

*[1]Faculty of Mechanical Engineering & Technology, Universiti Malaysia Perlis,
Kampus Alam UniMAP, Pauh Putra, 02600 Arau, Perlis, Malaysia*
*[2]Advanced Medical and Dental Institute, Universiti Sains Malaysia, Bertam, Jln. Tun Hamdan
Sheikh Tahir, 13200 Kepala Batas, Pulau Pinang, Malaysia*

[*]Corresponding author: ikman@unimap.edu.my

ABSTRACT: Topology optimisation is a prominent method to improve the performance of any systems by optimising geometrical factors to save materials without compromising the system functionality. Currently, there is limited published data discussing the topologically optimised dental implants that makes the matter still unclear. This study aimed to evaluate the mechanical and fatigue behaviours of regular and topologically optimised dental implant designs using 3-D FEA. Geometrical models were developed in accordance with ISO 14801 using SolidWorks 2020 before being analysed in ANSYS 18.1. The new implant design was created by topology optimisation analysis. The material properties of all parts were assumed to be isotropic, linearly elastic, and homogenous. Nine different compressive load values ranging from 100 to 500 N were applied on the loading structure as separated cases. The vertical and bottom surfaces of the holder were fully constrained. The results showed that the topologically optimised implant recorded about 12.3% lower implant stress than the regular implant. Both implant designs revealed a comparable displacement result with a percentage difference of only 2.3%. The optimised design was also found to produce longer fatigue life and approximately 12.3% higher safety factor compared to the regular design. The increase in the compressive load value has increased the stress and deformation, whilst decreased the fatigue life and safety factor in both designs. Although it was estimated that the volume of the new implant could be reduced to about 24% of the traditional one, the implant functionality may still be retained or even be improved.

ABSTRAK: Pengoptimuman topologi adalah kaedah utama bagi meningkatkan prestasi mana-mana sistem dengan mengoptimumkan faktor geometri bagi menjimatkan bahan tanpa menjejaskan fungsi utama sistem. Dewasa ini, terdapat kurang data diterbitkan berbincang mengenai implan gigi yang dioptimumkan secara topologi yang menjadikan perkara ini masih tidak jelas. Kajian ini bertujuan bagi menilai perlakuan mekanikal dan kelesuan bagi reka bentuk implant gigi biasa dan yang dioptimumkan secara topologi menggunakan 3-D FEA. Model geometri telah dibangunkan mengikut ISO 14801 menggunakan SolidWorks 2020 sebelum dianalisis dalam ANSYS 18.1. Reka bentuk implan baharu telah dibuat melalui analisis pengoptimuman topologi. Sifat pada semua bahagian bahan diandaikan sebagai isotropik, keanjalan linear, dan homogen. Sembilan nilai beban mampatan berbeza antara 100 hingga 500 N telah dikenakan pada struktur pembebanan sebagai kes berasingan. Permukaan menegak dan bawah pemegang dikekang sepenuhnya. Keputusan menunjukkan bahawa implan yang dioptimumkan secara topologi

merekodkan tegasan implan 12.3% lebih rendah daripada implan biasa. Kedua-dua reka bentuk implan menunjukkan hasil anjakan yang setanding dengan perbezaan peratusan hanyalah 2.3%. Reka bentuk yang dioptimumkan juga didapati menghasilkan hayat kelesuan yang lebih lama dan kira-kira 12.3% faktor keselamatan yang lebih tinggi berbanding reka bentuk biasa. Peningkatan dalam nilai beban mampatan telah meningkatkan tegasan dan perubahan bentuk, sementara mengurangkan hayat kelesuan dan faktor keselamatan dalam kedua-dua reka bentuk. Walaupun dianggarkan bahawa isipadu implan baru boleh dikurangkan kira-kira 24% daripada implan tradisional, fungsi implan masih boleh dikekalkan atau dipertingkatkan.

KEYWORDS: *deformation; dental implant; fatigue; finite element analysis; topology optimisation*

1. INTRODUCTION

Endosseous dental implants are a popular treatment choice for partially or completely edentulous patients to restore dental performance and improve dental aesthetic appearance [1]. This artificial surgical part is always in high demand and reported to possess a satisfactory success and survival rate [2]. Titanium is the most common material for dental implants due to its biocompatible properties allowing the implant to remain stable and function successfully. The root form implant is one of the types of the endosseous dental implant that occupies the vertical bone region mimicking the root structure of a natural tooth [3]. Among the typical configurations of the root form implant are cylindrical- and screw-type where both are available in solid or hollow form. The hollow-cylindrical implant is introduced to increase the total contact surfaces between the implant and the bone compared to the solid-cylindrical implant. The solid-cylindrical design is exposed to destructive shear force and dependent on surface treatment for improving retention to the bone. The solid-screw implants, alternatively, possess macroscopic retaining features along the body for primary bone anchorage. Apart from reducing the impact of shear load, the implant retentive elements may also decrease the potential of overloading at the bone-implant interface. The solid-screw implants offer an improved functional contact surface with the adjacent bone and easier surgical placement compared to hollow-cylindrical or hollow-screw designs.

Biomechanical compatibility of a dental implant is influenced by many factors such as the applied occlusal loading, implant geometry, implant dimensions, implant material stiffness, and bone quality and quantity [4]. The connection of the implant with the attached living tissues is indicated by a physiological phenomenon that is described as osseointegration. In the early stage of implantation, it is normally observed that peri-implant bone resorption occurs due to less mechanical stimulation. As the high stress value and concentration can exist in the vicinity of the implant, particularly at the implant neck region, this could also predict the occurrence of bone loss [5]. It has been suggested that in low-quality bone, the implant with large diameter, long, and straight body wall is highly preferable. Besides, the application of a platform switching concept in dental implants, where the diameter of the abutment is much smaller than the implant platform, produces lower tensile and compressive bone stresses than the platform-matched implant [6]. The platform-switched implants could preserve alveolar bone level by shifting the stresses from the area of compact bone to the area of trabecular bone. In terms of the applied load and bone quality, these two factors are challenging to alter for an optimum implant stress response at the bone-implant interface. Many attempts have been made to optimise the design of dental implants including by resembling the shape of a natural tooth. The optimised implant design with a tapered and wider neck exhibited considerably lower peak stresses relative to the traditional implant design [7].

Topology optimisation provides an optimised material allocation in a given design space based on a set of constraints and loadings [1,8,9]. In other words, it is a method for shape optimisation to establish material outline. An ideal design of structure demands an even stress dissipation within the body up to tolerable limit. There is a possibility to remove material at the minimally stressed regions such that the final design could achieve similar function with a reduced mass. For instance, a topologically optimised design of spinal cage managed to reduce the volume of existing design about 36% and increase the space for bone augmentation while maintaining afforded spinal stability [10]. Regarding dental implants, apparent high stress concentration is commonly observed at the regions interfacing with the compact bone with minimal stress accumulation at the implant apex. The mathematical algorithm in topology optimisation technique is in reference to objective function, design variables, and constraints [11]. Through finite element analysis (FEA), the optimisation method iteratively analyses the design performance to remove the redundant materials without affecting the structural stability and function. The FEA has been proven as a useful and widely accepted method to solve complex mathematical problems regarding stability and failure analyses in many fields such as structure, biomedicine, electronics, fluid, and heat. The use of FEA in implant dentistry began in 1973, and has since been increasingly used to predict responses that are challenging to obtain in experimental and clinical works. To date, limited published data could be found on the topologically optimised solid dental implant that makes this matter still unclear and inconclusive. In a study by Chang et al. [8], it was reported that the volume of the original implant design could be decreased about 17.9%. The study analysed the implant embedded in the maxillary first molar region and applied with a static occlusal load. A more recent investigation by Gupta et al. [1] demonstrated a volume reduction of 32 – 45% could be achieved with the implant still able to retain its function. The variation in the length and diameter of the implant and bone quality were studied under the application of constant static load. Dental implants are highly exposed to fatigue failure due to repetitive occlusal loading. Many existing computational studies have only evaluated the performance of traditional implants and analysed the optimum topology of an implant under static loading, without emphasising on fatigue characteristics. Critical grasp of fatigue conditions is important to comprehend the force transfer within a dental implant system.

The main objective of the present study was to examine and compare the mechanical and fatigue responses between a regular and topologically optimised dental implant designs under different vertical compressive load values using three-dimensional (3-D) FEA. Four types of result criteria, namely maximum equivalent von Mises stress, maximum total deformation, minimum fatigue life, and minimum safety factor were interpreted from the models. The null hypothesis of this study was that the traditional and new implant designs demonstrate insignificant difference in the response data investigated. The novelty of our study was to shed light on the determination of redundant material distribution in a threaded dental implant and generation of clearer quantitative mechanical and fatigue response data for regular and topologically optimised implant designs. This may give valuable insights for clinicians and/or implant manufacturers in the development of new implant macrogeometry design with reduced mass and acceptable strength. It is also expected that this study could provide an improved comprehension of the force transfer in different implant designs under the evaluation of fatigue conditions.

2. MATERIALS AND METHODS

2.1 Geometrical Model

The geometry of a commercial dental implant system (dual-fit (DFI)) from Alpha-Bio Tec, Petach Tikva, was used as a reference to build a 3-D model of a solid threaded dental implant. The length and diameter of the implant body are 11.5 and 3.75 mm, respectively. Besides, a 3.5-mm high straight abutment, and a 2.2-mm wide and 8.0-mm long abutment screw were also modelled. The abutment screw is used to rigidly attach the abutment to the implant body at the implant platform. In this study, the implant platform was designed with an internal hexagonal connection as stated in the implant manufacturer's catalogue. Apart from the implant components, the 3-D geometrical modelling also includes a holder (40.9 mm (L) x 39.8 mm (H) x 23.0 mm (W)), a hemispherical cap (6.5 mm (L) x 5.1 mm (W)), and a loading structure (12.0 mm (L) x 11.0 mm (W)). As the study covers the fatigue prediction, the 3-D model setup was prepared based on the fatigue testing standard ISO 14801 using a computer-aided design (CAD) software, SolidWorks 2020 (SolidWorks Corp., Concord, Massachusetts, USA). Related in-built geometrical shape creation tools provided in the software such as extrude, mirror, revolve, sweep, and/or loft were employed to develop the models.

All the models were assembled and placed under the testing condition described in ISO 14801. Figure 1(a) exhibits the exploded view of the final geometrical analysis model. The implant body was virtually embedded into a 4.5-mm diameter cylindrical hole created on the holder. Boolean subtraction option was adopted to place the implant body into the holder. A bone loss of 3.0 mm measured from the implant platform to the holder surface (Fig. 1(b)) was simulated, and the abutment and its screw were attached to the implant body. The hemispherical cap was then placed onto the abutment for loading. The central axis of the cap, abutment, abutment screw, and implant body were ensured to be aligned. The central point of the cap was set to be at 11.0 mm from the bone (holder) level to represent a moment arm. The assembled model was then transferred into ANSYS 18.1 software (ANSYS Inc., Houston, TX, USA) to generate the model mesh and pre-processing settings of the FEA.
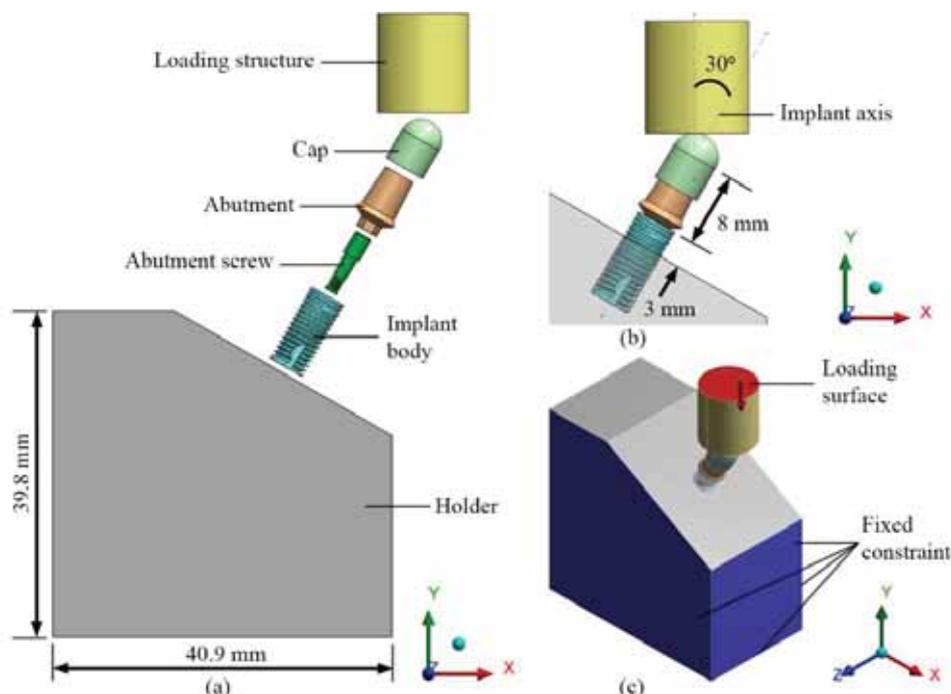

Fig. 1: (a) Exploded view of the model. (b) Dimensions for the implant placement. (c) Loading and support locations of the model.

2.2 Material Properties

Each component in the model was considered to be isotropic, homogenous, and linearly elastic. The abutment, abutment screw, and implant body were made of titanium alloy (Ti-6Al-4V) with a yield strength of 847 MPa [12]. Additionally, the hemispherical cap and loading structure were assigned with steel alloy. According to ISO 14801 specifications, the implant body must be fixed in a rigid clamping holder having the elastic modulus of at least 3.0 GPa. Therefore, in this study, we have chosen aluminium alloy as the material of the holder. A summary of the material properties for the model setup is shown in Table 1.

Table 1: Material properties for the model setup

Material	Part	Young's Modulus, E (GPa)	Poisson's Ratio, v	References
Ti-6Al-4V	Abutment, abutment screw & implant body	113.8	0.342	Yalçın et al. [13]
Aluminium alloy	Holder	71	0.33	Bayata et al. [14]
Steel alloy	Cap & loading structure	200	0.31	Yao et al. [15]

2.3 Interface Conditions

All the interfaces were modelled as perfectly bonded using contact and target elements in ANSYS software. This type of contact modelling was also implemented in many previous related studies [16]. The bonded contact type indicates no penetration or loosening among the interfaces of the implant body-holder, abutment screw-implant body, abutment-abutment screw, abutment-implant body, cap-abutment, and cap-loading structure.

2.4 Loading and Boundary Condition

As recommended by ISO 14801, to simulate the experimental test condition, a vertical load was applied on the flat surface of the loading structure model. This load configuration basically represents an oblique load at an inclination of 30° with respect to the implant longitudinal axis [17]. The applied compressive load was varied from 100 to 500 N with an increment of 50 N that signify the range of normal bite forces [18]. The location of the loads was at a distance of 11.0 mm from the holder inclined surface.

Fatigue analysis was performed on the implant complex to predict the mechanical responses and corresponding fatigue behaviours of the titanium alloy components. The analysis was executed by a repetitive simulated masticatory loading sets (as stated earlier) with alternating load value using fatigue algorithm based on Goodman fatigue theory in the elasticity mode using ANSYS software. The fatigue lives and potential failure regions were predicted for infinite fatigue life criteria.

For the model boundary condition, all the nodes on the vertical and bottom surfaces of the holder were constrained in all degrees of freedom. This denotes that the nodal displacement is set to zero at the stated surfaces. Figure 1(c) depicts the loading and boundary condition used in the analysis.

2.5 Topology Optimisation

In clinical practices, most dental implant designs used have a circular cross-sectional area. This kind of design possesses low resistance towards shear or torsional forces especially when the abutment screw is tightened or even when the implant is free-standing. The inclusion of antirotational features such as vent or hole in the apical part of the implant body is expected to increase inner space for bone ingrowth while not losing too much

stiffness. Topology optimisation tool in ANSYS software was employed to construct a new dental implant design that targeted to increase internal contact surfaces of the implant body. It is important to note that the strength of the implant body must be retained with the decreasing volume.

The main goal of topology optimisation is to reduce the structure compliance energy. Reducing the compliance is similar to increasing the global stiffness of the structure. In other words, the standard formulation of the method is minimising the structure compliance and at the same time satisfying a limitation on the structure volume. The topology optimisation problem is described as follows (Eq. (1) to Eq. (6)):

$$\text{objective function} = \text{minimise} \left(U_c \right) \tag{1}$$

$$\text{limitation} = 0 < \eta_i < 1 \; \left(i = 1,2,3,\ldots,n \right) \tag{2}$$

$$V \leq V_0 - V* \tag{3}$$

$$V = \sum \eta_i V_i \tag{4}$$

$$E_i = E\left(\eta_i \right) \tag{5}$$

$$\{\sigma_i\} = [E_i]\{\varepsilon_i\} \tag{6}$$

where U_c = energy of structure compliance; η_i = internal pseudodensities that are assigned to each finite element (i) in the topology problem; V = computed volume; V_0 = original volume; $V*$ = amount of material to be removed; V_i = volume of element i; E_i = elasticity tensor for each element; E is the elasticity tensor; i = stress vector of element i; ε_i = strain vector of element i.

The variables of density η range from 0 to 1 in which a value close to 1 indicates that the material should be kept, while a value close to 0 indicates material should be removed. In the optimisation analysis, the volume of the entire regular implant body (design domain) was set to be decreased by about 50% with 10 iterations as the pre-set response constraints. The convergence accuracy was defined as 0.1%. By applying 100-N vertical static force normal to the implant platform and fixedly constrained the outer surfaces along the implant body, we managed to obtain a reasonable distribution of material for a new implant design as illustrated in Fig. 2. It was observed that only the internal portion of the apical implant body volume was removed. A comparison of mechanical and fatigue behaviours between regular and topologically optimised implants were made and evaluated.
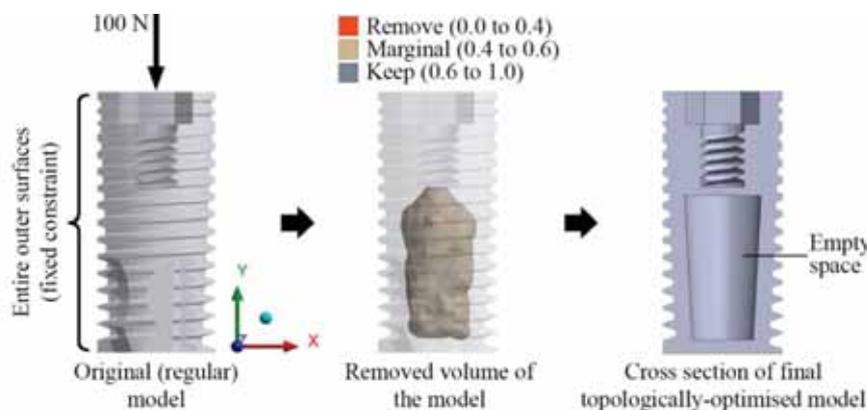

Fig. 2: Main stages of topology optimisation analysis.

2.6 Model Meshing

Since the geometrical shapes of the implant parts are irregular, complex, and small in dimensions, ten-node quadratic tetrahedral elements (SOLID187) were applied in this study to assure the continuity of the force and displacement transfers on the nodes, as considered in some earlier studies [19,20]. Prior to final model discretisation, it is worth mentioning that a mesh convergence analysis was performed so that the analysis results are independent of mesh condition. A smaller mesh size provides a better estimate for the exact solution. However, a high number of nodes and elements may increase computing time. Thus, an agreement must be achieved between the ideal mesh size and the demand for a reliable solution. In this mesh independence test, the details of material properties, contact modelling, and boundary conditions are as mentioned in the previous sections with the 100-N applied load. The maximum equivalent von Mises stress within the abutment-implant complex was recorded for variety mesh sizes. A total of nine relative characteristic mesh sizes was considered which are 2.2 mm (Tet A: ~34,000 elements), 2.0 mm (Tet B: ~43,000 elements), 1.8 mm (Tet C: ~67,000 elements), 1.4 (Tet D: ~114,000 elements), 1.2 mm (Tet E: ~117,000 elements), 1.0 mm (Tet F: ~300,000 elements), 0.95 mm (Tet G: ~349,000 elements), 0.8 mm (Tet H: ~581,000 elements), and 0.7 mm (Tet I: ~865,000 elements). Upon refining the mesh, the acceptability of the result is determined based on the variation of maximum stress value among the mesh sizes, which should be less than 5% [21, 22]. In general, there was a significant difference found on the stress values generated among the mesh sizes. After four refinements, the result seemed to yield at the mesh size of 1.2 mm (Tet-E) with the relative deviation of 1.2%. The total number of nodes and elements are approximately 250,000 and 176,500, respectively. The illustration of mesh sensitivity plot and mesh distribution in the coarse (Tet-A) and refined (Tet-E) models are shown in Fig. 3.
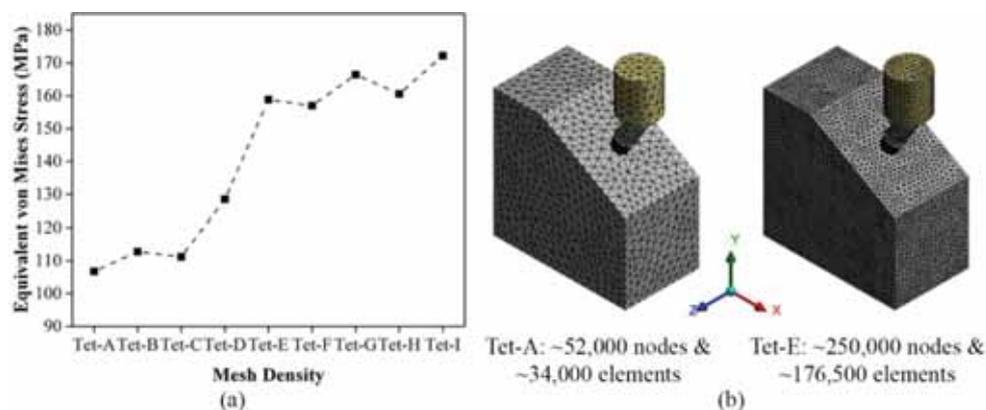
Fig. 3: (a) Mesh sensitivity plot for different element numbers. (b) Mesh distribution in the coarse (Tet-A) and refined (Tet-E) models.

2.7 Model Verification

A comparison was made between our proposed model with those published in earlier studies that considered similar analysis type, modelling design, and analysis software. The pre-processing settings used in the selected previous studies were replicated as closely as possible. It was found that our implant stress result was comparable with that of previous studies as shown in Table 2. A little difference was recorded that could possibly be due to different detailed geometrical features of the models and several assumptions made.

Table 2: Comparison of implant equivalent von Mises stress results between literature and our models

Past Studies	Literature Results	Proposed Model Results
Wang et al. [16]	368.4 MPa	326.19 MPa
Cheng et al. [23]	625.21 MPa	462.02 MPa
Bayata et al. [14]	368.68 MPa	392.02 MPa

3. RESULTS

3.1 Topology Optimisation Analysis

The new implant design was successfully developed from the topology optimisation analysis. As depicted in Fig. 2, only the apical third part of the implant body was affected, and the corresponding volume was removed. The reduction of volume was about 24% considering the change from 123.13 mm^3 (traditional implant) to 93.63 mm^3 (new implant). Meanwhile, the finite element model of the new complex implant was reconstructed and consisted of 249,500 nodes and 176,000 elements.

3.2 Mechanical Behaviour of the Regular and Topologically Optimised Implant Designs

The magnitude of the maximum equivalent von Mises stress and maximum total deformation of the implant-abutment complex for all loading values was extracted in the post analyses to predict the risk of implant failure. The colour contour outline of the results was also provided to clearly scrutinise the mechanical stimuli distribution.

3.2.1 Maximum Equivalent von Mises Stress Results

In comparison, the regular or standard implant design exhibited a greater maximum equivalent von Mises stress value within the implant-abutment assembly than the topologically optimised one, irrespective of loading levels (Fig. 4(a)). The increase in the load value led to the linear increase in the output stress for both designs with the peak values of 794.23 MPa and 707.12 MPa recorded in the regular and optimised designs, respectively, under 500-N load. Meanwhile, the lowest stress was generated by 100-N load with the value of 158.85 MPa (regular design) and 141.42 MPa (optimised design). It became evident that the regular design recorded about 12.3% higher implant stress as compared to the new design in all loading values.
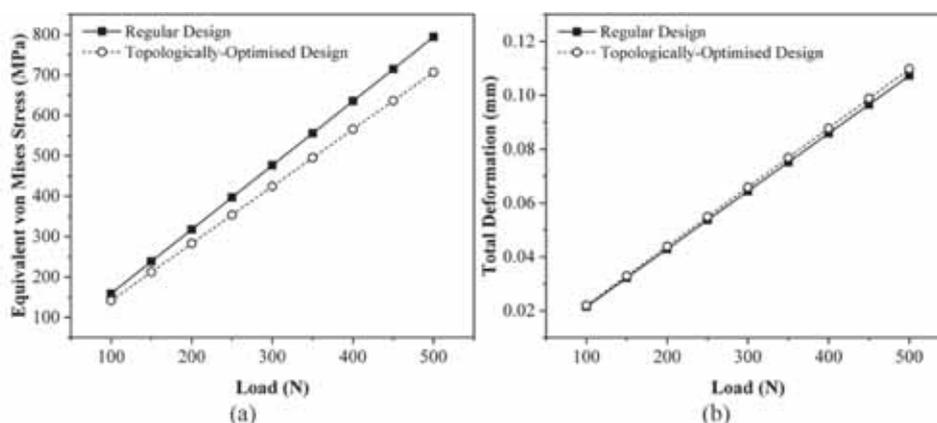
Fig. 4: (a) Maximum equivalent von Mises stress and (b) total deformation of the implant-abutment assembly for both designs under all loading conditions.

Figure 5 depicts the maximum stress concentration was found in the abutment specifically at the interface region of the implant platform. Besides, a high stress amplification region was also observed around the upper threads of the implant body and near the junction of the holder regardless of implant design. The distribution of stress within the assembly was comparable between the regular and new designs for all loading values.
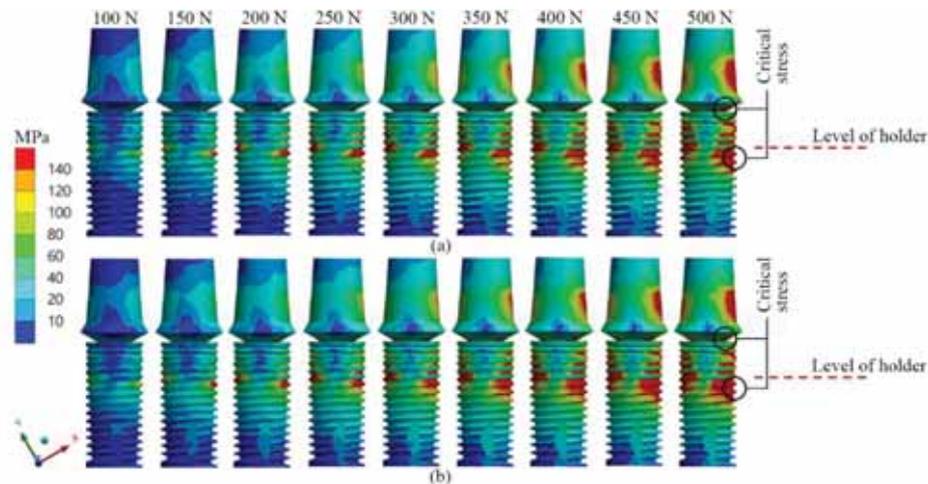
Fig. 5: Dissemination of equivalent von Mises stress in the implant-abutment complex for (a) regular and (b) topologically optimised designs.

3.2.2 Maximum Total Deformation Results

The displacement of the implant-abutment assembly for the regular form was relatively lower than that for the topologically optimised design in all loading values (Fig. 4(b)). Similar to the stress results, it was shown that the total deformation was proportionally increased with the increase in the loading value for both implant configurations. The load of 500 N resulted in the maximum deformation of the implant-abutment assembly with the value of 0.1073 and 0.1098 mm for the regular and new implant designs, respectively. The minimum deformations were 0.0215 mm (regular design) and 0.022 mm (optimised design) generated under 100-N load. The displacement of the new model was only increased by about 2.3% compared to that of the regular model regardless of loading values.

Fig. 6: Distribution of total deformation of the implant-abutment complex for (a) regular and (b) topologically optimised designs.

3.3 Fatigue Behaviour of the Regular and Topologically Optimised Implant Designs

The minimum value of fatigue lives and safety factor of both implant designs under different loading magnitudes were presented. The outcomes were also interpreted in the form of spectrum colouring scale for a clearer understanding.

3.3.1 Minimum Fatigue Life Results

Figure 7(a) shows a comparison of minimum fatigue life (cycles) between the regular and topologically optimised designs. Our results demonstrated that fatigue failure was only predicted for the applied loading values ranging from $350 – 500$ N by considering the fatigue life limit of 5×10^6 cycles based on existing experimental fatigue testing. These loads were equivalent to the resulting stresses of $555.96 – 794.23$ MPa and $494.99 – 707.12$ MPa for the regular and new implant designs, respectively. It was observed that the regular design showed a shorter estimated life with the lowest cycle number of 0 as compared to the optimised design that yielded at 2.69×10^5 cycles under 500-N load. There was no difference in the number of cycles (1×10^7 cycles) at the lower loading values ($100 – 300$ N) among both designs.

Fig. 7: (a) Minimum fatigue life and (b) safety factor of the implant-abutment assembly for both implant designs under all loading conditions.

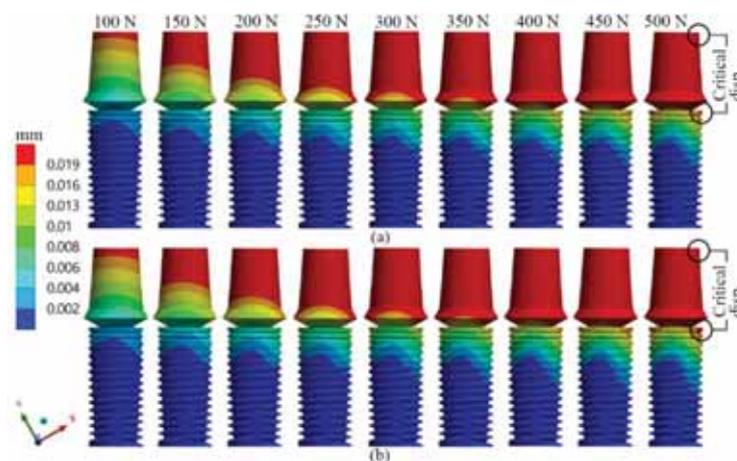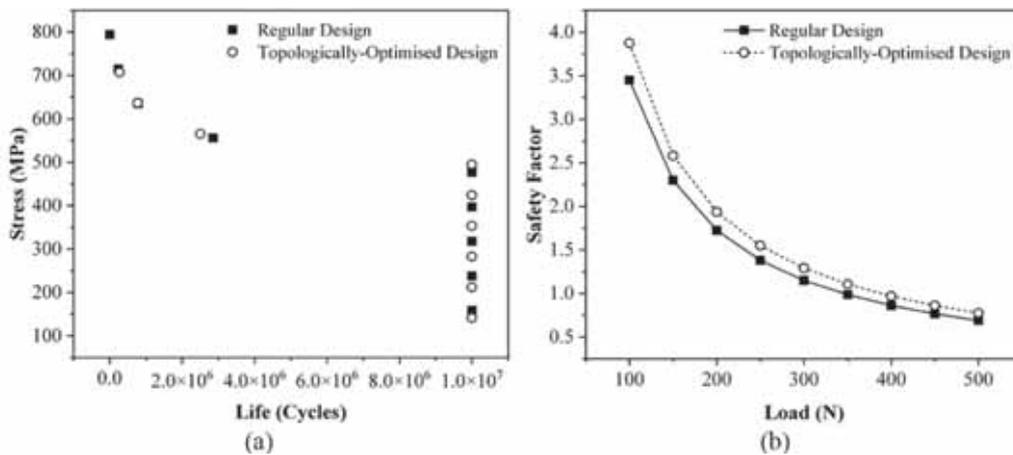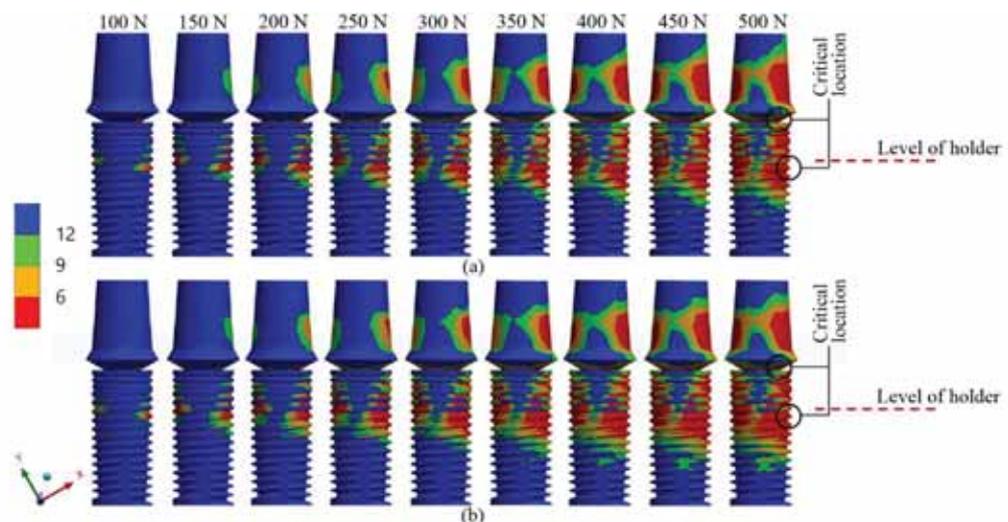

Fig. 8: Distribution of safety factor of the implant-abutment complex for (a) regular and (b) topologically optimised designs under all loading conditions.

3.3.2 Minimum Safety Factor Results

The safety factor values for fatigue life were computed based on Goodman fatigue theory in accordance with infinite fatigue life criteria. The plot of minimum safety factor in the implant-abutment complex versus the applied loads is depicted in Fig. 7(b). The safety factors for the new design were greater than those for the regular design irrespective of loading values. The maximum safety factors were 3.45 and 3.87 for the regular and optimised designs, respectively, under the load of 100 N. It is noted that the safety factors were below 1 when the models were subjected to the load from 350 N for the regular design, and from 400 N for the optimised design. This finding was consistent with the fatigue life results where the load lower than 350 N was predicted to be safe against fatigue. It was observed that critical regions that could attribute implant to failure were the abutment-implant connection and the middle threads of the implant body as illustrated in Fig. 8. Safe regions were more dominant at the apical portion of implant body and top part of the abutment.

4. DISCUSSION

Dental implants are used to transmit the occlusal force to the neighbouring bones. Therefore, the main functional design objective can be associated with the distribution of external loads by improving the function of implant-supported prosthesis. The integration of fundamental scientific knowledge related to geometrical features and force with engineering solutions may accomplish the targeted clinical goals. Presently, there is a high number of studies investigating the perseverance of bone-implant attachment in order to obtain optimal mechanical stimuli transfer. However, low emphasis is placed, and limited information available, on the structural optimisation design of dental implants, particularly for fatigue prediction. Our topology optimisation analysis showed that approximately 24% of the whole implant body volume was redundant material. Only the internal volume of the apical part was removed while the other regions were kept intact. The suggested regions to be removed also covered the vicinity underneath the threaded hole for the abutment screw attachment. However, this was not completely removed so as to preserve the appropriate volume in sustaining the screw integrity and to minimise the potential of high stress concentration in that region. One study performed by Chang et al. [8] reported that the decreased volume of the traditional implant as a result of topology optimisation was lower compared with the one of our findings where the reduction percentage was only about 17.9%. In contrast, the finding from Gupta et al. [1] predicted a far greater volume reduction of 32 – 45% for a topology study that also considered the effect of the implant macrogeometry and bone quality. The differences could possibly be due to the different geometrical shape of the models and pre-processing settings of the analysis employed. Nevertheless, the consistency was found for the location of the removed material which is the apical part of the implant body. To examine the mechanical and fatigue behaviours of the regular and topologically optimised implant designs, linear static FEA was performed under a variation of occlusal loading values to analyse the stress and displacement of the implant-abutment complex.

The level of stress in the newly modified implant was about 12.3% lower than the one in the traditional implant. However, the pattern of stress dissemination over the regions in the assembly seemed to be identical in both designs. Our findings disagreed with the results of Chang et al. [8] where the optimised design (128.1 MPa) generated 13% higher implant stress value compared to the original one (113.3 MPa) under the 204-N oblique load. In contrast, the findings of Gupta et al. [1] were consistent with ours by which a lower implant

stress was recorded by the modified design (33.21 MPa) than the traditional design (54.26 MPa) at the imposed volume reduction of 50%. A possible explanation of our results is that the reduction of implant material tends to minimise the stress shielding effect. Therefore, the mechanical stress transfer from the high stiff implant to the surrounding region having lower stiffness is increased. As a result, the implant-abutment embodiment for the optimised design sustained less stress. In terms of the region of stress distribution, apparent stress amplification occurred at the connecting part of the abutment, and at the implant neck near the junction of the holder. This is parallel with the stress colour contour plots in past studies [1, 8]. One can conclude that the removal of volume from the end of the implant body up to its mid region can be made. A study by Shi et al. reported that alternative implant design with a larger and tapered crestal part radius was preferred compared to the commercially available implant due to a lower peak stress generated [7]. Besides, the topologically optimised implant design could increase the area for osseointegration ingrowth by providing more spaces in the apical region [24]. To relate our results with more realistic studies that concern biological situations, it is commonly found that high stress intensity is located at the interface with the compact bone or peri-implant bone region. A bone stress higher than the strength of the cortical bone, 170 MPa, means a prediction of failure. In implant dentistry, peri-implant bone loss is one of the main manifestations regarding osseo-insufficiency which could lead to detrimental implications such as patient complaints, aesthetic compromise, soft and hard tissue deformation, and implant removal [25]. A greater decrease in the bone level was noticed in the first year of implantation of about 1.0 mm; however, it was reduced to 0.2 mm on average in the following years. The loosening and fracture of the implant due to the resorption of supporting bone could cause eventual implant failure. As such, the effect of bone loss should be minimised by achieving optimum mechanical stress transfer at the bone-implant interface. In this study, titanium alloy Ti-6Al-4V is the main material used for the implant components in both implant designs. We found that the peak implant stress regardless of implant designs under all loading values (141. 42 – 794.23 MPa) was lower than the yield strength of Ti-6Al-4V (847 MPa) [12]. This would predict the success of the implants or low tendency of failure.

In implant dentistry, rigid fixation of an implant is desired as it is the criterion of successful treatment. Rigid fixation is defined as the attainment of implant stability without clinical mobility or the displacement evaluated with horizontal or vertical forces, similar to analysing natural teeth. A healthy natural tooth moves about 56 – 73 μm [26], while a firm implant displaced less than 75 μm [27] with no clinical movement. The implant is exposed to a higher risk of failure than a natural tooth if its motion is higher than 0.5 mm horizontally. Our results revealed that a higher displacement occurred in the abutment rather than in the implant body in both designs which could be due to less resistance towards direct loading imposed from the cap. The optimised design showed a slightly greater abutment deformation relative to the regular one that possibly caused by the reduced retention on the apical portion of the implant body. The loss of volume in that region attributed the implant body to instability, thus increasing the tendency of dislocation. Concerning the displacement of the implant body itself, it was evident that the peak value of the optimised and regular designs was 22.2 and 20.7 μm, respectively, under the 500-N applied load. Compared to the motion of a healthy implant, the values were relatively 70.4% (optimised) and 72.4% (regular) lower than 75 μm. The optimised design still offered acceptable rigid fixation to maintain implant stability although it recorded about 2.3% greater deformation than the regular design.

For the fatigue behaviour assessment, one cyclic loading consisted of one 30° oblique loading [17]. The traditional and topologically optimised models survived up to 1×10^7

cycles under different loading limits which are 300 N and 350 N, respectively. Further to that, under the highest loading value of 500 N, a total failure was found for the regular implant, while the new implant yielded at 2.69 x 10^5 cycles. According to ISO 14801, a dental implant can sustain a life of at least 5 x 10^6 without demonstrating any damage [14]. Our results therefore corresponded well with the standard. The optimised implant design was expected to have higher resistance to fatigue failure in their life as compared to the regular design.

The safety factor of the implant-abutment assembly was above 1 for the loading magnitudes up to 300 N and 400 N in the regular and optimised models, respectively. It is determined by the relation between mean stress, σ_m and alternating stress, σ_a according to the modified Goodman theory as,

$$\sigma_m = \frac{\sigma_{max} + \sigma_{min}}{2} \tag{7}$$

$$\sigma_a = \frac{\sigma_{max} - \sigma_{min}}{2} \tag{8}$$

$$\frac{\sigma_a}{S_e} + \frac{\sigma_m}{S_u} = \frac{1}{N_f} \tag{9}$$

where S_e = endurance limit, S_u = ultimate tensile strength of the material, and N_f = the safety factor for fatigue life in a loading cycle. The safety factor values were decreased in the regular implant compared to the optimised implant due to the increased stress level in the regular implant. The most critical value of 0.69 was recorded in the regular implant under the applied load of 500 N. Thus, considering the safety factor in the mechanical aspect, the topologically optimised implant is more favourable for dental implant application.

Albeit that robust outcomes of the study were found, the quantitative result data were still predictive and could be relying on several limitations of the analyses. In fact, the implant was placed in the non-living material under static loading, while the real applied one is embedded in the complex living tissues without a definite pattern, and this argument may somewhat affect the results. Besides, this study only analysed the single restoration type, thus the findings were attributed to this kind of treatment. Several important aspects can be considered in future topology optimisation studies such as assessing different material stiffnesses and dimensions of the implant body, employing more complex geometrical models, considering more realistic dynamic occlusal loading, and simulating implant extraction from the bone socket. Although the outcomes of the present study could not directly be inferred to actual clinical situation, they managed to reveal a difference in the mechanical and fatigue behaviours through simulation analysis. *In vitro* and *in vivo* studies are needed in order to validate the prognosis of implant perseverance even at low levels. The null hypothesis of the study was rejected as the traditional and topologically optimised implant designs demonstrated a significant difference in the responses obtained.

5. CONCLUSION

The new implant designed by topology optimisation analysis has reduced the volume of the traditional implant by about 24%. However, this topologically optimised design was still able to maintain the implant fixation in sustaining the loadings.

- It was shown that the newly modified implant generated nearly 12.3% lower implant stress than the regular one.

- In terms of displacement, insignificant difference was found among both designs where the percentage deviation was merely 2.3%.

- The optimised implant design also promoted a longer fatigue life and an improved safety factor by approximately 12.3%.

- The manufacturing cost of the new implant might be high using conventional machining processes. However, it could alternatively be manufactured using additive manufacturing methods, which is expected to offer much lower cost.

ACKNOWLEDGMENT

REFERENCES

[1] Gupta Y, Iyer R, Dommeti VK, Nutu E, Rana M, Merdji A, Biswas JK, Roy S. (2020) Design of dental implant using design of experiment and topology optimization: a finite element analysis study. Proc Inst Mech Eng H: J Eng Med., 235(2): 157-166. https://doi.org/10.1177/0954411920967146

[2] Agustin-Panadero R, Leon-Martinez R, Labaig-Rueda C, Faus-Lopez J, Sola-Ruiz MF. (2019) Influence of implant-prosthetic connection on peri-implant bone loss: a prospective clinical trial with 2-year follow-up. Int J Oral Maxillofac Implants., 34(4): 963-968. https://doi.org/10.11607/jomi.7168

[3] Samieirad S, Mianbandi V, Shiezadeh F, Hosseini-Abrishami M, Tohidi E. (2019) Tapered versus cylindrical implant: which shape inflicts less pain after dental implant surgery? A clinical trial. J Oral Maxillofac Surg., 77(7): 1381-1388. https://doi.org/10.1016/j.joms.2019.02.027

[4] Bidez MW, Misch CE. (2015) Chapter 5 - Clinical biomechanics in implant dentistry. In Dental Implant Prosthetics. 2nd Edition. St. Louis, Mosby; pp 95-106.

[5] Gil J, Sandino C, Cerrolaza M, Pérez R, Herrero-Climent M, Rios-Carrasco B, Rios-Santos JV, Brizuela A. (2022) Influence of bone-level dental implants placement and of cortical thickness on osseointegration: In silico and in vivo analyses. J Clin Med., 11(4): 1027 https://doi.org/10.3390/jcm11041027

[6] Chang CL, Chen CS, Hsu ML. (2010) Biomechanical effect of platform switching in implant dentistry: a three-dimensional finite element analysis. Int J Oral Maxillofac Implants., 25(2): 295-304.

[7] Shi L, Li H, Fok AS, Ucer C, Devlin H, Horner K. (2007) Shape optimization of dental implants. Int J Oral Maxillofac Implants., 22(6): 911-20.

[8] Chang C-L, Chen C-S, Huang C-H, Hsu M-L. (2012) Finite element analysis of the dental implant using a topology optimization method. Med Eng Phys., 34(7): 999-1008. https://doi.org/10.1016/j.medengphy.2012.06.004

[9] Seitz KF, Grabe J, Köhne T. (2018) A three-dimensional topology optimization model for tooth-root morphology. Comput Methods Biomech Biomed Engin., 21(2): 177-185. https://doi.org/10.1080/10255842.2018.1431778

[10] Zhong Z-C, Wei S-H, Wang J-P, Feng C-K, Chen C-S, Yu C-H. (2006) Finite element analysis of the lumbar spine with a new cage using a topology optimization method. Med Eng Phys., 28(1): 90-98. https://doi.org/10.1016/j.medengphy.2005.03.007

[11] Zhang K, Cheng G, Xu L. (2019) Topology optimization considering overhang constraint in additive manufacturing. Comput Struct., 212: 86-100. https://doi.org/10.1016/j.compstruc.2018.10.011

[12] Lee H, Jo M, Noh G. (2021) Biomechanical effects of dental implant diameter, connection type, and bone density on microgap formation and fatigue failure: A finite element analysis. Comput Methods Programs Biomed., 200: 105863. https://doi.org/10.1016/j.cmpb.2020.105863

[13] Yalçın M, Kaya B, Laçin N, Arı E. (2019) Three-dimensional finite element analysis of the effect of endosteal implants with different macro designs on stress distribution in different bone qualities. Int J Oral Maxillofac Implants., 34(3): e43–e50. https://doi.org/10.11607/jomi.7058

[14] Bayata F, Yildiz C. (2020) The effects of design parameters on mechanical failure of Ti-6Al-4V implants using finite element analysis. Eng Fail Anal., 110: 104445. https://doi.org/10.1016/j.engfailanal.2020.104445

[15] Yao K-T, Chen C-S, Cheng C-K, Fang H-W, Huang C-H, Kao H-C, Hsu M-L. (2018) Optimization of the conical angle design in conical implant–abutment connections: a pilot study based on the finite element method. J Oral Implantol., 44(1): 26-35. https://doi.org/10.1563/aaid-joi-D-17-00149

[16] Wang K, Geng J, Jones D, Xu W. (2016) Comparison of the fracture resistance of dental implants with different abutment taper angles. Mater Sci Eng C., 63, 164-171. https://doi.org/10.1016/j.msec.2016.02.015

[17] Zhang X, Mao J, Zhou Y, Ji F, Chen X. (2020) Study on statics and fatigue analysis of dental implants in the descending process of alveolar bone level. Proc Inst Mech Eng H., 0954411920926080. https://doi.org/10.1177/0954411920926080

[18] Levartovsky S, Peleg G, Matalon S, Tsesis I, Rosen E. (2022) Maximal bite force measured via digital bite force transducer in subjects with or without dental implants - A pilot study. Appl Sci., 12(3): https://doi.org/10.3390/app12031544

[19] Amruta G, Rajeev S, Gaurang M, Simran M, Mishal Ds, Kunal M. (2022) Comparative evaluation of stress on the peri-implant bone of different densities in mandible restored by various diameter of implants - a finite element analysis. J Electron Sci., 12(11): 161-177.

[20] Satheesh Kumar P, Satheesh KK, John J, Patil G, Patel R. (2013) Force transfer and stress distribution in an implant-supported overdenture retained with a hader bar attachment: a finite element analysis. ISRN Dent., 2013, 369147. https://doi.org/10.1155/2013/369147

[21] Niroomand MR, Arabbeiki M. (2020) Implant stability in different implantation stages: analysis of various interface conditions. Inform Med Unlocked, 19: 100317. https://doi.org/10.1016/j.imu.2020.100317

[22] Jeng M-D, Lin Y-S, Lin C-L. (2020) Biomechanical evaluation of the effects of implant neck wall thickness and abutment screw size: a 3D nonlinear finite element analysis. Appl Sci., 10(10), 3471: https://doi.org/10.3390/app10103471

[23] Cheng Y-C, Jiang C-P, Lin D-H. (2019) Finite element based optimization design for a one-piece zirconia ceramic dental implant under dynamic loading and fatigue life validation. Struct Multidiscipl Optim., 59(3): 835-849. https://doi.org/10.1007/s00158-018-2104-2

[24] Lawrence AA, Singh N, Davis R, Sahil Ansari M, Tewari YV, Francis V. (2022) Chapter 13 - Topological design optimisation of dental implant. In Optimization of Industrial Systems. 1st Edition. Edited by Panchal D, Tyagi M, Sachdeva A, Pamucar D. Scrivener Publishing LLC.

[25] Naveau A, Shinmyouzu K, Moore C, Avivi-Arber L, Jokerst J, Koka S. (2019) Etiology and measurement of peri-implant crestal bone loss (CBL). J Clin Med., 8(2): 166. https://doi.org/10.3390/jcm8020166

[26] Rudd KD, O'Leary TJ, Stumpf AJ Jr. (1964) Horizontal tooth mobility in carefully screened subjects. Tech Doc Rep US Air Force Syst Command Electron Syst Div, 94: 1-6.

[27] Sekine H, Komiyama Y, Potta H, Yoshida K. (1986) Mobility characteristics and tactile sensitivity of osseointegrated fixture-supporting systems. In Tissue Integration in Oral Maxillofacial Reconstruction. 1st Edition. Edited by Van Steeberghe D. Amsterdam, Excerpta Medica; pp. 326–332.

HIGH ACCURACY HUMAN MOTION TRAJECTORY GENERATION FOR EXOSKELETON ROBOT USING CURVE FITTING TECHNIQUE

Muhammad Abdul Jalil[1*], Muhammad Fahmi Miskon[2], Bazli Bin Bahar[2]

[1]Mechanical Engineering Department, Polytechnic Port Dickson, Port Dickson, Malaysia
[2]Center of Excellence in Robotic and Industrial Automation (CeRIA),
Faculty of Electrical Engineering, Universiti Teknikal Malaysia, Melaka, Malaysia

*Corresponding author: muhammadaj@polipd.edu.my

ABSTRACT: Robotic systems often require trajectory planning algorithms that can generate natural human-like movements for tasks such as grasping and manipulation. However, conventional trajectory planning methods may not accurately capture the complex movement patterns observed in humans. In this paper, we present a trajectory planning algorithm based on polynomial curve fitting that aims to address this issue. The algorithm determines the polynomial coefficient values that accurately match the natural human trajectory profile and is evaluated using MATLAB simulations. We compare the proposed algorithm to the conventional quintic polynomial trajectory method, analysing the accuracy, precision, and via-point continuity. The result shows that the algorithm has the ability to generate a trajectory profile with accuracy of 99.8% and a precision of 0.002°. However, the result for via-point continuity shows an error on every sub-phase transition, with the lowest error of 0.0031 between the transition of sub-phases 1 and 2. The result also shows that the lowest fitting error recorded is 0.00014°. The results demonstrate that our algorithm can generate trajectory profiles with higher accuracy and naturalness, potentially improving the performance and usability of robotic systems.

ABSTRAK: Sistem robotik sering memerlukan algoritma perancangan trajektori yang dapat menghasilkan gerakan semulajadi seperti manusia bagi tugas seperti memegang dan memanipulasi objek. Walau bagaimanapun, kaedah perancangan trajektori konvensional mungkin tidak dapat merekodkan pola gerakan kompleks seperti yang dihasilkan manusia secara tepat. Kajian ini adalah berkenaan algoritma perancangan lintasan berdasarkan penyepaduan lengkung polinomial bagi menyelesaikan masalah ini. Algoritma ini menentukan nilai pekali polinomial yang sepadan dengan profil gerakan semulajadi manusia dan dinilai menggunakan simulasi MATLAB. Algoritma yang dicadangkan ini telah dibandingkan dengan kaedah perancangan lintasan polinomial kuintik konvensional, dianalisis kejituan, ketepatan, dan keberterusan titik lalu. Keputusan menunjukkan bahawa algoritma tersebut mampu menghasilkan profil lintasan dengan kejituan sebanyak 99.8% dan ketepatan sebanyak 0.002°. Walau bagaimanapun, dapatan kajian mengenai keberterusan titik lalu menunjukkan ralat pada setiap peralihan fasa-sub dengan ralat terendah sebanyak 0.0031 pada peralihan antara fasa-sub 1 dan fasa-sub 2. Dapatan kajian juga menunjukkan bahawa ralat penyepaduan terendah yang direkodkan adalah sebanyak 0.00014°. Keputusan ini menunjukkan bahawa algoritma ini mampu menghasilkan profil lintasan dengan ketepatan dan sifat semula jadi yang lebih tinggi, berpotensi meningkatkan prestasi dan kegunaan sistem robotik.

1. INTRODUCTION

Trajectory refers to a time history of the position, velocity, and acceleration for each degree of freedom (DOF) [1]. In trajectory generation, the desired trajectory for motions is generated based on the prediction of how the system (robot) responds to the input trajectory. Generally, trajectory generation deals with the problem of 1) how to specify a trajectory with a simple description, 2) how the trajectory is represented, 3) how to generate the trajectory in real-time [1], and 4) how to find a relationship between two domains: time and space [2].

According to Miskon et al. [3], there are three strategies for generating a trajectory for robot applications: off-line, on-line, and combined or hybrid. The off-line strategy is a strategy that uses either a mathematical model such as a polynomial equation [4], Fourier Transform [5], Central Pattern Generation (CPG) [6], Neural Oscillator [7], or uses recorded or normalized human motion data [8]. The advantage of this strategy is that it does not require any dynamic relationship between the robot and the environment. However, the main drawback of this strategy is adaptation due to environmental uncertainties. Also, some of the off-line methods required very accurate robot modelling before the trajectory was planned.

On the other hand, the on-line strategy does not need a predefined trajectory to generate the motion. It has the capability to produce the trajectory according to the working space environment. The implementation of Neural Network [9] and Fuzzy Logic [10] attached with an additional sensor [11,12] is used to improve the accuracy of the generating trajectory. However, this strategy must consider issues such as the fastest time adapting for real-time application and accuracy. Meanwhile, the hybrid strategy used both advantages of off-line and on-line strategy to generate a trajectory for the robot.

The normative human trajectory has been studied to ensure that the robot (in this case, bipedal or exoskeleton) can follow the nature of human motion. Two types of trajectory approaches can generate trajectory profiles using the cartesian and joint space approaches. The cartesian space trajectory generation approach will involve inverse kinematics. There will be multiple solution problems in which matching with human motion will be an issue. Craig [1] stated that there are three general problems in the cartesian space trajectory scheme: first, unreachable intermediate points; second, high joint rates near a singularity; and lastly, start and goal being reachable in different solutions. Meanwhile, in the joint space trajectory approach, the robot's motion design is to be made using joint space values such as joint space position, velocity, etc. The general problem of this approach is the lack of visualization of joint motion and position of the end-effector during the time.

Various methods of implementation have already been established. Within that, the accuracy of the trajectory is an important parameter when designing the trajectory of the robot that operates alongside humans, such as an exoskeleton. The trajectory's accuracy can be viewed as an error between generated (target) trajectory and the reference robot trajectory. The less error shows a higher accuracy of the generating trajectory. The more accurate trajectory generated, the more naturally it follows human motion.

Many methods have been discussed to improve the accuracy of trajectory in exoskeleton robots. In Gomes et al. [13], Gait Pattern Adaptation (GPA) was designed to generate natural human motion for the LOKOMAT rehabilitation robot. The GPA adjusted the generating trajectory to suit the desired trajectory based on torque interaction between humans and robots. This method was also studied in different approaches [14-16]. The NaTUre-gaits is also a technique that is used to generate the trajectory that follows natural

human motion for rehabilitation studies [17]. These two methods used predefined trajectories generated from mathematical (cubic) or recorded motion data. Other methods like Complementary Limb Motion Estimation (CLME) [18], Neural Oscillator [7], Gait Phase Switching Algorithm (GPSA) [6], Radial Basis Function (RBF) [19], Neural Network [9], polynomial [20,21], Probabilistic Foam Method (PFM) [22] have also been used to improve accuracy of the generated trajectory profiles to the wearer.

However, all the methods discussed so far require redefining constraint parameters (i.e., start time, stop time, start velocity, stop velocity, etc.) before implementing the method in the robot can be done. There are advantages of having many constraints that need to be considered, such as smooth trajectory, accuracy, etc. However, these constraints can increase the trajectory generator's computation cost and computation error [2].

This paper presents the trajectory generation algorithm that generates an accurate human-like trajectory profile to overcome the limitations stated. Unlike other methods, this method does not require additional parameters such as velocity or acceleration at the trajectory via-point to design trajectory profile. The only parameter required in this design is the joint displacement of the actuator. The complete cycle of time series data consisting of natural human motion data is mapped using the curve fitting approach; then, the quintic polynomial coefficient is produced. The polynomial trajectory is used because of its high precision and ability to calculate the kinematics, dynamics, and control parameters [23].

This paper's main contribution lies in proposing a trajectory generation algorithm that generates highly accurate and natural human-like trajectory profiles. This algorithm can potentially improve the performance and usability of robotic systems in a range of applications, such as grasping and manipulation tasks.

2. POLYNOMIAL CURVE FITTING TRAJECTORY PLANNING (PCFTP)

In robotics, trajectory planning is designed by assigning the initial (t_0) and final time (t_f) and other constraints on position, velocity, acceleration and so on at t_0 and t_f [2]. For the robot application that works alongside humans, such as an exoskeleton, the performance of the trajectory generator is how close the generated trajectory is to the human reference trajectory. Figure 1 shows the application of an exoskeleton robot in the rehabilitation process taken from [24]. In this application, the robot must generate a similar joint trajectory profile to the normal human walking joint to ensure the patient recovers quickly and avoids discomfort.
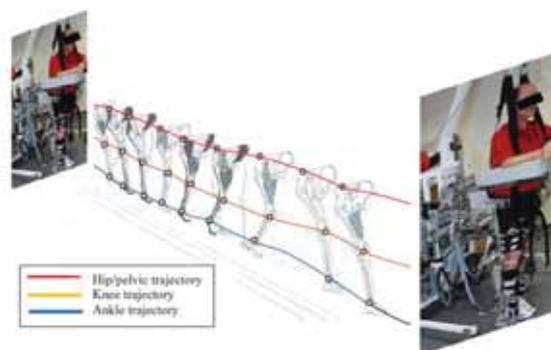

Fig. 1: Illustration of hip, knee and ankle joint trajectory in exoskeleton robot application taken from [24].

The trajectory generation technique requires some parameters to be set up before generation of the required trajectory. Table 1 shows the initial trajectory parameters. The trajectory generation technique requires the generation of a reference trajectory. It is shown that the current trajectory technique requires multiple initial setup trajectory parameters. Having multiple parameters in the initial setup will improve the accuracy performance of the trajectory generator. However, this will burden the system for processing and require exact mathematical modeling of the robot.

Table 1: Initial parameter setup for trajectory generation technique

Trajectory Generation Technique	Initial Trajectory Parameters
Quintic Polynomial	Start Boundary Parameter: i.e., Starting angular position, velocity, acceleration, jerk, t_{Start}
	Stop Boundary Parameter: i.e., Stoping angular position, velocity, acceleration, jerk, t_{Stop}
Neural Oscillator	Neuron Parameters: Frequency, coupling coefficients, different phase matrix
	CoM Trajectory Parameters:
Gait Phase Switching Algorithm (GPSA)	Neuron Parameters: Frequency, coupling coefficients, different phase matrix
	Gait Parameters: i.e., amplitude, regulating signal

Figure 2 shows the algorithm based on a curve fitting approach to identify the polynomial coefficients. This algorithm can be applied to all strategies discussed in [3] since it requires complete data to represent the motion. These data can be off-line, obtained from biomechanical studies, or come from an on-line time-series sequence received from the actuator encoder. However, in this paper, our algorithm will implement and evaluate only off-line situations. This paper used the human hip motion profile data from a biomechanical study [24] as a reference trajectory.

```
Algorithm 1: PCFTP
1:  Determine Data size of motion profile, mData
2:  Rearrange motion data followed by phases boundary parameter i = phases number
3:  Determine the size of phases data mSi
4:  while (i < imax)
5:      Select polynomial power (n)
6:      while (∑mSi < mData)
7:          Calculate the polynomial coefficient for selected n, mData, and mSi
8:      end
9:  end
```

Fig. 2: PCFTP Algorithm.

Step 1: Data Acquisition

The dataset used in this paper was based on a biomechanics study conducted by [25], which employed 40 healthy subjects. The dataset consisted of two different groups, adult and young, with different walking speeds recorded and labeled as normal walking (N), very slow walking (XS), slow walking (S), medium walking (m), fast walking (L), walking on toe (T), walking on heel (H), stair ascending (U) and stair descending (D). In this paper, we used an adult hip normal walking profile. This paper does not cover how joint trajectory is

generated from 40 healthy subjects and the criteria for selecting their subject. Other datasets can also be used and are not rigid from the stated researcher.

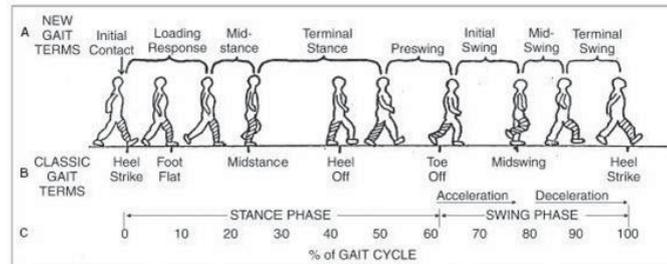**Step 2: Divided Motion Data Based on Biomechanical Gait Cycle**

Fig. 3: Phases of human walking gait.

In human biomechanical studies, gait is defined as any method of locomotion characterized by periods of loading and unloading of the limbs [26]. This definition is not restricted to walking; it includes running, hopping, skipping, swimming, and cycling. Since walking is the most frequently used gait in the activities of daily living (ADLs), most of the research definition of gait refers to walking.

The gait cycle in walking is defined as a time interval or sequence of motion occurring from heel strike to heel strike [27]. Fig. 3 shows the complete phases of the human walking gait, with seven phases in the human walking motion [28]. In general, human walking data shows that there is smooth and continuous motion between phase transitions (via-point section). This became the most challenging parameter for both methods to design trajectory profiles while accurately ensuring motion continuity and smoothness.

Table 2 shows the complete walking cycle's human hip motion parameter phase. Each phase has its own parameter starting and ending. Based on [1], these starting and ending parameters are critical when designing the via-point.

Table 2: Human hip phases joint angle parameters based on [28] and [25]

Gait Phases	Gait Cycle (%)	Starting Angle (Deg)	Ending Angle (Deg)	Starting time (s)	Ending time (s)
Initial Contact	0 – 10	25.20	21.00	0.00	0.10
Mid Stance	10 – 30	21.00	-1.30	0.10	0.30
Terminal Stance	30 – 50	-1.30	-16.10	0.30	0.50
Pre Swing	50 – 60	-16.10	-10.40	0.50	0.60
Initial Swing	60 – 73	-10.40	15.10	0.60	0.73
Mid Swing	73 – 87	15.10	25.10	0.73	0.87
Terminal Swing	87 - 100	25.10	24.10	0.87	1.00

Step 4: Polynomial Curve Fitting Trajectory Planning

The curve fitting technique is the main core of our proposed algorithm. It determines the polynomial coefficient from the normative hip walking profile. The curve fitting method is chosen because of its ability to deal with a series of data. There are two approaches in the curve fitting technique that have the ability to generate a polynomial coefficient. The first is Least-squares regression (LSR), or regression in short, and the second is an interpolation. This paper used the least-square regression approach to generate a polynomial coefficient.

The polynomial regression curve fitting technique is formulated in Eq. (1). Where a_i is a polynomial coefficient. The core idea of this technique is the regression that is used to

minimize the sum of the squares of residuals, S_r between the desired y_i and the forecast y'
as shown in Eq. (2), where all summation from $i = 1$ through m (number of data) and n
(polynomial order).

$$\theta(t) = \sum_{i=0}^{n} a_i t^i \tag{1}$$

$$S_r = \sum_{i=1}^{m} \sum_{j=0}^{n} \left(y_i - a_j x_i^j \right)^2 \tag{2}$$

Equation (3) shows the partial derivative of the sum of the residuals over the coefficient
$\left(\frac{\partial S_r}{\partial a_j}\right)$ in Eq. (2). and also in a matrix form of Eq. (3) as shown in Eq. (4).

$$\frac{\partial S_r}{\partial a_j} = -2 \sum_{i=1}^{m} \sum_{j=0}^{n} x_i^j \left(y_i - a_j x_i^j \right) \quad \begin{cases} i=1,2,\dots..m \\ j=1,2,\dots..n \end{cases} \tag{3}$$

$$\begin{bmatrix} m & \sum x_i^j & \dots & \sum x_m^n \\ \sum x_i^j & \dots & \dots & \sum x_m^{n+1} \\ \vdots & \dots & \dots & \vdots \\ \sum x_i^n & \dots & \dots & \sum x_m^{2n} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \vdots \\ \sum x_i^n y_i \end{bmatrix} \tag{4}$$

The standard error $S_{y/x}$, Eq. (5) and the coefficient correlation, Eq. (6), is used as a
performance evaluation for the PCFTP method.

$$S_{y/x} = \sqrt{\frac{S_r}{m-(n+1)}} \tag{5}$$

$$r^2 = \frac{\sum (y_1 - \bar{y}) - \sum S_r}{\sum (y_1 - \bar{y})} \tag{6}$$

3. SIMULATION SETUP

We used MATLAB to demonstrate the accuracy of the trajectory generated from the
PCFTP algorithm and the effectiveness of generation using the trajectory algorithm
compared to the quintic polynomial method.

3.1 Comparison of Quintic Trajectory Generation for Generating Hip Joint Trajectory Profile

The quintic polynomial is a fifth-degree polynomial formulated in Eq. (7). Besides
position, other constraints, such as velocity and acceleration, must be considered when
designing the trajectory. A higher-order polynomial gives more constraints (time
derivatives) that can be used to adjust the trajectory to suit the application requirement.
However, the more constraints are considered, the more complex it is to determine the
unknown coefficients. This constraint provided high computation cost and produced a
numerical error for a higher value of polynomial degree [2].

The quintic polynomial can provide five constraints instead of three, as mentioned in
Eqs. (7) to (9). These three constraints are enough to obtain a smooth trajectory profile. The
suitable initial and final constraints for the position Eq. (7), velocity Eq. (8), and acceleration

Eq. (9) [2] are determined first. Besides, the number of boundary constraints is usually even, and the degree of the polynomial function is odd.

$$\theta(t) = \sum_{i=0}^{n} a_i t^i \tag{7}$$

$$\dot{\theta}(t) = \sum_{i=1}^{n} i a_i t^{i-1} \tag{8}$$

$$\ddot{\theta}(t) = \sum_{i=2}^{n} i(i-1) a_i t^{i-2} \tag{9}$$

Therefore, from Eq. (7) to (9), there are six boundary constraints (starting position, θ_0, ending position, θ_f, starting velocity, v_0, ending velocity, v_0, starting acceleration, s_0 and ending acceleration, s_f) of a quintic polynomial as shown in Eq. (10), that needs to be determined.

$$\begin{aligned}
\theta(t_0) &= \theta_0 \\
\theta(t_f) &= \theta_f \\
\dot{\theta}(t_0) &= v_0 \\
\dot{\theta}(t_f) &= c \\
\ddot{\theta}(t_0) &= s_0 \\
\ddot{\theta}(t_f) &= s_f
\end{aligned} \tag{10}$$

This constraint is then used to determine the quintic coefficient, as shown in Eq. (11). The first three coefficients, a_0, a_1, and a_2, are the initial values of the generated trajectory's position, velocity, and acceleration.

Table 3 shows the polynomial coefficients generated using quintic polynomial trajectory planning for each sub-phase. Table 4 shows the polynomial coefficients using curve fitting. All these coefficients are then used again to generate the trajectory. This trajectory is then compared to the recorded trajectory profile to validate the accuracy and effectiveness of the proposed method.

$$\begin{aligned}
a_0 &= \theta_0 \\
a_1 &= \dot{\theta}_0 \\
a_2 &= \frac{\ddot{\theta}_0}{2} \\
a_3 &= \frac{20\theta_f - 20\theta_0 - (8\dot{\theta}_f + 12\dot{\theta}_0)t_f - (3\ddot{\theta}_f - \ddot{\theta}_0)t_f^2}{2t_f^3} \\
a_4 &= \frac{30\theta_f - 30\theta_0 - (14\dot{\theta}_f + 16\dot{\theta}_0)t_f - (3\ddot{\theta}_f - 2\ddot{\theta}_0)t_f^2}{2t_f^4} \\
a_5 &= \frac{12\theta_f - 12\theta_0 - (6\dot{\theta}_f + 6\dot{\theta}_0)t_f - (\ddot{\theta}_f - \ddot{\theta}_0)t_f^2}{2t_f^5}
\end{aligned} \tag{11}$$

4. RESULTS

This section comprehensively analyses a significant problem known as the via-point disjointed problem. Within this context, we explore trajectory accuracy, employing two distinct methods: the quintic polynomial method and the PCFTP method. By thoroughly examining and comparing the outcomes generated by these methods, we aim to understand better their respective strengths, limitations, and overall effectiveness in achieving accurate trajectories. **Error! Reference source not found.** shows the hip joint motion simulation r esult based on the quintic polynomial trajectory method, and **Error! Reference source not found.** shows the hip joint motion simulation result based on the polynomial curve fitting method.
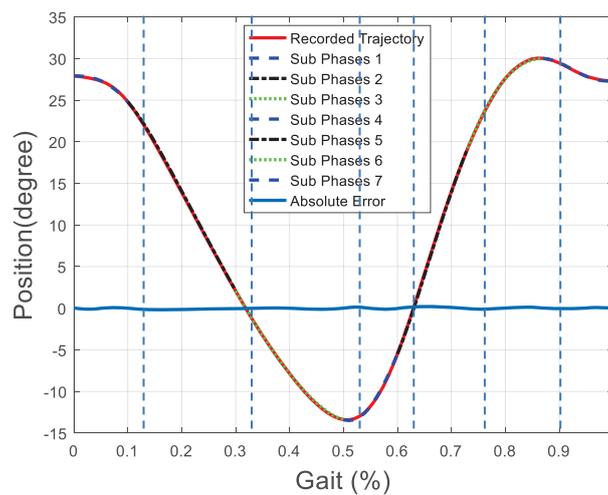
Fig. 4: Hip joint trajectory profile generated using quintic polynomial.

4.1 Via-point Disjointed Analysis

Via-point is an intermediate position that the system or object needs to pass through during its movement from the initial position to the final destination. The via-point is located at every phase's transition. Based on Winter [28], the seven phases of human walking require six via points for the complete cycle. As mentioned in the previous section, the ending position for Phase 1 is a starting position for Phase 2.

Fig. 5: Hip joint trajectory profile generated using PCFTP.

Abdul Jalil et al.

Table 3: Coefficient of a polynomial using quintic polynomial trajectory method

Phases	Coefficient					
	a_5	a_4	a_3	a_2	a_1	a_0
1	-8.102×10^5	2.108×10^5	17178	0	0	25.2
2	32128	-35861	15985	-3572.4	285.07	15.497
3	-2268.6	4303.8	-2828.9	958.46	-298.37^4	48.981
4	-8.039×10^6	-2.216×10^6	2.439×10^6	1.339×10^6	3.666×10^5	-40040
5	1.932×10^5	-6.268×10^5	8.072×10^5	-5.152×10^5	1.630×10^5	-20480
6	-99297	4.0306×10^5	-6.5399×10^5	5.2952×10^5	-2.136×10^5	34301
7	64871	$--3.364 \times 10^5$	6.915×10^5	-7.052×10^5	3.571×10^5	-71816

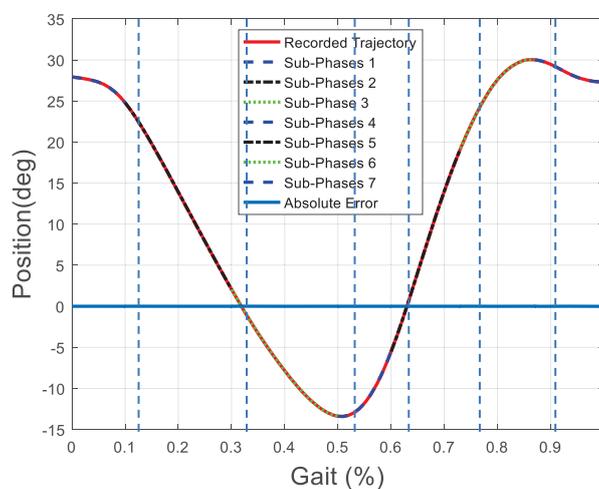Table 4: Coefficient of a polynomial using the PCFTP method

Phases	Coefficient					
	a_5	a_4	a_3	a_2	a_1	a_0
1	75178	-82804	14364	-1158.9	6.0584	25.193
2	91252	-88042	32870	-5978.7	424.14	13.481
3	-88293	1.72×10^5	-1.32×10^5	50020	-9507.7	732.37
4	2.140×10^5	-5.62×10^5	5.90×10^5	-3.09×10^5	80235	-8293.4
5	1.61×10^5	-5.34×10^5	7.02×10^5	-4.58×10^5	1.48×10^5	-19082
6	84648	-3.37×10^5	5.33×10^5	-4.22×10^5	1.67×10^5	-26431
7	2.17×10^5	-1.06×10^6	2.08×10^6	-2.02×10^6	9.84×10^5	-1.91×10^5

Table 5 compares generated via-point for each sub-phase between the quintic polynomial and PCFTP. It shows that the quintic polynomial method handles via-points without encountering problems during each phase. However, the PCFTP method demonstrates a notable concern with disjointed via-points in the phase transitions. This disparity highlights the quintic method's ability to seamlessly incorporate via-points across all phases, while the PCFTP method experiences challenges in maintaining continuity and cohesiveness in the trajectory when navigating through different phases.

Table 5: Start and Stop Position at via-point

Phases	Quintic Polynomial (deg)		PCFTP (deg)	
	Start	*Stop*	*Start*	*Stop*
1	25.200	21.000	25.193	21.046
2	21.000	-1.300	21.087	-1.265
3	-1.300	-16.100	-1.303	-16.167
4	-16.100	-10.400	-16.143	-10.443
5	-10.400	15.100	-10.453	13.473
6	15.100	25.100	13.480	25.316
7	25.100	24.100	25.310	24.086

The issue with our algorithm is rooted in the fundamental characteristics of the curve fitting technique. This method utilizes a best-fitting strategy to effectively incorporate and precisely depict the data within a designated interval, as outlined in Eq. (5) and Eq. (6). On the other hand, the quintic method integrates the initial and final values as constraints in Eq. (11). The utilization of constraints guarantees that the trajectory's initiation and termination occur at predetermined positions, thereby enabling accurate management of the start and end points.

In Table 6, the absolute errors associated with each via-point are presented, reflecting the performance of the PCFTP method. Notably, the most significant error is observed during the transition from phase 4 to phase 5, specifically for the stop position, with an error magnitude of approximately 1.627°. Additionally, the transition from phase 5 to phase 6 exhibits an error of approximately 1.620°. It is important to note that these errors directly impact the overall smoothness and continuity of the trajectory, underscoring the need for further analysis and potential improvements in the PCFTP method to mitigate such errors and enhance trajectory quality.

Table 6: Absolute Error of a Via-point in Each Phases Transition for Curve Fitting

Phases Transition	Absolute Error Start Position (deg)	Absolute Error Stop Position (deg)
1 – 2	0.007	0.046
2 – 3	0.087	0.035
3 – 4	0.003	0.067
4 – 5	0.043	1.627
5 – 6	1.620	0.216
6 – 7	0.210	0.014

4.2 Overall Fitting Performance Analysis

An extensive evaluation was performed to compare the fitting errors of the quintic polynomial method and PCFTP methods' fitting errors to obtain significant insights into the precision and efficiency of trajectory generation techniques. Fig. 6 compares fitting errors between the quintic method and the curve fitting method. Based on Fig. 6, the PCFTP approach has a lower fitting error than the quintic trajectory generation.
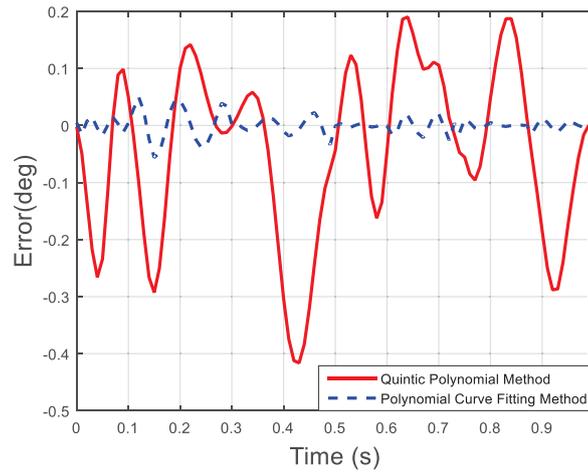
Fig. 6: Fitting error comparison between quintic polynomial and PCFTP methods.

Table 7 shows the comparison of MAE for each phase for both methods. MAE is the mean of fitting error for each phase. Table 6 shows that the polynomial curve fitting method gives a high efficiency with the lowest generation error 0.002° phase. However, the lowest generation error for the quintic polynomial approach is 0.086° phase. The overall MAE for the quintic polynomial approach is 0.112° phase and 0.011° phase for the polynomial curve fitting approach. The MAE represents the difference between the desired trajectory and generated trajectory.

Table 8 shows the Root Mean Square Error (RMSE) comparison between the quintic polynomial and PCFTP methods. These values indicate the accuracy of the trajectory profile. It shows that the PCFTP method exhibits superior precision in trajectory profile compared to the quintic polynomial method.

Table 7: Mean absolute error (MAE) comparison between quintic and PCFTP methods

Sub-Phase	Method	
	Quintic (deg)	*Curve Fitting (deg)*
1	0.115	0.009
2	0.101	0.032
3	0.164	0.012
4	0.086	0.002
5	0.101	0.013
6	0.089	0.004
7	0.129	0.006

Table 8: RMSE comparison between quintic polynomial and PCFTP methods

Sub-Phase	Method	
	Quintic (deg)	*Curve Fitting (Deg)*
1	7.030×10^{-2}	6.783×10^{-15}
2	2.900×10^{-2}	4.289×10^{-14}
3	1.431×10^{-1}	4.602×10^{-13}
4	1.440×10^{-2}	4.151×10^{-11}
5	9.210×10^{-2}	1.221×10^{-11}
6	3.730×10^{-2}	6.500×10^{-11}
7	1.267×10^{-1}	3.035×10^{-10}

Considering the RMSE values obtained, we can observe that the PCFTP method consistently outperforms the quintic polynomial method regarding trajectory accuracy. This outcome underscores the efficacy of the PCFTP method in achieving a closer fit to the desired trajectory, resulting in minimized deviations from the intended motion path.

These findings highlight the potential advantages of adopting the PCFTP method over the quintic polynomial method, particularly in applications with high trajectory accuracy, such as robotics, motion planning, and autonomous navigation systems.

5. CONCLUSION

In conclusion, the proposed PCFTP algorithm generates natural human motion profiles more accurately than the quintic polynomial trajectory generation method. The algorithm does not require constraint parameters such as angular position, velocity, and acceleration, which is a significant advantage. However, it requires the complete datasets or trajectory profiles to be modelled, which can sometimes be challenging. The proposed algorithm achieved a trajectory profile accuracy of 99.8% and precision of 0.002°, significantly improving over existing methods. Moreover, the performance validation is based on human biomechanics walking data from [25]. This walking data is not based on a single person but on the normalization of 40 healthy people. Different biomechanics data can be used as well instead of this data. However, the result might vary slightly due to a few factors. First, the curve fitting technique performance is based on data trends in each segment or phase. If the data trend is well segmented, the fitting performance is high. Second, the polynomial degrees are used to model the trajectory data.

However, an open question for a proposed method is how to improve, especially in disjointed via-point problems on every phase transition. This disjointed issue will affect the overall smoothness of the motion profile. The future work for this research is the new extension of this method that determines the numbers of via-points based on best fitting accuracy (Eqs. (5) and (6)) and continuity. Also, implementing the proposed algorithm to real-time data will guide us to a new dimension of using the curve fitting approach in trajectory planning since the curve fitting approach works only if all fitting data are complete.

ACKNOWLEDGMENT

REFERENCES

[1] Craig JJ. (2014) Introduction to Robotics: Mechanics and Control, 3rd Edition Pearson.

[2] Biagiotti L, Melchiorri C. (2009) Trajectory Planning for Automatic Machines and Robots. Berlin, Heidelberg, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-540-85629-0.

[3] Miskon MF, Yusof MBAJ. (2014) Review of trajectory generation of exoskeleton robots., in 2014 IEEE International Symposium on Robotics and Manufacturing Automation (ROMA), IEEE, 12-17. doi: 10.1109/ROMA.2014.7295854.

[4] Wang H, Zhao Q, Li H, Zhao R. (2022) Polynomial-based smooth trajectory planning for fruit-picking robot manipulator. Information Processing in Agriculture, 9(1): 112-122. doi: 10.1016/j.inpa.2021.08.001.

[5] Kulak T, Silverio J, Calinon S. (2020) Fourier movement primitives: an approach for learning rhythmic robot skills from demonstrations., in Robotics: Science and Systems XVI, Robotics: Science and Systems Foundation. doi: 10.15607/RSS.2020.XVI.056.

[6] Wei S, Wu H, Liu L, Zhang Y, Chen J, Li Q. (2022) A CPG-based gait planning and motion performance analysis for quadruped robot. Industrial Robot: the international journal of robotics research and application, 49(4): 779-797. doi: 10.1108/IR-08-2021-0181.

[7] Liu C, Yang J, Bu W, Chen Q. (2016) A trajectory generation method for biped walking based on neural oscillators., in 2016 IEEE 13th International Conference on Networking, Sensing, and Control (ICNSC), IEEE, 1-6. doi: 10.1109/ICNSC.2016.7478995.

[8] Wang Q, Qian J, Zhang Y, Shen L, Zhang Z, Feng Z. (2007) Gait trajectory planning and simulation for the powered gait orthosis., in 2007 IEEE International Conference on Robotics and Biomimetics (ROBIO), IEEE, 1693-1697. doi: 10.1109/ROBIO.2007.4522420.

[9] Watanabe T, Johnson EN. (2018) Trajectory Generation using Deep Neural Network. 2018 AIAA Information Systems-AIAA Infotech @ Aerospace, (January): 1-12. doi: 10.2514/6.2018-1893.

[10] Aly AA, Vu MT, El-Sousy FFM, Alotaibi A, Mousa G, Le D, Mobayen S. (2022) Fuzzy-Based Fixed-Time Nonsingular Tracker of Exoskeleton Robots for Disabilities Using Sliding Mode State Observer. Mathematics, 10(17): 3147. doi: 10.3390/math10173147.

[11] Kundu AS, Mazumder O, Chattaraj R, Bhaumik S, Lenka PK. (2014) Trajectory generation for myoelectrically controlled lower limb active knee exoskeleton., in 2014 Seventh International Conference on Contemporary Computing (IC3), IEEE, 230-235. doi: 10.1109/IC3.2014.6897178.

[12] Schuy J, Mielke T, Steinhausen M, Beckerle P, Rinderknecht S. (2015) Design and evaluation of a sensor minimal gait phase and situation detection Algorithm of Human Walking. IEEE-RAS International Conference on Humanoid Robots, 2015-Dec: 20-25. doi: 10.1109/HUMANOIDS.2015.7363517.

[13] Gomes MA, Silveira GLM, Siqueira AAG. (2011) Gait pattern adaptation for an active lower-limb orthosis based on neural networks. Advanced Robotics, 25(15): 1903-1925. doi: 10.1163/016918611X588899.

[14] Siqueira AAG, Jardim B, Vilela PRC, Winter TF. (2008) Analysis of gait-pattern adaptation algorithms applied in an exoskeleton for lower limbs., in 2008 16th Mediterranean Conference on Control and Automation, IEEE, 920–925. doi: 10.1109/MED.2008.4602081.

[15] Gomes MA, Silveira GLM, Siqueira AAG. (2009) Gait-pattern adaptation algorithms based on neural network for lower limbs active orthoses., in 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009, IEEE, 4475-4480. doi: 10.1109/IROS.2009.5354232.

[16] Silveira GLM, Gomes MA, Siqueira AAG. (2010) First experimental results on adaptation algorithms for lower limbs active orthoses., in 18th Mediterranean Conference on Control and Automation, MED'10, IEEE, 1316-1321. doi: 10.1109/MED.2010.5547868.

[17] Wang P, Low KH, McGregor AH. (2011) A subject-based motion generation model with adjustable walking pattern for a gait robotic trainer: NaTUre-gaits., in 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 1743-1748. doi: 10.1109/IROS.2011.6048809.

[18] Vallery H, van Asseldonk EHF, Buss M, van der Kooij H. (2009) Reference Trajectory Generation for Rehabilitation Robots: Complementary Limb Motion Estimation. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 17(1): 23-30. doi: 10.1109/TNSRE.2008.2008278.

[19] Daachi ME, Madani T, Daachi B, Djouani K. (2015) A radial basis function neural network adaptive controller to drive a powered lower limb knee joint orthosis. Applied Soft Computing, 34: 324-336. doi: 10.1016/j.asoc.2015.04.034.

[20] Mohammed MQ, Miskon MF, Ali SA. (2017) The effectiveness of kinematic constraints on the accuracy of trajectory profile of human walking using PSPB technique. International Journal of Mechanical and Mechatronics Engineering, 17(6): 86-96

[21] Sapiee MR, Marhaban MH, Ishak AJ, Miskon MF. (2020) Trajectory Generation for Hip Rehabilitation Exoskeleton Using Trajectory Morphing Method. International Journal of Human and Technology Interaction, 4(1): 45-51

[22] Nascimento LBP, Barrios-Aranibar D, Alsina PJ, Santos VG, Fernandes DHS, Pereira DS. (2020) A Smooth and Safe Path Planning for an Active Lower Limb Exoskeleton. Journal of Intelligent and Robotic Systems: Theory and Applications, 99(3-4): 535-553.
doi: 10.1007/s10846-019-01134-7.

[23] Sidobre D, Desormeaux K. (2019) Smooth Cubic Polynomial Trajectories for Human-Robot Interactions. Journal of Intelligent & Robotic Systems, 95(3-4): 851-869.
doi: 10.1007/s10846-018-0936-z.

[24] Low KH. (2011) Subject-oriented overground walking pattern generation on a rehabilitation robot based on foot and pelvic trajectories. Procedia IUTAM, 2: 109-127.
doi: 10.1016/j.piutam.2011.04.012.

[25] Bovi G, Rabuffetti M, Mazzoleni P, Ferrarin M. (2011) A multiple-task gait analysis approach: Kinematic, kinetic and EMG reference data for healthy young and adult subjects. Gait & Posture, 33(1): 6-13. doi: 10.1016/j.gaitpost.2010.08.009.

[26] Kirtley C. (2006) Introduction., in Clinical Gait Analysis, Elsevier, 5-14.
doi: 10.1016/B978-0-443-10009-3.50004-7.

[27] Malanga G., Delisa J. (1998) Clinical Observation. In Gait analysis in the science of rehabilitation. Edited by Delisa J. Baltimore, Diane Publishing; pp 1-10.

[28] Winter DA. (2009) Biomechanics and Motor Control of Human Movement, 4th ed.: New Jersey: John Wiley & Sons, Inc. doi: 10.1002/9780470549148.

TRIPLE NONLINEAR HYPERBOLIC PID WITH STATIC FRICTION COMPENSATION FOR PRECISE POSITIONING OF A SERVO PNEUMATIC ACTUATOR

**KHAIRUN NAJMI KAMALUDIN[1], LOKMAN ABDULLAH[1*],
SYED NAJIB SYED SALIM[2], ZAMBERI JAMALUDIN[1], NUR AIDAWATY RAFAN[1],
MOHD FUA'AD RAHMAT[3] AND RPRAKASH RAMANATHAN[1]**

[1]*Faculty of Manufacturing Engineering, Universiti Teknikal Malaysia Melaka,
Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia*
[2]*Faculty of Electrical and Electornic Engineering Technology, Kampus Teknologi,
Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal,
Melaka, Malaysia*
[3]*Department of Control and Mechatronic Engineering (CMED), Faculty of Electrical
Engineering, Universiti Teknologi Malaysia Johor Bahru, 81310, Malaysia*

[*]Corresponding author: lokman@utem.edu.my

ABSTRACT: Accurate and precise positioning control is critical in designing a positioning servo pneumatic system. The internal friction force of the pneumatic is one of the disturbances that make it challenging to achieve accurate and precise positioning. Dynamic friction identification and modelling are usually very complex and computationally exhaustive. In addition, pneumatic actuators are nonlinear systems, and applying linear control to the system is a mismatch. This study proposes an enhanced triple nonlinear hyperbolic PID controller with static friction (T-NPID+F_{SS}) feedback module. T-NPID is integrated with nonlinear hyperbolic functions at each PID gain, hence the name. The reference in designing the T-NPID is the Popov stability criterion. Meanwhile, static friction (comparatively more straightforward than dynamic friction) is identified by measuring the actuator's internal friction at various velocities and applying it to the static friction model. T-NPID+F_{SS} is compared to a classical PID, a PID with static friction (PID+F_{SS}), and T-NPID without the friction module. With the comparisons, the performance gains of each module are clear. While most previous research focuses on the sinusoidal wave tracking performance (measuring the maximum tracking error, MTE, and root mean square error, RMSE), the analysis in this research focuses on obtaining precise positioning; steady-state analysis is the primary measurement. However, transient response and integral of absolute error (IAE) analysis are also observed to ensure no significant drawback in the controller's performance. T-NPID+F_{SS} achieved the best precise positioning control, with 88.46% improvement over PID, 71.15% over PID+F_{SS}, and 59.46% over T-NPID. The final controller is also on par with T-NPID for transient responses compared to the base PID. Although the F_{SS} model caters to friction compensation, optimizing the F_{SS} parameter by applying artificial intelligence, such as Neural Networks (NN) and Genetic Algorithm (GA), will increase the friction modeling's accuracy, and improve the compensation.

ABSTRAK: Kawalan kedudukan tepat dan jitu adalah kritikal dalam mereka bentuk sistem pneumatik servo penentu. Daya geseran dalaman pneumatik merupakan salah satu gangguan yang menyukarkan bagi mencapai kedudukan yang tepat dan jitu. Penentuan dan pemodelan daya geseran dinamik kebiasaannya sangat kompleks dan melibatkan pengiraan menyeluruh. Tambahan, penggerak pneumatik adalah sistem tak linear, dan

menggunakan kawalan linear pada sistem adalah tidak sesuai. Kajian ini mencadangkan kawalan PID hiperbolik tiga fungsi tak linear yang dipertingkatkan dengan modul suapan-balik geseran statik (T-NPID+F_{SS}). T-NPID diintegrasikan dengan tiga fungsi hiperbolik tidak linear pada setiap pekali PID, seperti namanya. T-NPID direka bentuk dengan kriteria kestabilan Popov. Manakala geseran statik (secara perbandingan lebih mudah daripada geseran dinamik) dikenal pasti dengan mengukur penggerak geseran dalaman pada pelbagai halaju dan menerapkannya pada model geseran statik. T-NPID+F_{SS} dibandingkan dengan PID klasik, PID dengan geseran statik (PID+ F_{SS}) dan T-NPID tanpa modul geseran. Melalui perbandingan, prestasi peningkatan setiap modul adalah jelas. Walaupun kebanyakan kajian terdahulu memfokuskan pada prestasi penentuan gelombang sinusoidal (mengukur ralat penentuan maksimum, MTE dan ralat purata kuasa dua, RMSE), analisis kajian ini memberi tumpuan kepada mendapatkan kedudukan yang tepat; oleh itu, analisis keadaan akhir adalah ukuran utama. Namun, tindak balas sementara dan analisis kamiran ralat mutlak (IAE) juga diperhatikan bagi memastikan tiada kekurangan ketara dalam prestasi kawalan. T-NPID+F_{SS} mencapai kawalan penentuan kedudukan tepat terbaik, dengan peningkatan 88.46% berbanding PID, 71.15% berbanding PID+F_{SS} dan 59.26% berbanding T-NPID. Kawalan akhir yang dicadangkan juga adalah setanding dengan T-NPID bagi respons sementara berbanding PID asas. Walaupun model F_{SS} memenuhi pampasan geseran, mengoptimumkan parameter F_{SS} dengan menggunakan kecerdasan buatan (*artificial intelligence*, AI) seperti *Neural Networks*, NN dan *Genetic Algorithms*, GA akan meningkatkan ketepatan dan pampasan pemodelan geseran.

KEYWORDS: *servo pneumatic actuator; nonlinear control; PID controller; steady-state error; transient response; static friction*

1. INTRODUCTION

Many industrial applications such as those requiring manipulators, riveting machines, automobiles, pick-and-place devices, and others have made extensive use of pneumatic actuators. This is because pneumatic systems have a variety of benefits, including ease of maintenance, lower cost, low heat under steady load and many others [1-3]. The pneumatic actuator continues to garner ever more research interest as a result of these benefits. However, achieving the performance of great precision and accuracy yields obstacles. Nonlinear pneumatic actuators present challenges in the forms of control field with disturbances such as wide dead zones, air compression nonlinearity, low damping, and frictional forces [4]. The internal friction force is agreed to be one of the critical disturbances in an actuator that affects the performance needed to achieve precise positioning and trajectory tracking in a servo system [5-6]. In a pneumatic actuator, friction occurs between the seal and the internal cylinder wall [5,7]. According to a review by [8], friction force compensation is mainly compensated by scholars compared to other disturbances such as air pressure and dead zone, which shows that it is a major disturbance in this field.

The PID controller is a linear control scheme and modifying a PID controller by integrating other elements, such as fuzzy and nonlinear mathematical functions, ultimately changes the controller's characteristics into a nonlinear control scheme. It is still desirable to enhance and improve the classical PID controller since it is the most widely used in the industry [9-10]. Integrating a PID together with other industrial automation control such as Programmable Logic Control (PLC), SCADA and a few others is a proven workable solution [11-12]. Nonlinear controllers for servo pneumatic systems usually adapt sliding mode controllers (SMC) [13-15], as well as a few other controllers such as model reference adaptive controller (MRAC) [16], fuzzy logic controllers (FLC) [17] and neural network (NN) controller [18]. Nonlinear controllers are commonly complex and are less applied in the industrial field.

An example of a controller that adapts both nonlinear controls with a friction compensation is an SMC with generalized Maxwell Slip (GMS) dynamic friction as explored by [19], whereby the author adapts both dynamic and static friction. The GMS model is applied during near zero velocity while static friction is implemented during velocity. The system reduces tracking errors compared to the Stribeck friction model. However, obtaining the zero-velocity hysteresis model is complex, as other models need to be applied. Another research adapting multiple surface SMC (MSSMC) with a friction model is by [20]. A dynamic LuGre model is employed for the dynamic friction compensation. The MSSMC were compared with and without the friction observer proving that the observer yields a lower tracking error. The derivation of MSSMC is complex and challenging due to the inputs required other than position, such as velocity and pressure. A study by [21] presented a dynamic adaptive backstepping SMC (DAB-SMC) with LuGre friction model compensation. The controller with friction compensation resulted in up to 15% in root mean square error (RMSE) in step positioning and sine-wave tracking. No nonlinear-PID (N-PID) based controllers have been successfully developed adapting friction compensating modules. Some examples of N-PID, such as [22] developed an enhanced self-regulating nonlinear PID (SN-PID) in which the nonlinear function's variables were designed to be adaptive using a self-regulating function. The same author also developed a multi-rate nonlinear PID (MN-PID), where the nonlinear function varies with the control of fuzzy logic and the disturbance compensated is the valve dead zone. The difference before and after adapting the module compensator was not calculated so the improvement rate was not concluded. [23] proposed a nonlinear PI (N-PI) controller without any add-on disturbance compensating module. The research focused on overshoot reduction, validating a few N-PI controllers against the classical PI controller. However, the precise positioning or steady-state results were also not presented.

Owing to the recent research gap in advances for N-PID, this research attempts to explore and analyze a new strategy of N-PID in combination with the friction disturbance compensator. T-NPID was previously developed to improve the performance of a classical linear PID enhancing the sinewave trajectory tracking for precise positioning performance of an XY table ball-screw system and step response of a servo pneumatic plant [24-25]. T-NPID would be a better match to the nonlinear characteristics of the servo pneumatic plant as explained by [4-5]. The stability of the controller is validated via Popov stability criterion. Finally, the static friction compensation module will enhance the T-NPID by compensating for the internal friction, therefore increasing the final steady-state positioning performance of the pneumatic system.

2. METHODOLOGY

2.1 Experimental Plant Configuration and Plant Modeling

The experimental plant configuration and equipment model used in the plant are presented in Fig. 1 and Table 1. The input and output signals are connected to the main electrical junction box which is then connected to the data acquisition (DAQ) module, allowing for two-way communication with the personal computer (PC). The human-machine interface utilized in this research is Matlab with a Simulink environment.
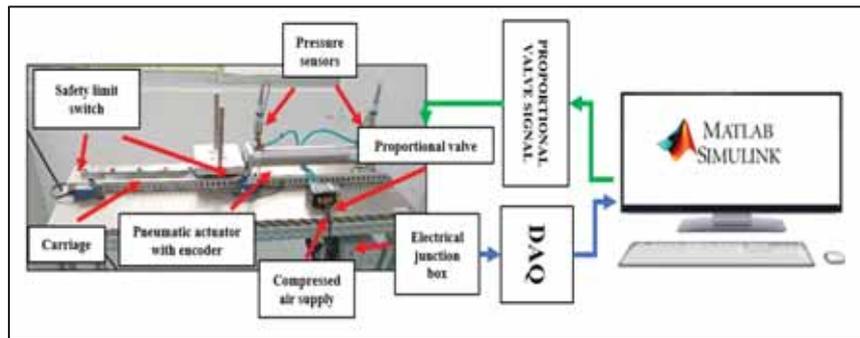
Fig. 1: Experimental plant setup and configuration.

Table 1: Equipment utilized for the experimental plant

Equipment	Model number	Specification
Proportional valve	Enfield LS-V15s	5/3 port, 0-10 Volts
Pneumatic actuator with integrated position encoder	Enfield ACTB-200-S10200	Double acting, 12 inch/ 304.80 mm stroke length
Pressure sensor	Gems Sensor 1200SGG	0 to 10 volts, 0-150 PSI
Safety limit switch	Tezuo AZ8104	Contact on
DAQ Box	National Instrument PCI-6221	37 Pin PCI
Matlab / SIMULINK	Version 2016b	N/A

In the control system field, modeling the experimental plant is the first predefined step before any other analysis or design is executed as agreed upon and performed by [26-27]. Modeling of the plant is performed using the system identification toolbox in Matlab. A predefined multi-sinewave signal in Volts is inserted into the experimental system in an open loop setup, and the response (in millimeters or voltage) of the system is recorded as explained by [28-29]. A pneumatic actuator is a nonlinear system; therefore, modeling of the plant is expressed in the state space matrix model. Linearization of the state space to a transfer function expression is also preferred for analysis simplification in the Matlab editor environment, such as gain margin, phase margin, and Nyquist stability analysis. The continuous state space matrix modeling is presented in Eq. (1) to Eq. (4). Linearization to the transfer function yields a third-order model as expressed in Eq. (5). The best-fit percentage obtained via system identification is 91.99%. Therefore, the model is valid for the analysis and design [28,30-31]. The block diagram presents the system's transfer function as a 'G' plant, as shown in Fig. 2.

$$A = \begin{bmatrix} -0.2921 & -0.02056 & -0.000085 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{1}$$

$$C = [0.03468 \quad 0.1468 \quad -0.002041] \, D = [0] \tag{2}$$

$$\dot{x} = Ax + Bu \tag{3}$$

$$y = Cx + Du \tag{4}$$

where A are the state vectors, B is the measured output, C are the measured input and D is noise.

$$G(s) = \frac{0.03468s^2 + 0.1468s - 0.002041}{s^3 + 0.2921\,s^2 + 0.02056s + 0.000085} \tag{5}$$
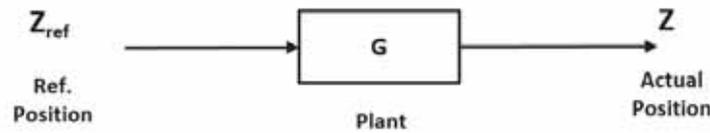

Fig. 2: Transfer function 'G' in a schematic block diagram.

2.2 Static Friction Identification and Modeling

Identifying the system's internal friction force is critical in order to compensate for the known disturbance. Identification of the static friction (also generally known as the sliding regime) requires the friction to be determined in the presence of velocity. Figure 3 presents the general friction force characteristics of velocity, with the presence of Stribeck and viscous friction effects [32]. The mathematical expression of the static friction, F_{SS}, is presented by Eq. (6) [19].
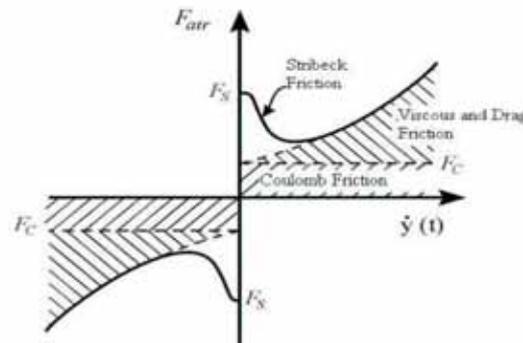
Fig. 3: Static friction model [7].

$$F_{SS}(v) = sgn(v).\left[F_c + (F_s - F_c)e^{-\left(\frac{v}{vs}\right)^2}\right] + F_v.v \tag{6}$$

where v, F_C, F_S, V_S and F_V are illustrated as velocity, Coulomb friction, static friction, Stribeck velocity and viscous friction, respectively.

To obtain each of the parameters expressed in Eq. (6) multiple voltages are applied to the valve of the system in an open loop configuration to actuate the pneumatic actuator to move in various constant velocities [32-33]. The pressure difference between the two pneumatic actuator chambers is recorded at each constant velocity. The difference in the pressure is converted to force by the equation $f = p \times a$, where f is the friction force, a is the effective area of the internal pneumatic piston, and p is the known pressure measured by the pressure sensor. The actuator's obtained friction force versus velocity is shown in Fig. 4.

Fig. 4: Experimental plot for velocity versus friction force.

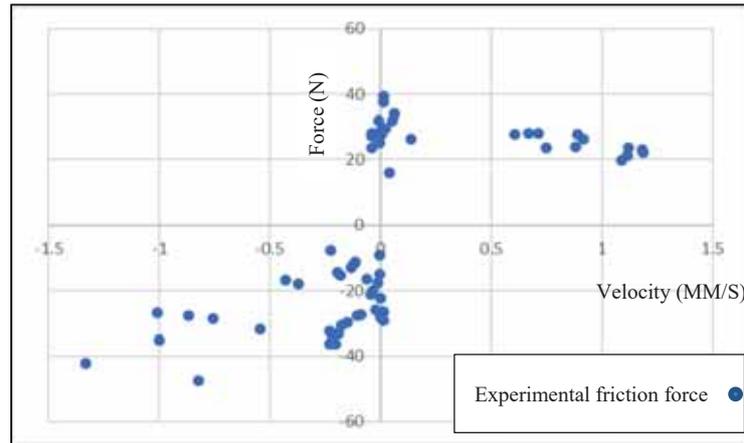A model of static friction is compared to the force versus velocity plot and optimized by a best-fit curve. A comparison of the model overlapped with the actual force from Fig. 4 is presented in Fig. 5 (a) and the final model applied as the F_{SS} model is shown in Fig. 5 (b). Table 2 displays the obtained static friction parameters in Eq. (6).

| (a) | (b) |
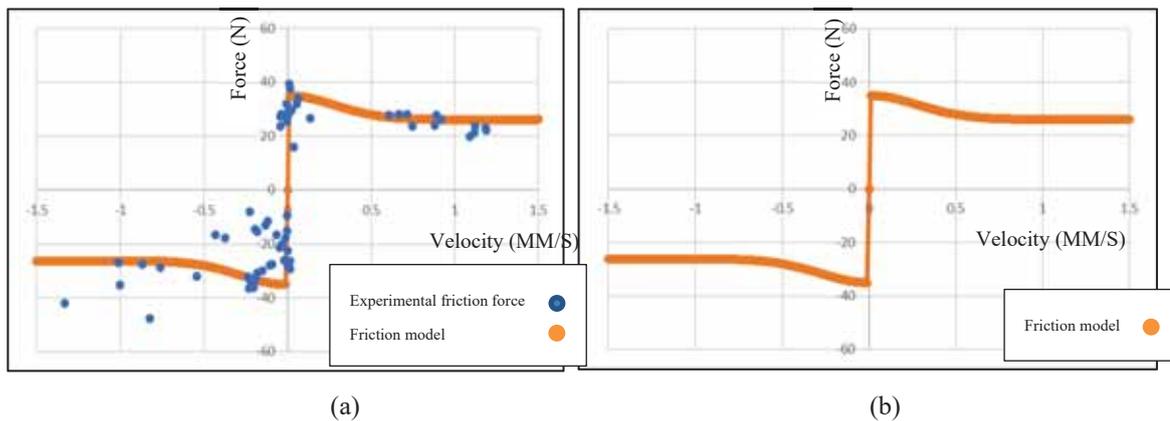
Fig. 5: (a) Comparison of the actual friction force and friction model; (b) applied friction model.

Table 2: Static friction parameter value

Static friction parameter	Parameter value
Coulomb friction, F_c (Newton)	26
Stribeck friction, F_s (Newton)	35
Stribeck velocity, V_s (mm/s)	0.4
Viscous friction, F_v (Newton·second / mm)	0.1

2.3 Validation of Friction Model with System Modeling

The static friction model is a feedback block to the transfer function, as presented in Fig. 6 [33]. Since the output of the transfer function block G is position (in mm), a derivative function is required to convert the displacement velocity of the actuator during motion as an input to the friction model, F_{SS}. A gain is used to convert the force value (in Newton) to voltage (volts). The procedure is conducted in a closed-loop system response by applying different known forces to the system multiple times. The counteract voltage by the controller is recorded. The gain obtained is, $K_f = -1/2020.20$ volt/Newton, as presented in Table 3. In

simple terms, the K_f gain translates to the system requiring 0.000495 volts to overcome 1 Newton of force.

Fig. 6: Modeling of the friction force with the servo pneumatic plant.

Table 3: Newton to voltage converter gain, K_f

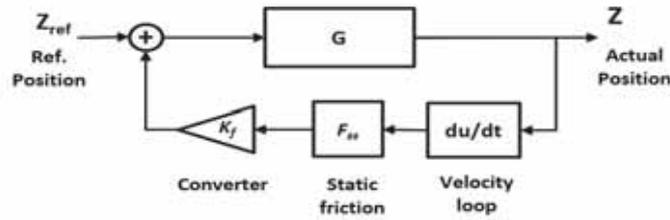| Gain type	Parameter value
K_f	-1/2020.20

Validation of the transfer function with the friction model is conducted by comparing the new plant and friction model in a simulation environment with the experimental response [34]. The same sinusoidal input performed during the system identification process is supplied to the new transfer function with the static friction model, and the output response data is captured. In order to observe the accuracy of modeling, the best-fit percentage is recalculated based on Eq. 7 [35-36] using the same equation applied by the Matlab system identification toolbox explained in Section 2.1. If the best-fit percentage drops below a predefined threshold of 90% accuracy, adjustment of the transfer function of plant 'G' will be required [32][37]. Figure 7 shows the estimation model and experimental data for the best-fit percentage calculation.

$$Best\ fit\ percentage = \left(1 - \frac{\sum |\hat{y}_{estimate} - y_{experimental}|}{\sum |y_{experimental} - y_{mean}|}\right) \times 100\% \tag{7}$$
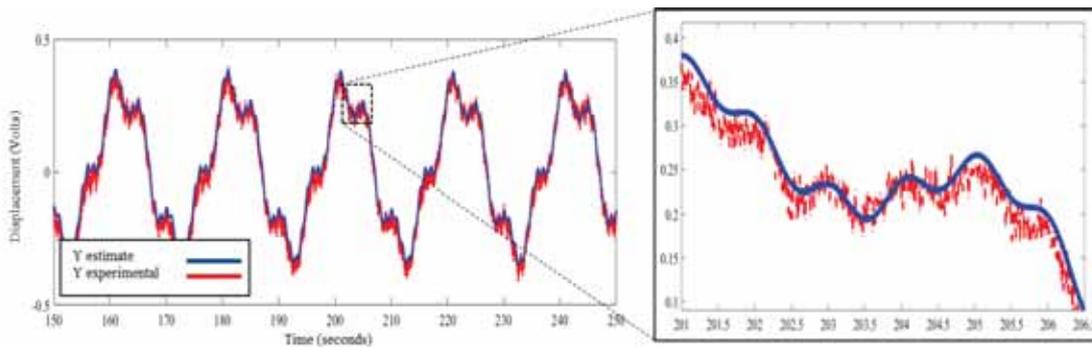
Fig. 7: Comparison of experimental data and model with friction data.

The best-fit percentage after including the static friction in the model is 91.373% which has decreased by 0.617% based on the initial system identification process. Although there is a reduction of modeling accuracy, the threshold of 90% is still in place; therefore, adjusting the transfer function is unnecessary in this case.

2.4 Design of PID Controller

Figure 8 presents a general PID block diagram for the plant. In this research, plant 'G'

is modeled with static friction ($G+F_{SS}$). The process flow related to tuning the PID controller gains follows the previous procedure performed by [25]. Table 4 tabulates the gain parameters obtained for the PID controller.
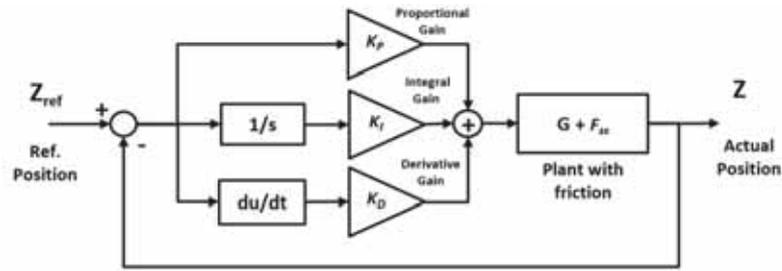
Fig. 8: General PID controller with plant block diagram.

Table 4: PID controller parameter

Gain type	Gain values
Proportional gain, K_P	10
Integral gain, K_I	0.1925
Derivative gain, K_D	0.35

2.5 Stability of PID Controller

Once the PID parameters are obtained, the Nyquist stability theorem is applied to the PID controller. An open loop configuration is observed in a Nyquist plot. Referring to Fig. 9, it is observed that the Nyquist plot does not encircle at -1 on the real axis, thus it is concluded that the system is stable.
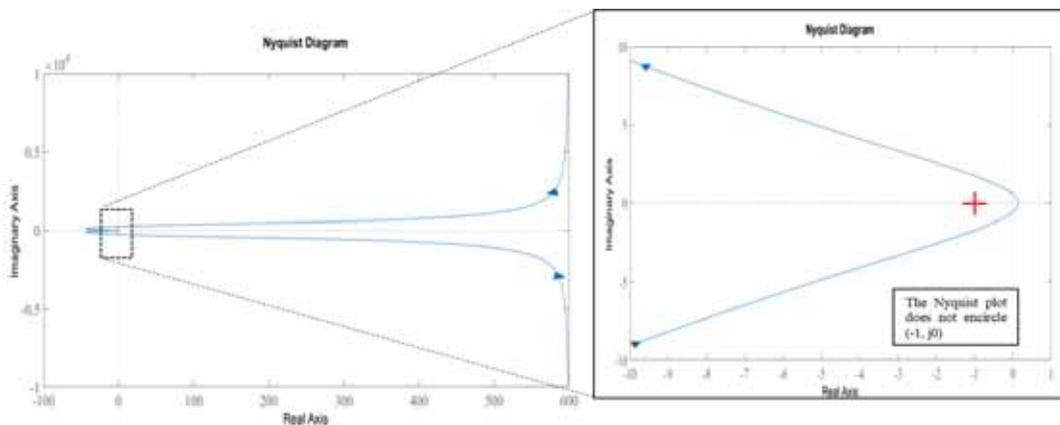
Fig. 9: (a) Nyquist plot for plant with PID controller; (b) close-up of Nyquist plot.

2.6 Design of the Triple Hyperbolic T-NPID Controller

Figure 10 presents the T-NPID block diagram. Three individual nonlinear hyperbolic functions are cascaded at each PID gain. The nonlinear hyperbolic functions are as presented in Eq. (8) to Eq. (10). Each nonlinear function will adapt accordingly to the error produced, multiplying the error exponentially if the error increases until the limit of predetermined maximum error, e_{Max}. Since there are three individual nonlinear functions, the errors are multiplied individually by the PID gains, adding flexibility to each gain of the PID.
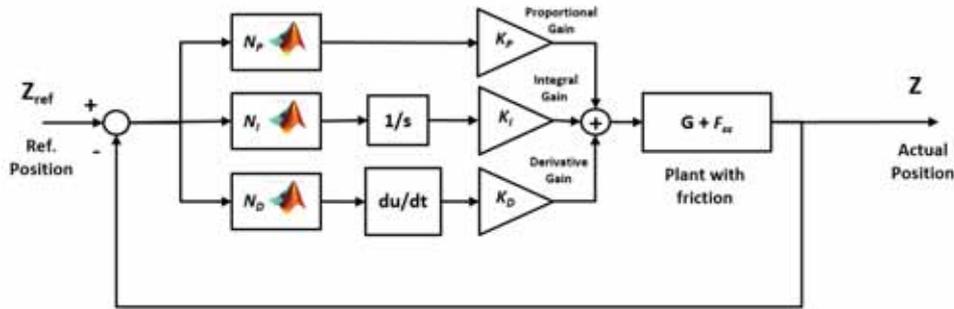
Fig. 10: Block diagram of a plant with friction model and T-NPID controller.

The errors are processed at each of the nonlinear function blocks before the nonlinear functions are multiplied by the error [24]. If the absolute error is smaller than or equal to e_{Max}, then the error is used by the nonlinear function. If the absolute error is larger than e_{Max}, then e_{Max} is multiplied by the signum of error and applied to the nonlinear function. The function written in each function block is as follows:

```
if
abs(error)<=eMax,
error = error;
else
e = (eMax)*sign(e);
end
```

$$K_P(e) = 1 + f \times [1 - \text{sech}(g \times e_P)] \tag{8}$$

$$K_I(e) = 1 \div [p + q \times (1 - \text{sech}(r \times e_I))] \tag{9}$$

$$K_D(e) = 1 + a \times [1 - \text{sech}(b \times e_D)] \tag{10}$$

The constants in Eq. (8) to Eq. (10) are tuned as per previous research [25]. Each parameter is tuned to achieve the best steady-state or actual positioning results. In addition, further tuning of the parameter was also performed to improve the transient response of the system. All the constants or parameter values are tabulated in Table 5.

Table 5: Nonlinear function parameter

Paramater	Value
f	1.65
g	30.1
e_P	0.095
p	34
q	24
r	10
e_I	0.35
a	1.9
b	1.75
e_D	0.25

Popov stability criterion is utilized in order to check the stability of the system with T-NPID controller. This criterion is a guideline so that the nonlinear gains operate within an allowable region of the controller and plant [38]. A reduced transfer function of the plant is applied for the analysis simplification [39]. In this research, three Popov plots from the criterion are developed as three nonlinear functions are applied at each PID gains. For each Popov plot, different equations are applied. The individual PID components Popov equations from Eq. (11) to Eq. (16) are summarized from the detailed analysis of [40].

The nonlinear P component of the Popov plot expression is in Eq. (11) and Eq. (12):

$$\mathcal{R}eW(j\omega) = \frac{k(-K_P \cdot \omega^2 + j \cdot K_P)}{[d^2 \cdot \omega^2 + (j - \omega^2)^2]} \tag{11}$$

$$\omega ImW(j\omega) = \frac{-k(d \cdot K_P \cdot \omega^2)}{[d^2 \cdot \omega^2 + (j - \omega^2)^2]} \tag{12}$$

The nonlinear I component of the Popov plot expression is in Eq. (13) and Eq. (14):

$$\mathcal{R}eW(j\omega) = \frac{k(d \cdot K_I)}{[d^2 \cdot \omega^2 + (j - \omega^2)^2]} \tag{13}$$

$$\omega ImW(j\omega) = \frac{-k(-K_I \cdot \omega^2 + j \cdot K_I)}{[d^2 \cdot \omega^2 + (j - \omega^2)^2]} \tag{14}$$

The nonlinear D component of the Popov plot expression is in Eq. (15) and Eq. (16):

$$\mathcal{R}eW(j\omega) = \frac{k(d \cdot K_I)}{[d^2 \cdot \omega^2 + (j - \omega^2)^2]} \tag{15}$$

$$\omega ImW(j\omega) = \frac{-k(K_D \cdot \omega^4 - j \cdot K_D \cdot \omega^2)}{[d^2 \cdot \omega^2 + (j - \omega^2)^2]} \tag{16}$$

2.7 Design of the T-NPID with Static Friction Compensator (T-NPID+F_{SS}) Controller

The final T-NPID+F_{SS} controller is shown in Fig. 11. The friction compensator (within the box) is the same basic schematic configuration as the static friction modeled in Fig.6. The difference is the summing junction attached to the plant, where the T-NPID+F_{SS} is subtracting the simulated friction, in other words, compensating for the friction. This configuration is in line with the general friction observer and compensator reviewed by [8]. For schematic simplifying, the friction compensator is combined as a simple block, F_{SS}, as shown in Fig. 12.

In experimental mode, the basic schematic is shown in Fig. 13. The modeled plant and friction in the simulation, as shown in the previous basic schematics, are changed to the proportional valve and the actuator position encoder. The output of the TNPID+F_{SS} is connected to the plant's proportional valve, and the plant's output, the actual position (Z), is obtained from the actuator position encoder. Z is feedback for both friction compensation and a closed-loop controller system.
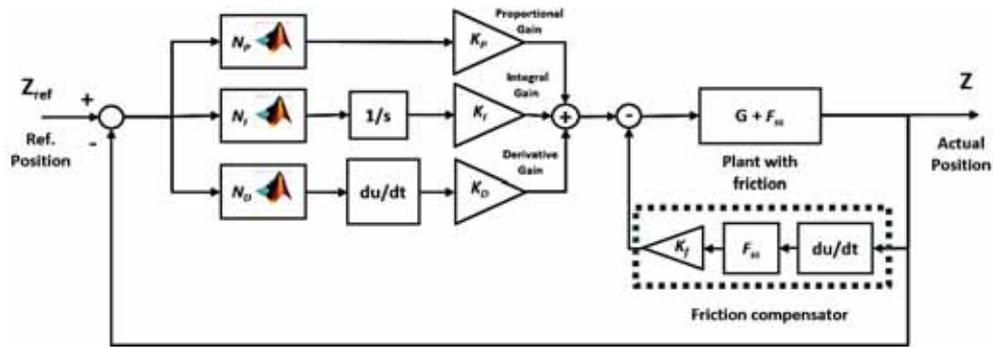
Fig. 11: The T-NPID+F_{SS} controller scheme.
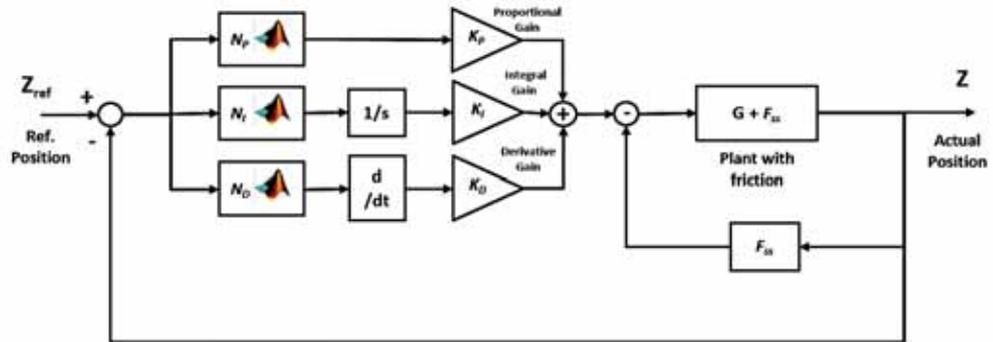
Fig. 12: Simplifying the friction compensator basic schematic.
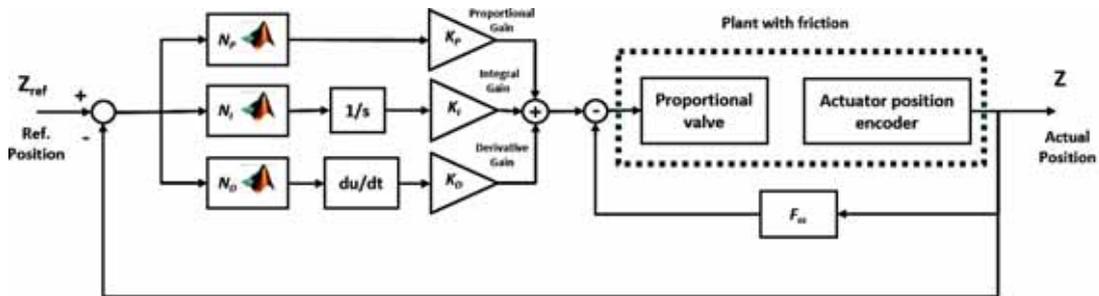
Fig. 13: The basic schematic in experimental mode T-NPID+F_{SS} and servo pneumatic plant.

2.8 Stability of the Triple Hyperbolic T-NPID Controller

The Popov stability criterion application has been described in detail by [38]. The Popov plot for P, I, and D components are shown in Figs. 14, 15, and 16, respectively. Both the P and D controller Popov plot does not cross the real axis; therefore, the maximum allowable gain K is $(0, \infty)$. The real axis is the y-axis (imaginary axis) at 0, as shown in Fig 14. The closed loop nonlinear gain is always stable for the P and D components. The Popov plot crosses the real axis at (-34.1202,0) for the I component. To obtain the maximum allowable gain, Eq. (17) is applied; therefore, the $K_I(e)$ value must be between 0 and 0.0293 $(0 < K_I(e) < 0.0293)$.

$$K(e_{Max}) = -\frac{1}{\mathcal{R}e\,\omega(j\omega_0)} \tag{17}$$

From Eq. (8) to Eq. (10), the value of K_P (e_{Max}), K_I (e_{Max}), and K_D (e_{Max}) are obtained as 2.4616, 0.028, and 1.1683, respectively, based on the range of allowable nonlinear gain

and the best steady-state position obtained. These values are obtained after the determination of the hyperbolic algorithm parameters.
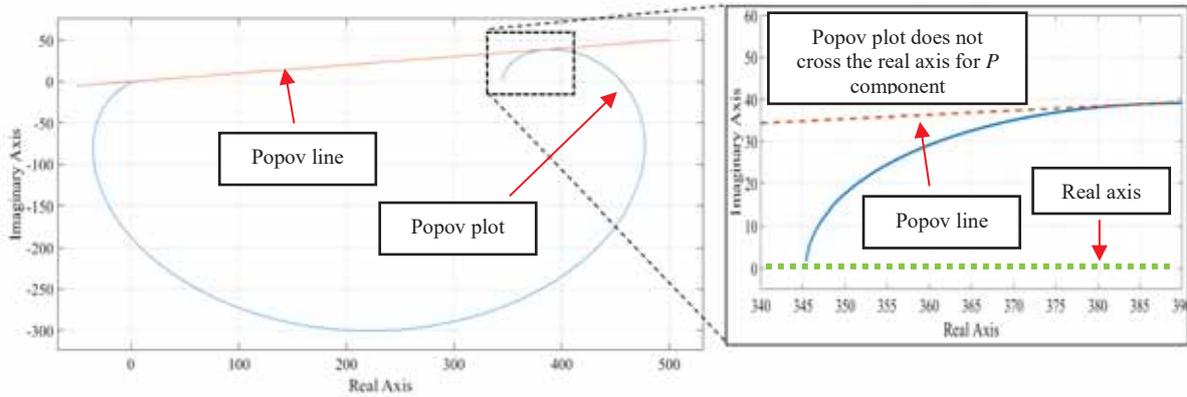
Fig. 14: The P component Popov plot.
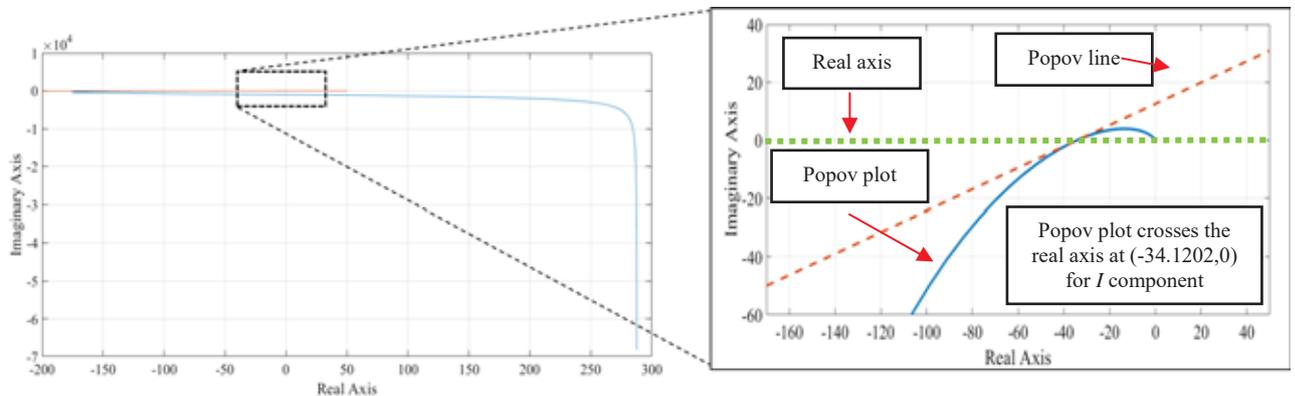
Fig. 15: The I component Popov plot.
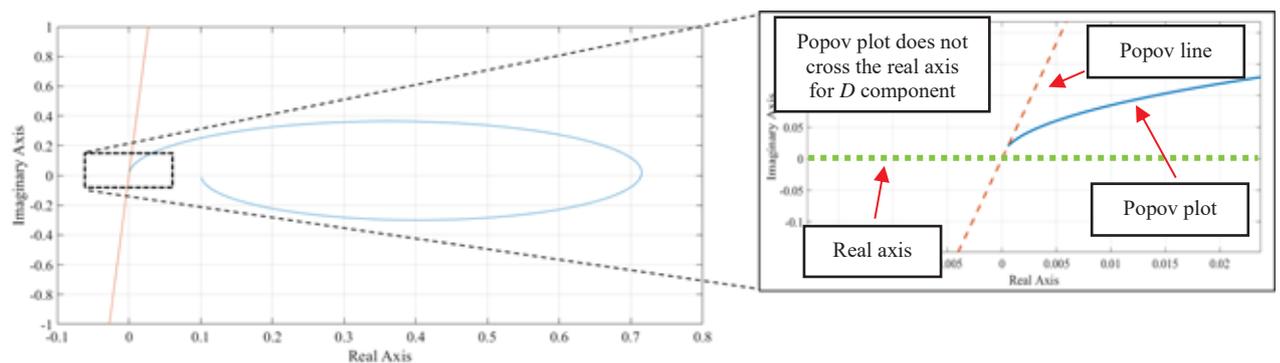
Fig. 16: The D component Popov plot.

Each method for designing the compensators has been presented in this section. Altogether three compensators were analyzed and compared to the typical PID controller. The first compensator is PID with the added friction force compensator, PID+F_{ss}. By adding F_{SS} only, the performance gains will be able to be analyzed and discussed. The next compensator is the T-NPID. T-NPID is anticipated to improve significantly even without friction compensation due to its system error-handling characteristic. The final compensator is T-NPID+F_{SS}. Combining a nonlinear PID with friction compensation, the system's performance will excel in both steady state and transient response analysis.

3   RESULTS AND DISCUSSION

3.1 Simulation Analysis

A step input is given to the compensated system. For the first simulation, a value of 1, equivalent to 1 volt, is set as reference input. Figure 17 shows the transient response and the steady-state phase of all the compensators analyzed. Figure 18 presents a closeup of the transient response from Fig. 14. The steady-state performance of the controllers is tabulated in Table 6, and the transient response performance is tabulated in Table 7.

Fig. 17: Simulation of system response of all controllers.

Table 6: Simulation results of the steady-state performance comparison

Steady-state parameter	PID	PID+F_{SS}	T-NPID	T-NPID+F_{SS}
Desired output (volts)	1	1	1	1
Actual output (volts)	0.987	0.987	0.993	0.993
Steady-state error, SSE (volts)	0.013	0.013	0.007	0.007
% Steady-state error, %SSE (%)	1.3	1.3	0.7	0.7

For the data tabulated in Table 6, the T-NPID without and with the F_{SS} presents better steady performance than base PID controllers. Although the value of the actual output of the controllers with friction compensators compared to the controllers without the friction compensators is the same, the close-up of the steady-state condition shows that the controllers with the friction compensators can compensate for the small oscillation due to the estimated friction. This oscillation due to friction is also known as the position-hunting phenomenon [33].

As tabulated in Table 7, the transient response data presents no significant compromise due to the integration of F_{SS} to both PID and T-NPID. T-NPID shows a significant improvement of the transient responses over the classical PID, especially to the Maximum overshoot, C_{MAX}, corresponding to the reduction and improvement of Percent overshoot, % OS. The results parallel with the nonlinear PID controller's response as presented by Salim et al. [22] and Jamian et al. [23], where the overshoot is drastically reduced.
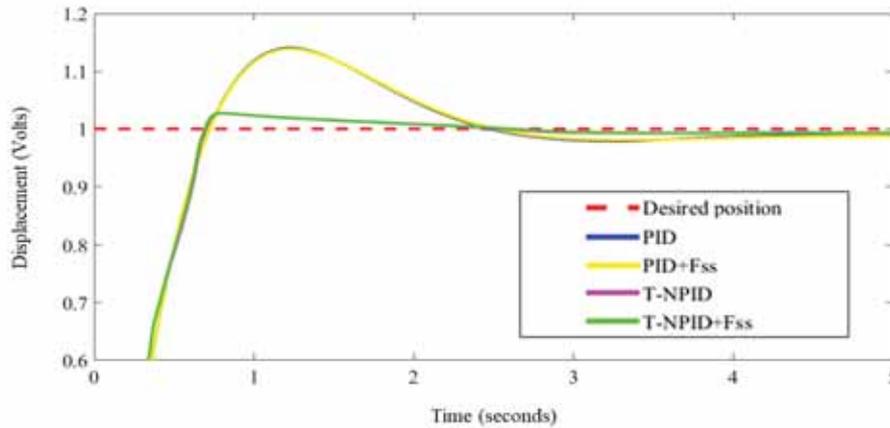
Fig. 18: The close-up transient response of each controller in simulation.

Table 7: Simulation results of the transient response performance comparison

Transient response parameter	PID	PID+F_{SS}	T-NPID	T-NPID+F_{SS}
Rise time, T_R (seconds)	0.49	0.49	0.53	0.53
Percent overshoot, %OS (%)	14.00	13.90	2.7	2.7
Maximum overshoot, C_{MAX} (volts)	1.140	1.139	1.027	1.027
Peak time, T_P (seconds)	1.23	1.23	0.8	0.8
Settling time, T_S (seconds)	2.35	2.36	1.621	1.653

3.2 Experimental Analysis

In the experimental analysis, the input given to the system is 100 mm, compared to 1 volt in the simulation analysis. For the Enfield actuator and encoder, a 1-volt conversion to mm is 30.48 mm (1 inch) of displacement, as mentioned in the product specification sheet. A displacement of 100 mm is given as the input rather than 30.48 mm as the unit is significantly presentable. Figure 19 presents the overall transient and steady-state response for all the controllers. Figure 20 shows a closeup of the controller's transient response. The data obtained is noticeably "noisy" compared to the simulation figures due to encoders or position sensors. Visually observed, the PID controller performs the least during transient response as the overshoot is the highest. Other controllers can control and reduce the overshoot during transient response. Besides PID, other controllers were also able to maintain the final position at the steady state, as the closeup Fig. 20 shows. PID controller offsets significantly at the steady state compared to other controllers.

Table 8 tabulates the steady-state performance of the controllers. T-NPID+F_{SS} presents the best performance of precise positioning at 0.3% SSE. A slight increase in the steady-state error is seen for the T-NPID. In the author's view, this is on par with T-NPID+F_{SS}. The small performance gain from T-NPID of 0.74% SSE to 0.30% is partially due to the noise measurement in the error sensor, especially at the time of steady-state (*t=50 seconds*). An additional performance measure, such as the IAE index performance, will present a complete conclusion of the controller's performance as all transient response, steady-state, and position hunting effect is calculated at once. The PID+F_{SS} controller alone can increase the precise positioning performance of the controller (from 2.60% SSE to 1.04% SSE). Figure 20 shows the close-up transient response performance of each controller. PID controller presents the highest overshoot, which means the controller performs the least.

PID+F_{SS} is not expected to improve the transient response as observed during simulation; however, in the experimental stage, PID+F_{SS} is able to slightly improve on the transient response, bringing the performance closer to T-NPID and T-NPID+F_{SS}. The transient response performance is tabulated in Table 9.
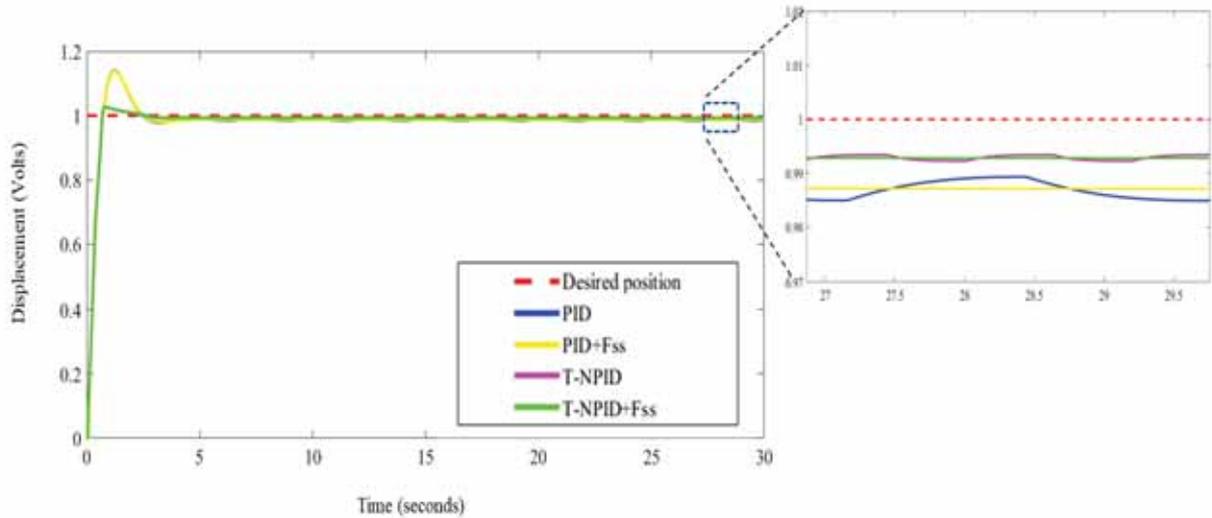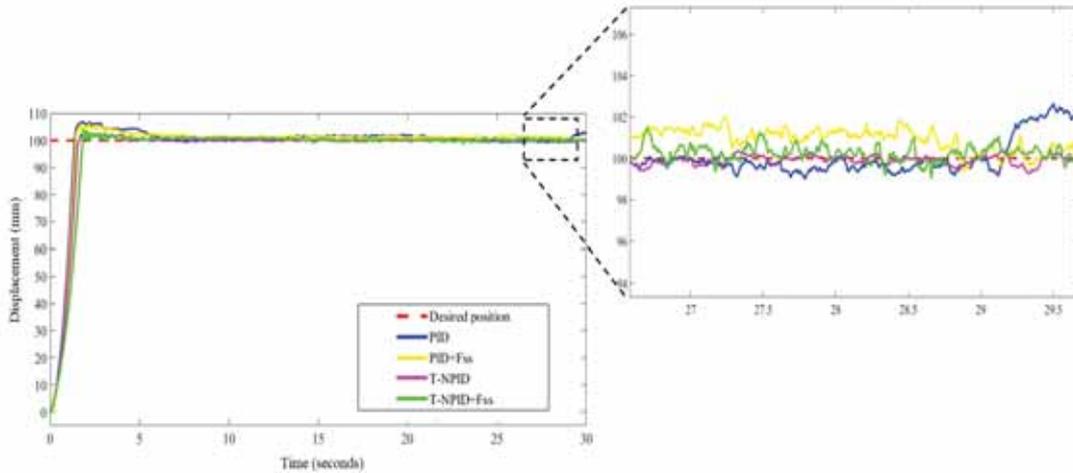

Fig. 19: Experimental of system response of all controllers.

Table 8: Experimental results of the steady-state performance comparison

Steady-state parameter	PID	PID+F_{SS}	T-NPID	T-NPID+F_{SS}
Desired output (mm)	100	100	100	100
Actual output (mm)	102.60	98.96	99.26	100.30
Steady-state error, SSE (mm)	2.60	1.04	0.74	0.30
% Steady-state error, %SSE (%)	2.60	1.04	0.74	0.30
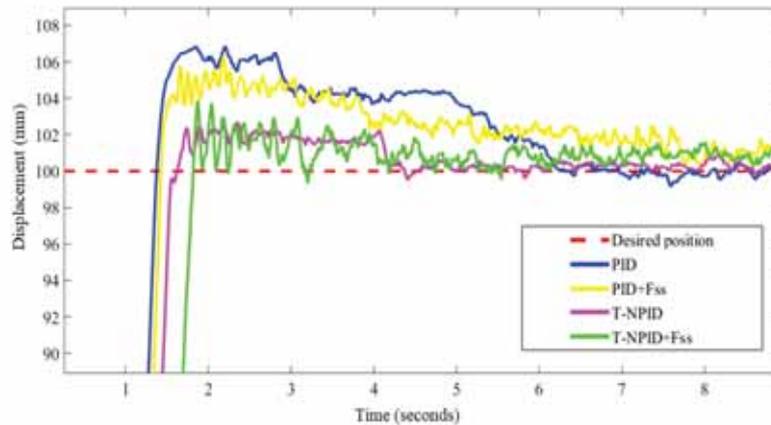
Fig. 20: Transient response experimental.

Figure 21 shows the single response of the PID controller. There is a noticeable position hunting at the desired position line. The phenomenon is in line with the previous research that internal friction will manage to cause the effect of position hunting in a system.

Figure 22 shows the PID+F_{SS}. The noticeable position hunting has been minimized. This improvement is validated by the measurement of IAE, where a significant reduction in the IAE for PID+F_{SS} compared to PID as tabulated in Table 10. The improvement is from 128 (volt^2) to 113.7 (volt^2), a reduction of 11.17%.
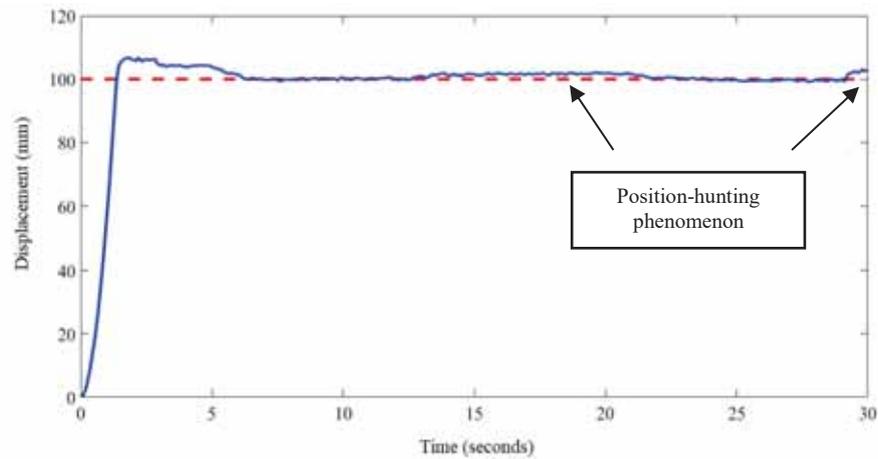
Fig. 21: The PID controller response.
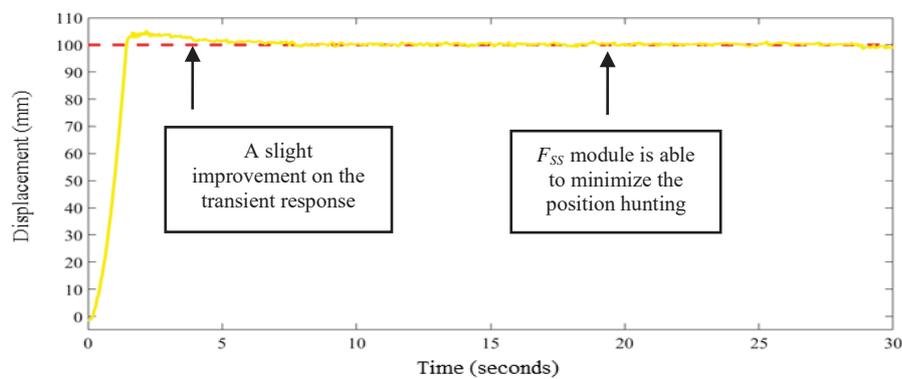
Fig. 22: The PID+F_{SS} controller response.

Figure 23 shows the T-NPID+F_{SS} system response. The controller and module can maintain the precise positioning performance and reduce the overshoot, reflecting improvement in the transient response.
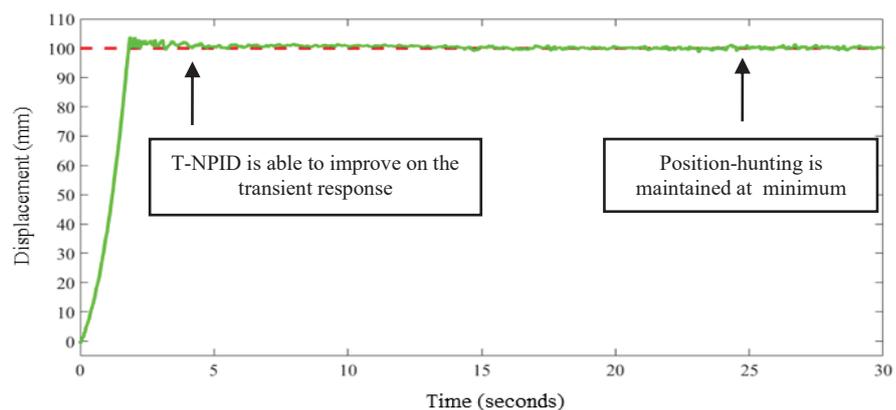
Fig. 23: The T-NPID+F_{SS} controller response.

Most previous research in the literature focuses on friction compensation to sinusoidal tracking performance. The closest to a step response analysis is a multi-steps positioning tracking performance by Rahman et al. [21]. The author analyzed both sinewave trajectories and multi-steps polynomial trajectory. The multi-steps reference input is trapezoidal shapes

multi-steps, as shown in Fig. 24. Therefore, the performance measures were RMSE and mean steady-state errors (SSE). Transient response analyses were not presented. However, the comparison of controllers with and without the LuGre friction compensator was presented, where the proposed DAB-SMC controller improved by 1.77 mm in mean SSE with the friction observer active (in terms of percentage improvement, the numbers were unclear). The performance of controllers without friction observer is shown in Fig. 25. The SSE results without the observer are visible.
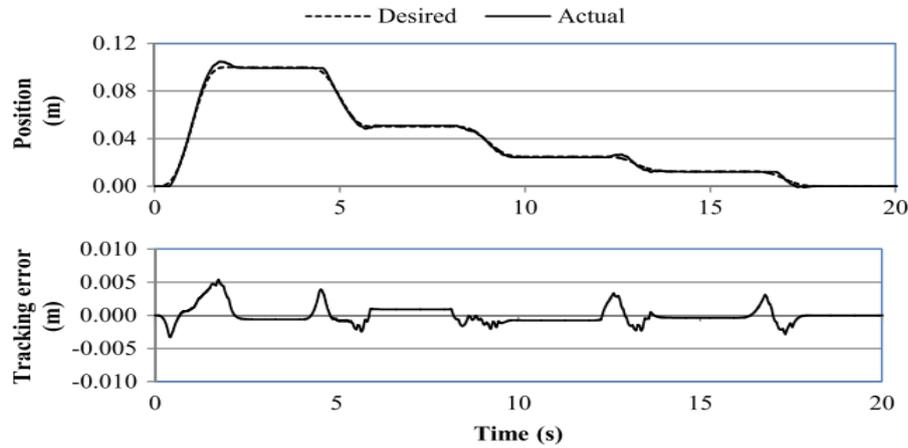
Fig. 24: Multi-steps trajectory performance analysis for DAB-SMC with LuGre friction observer [21].
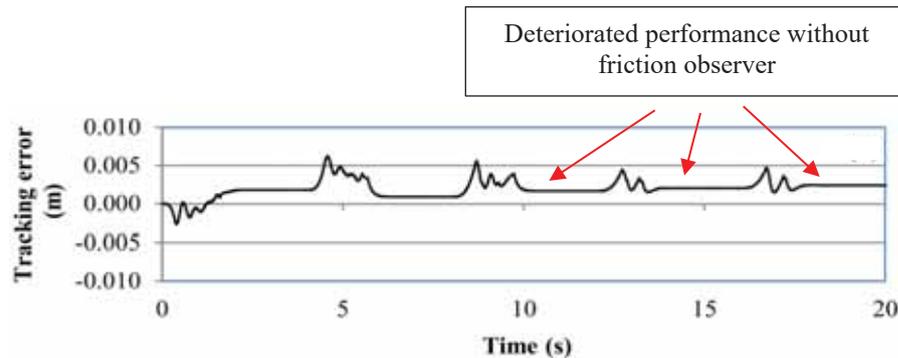
Fig. 25: Tracking error for DAB-SMC without friction observer [21].

Table 9: Experimental results of the transient response performance comparison

Transient response parameter	PID	PID+F_{SS}	T-NPID	T-NPID+F_{SS}
Rise time, T_R (seconds)	0.895	0.943	1.068	1.171
Percent overshoot, %OS (%)	6.80	4.9	2.70	3.30
Maximum overshoot, C_{MAX} (mm)	106.8	104.9	102.7	103.3
Peak time, T_P (seconds)	1.84	2.18	2.35	1.58
Settling time, T_S (seconds)	5.573	5.215	2.377	3.253

Table 9 tabulates the transient response of each controller. The controller with F_{SS}, PID+F_{SS}, and T-NPID+F_{SS} does not necessarily improve overshoot, respectively, as shown for the T-NPID+F_{SS} controller. PID+F_{SS} is observed to slightly improve on the overshoot (from 6.8% to 4.9%, an improvement of 1.9%), whereas for T-NPID, the overshoot slightly increases (from 2.7 % to 3.3%). The 0.6% decrease is again due to the noise of the sensor signal. As shown in Fig. 26, comparing T-NPID and T-NPID+F_{SS}, the overshoots are

comparably on par. There is a slight suppression during overshoot for T-NPID+F_{SS} due to the friction module compensating. The settling time, T_S, for both T-NPIDs is significantly improved over PID by around 40 to 60%.
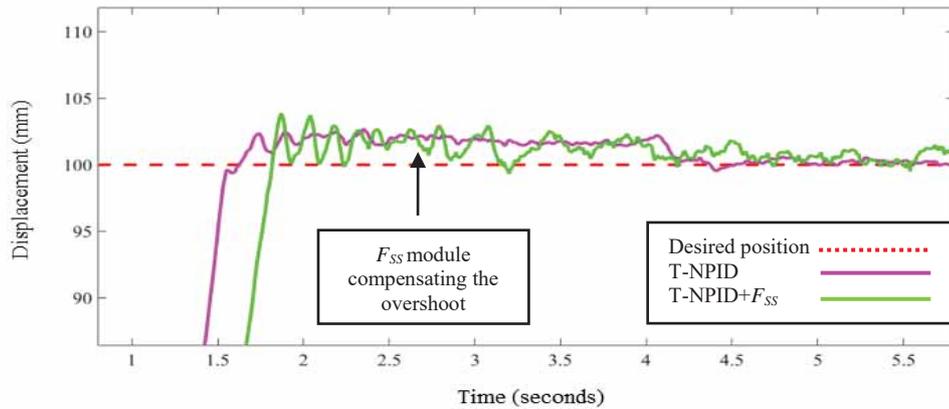
Fig. 26: The T-NPID and T-NPID+F_{SS} overshoot comparison.

Table 10 tabulates the IAE performance index of each controller in the experimental condition based on Eq. (18). The IAE for each component of the compensator adds improvement. PID+F_{SS} reduces the IAE from 128 to 113.7 (mm). The final T-NPID+F_{SS} controller produces the least IAE, which shows this controller produces the best transient response and the steady-state position. Figure 27 shows the IAE performance index in graphical form.

$$IAE = \int_0^\infty |r(t) - y(t)| dt \qquad (18)$$

where r(t) and y(t) are desired and actual positions, respectively.

Table 10: IAE performance comparison

Performance measure	PID	PID+F_{SS}	T-NPID	T-NPID+F_{SS}
IAE (mm)	128	113.7	108.2	104.9

Table 11 tabulates the percentage of increased performance of important parameters in steady-state and transient response, compared to the base PID controller. The percentage of improvement of the parameters is calculated based on Eq. (19) [41]. A higher percentage denotes a better improvement of the parameter.

$$Percent\ Improved\ Value\ Reduction = \left[\frac{Improved\ value - Benchmark\ value}{Benchmark\ value}\right] \times 100\% \qquad (19)$$

Table 11: Percentage of reduction of parameter performance of each controller

Performance parameter	PID+F_{SS} with PID	T-NPID with PID	T-NPID+F_{SS} with PID
Steady-state error, SSE (%)	60	70.77	88.46
IAE (%)	11.17	15.47	18.05
Percent overshoot, %OS (%)	29.94	60.29	51.47
Settling time, T_S (%)	6.42	57.35	41.63
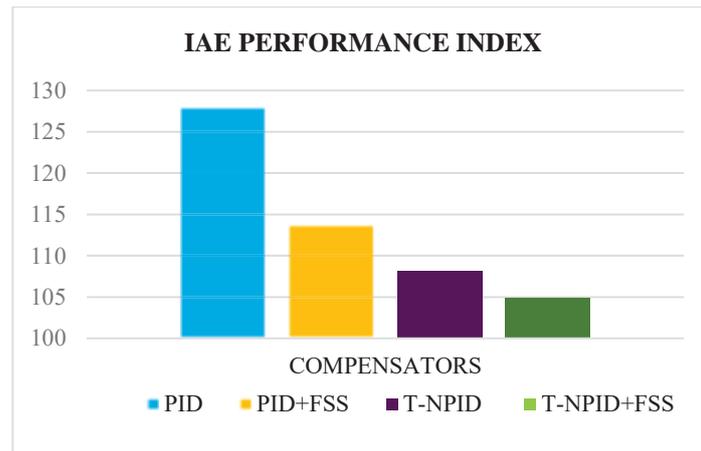
Fig. 27: IAE performance comparison experimental.

SSE recorded the best improvement among the parameters measured by the T-NPID+F_{SS} at 88.46% improvement. The performance follows T-NPID (70.77%) and PID+F_{SS} (60%). PID+F_{SS} alone is respectable as a high percentage of improvement is seen by integrating the F_{SS} module. IAE index performance also presents an improvement of 18.05% by the final controller, overcoming the other controllers. The transient response, such as %OS and T_S, T-NPID, slightly overcomes the T-NPID+F_{SS}. As explained in the previous discussion, this is highly due to the noise in the measurement sensor during measurement (point-to-point measurement), but the performance is on-par. Visually analyzed, these two transient response parameters are on par for T-NPID and T-NPID+F_{SS}. IAE validates that overall performance as TNPID+F_{SS} produced the least error and therefore is the best controller in this case.

Comparing the analysis to the previous research, as stated earlier, the included literature primarily analyzes sinewave tracking trajectories with different amplitudes and frequencies. In this research, step positioning performance is analyzed. Friction compensation is also significant in step positioning on top of trajectories. As presented by Soleymani et al. [19], Tran et al. [20], and Rahman et al. [21], the RMSE errors for sinewave tracking were reduced from 15% to up to 50%. Figure 28 shows the tracking error of a sinewave trajectory comparison between two different friction models applied where GMS model performs better than Stribeck model. This research and the previous literature prove that friction compensation is highly important and is able to add performance to any primary controller. Even though the static friction model is relatively simple compared to the dynamic models, the model is significant to achieve better precise positioning of the system.
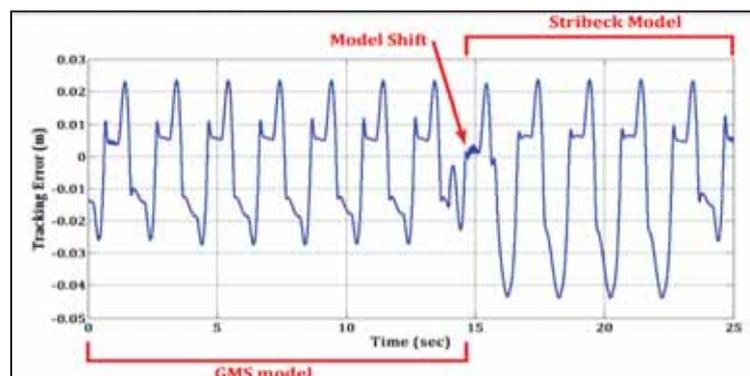
Fig. 28: Tracking error analysis performed by Soleymani et al. [19].

4. CONCLUSION

A proposed static friction compensator module has been designed, applied to several proposed controllers, and analyzed to achieve the best actual position versus the desired position for a pneumatic actuator. The friction was successfully identified, modeled, and compensated by applying the static friction model. Combining the static friction model compensation with the T-NPID, T-NPID+F_{SS} has been proven to obtain an exceptional precise positioning for the pneumatic system. The precise positioning performance has been improved by 88.46% to the reference controller. A sinewave tracking performance will also need to be analyzed, as this analysis will conclude the performance of this research to the previous related research. A tracking differentiator algorithm is desirable to be integrated into a feedforward configuration and has been proven to improve servo pneumatic trajectory performance [42]. Further enhancing this research, the friction model can be identified by adopting AI optimization tools, such as Genetic Algorithm (GA) and Neural Networks (NN) [43-44]. By adopting the tools, the model's accuracy in predicting actual friction is also improved. A higher accuracy friction model will be able to cater to better friction compensation, resulting in a higher precise positioning system.

ACKNOWLEDGMENT

REFERENCES

[1] Hildebrandt A, Neumann R, Sawodny O. (2010) Optimal system design of SISO-servopneumatic positioning drives. IEEE Transaction on Control Systems Technology, 18(1): 35-44.

[2] Noor SBM, Ali HI, Marhaban MH. (2011) Design of combined robust controller for a pneumatic servo actuator system with uncertainty. Scientific Research and Essays, 6(4): 949-965 3.

[3] Lu CH, Hwang YR. (2012) A model reference robust multiplesurfaces design for tracking control of radial pneumatic motion systems. Nonlinear Dynamics, 67(4): 2585-2597.

[4] Salim SN, Rahmat MF, Faudzi AAM, Ismail ZH, Sunar N. (2014) Position control of pneumatic actuator using self-regulation nonlinear PID. Mathematical Problems in Engineering, 2014.

[5] Beater P. (2007) Pneumatic Drives. Springer-Verlag Berlin Heidelberg.

[6] Faisal WT, Albagul A. (2004) Anti-windup scheme for practical control of positioning systems. IIUM Engineering Journal, 5(2): 1-15.

[7] Andrighetto PL, Valdiero AC, Carlotto L. (2006) Study of the friction behavior in industrial pneumatic actuators. In ABCM Symposium series in mechatronics, 2: 369-376.

[8] Saravanakumar D, Mohan B, Muthuramalingam T. (2017) A review on recent research trends in servo pneumatic positioning systems. Precision Engineering, 49 :481-492.

[9] Johnson MA, Moradi MH. (2005) PID control. London, UK, Springer-Verlag London Limited.

[10] Vinagre BM, Monje CA, Calderón AJ, Suárez JI. (2007) Fractional PID controllers for industry application. A brief introduction. Journal of Vibration and Control, 13(9-10): 1419-1429.

[11] Zainal A, Wahab NA, Yusof MI. (2023) PLC-based PID controller for real-time pH neutralization process using Palm Oil Mill Effluent. IIUM Engineering Journal, 24(1): 244-255.

[12] Priyanka EB, Maheswari C, Ponnibala M, Thangavel S. (2019) SCADA based remote monitoring and control of pressure & flow in fluid transport system using IMC-PID controller. Advances in Systems Science and Applications, 19(3): 140-162.

[13] Ashmi M, Anila M, Sivanandan KS. (2021) Comparison of SMC and PID controllers for pneumatically powered knee orthosis. Journal of Control, Automation and Electrical Systems, 32(5): 1153-1163.

[14] Jouppila VT, Gadsden SA, Bone GM, Ellman AU, Habibi SR. (2014) Sliding mode control of a pneumatic muscle actuator system with a PWM strategy. International Journal of Fluid Power, 15(1): 19-31.

[15] Richer E, Hurmuzlu Y. (2000) A high performance pneumatic force actuator system: part II nonlinear controller design. Journal of Dynamic Systems, Measurement and Control, 122(3): 426-34.

[16] Zhu Y, Barth EJ. (2010) Accurate sub-millimeter servo-pneumatic tracking using model reference adaptive control (MRAC). International Journal of Fluid Power, 11(2): 43-55.

[17] Takosoglu JE, Laski PA, Blasiak S. (2012) A fuzzy logic controller for the positioning control of an electro-pneumatic servo-drive. Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, 226(10): 1335-1343.

[18] Ren HP, Jiao SS, Wang X, Li J. (2020) Adaptive RBF Neural Network Control Method for Pneumatic Position Servo System. IFAC-PapersOnLine, 53(2): 8826-8831.

[19] Soleymani F, Rezaei SM, Sharifi S, Zareinejad M. (2016) Position control of a servo-pneumatic actuator using generalized Maxwell-Slip friction model. In 2016 4th International Conference on Robotics and Mechatronics (ICROM): 26 Oct 2016; Edited by IEEE, pp 246-251.

[20] Tran XB, Nguyen VL, Nguyen NC, Pham DT, Phan VL. (2019) Sliding mode control for a pneumatic servo system with friction compensation. In Advances in Engineering Research and Application: Proceedings of the International Conference on Engineering Research and Applications, ICERA 2019: 1 Dec 2019; Edited by Springer International Publishing, pp 648-656.

[21] Rahman RA, He L, Sepehri N. (2016) Design and experimental study of a dynamical adaptive backstepping–sliding mode control scheme for position tracking and regulating of a low-cost pneumatic cylinder. International Journal of Robust and Nonlinear Control, 26(4): 853-875.

[22] Salim SN, Amran AC, Faudzi AAM, Ismail ZH, Rahmat MF, Sunar NH, Shamsudin SA. (2015). A study on tracking performance of the pneumatic system with enhanced NPID controller. In 2015 10th Asian control conference (ASCC): 31 May 2015; Edited by IEEE, pp 1-6.

[23] Jamian S, Salim SN, Junoh SCK, Kamarudin MN, Abdullah L. (2020) Nonlinear Proportional Integral (NPI) Double Hyperbolic Controller for Pneumatic Actuator System. In Advances in Electronics Engineering: Proceedings of the ICCEE 2019: 29-30 April 2019; Kuala Lumpur, Malaysia. Edited by Springer Singapore, pp 221–229.

[24] Junoh SCK, Salim SN, Abdullah L, Anang NA, Chiew TH, Retas Z. (2017) Nonlinear PID triple hyperbolic controller design for XY table ball-screw drive system. International Journal of Mechanical and Mechatronics Engineering, 17(3): 1-10.

[25] Kamaludin KN, Abdullah L, Salim SN, Jamaludin Z, Maslan MN, Rahmat MF. (2022) Comparison of a Double and Triple Nonlinear Hyperbolic Proportional-Integral-Derivative (PID) compensator for a servo pneumatic actuator. In 2022 13th Asian Control Conference (ASCC): 4 May 2022; Jeju, Korea, Edited by IEEE, pp 1775-1781.

[26] Abdullah L, Jamaludin Z, Chiew TH, Rafan NA, Mohamed MS. (2012) System identification of XY table ballscrew drive using parametric and non parametric frequency domain estimation via deterministic approach. Procedia Engineering. 41: 567-574.

[27] Dermawan D, Abbas H, Syam R, Djafar Z, and Muhammad AK. (2020) Dynamic modeling of a single-link flexible manipulator robot with translational and rotational motions. IIUM Engineering Journal. 21(1): 228-239.

[28] Rahmat MF, Sunar NH, Salim SN, Abidin MS, Fauzi AM, Ismail ZH. (2011) Review on modeling and controller design in pneumatic actuator control system. International Journal on Smart Sensing and Intelligent Systems. 4(4): 630-661.

[29] Jamian S, Salim SN, Kamarudin MN, Abdullah L, Hanafiah MA. (2018) Modeling of a single rod double acting pneumatic actuator system. In Proceedings of Symposium on Electrical, Mechatronics and Applied Science (SEMA ,18): 8 December 2018; Melaka, Malaysia. pp 45-46.

[30] Ljung L. (1995) System Identification Toolbox: User's Guide. Natick, MA, USA: MathWorks Incorporated.

[31] Sulaiman SF, Rahmat MF, Faudzi AA, Osman K. (2016) Linear and nonlinear ARX model for intelligent pneumatic actuator systems. Jurnal Teknologi. 78(6): 21-28.

[32] Chiew TH, Jamaludin Z, Hashim AB, Rafan NA, Abdullah L. (2013) Identification of friction models for precise positioning system in machine tools. Procedia Engineering. 53: 569-578.

[33] De Wit CC, Olsson H, Astrom KJ, Lischinsky P. (1995) A new model for control of systems with friction. IEEE Transactions on automatic control. 40(3): 419-425.

[34] Meng D, Tao G, Chen J, Ban W. (2011) Modeling of a pneumatic system for high-accuracy position control. In Proceedings of 2011 International Conference on Fluid Power and Mechatronics: 17-20 August 2011; Beijing, China. pp 505-510.

[35] Salim SN. (2014) Modeling and Control Design of an Industrial Pneumatic Actuator System. PhD thesis, Universiti Teknologi Malaysia.

[36] Junoh SCK. (2019) NPID Double Hyperbolic Controller for Improving Tracking Performance of XY Table Ballscrew Drive System. MSc thesis, Universiti Teknikal Malaysia Melaka.

[37] Jamaludin Z. (2008) Disturbance Compensation for Machine Tools with Linear Motors. PhD thesis. Ketholieke Universiteit Leuven, Belgium.

[38] Seraji H. (1997) A new class of nonlinear PID controllers. In IFAC Proceedings Volumes 30.20: 1 September 1997; pp 65-71.

[39] Albatran S, Alatoum A, Al Khalaileh AR. (2021) Informative Order-Reduction of Underdamped Third-Order Systems. IEEE Access. 9: 88512-88523.

[40] Su YX, Sun D, Duan BY. (2005) Design of an enhanced nonlinear PID controller. Mechatronics. 15(8): 1005-1024.

[41] Abdullah L. (2014) A New Control Strategy for Cutting Force Disturbance Compensation for XY Table Ball Screw Driven System. PhD thesis. Universiti Teknikal Malaysia Melaka.

[42] Zhao L, Xia Y, Yang Y, Liu Z. (2017) Multi controller positioning strategy for a pneumatic servo system via pressure feedback. IEEE Transactions on Industrial Electronics. 64(6): 4800-4809.

[43] Carducci G, Giannoccaro NI, Messina A, Rollo G. (2006) Identification of viscous friction coefficients for a pneumatic system model using optimization methods. Mathematics and Computers in Simulation. 71(4-6): 385-94.

[44] Wang J, Wang JD, Daw N, Wu QH. (2004) Identification of pneumatic cylinder friction parameters using genetic algorithms. IEEE/ASME transactions on mechatronics. 9(1): 100-107.

APPLICATION OF DRIVING BEHAVIOR CONTROL SYSTEM USING ARTIFICIAL NEURAL NETWORK TO IMPROVE DRIVING COMFORT BY ADJUSTING AIR-TO-FUEL RATIO

ARIS TRIWIYATNO[1]*, **SUROTO MUNAHAR[2,3], M. MUNADI[3]**
AND JOGA DHARMA SETIAWAN[3]

[1]*Department of Electronics Engineering, Diponegoro University, Semarang, Indonesia*
[2]*Department of Automotive Engineering,*
Universitas Muhammadiyah Magelang, Magelang, Indonesia
[3]*Department of Mechanical Engineering, Diponegoro University,*
Semarang, Indonesia

**Corresponding author: aristriwiyatno@elektro.undip.ac.id*

ABSTRACT: Energy-efficient engines were introduced due to limited amount of global energy and the need for engine power to carry vehicle loads. It was discovered that the power factor of these engines was essential in developing automotive technology with subsequent significant effect on driving comfort. Moreover, it was possible to control the power and energy savings of vehicle engines by adjusting the Air to Fuel Ratio (AFR). Therefore, this study focused on achieving AFR values in the stoichiometric range of 14.7 in order to produce good emissions. The technology applied was observed to have some drawbacks, specifically in fulfilling engine power when the vehicle operates with a large load. This led to the development of a new method by designing an AFR control system with due consideration for driving behavior using an Artificial Neural Network (ANN). The aim was to overcome the problem of meeting engine power and ensuring better efficiency. The driving behavior was classified into through categories including the sporty, standard, and eco schemes. The eco scheme was the smooth behavior of a driver during the movement of the vehicle in a busy urban area, the sporty scheme was the responsive driving behavior when the vehicle operates on the highway at speeds above 80 km/h, and the standard scheme was the behavior between the eco and sporty schemes. Furthermore, the driving behavior in a sporty scheme required the addition of fuel to increase engine power while eco-scheme focused on reducing fuel to increase fuel economy. The findings showed that control system designed was able to improve driving comfort in terms of fuel economy during the eco scheme with an average AFR value of 15.68. The system further reduced the value to 13.66 during the sporty scheme. Furthermore, the AFR under stoichiometry was discovered to have produced the maximum engine power. The system was expected to be incorporated into electric, gas-fired and fuel cell vehicles in the future.

ABSTRAK: Faktor kuasa enjin dan enjin cekap tenaga adalah penting dalam membangunkan teknologi automotif. Mesin penjimat tenaga diperlukan kerana jumlah tenaga global yang terhad. Manakala kuasa enjin digunakan bagi membawa muatan kenderaan. Kedua-dua faktor ini sangat mempengaruhi keselesaan pemanduan. Penjimatan kuasa dan tenaga dalam enjin kenderaan boleh dipenuhi dengan mengawal Nisbah Angin kepada Minyak (AFR). Tumpuan kajian semasa adalah berorientasikan ke arah mencapai nilai AFR dalam julat stoikiometri (14.7) atas sebab ingin mencapai pelepasan terbaik. Namun begitu, teknologi ini mempunyai kelemahan terutama dalam

memenuhi kuasa enjin apabila kenderaan beroperasi dengan muatan besar. Oleh itu, kajian ini adalah berkaitan kaedah baharu bagi mengatasi masalah memenuhi kuasa enjin dan mencapai enjin cekap tenaga dengan mereka bentuk sistem kawalan AFR yang mempertimbangkan tingkah laku pemanduan menggunakan Rangkaian Neural Buatan (ANN). Tingkah laku pemanduan direka bentuk kepada tiga skim: sporty, standard dan eko. Skim eko adalah kelancaran tingkah laku pemandu apabila kenderaan bergerak di kawasan bandar yang sibuk. Skim sporty ialah tingkah laku pemanduan responsif apabila kenderaan beroperasi di lebuh raya pada kelajuan melebihi 80 km/j, dan skema standard ialah tingkah laku antara skim eko dan sporty. Tingkah laku pemanduan dalam skema sporty memerlukan penambahan bahan api bagi meningkatkan kuasa enjin. Sementara itu, tingkah laku pemanduan dalam skim eko memerlukan pengurangan bahan api bagi meningkatkan penjimatan bahan api. Hasil kajian menyatakan sistem kawalan yang direka mampu meningkatkan keselesaan pemanduan dari segi penjimatan bahan api apabila tingkah laku pemandu memasuki skim eko. AFR dicapai pada nilai purata 15.68. Apabila tingkah laku pemandu memasuki skim pemanduan sporty, sistem kawalan boleh mengurangkan AFR dengan nilai purata 13.66. AFR di bawah stoikiometri menghasilkan kuasa enjin maksimum. Pada masa hadapan, sistem ini berpotensi untuk dibangunkan pada kenderaan elektrik, menggunakan gas dan sel bahan api.

KEYWORDS: *driving behavior; AFR; ANN; engine power; fuel saving*

1. INTRODUCTION

Vehicle comfort is a very important topic in the development of automotive technology and it can be increased through suspension control [1], steering system power transfer, engine power, and fuel economy. It is also significantly related to the aspects of engine power and fuel economy, and these two concepts have become international issues presently focused on by world researchers [2] due to the limited energy problem [3,4] and the very significant increase in global energy consumption.

The current high increase in the number of vehicles using gasoline engines has made fuel-saving technology necessary. Meanwhile, there is also the need to have the appropriate engine power to move the vehicle, especially when it is carrying a heavy load. It is important to note that vehicle comfort can be achieved from the engine power and fuel economy aspects by controlling the Air to Fuel Ratio (AFR) which is categorized into three including lean mixture with AFR values above 14.7, ideal or stoichiometry ratio which is 14.7, and rich mixture with values below 14.7. The lean category usually produces low engine power with high fuel economy, the rich mixture is the opposite, and the stoichiometry category normally has optimal engine power and fuel economy.

Several studies have been conducted on AFR using different variables such as the application of fuzzy logic which is Artificial Intelligence (AI) to set AFR in a simulation [5]. It was discovered that the system was able to control the AFR with inaccurate signals or data but it has not been applied to real conduction. Another different study tried to add a variable turbo charge to increase the volumetric efficiency of the engine and changes were observed in the AFR because the turbo charge added forced air into the engine based on the power generated by the exhaust gas pressure [6]. However, the research has not yet integrated an intelligent control system to modify the AFR. Some studies have also started focusing on AFR stoichiometry control which was denoted by the achievement of lambda (λ) = 1 [7]. It is important to note that Lambda (λ) is a comparative value between the actual AFR and the stoichiometry of the theoretical AFR. This method also has some problems which are making the vehicle users complain, especially due to the low engine power when the vehicle is operating on a steep or inclined surface or carrying a large load.

Driving behavior has been empirically discovered to have a great influence on the fuel system (AFR) dynamics but it was discovered that no study has focused on controlling AFR through driving behavior. Previous studies have only focused on using several variables such as the development of a car using systems in autonomous vehicles, application of a simulation scale Connected Vehicle Driving Strategy with Intelligent Model (CVDS-IDM), design of safety systems on vehicles, analysis of driving behavior, vision technology or graphical display, and the vehicle emissions dynamics when at a traffic light.

Driving behavior in car-following has also been studied with several variations such as autonomous vehicle dynamics, the use of CVDS-IDM, and a safe car-following behavior system. Those related to the dynamics of the autonomous vehicle showed that the driver following a driverless autonomous vehicle was significantly affected psychologically [8]. Moreover, CVDS-IDM was defined as a simulator tool to predict driving behavior in real time by providing a driver a stimulus with various variables, especially when driving behind another vehicle [9]. It was discovered that the CVDS-IDM was able to model the car-following mechanism through the method applied by the driver. Another study also focused on safe driving behavior in order to develop a safety system based on the behavior of car-following drivers [10]. However, it was discovered that none of these three studies discuss the relationship between driving behavior and fuel consumption.

Further studies focused on driving behavior in relation to the safety systems design on vehicles such as the usage of cameras, vehicle theft prevention, and adaptive Forward Vehicle Collision Warning (FCW) systems [11] as well as the analysis of the driving behavior when the driver was sleepy [12]. The cameras were installed on the vehicle to record ongoing activities, hence it provides feedback for a sleepy driver in order to control the vehicle immediately and properly with the intention of reducing the occurrence of accidents. Martinelli [13] also prevented vehicle theft by comparing driving behavior against a database. The research was initiated by learning and storing the driver's behavior in a control system and was later compared with the pattern through which an individual drive. The engine of the vehicle was programmed to stop when the driving behavior stored and observed are not similar to protect the vehicle from theft. Another study by Yuan [14] used an adaptive Forward Vehicle Collision Warning (FCW) to avoid collisions when the driver is driving abnormally. This was achieved by detecting the distance between the vehicle and the monocular. It is important to note that all these studies did not also consider the relationship between fuel usage systems and driving behavior.

Scholars also researched the improvement of driving safety through several variables which include analyzing the attitude and performance of drivers during sudden braking or steering operations and determining their safety rating [15]. Another study also observed and assessed driver control and safe practices during the process of driving a vehicle using risky driving behavior as the research variable [16]. Moreover, Mafeni [17] and Takashi Bando [18] evaluated driving behavior errors with a focus on abnormal braking and vehicle speed operation when traveling at very high speeds. It was also observed that even though the studies have different objectives, none focused on energy consumption in evaluating driving behavior.

Subsequent research also focuses on driving behavior in autonomous and electric vehicles (EVs). This was observed from the design of a lateral system control based on autonomous vehicles by Hongbo [19] to trace the vehicle's turning angle accurately and provide a better steering control effect compared to conventional steering lateral control. Moreover, Ashkrof [20] developed EVs control strategies to select travel routes and determine battery charging locations based on driving behavior. These studies do not also discuss variables related to energy use.

An automatic control system was introduced to recognize driving behavior by Silver [21] and Yansong [22] to alert drivers when they exceed safety limits. Julian [23] also used vision sensors for the same purpose with the recognition results stored in a database and exclusively patented. These studies have interesting concepts but this could have been more exciting if the driving behavior recognition factor was used to control energy. Unfortunately, it was not included.

The other aspects studied in relation to driving behavior are traffic lights and graphical displays. This was observed in the analysis of emissions produced by vehicles at traffic lights, especially at road junctions, by Stagios [24] through the observations made about the changes in vehicle emissions on the highway due to driving behavior. It was discovered that it is possible to reduce emissions by up to 26% when the driver behaves in certain conditions. However, this study only focused on changes in emissions without considering the energy consumption factor in behavioral changes.

Kohl [25] also showed that graphic displays, especially images/messages that frequently appear, have a significant influence on safe driving behavior. It was discovered that the frequent graphic display reduces the focus of drivers while driving and this is very dangerous for safety. This is considered very interesting but this study did not include the energy control influenced by driving behavior. This was followed by studies conducted on the relationship between energy consumption and driving behavior. Xing [26] predicted future energy consumption by observing the behavior of the driver but the study does not lead to energy control.

This simply implies several variables have been considered in studying driving behavior such as car-following systems, CVDS-IDM simulations, vehicle safety systems design, driving behavior analysis, vision technology or graphical display, and the vehicle emissions dynamics at traffic lights. It has also been discovered from previous studies that fuel energy control can be denoted by AFR used to generate stoichiometry values, machine learning or artificial intelligence applications, and the use of turbochargers. However, research has not been conducted on using driving behavior to control fuel even though driving behavior has been empirically proven to have a significant influence on fuel consumption. Therefore, this research was used to discuss the application of a driving behavior system to control fuel using an Artificial Neural Network as a continuation of the previous study [26].

Description:
a. Steering speed sensor.
b. Braking speed sensor.
c. Throttle valve speed sensor.
d. Engine.
e. ECU with driving behavior control system using ANN.
f. Fuel pump.
g. Fuel tank.
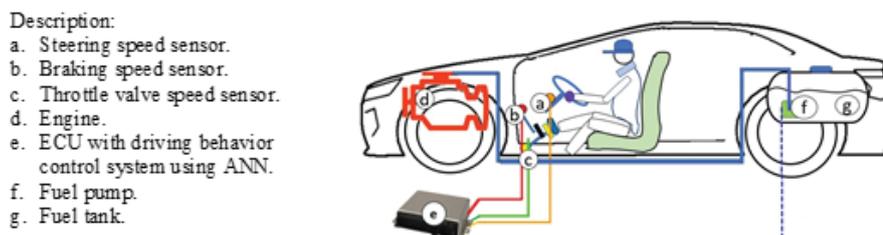
Fig. 1: The concept of the designed control system.

The control system developed based on the driving behavior using Artificial Neural Network (ANN) was successfully applied to real vehicles and presented in the following Figure 1. It was designed to work on three inputs which include the steering speed sensor (a), braking speed sensor (b), and throttle valve opening speed sensor (c) as well as supporting components to assist its performance which include an engine to generate power to drive the vehicle (d) and engine control unit – ECU (e) designed with ANN to recognize driving behavior in real time. The results obtained from the ANN were used to control fuel consumption through the fuel pump (f) placed in the fuel tank (g) in order to deliver fuel to the injectors. It is important to

note that the speed of the fuel pump was controlled by the ECU with the ANN applied to recognize driving behavior.

2. METHOD

2.1 Control System by Driving Behavior

The behaviors of drivers are usually different when driving based on their needs, mood, and inclinations, and this greatly influences vehicle operation and fuel consumption. This led to the design of three schemes to control the vehicle which include eco-driving, standard or stoichiometry, and sporty behaviors. The eco-driving is the behavior associated with vehicles operating smoothly on busy urban roads [27] and allows the driver to operate the throttle valve opening, braking, and steering smoothly at a speed of approximately 40-60 km/hour. The sporty scheme is a responsive driving behavior that involves the vehicle operating on a toll road [28] and the driver is expected to generally drive responsively at a vehicle speed of more or less than 80 km/hour. Meanwhile, the standard scheme is a driving behavior between the eco and sporty schemes.

Figure 2 shows the design of the control system developed in this study to regulate fuel using ANN with due consideration for the driving behavior. The ANN designed to be embedded in the control system is used to recognize driving behavior in real-time. Moreover, the algorithm method applied in the ANN was Levenberg Marquardt type learning (trainlm) while the learning performance was based on the Mean Squared Error (MSE) value using 300 nodes, one output layer, one hidden layer, and three inputs.
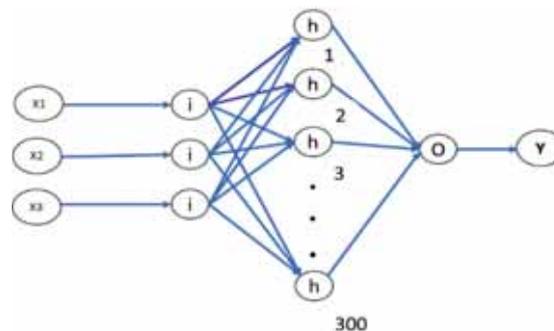
Fig. 2: ANN system design embedded in the control system.

The three ANN inputs consist of signals generated by the throttle valve opening speed sensor (X1), steering speed sensor (X2), and braking speed sensor (X2). These three sensors produce a speed signal which in the next process becomes the average acceleration value (MA) which is processed using Eqs. (1), (2), and (3). The ANN (Y) output is a driver behavior scheme to set the fuel pump.

The types of layers used include logsig, tansig, and purelin models while the number of neural network learning used was 20,000 epochs with a learning success rate of 99.93%. Furthermore, the neural network learning was conducted in the MATLAB Simulink software and later embedded in the control system (microcontroller mega 2560). This study uses 3285 data for ANN training. The training data used in this research is in the form of signal data generated by the throttle valve opening, steering, and braking sensors presented in Fig. 3. This training data is obtained from measurements of real driving behavior characteristics. Data from sensors is taken through data acquisition that has been processed beforehand using Eqs. (1), (2), and (3).
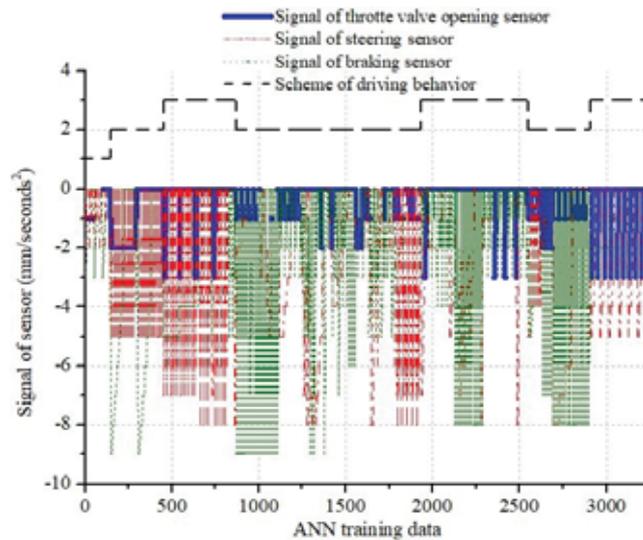
Fig. 3: Data used for ANN training.

Mapping driving behavior recognition with ANN has 27 levels. The driving behavior scheme was designed using the three schemes presented in Table 1. Driving behavior recognition mapping consists of three schemes, including eco-driving (A), stoichiometry/normal driving (B), and sporty driving (C) schemes. Determination of the driving behavior scheme serves to assess the characteristics of each driver. The detailed determination of the driving behavior scheme is presented in Table 4. Table 4 and it was discovered that it was designed on the operation of opening the throttle valve (X2), steering (X2), and braking (X3). The throttle valve opening operation has three low, medium, and high clusters for steering and braking operations. The findings of each cluster for throttle valve opening, steering and braking operations are presented in Table 3.

2.2 Equations Used in Control Systems

The throttle valve opening, steering, and braking sensors generated speed data which were later converted into an acceleration function embedded in the control system (Microcontroller Ni MyRio). Moreover, the average acceleration value (Moving Average - MA) was applied to analyze the driver's behavior from opening the throttle valve, steering, and braking. This method was preferred due to its ability to reduce the value of erratic variations [29,30], thereby increasing its suitability to recognize the variations in the driving behavior.

2.2.1 Acceleration Function Equation

The speed data generated by the throttle valve, steering, and braking sensors were converted to the acceleration function presented in Eq. (1).

$$a_i = \frac{(v_i - v_{(i-1)})}{(t_i - t_{(i-1)})} \tag{1}$$

where: I = 1,2…180

a_i = Acceleration value at the time I (mm/$seconds^2$).

T = Time of speed change in seconds.

v_i = Speed value in period I (mm/$seconds$).

$v_{(i-1)}$ = Speed value in period I -1 (mm/$seconds$).

The next process after the speed has been changed into an acceleration function with Eq. (1) was to determine the average acceleration value.

Table 1: Mapping driving behavior recognition with ANN system

No.	Schematic of driving behavior (Y)	Driving operation type		
		Opening of the throttle valve (X1)	Steering (X2)	Braking (X3)
1.	A	1	1	1
2.	B	1	1	2
3.	B	1	2	2
4.	B	1	1	3
5.	B	1	2	2
6.	B	2	2	2
7.	B	2	1	1
8.	B	2	2	1
9.	B	2	1	2
10.	C	3	1	1
11.	C	3	2	1
12.	C	3	1	2
13.	C	3	2	2
14.	C	3	2	3
15.	C	3	3	2
16.	C	3	3	3
17.	C	3	1	3
18.	C	3	3	1
19.	B	1	2	3
20.	B	1	3	2
21.	B	1	3	3
22.	B	1	3	1
23.	B	2	2	3
24.	B	2	3	3
25.	B	2	3	2
26.	B	2	3	1
27.	B	2	1	3

Description: 1 = low, 2= moderate, 3 = high, A= Eco driving scheme,
B= Stoichiometry/standard driving scheme, C= Sporty driving scheme.

2.2.2 Calculation of Change in Average Acceleration

The acceleration value obtained from Eq. (1) was used to calculate the average value using the Moving Average (MA) method presented in Eq. (2).

$$MA_i = \frac{a_i + a_{i+1} + a_{i+2}}{3} \tag{2}$$

where:
$i = 1, 2, \ldots 180$
MA_i = Average acceleration (mm/$seconds^2$).
a_i = Acceleration value in period i (mm/$seconds^2$).
a_{i+1} = Acceleration value in period i+1 (mm/$seconds^2$)
a_{i+2} = Acceleration value in period i+2 (mm/$seconds^2$).

The MA value was successfully calculated using Eq. (2) while the average was calculated again using Eq. (3). It was discovered that the average MA had 178 constraints,

which were determined once every 15 minutes, and sent to the ECU to recognize driving behavior using the ANN embedded in the control system (microcontroller mega 2560).

$$\overline{MA} = \frac{MA_1 + MA_2 + \cdots MA_{178}}{178} \tag{3}$$

2.2.3 Control System Testing Set Up Developed

The control system test was designed by mounting an AFR sensor (1) on the exhaust gas line to detect the air and fuel mixture while a data acquisition microcontroller (2) was used to change the signal generated by the AFR sensor and the engine speed sensor (4). Moreover, the speed module (3) was employed to convert the pulse signal into a signal readable by the computer as indicated in Fig. 2. It was also observed that the control system (8) developed has two components which include a Ni MyRio microcontroller (a) and a control module (b) with a Mega 2560 microcontroller and an electronic circuit.

The speed data generated every 5 seconds by the throttle valve opening, braking, and steering speed sensors were converted into an acceleration function in the Ni MyRio Microcontroller (a). Moreover, the process to change the speed function into an acceleration function is presented in Eq. (1) after which the acceleration data obtained were averaged through Eq. (2) using the Moving Average – MA method. The MA average results were also calculated using Equation (3) every 15 minutes and sent to the ECU where an AI system with a neural network was used to perform clustering, analyze the different kinds of driving behavior, and send the cluster results to the speed controller module. The speed controller module was applied to control the fuel pump (2) to ensure the fuel is supplied to the engine according to the driving behavior. Additionally, an AI system with a neural network is embedded in the Mega 2560 microcontroller. The control system application setup is presented in Fig. 4.

Description:
1. AFR sensor
2. Microcontroller
3. Speed controller module
4. Speed sensor
5. Engine
6. Computer monitor
7. Computer control system
8. Engine Control Unit (ECU)
   a. NI MyRio
   b. Control module
9. Module speed controller
10. Fuel pump
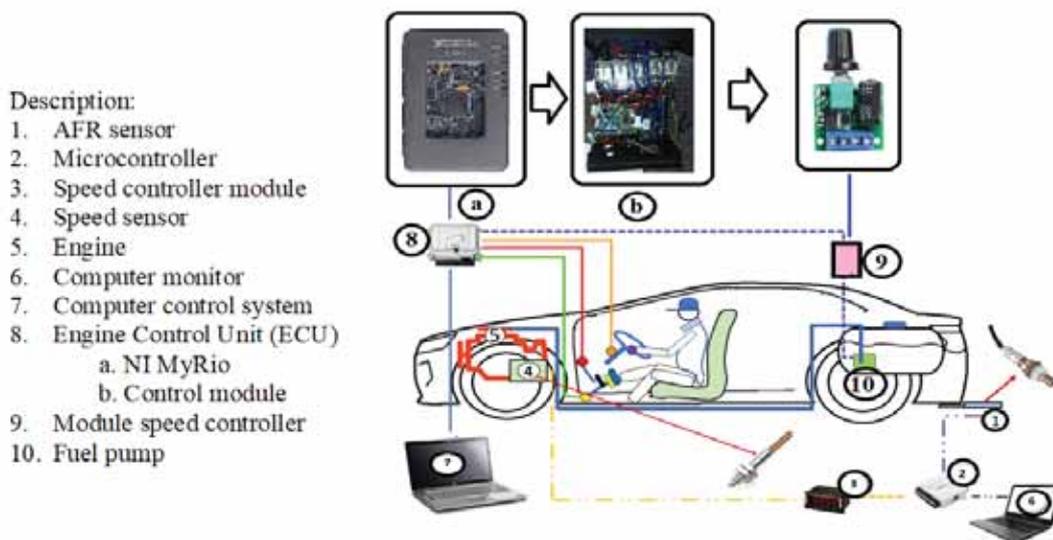
Fig. 4: Driving behavior control system test setup.

The installation of equipment to test the control system with driving behavior is presented in Fig. 5. The AFR sensor was used to measure the air and fuel mixture, the throttle valve speed sensor was mounted on one shaft with the throttle position sensor, while the braking speed sensor was fixed on the brake master shaft and actuated by the brake pedal. Moreover, the steering speed sensor was placed on the side of the steering shaft using a V belt. It is important to note that there was direct documentation of the behavioral recognition test while the speed sensor was used to measure engine speed. While the specifications of the equipment used in the study are presented in Table 2.

Fig. 5: Installation of AFR sensor (1), throttle valve speed sensor (2), braking speed sensor (3), steering speed sensor (4), driving operation (5), and engine speed sensor (6).

The installation of equipment to test the control system with driving behavior is presented in Fig. 5. The AFR sensor was used to measure the air and fuel mixture, the throttle valve speed sensor was mounted on one shaft with the throttle position sensor, while the braking speed sensor was fixed on the brake master shaft and actuated by the brake pedal. Moreover, the steering speed sensor was placed on the side of the steering shaft using a V belt. It is important to note that there was direct documentation of the behavioral recognition test while the speed sensor was used to measure engine speed. While the specifications of the equipment used in the study are presented in Table 2.

Table 2: Specifications of the equipment used

No.	Description	Specification
a.	Steering speed sensor	12-volt Direct Current (DC) motor with gear ratio and dimensions of $\emptyset\ 30 \times 60$ mm.
b.	Braking speed sensor.	
c.	Throttle valve speed sensor.	
d.	Engine.	Gasoline engine with injection type, which has a volume of 1500 cc.
e.	ECU with driving behavior control system using ANN.	NI MyRIO microcontroller and speed controller module equipped with atmega 2560 microcontroller.
f.	Fuel pump	12-volt DC rotary type
g.	Fuel tank	Capacity 44 liters
h.	Test vehicle	Saloon-type vehicle with a capacity of 4 passengers

3.   RESULTS AND DISCUSSION

3.1 Throttle Valve Opening Sensor Results

The control system designed was applied directly to assess driving behaviors in order to determine the characteristics of drivers in real time based on three schemes which include the eco, stoichiometry/standard, and sporty schemes. Each of these schemes was determined using the changes in acceleration generated by the throttle valve opening, steering, and

braking sensors as the input. Moreover, the throttle valve opening signal was calculated through four stages which include the measurement of the throttle valve opening speed, change of the speed to acceleration, calculation of the Moving Average (MA) value, and determination of the average MA.

The change in the throttle valve opening speed was monitored over three periods as indicated in Fig. 6 with the first period found to be between 0 - 900 seconds, the second was 901-1800 seconds, and the third was 1,801 – 2,700 seconds. The findings showed that the speed change for these periods was 0–2.5, 0–2.6, and 0–4.6 mm/second respectively. It was observed that the third period tends to have a higher speed change rate and density level.
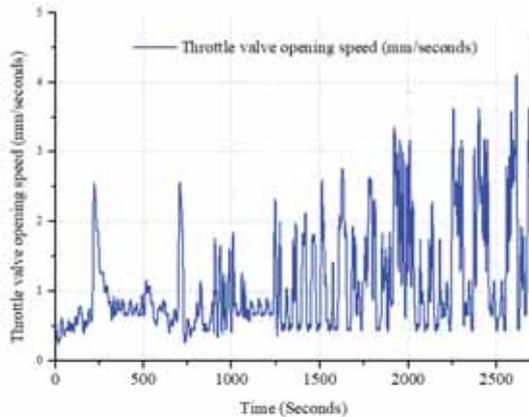
Fig. 6: Change of throttle valve opening speed.
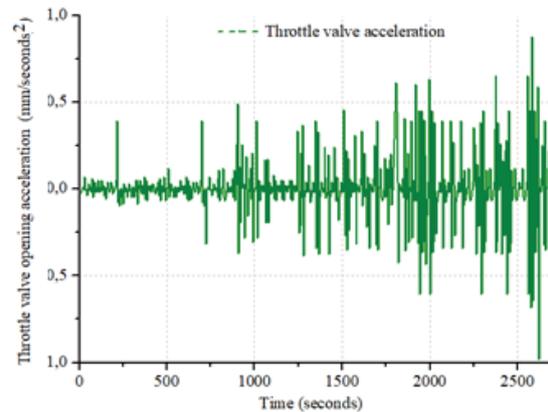
Fig. 7: Change of throttle valve opening acceleration.

The speed data presented in Fig. 6 was converted into an acceleration function using Eq. (1) and the results are presented in Fig. 7 as a function of the change in acceleration. It was discovered that there was an acceleration value of $0 – 0.5$ mm/seconds^2 and a negative acceleration or deceleration value of 0-(-0.45 mm/seconds^2) for $0 – 900$ seconds. The findings also showed that the second period, $901 – 1,800$ seconds, had $0 – 0.7$ mm/seconds^2 and 0-(-0.7 mm/seconds^2) while the third period, $1,801 – 2,700$ seconds, had $0 – 0.9$ mm/seconds^2 and 0-(-0.9 mm/seconds^2) respectively. However, it is important to note that the third period had a higher density.

The acceleration/deceleration values presented in Fig. 7 were used in the next process to determine the average using the MA method in Eq. (2). This was also followed by finding the average of the MA values using Eq. (3) and the results are presented in Fig. 8. It was discovered that the MA value for the $0 – 900$ seconds period ranged from $0 – 0.14$ mm/seconds^2 and a negative acceleration (deceleration) value of 0-(-0.14 mm/seconds^2). The results further showed that the $901 – 1,800$ seconds period had $0 – 0.12$ mm/seconds^2 and 0-(-0.12 mm/seconds^2) while $1,801 – 2,700$ seconds period had $0 – 0.2$ mm/seconds^2 and 0-(-0.22 mm/seconds^2) respectively. Meanwhile, the average MA for $0 – 900$ seconds was found to be 0.00 while $901 – 1,800$ seconds had 0.002, and $1801 – 2700$ seconds had 0.003. It is pertinent to restate that the average MA value was determined every 15 minutes and sent to the developed control system. It was discovered that the highest MA average value was recorded in the last period.

3.2 Steering Sensor Results

The changes observed in the steering speed of each driver are presented in Fig. 9. It was discovered that the drivers possess quite diverse characteristics as observed with the 0

– 17 mm/second recorded for the 0 – 900 seconds period, 0 – 14 mm/seconds for the 901 – 1,800 seconds period, and 0 – 43 mm/seconds for the 1,801 – 2,700 seconds period which is the highest compared to the others.
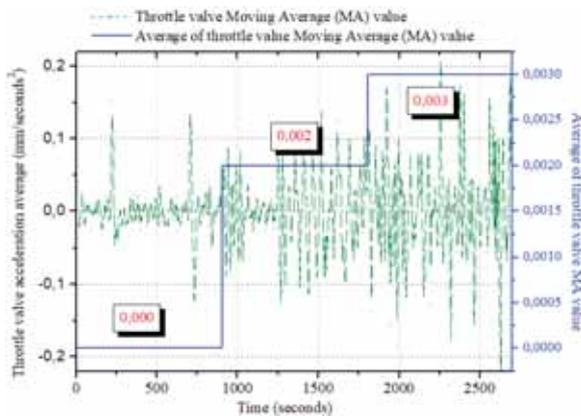
Fig. 8: Average MA value of driver throttle valve opening behavior.
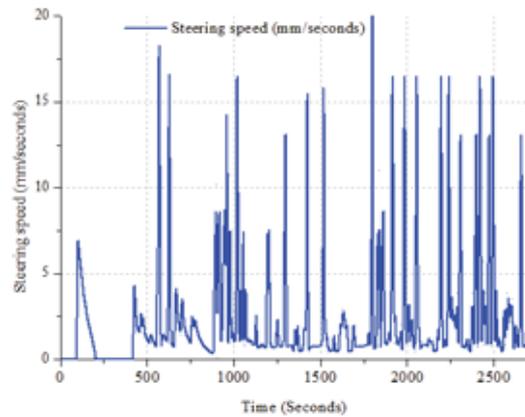
Fig. 9: Steering speed changes behavior.

The change in driver acceleration presented in Fig. 10 showed that the 0 - 900 seconds period had an acceleration value of $0 – 4$ mm/seconds^2 and a negative acceleration (deceleration) value of 0-(-4 mm/seconds^2) while 901 – 1,800 seconds period had $0 – 4.1$ mm/seconds^2 and 0-(-4.1 mm/seconds^2) and 1,801 – 2,700 seconds period had $0 – 5.8$ mm/seconds^2 and 0-(-3 mm/seconds^2) respectively. It was observed that the driver tends to add more steering acceleration in the third period.

The MA values recorded based on the changes in the driver's steering acceleration are presented in Fig. 11. It was discovered that the 0 – 900 seconds period had an MA range between 0 - 1.3 mm/seconds^2 and negative acceleration (deceleration) of 0-(-1.3 mm /seconds^2) while the 901 – 1,800 seconds period had $0 – 1.1$ mm/seconds^2 and 0-(-1.1 mm/seconds^2) and the 1,801 – 2,700 seconds period had $0 – 1.1$ mm/seconds^2 and 0-(-1.1 mm/seconds^2) respectively. Moreover, the average MA value in the 0 – 900 seconds period was found to be 0.0000, the 901 – 1,800 seconds period had 0.002, and the 1,801 – 2,700 seconds had 0.03 which is the highest.
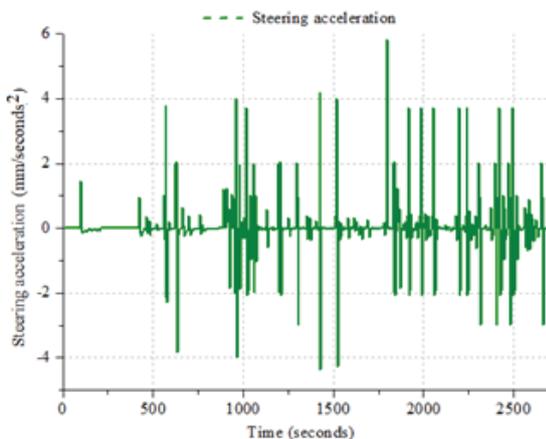
Fig. 10: Changes in steering acceleration behavior while driving.
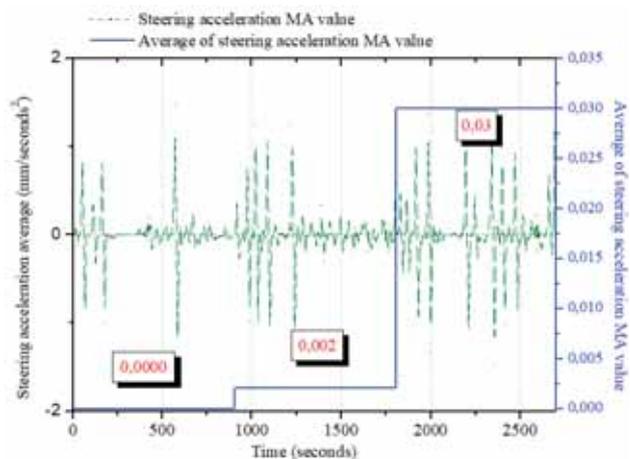
Fig. 11: Average MA value of driver steering behavior.

3.3 Braking Sensor Results

The changes in the driver's braking speed presented in Fig. 12 showed that the value at the 0 – 900 seconds period was 0–2.9 mm/second, the 901 – 1,800 seconds period had 0-2.6 mm/second, and the 1,801 – 2,700 seconds had 0 – 4.5 mm/second. This implies the 0 – 900 seconds period has the greatest braking speed and lower density frequency compared to the others.

The speed data presented in Fig. 10 were converted into an acceleration function through Eq. (1) to determine the changes in driver acceleration and the results are presented in Fig. 13. It was discovered that the pattern acceleration for the drivers at the 0 – 900 seconds period was in the range of 0 – 0.4 mm/seconds^2 and a negative acceleration (deceleration) value of 0-(-0.3 mm/seconds^2) while the 901 – 1,800 seconds period had 0 – 0.7 mm/seconds^2 and 0-(-0.6 mm/seconds^2 and 1,801 – 2,700 seconds period had 0 – 0.8 mm/seconds^2 and 0-(-0.8 mm/seconds^2) respectively.
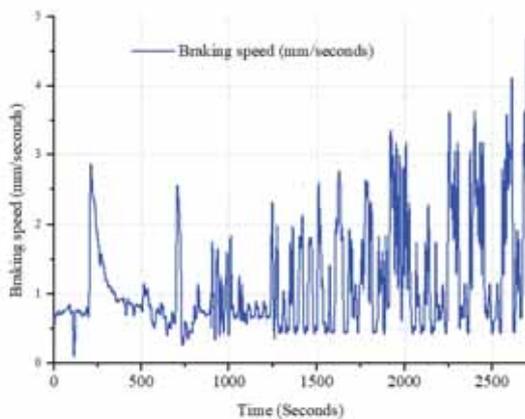
Fig. 12: Braking speed on observed driving behavior.
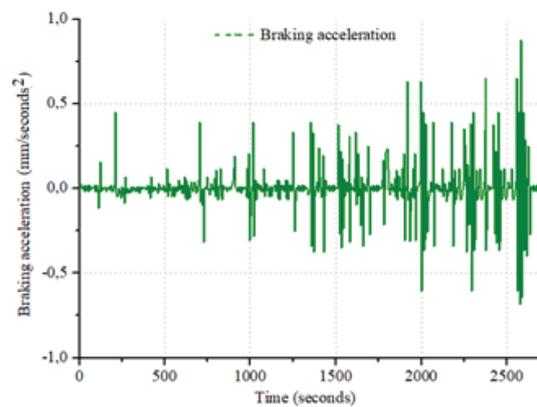
Fig. 13: Changes in the acceleration of the driver's braking behavior.

The MA values associated with the changes in the driver's braking acceleration are presented in Fig. 14. It was discovered that the 0 – 900 seconds period has a pattern where the MA values had the lowest frequency with values ranging between 0 – 0.15 mm/seconds^2 and negative acceleration (deceleration) values 0-(- 0.13 mm/seconds^2). The findings further showed that the 901 – 1,800 seconds period had 0 – 0.12 mm/seconds^2 and 0-(-0.12 mm/seconds^2) while the 1,801 – 2,700 seconds period had 0 – 2.1 mm/seconds^2 and 0-(-0.18 mm/seconds^2). These MA values were later used to determine the average MA and the results showed that the value for the second driver was 0.0003 for the 0 – 900 seconds period, 0.0016 for the 901 – 1,800 seconds period, and 0.0021 for the 1,801 – 2,700 seconds period. This implies the 0 – 900 seconds period has the lowest average MA.

3.4 Driving Behavior Recognition Range

The driving behavior control system was designed using several variables that cause a slight difference between the simulated conditions and the real control system. It was discovered from the control system that the real driving behavior has a negative value because of the deceleration process and this cannot be read by the microcontroller. Therefore, a normalization system was needed through the inclusion of several different variables in the simulation scale and real control systems but the concept remains the same. Normalization is a process of adding constants to signal conditioning so that the microcontroller can work according to its designation. The methods used to recognize the

driving behavior are listed in Table 3 while the real decision-making process using the control system designed was based on the cluster system developed with the average MA value as indicated in Table 4.
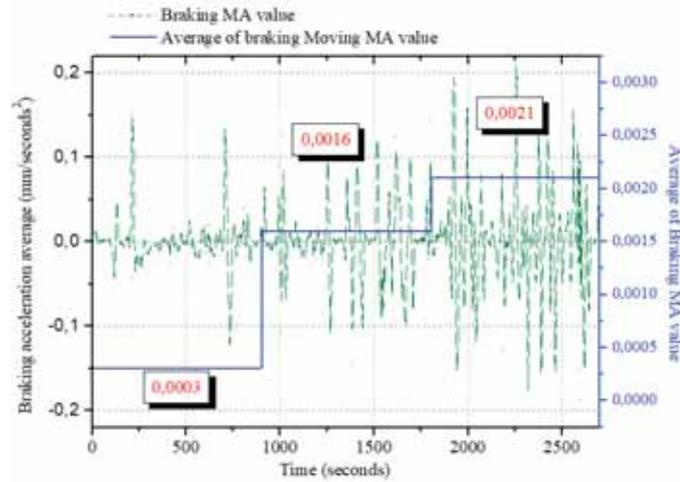
Fig. 14: Average MA value of the driver's braking behavior.

Table 3: Cluster average ($\overline{MA}$) for the throttle valve, steering, and braking

Description	Low (mm/$seconds^2$)	Middle (mm/$seconds^2$)	High (mm/$seconds^2$)
MA throttle valve acceleration	(-0.001)-0.001	0.002	0.003
MA from steering	(-0.02) - 0.02	(-0.05) – (-0.03) & 0.03 - 0.05	(-0.08) – (-0.06) & 0.06 - 0.08
MA of braking acceleration	(-0.0003)-0.0003	(-0.0009)–(-0.0006) & 0.0006 - 0.0009	(-0.005) – (-0.001) & 0.001 - 0.005

The average MA value generated by the sensor was analyzed to identify and recognize the driving behavior clusters using the ANN that has been trained and embedded in the controller system as presented in Table 3. It was discovered that the first period was in the eco scheme, the second period was in the stoichiometric scheme, and the third period was in the sporty scheme.

Table 4: Recognition of the driver's steering behavior in the control system

Description	First period (0-900s)	Second period (901-1,800s)	Third period (1801-2,700s)
MA throttle valve acceleration	Low (0.000)	Middle (0.002)	High (0.003)
MA acceleration steering	Low (0.0088)	Low (0.0101)	Low (0.0021)
MA braking acceleration	Low (0.0003)	High (0.0016)	High (0.0021)
Driving behavior scheme	*Eco*	*Stoichiometry*	*Sporty*

3.5 AFR Dynamics Results

The AFR values were measured using the data previously acquired and validated through an AFR meter and were subsequently applied to determine the influence of driving behavior on their variations based on the information in Table 3. The first period is in the eco scheme, the second is in the stoichiometric scheme, and the third is in the sporty scheme. It is important to note that vehicles experience acceleration and deceleration when operating on the highway. Acceleration is a condition associated with an increase in the speed of the vehicle by opening the throttle valve while deceleration involves reducing the vehicle's speed by closing the throttle valve. An increase in the vehicle speed by opening the throttle valve usually leads to a decrease in the AFR value and vice versa. The dynamics of the AFR value including the increase or decrease are presented in Figure 15 for the models with and without the driving behavior control system.

In the application of the driving behavior control system, the first period of driving behavior is known to have an AFR value between 14.8 – 17.7 with an average of 15.87, while the second period is between 14 – 16.6 with an average of 14.84, and the third period had the value is between 11.5 – 16 and the average is 13.66. This shows that the driving behavior in the first period allowed maximum fuel economy compared to the second and third periods. Meanwhile, the driving behavior in the third period led to the production of maximum power by the engine as indicated by the average AFR value recorded. Meanwhile, the dynamics of AFR without applying a driving behavior control system have quite a high difference in average values. The average AFR value without the developed control system is 14.78 for 2700 seconds. In the first period, the driver rarely decelerates, so the AFR value has a lower fluctuation range value when compared to the second and third periods.

The achievement of the highest fuel economy through the AFR value above the stoichiometry of 14.7. This is in line with a previous study that predicted the potential development of technology to achieve a lean AFR scale above the stoichiometry for commercialization purpose in order to have energy-efficient machines [31]. The research did not discuss the exact conditions to implement lean AFR but this current study considered its application based on driving behavior. It was discovered that the AFR above stoichiometry achieved in the first period, 0 – 900 seconds, led to smoother acceleration and this implies there is no need for large engine power.
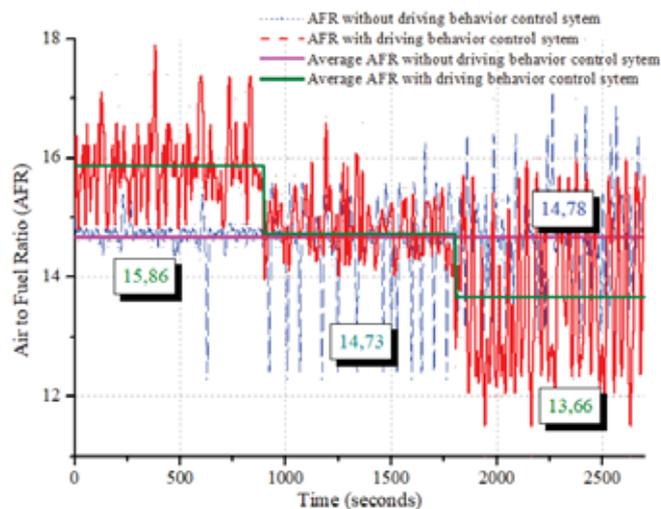
Fig. 15: AFR dynamics on observed driving behavior.

The AFR values in the second period, 901 – 1,800 seconds, are classified to be in the stoichiometric range (14.7) and this is in line with the findings of a previous study that AFR stoichiometry ensures the achievement of the most optimal value for power and fuel saving [7,32]. The research did not discuss the time and requirements to achieve maximum power saving but these are developed in this current study based on driving behavior.

The driving behavior in the third period, 1,801 – 2,700 seconds, was observed to have led to the fulfillment of maximum power as indicated by the average AFR stoichiometry value of 13.66. This agrees with the findings of a previous study that an AFR under a stoichiometry value of 12–14 usually produces the greatest engine power [33]. The control system showed that the driving behavior required high engine power and this led to the addition of more fuel at an average AFR value of 13.66.

This shows that the control system was successfully developed with due consideration for the driving behavior. This was able to improve driving comfort in terms of fuel economy as indicated by the AFR values above stoichiometry and fulfillment of the engine power requirements with values below stoichiometry. It was discovered that the fuel economy was satisfied in the first period when driving behavior is categorized as eco-scheme while the demand for the engine power was fulfilled in the third period when driving behavior is included in the sporty scheme.

4. CONCLUSION

The design of AFR management with driving behavior control using ANN has been successfully applied to actual vehicles. The designed control system can recognize driver behavior in real-time to control fuel and increase vehicle comfort. Driving comfort is achieved by meeting the need for fuel economy when the driving behavior is included in the eco-driving scheme and aspects of fulfilling engine power when the driving behavior is included in the sporty scheme. When the driver enters the eco-driving scheme, the control system can control AFR with an average value of 15.68 (entering the lean range). AFR above stoichiometry can improve fuel economy. The eco scheme is achieved in the first period (0–900 seconds), where this decision is based on the MA value of throttle valve acceleration in the low category (0.0000 mm/second^2), MA steering acceleration in the low category (0.0088 mm/second^2) and braking acceleration MA in the category low (0.0003 mm/s^2). When the driving behavior enters the sporty driving scheme, the control system can control AFR with an average value of 13.66 (below stoichiometry). AFR under stoichiometry produces maximum engine power. The sporty scheme is achieved in the third period (0–900 seconds), where this decision is based on the MA value of throttle valve acceleration in the high category (0.003 mm/second^2), MA steering acceleration in the low category (0.0021 mm/second^2) and braking acceleration MA in the high category (0.0021 mm/second^2). This research is applied to vehicles with gasoline engines and has not considered the road angle. For this reason, future research can be applied to electric, gas-fueled, and fuel-cell vehicles, considering the road angle.

ACKNOWLEDGMENT

REFERENCES

[1] Burdzik R. (2022) A comprehensive diagnostic system for vehicle suspensions based on a neural classifier and wavelet resonance estimators. Measurement, 200: 111602 https://doi.org/10.1016-/j.measurement.2022.111602.

[2] Uslu S, Celik MB.(2020) Performance and exhaust emission prediction of a si engine fueled with i amyl alcohol-gasoline blends : An ANN coupled RSM based optimization. Fuel, 265: 116922. https://doi.org/10.1016/j.fuel.2019.116922.

[3] Al-fattah SM. (2020) Non-OPEC conventional oil : Production decline, supply outlook and key implications. Journal of Petroleum Science and Engineering 189: 107049. https://doi.org-/10.1016/j.petrol.2020.1070.

[4] Kutlu O. (2020) Global oil production declines in June 2020. Energy. Available: https://www.-aa.com.tr/en/energy/international-organization/global-oil-production-declines-in-june-2020-/29901.

[5] Wang Y, Fan Y, Wang D, Liu Y, Qiu Z, Liu J. (2020) Optimization of the areas of solar collectors and photovoltaic panels in liquid desiccant air-conditioning systems using solar energy in isolated low-latitude islands. Energy, 198: 117324. https://doi.org/10.1016/j.energy.-2020.117324.

[6] Gagliardi G, Mari D, Tedesco F, Casavola A. (2021) An Air-to-Fuel ratio estimation strategy for turbocharged spark-ignition engines based on sparse binary HEGO sensor measures and hybrid linear observers. Control Engineering Practice, 107: 104694. doi:10.1016-/j.conengprac.2020.104694.

[7] Ahmed SFA. (2019) Analyzing and predicting the relation between air – fuel ratio (AFR), lambda (λ) and the exhaust emissions percentages and values of gasoline - fueled vehicles using versatile and portable emissions measurement system tool. SN Applied Science, 1(11): 1-12. https://doi.org/10.1007/s42452-019-1392–5.

[8] Zhao X, Wang Z, Xu Z, Wang Y, Li X, Qu X. (2020) Field experiments on longitudinal characteristics of human driver behavior following an autonomous vehicle. Transportation Research Part C: Emerging Technologies, 114: 205-224. https://doi.org/10.1016-/j.trc.2020.02.018.

[9] Sharma A, Zheng Z, Bhaskar A, Haque M. (2019) Modelling car-following behaviour of connected vehicles with a focus on driver compliance. Transportation Research Part B, 126: 256-279. https://doi.org/10.1016/j.trb.2019.06.008.

[10] Fadhloun K, Rakha H. (2020) Novel vehicle dynamics and human behavior car-following model : Model development and preliminary testing. International Journal of Transportation Science and Technology, 9: 14-28. https://doi.org/10.1016/j.ijtst.2019.05.004.

[11] Grove K, Soccolich S, Engström J, Hanowski R. (2019) Driver visual behavior while using adaptive cruise control on commercial motor vehicles. Transportation Research Part F, 60: 343-352. https://doi.org/10.1016/j.trf.2018.10.013.

[12] Fung KC, Dick TJ.(2016) System and method for responding to driver behavior. https://patents.google.com/patent/US9440646B2.

[13] Martinelli F, Mercaldo F, Orlando A, Nardone V, Santone A, Kumar A. (2020) Human behavior characterization for driving style recognition in vehicle system. Computers & Electrical Engineering, 83: 02504. https://doi.org/10.1016/j.compeleceng.2017.

[14] Yuan Y, Lu Y, Wang Q. (2020) Adaptive forward vehicle collision warning based on driving behavior. Neurocomputing, 408: 64-71. https://doi.org/10.1016/j.neucom.2019.11.02.

[15] Raz O, Fleishman H, Mulchadsky I. (2008) System and method for vehicle driver behavior analysis and evaluation. https://patents.google.com/patent/US7389178B/en.

[16] Hong Z, Chen Y, Wu Y. (2020) A driver behavior assessment and recommendation system for connected vehicles to prod Prevention, 139: 105460. https://doi.org/10.1016/j.aap.2020.-105460.

[17] Mafeni J, Majid S, Mesgarpour M, Torres M, Figueredo GP, Chapman P. (2020) Evaluating the impact of Heavy Goods Vehicle driver monitoring and coaching to reduce risky behaviour, Accident Analysis & Prevention, 146: 105754. https://doi.org/10.1016/j.aap.2020.105754.

[18] Bando T. (2015) System for detecting abnormal driving behavior. https://patents.google.com-/patent/US9111400B2/en.

[19] Hongbo G, Guotao X, Hongzhe L, Xinyu Z. (2017) Lateral control of autonomous vehicles based on learning driver behavior via cloud model. The Journal of China Universities of Posts and Telecommunications, 24(2): 10-17. http://dx.doi.org/10.1016/S1005-8885(17)601.

[20] Ashkrof P, Homem G, Correia DA, Van Arem B. (2020) Analysis of the effect of charging needs on battery electric vehicle drivers' route choice behaviour : A case study in the Netherlands. Transportation Research Part D, 78: 102206. https://doi.org/10.1016/j.trd.2019.-102206.

[21] Silver A, Lewis L. (2015) Automatic identification of a vehicle driver based on driving behavior. https://patents.google.com/patent/US9201932B2/en.

[22] Yansong R, O'Gorman L, Wood TL. (2019) Driver behavior monitoring systems and methods for driver behavior monitoring. https://patents.google.com/patent/US9201932B2/en.

[23] Julian DJ and Agrawal A. (2017) Driver behavior monitoring. https://patents.google.com-/patent/US10460600B2/en.

[24] Stogios C, Kasraian D, Roorda MJ, and Hatzopoulou M. (2019) Simulating impacts of automated driving behavior and traffic conditions on vehicle emissions. Transportation Research Part D, 76: 176-192. https://doi.org/10.1016/j.trd.2019.09.020.

[25] Kohl J, Gross A, Henning M, Baumgarten T. (2020) Driver glance behavior towards displayed images on in-vehicle information systems under real driving conditions. Transportation Research Part F: Traffic Psychology and Behaviour, 70: 163-174. https://doi.org/10.1016/j.trf.2020.01.017.

[26] Xing Y, Lv C, Cao D, Lu C. (2020) Energy oriented driving behavior analysis and personalized prediction of vehicle states with joint time series modeling. Applied Energy, 261: 114471. https://doi.org/10.1016/j.apenergy.2019.114.

[27] Vaezipour A, Rakotonirainy A, Haworth N. (2018) A simulator evaluation of in-vehicle human machine interfaces for eco-safe driving. Transportation Research Part A: Policy and Practice, 118: 696-713. https://doi.org/10.1016/j.tra.2018.10.022.

[28] Reinolsmann N, Alhajyaseen W, Brijs T, Pirdavani A, Hussain Q, Brijs K. (2019) Investigating the impact of dynamic merge control strategies on driving behavior on rural and urban expressways – A driving simulator study. Transportation Research Part F: Traffic Psychology and Behaviour, 65: 469-484. https://doi.org/10.1016/j.trf.2019.08.010.

[29] Kiat D, Poh H, Yee C, Zhen R, Markus C, Ping T. (2021) Internal quality control : Moving average algorithms outperform westgard rules. Clinical Biochemistry, 98: 63–69. https://doi.org/10.1016/j.clinbiochem.2021.09.007.

[30] Maverick JB. (2022) Advantages and disadvantages of the Simple Moving Average (SMA)?, Investopedia. Available: https://www.investopedia.com/ask/answers/013015/what-are-main-advantages-and-disadvantages-using-simple-moving-average-sma.asp. [Accessed: 05-May-2022].

[31] Zhu C, Wang P, Wang P, Liu Z, Wang P, Liu Z. (2018) Model Prediction Control of Fuel-Air Ratio for Lean-Burn Spark Ignition Gasoline. IFAC-PapersOnLine, 51(31): 640–645. https://doi.org/10.1016/j.ifacol.2018.10.150.

[32] Sardarmehni T, Aghili AA, Menhaj MB. (2019) Fuzzy model predictive control of normalized air-to-fuel ratio in internal combustion engines. Soft Computing, 23(15): 6169-6182. https://doi.org/10.1007/s00500-018–3270, 2019.

[33] Martyr A and Plint M. (2007) Third Edition Engine Testing Theory and Practice. Elsevier Ltd.

A ROBUST FRAMEWORK FOR DRIVER FATIGUE DETECTION FROM EEG SIGNALS USING ENHANCEMENT OF MODIFIED Z-SCORE AND MULTIPLE MACHINE LEARNING ARCHITECTURES

**RAFIUDDIN ABDUBRANI[1*], MAHFUZAH MUSTAFA[1]
AND ZARITH LIYANA ZAHARI[1,2]**

*[1] Faculty of Electrical and Electronics Engineering Technology,
Universiti Malaysia Pahang, Pekan 26600, Pahang, Malaysia
[2]Electronic Section, Universiti Kuala Lumpur British Malaysia Institute,
Gombak 53100, Selangor, Malaysia*

[*]Corresponding author: rafi128@yahoo.com

ABSTRACT: Physiological signals, such as electroencephalogram (EEG), are used to observe a driver's brain activities. A portable EEG system provides several advantages, including ease of operation, cost-effectiveness, portability, and few physical restrictions. However, it can be challenging to analyse EEG signals as they often contain various artefacts, including muscle activities, eye blinking, and unwanted noises. This study utilised an independent component analysis (ICA) approach to eliminate such unwanted signals from the unprocessed EEG data of 12 young, physically fit male participants between the ages of 19 and 24 who took part in a driving simulation. Furthermore, driver fatigue state detection was carried out using multichannel EEG signals obtained from O1, O2, Fp1, Fp2, P3, P4, F3, and F4. An enhanced modified z-score was utilised with features extracted from a time-frequency domain continuous wavelet transform (CWT) to elevate the reliability of driver fatigue classification. The proposed methodology offers several advantages. First, multichannel EEG analysis improves the accuracy of sleep stage detection, which is vital for accurate driver fatigue detection. Second, an enhanced modified z-score in feature extraction is more robust than conventional z-score techniques, making it more effective for removing outlier values and improving classification accuracy. Third, the proposed approach for detecting driver fatigue employs multiple machine learning classifiers, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Artificial Neural Networks (ANNs) that utilise Long Short-Term Memory (LSTM), and also machine learning techniques like Support Vector Machines (SVM). The evaluation of five classifiers was performed through 5-fold cross-validation. The outcomes indicate that the suggested framework attains exceptional precision in identifying driver fatigue, with an average accuracy rate of 96.07%. Among the classifiers, the ANN classifier achieved the most significant precision of 99.65%, and the SVM classifier ranked second with an accuracy of 97.89%. Based on the results of the receiver operating characteristic (ROC) and area under the curve (AUC) analysis, it was observed that all the classifiers had an outstanding performance, with an average AUC value of 0.95. This study's contribution lies in presenting a comprehensive and effective framework that can accurately detect driver fatigue from EEG signals.

ABSTRAK: Isyarat fisiologi, seperti elektroencefalogram (EEG), digunakan bagi memerhati aktiviti otak pemandu. Sistem EEG mudah alih menyediakan beberapa kelebihan, termasuk kemudahan operasi, keberkesanan kos, mudah alih dan sedikit sekatan fizikal. Namun, isyarat EEG mungkin sukar dianalisis kerana ia sering

mengandungi pelbagai artifak, termasuk aktiviti otot, mata berkedip dan bunyi yang tidak diingini. Kajian ini menggunakan pendekatan analisis komponen bebas (ICA) bagi membuang isyarat tidak diperlukan daripada data EEG yang belum diproses daripada 12 peserta lelaki muda, cergas fizikal berumur 19 hingga 24 tahun yang mengambil bahagian dalam simulasi pemanduan. Tambahan, pengesanan keadaan lesu pemandu telah dijalankan menggunakan isyarat EEG berbilang saluran yang diperoleh dari O1, O2, Fp1, Fp2, P3, P4, F3, dan F4. Penambah baik skor z digunakan dengan ciri diekstrak daripada transformasi wavelet berterusan (CWT) domain frekuensi masa bagi meningkatkan kebolehpercayaan klasifikasi keletihan pemandu. Metodologi yang dicadangkan menawarkan beberapa kelebihan. Pertama, analisis EEG berbilang saluran meningkatkan ketepatan pengesanan peringkat tidur, penting bagi pengesanan keletihan pemandu secara tepat. Kedua, penambah baik skor z dalam pengekstrak ciri adalah lebih teguh daripada teknik skor z konvensional, menjadikannya lebih berkesan bagi membuang unsur luaran dan meningkatkan ketepatan pengelasan. Ketiga, pendekatan yang dicadangkan bagi mengesan keletihan pemandu menggunakan pelbagai pengelas pembelajaran mesin, seperti Rangkaian Neural Konvolusi (CNN), Rangkaian Neural Berulang (RNN), Rangkaian Neural Buatan (ANN) yang menggunakan Memori Jangka Pendek Panjang (LSTM), dan juga teknik pembelajaran mesin seperti Mesin Vektor Sokongan (SVM). Penilaian lima pengelas dilakukan melalui pengesahan silang 5 kali ganda. Dapatan kajian menunjukkan cadangan rangka kerja ini mencapai ketepatan yang luar biasa dalam mengenal pasti keletihan pemandu, dengan kadar ketepatan purata 96.07%. Antara kesemua pengelas, pengelas ANN mencapai ketepatan paling ketara sebanyak 99.65%, dan pengelas SVM menduduki tempat kedua dengan ketepatan 97.89%. Berdasarkan keputusan analisis ciri operasi penerima (ROC) dan kawasan di bawah lengkung (AUC), didapati semua pengelas mempunyai prestasi cemerlang, dengan purata nilai AUC 0.95. Sumbangan kajian ini adalah terletak pada rangka kerja yang komprehensif dan berkesan mengesan keletihan pemandu secara tepat melalui isyarat EEG.

KEYWORDS: *driver fatigue; electroencephalogram (EEG); z-score; deep learning*

1. INTRODUCTION

According to statistics from the World Health Organization, roughly 127,000 individuals lose their lives in traffic accidents yearly, with nearly one-third of those casualties being teenagers and young adults [1]. Fatigue driving contributes to fatalities in road accidents, contributing to more than ten thousand deaths in a conservative estimate. Recently, some autonomous vehicles have proposed a warning system to prevent road accidents due to driver fatigue. The system would prompt drivers to take a break from prolonged driving by sounding an alarm in the vehicle, notifying the driver to stop driving and grab a coffee break.

Physiological signals such as electroencephalograms (EEG) are used to observe a driver's brain activities. A portable EEG system provides several advantages over other electroencephalography systems, including ease of operation, cost-effectiveness, portability, and few physical restrictions [2]. The presence of artefacts in EEG signals, such as muscle activity, eye blinking, and unwanted noise, can pose a significant challenge for analysis. Therefore, the current paper proposes using an independent component analysis (ICA) technique to eliminate such noise from the raw EEG signal. Numerous studies have suggested that an essential component of precise sleep stage detection is the analysis of multichannel EEGs [3]. Consequently, the present study considers multichannel EEG signals obtained from O1, O2, Fp1, Fp2, P3, P4, F3, and F4 for detecting driver fatigue states.

The features from a time-frequency domain, continuous wavelet transform (CWT) with enhanced modified z-score improved the accuracy of driver fatigue classification. It is important to choose the best features to get better results. The Morlet mother wavelet is a common practice in conventional CWT techniques due to its computational efficiency, surpassing other methods. This is because the Morlet wavelet involves fewer computations, most of which are performed through the fast Fourier transform, requiring less code [4].

In the field of data analysis and quality control, the identification of outliers is a crucial step in ensuring the accuracy and validity of statistical analyses. The z-score is a widely used method for detecting outliers in datasets, but it is susceptible to extreme values and is not considered robust in the presence of such outliers. The modified z-score was introduced to address this issue, which is less sensitive to outliers and has become a popular method for outlier detection in various applications. In recent years, the modified z-score has also been applied to feature extraction in machine learning and signal processing, where removing outlier values is crucial for accurate and robust analysis. This paper presents an enhancement of the modified z-score method for feature extraction in signal processing, specifically in driver fatigue detection using EEG signals.

Our proposed method has several strengths. First, using multichannel EEG analysis improves the accuracy of sleep stage detection, which is vital for accurate driver fatigue detection. Second, our use of enhanced modified z-score in feature extraction is more robust than conventional z-score techniques, making it more effective for removing outlier values and improving classification accuracy. Third, our approach utilises various machine learning classifiers, providing a comprehensive and accurate method for driver fatigue detection.

This paper presents a methodology for the precise identification of distinct levels of driver drowsiness by utilising diverse machine learning classifiers, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Artificial Neural Networks (ANNs) that incorporate Long Short-Term Memory (LSTM), and machine learning approaches like Support Vector Machines (SVM). A modified z-score technique to enhance the statistical feature of the classification was also introduced, which significantly improved the accuracy of the proposed method. To evaluate the effectiveness of this approach, a 5-fold cross-validation strategy was employed to distinguish between driver fatigue and normal states.

2. RELATED WORKS

Outlier detection is critical in various fields, including environmental monitoring, geology, epidemiology, and data mining. The modified z-score is a frequently employed technique for detecting outliers, which considers the weighted mean of adjacent data points to estimate the anticipated value of each point. Aggarwal et al. proposed a modified z-score method for detecting spatial outliers in datasets with spatial autocorrelation [5]. The technique improves the accuracy and robustness of the z-score test using a trimmed mean instead of the usual arithmetic mean. The study evaluated the method on simulated and real-world datasets and showed promising results in detecting spatial outliers. Although the modified z-score method proposed by Aggarwal et al. effectively detects spatial outliers with spatial autocorrelation, it may not perform well in datasets without spatial autocorrelation. Additionally, using a trimmed mean instead of the arithmetic mean may result in the loss of valuable information from the dataset.

Sandbhor et al. investigated the importance of detecting outliers in data mining and their effect on the quality and output of prediction models [6]. The study's primary objective was to determine the most effective approach for detecting outliers in neural networks (NN) to forecast real estate values. The authors assessed several univariate outlier detection methods, such as Tukey's Standard Deviation (SD), median, z-score, median absolute deviation (MAD), and modified z-score, on a set of 3,094 instances of property sales data. Based on the findings, it can be concluded that for this particular problem, the median technique proved to be the most efficient approach for detecting outliers. Although Sandbhor et al. found that the median technique was the most efficient approach for detecting outliers in neural networks for real estate value prediction, it is important to note that this conclusion may not necessarily apply to other types of datasets or outlier detection techniques. Moreover, the study only evaluated univariate outlier detection methods and did not consider multivariate techniques, which may be more effective in certain applications.

Leite et al. conducted a study to evaluate the effectiveness of the modified z-score as an indicator for identifying changes in entropy-based features to detect faults in bearings [7]. The research involved using 12 entropy-based features across the time, frequency, and time-frequency domains, in addition to three different entropy measures, namely Shannon entropy, Renyi entropy, and Jensen-Renyi divergence. The proposed technique was applied to process two real-bearing datasets obtained from experiments conducted until the point of failure. Furthermore, three bearings with different defects were examined to verify the performance of the entropy-based features. The results demonstrated that the modified z-score is a robust method for detecting changes in entropy-based features, highlighting its potential for early detection of anomalies in the vibration signals of bearings. This finding suggests that the proposed technique can be effectively utilised for fault diagnosis in bearings. However, it is important to note that the study only evaluated the effectiveness of the modified z-score method on two real-bearing datasets obtained from experiments conducted until the point of failure and three bearings with different defects.

Although outlier detection is a powerful tool for identifying unique data points, several limitations must be considered. For instance, in some cases, there may not be a clear definition of what constitutes an outlier, making it challenging to determine which data points to flag. Moreover, outlier detection methods may produce false positives or negatives, leading to incorrect conclusions and recommendations. Furthermore, choosing the appropriate outlier detection method for a specific dataset or problem can be complex, and there is no one-size-fits-all solution. Additionally, while outlier detection can identify anomalous data points, it may not always address the underlying cause of the outlier or provide a solution to the problem. Therefore, to get the most out of outlier detection, careful consideration of the goals and context of the analysis is essential. It is also important to use outlier detection in conjunction with other analytical tools and techniques to gain a more comprehensive understanding of the data and to develop effective solutions that address the root cause of any identified anomalies.

Several techniques have been suggested to identify the underlying mechanisms of fatigue in EEG signals. Among them, one method entails computing distinct types of entropies as feature sets based on a solitary channel [8]. Quintero-Rincon has presented a straightforward and efficient method for identifying driver fatigue in real-time systems using a single-channel EEG signal [9]. The algorithm selects the most significant channel and extracts four feature parameters to detect fatigue using an ensemble bagged decision trees classifier. By utilising data obtained from the Jiangxi University of Technology database, the proposed approach achieves an accuracy of 92.7% with a 1.8-second time delay. However, it is important to note that the study evaluated the method on a specific

dataset, and further research may be needed to determine its effectiveness on other datasets and under different conditions. Additionally, the time delay of 1.8 seconds may not be practical for real-time monitoring in some situations, and it is important to consider the potential impact on driver safety if there is a delay in detecting fatigue.

In another study, Jing et al. aimed to detect driving fatigue in low-voltage and hypoxia plateau environments using subjective and objective monitoring methods [10]. EEG signals from real-time driving tests were subjected to nonlinear and linear analyses to assess the signal trend during awake, critical, and fatigue states. The $(\alpha+\theta)/\beta$ and $(\alpha+\beta)/\theta$ energy features were identified as potential markers of driving fatigue in these environments, providing a basis for the development of a driving fatigue warning system. However, the study was limited to field driving fatigue tests in a specific environment, and further research is needed to validate the findings in other environments and driving conditions.

Additionally, Zhang et al. proposed an innovative approach known as clustering on brain networks (CBNs) to improve the performance of driver fatigue detection [11]. The CBNs approach employs a clustering algorithm to identify spatial nodes with unique connectivity features from electroencephalogram (EEG) data. The wavelet entropy features obtained from these nodes are then transformed into spatiotemporal images and examined using an image edge detection technique to differentiate between various stages of fatigue. This method reduces signal interference and detects fatigue before the onset of subjective feelings, making it a potentially useful tool for early warning and accident prevention. The research demonstrated the limitations of using EEG indicators in time and frequency domains for reliable detection of driver fatigue due to the challenge of signal mixing and limited sample size, lacking comparison with existing methods and validation in real-world driving scenarios. Then, the previous researcher proposed an intelligent system for automated driver fatigue detection utilising EEG signals [12]. This system comprises a feature generation network that utilises texture descriptors and a hybrid feature selection method to enhance detection accuracy. The proposed framework achieved an impressive classification accuracy of 97.29% for detecting fatigue using EEG signals, highlighting its potential for efficient driver fatigue detection. However, the proposed framework used traditional machine learning algorithms, which may limit its ability to adapt to complex and dynamic driving environments.

The proposed research introduces a novel approach for efficiently detecting driver fatigue using EEG signals [13]. The method employs a new channel selection algorithm based on correlation coefficients, an ensemble classifier using random subspace k-nearest neighbours (k-NN), and power spectral density (PSD) for feature extraction. The approach achieved an impressive accuracy of 99.99% for identifying driver fatigue using EEG signals in a 0.5-second time window. The proposed method demonstrates strong performance and can effectively detect EEG-based driver fatigue. However, due to its high computational complexity, a k-NN-based ensemble classifier may not be suitable for real-time applications. Hwang et al. proposed a subject-independent EEG-based driver fatigue state classification model in another study that addresses individual performance gaps [14]. The authors utilised an adversarial training approach to induce the misclassification of subject labels in the classification model. Additionally, they incorporated an Inter-subject Feature Distance Minimization (IFDM) technique to minimise performance discrepancies between individuals. Their method enabled training on EEG datasets with limited, subject labels and was evaluated on the SEED-VIG dataset, resulting in superior accuracy and decreased individual performance variability when classifying drowsiness. However, one of the major drawbacks is that EEG signals contain large differences between individuals, making it challenging to build a unified model that can perform well for all individuals.

The studies reviewed propose various methods for detecting driver fatigue using EEG signals, ranging from single-channel feature extraction to more complex machine learning models. One common approach involves using power spectral density and various entropy measures as feature sets, while others utilise clustering algorithms and image edge detection to distinguish different stages of fatigue. Several studies also address individual performance gaps and subject variability by employing adversarial training strategies and component-specific batch normalisation. These studies demonstrate the potential of EEG-based driver fatigue detection for early warning and accident prevention, achieving high accuracies and providing new possibilities for extracting more information from complex EEG data. However, the methods vary in computational complexity, the number of channels required, and the level of subject independence achieved, suggesting that further research is needed to identify the most efficient and effective approach for practical applications.

Wilapiprasitporn et al. proposed a deep learning approach that combines CNN and RNN to identify individuals using affective EEG data [15]. Their study used the Database for Emotion Analysis using Physiological Signals (DEAP) dataset and showed that the proposed method outperforms an SVM baseline system with a Correct Recognition Rate (CRR) of up to 99.90-100%. Recent research suggests that CNN-GRU models outperform CNN-LSTM models in identifying individuals using EEG data from the brain's frontal region, and they are effective at countering the impact of affective states. However, the proposed method relies on EEG signals, which may require specialised equipment and data collection and analysis expertise. Qin et al. proposed a deep learning model that combines CNN and LSTM to extract vein features from raw images for finger-vein biometrics [16]. The proposed model uses supervised encoding to eliminate binary vein texture, resulting in significantly improved verification accuracy when evaluated on a publicly available finger-vein database. However, deep learning models are prone to overfitting, learning the training data too well and failing to generalise to new data. Techniques such as regularisation and dropout can help prevent overfitting.

Mondal et al. developed a multitask learning framework using a CNN and a bidirectional long short-term memory (Bi-LSTM) model to analyse surgical workflows from video data [17]. Their framework included a joint distribution loss function for concurrent tool usage during phase identification. The proposed method demonstrated excellent tool and phase identification performance compared to previous approaches when evaluated on the Cholec80 dataset. However, the limitation of this study was that it was only evaluated on a single dataset, and it is unclear how well the proposed approach would generalise to other surgical datasets. Hu et al. proposed the Deep Complex Convolution Recurrent Network (DCCRN), a network architecture that can handle both CNN and RNN structures and replicate complex-valued operations [18]. In the Interspeech 2020 Deep Noise Suppression (DNS) challenge, DCCRN outperformed previous networks based on objective and subjective metrics and obtained the top rank for the real-time track and the second rank for the non-real-time track based on Mean Opinion Score (MOS). The proposed DCCRN network with 3.7M parameters proved highly effective in this task. However, the study focused on speech enhancement in clean environments and did not consider noisy or reverberant conditions common in real-world scenarios.

Researchers proposed a machine learning model that utilised CNN, U-net architecture, RNN, and LSTM architecture to create structural topology configurations that fulfilled minimum compliance and deformation criteria under various load conditions and volume fraction limitations. The model was trained using randomly generated finite element simulation data and a strategy to remove elements during training. The model outperformed traditional methods regarding time, cost, and practicality when applied to two-dimensional

and three-dimensional cantilever-beam structural topology designs. This data-driven approach can speed up preliminary structural design procedures [19]. However, the study's limitations include the need for training data and the lack of validation on real-world applications. Later, other researchers focused on improving solar radiation estimation models in agriculture meteorology due to limited data availability and low data quality [20]. Several neural network models (SVM, Extreme Learning Machine, CNN, and LSTM) were developed and tested in Southern Spain using different input variable configurations. Performance was analysed using various statistical indices. One limitation of this study is that it only focused on using temperature and relative humidity as input variables for solar radiation estimation. Other climatic variables that can affect solar radiation, such as atmospheric pressure, cloud cover, and wind speed, which were not included in this study. Incorporating these variables could potentially improve the accuracy of solar radiation estimation.

The previous works discussed different deep learning approaches for various applications, including affective EEG-based person identification, finger-vein biometrics, surgical workflow analysis, speech enhancement, and structural topology design. The proposed models showed significant accuracy, efficiency, and applicability improvements over previous methods. Different deep learning architectures, such as CNNs, RNNs, and LSTM, extracted features from raw data, such as EEG signals, video data, and simulation data. The models were evaluated on different datasets and achieved state-of-the-art results regarding recognition rate, mean average precision, and mean opinion score. Additionally, deep learning models were used to improve solar radiation estimation models in agriculture meteorology.

In conclusion, an outlier detection is a valuable tool for identifying anomalies in data. However, its limitations must be carefully considered, such as the lack of a clear definition for what constitutes an outlier, the possibility of false positives or false negatives, and the challenge of choosing the appropriate method for a specific dataset or problem. EEG-based driver fatigue detection has shown great potential for early warning and accident prevention using various deep learning methods, achieving high accuracies and extracting more information from complex EEG data. Moreover, deep learning has significantly improved accuracy, efficiency, and applicability for various applications, such as affective EEG-based person identification, finger-vein biometrics, surgical workflow analysis, speech enhancement, and structural topology design. Further research is needed to identify the most efficient and effective approach for practical applications in outlier detection and deep learning.

3. METHODOLOGY

The research design is structured into four major components, data acquisition, pre-processing, classifiers, and evaluation metrics, as illustrated in Fig. 1. The successful implementation of independent component analysis (ICA) was critical in enhancing frequency resolution and achieving improved energy conservation results in the proposed study. To eliminate unwanted noises and artefacts, such as muscle movements and eye blinks, ICA was employed as a pre-processing step for EEG signals. Additionally, ICA was utilised to obtain EEG amplitude and correctly position the signal in the appropriate coordinate system for further analysis. The resulting clean EEG signals were then divided into alpha, delta, and theta sub-bands.
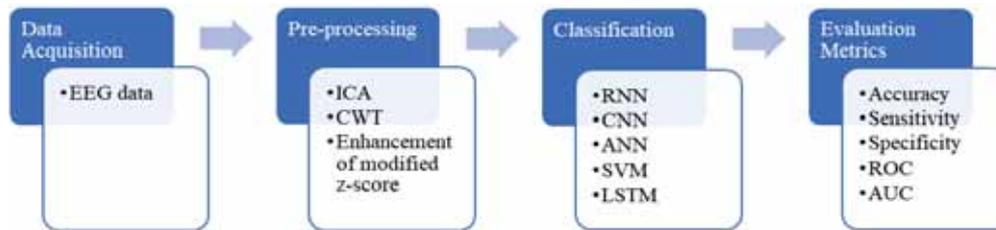
Fig. 1: The framework of enhancement of modified z-score.

This study [21] employed continuous wavelet transform (CWT) as the preferred technique for time-frequency domain analysis in the feature extraction stage. The Morlet wavelet was chosen as the mother wavelet, a combination of a complex sinusoid and a Gaussian envelope with a time scale of t and exhibited an inverse relationship between scale and frequency, leading to an increase in frequency as the scale decreased. The researchers then enhanced statistical features using the modified z-score in the feature extraction process to improve the accuracy of the classification process. In the final step, the classification process involved using CNN, RNN, and ANN, including LSTM, and machine learning methods like SVM using a 5-fold cross-validation strategy to distinguish between fatigue and normal states.

3.1 Data Acquisition

The dataset used in this study was obtained from a previous researcher's online database [22]. The dataset consisted of EEG recordings from 12 healthy male participants (19-24) who completed a driving simulator task for up to 2 hours. EEG data from eight specific channels (O1, O2, Fp1, Fp2, P3, P4, F3, and F4) were selected from a Neuroscan device that had 30 electrodes and operated at a sampling rate of 1000 Hz. The study was divided into two phases: a 5-minute normal and a 5-minute fatigued state. Fatigue was self-reported by the participants after 40-100 minutes of driving. The ZY-31D driving simulator, which features a wide-screen display consisting of three 24-inch screens, was utilised in the study. The driving environment was created using a Peking Ziguangjiye program ZG-601, resulting in a low traffic density scenario.

This study's driver fatigue detection system focused on EEG channel O1, which exhibited the highest correlation among the selected channels. Electrodes O1, O2, and Fp1 were chosen based on their correlation with fatigue and drowsiness. Combining Fp1 and Fp2 achieved an accuracy of 85% in classifying fatigue driving, which was higher than using Fp1 alone (79%) or Fp2 alone (68%) [23]. The EEG electrode F4 was chosen for analysis due to its high performance in the classification process. In addition, electrode P4 positively impacted drowsiness and poor driving performance. Other electrodes, such as O2, Fp2, F3, and P3, were also included for further research. Previous studies have used ten EEG channels, including Fp1, Fp2, F3, and F4, to obtain the best results [24]. According to the study, the electrodes Fp1 and P3 were the most effective for driver fatigue detection. Interestingly, these electrodes have also been successful in EEG emotion recognition in previous studies.

3.2 Pre-processing

The purpose of using MATLAB software is to create programs. Raw EEG signals typically contain interference and noise that must be removed before processing. In order to accomplish this, the independent component analysis (ICA) method was employed during the pre-processing stage. This method can eliminate unwanted noises and artefacts and

separate source signals from observed signals, all without prior knowledge of the mixture. The ICA mixture model can be represented as a vector matrix.

$$x = As \qquad\qquad (1)$$

Matrix A represents the mixing matrix or the linear combination of independent components that contribute to the observed mixed signals in EEG data. Each row in matrix A corresponds to a unique independent component, and the columns represent different time points. On the other hand, the observed mixed signals are represented by the rows in matrix s. Each row in s corresponds to a specific time point in the EEG data, and the columns represent different electrodes or sensors used to measure the electrical activity in the brain. In essence, matrix A and matrix s represent two views of the same EEG data. Matrix A provides information about the underlying independent components contributing to the observed signals, while matrix s includes information about the observed signals at each electrode or sensor.

To reduce redundant data in a dataset, a feature extraction technique known as continuous wavelet transform (CWT) creates a time-frequency distribution. In order to identify signs of fatigue, the signal was partitioned into distinct sub-bands, namely alpha (8-13 Hz), delta (0.5-4 Hz), and theta (4-8 Hz), that are recognised as significant in this regard. The CWT is a highly effective method for these tasks because it has superior computing performance, provided that enough wavelets of analytics are used. The CWT can be defined as:

$$CWT\{h(x)\} = S_{CWT}(x, \omega) = \int_{-\infty}^{\infty} h(\varepsilon)\psi * \big((\varepsilon - x)\omega\big)d\varepsilon \qquad\qquad (2)$$

The output obtained from this method is a spatially dependent spectral decomposition that exhibits the spectral response for both spatial frequency (ω) and spatial area (x). The CWT output is typically represented as a two-dimensional plot, where the x-axis represents time or position, and the y-axis represents frequency or scale. The plot shows the wavelet coefficients as a function of time and frequency, with higher coefficients indicating more robust signal content at that time and frequency. This plot is often referred to as a spectrogram or a scalogram.

The mother wavelet, denoted as ψ, plays a crucial role in the CWT process. It determines the shape and frequency characteristics of the wavelet used in the transformation and serves as the basis for analysing signals in various fields such as engineering, physics, and finance. The choice of the mother wavelet can significantly impact the results of the CWT analysis, as different wavelets are better suited for different types of signals or applications. In addition, the selection of the mother wavelet can also affect the computational efficiency and accuracy of the CWT algorithm. A wide range of mother wavelets are available, each with its unique advantages and disadvantages, and researchers continue to develop new wavelets with improved performance in various applications. As such, the study of mother wavelets remains an active and important area of research in signal processing.

In the context of EEG analysis, CWT can be used to analyse the spectral content of neural activity at different spatial locations on the scalp. By applying CWT to EEG data from multiple electrodes, researchers can generate a map of the spectral response for different spatial frequencies and locations, providing insights into the spatial distribution of neural activity at different frequency scales.

The paper proposes using a modified z-score to enhance statistical features in the classification process. Unlike the traditional z-score, which can be unreliable in the presence

of outliers, the modified z-score uses the median and median absolute deviation (MAD) instead of the mean and standard deviation to detect outliers in a dataset. The modified z-score is less affected by extreme values and is particularly useful for datasets with non-normal distributions or with extreme values. The formula for calculating the modified z-score is:

$$Modified\ z - score = k * (Xi - Median)/MAD \qquad (3)$$

where k is the constants, Xi is the observation, the median is the median of the dataset, and MAD is the median absolute deviation. The constant, k, was modified from 0.6745 to 0.33725 to account for the larger input data spread, making the modified z-score comparable to the z-score for normally distributed data. The larger spread of input data can be accounted for by lowering the value of k. When the value of k is reduced, the range of acceptable data points also decreases. This means that the modified z-score becomes less sensitive to outliers or extreme values, which may be more likely to occur in data with a larger spread.

By decreasing the value of k, the modified z-score can be more comparable to the z-score for normally distributed data. This can help ensure the analysis is accurate and reliable, particularly when working with large or complex data sets. Then, further improvement of the modified z-scored data is multiplied by a binary mask, which assigns a value of 1 to rows that correspond to the fatigue data and a value of -1 to the rows that correspond to the non-fatigue data. This allows the non-fatigue and fatigue data to be easily separated based on their modified z-scores, with the fatigue data having positive values and the non-fatigue data having negative values.

The modified z-score is a statistical technique commonly used to identify dataset outliers. It is more robust than other outlier detection methods because it is less sensitive to extreme values. This is particularly important in data analysis and quality control applications, where outliers can significantly impact the overall analysis or conclusions drawn from the data. The modified z-score is considered a more reliable approach to identifying outliers because it uses the median absolute deviation (MAD), a robust measure of variability not affected by extreme values. In contrast, traditional z-score methods use the sample standard deviation, which can be heavily influenced by extreme values and may not accurately represent the variability of the dataset. Therefore, the modified z-score is often preferred over other outlier detection methods in situations where robustness and reliability are critical, such as analysing EEG signals for driver fatigue detection.

The modified z-score is a powerful tool that enables researchers and analysts to identify and handle outliers more reliably and consistently. Using a robust measure of variability not influenced by extreme values, the modified z-score can accurately detect and flag outliers in a dataset. This, in turn, allows analysts to handle outliers more effectively, either by excluding them from the analysis or by applying appropriate statistical techniques to account for them. As a result, using a modified z-score can lead to more accurate and reliable statistical analysis and better-informed decision-making. In addition, the ability to detect outliers reliably and consistently is particularly important in fields where the impact of outliers can be significant, such as in EEG signal analysis for driver fatigue detection. Using the modified z-score, researchers and analysts can ensure that their analyses are robust and reliable and that conclusions drawn from the data are based on accurate and representative information.

3.3 Classifiers

On the other hand, RNNs, including LSTM, are better suited for sequential data, such as time series data, and have been used for fatigue detection based on physiological signals,

such as EEG and ECG. These signals provide valuable insights into the driver's physiological state and can help detect the onset of fatigue before the driver becomes visibly drowsy or sleepy. Additionally, ANNs and SVMs have been used for fatigue detection based on physiological and behavioural data, such as steering wheel movements, vehicle speed, and lane deviations. Overall, the use of machine learning and deep learning methods for fatigue detection shows great promise for improving road safety. These automated systems can provide real-time feedback to drivers and alert them when they are at risk of falling asleep at the wheel, thereby preventing accidents and saving lives.

Sequential data, like time-series data, is best processed by RNNs because they can identify the temporal dependencies between consecutive frames of video or physiological signals. This is particularly important in detecting driver fatigue. Meanwhile, ANN is a versatile type of neural network that can be utilised for various tasks, such as classification. Our study divided the dataset into two subsets for the ANN classifier, 80% for training and 20% for testing. ANN is often used as a baseline for comparing with other deep-learning models and is an appropriate starting point for developing a fatigue detection system [25]. When dealing with datasets with fewer features, SVM is a popular machine learning algorithm that can be used for classification and regression analysis [11]. Lastly, LSTM is an RNN architecture that solves the problem of vanishing gradients in traditional RNNs by enabling the neurons to give feedback on their output as inputs to the next neuron, allowing RNNs to identify sequential dependencies between different time steps in a time series.

Driver fatigue is a significant issue on the road that can lead to accidents and fatalities. Automated systems that detect and classify driver fatigue in real time are a potential solution to this problem. Deep learning methods like CNN, RNN, and LSTM and machine learning methods like ANN and SVM have shown promise in fatigue detection [26]. These classification algorithms were implemented using Python for its ease of use and flexibility. These classifiers were selected based on their suitability for analysing EEG signals and their proven success in related classification tasks. Precisely, CNNs effectively extract relevant features from multichannel EEG signals, RNNs and LSTMs are well-suited for handling sequential EEG data, and SVMs are known for their ability to handle high-dimensional data and perform well with limited sample sizes. ANNs are versatile and widely used types of classifiers that can be applied to a range of data types. We used multiple classifiers to compare their performances on the EEG dataset and evaluate their suitability for driver fatigue classification.

3.4 Evaluation Metrics

We aimed to evaluate the fatigue detection system's effectiveness using standard metrics typically employed to gauge classifier performance. Classification models in machine learning are commonly evaluated using metrics such as accuracy, sensitivity, and specificity. Accuracy indicates the model's ability to classify instances correctly and is calculated by determining the percentage of correctly classified cases. Sensitivity evaluates the true positive rate by determining the percentage of actual positives correctly identified by the model. Conversely, specificity measures the true negative rate by determining the proportion of real negatives correctly identified by the model. When false positives or negatives may have profound implications, combining sensitivity and specificity can provide a more comprehensive evaluation of the model's performance.

In evaluating the performance of our fatigue detection system, we utilised receiver operating characteristic (ROC) curves and area under the curve (AUC) in addition to standard metrics. ROC curves plot true and false positive rates at varying classification thresholds, providing a comprehensive evaluation of the classifier's performance, mainly

when the costs of false positives and negatives are distinct. Meanwhile, the AUC yields a single value summarising the classifier's overall performance, with a score of 1 denoting perfect classification and 0.5 indicating random classification. By incorporating these metrics, the system's performance could be effectively evaluated and areas for improvement identified.

In order to increase the reliability of these findings, a 5-fold cross-validation approach was employed. This process randomly partitions the dataset into five equally sized subsets, with four subsets utilised for training and one subset reserved for testing. This procedure was repeated five times, with each subgroup used once as a test set, to ensure each fold received equal representation. Each fold's evaluation metrics were calculated, including accuracy, sensitivity, specificity, ROC curves, and AUC, to assess the classifier's performance. By computing these metrics for each fold, the classifier's performance was evaluated more accurately across the entire dataset. Finally, the evaluation metrics were averaged over the five folds for a final classifier performance estimate.

To compare the performance of various classifiers, the Wilcoxon signed-rank test was utilised; this non-parametric statistical test compares paired data sets to determine if there is a significant difference between them. The aim was to compare the performance of different classifiers on the same dataset, considering a statistically significant difference if the p-value was less than 0.05, using a significance level of 0.05. The Wilcoxon signed-rank test was combined with a cross-validation strategy to obtain unbiased and dependable results, allowing the evaluation and comparison of different classifiers' performance objectively.

4. RESULTS AND DISCUSSION

This study analysed the patterns of brainwave activity in a group of participants using electroencephalography (EEG). Specifically, the focus was on the alpha (8-13 Hz), delta (0.5-4 Hz), and theta (4-8 Hz) frequency bands, which have been linked to different cognitive and emotional processes. The data were analysed using statistical and visualisation techniques, including box and scatter plots. The results revealed the presence of outliers in the data, which are data points that fall far outside the expected range and can significantly impact the overall pattern of results. These outliers were particularly evident in the alpha and delta frequency bands, with fewer outliers observed in the theta band.

Outliers are data points that fall far outside the expected range and can significantly impact the overall pattern of results when analysing a data set. One approach to identifying outliers is to use z-scores or modified z-scores, which are measures of how many standard deviations a given data point is away from the mean of the dataset. By plotting these scores visually, it is possible to detect outliers in a dataset and evaluate their impact on the overall data pattern. This method allows the identification of the magnitude of the outliers and assesses their potential influence on the analysis.

Figures 2 and 3 illustrate the distribution of data points, where the blue dots correspond to data points that fall within the expected range, and the red dots represent data points that fall outside the expected range and are considered outliers. In addition to the potential impact of outliers on statistical analyses, it is important to note that outliers can also affect the interpretation and generalisation of research findings. Outliers can lead to overestimation or underestimation of effect sizes, which can have significant implications for the practical significance of research results. Thus, appropriate identification and treatment of outliers

are essential to ensure research findings' integrity and applicability to the broader population or context of interest.
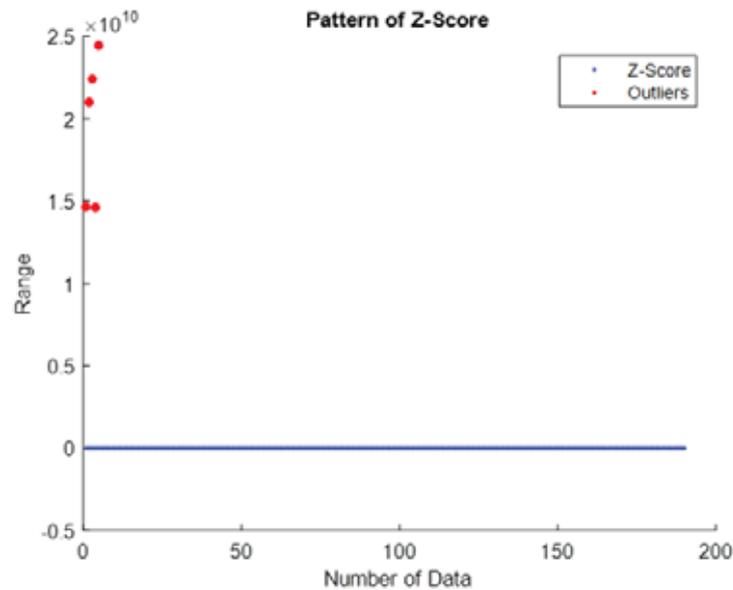
Fig. 2: Identification of outliers in z-score data.
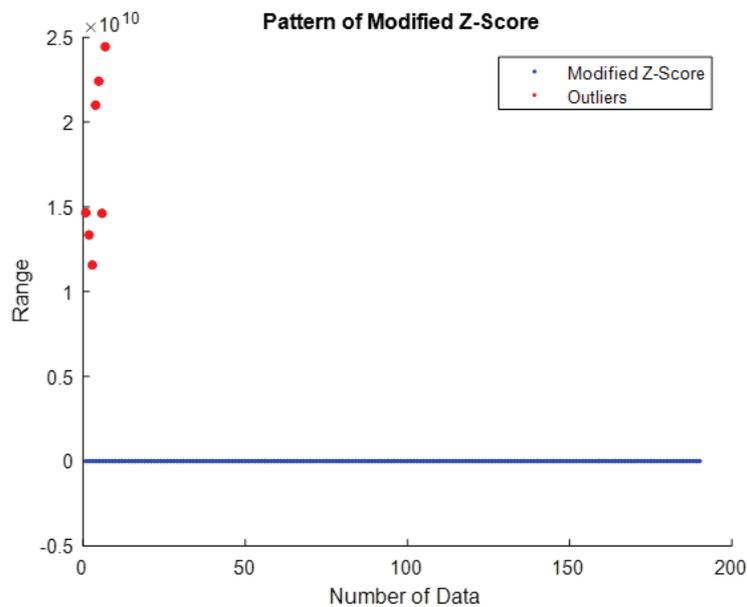
Fig. 3: Identification of outliers in modified z-score data.

In Fig. 4, the proposed method enhancing modified z-scores was influential in ensuring that all data points fell within the expected range, indicating that it successfully addressed the issue of outliers and improved the overall data quality. The method effectively managed and adjusted the data points to achieve the desired outcome without reducing any data classified as outliers during the process. This study evaluated the performance of five popular machine learning models, RNN, CNN, ANN, SVM, and LSTM, to determine the most effective classification task.
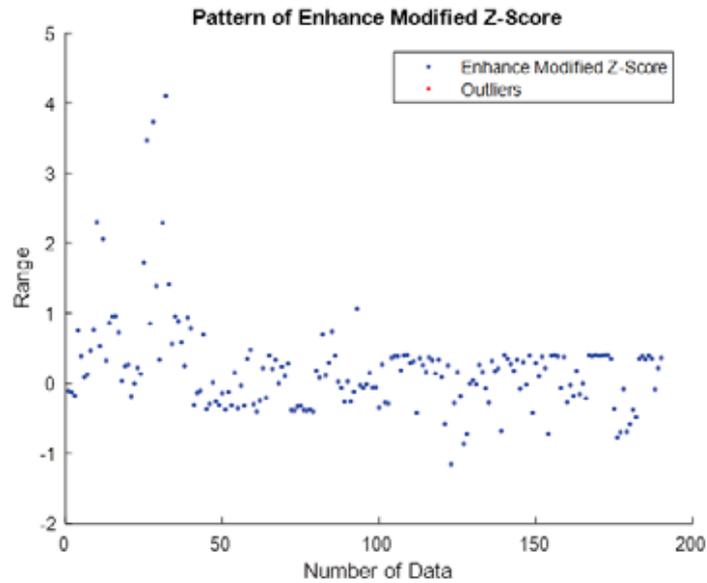
Fig. 4: Identification of outliers in the enhancement of modified z-score data.

The z-score is a method of standardisation that involves subtracting the mean from the data and dividing it by the standard deviation. However, datasets that include outliers, such as driver fatigue detection, may not be well-suited for this method. The MZS process uses the median and median absolute deviation instead of the mean and standard deviation to standardise the data. The MZS is more robust to outliers and has been used in several studies for driver fatigue detection. In this study, we further enhanced the MZS, the so-called enhancement of modified z-score (EMZS), by applying a scaling factor to adjust the sensitivity and specificity of the classifier. This enhancement resulted in the highest accuracy among the five classifiers tested: RNN, CNN, ANN, SVM, and LSTM.

Table 1: Classification results using z-score, modified z-score (MZS), and enhancement of modified z-score (EMZS)

| Classifier	Accuracy (%)			Sensitivity (%)			Specificity (%)		
	Z-Score	MZS	EMZS	Z-Score	MZS	EMZS	Z-Score	MZS	EMZS
RNN	61.69	61.16	93.85	0.00	0.00	84.20	100.00	100.00	100.00
CNN	63.16	64.74	97.19	n/a	n/a	95.68	63.16	66.21	100.00
ANN	63.09	64.15	**99.65**	88.04	84.17	**99.48**	23.02	32.85	**100.00**
SVM	61.69	61.16	97.89	0.00	0.00	96.62	100.00	100.00	100.00
LSTM	63.16	64.91	93.86	n/a	n/a	100.00	63.16	64.91	91.36

Based on the results in Table 1, the ANN performed the best among the five classifiers tested. The ANN demonstrated higher accuracy and specificity than the other classifiers, with values of 99.65% and 100.00%, respectively. However, its sensitivity value of 99.48% was slightly lower than that of different classifiers. Nonetheless, the results indicate that the ANN was still proficient at detecting driver fatigue. This result outperformed the accuracy of 96% and sensitivity of 94% reported by a previous researcher in the field [27]. Therefore, our study demonstrates the effectiveness of the proposed approach for accurate driver fatigue detection. The sensitivity of the classifier CNN and LSTM were not applicable (n/a) in this study due to the absence of negative instances in the test set. As a result, the denominator of the sensitivity calculation was zero, and the sensitivity value was undefined.

However, the high precision and specificity values achieved by the classifier suggest that it effectively identifies positive instances. Implementing the enhanced, modified z-score approach resulted in an accuracy improvement of over 30.00% compared to the traditional modified z-score method. The SVM accuracy result, 97.89%, indicates high accuracy compared to the previous researcher, who used the same machine learning with eight electrode channels [13].

According to the results, the proposed enhanced, modified z-score method was found to be more efficient than the traditional modified z-score approach in identifying driver fatigue from EEG signals. The accuracy of the improved process was significantly better than the conventional method, as indicated by the study's findings. The proposed approach effectively minimises false positives and negatives, which is critical for detecting driver fatigue with high precision. Notably, the accuracy results obtained through the proposed method were competitive with the findings of similar studies that used EEG signals for driver fatigue detection. Furthermore, the proposed framework's multi-classifier approach could be used to improve the detection of other physiological signals related to driver fatigue, such as eye movements and heart rate variability.

ROC curves were utilised to plot and evaluate the performance of each classifier to establish the ideal threshold for classification. The results indicated that the CNN and ANN classifiers exhibited the highest true positive and lowest false positive rates, suggesting they could accurately distinguish between normal and fatigued states. Additionally, the SVM classifier displayed a high true positive rate and a low false positive rate, indicating its effectiveness in classifying driver fatigue.
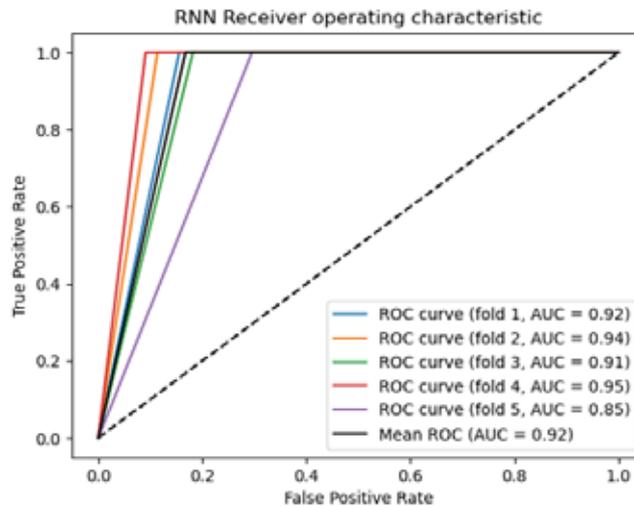
Fig. 5: Performance evaluation of RNN using ROC curve.

In contrast, the LSTM and RNN classifiers had slightly lower true positive rates and higher false positive rates than the other classifiers. This suggests that they were less effective in accurately classifying driver fatigue. Despite these limitations, the LSTM and RNN classifiers still represent promising approaches to detecting driver fatigue and should be further investigated to determine if their performance can be improved. The findings suggest that the CNN and ANN classifiers and the SVM classifier have significant potential for accurately detecting driver fatigue from EEG signals. These results have important implications for reducing the risks associated with drowsy driving and improving overall road safety.

Fig. 6: Performance evaluation of CNN using ROC curve.

Fig. 7: Performance evaluation of ANN using ROC curve.

Fig. 8: Performance evaluation of SVM using ROC curve.

The study's findings indicate that the enhanced modified z-score approach proposed, combined with multiple classifiers, was exceedingly influential in detecting driver fatigue from EEG signals with high accuracy. In particular, the ANN classifier outperformed the other classifiers, exhibiting the highest accuracy and specificity. On the other hand, the CNN and LSTM classifiers achieved perfect classification with the highest area under the curve (AUC) values. The ANN classifier is particularly well-suited for detecting driver fatigue from EEG signals due to its capacity to learn complex nonlinear relationships between input and output variables, flexibility in handling vast amounts of data, and ability to adapt to changing input patterns over time. Therefore, the ANN classifier represents a promising approach to accurately detecting driver fatigue and mitigating the risks associated with drowsy driving. ANN classifier has several benefits but has limitations, such as the need for high-quality input data, significant training data, and customisation to specific contexts. Despite these challenges, the ANN classifier remains a promising approach for accurately detecting driver fatigue.
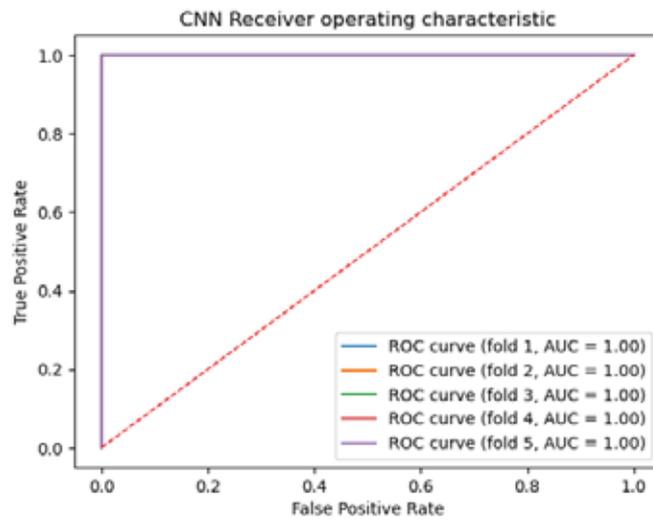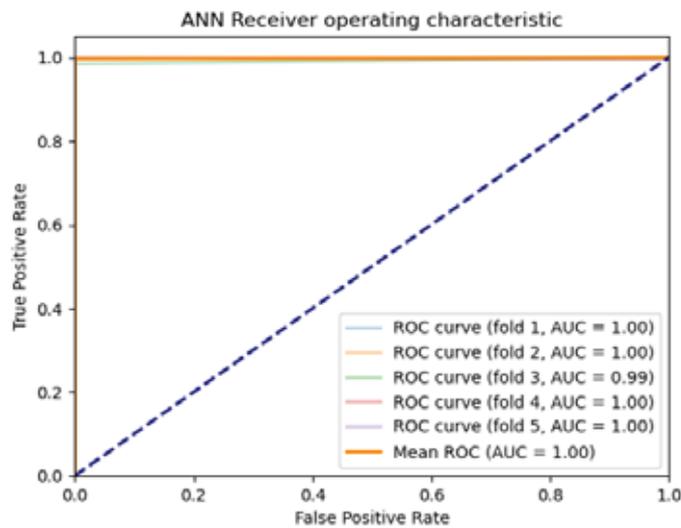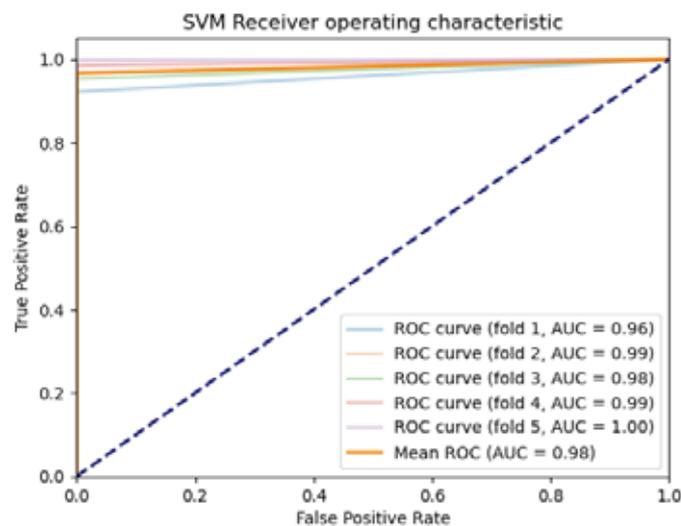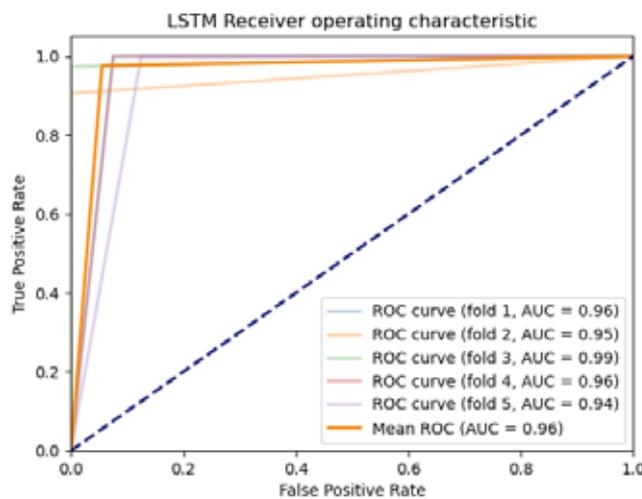

Fig. 9: Performance evaluation of LSTM using ROC curve.

Artificial Neural Networks (ANNs) are inspired by the structure and function of the human brain, consisting of interconnected nodes that process information and transmit signals. By adjusting the weights between nodes based on training data, ANNs can learn to identify complex patterns and make highly accurate predictions. In the case of detecting driver fatigue from EEG signals, ANNs can analyse the intricate patterns of brainwave activity associated with fatigue and accurately distinguish them from normal brainwave patterns. The proposed approach has exhibited significant improvements in accuracy as compared to the traditional modified z-score method, which is an important development in the field of driver fatigue detection. Moreover, the multi-classifier approach has shown potential in detecting other physiological signals related to driver fatigue, thereby enhancing the overall reliability of driver fatigue detection systems. These findings could have far-reaching implications in ensuring road safety by enabling the development of more accurate and reliable driver fatigue detection systems.

5. CONCLUSION

In conclusion, the proposed framework for driver fatigue detection using enhancement of modified z-score and multiple machine learning architectures was highly influential in accurately detecting driver fatigue from EEG signals. Our study demonstrated that the

framework achieved high accuracy for all classifiers, with the ANN achieving the highest accuracy of 99.65%. The SVM and CNN also achieved high accuracy, with 97.89% and 97.19%, respectively. Although the RNN and LSTM achieved slightly lower accuracy, they still achieved over 90.00% accuracy. Our study's contribution lies in presenting a comprehensive and effective framework that can accurately detect driver fatigue from EEG signals, surpassing the performance of previous approaches. Furthermore, the proposed framework can potentially be applied to practical driver fatigue detection systems, improving driving safety and reducing the number of road accidents caused by driver fatigue. Future research could build on this framework by investigating its performance in real-world settings and exploring ways to improve its sensitivity to subtle changes in EEG signals.

REFERENCES

[1] Gao Z, Wang X, Yang Y, Mu C, Cai Q, Dang W, Zuo S. (2019) EEG-Based Spatio-Temporal Convolutional Neural Network for Driver Fatigue Evaluation. IEEE Transactions on Neural Networks and Learning Systems, 30(9): 2755-2763. https://doi.org/10.1109/TNNLS.2018.2886414

[2] Mehreen A, Anwar SM, Haseeb M, Majid M, Ullah MO. (2019) A Hybrid Scheme for Drowsiness Detection Using Wearable Sensors. IEEE Sensors Journal, 19(13): 5119-5126. https://doi.org/10.1109/JSEN.2019.2904222

[3] Hasan MJ, Shon D, Im K, Choi HK, Yoo DS, Kim JM. (2020) Sleep state classification using power spectral density and residual neural network with multichannel EEG signals. Applied Sciences (Switzerland), 10(21): 1-13. https://doi.org/10.3390/app10217639

[4] Cohen MX. (2019) A better way to define and describe Morlet wavelets for time-frequency analysis. NeuroImage, 199: 81-86. https://doi.org/10.1016/j.neuroimage.2019.05.048

[5] Aggarwal V, Gupta V, Singh P, Sharma K, Sharma N. (2019) Detection of spatial outlier by using improved Z-score test. In 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 788-790). Tirunelveli, India.
doi: 10.1109/ICOEI.2019.8862582.

[6] Sandbhor S, Chaphalkar NB. (2019) Impact of Outlier Detection on Neural Networks Based Property Value Prediction. In: Satapathy, S., Bhateja, V., Somanah, R., Yang, XS., Senkerik, R. (eds) Information Systems Design and Intelligent Applications. Advances in Intelligent Systems and Computing, vol 862. Springer, Singapore. https://doi.org/10.1007/978-981-13-3329-3_45

[7] Leite GNP, Araújo AM, Rosas PAC, Stosic T, Stosic B. (2019) Entropy measures for early detection of bearing faults. Physica A: Statistical Mechanics and its Applications, 514: 458-472. https://doi.org/10.1016/j.physa.2018.09.052

[8] Hosseini M-P, Hosseini A, Ahi K. (2021) A Review on Machine Learning for EEG Signal Processing in Bioengineering. IEEE Reviews in Biomedical Engineering, 14: 204-218. https://doi.org/10.1109/RBME.2020.2969915

[9] Quintero-Rincon A, Fontecha ME, D'Giano C. (2019) Driver fatigue EEG signals detection by using robust univariate analysis. Retrieved from http://arxiv.org/abs/1912.13351

[10] Jing D, Liu D, Zhang S, Guo Z. (2020) Fatigue driving detection method based on EEG analysis in low-voltage and hypoxia plateau environment. International Journal of Transportation Science and Technology, 9(4): 366-376. https://doi.org/10.1016/j.ijtst.2020.03.008

[11] Zhang C, Sun L, Cong F, Kujala T, Ristaniemi T, Parviainen T. (2020) Optimal imaging of multichannel EEG features based on a novel clustering technique for driver fatigue detection. Biomedical Signal Processing and Control, 62: 102103. https://doi.org/10.1016/j.bspc.2020.102103

[12] Tuncer T, Dogan S, Ertam F, Subasi A. (2021) A dynamic center and multi threshold point based stable feature extraction network for driver fatigue detection utilising EEG signals. Cognitive Neurodynamics, 15(2): 223-237. https://doi.org/10.1007/s11571-020-09601-w

[13] Rashid M, Mustafa M, Sulaiman N, Abdullah NRH, Samad R. (2021) Random subspace K-NN based ensemble classifier for driver fatigue detection utilising selected EEG channels. Traitement Du Signal, 38(5): 1259-1270. https://doi.org/10.18280/ts.380501

[14] Hwang S, Park S, Kim D, Lee J, Byun H. (2021) Mitigating Inter-Subject Brain Signal Variability FOR EEG-Based Driver Fatigue State Classification. ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 990-994. https://doi.org/10.1109/ICASSP39728.2021.9414613

[15] Wilaiprasitporn T, Ditthapron A, Matchaparn K, Tongbuasirilai T, Banluesombatkul N, Chuangsuwanich E. (2020) Affective EEG-Based Person Identification Using the Deep Learning Approach. IEEE Transactions on Cognitive and Developmental Systems, 12(3): 486-496. https://doi.org/10.1109/TCDS.2019.2924648

[16] Qin H, Wang P. (2019) Finger-vein verification based on LSTM recurrent neural networks. Applied Sciences (Switzerland), 9(8): 1-18. https://doi.org/10.3390/app9081687

[17] Mondal SS, Sathish R, Sheet D. (2019) Multitask Learning of Temporal Connectionism in Convolutional Networks using a Joint Distribution Loss Function to Simultaneously Identify Tools and Phase in Surgical Videos. 1–15. Retrieved from http://arxiv.org/abs/1905.08315

[18] Hu Y, Liu Y, Lv S, Xing M, Zhang S, Fu Y, … Xie L. (2020). DCCRN: Deep complex convolution recurrent network for phase-aware speech enhancement. Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2020-Octob, 2472-2476. https://doi.org/10.21437/Interspeech.2020-2537

[19] Qiu C, Du S, Yang J. (2021) A deep learning approach for efficient topology optimisation based on the element removal strategy. Materials and Design, 212: 110179. https://doi.org/10.1016/j.matdes.2021.110179

[20] Bellido-Jiménez JA, Estévez J, García-Marín AP. (2021) Assessing Neural Network Approaches for Solar Radiation Estimates Using Limited Climatic Data in the Mediterranean Sea. 19. https://doi.org/10.3390/ecas2020-08116

[21] Abdubrani R, Mustafa M, Zahari ZL. (2023) Enhancement of Morlet Mother Wavelet in Time–Frequency Domain in Electroencephalogram (EEG) Signals for Driver Fatigue Classification . In: , et al. Advances in Intelligent Manufacturing and Mechatronics. Lecture Notes in Electrical Engineering, vol 988. Springer, Singapore. https://doi.org/10.1007/978-981-19-8703-8_13

[22] Kamti MK, Iqbal R. (2022) Evolution of Driver Fatigue Detection Techniques – A Review From 2007 to 2021. Transportation Research Record, 2676(12): 485-507. https://doi.org/10.1177/03611981221096118

[23] Luo H, Qiu T, Liu C, Huang P. (2019) Research on fatigue driving detection using forehead EEG based on adaptive multi-scale entropy. Biomedical Signal Processing and Control, 51: 50-58. https://doi.org/10.1016/j.bspc.2019.02.005

[24] Islam MR, Ahmad M. (2019) Wavelet Analysis Based Classification of Emotion from EEG Signal. 2nd International Conference on Electrical, Computer and Communication Engineering, ECCE 2019, 1-6. https://doi.org/10.1109/ECACE.2019.8679156

[25] Xie Y, Oniga S. (2020) A view of Processing Methods and Classification Algorithm for EEG Signal. Carpathian Journal of Electronic and Computer Engineering, 13(1): 23-29. https://doi.org/10.2478/cjece-2020-0004

[26] Sarkar A, Singh A, Chakraborty R. (2022) A deep learning-based comparative study to track mental depression from EEG data. Neuroscience Informatics, 2(4): 100039. https://doi.org/10.1016/j.neuri.2022.100039

[27] Parekh V, Shah D, Shah M. (2020) Fatigue Detection Using Artificial Intelligence Framework. Augment Hum Res 5: 5. https://doi.org/10.1007/s41133-019-0023-4

3D COLLISION AVOIDANCE SYSTEM FOR UNMANNED AERIAL VEHICLE (UAV) WITH DECENTRALIZED APPROACH

**MOHAMAD HANIFF HARUN[1,2,3]*, SHAHRUM SHAH ABDULLAH[1],
MOHD SHAHRIEEL MOHD ARAS[3], MOHD BAZLI BAHAR[1,3]
AND FARIZ ALI@IBRAHIM[3]**

*[1]Department of Electronic System Engineering,
Malaysia-Japan International Institute of Technology,
UTM KL Campus, Jalan Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia
[2]Faculty of Electric and Electronic Engineering Technology,
Universiti Teknikal Malaysia Melaka,
Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
[3]Underwater Technology Research Group (UTeRG),
Center for Robotics and Industrial Automation (CERIA), Faculty of Electrical Engineering,
Universiti Teknikal Malaysia Melaka, 76100 Durian Tunggal, Melaka, Malaysia.*

*Corresponding author: haniff@utem.edu.my

ABSTRACT: Unmanned aerial vehicles UAVs have been developed and refined for decades. Using an integrated software system, autonomous unmanned aerial vehicles (UAVs) perform missions automatically and return to a pre-programmed point. Malaysia has a lot of unoccupied airspace, yet autonomous UAV applications and research are still rare. In critical conditions, autonomous UAVs must deal with a variety of environmental and flight issues. This project involves a decentralized 3D collision avoidance system for an autonomous UAV. Ultrasonic, infrared, and laser rangefinders were chosen for the 3D collision avoidance system. The UAV's obstacle recognition and collision avoidance performance are also tested in four experiments. In various flight conditions, the 3D collision avoidance system can identify several material types and opacities by integrating selected rangefinders. Finally, the 3D collision avoidance system quickly reacts to obstacles in the X, Y, and Z axes.

ABSTRAK: Kenderaan udara tanpa pemandu (UAV) telah dibangunkan dan diperhalusi selama beberapa dekad. Menggunakan sistem perisian bersepadu, kenderaan udara tanpa pemandu (UAV) autonomi melaksanakan misi secara automatik dan kembali ke titik pra-diprogramkan. Malaysia mempunyai banyak ruang udara yang tidak berpenghuni, namun aplikasi dan penyelidikan UAV autonomi masih jarang berlaku. Dalam keadaan kritikal, UAV autonomi mesti menangani pelbagai isu alam sekitar dan penerbangan. Projek ini melibatkan sistem pengelakan perlanggaran 3D terpencar untuk UAV autonomi. Pencari jarak ultrasonik, inframerah dan laser telah dipilih untuk sistem pengelakan perlanggaran 3D. Prestasi pengecaman halangan dan pengelakan perlanggaran UAV juga diuji dalam empat eksperimen. Dalam pelbagai keadaan penerbangan, sistem pengelakan perlanggaran 3D boleh mengenal pasti beberapa jenis bahan dan kelegapan dengan menyepadukan pencari jarak terpilih. Akhir sekali, sistem pengelakan perlanggaran 3D bertindak balas dengan cepat terhadap halangan dalam paksi X, Y dan Z.

1. INTRODUCTION

The IR 4.0 is increasing interest in autonomous UAVs. A UAV is an aerial vehicle that may be piloted from a distance [1]. UAVs are of use in military and civilian/commercial applications due to their wide range, inexpensive maintenance, quick deployment, mobility, and ability to hover [2]. The military uses UAVs for border security surveillance, reconnaissance, and target removal. Unmanned aerial vehicles are also used for search and rescue, parcel delivery, precision horticulture, and pharmaceutical transport. Multi-rotor drones, fixed-wing drones, single-rotor helicopters, and fixed-wing Vertical Take-Off and Landing (VTOLs) are the four main drone types, as shown in Fig. 1.

The fact remains that each key drone type has its own benefits and drawbacks. For example, multi-rotor drones may hover due to their vertical direction drive framework, while single-rotor helicopters can hover and rotate due to their vertical direction drive framework. To put it another way, this type of aerial vehicle has slow movement speeds and requires extra energy to fulfil tasks [3]. However, fixed-wing drones and fixed-wing hybrid VTOLs can travel long distances due to their propulsion system and aerodynamic surfaces. Fixed-wing VTOLs (vertical takeoff and landing) can launch and land vertically without a runway [3]. However, fixed-wing drones have many drawbacks. Fixed-wing aircraft require a lot of airspace to maneuver and direct themselves. Other than that, the aerodynamic wing surfaces produce lifting power through air impacts.

UAVs have both military and commercial uses. They must be built to work in difficult settings, such as huge open spaces, dense tree clusters, or rocky mountain slopes [4]. As a result, UAVs will struggle to achieve their missions in these hostile environments. Outdoor operation of UAVs is even more problematic due to GPS errors, poor communication, and bad weather [5]. Sensors enable UAVs to scan their surroundings and avoid collisions. Collecting accurate environmental data is important to avoiding crashes and completing the mission. Due to the increasing utilization of multi-UAV cooperative operations, 3D collision avoidance is crucial. When functioning centrally, an unmanned aerial vehicle assesses all colleagues' inputs and outputs to avoid collisions. While a central communication hub is beneficial, it is difficult to expand [6-8]. Decentralized systems can also handle more UAV teams in congested situations. These are more stable and robust.

Malaysia is believed to be falling behind in aerial vehicle research and development compared to countries like the US, China, and Iran. Limitations include lack of aeronautical knowledge, high sensor and communication system costs, and limited airspace. To boost Malaysia's aeronautics industry, aerial vehicle research and development must continue. Currently, only a few Malaysian aerospace companies are researching and developing autonomous UAVs to solve industrial and military concerns.

In many situations, rapid deployment of aerial vehicles can change a crisis situation into a non-critical situation since they can gather information faster than humans. Wildfires, medication delivery, crime fighting, and search and rescue are examples [9]. Wildfires, like the 2019 Amazon rainforest fires, can erupt at any time throughout the dry season [10]. Also, the Amazon rainforest covers 5.5 million km^2, making human patrolling and scouting impractical. Using flying vehicles to monitor the Amazon rainforest could help limit the damage and mortality caused by wildfires.

Fig. 1: Types of drones.

2. THEORETICAL BACKGROUND

Numerous theories and principles must be explored and incorporated into the design and development of the unmanned aerial vehicle with decentralized 3D collision avoidance system. Aerodynamic forces, center of gravity (CG), and time of flight (ToF) are all significant theories and principles of flight.

2.1 Aerodynamic Force

All aircraft in flight are subjected to the forces of thrust, drag, lift, and weight. Understanding how these forces interact and utilizing power and flight controls to manage them is critical for flight. Thrust, drag, lift, and weight are the four forces that act on an airplane in level, unaccelerated flight [11]. These are the terms that are used to describe them:

a) Thrust—the force applied forward by the engine/propeller or rotor. It opposes or overcomes drag. By and large, it acts perpendicular to the longitudinal axis.

b) Drag—a rearward, retarding force created by the wing, rotor, fuselage, and other projecting items disrupting airflow. Drag, on the other hand, opposes thrust and acts in the opposite direction of the relative wind.

c) Lift—is a force generated by the dynamic action of the air on the airfoil that acts perpendicular to the flight path through the center of lift (CL) and perpendicular to the lateral axis. Lift opposes the downward force of weight in level flight.

d) Weight—the total weight of the aircraft, crew, fuel, and cargo or baggage. Weight is a force that acts as a drag on the aircraft due to the force of gravity. It acts vertically downward through the aircraft's center of gravity (CG) and opposes lift.

The sum of these opposing pressures is always zero in steady flight. According to Newton's Third Law, there can be no unbalanced forces in steady, straight flight, because every action or force has an equal but opposing response or force. This is true regardless of whether you are flying level or ascending or descending.

This is not to say that the four forces are equal. This signifies that the opposing forces are equivalent to one another and hence cancel out their effects. In Fig. 2(a), the thrust, drag, lift, and weight force vectors appear to be identical in magnitude. The conventional explanation claims (without specifying) that thrust equals drag and weight equals lift. While this statement is true, it might be misleading. It should be known that the opposing

lift/weight forces are equal in straight, level, unaccelerated flight. Additionally, they are greater than the opposing thrust/drag forces, which are equal only to one another.

This revision of the conventional "thrust equals drag; lift equals weight" formula explains that during climbs and slow flight, a component of force is directed upward and functions as lift, while a piece of weight is directed backward in the opposite direction of flight and acts as drag. In sluggish flight, thrust is directed upward. However, because the aircraft is in level flight, weight has no effect on drag, as illustrated in Fig. 2(b).

(a)   (b)
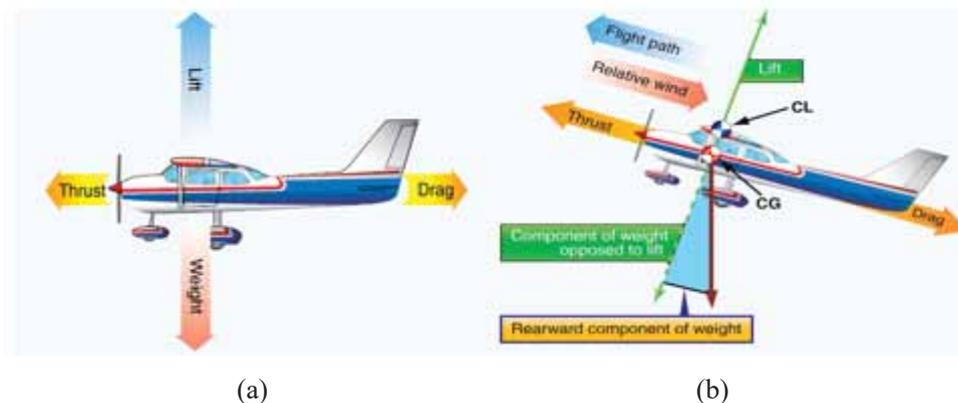
Fig. 2: (a) Relationship of forces acting on an aircraft (b) Force vectors during a stabilized climb.

2.2 Centre of Gravity (CG)

A body in a gravitational field has an unlimited number of particles of varying sizes, each with its own weight. The total weight of the body is the resultant parallel force system created by those weights [12]. Moreover, when the gravitational field is uniform over a body, the center of gravity is also the center of mass. The UAV's center of gravity influences the UAV's balance while in flight. To improve overall flying performance and minimize any instability issues, the UAV's center of gravity must be on the same horizontal plane as the propellers.

2.3 Vibration

Vibration is the oscillating motion of a system or body of attached bodies [13]. Free vibration and forced vibration are the two main types of vibration. Forced vibration is created by an external periodic or intermittent force, whereas free vibration is caused by gravitational or elastic restoring forces. A UAV's vibrations come from its aerodynamic, mechanical, and flying motions. Due to electronic components like flight controllers, inertial measurement units (IMUs), and sensors are susceptible to vibrations, multirotor UAVs have motor-propeller produced vibrations. It is also vital to consider vibrations because a built UAV vibrating at its normal frequency would be disastrous. So, it is critical to isolate electrical components from vibrations and limit UAV vibrations as much as feasible.

2.4 Time of Flight (ToF)

The Time-of-Flight concept measures the distance between a sensor and an item by the time it takes for a signal to travel from the sensor to the object and back. The Time-of-Flight principle works with several signals (carriers), the most prevalent being sound and light [14]. Light is the carrier for TeraRanger sensors because it combines speed, range, weight,

and eye-safety. As a result, infrared sensors perform better than other distance sensors of similar size and weight.

(a)　　　　　　　　　　　　　　　　　　(b)
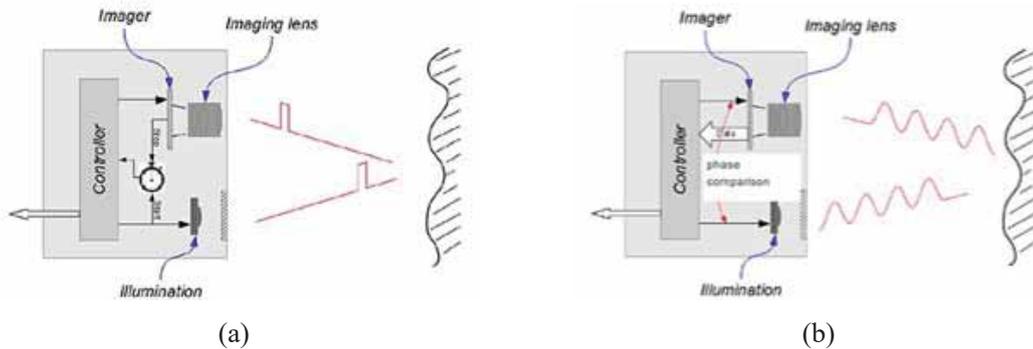
Fig. 3: (a) Direct ToF (b) Indirect ToF.

Time-of-Flight (ToF) sensors detect distances by measuring the time it takes photons to travel from the sensor's emitter to a target and back. As illustrated in Fig. 3, both indirect and direct ToF have particular advantages. Both can measure a pixel's intensity and distance. Direct ToF sensors give out nanosecond-long light pulses and time how long it takes for some of the light to return. Indirect ToF sensors use continuous modulated light to calculate the distance to objects. The object distance from the sensor can be calculated using Eq. 1, where d is the object distance and c is the sound or light speed.

$$d = \frac{c \times ToF}{2}$$

(1)

3. METHODOLOGY

This section contains a mechanical description and the concerns related to the design of the fuselage, landing gear, and propulsion system are described in detail. The electrical and electronic description includes in-depth discussions of the thought processes behind and justifications for the choice of components including the power supply, ESC, flight controller, RC transmission, and range finder. Finally, the experimental procedures are presented, including the setup of all four experiments.

3.1 Mechanical Part - Fuselage

SolidWorks is used to design the multirotor fuselage. The fuselage design defines the type of multirotor aircraft, such as the hexa-copter depicted in Fig. 4(a). Thus, multirotor aircraft have 6 motors and 6 propellers for propulsion. The extra two motors in the Hexa X rotor arrangement provide more lift force. A hexa-copter has higher stability and can continue to fly even if one of its motors fails. Also, the bare fuselage design is more convenient and versatile for research and development. These motor mounts are colored differently from the other motor mounts to show the multirotor aircraft's heading. The fuselage is also constructed such that the motor arms may be stored and transported easily. The multirotor aircraft also features a cargo mechanism at the bottom for extra electronics or sensors.

To build the multirotor aircraft, the SolidWorks file of the bottom plate is exported as a drawing exchange format (DXF) file. This ensures the bottom plate is accurate and closely

matches the drawing. The carbon fiber tube is cut to the desired length using a grinder, where the intended motor-to-motor length is 685 mm.

3.2 Mechanical Part - Landing Gear

Similarly, SolidWorks is used to design the multirotor aircraft landing gear, which is shown in Fig. 4(b). This style of landing gear also allows for extra cargo area. The landing gear must be designed to support the overall weight of the multirotor aircraft plus any additional payloads. The landing gear can also be retracted horizontally for storage and transit.

Similar to the fuselage, the landing gear is made from carbon fiber plates and tubes, as seen in Fig. 4(b). The desired payload height clearance from the ground is 180 mm.

(a)   (b)
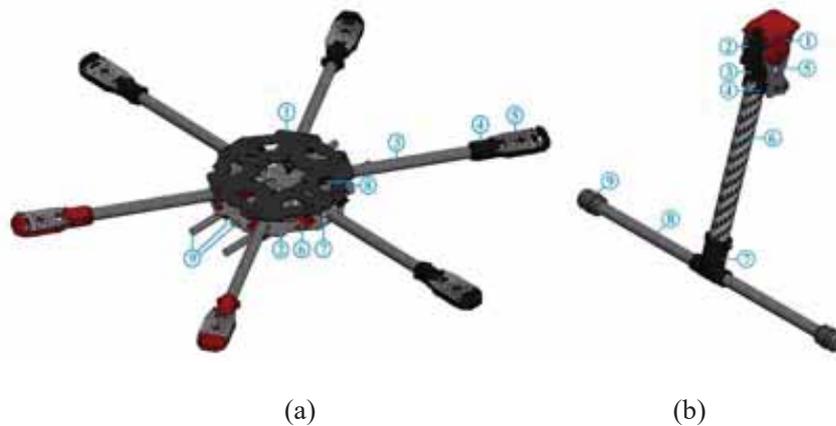
Fig. 4: (a) Fuselage design (b) Landing gear.

3.3 Mechanical Part - Propulsion System

The mechanical components of a multirotor aircraft propulsion system include BLDC motors and propellers. BLDC motors and propellers for multirotor aircraft must be chosen based on the above requirements and the aircraft's overall takeoff weight. The weight of the multirotor aircraft is the key element impacting the BLDC motor choices. To reach the 2:1 thrust to weight ratio, the motors' total thrust must be at least twice the aircraft's entire takeoff weight [15]. The required motor thrust can also be determined using Eq. 2 from [16].

$$Thrust = \frac{(Total\ takeoff\ weight\ of\ aircraft) \times 2}{Number\ of\ motors} \tag{2}$$

The datasheet for a BLDC motor frequently includes the motor's thrust. Important information such as Kv ratings (RPM per supplied volt) and current ratings (maximum current drawable safely) are also included in the datasheet [17]. Additionally, to maximize motor efficiency, most BLDC motors recommend a propeller size and design. The multirotor aircraft constructed for this project is expected to weigh 3 kg in total. Tarot 4108 Brushless DC was finally selected. Six lithium-ion batteries (25.2 V), a 380 Kv motor with 1 kg of torque, and 1355 carbon fiber propellers were also selected. The UAV is 13 inches in diameter and has a 5.5-degree pitch. Therefore, the multirotor should include axial propulsion, puller configuration, and two-blade propellers for better performance.

3.4 Electrical and Electronic Part – Power Supply

A multirotor aircraft's power supply is chosen to meet the propulsion system's voltage and current needs before being stepped down to meet the electronics components' voltage and current requirements. After settling on the motors and propellers, the series and parallel configuration of battery cells can be estimated. This is because one 18650 Li-ion cell delivers 4.2 V when completely charged and may safely draw up to 20 A (depending on the model). For example, to provide 24 V for the motors, at least 6 cells must be connected in series. Batteries are rated in mAh and C, where mAh is the maximum power stored and C is the rate of discharge relative to maximum capacity. According to Eqs. 3 and 4 from [18], a 2000 mAh Li-ion cell at 10 C may discharge at 20 A for 6 minutes.

$$I = Q \times C \tag{3}$$

$$Discharge\ time\ (minutes) = \frac{I}{C} \times 60 \tag{4}$$

Hence, the multirotor aircraft power supply will be LG HG2 3000 mAh 20 C 18650 Li-ion batteries (6S4P) as shown in Fig. 5(a) with a total capacity of 12000 mAh and a maximum voltage of 25.2 V. With this Li-ion battery combination, the motors can run for 12 minutes at maximum discharge current of 20 A. Because the motor can only draw 12.4 A at full throttle, the multirotor aircraft can only fly for 20 minutes. Nickel strips are used to spot weld the Li-ion cells into 6S4P arrangement.

3.5 Electrical and Electronic Part – Electronic Speed Control (ESC)

The ESC controls the motor speed of the multirotor aircraft, affecting the thrust force generated. The key problem in selecting an ESC is that it can handle the voltage and current draw from the power supply to the BLDC motors at full throttle. As indicated in Fig. 5(b), the multirotor aircraft ESC chosen is the HobbyWing XRotor-40A. A safety factor larger than 2 means the ESC's continuous current rating is substantially higher than the motors' maximum current demand.

3.6 Electrical and Electronic Part – Flight Controller

The flight controller of a multirotor aircraft receives input signals and converts them to output signals for the motors. So, one way to find a good flight controller is to look at its specs. The Pixhawk 2/Cube flight controller was chosen for the multirotor aircraft since it is open source and popular among UAV developers. The Cube also has a double redundant microprocessor, triple redundant IMU, and can accommodate up to three GPS modules. The Cube's (I/O) ports are also attached to a carrier board that can integrate with the Intel® Edison single board computer. It also includes 14 PWM pins, 8 of which have a failsafe and manual override.

Thus, redundant components in The Cube ensure that the flight controller does not fail. The carrier board also allows for easy attachment of sensors and other electronic components to the flight controller. Because Pixhawk 2 is an OSH, the dimensions and CAD model of the flying controller depicted in Fig. 5(c) are simply obtained. So, the flight controller may be exactly positioned in the multirotor center to ensure accurate IMU and accelerometer data.

3.7 Electrical and Electronic Part – DC Transmission

The multirotor aircraft uses RC transmission to receive radio signals from the transmitter. The transmission range and frequency are important factors to consider when

choosing an RC transmission system. The TTSRC X9 Remote Control 2.4 G 9CH Transmitter with Receiver X9D is chosen for manual mode since it has a transmission range up to 1 km and runs at 2.4 GHz. In autonomous mode, the 3DR 100 mW Radio Telemetry is chosen due to its 1.5 km transmission range and 915 MHz transmission frequency, as seen in Fig. 5(d) and 5(e). Moreover, both transmitting frequencies are legal.

Fig. 5: (a) Li-ion battery (b) ESC XRotor-40A (c) Pixhawk 2 CAD model
(d) TTSRC X9 (e) 3DR radio telemetry

3.8 Electrical and Electronic Part – Rangefinder

The primary purpose of rangefinders is to determine the distance between an object and the rangefinder, but they are also utilized in this project to identify obstacles and avoid collisions. Table 1 compares several types of rangefinders.

Table 1: Comparison between rangefinders

Parameters	Ultrasonic Sensor	Infrared sensor	LiDAR sensor
Manufacturer	WAVGAT	SHARP	Benewake
Model	GY-US42V2	GP2Y0A710K0F	TFMini
Medium	Ultrasound	Infrared	Laser
Communication interface	Serial UART	Analog	Serial UART
Distance measuring range	20 cm ~ 720 cm	100 cm ~ 550 cm	30 cm ~ 1200 cm

To construct the decentralized 3D collision avoidance system, all of the rangefinders listed in Table 1 will be employed for sensor fusion, allowing the advantages of each rangefinder to outweigh the shortcomings of the others. Additionally, several simulations

and experiments will be undertaken to determine each rangefinder's performance in a variety of different environments.

3.9 Electrical and Electronic Circuit Diagram

Figure 6(a) illustrates the relationships between electrical components that are necessary for the building of the multirotor aircraft circuit diagram depicted in Fig. 6(b). The circuit schematic is created using Fritzing software to aid in the rapid and exact production of the electrical and electronic system.
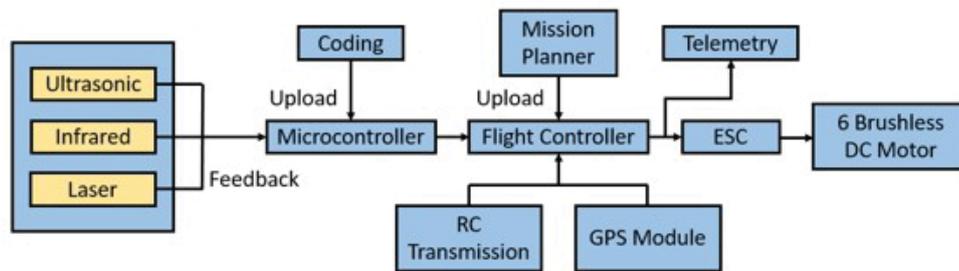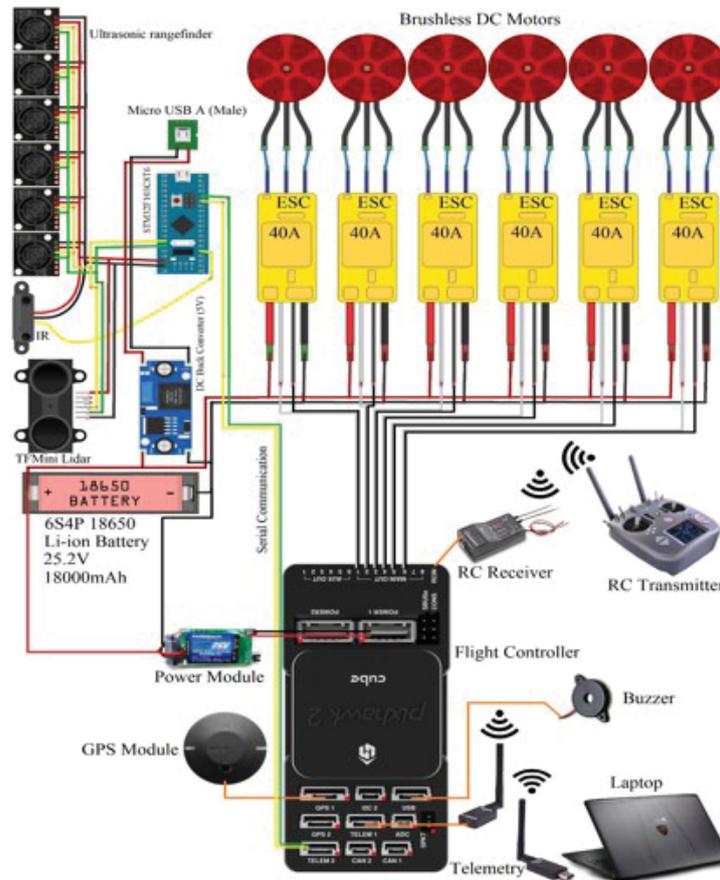
Fig. 6: (a) Electronic components relations, (b) Multirotor aircraft circuit diagram.

Multiple GPIO pins are available on microcontrollers and single-board computers to read analogue and digital inputs from sensors and to output analogue and digital voltage signals. This is to respond to the mobile robot's continually changing electrical requirements. The

microcontroller is used to read data from ultrasonic sensors that generate and receive digital signals. Additionally, through Universal Asynchronous Receiver and Transmitter (UART) connectivity, the single-board computer receives GPS coordinates from the GPS Module. Additionally, single-board computers are utilized to get photographs of railway track defects and video recordings of the railway track from cameras using serial connection protocol. C++ and Python will be used to program the microcontroller and single-board computer, respectively.

3.10 Experimental Procedures

Various experiments are conducted to obtain the required results to achieve the project's objectives. The results are based on the specific performance index to determine the success of the research. One of them is the percentage rate of obstacle detection based on the sensor fusion algorithm. This parameter is essential to show how quickly the system detects obstacles and the percentage of successful detection. The other parameter is the reaction time for the system to perform collision avoidance when an obstacle is detected. The reaction time is vital to ensure the system's robustness in all situations. Four experiments will be conducted to evaluate all the performance parameters for the system.

A. Experiment 1: Obstacle Detection on Multiple Type and Opacity of Material Test

Objective	:	To measure and determine the capabilities for obstacle detection on multiple type of material based on LiDAR, Sonar, and IR sensors
Parameters	:	Manipulated Variable: Type of rangefinder and material
		Responding Variable: Detection of obstacle
Apparatus	:	White polystyrene, cotton hoodie, clear plastic, translucent plastic, wooden plank, black box and rangefinders
Procedure	:	1. Ultrasonic, laser, and infrared rangefinders are positioned at a fixed distance from the wooden plank.
		2. The microcontroller controlling the rangefinders is programmed to obtain the distance between the rangefinders and obstacle.
		3. The distance obtained is recorded and compared with the fixed distance set between the rangefinders and obstacle.
		4. Step 1 to 3 is repeated by changing the type of material to wooden plank, cotton hoodie, clear plastic, translucent plastic, white polystyrene and black box.
		5. The results obtained are tabulated in table form.

B. Experiment 2: Obstacle Detection Range in Different Environments Test

Objective	:	To verify and evaluate the obstacle detection range for each type of rangefinder in various environments.

Parameters	:	Manipulated Variable: Flight environment.
		Responding Variable: Obstacle detection range of rangefinder
Apparatus	:	Wooden plank, rangefinders, and multirotor aircraft
Procedure	:	1. Ultrasonic, infrared, and laser rangefinders are mounted at the same position on the multirotor aircraft.

2. Multirotor aircraft is set up in a sunny environment.

3. The wooden plank is positioned at a fixed distance of 10m from the multirotor aircraft.

4. The distance between the wooden plank and multirotor aircraft is decreased to 9 m, 8 m, 7 m, 6 m, 5 m, 4 m, 3 m, 2 m, and 1 m.

5. The readings of the rangefinders are recorded.

6. Step 2 to 6 is repeated by setting up the multirotor aircraft in a cloudy environment and a night environment.

7. The results obtained are tabulated in table form.

C. Experiment 3: Collision Avoidance Reaction Time Test

Objective	:	To measure and study the reaction time of the multirotor aircraft on collision avoidance.
Parameters	:	Manipulated Variable: Velocity of incoming multirotor aircraft
		Responding Variable: Reaction time of multirotor aircraft
Apparatus	:	Rangefinders, radio controller, telemetry, laptop, and multirotor aircraft
Procedure	:	1. Multirotor aircraft with rangefinders mounted at forward facing position is set up at an empty and open field.

2. Mission Planner is launched on laptop and the flight path of the multirotor aircraft is planned.

3. The flight controller is configured to record the velocity of multirotor aircraft, rangefinder readings, and throttle level of each BLDC motor.

4. The multirotor aircraft took off manually in position hold mode using a radio controller and was controlled to move at a direction parallel to the surface obstacle at a speed of 2 m/s.

5. The data logged in graph form is tabulated in table form.
6. Step 4 to 6 is repeated by changing the speed of the multirotor aircraft to 4 m/s and 6 m/s.

D. Experiment 4: 3D Collision Avoidance Test

Objective : To evaluate and demonstrate the capability of the multirotor aircraft on 3D collision avoidance.

Parameters : Manipulated Variable: Position of obstacles.

Responding Variable: Collision avoidance of multirotor aircraft

Apparatus : Wooden plank, rangefinders, telemetry, laptop, and multirotor aircraft

Procedure : 1. Multirotor aircraft with mounted rangefinders is set up at an empty and open field.

2. Mission Planner is launched on laptop and the flight path of the multirotor aircraft is planned.

3. The multirotor aircraft took off in autonomous mode and executed the flight path planned.

4. Wooden plank is introduced along X-axis, Y-axis and Z-axis of the multirotor aircraft.

5. The reaction of the multirotor aircraft is observed and recorded.

6. The observations obtained are tabulated in table form.

4. RESULTS AND DISCUSSIONS

Results are based on the design of the multirotor aircraft and the 4 real experiments are conducted using the proposed multirotor aircrafts throughout this research.

4.1 Design of Multirotor Aircraft

As seen in Fig. 7(a) and 7(b), the built multirotor aircraft includes the following characteristics: Hexa X, typical fixed skid landing gear, two blade propellers, puller arrangement, and axial propulsion. Due to the cylindrical form of the motor arm, motor mount, and landing gear, clamping is needed to secure them in place; drilling a hole through the carbon fiber tube will reduce its strength. Silicon rubber dampers are installed on the carbon fiber tube that will come into touch with the ground to absorb impact forces on the landing gear system. The 6S4P 18650 Li-ion battery pack is arranged in Fig. 7(c), and it is designed to be sealed using a non-conductive waterproof sealing tape.

These dimensions are critical because they reflect the multirotor aircraft's virtual cylinder dimension for 3D collision avoidance. Thus, the width of the multirotor aircraft with propellers and the height of the multirotor aircraft with battery pack are utilized to establish the proportions of the virtual cylinder for 3D collision avoidance. As a result, the virtual cylinder has a diameter of 1014.7 mm and a height of 300.36 mm as shown in Table 2, which is used to construct a 3D collision avoidance system.

4.2 Obstacle Detection on Multiple Types and Opacity of Material Test

The performance of the collision avoidance system for multirotor aircraft is directly affected by the performance of the several rangefinder types on identifying multiple types

and opacity of material. This experiment will also assess each rangefinder's ability to detect obstacles. Table 2 lists the findings.

As shown in Table 3, ultrasonic sensors can successfully detect all materials used. Also, both laser and infrared rangefinders can detect all materials except transparent materials like clear plastic where incorrect readings occur. Contrary to expectations, ultrasonic sensors were able to detect soft materials like cotton hoodies with ease. However, humans typically wear cotton-based clothes, thus it is critical that the rangefinder utilized can detect it.

Fig. 7: (a) Multirotor aircraft frame design, (b) Modelling of multirotor aircraft design, (c) Modelling of 18650 Li-ion battery pack.

Table 2: Dimension of multirotor aircraft

Dimension	Length (mm)
Motor to motor length	684.50
Breadth (Without propellers)	669.77
Width (Without propellers)	761.45
Height (Without battery pack)	229.00
Breadth (With propellers)	922.99
Width (With propellers)	1014.70
Height (With battery pack)	300.36

Because laser and infrared rangefinders use light to measure distance between objects and the rangefinder, they have trouble identifying transparent materials. Light does not reflect from clear plastic since it is transparent and has low reflectivity. However, when using translucent materials such as translucent plastic, both the laser and infrared rangefinder may detect it due to the material's reflection.

Thus, the advantages of the ultrasonic rangefinder over laser and infrared rangefinder can be exploited to overcome the disadvantages. However, ultrasonic rangefinders have drawbacks including loud readings and limited obstacle detection range that can be overcome by laser and infrared rangefinders.

Table 3: Obstacle Detection on Multiple Type and Opacity of Material Test

Type of Material	Material Used	Type of Rangefinder (RF)					
		Ultrasonic RF		Infrared RF		Laser RF	
		Detection	Observation	Detection	Observation	Detection	Observation
Hard	Wooden plank	√	Good	√	Good	√	Good
Soft	Cotton hoodie	√	Good	√	Good	√	Good
Transparent	Clear plastic	√	Good	X	Bad	X	Bad
Translucent	Translucent plastic	√	Good	√	Good	√	Good
Bright	White polystyrene	√	Good	√	Good	√	Good
Dark	Black box	√	Good	√	Good	√	Good

Table 4: Obstacle Detection Range in different Flight Environments Test Using Ultrasonic, Infrared and Laser Rangefinder

Flight Environment	Detection Range (m)																													
	Ultrasonic RF										Infrared RF										Laser RF									
	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
Sunny	√	√	√	√	√	√	√	X	X	X	√	√	√	X	X	X	X	X	X	X	√	√	√	√	√	√	X	X	X	X
Cloud	√	√	√	√	√	√	√	X	X	X	√	√	√	X	X	X	X	X	X	X	√	√	√	√	√	√	√	√	X	X
Night	√	√	√	√	√	√	√	X	X	X	√	√	√	X	X	X	X	X	X	X	√	√	√	√	√	√	√	√	√	√

4.3 Obstacle Detection Range in Different Environments Test

While an ultrasonic rangefinder is generally unaffected by its environment (such as illumination, brightness, and weather), it is tested to confirm its function in a variety of conditions. The infrared and laser rangefinders' functionality will be impacted by the illumination conditions in the environment. In bright flight situations, the average illuminance on the surface of the obstacle is between 32 Klux and 100 Klux, whereas in cloudy flight environments, the average illuminance on the surface of the obstacle is between 1 Klux and 20 Klux. Needless to add, the average value of illuminance on the surface of obstacles during night flight conditions is less than 1 lux. Table 4 summarizes the findings.

According to Table 3, the ultrasonic rangefinder can detect obstacles from 1m to 7m in various flight situations. Also, because ultrasonic rangefinders use air to measure distances between themselves and obstacles, the weather does not affect their effectiveness. The performance of the laser rangefinder varies with flight conditions since the surface of the barrier is illuminated differently in different flight situations. The laser rangefinder has a 6 m obstacle detection range in sunny conditions, 8 m in overcast conditions, and 10 m in night flight. The obstacle detection ranges of a laser rangefinder increase with decreasing surface illuminance. Nevertheless, the laser rangefinder's obstacle detection range in Table 1 matches its performance during night flight.

The infrared rangefinder's performance does not match the criteria in Table 1, which only identifies obstacles 1 to 3 meters away. Surprisingly, the degree of illuminance does not impact the infrared rangefinder's effectiveness in sunny, overcast, or night flight conditions. This is true for this particular infrared rangefinder device. The chosen ultrasonic and laser rangefinders are suited for obstacle identification in the X- and Y-axis of the multirotor aircraft frame because they can identify obstacles at a distance. The ultrasonic rangefinder also offers a larger obstacle detecting area, increasing the success rate of the collision avoidance system. Due to the design of the multirotor aircraft's frame, the infrared rangefinder can be employed for obstacle detection in the Z-axis. The Z-axis multirotor aircraft outermost components are the battery and landing gear, whereas the X- and Y-axis outermost components are the propellers.

4.4 Collision Avoidance Reaction Time Test

The reaction time of a multirotor aircraft is the time it takes to avoid an obstacle. Preventing collisions requires quick response from multirotor aircraft when obstacles are detected. This experiment is designed to estimate the reaction time of a multirotor aircraft so that it can react quickly when an obstruction is detected. Table 5 shows the results.

Table 5 shows that when travelling at 2 m/s, the multirotor aircraft reacts promptly when an obstruction is detected. However, at 4 m/s, the multirotor aircraft made evasive maneuvers 30 milliseconds after recognizing the obstruction. The multirotor aircraft also experienced a 35 milliseconds delay before performing evasive moves when going at 6 m/s.

When an obstruction is detected, the multirotor aircraft's reaction time is sufficient to avoid it. Rangefinders also read distances at 100 Hz, or 10 milliseconds each reading. As a result, the reaction time for the multirotor aircraft is 0–35 milliseconds. That's because the laser rangefinder's maximum obstacle detection range in diverse flight conditions is 6 meters, so the multirotor maximum speed is limited. It will also be unable to avoid a collision with an obstruction if it exceeds 6 m/s.

4.5 3D Collision Avoidance Test

The multirotor aircraft should be able to avoid obstacles on the X, Y, and Z axes. Figure 8 (a) shows the intended autonomous mission flight path with an impediment noted in red. The

multirotor aircraft must also avoid an impediment at waypoint 3 on the map upon landing. The mission speed for the multirotor aircraft is set to 1 m/s and the altitude to 2.5 m. Table 6 contains the results seen in Fig. 8(b)–(d).

Table 5: Collision avoidance reaction time test results

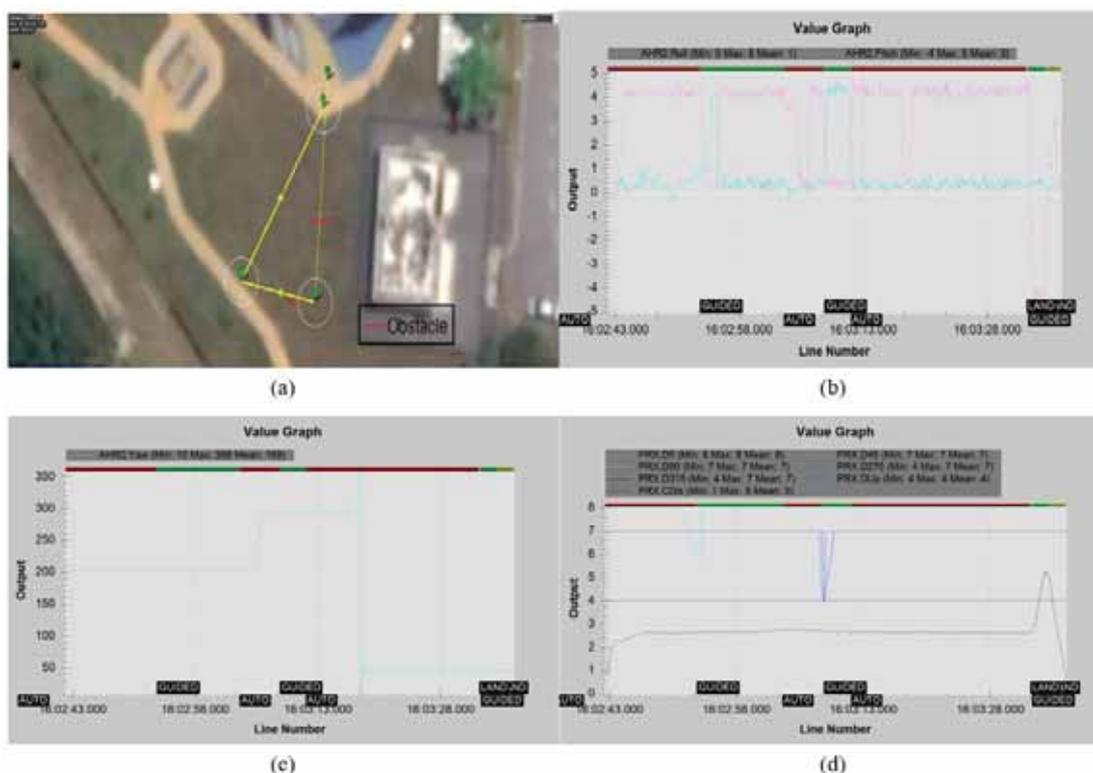| Velocity of multirotor aircraft (m/s)	Reaction time of multirotor aircraft (1×10^{-3} s)
2	0
4	30
6	35

Fig. 8: (a) Autonomous mission flight path with obstacle position marker (b) Roll and pitch angle data logged of multirotor aircraft (c) Yaw angle data logged of multirotor aircraft (d) Rangefinders data logged of multirotor aircraft

Table 6: 3D Collision Avoidance Test Results

Position of obstacle	Observation
Along X-axis	Successful avoidance
Along Y-axis	Successful avoidance
Along Z-axis	Successful avoidance

Figure 8(b) and 8(c) show the multirotor aircraft's attitude and heading reference system data logged during the autonomous mission. Moreover, positive pitch indicates forward motion, while positive roll indicates rightward motion. Also, the multirotor yaw angle shows the multirotor heading. On the other hand, Fig. 8(d) shows the rangefinder data logged during the autonomous mission, with D0 being the forward rangefinder reading, D45 the northeast

rangefinder reading, D90 the east rangefinder reading, and D315 the northwestern rangefinder reading. Also, Dup is the rangefinder reading for upwards and CDIS is the rangefinder reading for downwards. The multirotor aircraft avoided collisions along the X, Y, and Z-axis in Table 6.

To get from waypoint 0 to 1, the multirotor aircraft first pitches forward. The program for this experiment makes the multirotor aircraft conduct evasive maneuvers when it detects a forward obstruction. The forward barrier is between waypoints 0 and 1. Also, when a multirotor aircraft detects a forward obstruction, it will roll to the right until it no longer detects the barrier and then proceed forward. The multirotor aircraft then yaws 90° and pitches forward to move from waypoint 1 to 2. The back obstacle is also parallel to the path from waypoint 1 to 2. When a multirotor aircraft detects a rear obstacle, it is programmed to roll away from it and pitch ahead. To proceed from waypoint 2 to waypoint 3, the multirotor aircraft must yaw for roughly 120° and pitch forward. Also, when the multirotor aircraft reaches waypoint 3, it immediately enters LAND mode. An obstacle is positioned at waypoint 3 underneath the multirotor aircraft, which will undertake evasive maneuvers, such as pitching backward, until the object is not spotted, then land. Thus, the multirotor aircraft can avoid obstacles on the X, Y, and Z axes.

5. CONCLUSIONS

The 3D UAV collision avoidance system has been built and tested. The 3D collision avoidance system can also detect several material types and opacities in varied flight environments by combining selected rangefinders. Finally, the 3D collision avoidance algorithm can react quickly to obstructions in the X, Y, and Z axes.

In the future, the hexa-rotor design can be improved with better obstacle detection systems such as 360° LIDAR and solid-state LIDAR, or vision-based obstacle detection systems such as stereo cameras and depth sensors. As obstacle detection coverage and performance improve, better collision avoidance algorithms can be built. Suitable filters should also be attached to the obstacle detection system to eliminate inconsistent readings owing to electrical disturbances and sensor interference.

ACKNOWLEDGMENTS

REFERENCES

[1] Shokirov R, Abdujabarov N, Jonibek T, Saytov K, Bobomurodov S. (2020) Prospects of the development of unmanned aerial vehicles (UAVs). Tech. Sci. Innov., 2020(3): 4-8. doi: 10.51346/tstu-01.20.3-77-0069.

[2] Shakhatreh H, Sawalmeh AH, Al-Fuqaha A, Dou Z, Almaita E, Khalil I, Othman NS, Khreishah A, Guizani M. (2019) Unmanned aerial vehicles (UAVs): A survey on civil applications and key research challenges. IEEE Access, 7:48572–48634. doi: 10.1109/ACCESS.2019.2909530.

[3]   Ke Y, Wang K, Chen BM. (2018) Design and implementation of a hybrid UAV with model-based flight capabilities. IEEE/ASME Trans. Mechatronics, 23(3): 1114-1125. doi: 10.1109/TMECH.2018.2820222.

[4]   Magree D, Mooney JG, Johnson EN. (2014) Monocular visual mapping for obstacle avoidance on UAVs. J. Intell. Robot. Syst. Theory Appl., 74(1-2): 17-26. doi: 10.1007/s10846-013-9967-7.

[5]   Ferrera E, Alcántara A, Capitán J, Castaño AR, Marrón PJ, Ollero A. (2018) Decentralized 3D collision avoidance for multiple UAVs in outdoor environments. Sensors (Basel), 18(12): 1-20. doi: 10.3390/s18124101.

[6]   Eduardo Ferrera PJM, Capitan J, Castano AR. (2017) Decentralized safe conflict resolution for multiple robots in dense scenario. Rob. Auton. Syst., 91: 179-193.

[7]   Harun MH, Abdullah SS, Aras MSM, Bahar MB. (2022) Sensor fusion technology for unmanned autonomous vehicles (UAV): A review of methods and applications. IEEE International Conference on Underwater System Technology: Theory and Applications (USYS), 2020:1-8.

[8]   Harun MH, Abdullah SS, Aras MSM, Bahar MB. (2021) Collision avoidance control for unmanned autonomous vehicles (UAV): Recent advancements and future prospects. Indian Journal of Geo Marine Sciences, 50(12): 873-883.

[9]   Sarin P. (2013) DigitalLab Course Project EP 315, 1–22.

[10] Hughes R. (2019) Amazon fires: What's the latest in Brazil?, BBC (Oct, 2019).

[11] Administration FA. (2000) Aerodynamics of flight, Pilot Handbook Aeronautic Knowledge.

[12] Hibbeler RC. (2013) Engineering Mechanics Statics, Thirteenth Edition.

[13] Hibbeler RC. (2012) Engineering Mechanics Dynamics, Thirteenth Edition.

[14] Pikalov S, Azaria E, Sonnenberg S, Ben-Moshe B, Azaria A. (2021) Vision-less sensing for autonomous micro-drones. Sensors, 21(16):5293.

[15] Hamaza S, Georgilas I, Heredia G, Ollero A, Richardson T. (2020) Design, modeling, and control of an aerial manipulator for placement and retrieval of sensors in the environment. J. F. Robot., 37(7):1224–1245.

[16] Suprapto BY, Heryanto MA, Suprijono H, Muliadi J, Kusumoputro B. (2018) Design and development of heavy-lift hexacopter for heavy payload. International Seminar on Application for Technology of Information and Communication (iSemantic), 2018-January:242–246. doi: 10.1109/ISEMANTIC.2017.8251877.

[17] Ahmed MDF, Mohanta JC, Zafar MN. (2018) Development of smart quadcopter for autonomous overhead power transmission line inspections. Materials Today: Proceedings, 51:261-268.

[18] Zhao T. (2018) Propulsive battery packs sizing for aviation applications. Master Theses.

EMG BASED CONTROL OF WRIST EXOSKELETON

MOHD SAFIRIN KARIS[1*], HYREIL ANUAR KASDIRIN[2], NORAFIZAH ABAS[2], WIRA HIDAYAT MOHD SAAD[3] AND MOHD SHAHRIEEL MOHD ARAS[2]

[1]*Faculty of Electrical and Electronic Engineering Technology, Universiti Teknikal Malaysia Melaka, 75450 Ayer Keroh, Melaka, Malaysia*
[2]*Faculty of Electrical Engineering, Universiti Teknikal Malaysia Melaka, 76100 Durian Tunggal, Melaka, Malaysia, Malaysia*
[3]*Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka, 76100 Durian Tunggal, Melaka, Malaysia, Malaysia*

[*]*Corresponding author: safirin@utem.edu.my*

ABSTRACT: The significance of human motion intentions in a designed exoskeleton wrist control hand is essential for stroke survivors, thus making EMG signals an integral part of the overall system is critically important. However, EMG is a nonlinear signal that is easily influenced by several errors from its surroundings and certain of its applications require close monitoring to provide decent outcomes. Hence, this paper proposes to establish the relationship between EMG signals and wrist joint angle to estimate the desired wrist velocity. Fuzzy logic has been selected to form a dynamic modelling of wrist movement for a single muscle at different MVC levels and double muscles at a similar MVC level. The physical model of the exoskeleton hand using Simmechanics Matlab software has been developed to validate the performance of the fuzzy logic output result from both dynamic modelling approaches. A PID controller has been developed to smooth the exoskeleton hand movement fluctuations caused by the fuzzy logic decision-making process. As a conclusion, results showed a strong relationship between EMG signals and wrist joint angle improved the estimation results of desired wrist velocity for both dynamic modelling approaches hence strengthened the prediction process by providing a myoelectronic control device for the exoskeleton hand.

ABSTRAK: Kepentingan dalam mengetahui kehendak gerakan pergelangan tangan manusia adalah penting untuk pesakit strok yang terselamat, justeru menjadikan isyarat EMG amat penting pada keseluruhan sistem. Walau bagaimanapun, EMG adalah isyarat tidak linear yang mudah dipengaruhi ralat sekitaran dan memerlukan pemantauan rapi bagi hasil yang baik. Oleh itu, kajian ini mencadangkan kewujudan hubungan antara isyarat EMG dan sudut sendi pergelangan tangan bagi menganggarkan halaju pergelangan tangan yang dikehendaki. Logik kabur (*fuzzy logic*) telah dipilih bagi membentuk model dinamik pergerakan pergelangan tangan pada otot tunggal di tahap MVC yang berbeza dan otot berganda pada tahap MVC yang serupa. Model fizikal rangka luar tangan menggunakan perisian Matlab Simmekanik telah dibangunkan bagi mengesahkan prestasi Logik Kabur daripada kedua-dua pendekatan model dinamik. Pengawal PID telah dibangunkan bagi melicinkan gerakan turun naik tangan yang disebabkan proses membuat keputusan oleh Logik Kabur. Sebagai kesimpulan, dapatan kajian menunjukkan hubungan yang kukuh antara isyarat EMG dan sudut sendi pergelangan tangan. Ini meningkatkan anggaran dapatan halaju pergelangan tangan yang dikehendaki bagi kedua-dua pendekatan model dinamik seterusnya mengukuhkan proses ramalan melalui peranti kawalan mioelektronik rangka tangan.

1. INTRODUCTION

Good health is a vital component for humans to emphasise life's quality by giving a person the best potential to face all of life's problems. This element might be degraded due to so many factors and one of them is stroke. In 2019, stroke overtook heart disease and pneumonia to become the third-leading cause of mortality globally and the second-leading cause of disability in Malaysia [1]. The most prevalent causes include hypertension (high blood pressure), diabetes, high cholesterol, irregular heartbeats, smoking, excessive alcohol use, inactivity, obesity, stress, and family history [1]. 75% of stroke survivors experience upper limb weakness, which includes constrained hand and wrist movement [2]. Based on this issue, stroke can disrupt a person's lifestyle, especially if it results in slow hand response, which restricts hand movement.

Degradation of hand function, which can be caused by stroke attack, may distract a brain signal from channelling user motion intention to control the hand movement. In general, hand structure is highly articulated, built up with many muscles' composition and bony segments. The main components of the hand are referred to in human anatomy as the palm, fingers, and wrist. There are various classifications of wrist movement, including neutral, flexion, extension, adduction, abduction, pronation, and supination [3]. These movements are directed by the user's motion intention which ultimately causes the user's hand muscles to contract, causing the user's hands to move [4]. This muscle excitation causes the Electromyography (EMG) signal to get excited. The EMG signal is a measure of electrical activity that spreads throughout the muscles and activates them. However, the desired wrist velocity and wrist joint angle position are correlated with this EMG signal strength [5].

To analyse the relationship between EMG signals, wrist joint angle position and desired wrist velocity, a physical model of the exoskeleton hand has been built up to mimic the human hand's articulated movement. The EMG signals collected from forearm muscles has served as the exoskeleton hand's primary input, and the wrist joint's placement angle has been increasing the accuracy of predictions of the desired wrist velocity. Given the significance of this relationship, it is important to include all of these factors as part of the control input so that the hand can move predictably whenever the muscles are flexed.

This written paper aims to analyse the relationship between EMG signals, wrist joint angle positioning, and desired wrist velocity. The output from this relationship has been expected to improve the output results for single muscles at different MVC levels and double muscles at similar MVC levels in wrist dynamic modelling approaches. After the decision-making process has been made, it has been tested on a virtual exoskeleton hand focusing on the desired wrist velocity.

2. MECHANICAL HAND DESIGN

The exoskeleton hand is designed to closely emulate the actual human hand that comprises of 5 fingers, a palm and a wrist that is attached to the forearm. The fingers comprise of 4 medial fingers with 3 bony segments and 1 thumb that has 2 bony segments. Based on hand anatomy, several hand measurements from ten male Asian subjects' aged from 21 to 40 years old following the anthropometric hand measurement have been collected. The mean of each measurement is illustrated in Table 1 below.

Table 1: Terminology on hand measurement

No	Finger	Hand Part	Mean [cm]
1	**Thumb**	Distal Phalanges Length	31.41
2		Proximal Phalanges Length	25.59
3		Distal Phalanges Breadth	18.26
4		Proximal Phalanges Breadth	16.67
5	**Index**	Distal Phalanges Length	23.94
6		Intermediate Phalanges Length	25.06
7		Proximal Phalanges Length	26.16
8		Distal Phalanges Length	14.55
9		Intermediate Phalanges Breadth	14.35
10		Proximal Phalanges Breadth	14.27
11		Distal Phalanges Length	25.15
12	**Middle**	Intermediate Phalanges Length	26.78
13		Proximal Phalanges Length	26.90
14		Distal Phalanges Breadth	14.01
15		Intermediate Phalanges Breadth	16.79
16		Proximal Phalanges Breadth	16.79
17	**Ring**	Distal Phalanges Length	24.36
18		Intermediate Phalanges Length	25.68
19		Proximal Phalanges Length	24.20
20		Distal Phalanges Length	13.62
21		Intermediate Phalanges Breadth	14.69
22		Proximal Phalanges Breadth	15.58
23	**Little**	Distal Phalanges Length	22.46
24		Intermediate Phalanges Length	15.28
25		Proximal Phalanges Length	15.89
26		Distal Phalanges Breadth	12.22
27		Intermediate Phalanges Breadth	13.26
28		Proximal Phalanges Breadth	14.54
29	**Palm**	Centre of Wrist crease to Middle	100.36
30		Centre of metacarpal	81.10
31		Hand Depth	47.60

The exoskeleton hand has been designed using measurements based on the collected anthropometry data. The design process began with the creation of each component, followed by their assembly into a single model or prototype as shown in Fig. 1. The exoskeleton hand design has been able to mimic natural human hand movement.
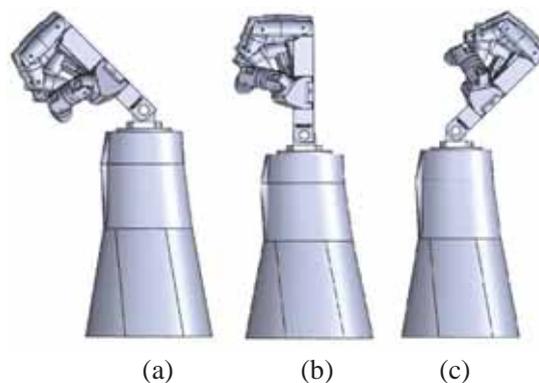

(a)   (b)   (c)

Fig. 1: Solidwork 3D Hand Design (a) Flexion Position, (b) Neutral Position, (c) Extension Position.

In this exoskeleton hand design, one degree of freedom (DoF) of wrist joint angle position has been focused on. This wrist movement covered three wrist hand joint angle positions namely, neutral (0°), flexion (-45°) and extension (45°). The exoskeleton hand has been designed under the fully actuated category since the exoskeleton hand wrist design can be moved towards achieving the desired wrist velocity. The flexibility of the hand has been designed to provide practical, typical hand movements.

3. EMG DATA COLLECTION

The samples came from male participants between the ages of 21 and 40. No neurological illnesses were present in any of the individuals, who were all in good health. The dominant hand has been chosen to collect the muscle movement-related EMG data. The palpate Scaphoid technique has been utilised to determine the wrist movement, and the medial epicondyle has been used to pinpoint the muscles' location [6]. For each position of wrist joint angle movement, a particular muscle has been activated in response to it. In this paper, the EMG signal values of the Flexor Carpi Radialis (FCR) and Extensor Carpi Radialis Longus (ECRL) has been used to represent each wrist joint angle movement [7].

Ten male subjects agreed to do the hand grip pattern experiment at different wrist joint angle positions by signing the consent form provided by the researcher. The experiment started as all the subjects were given detailed explanation of the procedure involved. Each experiment repeated three times [8]. The equipment used in this experiment, such as a hand dynamometer, LabQuest Mini data acquisitions by Vernier, EMG sensors by Vernier, a personal computer with Logger Lite data-collection software, a stopwatch, a protector and Kendall5400 diagnostic tab electrodes. All the experimental procedures were subject to the approval of the appropriate University Ethical Committee or Centre for Research and Innovation Management (CRIM) University Technical Malaysia Melaka (UTeM) Malaysia.

For the experimental procedures, subjects were instructed to use full hand grip strength to hold the hand dynamometer for five seconds, as shown in Fig. 2. The subject's greatest voluntary contraction was measured as the hand grip's maximal force (MVC). FCR and ECRL's muscles had electrode patches applied to the top of their abdominal muscles. The subjects were told to hold the hand dynamometer for five seconds while various hand grip strengths (20, 40, 60, 80, and 100% MVC) were utilised [8]. Rest periods were integrated for around two seconds after each grip. The Logger Lite software was used to record the extracted raw EMG signals. Then subjects repeated the same step as at 0° of wrist joint angle for different wrist positions (at 45° and -45°), as shown in Fig. 3 [9]. Figure 4 shows the EMG signals for the muscle excitation data parameters at different wrist joint angle hand positions.

4. EMG SIGNAL PROCESSING

The data that was collected was processed through several steps including detrend, filtering, feature extraction, and regression method. In the detrending process, data that did not begin at zero was moved to zero before proceeding to the following procedure. This approach improved amplitude readings by eliminating the inaccuracy caused by DC offset during the experimental procedure [10].

The feature of pre-processing EMG signals was extracted using a time domain method called Waveform Length (WL). In this study, WL technique has been used to extract the time-domain-based feature and measure the complexity of EMG signal excitation [11]. WL

is also defined as the cumulative length of waveform over the segment. Equation (1) below showed the equation modelled for the WL method.
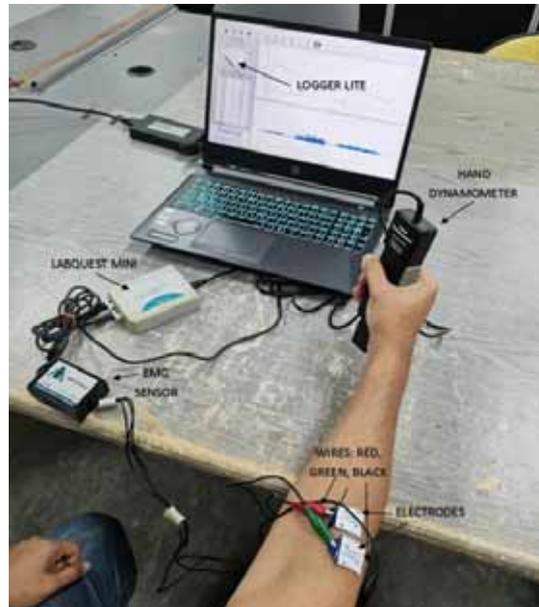
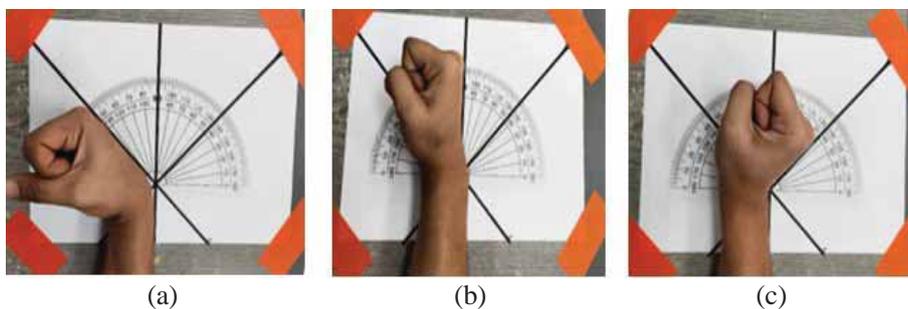$$WL = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{1}$$

Fig. 2: Experimental set-up.

| (a) | (b) | (c) |

Fig. 3: Hand grasping positions for various wrist angle positions:
(a) at -45°, (b) at 0°, and (c) at 45°.

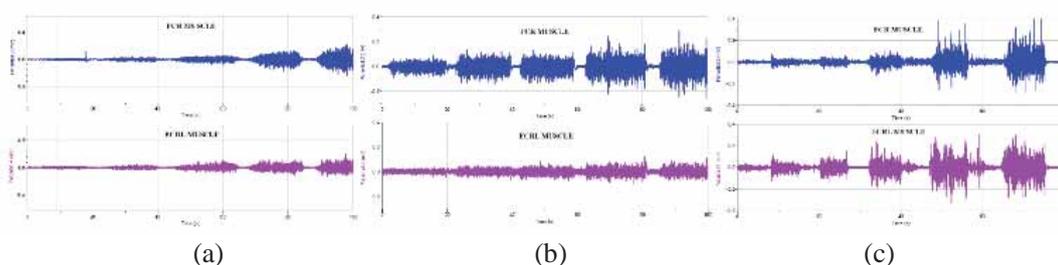| (a) | (b) | (c) |

Fig. 4: EMG muscle excitations at different wrist joint angle positions,
(a) neutral, (b) flexion, (c) extension.

The sampling frequency was selected to be 1 kHz (i.e., T=1 ms) [8] as the acceptable sampling time ranges between frequencies of 200 and 2000 Hz. The recorded EMG voltages

were successfully processed within a 128 ms (sliding) time span and the analysis window increment was set to 50%, which is the interval between two adjacent windows.

The EMG signal was filtered using a second-order band-pass Butterworth filter with a cut-off frequency of 2 Hz (10-350Hz) [12]. A Butterworth filter rejects particular frequencies while allowing signals between two designated frequencies to flow through. Because the EMG signal ranged from 2-45Hz, 10-350Hz filters were applied.

To standardise EMG measurements, the maximum voluntary contraction (MVC) recorded individually from each subject was employed. This mechanism regulated the scaled value between 0 and 1 based on the measurement [13]. The data for each joint angle was normalised by subtracting the smallest joint angle for each sample and dividing it by its mean value. MVC-normalizing is one of the most extensively used normalisation algorithms [14].

5. DYNAMIC MODELLING OF WRIST MOVEMENT USING FUZZY LOGIC

Input signals derived from EMG signals, wrist joint angle position, and desired wrist velocity were predicted to be related. Dynamic modelling was utilised to model the relationship between these three parameters to analyse this relationship. According to the experimental results, EMG signals generated by muscle excitement are directly proportional to wrist joint angle position. As a result, the appropriate wrist desired velocity was determined by this relationship. Fuzzy logic was utilised to perform the prediction mapping between these three parameters.

5.1 Dynamic Modelling on Single Muscle at Different MVC Level

The FL's three inputs recorded EMG signals from the forearm muscles and wrist joint angle position. The desired velocity of the hand wrist was chosen as the FL's output since it has one of the most common control theories for exoskeleton hand systems [4]. The activation rate of the extensor and flexor muscles was continuously assessed utilising EMG signal data and WL feature extraction algorithms.

The FLC's structure consists of fuzzification, a rule table, a fuzzy inference system, and a defuzzification algorithm [7]. As seen in Fig. 5, fuzzy logic first fuzzifies the input variables before building the rule table with membership functions. A Mamdani-type fuzzy inference system was used to map the inputs to the output by merging all of the linguistic claims in the rule table. Finally, the defuzzification operation was carried out using a "centroid" technique that returns the area centre beneath the curve [7]. FCR and ECRL muscle values were analysed separately in this design based on their performance at different MVC (20%, 60%, and 100%) with various wrist joint angle positions to predict the wrist desired velocity.
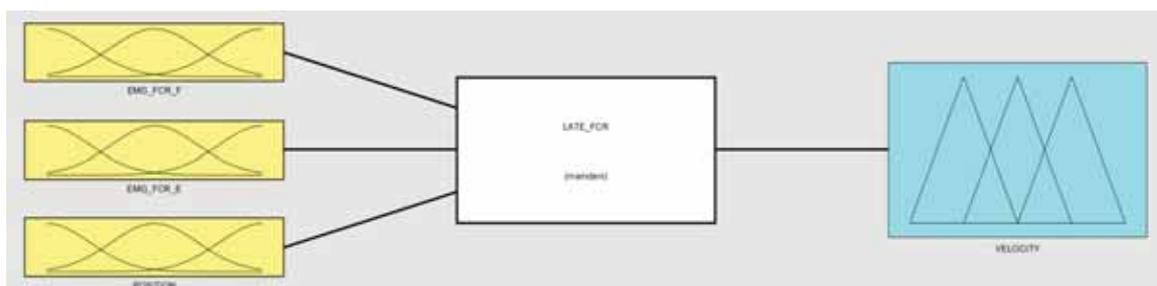

Fig. 5: Fuzzy logic single muscle design.

As the first stage in the fuzzy logic process, the input variables have been fuzzified using membership functions that can take several forms. Figure 6(a), 6(b), and 6(c) below showed three triangular-shaped membership functions that have been used to fuzzify EMG signals and two trapezoidal-shaped membership functions that have been used to fuzzify wrist join angle position. The input to the FCR and ECRL muscles has been analysed in two stages: EMG_F (EMG value while the wrist was in flexion position) and EMG_E (EMG value while the wrist was in extension position). EMG signals are classified as SMALL (S), MEDIUM (M), and HIGH (H). Input triangular membership functions are defined based on the testing data after passing through the feature extraction process. For input, the highest and lowest data numbers were taken from their own group dataset to form this triangular membership function.
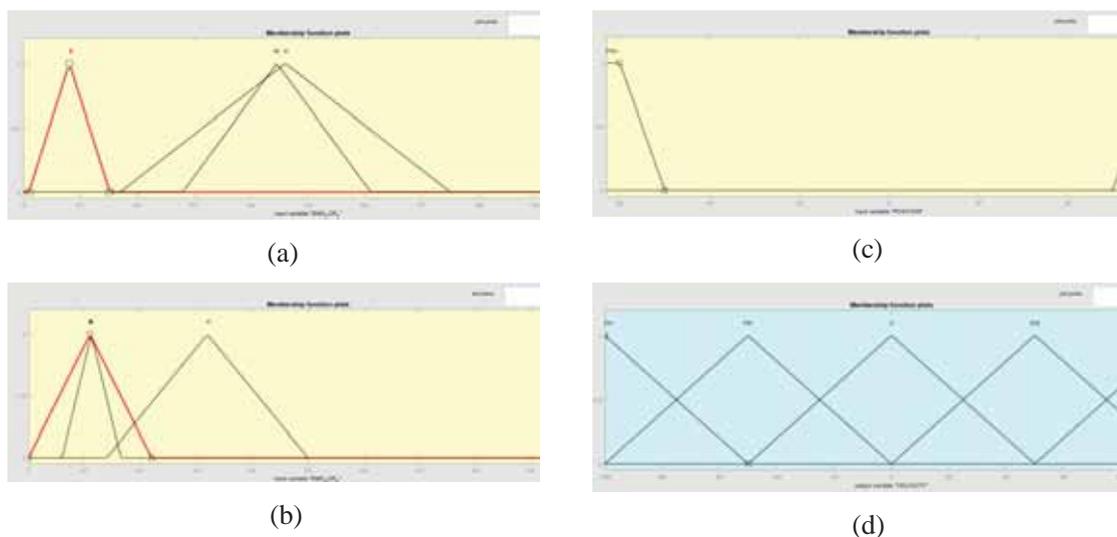

(a)

(c)

(b)

(d)

Fig. 6: (a) Function of triangular-shaped membership used for EMG signal (EMG_F), (b) Function of triangular-shaped membership used for EMG signal (EMG_E), (c) Function of trapezoidal-shaped membership used for wrist joint angle (POSITION), (d) Function for triangular membership used for the desired wrist velocity.

The output desired wrist velocity of fuzzy logic used five triangular-shaped membership functions. The performance of the fuzzy technique, which was chosen as the wrist desired velocity, was categorized as FH (flexion high: 100°/s), FM (flexion medium: 50°/s), Z (zero velocity: 0°/s), EM (extension medium: 50°/s) and EH (extension high: 100°/s) as illustrated in Fig. 6(d).

Table 2 shows the rule table for 15 IF-THEN rule statements. EMG_F and EMG_E are signals generated during the wrist joint angle for flexion and extension. The EMG signals produced by both flexion and extension actions should yield zero wrist velocity, as shown in the first three basic criteria in Table 2. The next six rules examined the different EMG signal excitations to predict the wrist desired velocity. To round out the fifteen criteria, there were rules that limited the device's wrist joint angle as well as provided further safety measures for the desired wrist velocity.

The 15 linguistic statements in the rule table were all combined to form a Mamdani-style fuzzy inference method that maps the inputs to the output. The conjunction expression AND in these sentences corresponds to a minimal operator that selects only the lowest of the fuzzified inputs. For each IF-THEN rule statement, an implication operation was performed using the "minimum" method, which directly truncates the fuzzy output sets. Instead, at the

end, a complex-shaped curve was generated by aggregating all the trimmed fuzzy output sets using the "maximum" method.

Table 2: 15 IF-THEN rule statements [7].

No		EMG_F		EMG_E		Wrist Position		Wrist Velocity
1	if	S		S			then	Z
2	if	M		M			then	Z
3	if	H		H			then	Z
4	if	M	and	S	and	not Elim	then	FM
5	if	H	and	M	and	not Elim	then	FM
6	if	H	and	S	and	not Elim	then	FH
7	if	S	and	M	and	not Elim	then	EM
8	if	M	and	H	and	not Elim	then	EM
9	if	S	and	H	and	not Elim	then	EH
10	if	M	and	S	and	Elim	then	Z
11	if	H	and	M	and	Elim	then	Z
12	if	H	and	S	and	Elim	then	Z
13	if	S	and	M	and	Elim	then	Z
14	if	M	and	H	and	Elim	then	Z
15	if	S	and	H	and	Elim	then	Z

5.1.1 Fuzzy Logic Output Design for Dynamic Modelling on Single Muscle at Different MVC Level
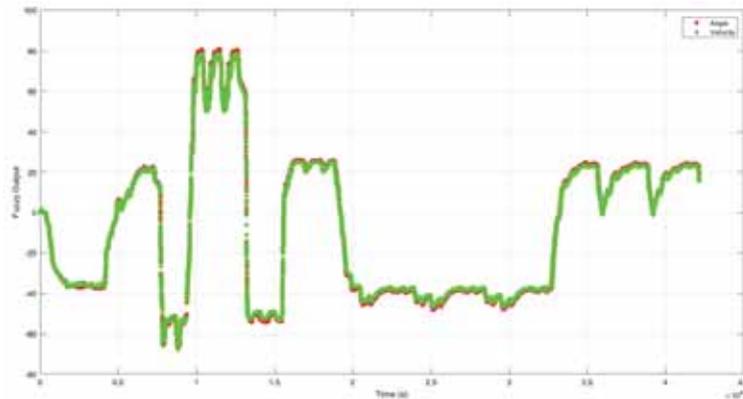

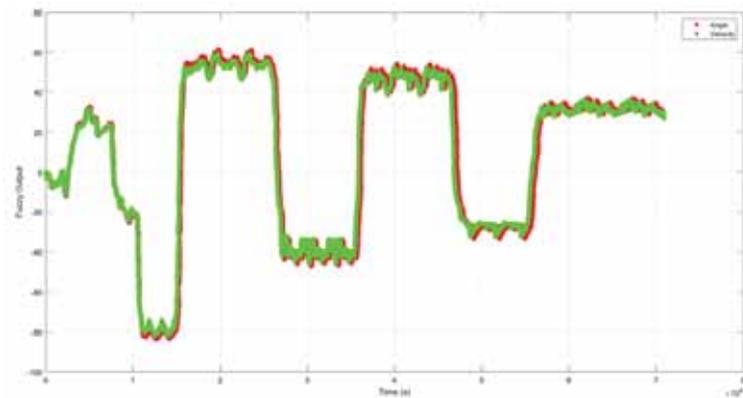Fig. 7: FCR muscle response from fuzzy logic output.

Fig. 8: ECRL muscle response from fuzzy logic output.

Figures 7 and 8 depict the output response of fuzzy logic depending on various MVC muscle levels and wrist joint angle movements. Based on these output results, identical levels of MVC should generate a velocity value of zero. As the Input signal from EMG_F became greater than EMG_E, the velocity of the hand shifted from extension to flexion and vice versa. When the tested data was compared to the bigger or lesser value than the other, it generated different amplitude values for the wrist's desired velocity.

5.2 Dynamic Modelling on Double Muscles at Similar MVC Level

The FL's two inputs recorded EMG signals from two forearm muscles. As one of the most popular control theories for exoskeleton hand systems, the desired velocity of the hand wrist was selected as the FL's output [4]. The activation rate for extensor and flexor muscles of the EMG signal were continuously assessed using the WL feature extraction algorithms.

The structure of the FLC comprises fuzzification, a rule table, a fuzzy inference system, and a defuzzification mechanism [7]. Fuzzy logic first fuzzifies the input variables before designing the rule table using membership functions, as depicted in Fig. 9. A fuzzy inference system of the Mamdani type is used to map inputs to outputs by merging all linguistic statements in the rule table. The defuzzification step was completed using a "centroid" technique that returned the area's centre beneath the curve [7]. In this architecture, FCR and ECRL muscle values were simultaneously analysed based on their performance at different MVC levels (20% and 60%) and wrist joint angle positions to estimate the wrist's desired velocity.
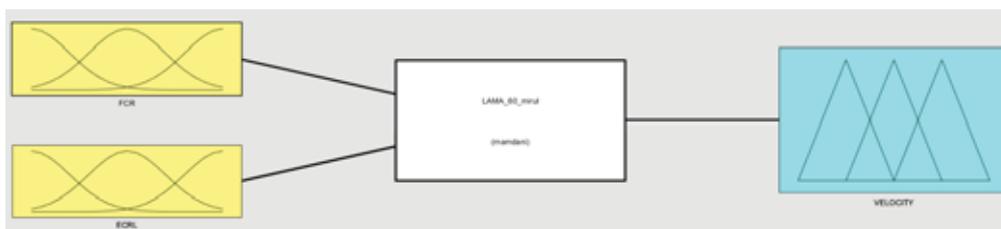

Fig. 9: Fuzzy logic double muscles design.

As the initial step of fuzzy logic, the input variables are first fuzzy - tuned using membership functions, which can take on a variety of shapes. As depicted in Fig. 10(a) and 10(b), three triangular-shaped membership functions are utilised to fuzzify EMG signals. Following the feature extraction procedure input triangular membership functions were defined based on the testing data. EMG signals were classified as SMALL (S), MEDIUM (M), and HIGH (H). To build this triangular membership function, the highest and lowest data numbers from their respective group dataset were used as input.

Five triangular-shaped membership functions were used to fuzzify desired wrist velocity fuzzy output depicted in Fig. 10(c). The performance of the desired wrist velocity was labelled as FH (flexion high: 100°/s), FM (flexion medium: 50°/s), Z (zero velocity: 0°/s), EM (extension medium: 50°/s) and EH (extension high: 100°/s).

Table 3 displays the rule table for 9 IF-THEN rule statements. During wrist flexion position, the FCR muscle gave most responsive signal level hence the parameter data was recorded, whereas the ECRL muscle signal level response was the most responsive during wrist extension position so the data being documented represented the extension state. The wrist velocity was suspected to produce zero state if the EMG signals generated by flexion and extension actions of wrist movement had similar value. The wrist velocity was determined by comparing the EMG signal level. If FCR value was greater than ECRL, FM was created, whereas if ECRL was greater than FCR, EM was produced. FH was formed

when the value of FCR was extremely bigger than ECRL, and EH was produced when the value of ECRL was extremely greater than FCR.
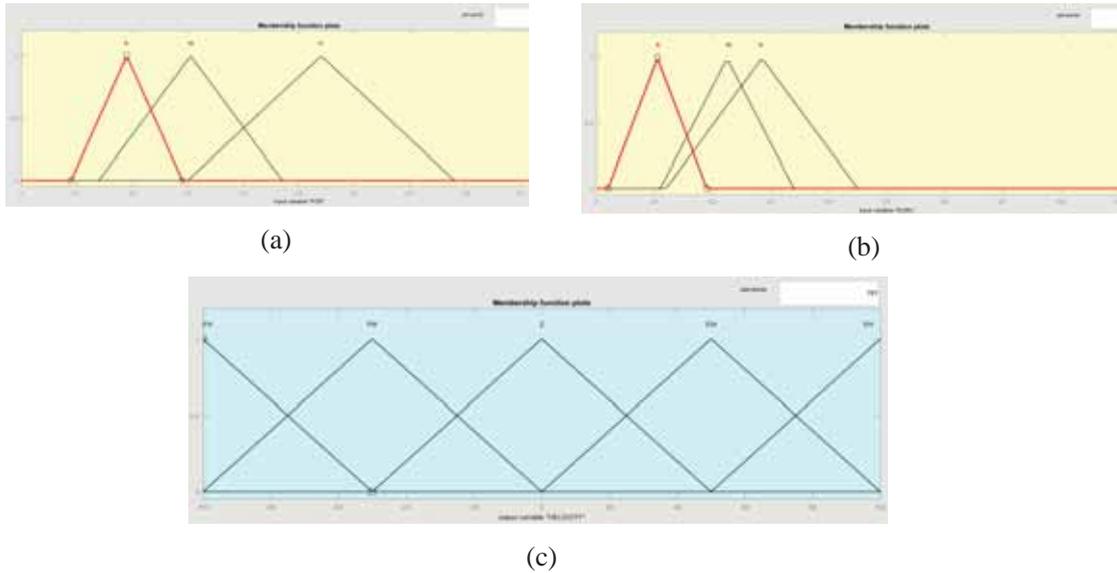

(a)

(b)

(c)

Fig. 10: (a) Function of triangular-shaped membership used for FCR Muscle, (b) Function of triangular-shaped membership used for ECRL Muscle, (c) Function for triangular membership used for the desired wrist velocity.

The 9 linguistic claims stated in rule Table 3 were synthesised into a Mamdani-style fuzzy inference system that mapped inputs to outputs. In these phrases, the conjunction AND represents a minimum operator that selects only the lowest of the fuzzy inputs. For each IF-THEN rule statement, an implication operation was performed using the "minimum" method, which directly truncated the fuzzy output sets. Instead, a complex-shaped curve was generated at the conclusion by aggregating all fuzzy output sets that have been trimmed using the "maximum" method.

The FCR muscle was more aggressively stimulated when the wrist was in the flexion position, but the ECRL responded more aggressively when the wrist was in the extension position. In other instances, despite being in the opposite position, both muscles still produced EMG output signal voltages, but the reading was lower than the others. In contrast, when the wrist was in a neutral position, both EMG signal muscles should produce identically stimulated voltages of signal strength from both muscles.

Table 3: 9 IF-THEN rule statements

No		FCR Muscle		ECRL Muscle		Wrist Velocity
1	if	S	and	S	then	Z
2	if	M	and	M	then	Z
3	if	H	and	H	then	Z
4	if	M	and	S	then	FM
5	if	H	and	M	then	FM
6	if	H	and	S	then	FH
7	if	S	and	M	then	EM
8	if	M	and	H	then	EM
9	if	S	and	H	then	EH

The FCR muscle was more aggressively stimulated when the wrist was in the flexion position, but the ECRL responded more aggressively when the wrist was in the extension position. In other instances, despite being in the opposite position, both muscles still produced EMG output signal voltages, but the reading was lower than the others. In contrast, when the wrist was in a neutral position, both EMG signal muscles should produce identically stimulated voltages of signal strength from both muscles.

5.2.1 Fuzzy Logic Output Design for Dynamic Modelling on Double Muscle at Similar MVC Level

Figures 11 and 12 show both fuzzy output results for wrist desired velocity. As illustrated by Fig. 11, at a signal level of 20% MVC, it can be estimated that the data division readings represented on each wrist joint angle position tended to overlap with each other, hence producing an unstable desired wrist velocity output result. Fig. 12 depicts a considerable desired wrist velocity fuzzy output result, as each component of the wrist position can be easily recognized at 60% MVC. This contributed when the EMG signal value excitation began to exhibit its considerable group uniqueness characteristic during the feature extraction process, thus making the fuzzy desired wrist velocity output data easier to comprehend. The fuzzy logic output for Fig. 7, 8, 11, 12 showed a certain range that was significantly comparable for wrist desired velocity and wrist joint angle data collected in the experiment.
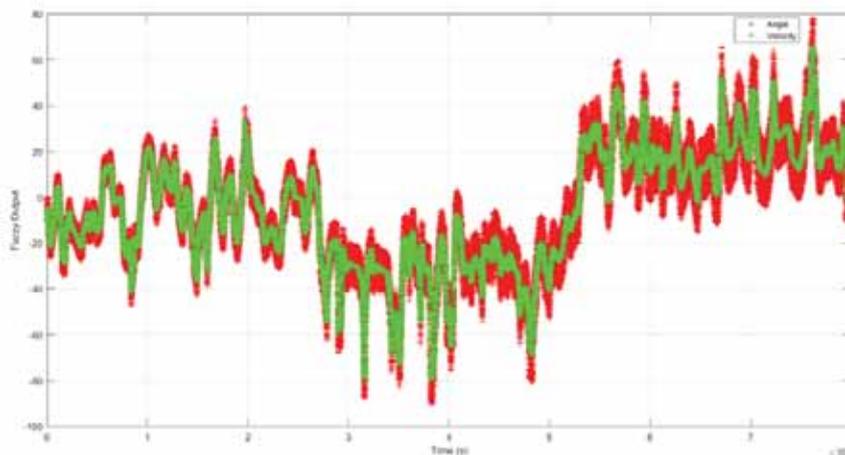

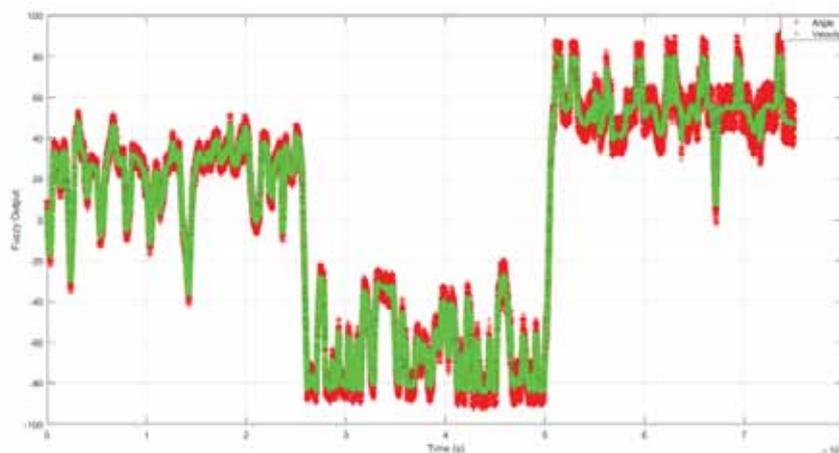Fig. 11: Fuzzy output for 20% MVC.

Fig. 12: Fuzzy output for 60% MVC.

6. PID CONTROL

PID controllers, which stands for proportional, integral, and derivative, employ a control loop feedback method to regulate process variables. The end-effector controller is configured according to Eq. (2):

$$f = J_a^T T_a = K_P{}^e + K_D^{\dot{e}} + K_I \int e\, dt \qquad (2)$$

where Kp, K_D, K_I are some positive definite tuned parameter matrices. To compute the actuator torque Ta, use $(J_a^T)^+$ to obtain as shown in Eq. (3):

$$T_a = (J_a^T)(K_{Pe} + K_{D\dot{e}} + K_I \int e\, dt \qquad (3)$$

The integral terms include eliminating or reducing the steady state error for tracking a step input signal. In actual implementation, the integral term is usually limited by some bounds. The main function of PID is to compensate for the fluctuation signal coming from the output mapping process, hence provide more stable continuous output control for the exoskeleton hand. In this paper, metaheuristic algorithm has been applied to obtain the value for Kp, K_I and K_D [15]. The value for Kp = 3.97 x 10^{-9}, K_I = 0.11 x 10^0 and K_D = 1.44 x 10^{-6}.

7. RESULTS AND DISCUSSIONS

7.1 Analysis of the Design Performance

The following step involves importing the SolidWorks file into MATLAB. At this point, any Simulink block, including Fuzzy Logic and PID, can be utilized to build a system to estimate the desired wrist velocity for the exoskeleton hand [16]. PID was utilised to control the output from the fuzzy logic to estimate the desired wrist velocity. For both fuzzy dynamic modelling approaches, identical plant was employed. The modifications were accomplished by modifying the input and fis files for fuzzy logic. Figure 13 depicts the completed development of the SimMechanics testing system for physical modelled of exoskeleton hands.
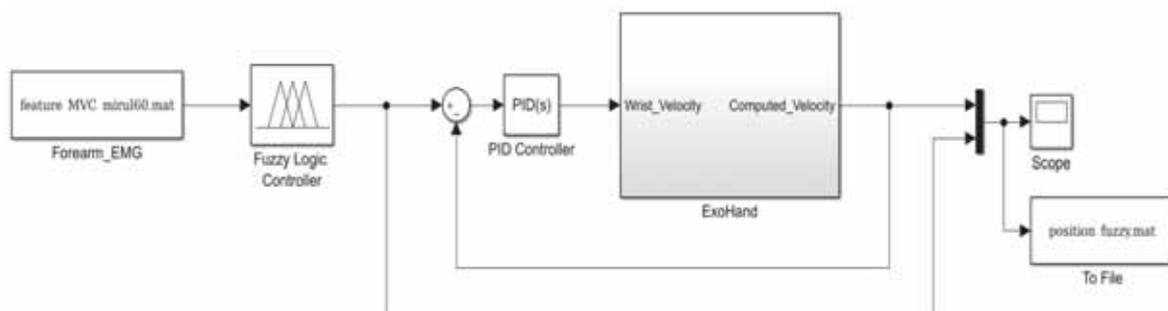

Fig. 13: Full system of exoskeleton hand for wrist movement with FL and PID controller.

7.2 Analysis of the Design Performance Using PID

7.2.1 Analysis of Dynamic Modelling on Single Muscle at Different MVC Level

Figures 14 and 15 show the Fuzzy Logic and PID controller and fuzzy logic output result based on the performance of a single muscle affected on predicting wrist desired velocity of the exoskeleton hand system. Initially, fuzzy logic was used as the system's dynamic model. However, the output from the FL mapping process was fed to the exoskeleton hand modelled resulted in an unstable wrist velocity output result. The red line from both graphs originated from the FL decision making process. In Fig. 14, an obvious

case where an overshoot signal happened located between 200s to 400s while undershoot happened at 1400 s to 1600 s. In Fig. 15, a clear case of overshoot happened at 700s to 900s while undershoot case happened between 900 s to 1100 s. Due to this instability in the FL result, the PID controller was utilised to compensate for the fluctuation of the predicted output value. The blue line graph in each figure illustrates the exoskeleton hand output result after the PID controller was developed. The results were a more stable output for the exoskeleton hand, compensating for the fluctuations happening in the FL mapping process thus smoothing the movement at the desired exoskeleton hand wrist velocity.
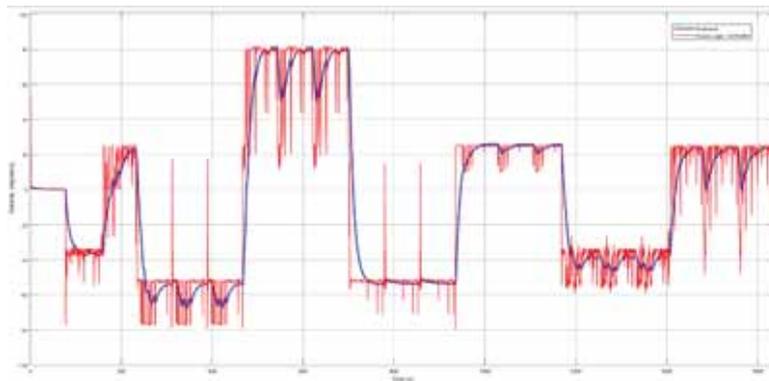

Fig. 14: Fuzzy Logic and PID controller for Estimated wrist desired velocity of Exoskeleton hand using FCR muscle.
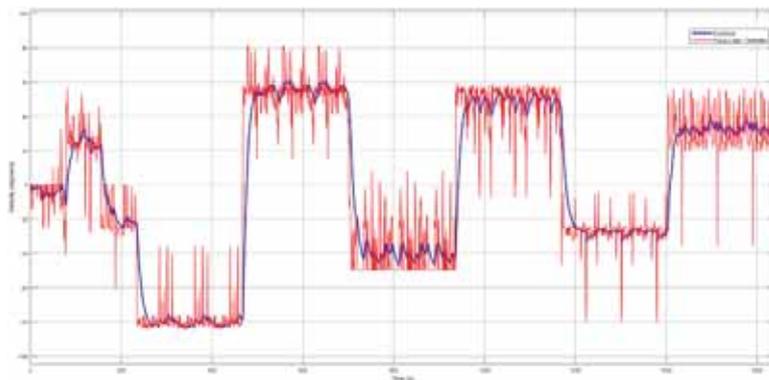

Fig. 15: Fuzzy Logic and PID controller for Estimated wrist desired velocity of Exoskeleton hand using ECRL muscle.

7.2.2 Analysis of Dynamic Modelling on Double Muscles at Similar MVC Level

Figures 16 and 17 show the desired wrist velocity output of the Fuzzy Logic and PID controller based on the performance of double muscles for the exoskeleton hand system. Fuzzy logic was chosen as the mapping technique for the exoskeleton hand system. This chosen method compared measurement data from each muscle's excitation group and used the results to predict the desired wrist velocity. Fig. 16 and 17 illustrate a red line graph representing a fluctuation of fuzzy output results. For both graphs, the pattern in each wrist desired velocity were related to EMG signal generation. In Fig. 16, the pattern generations were not too noticeable since the signal values were produced in closer range separation while in Fig. 17, the signal values produced were for larger range separation. To compensate for the fluctuations generated from Fuzzy logic's output, the PID was chosen as a controller. Through the use of the PID controller, the blue line graph in both Fig. 16 and 17 existed to smooth the fluctuation generated by the FL decision making process, hence producing

distinct wrist desired velocity output results at various wrist joint angle positions for the exoskeleton hand system.
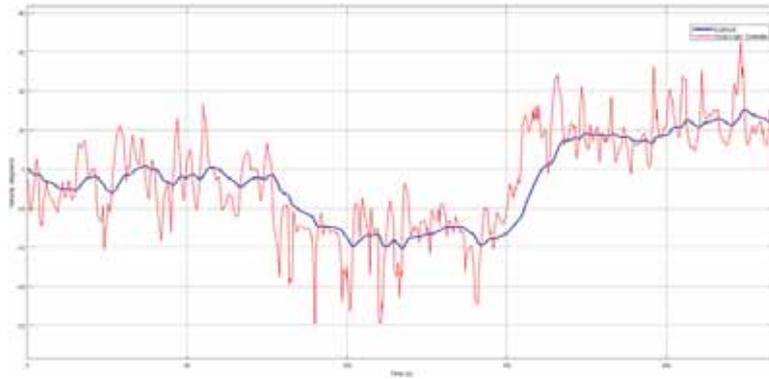

Fig. 16: Fuzzy Logic and PID controller for Estimated wrist desired velocity of Exoskeleton hand using FCR and ECRL muscles at 20% MVC.
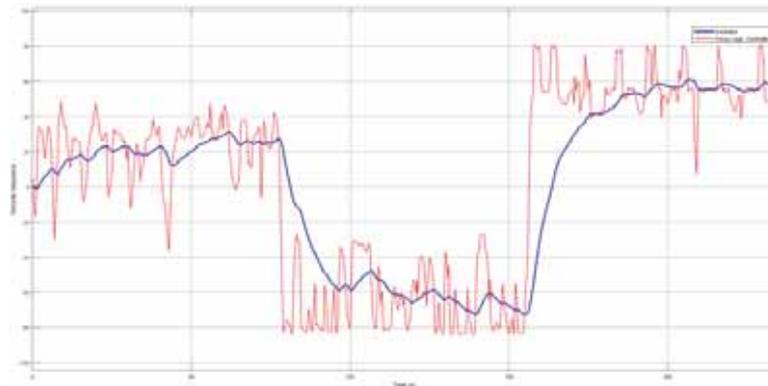

Fig. 17: Fuzzy Logic and PID controller for Estimated wrist desired velocity of Exoskeleton hand using FCR and ECRL muscles at 60% MVC.

8. CONCLUSION

Wrist hand position movement has become an integral component of human intentions in daily activities. Realizing the significance of this component, it is essential to emphasise the use of this wrist desired velocity estimation that follows the user's motion intention as employed in this paper, where it is also utilised for an exoskeleton hand and other rehabilitation applications. The relationship between and wrist joint angle hand position and EMG signal excitation can also be used to predict the assessment of the desired wrist velocity for the exoskeleton hand system.

Along the process to estimate the wrist desired velocity of the exoskeleton hand system, the feature extraction method needed to be selected properly. The collected data must pass through this process since it has been used to extract the usable data set to represent the EMG signals that are proportionally related to wrist joint angle position and desired wrist velocity of exoskeleton hand movement. Based on the calculation of standard deviation, Waveform length (WL) was chosen for the feature extraction method since it provides lower standard deviation value than Root Mean Square (RMS), Mean Absolute Value (MAV), Integrated EMG (IEMG), and Zero Crossing (ZC) [17]. By analysing these standard deviation values, the tabulation of data located in close proximity to one another produced more accurate estimates of the desired wrist velocity prior to the mapping stage [18].

The use of fuzzy logic as a dynamic modelling of the system gave the author an authority to design the exoskeleton hand system based on author logic control. However, the mapping process produced a fluctuation in estimating the desired wrist velocity output results. To compensate for the fluctuation generated, PID was chosen as a controller approach for this paper because of its simplicity and dependability in industrial applications. This controller has proven to reduce output volatility and stabilised desired wrist velocity output results. However, the exoskeleton hand's final results still contained some errors, as the hand's desired wrist velocity changed. An optimization method, such as Particle Swarm Optimization (PSO), Invasive Weed Optimization (IWO) or Firefly Algorithm (FA), could be utilised in future planning to adapt to nonlinear input derived from EMG data and wrist joint angle position and desired wrist velocity.

ACKNOWLEDGMENT

REFERENCES

[1] McGowan B, "Industrial exoskeletons: what you're not hearing," Occup. Heal. Saf. Mag., vol. 1, 2018.
[2] Wahit MAA, Ahmad SA. (2018) Design and development of low-cost exoskeleton hand robot structure. IEEE Student Conf. Res. Dev. Inspiring Technol. Humanit. SCOReD 2017 - Proc., vol. 2018-Jan, pp. 45-49. doi: 10.1109/SCORED.2017.8305423.
[3] Asogbon MG, Samuel OW, Jiang Y, Wang L, Geng Y, Sangaiah AK, Chen S, Fang P, and Li "Appropriate feature set and window parameters selection for efficient motion intent characterization towards intelligently smart emg-pr system," Symmetry (Basel)., vol. 12, no. 10, pp. 1–20, 2020, doi: 10.3390/sym12101710.
[4] Zhang L, Qi W, Hu Y, and Chen Y, "Disturbance-observer-based fuzzy control for a robot manipulator using an EMG-driven neuromusculoskeletal model," Complexity, vol. 2020, 2020, doi: 10.1155/2020/8814460.
[5] Chen C, Guo W, Ma C, Yang Y, Wang Z, and Lin C, "Semg-based continuous estimation of finger kinematics via large-scale temporal convolutional network," Appl. Sci., vol. 11, no. 10, 2021, doi: 10.3390/app11104678.
[6] Mendez V, Pollina L, Artoni F and Micera S, "Deep learning with convolutional neural network for proportional control of finger movements from surface EMG recordings," Int. IEEE/EMBS Conf. Neural Eng. NER, vol. 2021-May, pp. 1074–1078, 2021, doi: 10.1109/NER49283.2021.9441095.
[7] Kilic E, Dogan E. (2017) Design and fuzzy logic control of an active wrist orthosis. Proc. Inst. Mech. Eng. Part H J. Eng. Med., 231(8): 728-746. doi: 10.1177/0954411917705408.
[8] Xu F, Zheng Y, and Hu X, "Estimation of joint kinematics and fingertip forces using motoneuron firing activities: A preliminary report," Int. IEEE/EMBS Conf. Neural Eng. NER, vol. 2021-May, pp. 1035–1038, 2021, doi: 10.1109/NER49283.2021.9441433.
[9] Behrens M, Husmann F, Moeller AM, Schlegel J, Reuter EM, and Zschorlich1 VR, "Neuromuscular properties of the human wrist flexors as a function of the wrist joint angle," Front. Bioeng. Biotechnol., vol. 7, no. AUG, 2019, doi: 10.3389/fbioe.2019.00181.
[10] Kim S, Shin DY, Kim T, Lee S, Hyun JK, Park SM. (2022) Enhanced recognition of amputated wrist and hand movements by deep learning method using multimodal fusion of electromyography and electroencephalography. Sensors, 22(2): 1-17. doi: 10.3390/s22020680.

[11] Lara J, Paskaranandavadivel N, and Cheng LK, "HD-EMG Electrode Count and Feature Selection Influence on Pattern-based Movement Classification Accuracy," Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, vol. 2020-July, pp. 4787–4790, 2020, doi: 10.1109/EMBC44109.2020.9175210.

[12] Hayashi H, Shibanoki T, and Tsuji T, "A Neural Network Based on the Johnson SU Translation System and Related Application to Electromyogram Classification," IEEE Access, vol. 9, pp. 154304–154317, 2021, doi: 10.1109/ACCESS.2021.3126348.

[13] Subasi A and Qaisar SM, "Surface EMG signal classification using TQWT, Bagging and Boosting for hand movement recognition," J. Ambient Intell. Humaniz. Comput., vol. 13, no. 7, pp. 3539–3554, 2022, doi: 10.1007/s12652-020-01980-6.

[14] Copaci D, Serrano D., Moreno L, and Blanco D, "A high-level control algorithm based on sEMG signalling for an elbow joint SMA exoskeleton," Sensors (Switzerland), vol. 18, no. 8, 2018, doi: 10.3390/s18082522.

[15] Joseph SB, Dada EG, Abidemi A, Oyewola DO, and Khammas BM, "Metaheuristic algorithms for PID controller parameters tuning: review, approaches and open problems," Heliyon, vol. 8, no. 5, p. e09399, 2022, doi: 10.1016/j.heliyon.2022.e09399.

[16] Tageldeen MK, Perumal N, Elamvazuthi I, and Ganesan T, "Design and control of an upper arm exoskeleton using Fuzzy logic techniques," 2016 2nd IEEE Int. Symp. Robot. Manuf. Autom. ROMA 2016, pp. 1–6, 2017, doi: 10.1109/ROMA.2016.7847838.

[17] Jia G, Lam HK, Liao J, Wang R. (2020) Classification of electromyographic hand gesture signals using machine learning techniques. Neurocomputing, 401: 236-248. doi: 10.1016/j.neucom.2020.03.009.

[18] Bardizbanian B, Zhu Z, Li J, Huang X, Dai C, Luna CM, McDonald BE, Farrell TR, and Clancy EA, "Efficiently Training Two-DoF Hand-Wrist EMG-Force Models," Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, vol. 2020-July, pp. 369–373, 2020, doi: 10.1109/EMBC44109.2020.9175675.

FEATURE EXTRACTION AND SUPERVISED LEARNING FOR VOLATILE ORGANIC COMPOUNDS GAS RECOGNITION

NOR SYAHIRA MOHD TOMBEL[1], HASAN FIRDAUS MOHD ZAKI[2,3]*
AND HANNA FARIHIN BINTI MOHD FADGLULLAH[3]

*[1]Dept. of Science (Computational and Theoretical), Kulliyyah of Science,
International Islamic University of Malaysia, Kuantan, Pahang
[2]Centre for Unmanned Technologies (CUTe), Kulliyyah of Engineering,
International Islamic University of Malaysia, Gombak, Kuala Lumpur
[3]Dept. of Mechatronic, Kulliyyah of Engineering,
International Islamic University of Malaysia, Gombak, Kuala Lumpur*

*Corresponding author: hasanzaki@iium.edu.my

ABSTRACT: The emergence of advanced technologies, particularly in the field of artificial intelligence (AI), has sparked significant interest in exploring their potential benefits for various industries, including healthcare. In the medical sector, the utilization of sensing systems has proven valuable for diagnosing pulmonary diseases by detecting volatile organic compounds (VOCs) in exhaled breath. However, the identification of the most informative and discriminating features from VOC sensor arrays remains an unresolved challenge, essential for achieving robust VOC class recognition. This research project aims to investigate effective feature extraction techniques that can be employed as discriminative features for machine learning algorithms. A preliminary dataset was used to predict VOC classification through the application of five supervised machine learning algorithms: k-Nearest Neighbors (kNN), Random Forest (RF), Support Vector Machines (SVM), Logistic Regression (LR), and Artificial Neural Networks (ANN). Ten feature extraction methods were proposed based on changes in sensor response as inputs to classify three types of gases in the dataset. The performance of each model was evaluated and compared using k-Fold cross-validation (k=10) and metrics derived from the confusion matrix. The results demonstrate that the RF model achieved the highest mean accuracy and standard deviation, with values of 0.813 ± 0.035, followed closely by kNN with 0.803 ± 0.033. Conversely, LR, SVM (kernel=Polynomial), and ANN exhibited poor performances when applied to the VOC dataset, with accuracies of 0.447 ± 0.035, 0.403 ± 0.041, and 0.419 ± 0.035, respectively. Therefore, this paper provides evidence that classifying VOC gases based on sensor responses is feasible and emphasizes the need for further research to explore sensor array analysis to enhance feature extraction techniques.

ABSTRAK: Perkembangan teknologi canggih, khususnya dalam bidang kecerdasan buatan (AI), telah mencetuskan minat yang ketara dalam menerokai manfaatnya untuk pelbagai industri, termasuk bidang kesihatan. Dalam sektor perubatan, penggunaan sistem penderiaan telah terbukti bernilai untuk mendiagnosis penyakit paru-paru dengan mengesan sebatian organik meruap (VOC) dalam nafas yang dihembus manusia. Walau bagaimanapun, pengenalpastian ciri yang paling bermaklumat dan mendiskriminasi daripada penderia VOC kekal sebagai cabaran yang tidak dapat diselesaikan, penting untuk mencapai pengiktirafan kelas VOC yang kukuh. Projek penyelidikan ini bertujuan untuk menyiasat teknik pengekstrakan ciri yang berkesan yang boleh digunakan sebagai ciri diskriminatif untuk

algoritma pembelajaran mesin. Set data awal digunakan untuk meramalkan klasifikasi VOC melalui aplikasi lima algoritma pembelajaran mesin yang diselia: k-Nearest Neighbors (kNN), Random Forest (RF), Support Vector Machines (SVM), Logistic Regression (LR), dan Artificial Neural Networks (ANN). Sepuluh kaedah pengekstrakan ciri telah dicadangkan berdasarkan perubahan dalam tindak balas penderia sebagai input untuk mengklasifikasikan tiga jenis gas dalam set data. Prestasi setiap model telah dinilai dan dibandingkan menggunakan pengesahan silang *k-Fold* (k=10) dan metrik yang diperoleh daripada *confusion* matriks . Keputusan menunjukkan bahawa model RF mencapai ketepatan minima tertinggi dan sisihan piawai, dengan nilai 0.813 ± 0.035, diikuti oleh kNN dengan 0.803 ± 0.033. Sebaliknya, LR, SVM (kernel=Polinomial), dan ANN mempamerkan prestasi yang lemah apabila digunakan pada dataset VOC, dengan ketepatan masing-masing 0.447 ± 0.035, 0.403 ± 0.041 dan 0.419 ± 0.035. Oleh itu, kertas kerja ini memberikan bukti bahawa mengklasifikasikan gas VOC berdasarkan tindak balas penderia adalah boleh dilaksanakan dan menekankan keperluan untuk penyelidikan lanjut untuk meneroka analisis tatasusunan penderia untuk meningkatkan teknik pengekstrakan ciri.

KEYWORDS: *Supervised machine learning; Volatile Organic Compound; VOC Sensor; Gas classification; feature extraction*

1. INTRODUCTION

Volatile organic compounds (VOC) have been used as preclinical biomarkers in breath analysis to monitor health and diagnose various pulmonary diseases such as asthma and lung cancer [1] [2] [3][4]. An array of sensors, or electronic nose (e-nose) is known as the alternative for a non-invasive method of detecting volatile organic compounds (VOC). E-nose is a device inspired by the olfactory system of humans or mammals (sense of smell), composed of a collection of an array of gas sensors with a pattern recognition system designed to detect and differentiate a wide variety of gas compounds [5].

The advancement of nanosensor arrays with pattern recognition involving pre-processing, feature extraction and machine learning algorithms makes it a powerful tool for the detection and recognition of gas samples with concentration estimation. Feature extraction is an essential technique used to extract significant information from the sensor response signal [6] [7] to optimize the performance of pattern recognition algorithms for gas classification [6] [8].

However, the detection of VOC using nanosensor technologies still has some constraints in its detection system. The VOC sensor as a sensing unit faced a few limitations such as lack of sensitivity and selectivity [9] [10]. Besides, it is still not clear which type of features from VOC sensor arrays are the most descriptive and discriminative leading to a robust recognition of the VOC classes. Data collection from a gas sensor array can also be cumbersome and time-consuming which poses a nuisance in employing data-hungry machine learning algorithms.

Therefore, this paper proposes employing supervised machine learning algorithms to classify the preliminary data of the individual sensors for VOC recognition. The VOC detection was performed on a chemiresistive sensor from various functionalised reduced Graphene Oxide (rGO) as a sensing layer. The targeted VOC gases used are acetone, toluene and isoprene which have been suggested as pulmonary disease-related biomarkers [11] with concentration levels ranging from 1 to 6 ppm.

Concretely, we explore 10 feature sets that were extracted from the sensor's original response curve. Then, we analyse the effect of these features towards VOC classification with five benchmark machine learning algorithms including K-Nearest Neighbours (kNN), Random

Forest (RF), Logistic Regression (LR), Support Vector Machine (SVM) and Artificial Neural Network (ANN). The recognition models were then put into comparison to determine the one which provides the best evaluation and high accuracy in performing the classification of the targeted VOC gases using k-Fold Cross Validation (k=10) and Confusion Matrix.
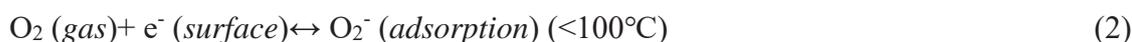
2. GAS SENSING MECHANISM

Sensing mechanism of the sensor is first studied, to understand the VOC detection on the sensor. The thin film is comprised of rGO which is one of the most promising materials for detecting low VOC concentrations at room temperature [12]. Graphene is a two-dimensional building block made up of a one-atom-thick sheet of a carbon atom.

Graphene can work well at room temperature because it has enormously high mobility [13]. Researchers are interested in modifying graphene into reduced Graphene Oxide as a sensing element because of its excellent electrical, high thermal conductivity, and mechanical properties [14] [15]. The functionalisation of rGO with nanoparticles and plasma treatment can improve sensor functionality and selectivity in distinguishing different vapours [11]. Different functionalisation of sensing elements is a good technique to improve the gas sensor's sensitivity and characteristics.

In this research, the sensing layer was deposited on Ti/Pt Interdigitated Electrode (IDE). The electrode was used to supply current flow from the power source to the device, which improved the sensing material's catalytic properties towards a specific gas [16]. Furthermore, the VOC sensor employed is a resistive type, which produces a signal based on a change in resistance in response to gas exposure. In general, VOC gas detection on a sensor is caused by the adsorption and desorption processes that occur between analytes and the sensor surface [17].

Oxygen ion species were absorbed on the sensor surface in the presence of air (humidity) and lowered the electron from the conduction band [18]. The electron density is falling off and forming an electron depletion layer and barrier potential on the surface. Electron removal causes an increase in the depletion layer. The related equation for chemisorbed oxygen at temperatures less than 100°C [19] is as follows:

$$O_2 \ (gas) \rightarrow O_2 \ (adsorption) \tag{1}$$

$$O_2 \ (gas) + e^- \ (surface) \leftrightarrow O_2^- \ (adsorption) \ (<100°C) \tag{2}$$

When VOC gas was introduced into the chamber, the gas molecules started to react with the absorbed oxygen ions and released electrons back into the conduction band. The predominant carrier in the sensors was modified by the reaction of the VOC gas (oxidizing or reducing agent) with the molecules in the sensing layer, resulting in an increasing or decreasing in the resistance measurement as the output [19].

Reduced Graphene Oxide has been reported to exhibit p-type behaviour [20]. However, the functionalised sensor was shown to be an n-type semiconductor in this VOC test, and the VOC analytes acted as reducing gases [18]. The sensor experienced an electron carrier majority, causing a decrease in depletion width and potential barrier. As a result, the sensor resistance decreased in the presence of VOC gas [21].

3. RELATED WORKS

3.1. Feature Extraction

Feature extraction is a technique that is used to extract significant information from the sensor response graph [6] to ensure better performance of machine learning algorithms in pattern recognition [8]. The information is deemed relevant when the derived value extracted from the measured data is non-redundant, not correlated with other features and projects the decisive features [22]. Other than that, feature selection is also related to the dimensionality reduction process of transforming high dimensional data into a low dimensional feature [8].

Detection of VOCs using gas sensors commonly used real-time analysis and discrimination of "breath prints" to perform the gas classification process [2]. In 2012, Vergara and his team applied 8 feature extraction from the time-series sensor, which are the change of the maximal resistance change (ΔR), the normalized resistance change ($\|\Delta R\|$), minimum and maximum exponential moving average(ema) with a value of $\alpha = 0.001, 0.01, 0.1$ each [10].

On the other hand, many features can be extracted from raw signals and applied in electronic nose applications. Commonly extracted features from gas original response curves such as maximum response, the response of special time, time of special response, area, integral, derivative, difference and second derivative [6]. Table 1 shows a few lists of feature extraction from electronic nose sensor data for wound detection [23].

Table 1: List of Feature Extraction for Wound Detection based Electronic Nose [23]

Feature	Description
Normalization	Preprocessing the sensor data for features from the steady-state response, eliminate the effect of a concentration difference on recognition.
Integral and derivative methods	Integrals may represent the accumulative total of the reaction degree change and derivatives may represent the rate at which the sensor reacts to the odour.
The Root Mean Square Error (RMSE) of curve fitting	Depends on the type of model and the number of parameters in the model.
Fourier transform and wavelet transform	Fourier transform decomposes the original response curve into a superposition of the DC component and different harmonic components.

3.2. Supervised Machine Learning

There were few studies which implemented the detection of different gases by Supervised learning models such as k-Nearest Neighbour (kNN), Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest (RF) and Logistic Regression. The findings were summarised in Table 2.

There are two gaseous flows in the system: for carrier gas and VOC gas. Clean Dry Air (CDA) was used as carrier gas, while isoprene, toluene, and acetone as the targeted VOC gas. The temperature of the gas and temperature chuck in the sensor chamber were controlled using a Cellkraft Humidifier P-10 and a Nextron Temperature controller module (Nextron Microprobe Station, with platier heater and 4 probe needles). Agilent SMU 34410A was used to drive the voltage and input current. A data acquisition (DAQ) system that is used to convert the output/measured signal from the sensor system into the computer is via a user interface software that is programmed using the LabVIEW program, provided by MIMOS.

Table 2: Summary of Implementation of Supervised Learning Algorithms on Different Types of Gas Detection.

Model	Description	References
k-Nearest Neighbour (Knn)	kNN is widely used in the classification of mixed gas and for gas discrimination systems. The kNN model is advantageous because it is comprehensible, insensitive to noise, low cost for retaining and good combination with other algorithms. However, this model is sensitive to sample distribution, it has a slow speed for recognition, high spatial complexity, heavy calculation burden and poor interpretability.	[24]
Support Vector Machines (SVM)	SVM of classifiers can cope well with gas sensor drift and perform better than the baseline competing methods on the extensive dataset. However, the SVM model requires a long learning time and poor application for larger data. Choice of kernel function is important as it is the key for feature space in SVM.	[10],[24]
Artificial Neural Network (ANN)	ANN is the frequently used method in predicting and analysing complex gas (Hashoul & Haick, 2019). It has good learning ability, good parallel processing capability and detecting compatibility error. However, this model has poor interpretability for output, long time learning and is easy to overfit. Therefore, weight, activation function and the number of hidden layers are important to develop an ANN algorithm in performing the classification of targeted output.	[24],[25]
Random Forest (RF)	RF model is used in a lot of feature datasets as it can prevent overfitting from a decision tree algorithm. In the Random Forest algorithm, the number of trees affected the accuracy of the model, as each tree has a classification result and the final result is based on the majority decision trees vote	[26], [36]
Logistic Regression	LR is a classification algorithm that calculates linear output and statistical function through the regression output. Logistic regression can perform multiclass classification problems by using one-vs-rest or one-vs-one wrapper models. The algorithm can be applied to a non-linear classification problem with a proper feature selection. LR model can produce high accuracy as it is a good signal to noise ratio.	[27]

4. EXPERIMENTAL SETUP

As illustrated in Fig 1, the gas sensing system for this study comprises a gas supply system, a sensor chamber, a temperature and humidity controller module, and data collection system [28].

Fig 2 showed the DAQ board of the LabVIEW that contain all the controller variables for the test measurement such as flow of the CDA, flow of the VOC gas, input current, input voltage, temperature inside the chamber, temperature of the sensor's heater, relative humidity, and system ramp rate. Then, the sensor was tested individually with the targeted VOC gas and the sensor's responses were recorded to study performance of the individual sensor.
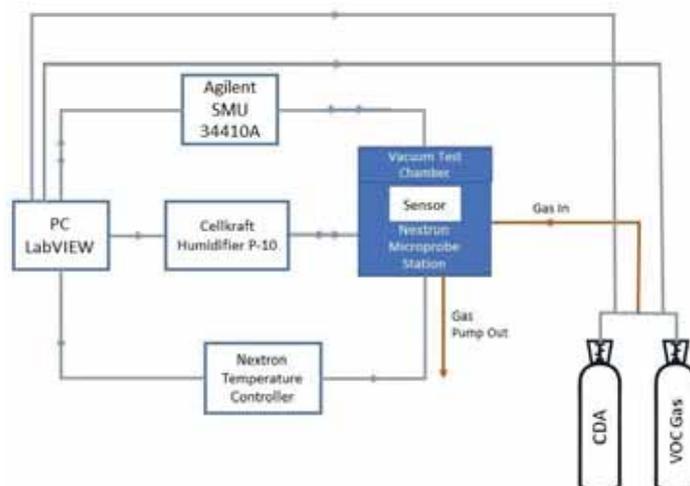
Fig. 1. Schematic Diagram of the Experimental Setup for Gas Detection System.

Fig. 2. Manual control of LabVIEW Gas System.

4.1. VOC Sensor

The gas sensor used in this study is called a VOC sensor, which is prepared, fabricated and functionalised by the engineering team at MIMOS Bhd. Reduced Graphene Oxide (rGO) as a sensing membrane was deposited on the Platinum-titanium interdigitated electrode (Pt/Ti IDE) on a silicon and silicon dioxide (Si/SiO$_2$) substrate. The rGO was functionalised with nanoparticles such as; gold (Au), silver (Ag) and platinum (Pt) and plasma treatment such as; hydrogen (H$_2$) and Octafluorocyclobutane (C$_4$F$_8$).

The sensor was fabricated using a standard semiconductor process using Chemical Vapor Deposition (CVD) by a standard lithography process for the functionalisation with different recipes. The rGO was functionalised with nanoparticles at a different duration of sputtering and Relative Frequency (RF) power, while functionalisation with plasma treatment at a variety of plasma power and temperature. Therefore, there are 21 individual VOC gas sensors used in this study and the details are according to Table 3. Next, the pre-processed signal proceeded with a feature extraction method to extract pertinent information to be input for supervised machine learning at classifying the gas components into targeted gas output. The features were

decided to extract from the original gas response involving measured resistance in the absence and presence of the VOC gas.

Table 3: List of Functionalisation and Recipe of VOC Sensors

Sample no.	Nanoparticles	Power [W_{RF}]	Time [sec]	Remarks
1	Reference rGO film			Bare Rgo
2	Au	30	15	rGO/Au (30W 15s)
3	Au	30	75	rGO/Au (30W 75s)
4	Au	70	15	rGO/Au (70W 15s)
5	Au	70	75	rGO/Au (70W 75s)
6	Pt	30	15	rGO/Pt (30W 15s)
7	Pt	30	75	rGO/Pt (30W 75s)
8	Pt	70	15	rGO/Pt (70W 15s)
9	Pt	70	75	rGO/Pt (70W 75s)
10	Ag	30	15	rGO/Ag (30W 15s)
11	Ag	30	75	rGO/Ag (30W 75s)
12	Ag	70	15	rGO/Ag (70W 15s)
13	Ag	70	75	rGO/Ag (70W 75s)
Sample no.	Plasma Treatment	Power (W_{RF})	Temperature (°C)	Remarks
14	H_2	-	RT	rGO/H_2 RT °C
15	H_2	-	200	rGO/H_2 200 °C
16	H_2	-	400	rGO/H_2 400 °C
17	H_2	-	700	rGO/H_2 700 °C
18	C_4F_8	150	-	rGO/C_4F_8 150 °C
19	C_4F_8	200	-	rGO/C_4F_8 200 °C
20	C_4F_8	250	-	rGO/C_4F_8 250 °C
21	C_4F_8	300	-	rGO/C_4F_8 300 °C

4.2. Data Collection

The sensors were tested individually with each of the selected VOC gas. The sensor was placed in a chamber with 30°C of temperature and presence of 40% relative humidity (RH). The voltage and current input were set at 1V and 1.2A respectively. CDA was maintained at 1 L/min for 5 minutes to stabilize the baseline reading. Then, the VOC gas was purged into the chamber with a gradual increase of concentrations, from 1 to 6 ppm in 12 minutes (2 minutes for each concentration). The sensor responses were analysed from the resistance changes of individual sensors that undergo the VOC test.

5. DATA PRE-PROCESSING AND FEATURE EXTRACTION

In this phase, the analytes of the VOC gas were reacting with the sensing element, thus leading to a change in resistance. The sensor response was determined by analysing the measured resistance as a signal output from each sensor. However, the parameter setup was not in optimal condition and the output signal contained unexpected noise from the SMU system. A typical sensor response could not be seen clearly from the graph of resistance versus time.

As a result, the signal was pre-processed by applying filter and smoothing methods to denoising the signal and reduce the influence of random variation caused by instrumental conditions and atmospheric effects [29]. The data was filtered using a moving average (MA

length = 3) and smoothed with Minitab software using a single exponential method with a constant = 0.02 value. The sensor response was determined by using the formula [30][31]:

$$\Delta R = R_g - R_a \tag{3}$$

$$S = \frac{\Delta R}{R_a} = \frac{R_g - R_a}{R_a} \times 100(\%) \tag{4}$$

Where,

R_a = resistance in clean dry air, without VOC gas

R_g = resistance with the exposure of VOC gas

Next, the pre-processed signal proceeded with a feature extraction method to extract pertinent information as input for supervised machine learning at classifying the gas components into targeted gas output. The features were decided to extract from the original gas response involving measured resistance in the absence and presence of the VOC gas. The 10 selected features as listed in Table 4.

Table 4: List of Features and Formula for the 10 Selected Features

No.	Feature	Formula
1	R_{gas}	Resistance at steady-state phase
2	R_0	Baseline resistance
3	Sensor Response, S	$\Delta R = \dfrac{R_{gas} - R_{air}}{R_{air}}$
4	Difference, ΔR	$\Delta R = R_{gas} - R_{air}$
5	Relative difference	$\dfrac{R_{gas}}{R_{air}}$
6	Log relative resistance value	$log(\dfrac{R_{gas}}{R_{air}})$
7	Normalisation	$\left|\dfrac{R_{gas} - R_{air}}{R_{air}}\right|$
8	G R_{gas}	$\dfrac{1}{R_{gas}}$
9	G R_0	$\dfrac{1}{R_{air}}$
10	Conductance difference	$G_{gas} - G_{air}$

The VOC dataset comprises ten feature extraction values and is organised into three categories. In summary, there are 918 total samples, with 252, 324, and 342 each for acetone, toluene, and isoprene gas, respectively.

6. SUPERVISED LEARNING FOR VOC GAS CLASSIFICATION

Five supervised learning models including k-Nearest Neighbour (kNN), Random Forest (RF), Logistic Regression (LR), Support Vector Machines (SVM) and Artificial Neural Networks (ANN) were benchmarked for the VOC gas classification. The model was implemented using the Python and Scikit-learn library. Each model's parameter settings are described in Table 5.

To avoid bias in the analysis, the dataset was first standardised in the range 0 to 1 to uniform the values with different scales by using the min-max normalisation technique [32].

Following that, the dataset was divided into 70% for the training set and 30% for the testing set. The performances of each model are then evaluated from confusion matrix-based measures in terms of accuracy, precision and using k-fold cross-validation technique, where k =10. K-fold Cross Validation is a cross validation technique used to evaluate the performance of a machine learning model by the resampling procedure. The training of the models proceeds using the k-1 parts and validation or testing errors from the remaining part [33].

Table 5: Parameter Setting of the Approached Supervised Machine Learning

Model	Parameter
K-nearest Neighbour	The K-value is decided as one, (k=1) and distance between two points is calculated by applying the distance metric formula (2), (mentioned in chapter 2), with p = 2, to manipulate the generalised distance to Euclidean Distance.
Random Forest	Grid search for the setting parameter, with n-estimator:100.
Artificial Neural Network	A shallow Neural Network was implemented, with a standard three-layer feed-forward network. For the hidden layer, the size was set up to 50 and used the ReLU activation function. While Softmax activation function for output layer with learning rate 0.001 and 50 epochs.
Support Vector Machine (RBF kernel)	Radial basis function (RBF) was selected as a kernel function for this SVM model, as defined in equation (6) (in Chapter 2) The kernel function, σ and regularisation, C used GridSearchCV from sci-kit- learn library to perform grid search for parameter setting.
Logistic Regression	The model used 'l2' for regularisation (penalty) and solver 'lbfgs'.

7. RESULT

Figures 3 a) to e) showed the results of the confusion matrices for each proposed supervised learning method. Meanwhile, Table 6 below showed the accuracies from the 10-fold cross validation from each model.

The diagonal values in the confusion matrix denoted the accuracy values of the gas classification to the targeted output [34]. Figure 3 shows that the kNN and RF models performed well in classifying each of the targeted VOCs. The kNN model correctly predicted all three gases with greater than 80% accuracy, while the Random Forest model predicted them with greater than 70% accuracy. On the other hand, Logistic Regression, Support Vector Machine (kernel = RBF) and Artificial Neural Network performed poor classification on the 3 VOCs gases. The SVM and ANN models misclassified the gas more to isoprene gas while LR model misclassified the toluene and isoprene gas.

It is noticeable in Table 6, the RF model showed the highest mean accuracy with 0.813 ±0.035, followed by the kNN model with 0.803 ±0.033. RF model are known to have advantages in the process of random sampling which can ensure randomness and avoid overfitting. Besides, this model is also robust to noise [5] and it is good at handling missing data and imbalance classes [4].

Whereas, the kNN model has limitations in understanding the relationship between the features and the class (output) thus easily producing the wrong classification for a multiclass problem [4]. Therefore, the highest accuracy achieved by kNN in this study showed that the features were related well to the output class.
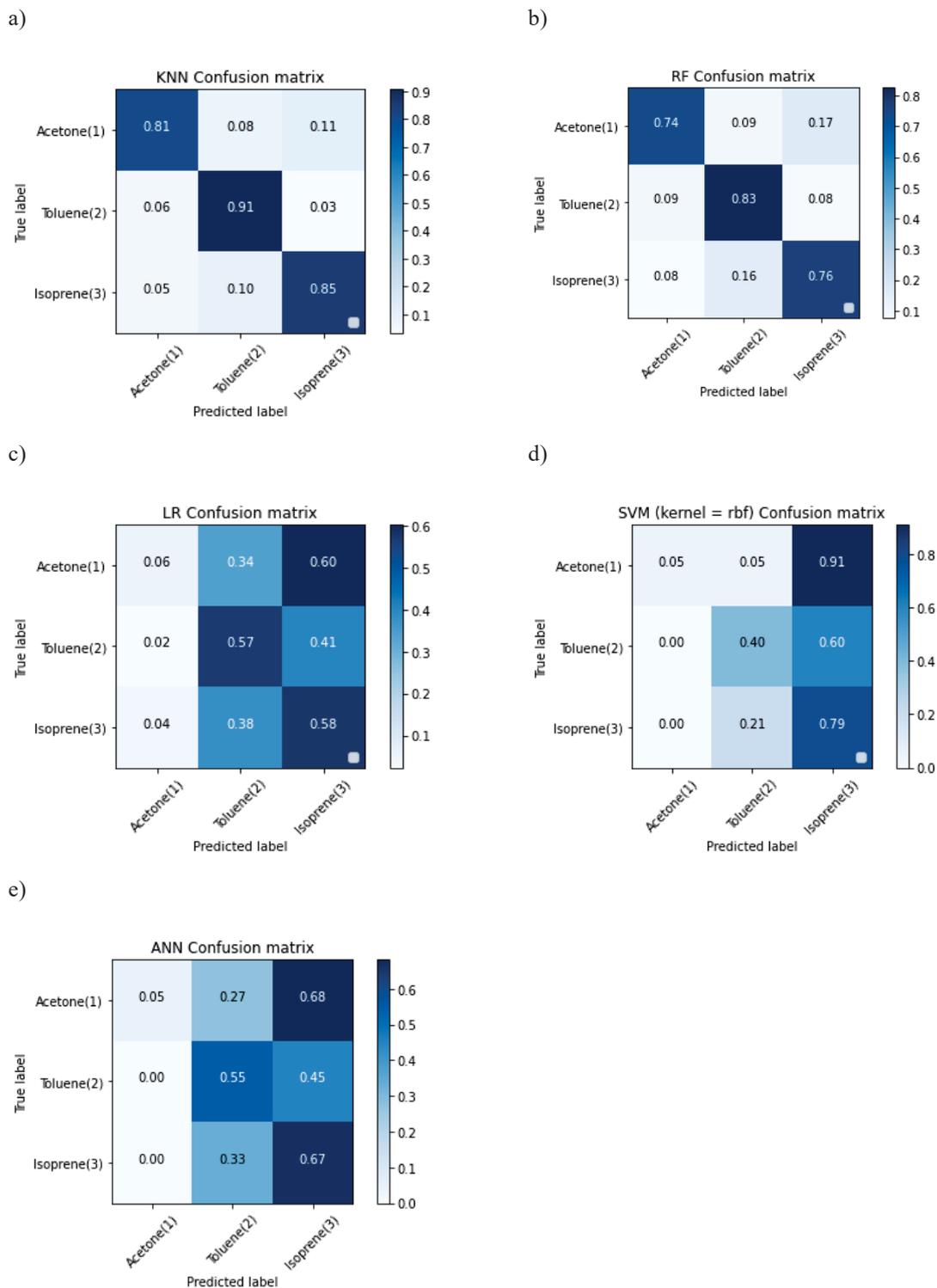
a)

b)

c)

d)

e)

Fig. 3. Comparison of Normalised Confusion Matrix of a) kNN, b) RF, c) LR, d) SVM
(kernel= RBF) and e) ANN.

However, the ANN, LR and SVM (Polynomial kernel) models had poor performance compared with the RF and kNN model with accuracies of 0.447 ± 0.035, 0.403 ± 0.041 and 0.419 ± 0.035 respectively. Meanwhile, the SVM model with Grid search parameters showed

high accuracy by using a Polynomial kernel at degree = 3, compared with other kernels such as Linear and RBF.

Table 6: Model Performance based on 10-Fold Cross-Validation Technique.

Model	10-Fold Cross-Validation (Mean Accuracy $\pm$ Standard Deviation)
Random Forest	0.813 $\pm$ 0.035
K-Nearest Neighbors	0.803 $\pm$ 0.033
Logistic Regression	0.403 $\pm$ 0.041
Support Vector Machine (Kernel = Polynomial)	0.419 $\pm$ 0.035
Support Vector Machine (Kernel = Linear)	0.401 $\pm$ 0.039
Support Vector Machine (Kernel = RBF)	0. 408 $\pm$ 0.055
Artificial Neural Network	0.447 $\pm$ 0.035

The poor performance from the ANN, LR and SVM (Polynomial kernel) models might be due to their weakness, in which they are very prone to overfitting training data [33] and they required testing with various kernels and model parameters [4]. The ANN model also is a learning-based algorithm and is more complex in architecture. Thus, it has more hyperparameters required to be tuned [33] and it needs enough samples for training.

Other than that, the ROC curve and AUC value are another way of visualising the output performances from the computed confusion matrix. Evaluation of the Receiver Operating Characteristics (ROC) and Area Under Curve (AUC) was done to analyse the performance of the classifiers. The highest value of the AUC showed good value prediction of the model to assign a larger probability to a random positive example than a random negative example [35]. The AUC value should be between 0.5 and 1.0. The ROC curves for each classifier are illustrated in Figure 4. As the minimum is 0.541, it can be said that the SVM classifier does not predict our dataset very well and could not differentiate the classes, while the highest value of AUC goes to the KNN classifier which is equal to 0.886 and 0.885 for the RF model. As a result, in this research, the kNN and RF are two models that can deal with the selected features in the VOC dataset as they obtained the highest accuracy for the gas classification.
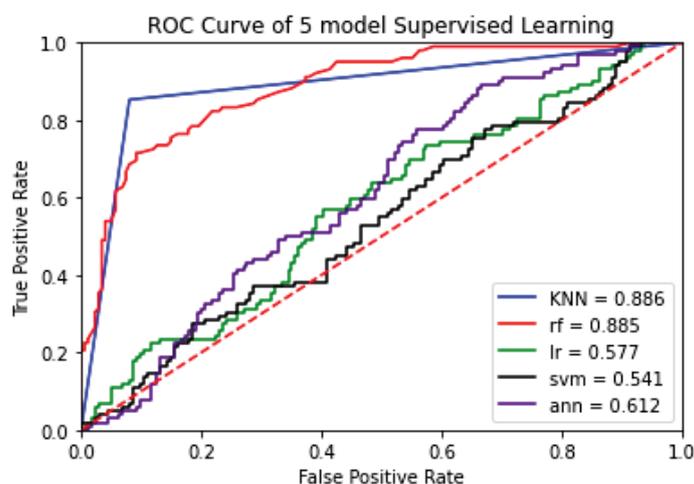

Fig. 4. ROC Curve of the 5 Supervised Learning Model

8. CONCLUSION

The gas sensor data was collected at a preliminary stage and has been used for the machine learning part, which involved pre-processing, feature extraction and classification algorithm. Each sensor was performing well at low operating temperatures and in the presence of humidity. The sensors response on different targeted VOC gas from 1 to 6 ppm were collected. Then, feature extracted were performed on the resistance-based data.

Then, feature extracted were performed on the resistance-based data. Ten featured were proposed as inputs to five supervised learning algorithms to accurately recognise and classify the selected VOC gas based on the labelled output. The confusion matrix and 10-Fold Cross Validation were used to evaluate each model's performance. As a result, the RF and kNN models have higher accuracy with 0.813 ± 0.035 and 0.803 ± 0.033, compared with LR, SVM and ANN with the accuracy of 0.447 ± 0.035, 0.403 ± 0.041 and 0.419 ± 0.035 respectively. The two highest accuracies achieved by RF and kNN models demonstrated that they distinguished the gas well from the VOC dataset.

Despite the gas sensor's shortcomings, such as low sensitivity, selectivity, and noise in the sensor signal output, the findings of this study can be utilised as a guide for selecting the optimum algorithm for dealing with a gas sensor array. The performance of the kNN and RF models is the Proof of Concept that the algorithm can perform gas classification tasks from the simplest feature selected from the steady-state phase. The feature extraction approach, on the other hand, can be discovered more from the raw signal to build a dataset with more significant features and relevant information to improve the algorithm's performance.

ACKNOWLEDGEMENT

REFERENCES

[1] Krisher S, Riley A, Mehta K. (2014) Designing breathalyser technology for the developing world: How a single breath can fight the double disease burden. Journal of Medical Engineering and Technology, 38(3), 156–163. doi:10.3109/03091902.2014.890678.

[2] Dragonieri S, Pennazza G, Carratu P, and Resta O. (2017) Electronic Nose Technology in Respiratory Diseases. Lung, 195 (2):157–165. doi:10.1007/s00408-017-9987-3.

[3] Thriumani R, Zakaria A, Hashim YZH, Jeffree AI, Helmy KM, Kamarudin LM, Omar MI, Shakaff AYM, Adom AH, Persaud KC. (2018) A study on volatile organic compounds emitted by in-vitro lung cancer cultured cells using gas sensor array and SPME-GCMS. BMC Cancer, 18:362 doi:10.1186/s12885-018-4235-7.

[4] Chen C, Lin W, Yang H. (2020) Diagnosis of ventilator-associated pneumonia using electronic nose sensor array signals: solutions to improve the application of machine learning in respiratory research. Respiratory Research, 21:45. doi:10.1186/s12931-020-1285-6.

[5] Hu W, Wan L, Jian Y, Ren C, Jin K, Su X, Bai X, Haick H, Yao M, Wu W. (2018) Electronic Noses: From Advanced Materials to Sensors Aided with Data Processing. Advanced Materials Technologies, 4(2),1-38. doi: 10.1002/admt.201800488.

[6]   Yan J, Guo X, Duan S, Jia P, Wang L, Peng C, Zhang S. (2015) Electronic Nose Feature Extraction Methods: A Review. Sensors, 15, 27804-27831. doi:10.3390/s151127804 27804–27831.

[7]   Zulkhairi MA, Mustafah YM, Abidin ZZ, Zaki HFM, Rahman HA. (2019) Car Detection Using Cascade Classifier on Embedded Platform. 7th International Conference on Mechatronics Engineering (ICOM), Putrajaya, Malaysia, pp. 1-3, doi: 10.1109/ICOM47790.2019.8952064.

[8]   Ansari AQ, Khusro A, Ansari MR. (2016) Performance evaluation of classifier techniques to discriminate odors with an E-Nose. 12th IEEE International Conference Electronics, Energy, Environment, Communication, Computer, Control: (E3-C3), INDICON, 2325-9418. doi:10.1109/INDICON.2015.7443838.

[9]   Yi Z, Li C. (2019) Anti-Drift in Electronic Nose via Dimensionality Reduction: A Discriminative Subspace Projection Projection Approach. IEEE Access, 7, 170087–170095. doi:10.1109/ACCESS.2019.2955712.

[10] Vergara A, Vembu S, Ayhan T, Ryan MA, Homer ML, Huerta R. (2012) Chemical gas sensor drift compensation using classifier ensembles. Sensors and Actuators, B: Chemical, 166–167, 320–329. doi: 10.1016/j.snb.2012.01.074.

[11] Liu B, Huang Y, Kam KW, Cheung WF, Zhao N, Zheng B. (2019) Functionalized graphene-based chemiresistive electronic nose for discrimination of disease-related volatile organic compounds. Biosensors and Bioelectronics: X, 1, 100016. doi: 10.1016/j.biosx.2019.100016.

[12] Gargiulo V, Alfano B, Capua RD, Alfè M, Vorokhta M, Polichetti T, Massera E, Miglietta ML, Schiattarella C, Francia GD. (2018) Graphene-like layers as promising chemiresistive sensing material for detection of alcohols at low concentration. Journal of Applied Physics, 123, 024503. doi:10.1063/1.5000914.

[13] Lu G, Ocola LE, Chen J. (2009) Reduced graphene oxide for room-temperature gas sensors. Nanotechnology, 20, 445502, 1-9. doi:10.1088/0957-4484/20/44/445502.

[14] Lee K, Yoo YK, Chae MS, Hwang KS, Lee J, Kim H, Hur D, Jeong HL. (2019) Highly selective reduced graphene oxide (rGO) sensor based on a peptide aptamer receptor for detecting explosives. Sci Rep, 9, 10297. Scientific Reports, 9. doi:10.1038/s41598-019-45936-z.

[15] Tian W, Liu X, Yu W. (2018) Research progress of gas sensor based on graphene and its derivatives: A review. Applied Sciences (Switzerland), 8(7). doi:10.3390/app8071118.

[16] Lee SP. (2017) Electrodes for Semiconductor Gas Sensors. Sensors, 17, 683; doi:10.3390/s17040683.

[17] Wang C, Yin L, Zhang L, Xiang D, Gao R. (2010) Metal Oxide Gas Sensors: Sensitivity and Influencing Factors. Sensors, 10, 2088-2106. doi:10.3390/s100302088.

[18] Baharuddin AA, Ang BC, Hoong Y, Haseeb ASMA, Wong YC. (2019) Materials Science in Semiconductor Processing Advances in chemiresistive sensors for acetone gas detection. Materials Science in Semiconductor Processing, 103, 104616. doi: 10.1016/j.mssp.2019.104616.

[19] Amiri V, Roshan H, Mirzaei A, Neri G, Ayesh AI. (2020) Nanostructured Metal Oxide-Based Acetone Gas Sensors: A Review. Sensors, 20, 3096. doi:10.3390/s20113096.

[20] Norizan MN, Zulaikha S, Demon N, Halim NA. (2021) The frontiers of functionalized graphene-based nanocomposites as chemical sensors. Nanotechnology Reviews, 10: 330-369. doi:10.1515/ntrev-2021-0030.

[21] James F, Fiorido T, Bendahan M, Aguir K. (2017) Comparison between MOX sensors for low VOCs concentrations with interfering gases ALLSENSORS, pp.39-40.

[22] Phillips CO, Syed Y, Parthaláin NM, Zwiggelaar R, Claypole TC, Lewis KE. (2012) Machine learning methods on exhaled volatile organic compounds for distinguishing COPD patients from healthy controls. Journal of Breath Research, 6(3). doi: 10.1088/1752-7155/6/3/036003.

[23] Yan J, Tian F, He Q, Shen, Y. (2012) Feature Extraction from Sensor Data for Detection of found Pathogen Based on Electronic Nose. Sensors and Materials, 24(2), 57–73.

[24] Feng S, Farha F, Li Q, Wan Y, Xu Y, Zhang T, Ning H. (2019) Review on smart gas sensing technology. Sensors (Switzerland), 19(17), 1–22. doi:10.3390/s19173760.

[25] Hashoul D, Haick H. (2019) Sensors for detecting pulmonary diseases from exhaled breath. European Respiratory Review, 28(152). doi:10.1183/16000617.0011-2019.

[26] Xu Y, Zhao X, Chen Y, Yang Z. (2019) Research on a mixed gas classification algorithm based on extreme random trees. Applied Sciences (Switzerland), 9(9). doi:10.3390/app9091728.

[27] Thorson J, Collier-Oxandale A, Hannigan M. (2019) Using A Low-Cost Sensor Array and Machine Mixtures and Identify Likely Sources. Sensors, 19, 3723. doi:10.3390/s19173723.

[28] Tombel NSM, Badaruddin SAM, Yakin FSM, Zaki HFM, Syono MI (2021) Detection of low PPM of volatile organic compounds using nanomaterial functionalized reduced graphene oxide sensor. AIP Conference Proceedings, 2368 vol. 020004. doi.org/10.1063/5.0057775.

[29] Smolinska A. (2014) Current breathomics - A review on data pre-processing techniques and machine learning in metabolomics breath analysis. Journal of Breath Research, 8, 027105: 22. doi:10.1088/1752-7155/8/2/027105.

[30] Das S, Jayaraman V. (2014) Progress in Materials Science SnO2: A comprehensive review in structures and gas sensors. Progress in Materials Science, 66, pp. 112-255. doi: 10.1016/j.pmatsci.2014.06.003.

[31] Maity A, Raychaudhuri AK, Ghosh B. (2019) High sensitivity NH3 gas sensor with electrical readout made on paper with prevoskite halide as sensor material. Scientific Report, 9, 7777. doi: 10.1038/s41598-019-43961-6.

[32] Webb AR (2002) Statistical Pattern Recognition Statistical Pattern Recognition Second Edition. John Wiley & Sons,Ltd, vol. 9.

[33] Bastuck M. (2019) Improving the Performance of Gas Sensor Systems with Advanced Data Evaluation, Operation and Calibration Methods. Linköping University Electronic Press, 298.

[34] Kong C, Zhao S, Weng X, Liu C, Guan R, Chang Z. (2019) Weighted Summation: Feature Extraction of Farm Pigsty Data for Electronic Nose. IEEE Access, vol. 7, pp. 96732-96742. doi:10.1109/access.2019.2929526.

[35] Allwright S. (2022) How to interpret AUC score. Retrieved from https://stephenallwright.com/interpret-auc-score/.

[36] Nyssa SSC, XueVZ, Justin TN, V. R. Saran KC, Raia CF, Michael JL, Thomas LW, Mike F, Gregory JS, Philippe B, Christopher JH, Steven JK. (2022) Machine Learning-Based Rapid Detection of Volatile Organic Compounds in a Graphene Electronic Nose. ACS Nano 2022 16 (11), 19567-19583. doi: 10.1021/acsnano.2c10240.