ISSN: 2302-9285, DOI: 10.11591/eei.v13i1.5577

Pre-processing technique of Aquilaria species from Malaysia for four different qualities

Siti Mariatul Hazwa Mohd Huzir¹, Anis Hazirah 'Izzati Hasnu Al-Hadi², Amir Hussairi Zaidi², Nurlaila Ismail², Zakiah Mohd Yusoff¹, Mohamad Hushni Haron³, Ali Abd Almisreb⁴, Mohd Nasir Taib^{2,5}

¹Electrical Engineering Studies, College of Engineering, Universiti Teknologi MARA, Shah Alam, Malaysia
²School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Shah Alam, Malaysia
³School of Science and Technology, Wawasan Open University, Georgetown, Malaysia
⁴Faculty of Computer Science and Engineering, International University of Sarajevo, Sarajevo, Bosnia and Herzegovina
⁵Malaysia Institute of Transport (MITRANS), Universiti Teknologi MARA, Shah Alam, Malaysia

Article Info

Article history:

Received Dec 21, 2022 Revised Jul 13, 2023 Accepted Aug 2, 2023

Keywords:

Agarwood oil Aquilaria species Boxplot analysis Four grades Grading classification

ABSTRACT

The paper interprets data distribution by using boxplot pre-processing in classify the quality of Agarwood oil for eleven chemical substances into four different qualities. The varieties usage of Agarwood oil makes it considered as an expensive and valuable product on the essential oil market. Perfumes, fragrances, incense, aromatherapy, and traditional medicine are the most popular Agarwood oil applications. However, the classification of Agarwood oil grades does not yet have standard grading method. This because it has been graded manually into different qualities by using human sensory evaluation. Boxplot analysis involving eleven chemical subtances that will be focusing in this study by concerned the quality for low, medium low, medium high and high. γ-eudesmol, ar-curcumene, β-dihydro agarofuran, Υcadinene, α-agarofuran, allo aromadendrene epoxide, valerianol, α-guaiene, 10-epi-γ-eudesmol, β-agarofuran, and dihydrocollumellarin compounds are the selected significant compounds that represent the input for boxplot. Agarwood oil consist 660 data samples from low, medium low, medium high, and high quality. The result in this study showed that the four selected significant compounds (γ-eudesmol, 10-epi-γ-eudesmol, β-agarofuran, and dihydrocollumellarin) are important as a marker for Agarwood oil quality classification. The identification of chemical substances on high quality done as reference for future research studies.

This is an open access article under the CC BY-SA license.



152

Corresponding Author:

Zakiah Mohd Yusoff Electrical Engineering Studies, College of Engineering, Universiti Teknologi MARA Johor Branch, Pasir Gudang Campus, Shah Alam, Malaysia Email: zakiah9018@uitm.edu.my

1. INTRODUCTION

Agarwood oils are the essential oil extracted from Agarwood trees from the *Aquilaria* species which belong to the genus of *Thymelaeaceae* family. Resin is a type of volatile chemical substance that permeates the heartwood of Agarwood trees [1]–[4]. This high resin formation resulting for a good quality of Agarwood oils [2], [5]–[7]. Agarwood oil is one of the most useful oils with a major use for everything from perfumery, medical industry, fragrances and also religious ceremonies [8]–[11]. It has been applied in wide areas where contribute to their popularity in essential oil worldwide market including Malaysia.

Technically, the quality of essential oils had been manually evaluated and graded using sensory evaluation based on their physical properties [12]–[14]. According to human perception and experience, an essential oil with the highest quality has a lot of resin, a dark oil color, a powerful odor, and a long-lasting

Journal homepage: http://beei.org

П

scent [4], [5], [10]. However, various people may have different impressions and conclusions regarding the technique. The sensory evaluation method is somewhat inaccurate. There is no guarantee that grading essential oils by human sensory evaluation would ensure its purity or quality. Due to the continuous process when dealing with a large number of samples at once, the human trained grader technique has a significant disadvantage in terms of objectivity and consistency, leading to a labor-intensive and time-consuming procedure [12], [15]–[17].

Numerous approaches have been proposed and used to verify the quality of essential oils using intelligent methods [13]–[15], [18]–[20]. There is a platform where Agarwood oil quality classification can be done solely based on their chemical subtances allowing essential oils to be classified into their respective classes (low, medium low, medium high, and high quality) by using today's data analysis technology. The accurate result can be measured by using statistical work. The highlight of this paper is the boxplot (also called box-and-whisker plots) analysis of Agarwood oils as well as its outlier labelling when dealing with the raw data. The abundance of Agarwood oil chemical compounds will be the input for the boxplot analysis. This introduction is imperative to highlight the objective of this study which is to present data distribution by using boxplot analysis on the Agarwood oils chemical substances.

2. THEORETICAL WORK

An easy way to interpret this research is to use images or visuals that describe the results more precisely. Boxplot analysis is one of the visuals used in joint display [21]–[24]. Visualization using joint display possibly provide structure for comparing many input data between its group [25]. Boxplot has become efficient tool in the industry standard for summarize the observation value, lowest quartile, median, highest quartile, greatest observation value and outliers in one diagram [26]–[28].

Figure 1 demonstrates the elements in boxplot visualization. The red '+' sign symbol represent an outlier which also known as extreme value and located above or below the whisker. The 25th percentile of the lowest data is the lowest quartile (Q1) while 75th percentile compute from the data is the highest quartile (Q3). There is a red line in the middle of the box which indicate the sample median. Minimum and maximum is the range values in the sample data [28]–[30].

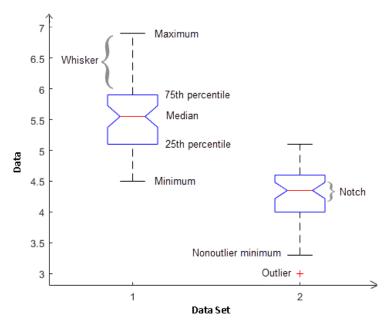


Figure 1. The MATLAB overview of elements in boxplot [29]

Primarily, the shorter the whiskers, the more uniform the distribution of data. The result will be accepted if 50% of the median is in the group [28], [31]. The whiskers are calculated based on the interquartile range (IQR) as mathematically shown in (1)-(3) [31].

$$IQR = Q3 - Q1 \tag{1}$$

154 □ ISSN: 2302-9285

Where Q3 indicates the highest quartile to be minus with lowest quartile (Q1) to get the IQR. Besides, the minimum and maximum range of dataset are (2) and (3):

$$Minimum \ range = Q1 - 1.5 \ (IQR) \tag{2}$$

$$Maximum \, range = Q3 + 1.5 \, (IQR) \tag{3}$$

3. METHOD

In this section, it is explained the research chronological and at the same time is given the comprehensive technique used for data acquisition. Method can be presented in the form of Pseudocode, tables and others that make the reader understand easily. The research method can be made in several sub-sections.

3.1. Sample acquisition

The Agarwood oil samples were prepared by Forest Research Institute Malaysia (FRIM) and Bioaromatic Research Centre of Excellence (BARCE), Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA). The targeted species are focused to only Aquilaria Malaccensis species but with varies oil volume, origin and age. FRIM and BARCE use gas chromatography-mass spectrometry (GC-MS) apparatus to obtain their chemical substances [9]. There are 660 samples starting from low, medium low, medium high and high different qualities of Agarwood oils. The eleven important compounds as tabulated in Table 1 have been used as input and the grades of Agarwood oil has been used as an output to the classification system.

Table 1. Data samples based on quality classification

11 chemical compounds	Grades	Value	Percentage (%)
r-Eudesmol, ar-curcumene, β-dihydro agarofuran, Υ-	Low	209	31.91
cadinene, α-agarofuran, allo aromadendrene epoxide,	Medium low	90	13.74
valerianol, α-guaiene, 10-epi-γ-eudesmol, β-agarofuran,	Medium high	30	4.58
and dihydrocollumellarin	High	326	49.77

3.2. Pre-processing technique

The first step to create boxplot is sorting the data samples into four groups which is low, medium low, medium high, and high quality. There are two columns involve where the x-axis listing the name of eleven important compounds of Agarwood oil (independent variable) and the y-axis is the abundances of important compound in percentage (dependent variable). The boxplot's performance was then evaluated. The flowchart of the experimental analysis in Figure 2 was employed to implement the distribution analysis on Agarwood oil sample. This paper focus on boxplot analysis of four oil qualities.

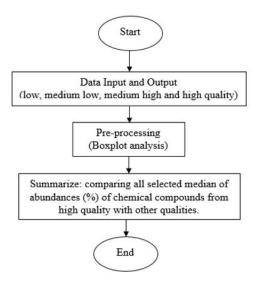


Figure 2. Flowchart of experimental analysis

4. RESULTS AND DISCUSSION

This section will summarize the success of boxplot analysis that support the research objective. The principal component analysis (PCA) filtered 660 data samples and 11 chemical substances to determine the important compounds. These 11 chemical substances became the inputs to boxplot. All samples are placed in the column using MATLAB software.

Figure 3 describe the distribution of chemical substances based on high grade using boxplot graphical. The whiskers for 11= ϕ -cadinene are very small between each other but contain some outliers that locate very far from the central tendency of each feature. The IQR for 1=10-epi- ϕ -eudesmol, 2= α -agarofuran, 3= ϕ -eudesmol, 4= β -agarofuran, 5=dihydrocollumellarin, 6=ar-curcumene, 7=valerianol, 8= β -dihydroagarofuran, 9= α -guaiene, and 10=allo aromadendrene epoxide is larger as shown by the size of box shape. The data distributed uniformly. Generally, the boxplot in Figure 3 can be conclude that only 11= ϕ -cadinene have a very small and indistinguishable IQR while the others chemical substances have large IQR.

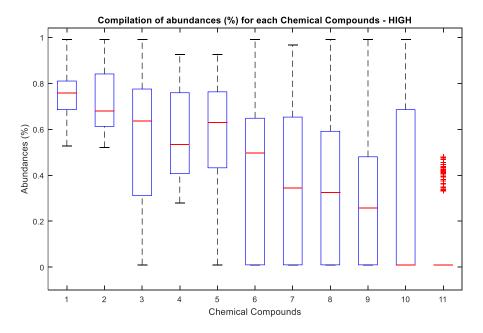


Figure 3. The boxplot of Agarwood oil chemical substances for high grade

The descriptive statistic could give more informative summaries over hundreds or thousands of "raw numbers" from previous data collected. The output of data will be summarize based on the mean satisfaction data, minimum, and maximum sample and also its standard deviation as shown in Table 2. As can see in the Table 2, the average percentage values for top 7 chemical substances are around 0.01 to 0.99.

Table 2. Distribution abundance chemical substances for high grade

7 chemical compounds	Statistic	Abundance percentage (%)
10-epi-γ-eudesmol	Min	0.5269
	Max	0.9900
α-agarofuran	Min	0.5198
	Max	0.9900
γ-Eudesmol	Min	0.0100
	Max	0.9900
β-agarofuran	Min	0.2800
	Max	0.9250
dihydrocollumellarin	Min	0.0100
	Max	0.9250
ar-curcumene	Min	0.0100
	Max	0.9900
valerianol	Min	0.0100
	Max	0.9668

156 □ ISSN: 2302-9285

The comparison between four qualities which is high, medium high, medium low and low analysed as shown in Figure 4. The highlight of this line graph is 10-epi- φ -eudesmol and φ -eudesmol discovered to be important chemical substances with median range between 0.4% to 0.8%. It was declared to be important chemical substances since all of the four grades has its own median value compared to others. Then, the highest median which is 50% of IQR range is belong to β -agarofuran for high grade with 0.99%.

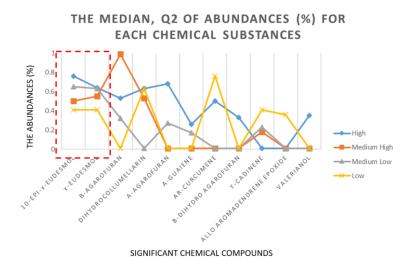


Figure 4. The median value of the Agarwood oil chemical substances for Aqualaria species

From Figure 5, referring to the further analysis, it can be seen that the range for highest quartile (Q1) of 10-epi- ϕ -eudesmol and ϕ -eudesmol is between 0.4–0.81. For 10-epi- ϕ -eudesmol, the highest Q1 is high grade with 0.8102% then followed by medium low, medium high and high grade with 0.7578%, 0.5320%, and 0.4961%, respectively. Next, chemical substance of ϕ -eudesmol, it can be seen that the highest Q1 is medium low grade followed by high, low then lastly medium high grade with value of 0.7764%, 0.7755%, 0.6549%, and 0.5833%, respectively. Lastly, the highest Q1 which is 75% compute from the data is belong to β -agarofuran for high grade with 0.99% value. These two significant compounds (10-epi- ϕ -eudesmol and ϕ -eudesmol) confirmed to be the most important out of others nine compounds.

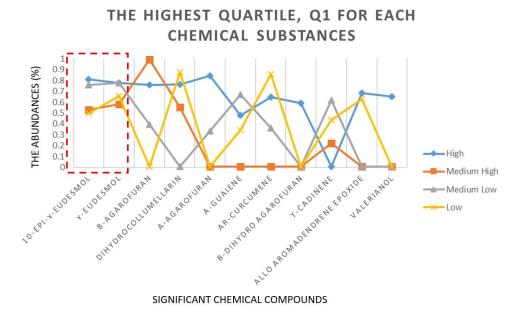


Figure 5. The highest quartile value of the Agarwood oil chemical substances

5. **CONCLUSION**

The research work in this paper has successfully achieved the objective by analysing the Agarwood oil of Aquilaria species boxplot between four grades (low, medium low, medium high, and high). Boxplot method was chosen because it is able to show the shape of distribution, data variability, and its significant value. The boxplot method with its median, the upper and lower quartiles of the range produce accurate differentiations between type data of samples of the Agarwood essential oil four grades. Moreover, the visualization graph is the most suitable and excellent technique to describe samples quality classification. The input is the abundance of eleven significant chemical compounds which are γ-eudesmol, ar-curcumene, β-dihydro agarofuran, Y-cadinene, α-agarofuran, allo aromadendrene epoxide, valerianol, α-guaiene, 10-epi-γ-eudesmol, β-agarofuran and dihydrocollumellarin, and the output is the grades of Agarwood oil. Overall, it is can be summarized that the boxplot and graph give the results of 10-epi-x-eudesmol and γ-eudesmol as important chemical substances for future analysis with suggestion into five and six grades. Chemical compound of β -agarofuran also recommended to be considered in future work since it gives high quartiles values for medium high quality.

ACKNOWLEDGEMENTS

The authors would like to express the gratitude to School of Electrical Engineering, UiTM Shah Alam, Selangor and excellent fund from Universiti Teknologi MARA (UiTM), Cawangan Johor for providing the financial support named GERAN BESTARI FASA 1/2023, 600-TNCPI 5/3/DDN (01) (001/2023) throughout this research.

REFERENCES

- Y. H. Mohammad et al., "Agarwood formation in Aquilaria beccariana and Aquilaria microcarpa in response to inoculation of newly isolated fungi from Brunei Darussalam," Biodiversitas Journal of Biological Diversity, vol. 22, no. 10, Sep. 2021, doi: 10.13057/biodiv/d221002.
- M. Mohamad et al., "kNN: Classification of Agarwood Types in Oil and Wooden Using E-nose," in Proceedings of the 6th International Conference on Electrical, Control and Computer Engineering, 2022, pp. 575-586.
- M. Sahlan et al., "Encapsulation of Agarwood Essential Oil with Maltodextrin and Gum Arabic," International Journal of Technology, vol. 10, no. 8, p. 1541, Dec. 2019, doi: 10.14716/ijtech.v10i8.3485.
- P. S. Naziz, R. Das, and S. Sen, "The Scent of Stress: Evidence From the Unique Fragrance of Agarwood," Frontiers in Plant Science, vol. 10, Jul. 2019, doi: 10.3389/fpls.2019.00840.
- M. N. Azah, S. S. Husni, J. Mailina, L. Sahrim, J. A. Majid, and Z. M. Faridz, "Classification of agarwood (gaharu) by resin content," Journal of Tropical Forest Science, pp. 213-219, 2013.
- S. R. Adhikari, K. Pokhrel, and S. D. Baral, "Economic Value of Agarwood and Its Prospects of Cultivation," International Journal of Applied Sciences and Biotechnology, vol. 9, no. 1, pp. 23-31, Mar. 2021, doi: 10.3126/ijasbt.v9i1.35984.
- W. Hou et al., "Chi-Nan agarwood germplasms constitute a new chemotype of Aquilaria sinensis (Lour.) Spreng," Industrial
- Crops and Products, vol. 187, p. 115494, Nov. 2022, doi: 10.1016/j.indcrop.2022.115494.

 M. A. H. Ador, F. Farabi, R. Ahmed, R. Khatun, and M. M. U. Haque, "Agar (Aquilaria agallocha Roxb.) based small-scale enterprises in Bangladesh: Management, production, marketing and role in socio-economic development," Trees, Forests and People, vol. 6, p. 100141, Dec. 2021, doi: 10.1016/j.tfp.2021.100141.
- M. H. Haron, M. N. Taib, N. Ismail, N. A. M. Ali, and S. N. Tajuddin, "Statistical Analysis of Agarwood Oil Compounds based on GC-MS Data," in 2018 9th IEEE Control and System Graduate Research Colloquium (ICSGRC), Aug. 2018, pp. 27-30, doi: 10.1109/ICSGRC.2018.8657571.
- [10] R. Kalra and N. Kaushik, "A review of chemistry, quality and analysis of infected agarwood tree (Aquilaria sp.)," Phytochemistry Reviews, vol. 16, no. 5, pp. 1045–1079, Oct. 2017, doi: 10.1007/s11101-017-9518-0.
- [11] L. Peeraphong, "Medicinal uses of agarwood," Chinese Medicine and Culture, vol. 4, no. 4, pp. 260-265, 2021, doi: 10.4103/CMAC.CMAC_43_21.
- [12] N. Ismail, N. A. M. Ali, M. Jamil, M. H. F. Rahiman, S. N. Tajuddin, and M. N. Taib, "A review study of agarwood oil and its quality analysis," Jurnal Teknologi, vol. 68, no. 1, pp. 37-42, 2014.
- [13] N. S. J. Jailani, Z. Muhammad, M. H. F. Rahiman, and M. N. Taib, "Analysis of Kaffir Lime Oil Chemical Compounds by Gas Chromatography-Mass Spectrometry (GC-MS) and Z-Score Technique," in 2021 IEEE International Conference on Automatic Control & Intelligent Systems (I2CACIS), Jun. 2021, pp. 110-113, doi: 10.1109/I2CACIS52118.2021.9495909.
- [14] S. Lias, N. A. M. Ali, M. Jamil, M. S. Y. Tolmanan, and M. A. Misman, "A Study on the Application of Electronic Nose Coupled with DFA and Statistical Analysis for Evaluating the Relationship between Sample Volumes versus Sensor Intensity of Agarwood Essential Oils Blending Ratio," MATEC Web of Conferences, vol. 201, p. 02008, Sep. 10.1051/matecconf/201820102008.
- [15] N. Z. Mahabob, Z. M. Yusoff, N. Ismail, and M. N. Taib, "Preliminary Study on Classification of Cymbopogon Nardus Essential Oil using Support Vector Machine (SVM)," in 2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC), Aug. 2020, pp. 63-67, doi: 10.1109/ICSGRC49013.2020.9232661.
- [16] M. S. Ling and K. Phonetip, "Gender and employment in the Burapha plywood factory, Hin Heup District, Vientiane Province," Lao PDR, 2021.
- [17] A. N. M. Ramli, S. Yusof, P. Bhuyar, A. W. Aminan, and S. N. Tajuddin, "Fungi mediated agarwood (A. malaccensis) production and their pharmaceutical applications: A systematic review," International Journal of Plant Based Pharmaceuticals, vol. 2, no. 2, pp. 261–270, 2022, doi: 10.29228/ijpbp.8.
- M. E. M. Samad, N. Ismail, M. H. F. Rahiman, M. N. Taib, N. A. M. Ali, and S. N. Tajuddin, "Analysis of distance metric variations in KNN for agarwood oil compounds differentiation," in 2017 IEEE Conference on Systems, Process and Control (ICSPC), Dec. 2017, pp. 151-156, doi: 10.1109/SPC.2017.8313038.

158 □ ISSN: 2302-9285

[19] N. C. Radzi and F. A. Kasim, "Effect of Microwave Pretreatment on Gaharu Essential Oil Using Hydrodistillation Method," Indonesian Journal of Chemistry, vol. 20, no. 4, p. 960, Jun. 2020, doi: 10.22146/ijc.43191.

- [20] L. Wyn and N. Anak, "Wood for the trees: a review of the agarwood (gaharu) trade in," Kuala Lumpur: TRAFFIC Southeast Asia, 2010.
- [21] T. C. Guetterman, S. Fàbregues, and R. Sakakibara, "Visuals in joint displays to represent integration in mixed methods research: A methodological review," *Methods in Psychology*, vol. 5, p. 100080, Dec. 2021, doi: 10.1016/j.metip.2021.100080.
- [22] B. C. Kaufmann, S. E. J. Knobel, T. Nef, R. M. Müri, D. Cazzoli, and T. Nyffeler, "Visual Exploration Area in Neglect: A New Analysis Method for Video-Oculography Data Based on Foveal Vision," *Frontiers in Neuroscience*, vol. 13, Jan. 2020, doi: 10.3389/fnins.2019.01412.
- [23] J. E. McKenzie and S. E. Brennan, "Synthesizing and presenting findings using other methods," in Cochrane Handbook for Systematic Reviews of Interventions, Wiley, 2019, pp. 321–347.
- [24] A. Ribés, J. Pouderoux, and B. Iooss, "A Visual Sensitivity Analysis for Parameter-Augmented Ensembles of Curves," *Journal of Verification, Validation and Uncertainty Quantification*, vol. 4, no. 4, Dec. 2019, doi: 10.1115/1.4046020.
- [25] D. Buttarazzi, G. Pandolfo, and G. C. Porzio, "A boxplot for circular data," *Biometrics*, vol. 74, no. 4, pp. 1492–1501, Dec. 2018, doi: 10.1111/biom.12889.
- [26] M. Lotfollahi, F. A. Wolf, and F. J. Theis, "scGen predicts single-cell perturbation responses," *Nature Methods*, vol. 16, no. 8, pp. 715–721, Aug. 2019, doi: 10.1038/s41592-019-0494-8.
- [27] W. Sun, P. Zhang, Z. Wang, and D. Li, "Prediction of Cardiovascular Diseases based on Machine Learning," ASP Transactions on Internet of Things, vol. 1, no. 1, pp. 30–35, May 2021, doi: 10.52810/TIOT.2021.100035.
- [28] Reonaldo, M. D. Lauro, and D. E. Herwindiati, "The Simple Boxplot Method for an Effective Prediction," *IOP Conference Series: Materials Science and Engineering*, vol. 1007, no. 1, p. 012141, Dec. 2020, doi: 10.1088/1757-899X/1007/1/012141.
- [29] M. Trojanová, S. Hrehová, L. Knapčíková, and J. Husár, "MATLAB A tool for learning students in the field of analysis of data," in ICERI2021 Proceedings, Nov. 2021, pp. 3524–3532, doi: 10.21125/iceri.2021.0857.
- [30] B. Xu *et al.*, "GABA signalling modulates stomatal opening to enhance plant water use efficiency and drought resilience," *Nature Communications*, vol. 12, no. 1, p. 1952, Mar. 2021, doi: 10.1038/s41467-021-21694-3.
- [31] R. Dawson, "How Significant is a Boxplot Outlier?," Journal of Statistics Education, vol. 19, no. 2, Jul. 2011, doi: 10.1080/10691898.2011.11889610.

BIOGRAPHIES OF AUTHORS





Anis Hazirah 'Izzati Hasnu Al-Hadi D was born in Malaysia, in July 1998. She received her B. Eng (Hons) of Electronics Engineering from Universiti Teknologi MARA (UiTM) Shah Alam, Selangor, Malaysia. She is currently an ASIC Design Engineer at Microchip Operations (Malaysia) Sdn Bhd and at the same time, she is a full-time postgraduate student at School of Electrical Engineering, Universiti Teknologi MARA (UiTM) Shah Alam, Selangor, Malaysia. She can be contacted at email: anshzrh@gmail.com.







Zakiah Mohd Yusoff si sa senior lecturer who is currently working at UiTM Pasir Gudang. She received the B. Eng in Electrical Engineering and Ph.D in Electrical Engineering from UiTM Shah Alam, in 2009 and 2014, respectively. In May 2014, she joined UiTM Pasir Gudang as a teaching staff. Her major interests include process control, system identification, and essential oil extraction system. She can be contacted at email: zakiah9018@uitm.edu.my.



Mohamad Hushnie Haron is is currently a lecturer at the School of Science and Technology, Wawasan Open University. He received a diploma (FH) degree in Mechatronics from University of Applied Sciences Aalen, Germany, and a master's degree and doctorate doctor's degree in Electrical Engineering from Universiti Teknologi MARA Shah Alam, in Selangor, Malaysia. His research interests include signal processing and data analysis, artificial intelligence, embedded system, and internet of things (IoT). He can be contacted at email: mhushnieh@wou.edu.my.



Ali Abd Almisreb () () () is currently an associate professor at the Faculty of Computer Sciences and Engineering, Director of Graduate Council and Editor in Chief at International University of Sarajevo. He received a master's degree in Computer Science and Doctorate doctor's degree in Electrical Engineering/Computer Engineering from Universiti Teknologi MARA, Malaysia. His major interests include deep learning, machine learning, computer vision voice recognition and quantum computing. He can be contacted at email: alimes96@yahoo.com.



Mohd Nasir Taib received his Ph.D from UMIST, UK. He is a senior professor at Universiti Teknologi MARA (UiTM). He heads the Advanced Signal Processing Research Group at the School of Electrical Engineering, College of Engineering, UiTM. He has been a very active researcher and over the years had author and/or co-author many papers published in refereed journals and conferences. He can be contacted at email: dr.nasir@uitm.edu.my.