IOP Conference Series: Materials Science and Engineering

PAPER • OPEN ACCESS

Intelligent classification of waxy crude oil odor-profile at different temperature

To cite this article: M F R M Mawardzi *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **469** 012071

IOP Conf. Series: Materials Science and Engineering 469 (2019) 012071 doi:10.1088/1757-899X/469/1/012071

Intelligent classification of waxy crude oil odor-profile at different temperature

M F R M Mawardzi[1], A Japper-Jaafar[1], M S Najib[2], S M Daud[2] and T M Y S T Ya[1]

[1]Department of Mechanical Engineering, Universiti Teknologi PETRONAS, Persiaran UTP, 32610 Seri Iskandar, Perak, Malaysia.
[2]Faculty of Electrical and Electronics Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia.

*Corresponding author: mohd.fakrul.radzi@gmail.com

Abstract. Crude oil is one of the basic needs required for humans to ease their life. The quality of crude oil with the lowest wax content is very important, in order to sustain the transportation and production of crude oil from offshore to onshore. Based on literature from previous studies, the appearance of wax depends on the temperature which is called Wax Appearance Temperature (WAT). Hence, there is a need to propose a new method to classify the waxy crude oil at a different temperature. The main purpose of this paper is to classify Malaysian waxy crude oil odor profile at different temperatures using intelligent classification technique. There are 28,000 data measurement of the waxy crude oil that was taken using an electronic nose (E-nose). The data readings have been normalized and analyzed using a statistical method. Then, the odor profiles were classified using K-Nearest Neighbour. The classification performance shows that the technique was able to classify the Malaysian waxy crude oil odor profile at different temperatures with 100% accuracy.

1. Introduction

Nowadays, humans consume a lot of energy source to simplify daily life. One of the sources of human needs is from petroleum or crude oil. In addition, crude oil can also produce chemicals, plastics, and synthetic materials that are always used by humans. Fuel and gasoline are the by-products of crude oil that have the highest percentage of human daily usage. Due to the ever-increasing demand for crude oil products, the oil and gas industry has also shifted into the deeper oceans to increase the productivity of crude oil depletion.

Among the focus of the crude oil production, the major attention is in the flow assurance field in relation to waxy crude oil. This wax will cause other problems as well as reducing the productivity of crude oil production. High costs are required to solve this problem. In serious circumstances, the transportation pipelines will be blocked by the wax [1–2].

One of the studies on waxy crude oil is a petroleum wax classification of using electronic nose by Wang, Gao, & Wang [3]. The results show that the electronic nose is able to distinguish the quality of petroleum waxes based on their volatile profile. K-nearest neighbour, support vector machine and multilayer perceptron were used in the study [3].

An electronic nose (E-nose) has provided plenty of benefits to a variety of commercial industries, including the agricultural, biomedical, cosmetics, environmental, food, manufacturing, military, pharmaceutical, regulatory, and various scientific research fields [4–11]. The E-nose could be used as a tool to measure the smell of the waxy crude oil. It has the ability to read the odor data and analyze in a short period of time. In this paper, the E-Nose was used to obtain and classify an odor profile of waxy crude oil at different temperatures [9].

There are numerous artificial techniques that can be used to classify waxy crude oil odor profile data such as Principal Component Analysis (PCA) [12–14], Discriminant Factor Analysis (DFA) [15–17], K-Nearest Neighbor (K-NN) [18;20], Artificial Neural Network (ANN) [21–22] and Case-Based Reasoning (CBR) [5–8;10;11;20].

K-NN will be used to analyze data obtained from the E-nose in this study. There are a few stages consisting of k-NN classification which are data preparation, data splitting, data training and parameter optimization using k-variable, distance and rule. The features or patterns of each sample will be presented in 2D graphical illustration [9].

Hence, this research is a necessary study to classify the waxy crude oil at different temperatures. It presents a significant classification technique based on odor-profile using e-nose instrument and k-NN classifier. E-nose was used to measure the odor data while k-NN was used as an intelligent classification technique.

2. Methodology

Figure 1 below shows the overall flow used to study the waxy crude oil odor-profile at different temperatures. It starts with experimental setup sample preparation for the data measurement using the E-nose. After the odor data has been taken, data pre-processing was made with data normalization and mean calculation technique. Then, waxy crude oil odor patterns for each temperature are extracted and classified using an intelligent classification technique. There will be a linear regression analysis in determining the classification accuracy at the end of this paper.

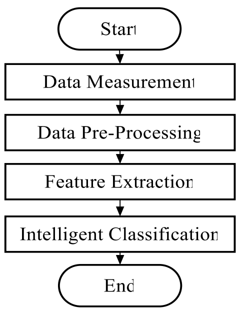

Figure 1. Overall flowchart for waxy crude oil odor-profile classification.

2.1. Experimental Setup and Data Measurement

Figure 2 below shows the sample of waxy crude oil for E-Nose data collection. The waxy crude oil sample used in this study was supplied by PETRONAS. It is from one of the Malay basins named PENARA. The samples were kept in a bottle to avoid any contamination. The volume of the sample needed for odor detection in this study is 10mL.

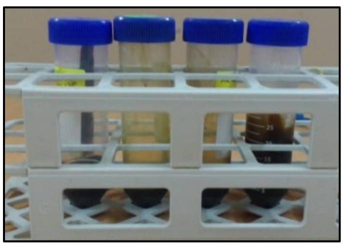

Figure 2. Sample preparation of waxy crude oil for E-nose data measurement.

IOP Conf. Series: Materials Science and Engineering 469 (2019) 012071 doi:10.1088/1757-899X/469/1/012071

Figure 3 below shows the experimental setup for data measurement. There is an internal pump in the E-nose to suck in and control the inlet source of the gas from the heated waxy crude oil sample. Four sensor array arrangement is used to take the odor reading from the heated sample. The sensors were arranged in a parallel position in the E-nose chamber. It was able to detect carbon monoxide (CO), LPG, CO, CH4, natural gas, propane, methane, i-butane, alcohol, hydrogen, and smoke. The data was then sent to a computer via USB cable. Arduino software was used as a user feature interface. The heating system was used to control the temperature since there were seven different temperatures involved. The temperature range for the heater is from 0°C to 300°C.

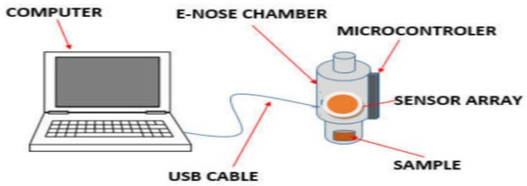

Figure 3. Electronic nose experimental setup.

All the data that has been taken by the E-Nose will be tabulated in Table 1 below. The table shows the example of the 1,000 data (5 times repeated experiment) which have been collected. DM is the data measurement of the waxy crude oil while S_1, S_2, S_3, and S_4 are the sensor array (S_1 for sensor 1, S_2 for sensor 2, S_3 for sensor 3, and S_4 for sensor 4). The subscript number for DM is the number of odors reading row and sensor number. For example, DM_{11} is the first data measure for sensor 1.

Table 1. Data measurement table for waxy crude oil odor-profile.

No. of Data Measured	S_1	S_2	S_3	S_4
1	DM_{11}	DM_{12}	DM_{13}	DM_{14}
2	DM_{21}	DM_{22}	DM_{23}	DM_{24}
3	DM_{31}	DM_{32}	DM_{33}	DM_{34}
.	.	.	.	.
1000	DM_{10001}	DM_{10002}	DM_{10003}	DM_{10004}

2.2. Data Pre-Processing

All of the data collected will be normalized using Equation 1 below in the pre-processing step,

$$R' = \frac{R}{R_{max}}, \tag{1}$$

where R' is the normalized data, R_{max} is the highest odor data at each row and R is the odor data for each sensor. This step will automatically rescale all the data from zero to one (0-1). The normalized data will be tabulated in Table 2. ND is the normalized data. The subscript number for the ND represents the number of odor reading and sensor number. For example, ND_{11} is the first normalized data for sensor 1. The normalized data in Table 2 consists of 4,000 data (1,000 odor reading x 4 sensors) for each temperature used in this experiment. S_1, S_2, S_3, and S_4 represents sensor 1, sensor 2, sensor 3 and sensor 4.

IOP Conf. Series: Materials Science and Engineering 469 (2019) 012071 doi:10.1088/1757-899X/469/1/012071

Table 2. Data normalization for waxy crude oil odor-profile.

No. of Data Measured	S_1	S_2	S_3	S_4
1	ND_{11}	ND_{12}	ND_{13}	ND_{14}
2	ND_{21}	ND_{22}	ND_{23}	ND_{24}
3	ND_{31}	ND_{32}	ND_{33}	ND_{34}
.	.	.	.	.
1000	ND_{10001}	ND_{10002}	ND_{10003}	ND_{10004}

2.3. Intelligent Classification
All of the data analysis in the previous section together with the feature extraction is very important for the classification stage. K-NN has been selected and used in this study. It is widely known as a simple algorithm which stores and group all the cases in view of similarity distance measure. There are several steps consisting of the classifier such as data input, data output, data splitting, data training, data testing and parameter optimization.

3. Results and Discussion

3.1. Raw Data Measurement
Figure 4-5 in page 4 and Figure 6-10 in page 5 shows the graph of data measurement against sensor array for seven different temperatures (20ºC, 30 ºC, 40 ºC, 50 ºC, 60 ºC, 70 ºC and 80 ºC). Y-axis indicates the reading data in resistance value (Ω) while the x-axis indicates the sensor array. Sensor array 1, 2, 3, and 4 are the four sensors in the E-nose. There is an indicated range and difference of odor reading for each sensor at each temperature but still does not cause changes to the odor pattern of the waxy crude oil. The data showed that the waxy crude oil sample will produce a different gas at different temperatures. The noticeable difference is in sensor 4. At low temperature (20ºC-70ºC), the odor reading is higher than sensor 2 and sensor 3. When the temperature reaches 80ºC, the odor reading for sensor 4 is lower than sensor 2 and 3. This concludes that at 80ºC the propane, methane, i-butane, alcohol, hydrogen, and smoke is reduced so that the reading is low.

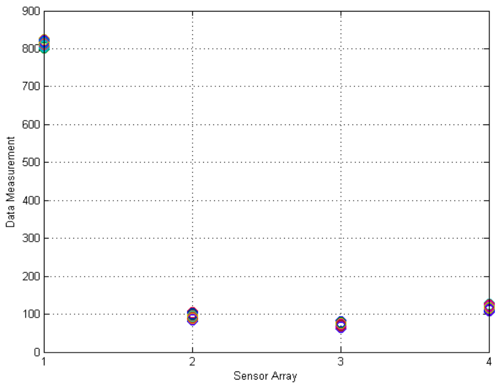

Figure 4. Raw data of waxy crude oil at 20ºC.

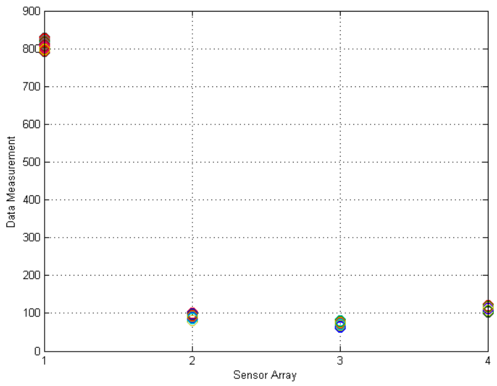

Figure 5. Raw data of waxy crude oil at 30ºC.

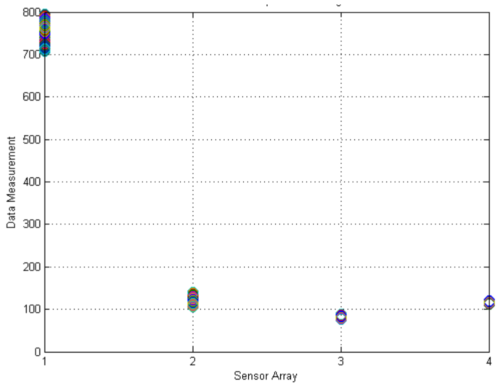

Figure 6. Raw data of waxy crude oil at 40^{o}C.

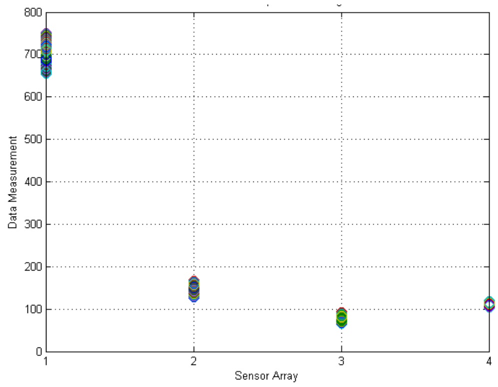

Figure 7. Raw data of waxy crude oil at 50^{o}C.

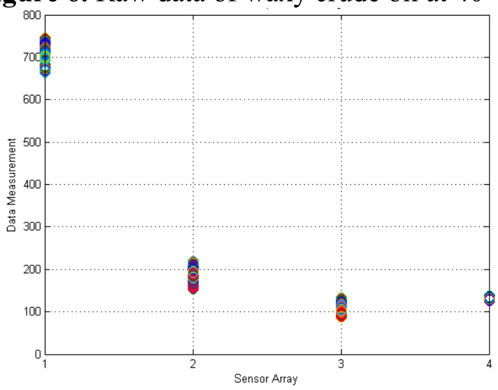

Figure 8. Raw data of waxy crude oil at 60^{o}C.

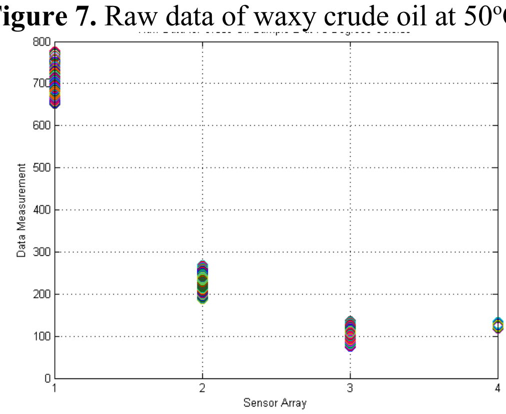

Figure 9. Raw data of waxy crude oil at 70^{o}C.

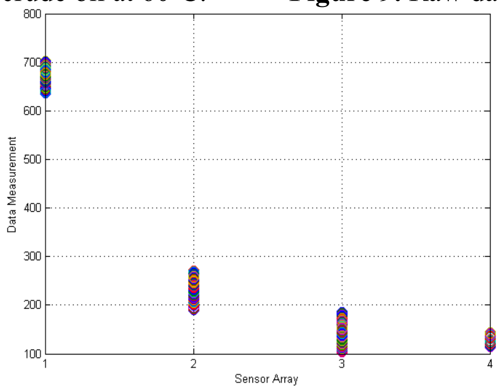

Figure 10. Raw data of waxy crude oil at 80^{o}C.

3.2. Data Pre-Processing

By dividing every data reading in every row with the highest value from its own row, all 28,000 measured data were normalized. It was represented in seven different temperatures. Figure 11-17 on page 6 shows the normalized data for all the data measurement. The collected data had been normalized into 0 to 1 interval. The graphs show that there are differences in odor pattern at a different temperature. Y-axis indicates the normalized value while the x-axis indicates the sensor array. The highest sensor reading is at sensor 1 while the lowest is at sensor 3. Based on the pattern recognition, all temperatures have shown different odor profile with each other. Even though the patterns are almost similar, there was still a distinctive difference between each other and can be calculated and also used in the classification stage. The changes in data pre-processing also showed that all the sensors in E-nose have the ability to detect and read the odor released.

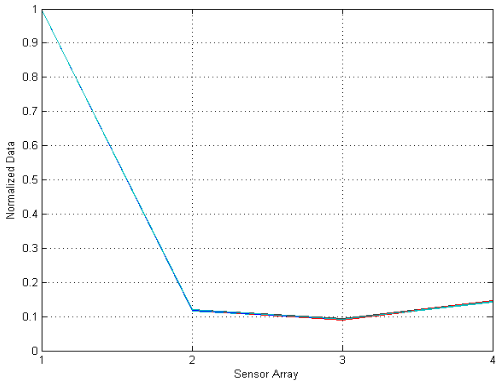

Figure 11. Odor pattern at 20ºC.

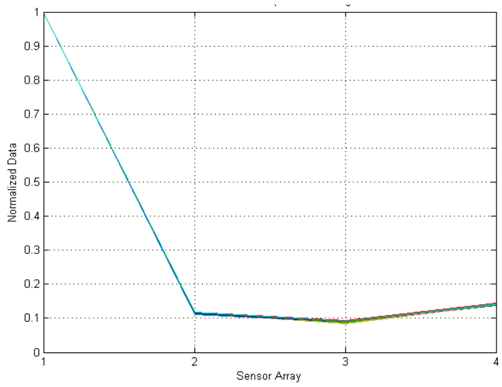

Figure 12. Odor pattern at 30ºC.

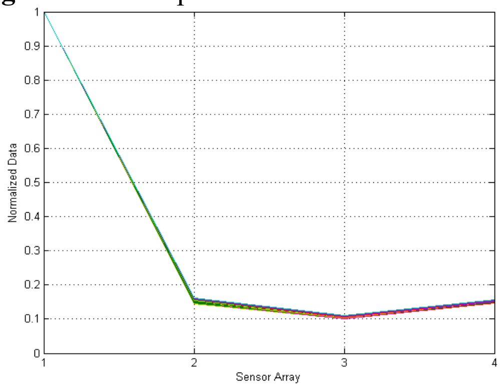

Figure 13. Odor pattern at 40ºC.

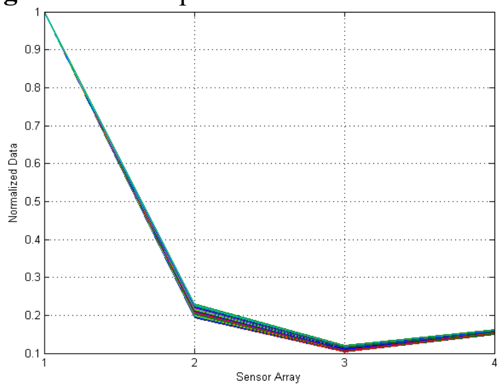

Figure 14. Odor pattern at 50ºC.

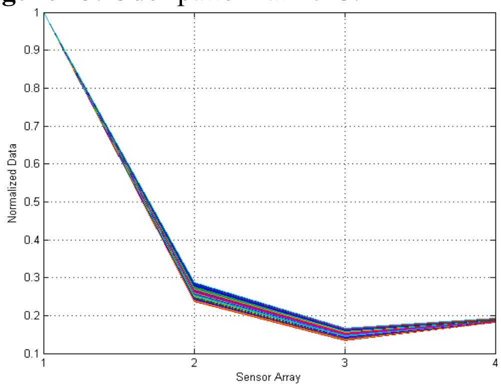

Figure 15. Odor pattern at 60ºC.

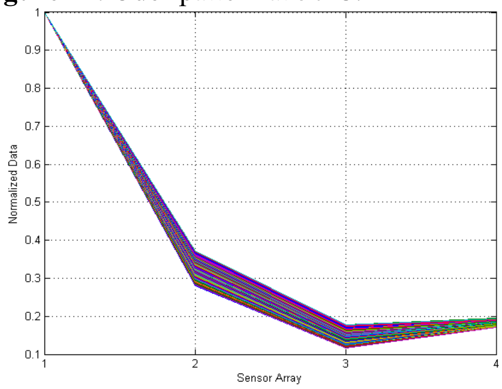

Figure 16. Odor pattern at 70ºC.

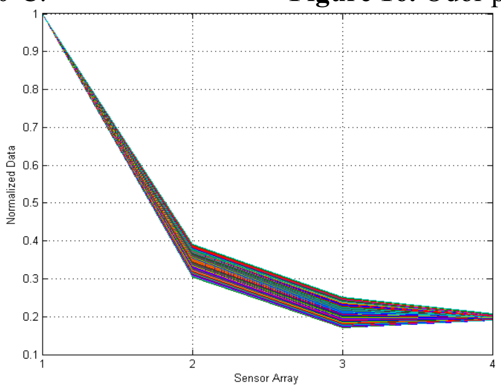

Figure 17. Odor pattern at 80ºC.

3.3. Boxplot Analysis

Figure 18-23 in page 6 and Figure 24 in page 7 shows the boxplot of waxy crude oil sample for 20°C, 30 °C, 40 °C, 50 °C, 60 °C, 70 °C and 80 °C respectively. Y-axis represents the normalized data while the x-axis is the sensor array in the E-Nose. Each boxplot consists of a median, first quartile, third quartile, a maximum and minimum value of the data. Based from Figure 6 (a-g), sensor 1 shows the best data result because it was the most consistent and has no range of variation. For sensors 2, 3 and 4 the boxplot width is smaller at a lower temperature and became bigger due to the increase in temperature, especially for sensors 2 and 3.

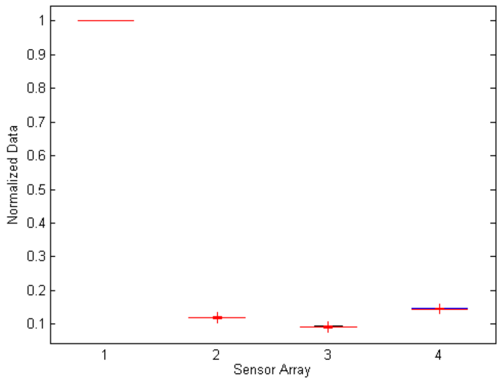

Figure 18. Boxplot for waxy crude oil at 20°C.

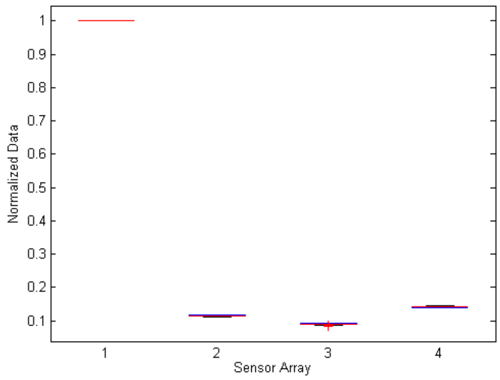

Figure 19. Boxplot for waxy crude oil at 30°C.

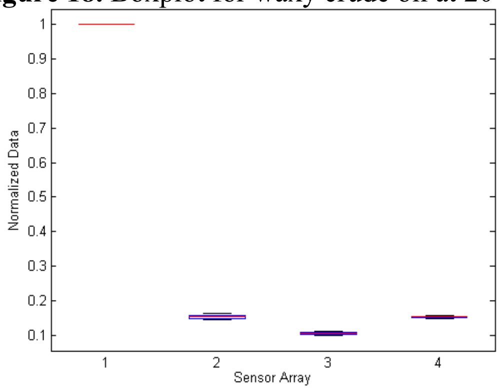

Figure 20. Boxplot for waxy crude oil at 40°C.

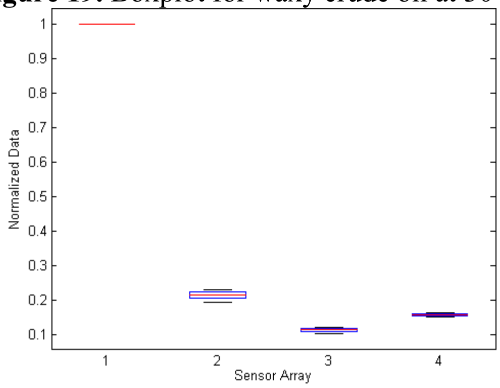

Figure 21. Boxplot for waxy crude oil at 50°C.

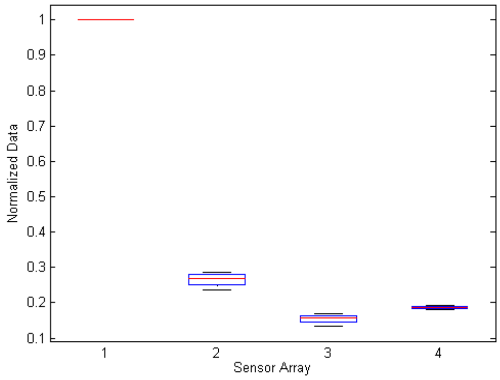

Figure 22. Boxplot for waxy crude oil at 60°C.

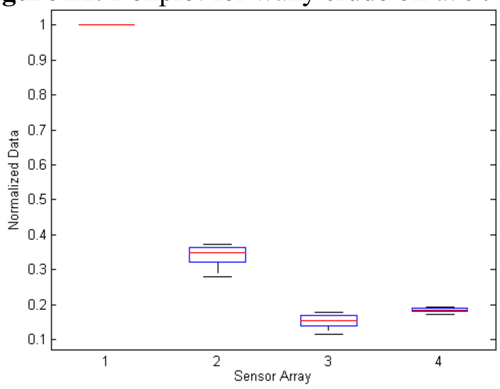

Figure 23. Boxplot for waxy crude oil at 70°C.

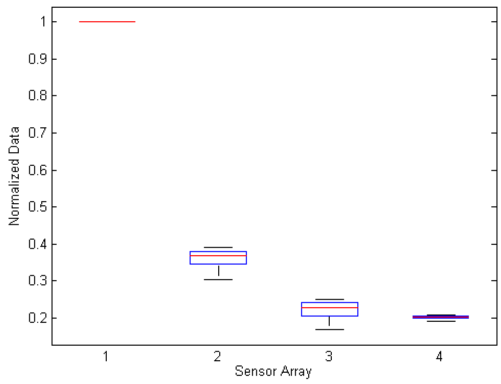

Figure 24. Boxplot for waxy crude oil at 80°C.

3.4. K-Nearest Neighbour (K-NN)

K-NN is a well-known classification technique which produces an excellent classification result. In this study, 9 types of splitting ratio for training and testing data which are 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20 and 90:10 were used in parameter optimization. Table 3 below shows the result of parameter optimization at data splitting 90:10 due to the excellent classification result. The first column shows the distance where every distance has their own algorithm. The distances are Euclidean, Cityblock, Cosine, and Correlation. The second column is a rule used for each distance (Nearest, Random and Consensus) and the last column shows the types of K used (K=1, K=2, and K=3). Table 3 below shows the result of classification at 90:10 ratio. It shows that the K-NN is able to produce 100% classification of waxy crude oil at different temperatures.

Table 3. Parameter optimization table for waxy crude oil classification using K-NN at 90:10 splitting data.

Distance	Rule	Ratio 90:10 Percentage similarity (%)		
		K=1	K=2	K=3
Euclidean	Nearest	100	100	100
	Random	100	100	100
	Consensus	100	100	99.29
Cityblock	Nearest	100	100	100
	Random	100	100	100
	Consensus	100	100	100
Cosine	Nearest	100	100	100
	Random	100	100	100
	Consensus	100	100	99.29
Correlation	Nearest	86.43	86.43	87.14
	Random	86.43	89.29	87.14
	Consensus	86.43	84.29	83.57

3.5. Linear Regression

Linear regression is a technique used to validate the training and testing data. Figure 25 on page 9 shows the regression plot for temperature classification. The training data is in a blue line with '+' sign, while the testing data is in the red line with 'O' sign. It can be seen that the data was divided into 7 classes which are 20°C, 30 °C, 40 °C, 50 °C, 60 °C, 70 °C, and 80 °C respectively. From the figure, it clearly shows that the K-NN model able to classify waxy crude oil at different temperature with zero percentage of error.

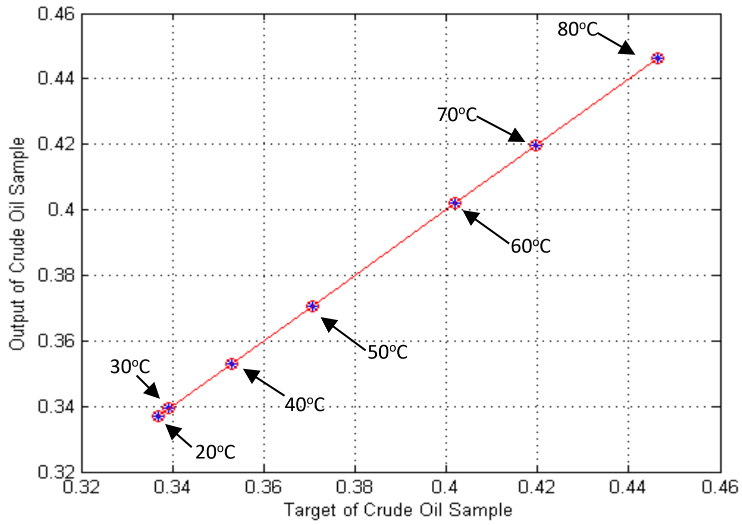

Figure 25. Regression Plot of Temperature for Waxy Crude Oil.

4. Conclusion

This paper has successfully presented the odor classification of waxy crude oil for seven different temperature ranges from 20oC to 80oC. This study has shown that the E-nose is effectively able to collect and measure the odor data of the waxy crude oil. It also shows that the waxy crude oil releases a different odor at a different temperature. The differences of the odor profile were caused by the changes of the waxy crude oil aroma which were influenced by the changes in the properties of the sample. All the odor data was normalized using boxplot statistical analysis and the odor profile of the waxy crude oil at each temperature was uniquely extracted. The intelligent classification technique of the odor profile using k-NN has successfully achieved 100% classification. This study can be extended later to correlate between E-Nose with the waxy crude oil properties.

5. References

[1] T. B. Petrus, "Viscosity Prediction for Waxy Crude Oil - Effects of Physico-Chemical Parameters," Universiti Teknologi PETRONAS, 2017.

[2] N. Ridzuan, "Experimental Investigation and Modelling of Wax Deposition Inhibition in Pipeline Transportation of Crude Oil," Universiti Malaysia Pahang, 2016.

[3] J. Wang, D. Gao, and Z. Wang, "Quality-grade evaluation of petroleum waxes using an electronic nose with a TGS gas sensor array," *Meas. Sci. Technol.*, vol. 26, no. 8, p. 85005, 2015.

[4] M. Falasconi, I. Concina, E. Gobbi, V. Sberveglieri, A. Pulvirenti, and G. Sberveglieri, "Electronic Nose for Microbiological Quality Control of Food Products," *Int. J. Electrochem.*, vol. 2012, pp. 1–12, 2012.

[5] M. F. Zahari, "Intelligent Classification of Ammonia Concentration Based on Odor Profile."

[6] M. S. Najib, N. H. Zamberan, N. Zahed, F. A. Halim, M. F. Zahari, W. M. A. Mamat, and H. Manap, "Fish quality study using odor-profile case-based reasoning (CBR) classification technique," *ARPN J. Eng. Appl. Sci.*, vol. 11, no. 10, pp. 6691–6696, 2016.

[7] N. Zahed, M. S. Najib, N. Fatin, N. Nik, and M. Azhani, "Classification of Honey Odor-Profile Using Case-Based Reasoning Technique (Cbr)," vol. 11, no. 10, pp. 6675–6679, 2016.

[8] M. F. Zahari, T. A. Julius, F. A. Halim, M. S. Najib, K. H. Ghazali, and A. A. M. Azoddein, "Intelligent Classification Hazardous Gas Using Sensors Array," *J. Adv. Inf. Technol.*, vol. 6, no. 4, pp. 233–237, 2015.

[9] M. F. C. J. and M. I. I. N. S. M. Daud, M. S. Najib , N. Zahed, M. F. M. Jusof, *, and Hassim, "Classification of Lubricant Oil Adulteration Level Using Case-Based Reasoning," *J. Fundam. Appl. Sci.*, vol. 9, no. 4S, pp. 256–275, 2017.

[10] M. F. Zahari, M. S. Najib, K. H. Ghazali, F. A. Halim, and A. A. Mohd, "Classification of Ammonia in water for Oil and Gas Industry using Case Based Reasoning (CBR)," *Colloq. Robot. Unmanned Syst. Cybern. 2014 (CRUSC 2014)*, vol. 2014, no. Crusc, pp. 12–16, 2014.
[11] N. Zahed, M. S. Najib, and N. F. N. N. M. Azhani, "Classification of honey odor-profile using case-based reasoning technique (CBR)," *ARPN J. Eng. Appl. Sci.*, vol. 11, no. 10, pp. 6675–6679, 2016.
[12] Z. Haddi, S. Mabrouk, M. Bougrini, K. Tahri, K. Sghaier, H. Barhoumi, N. El Bari, A. Maaref, N. Jaffrezic-Renault, and B. Bouchikhi, "E-Nose and e-Tongue combination for improved recognition of fruit juice samples," *Food Chem.*, vol. 150, pp. 246–253, 2014.
[13] O. S. Papadopoulou, C. C. Tassou, L. Schiavo, G.-J. E. Nychas, and E. Z. Panagou, "Rapid Assessment of Meat Quality by Means of an Electronic Nose and Support Vector Machines," *Procedia Food Sci.*, vol. 1, no. Icef 11, pp. 2003–2006, 2011.
[14] A. Romero-Flores, L. L. McConnell, C. J. Hapeman, M. Ramirez, and A. Torrents, "Evaluation of an electronic nose for odorant and process monitoring of alkaline-stabilized biosolids production," *Chemosphere*, vol. 186, pp. 151–159, 2017.
[15] M. Peris and L. Escuder-Gilabert, "A 21st century technique for food control: Electronic noses," *Anal. Chim. Acta*, vol. 638, no. 1, pp. 1–15, 2009.
[16] S. Kiani, S. Minaei, and M. Ghasemi-Varnamkhasti, "Application of electronic nose systems for assessing quality of medicinal and aromatic plant products: A review," *J. Appl. Res. Med. Aromat. Plants*, vol. 3, no. 1, pp. 1–9, 2016.
[17] A. Sanaeifar, H. ZakiDizaji, A. Jafari, and M. de la Guardia, "Early detection of contamination and defect in foodstuffs by electronic nose: A review," *TrAC - Trends Anal. Chem.*, vol. 97, pp. 257–271, 2017.
[18] S. Güney and A. Atasoy, "Multiclass classification of n-butanol concentrations with k-nearest neighbor algorithm and support vector machine in an electronic nose," *Sensors Actuators, B Chem.*, vol. 166–167, pp. 721–725, 2012.
[19] S. Güney and A. Atasoy, "Multiclass classification of n-butanol concentrations with k-nearest neighbor algorithm and support vector machine in an electronic nose," *Sensors Actuators, B Chem.*, vol. 166–167, pp. 721–725, 2012.
[20] F. A. Halim, "Classification of Ammonia from Water Based on Odor-Profile using K-NN and CBR," Universiti Malaysia Pahang, 2016.
[21] Y. González Martín, M. Concepción, C. Oliveros, J. Luis, P. Pavón, C. G. Pinto, and B. M. Cordero, "Electronic nose based on metal oxide semiconductor sensors and pattern recognition techniques: characterisation of vegetable oils," *Anal. Chim. Acta*, vol. 449, pp. 69–80, 2001.
[22] Q. Chen, J. Zhao, Z. Chen, H. Lin, and D. A. Zhao, "Discrimination of green tea quality using the electronic nose technique and the human panel test, comparison of linear and nonlinear classification tools," *Sensors Actuators, B Chem.*, vol. 159, no. 1, pp. 294–300, 2011.