PAPER • OPEN ACCESS

Classification of waxy crude oil odor-profile using gas sensor array

To cite this article: M F R M Mawardzi *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **469** 012073

View the article online for updates and enhancements.

IOP ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

Classification of waxy crude oil odor-profile using gas sensor array

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. Nowadays, waxy crude oil becomes one of the major concerns in the oil and gas industry. The waxy crude oil affects the production and transportation of the crude oil from offshore to onshore. Differentiation by image visualization in determining the type of waxy crude oil, with or without wax sometimes appear rather similar between each other. Hence, a new method needs to be used to differentiate and classify the waxy crude oil type. An electronic nose (E-Nose) is one of the devices that could detect and measure the odor data of the waxy crude oil type using gas sensor array. This paper aims to classify four types of Malaysian waxy crude oil from different fields at room temperature. There are 16,000 odor data that has been collected by using the E-Nose. Then, the measured data were normalized and analyzed using boxplot analysis. The unique odor-profile for each type of waxy crude oil sample has been extracted and classified using intelligent classification technique. The four types of waxy crude oil have been classified 100% using k-NN intelligent classification technique with zero percentage of error in this paper.

1. Introduction

Numerous researchers have investigated the flow assurance issues in the oil and gas industry. One of the major problems was found in the wax deposition on the pipe wall. On top of that, wax deposition does not only occur in the pipeline but also in the reservoir. The deposited wax will then cause other problems to arise, as well as reducing the productivity of crude oil. A high cost of spending will be needed to rectify this problem. In Malaysia, this issue became one of the main concerns in flow assurance study. In the worst case, the deposited wax will cause the total blockcage of transportation pipelines [1]. Among the examples of Malaysian fields faced with this problem are Penara, Angsi, and Dulang. These fields are located 60m to 70m below sea level where the temperature on the surface is 34°C and the seabed temperature at the depth of 61m is 25°C [2].

The presence or formation of wax from crude oil in seabed due to low temperature in wells has become a critical issue in the oil and gas industry. Formation or presence of wax can occur along the transportation pipeline from offshore to onshore [1]. In literature, wax can occur if the well temperature and pipeline are below the wax appearance temperature (WAT) where long linear n-paraffin chains of waxy crude oil will come into contact with some areas of the pipeline wall. The presence of this wax will also change the rheological behavior of the crude oil, thus reducing production capacity and block [1,3].

The waxy crude oil can be detected by measuring the odor. An E-Nose is a tool that can be used to measure the smell of waxy crude oil in a short period of time [4]. It consists of an array of sensors and has been widely used for the analysis of volatile compounds and odors. It contains artificial intelligence and pattern recognition subsystems. The E-Nose can perceive not only what odor molecules consist of, but also the concentration of molecules in odors [5].

A number of artificial techniques can be used to classify waxy crude oil odor-profile data such as k-Nearest Neighbor (k-NN) [5–7], Artificial Neural Network (ANN) [8,9], Principal Component Analysis (PCA) [12–14], Discriminant Factor Analysis (DFA) [9] and Case-Based Reasoning (CBR) [4, 13–16]. As a robust classifier, k-NN is widely used in pattern recognition. It is a non-parametric method for classification and regression. It consists of a few steps which are data splitting (training and testing) and parameter optimization (distance, rule, and k-variable) [17].

Hence, this research is to classify four types of waxy crude oil using intelligent classification technique. The E-Nose was used as the instrument to measure the odor data for each type of waxy crude oil, while k-NN was used as a classifier.

2. Methodology

Figure 1 below shows the overall flowchart in this study. There are four main steps involved in the classification of waxy crude oil. It starts with the data collection using the E-Nose. After the data collection, the data pre-processing has been done to extract the odor features of each waxy crude oil. The features were then classified using an intelligent classification technique which is k-Nearest Neighbour (k-NN).
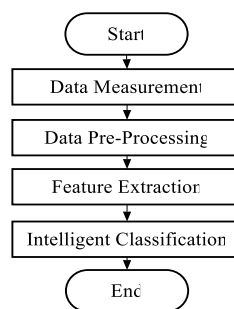
Figure 1. Overall flowchart for waxy crude oil odor-profile classification.

2.1. Waxy Crude Oil Sample

Four types of waxy crude oil have been selected for this study. All the samples were supplied by PETRONAS from four basins named ANGSI (Sample A), PENARA (Sample B), TAPIS (Sample C), and SEPAT (Sample D). All supplied samples were stored in a closed container to avoid any possibility of contamination. During the data collection, the samples were kept in a bottle with 10mL in volume. The waxy crude oil sample for E-Nose data collection is shown in Figure 2 below.
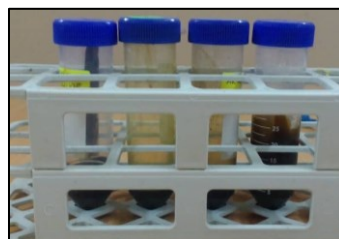
Figure 2. Sample preparation of waxy crude oil for E-nose data measurement.

2.2. Electronic Nose Setup

Figure 3 below shows the E-Nose setup to measure the odor data. There are 4 sensor arrays in a parallel position located inside the E-nose chamber. An internal pump in the E-Nose is used to suck in and control the inlet source of the gas from the waxy crude oil sample. The sensors were able to detect carbon monoxide (CO), LPG, CO, CH4, natural gas, propane, methane, i-butane, alcohol, hydrogen, and smoke. During the experiment, the odor data from the four sensors array will be sent to the computer with Arduino software as a feature. The heating system was applied in order to maintain the sample temperature at 30°C.
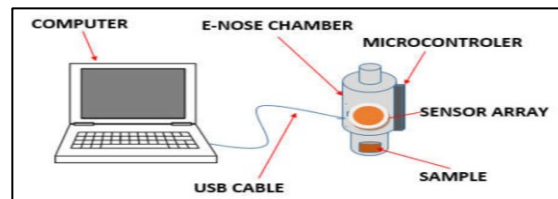
Figure 3. Electronic nose experimental setup.

The measurement phase for each experiment will require a period of 2 minutes while the cleaning phase requires 3 minutes. In the cleaning phase, the sample was removed from the E-nose chamber while the system and pump are running to clean the odor from the previous sample before running the next experiment. The cleaning phase is required to reduce the reading error. There are 200 data collected for each experiment. The experiment was repeated five times for each sample. All of the collected data were then tabulated in Table 1 below. S_1, S_2, S_3, and S_4 represents sensor 1, sensor 2, sensor 3, and sensor 4 respectively while DM is the data measurement. The subscript number for DM is the number of odor reading and sensor number. For example, DM_{11} is the first data measurement 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}

2.3. Data Pre-Processing

In the data processing stage, the collected data will be normalized by using Equation 1 shown below,

$$R^{'} = \frac{R}{R_{max}},$$

(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. Based on Table 2 below, all the data will be divided with the biggest value in their own row. The main purpose of normalization is to rescale the data from zero to one (0 to 1). Table 2 below shows the tabulated data normalization for waxy crude oil. ND is the normalized data while S_1, S_2, S_3, and S_4 represents sensor 1, sensor 2, sensor 3 and sensor 4 respectively. The subscript number for ND represents the number of odor reading and sensor number. For example, ND_{11} is the first normalized data measurement for sensor 1.

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}

2.4. K-NN Classification Technique
After the normalization and feature extraction, intelligent classification technique using k-NN will be used to classify the types of waxy crude oil samples. K-NN is widely known as a simple algorithm which stores and groups all the cases in view of similarity distance measure. Data input, data output, data splitting, data training, data testing and parameter optimization steps were used for K-NN classification analysis.

3. Results and Discussion

3.1. Raw Data Measurement
The data measurement for all types of waxy crude oil samples is shown in Figure 4-5 in page 4 and Figure 6-7 in page 5 below. Y-axis represents the odor reading data in resistance value (Ω) while the x-axis represents the four sensors used in the E-nose (S_1=sensor 1, S_2=sensor 2, S_3=sensor 3, and S_4=sensor 4). There are 1000 data plotted for each sensor. The total odor reading for each type of sample in each graph is 4000 data. Although each sensor has their own range of data measured, it will not change the odor-pattern of the waxy crude oil. The graph also shows that for each type of waxy crude oil sample from different fields will have their own odor pattern. The most fascinating part is at sensor 1, it shows that all types of waxy crude oil sample have released carbon monoxide (CO) which could be detected using sensor 1.
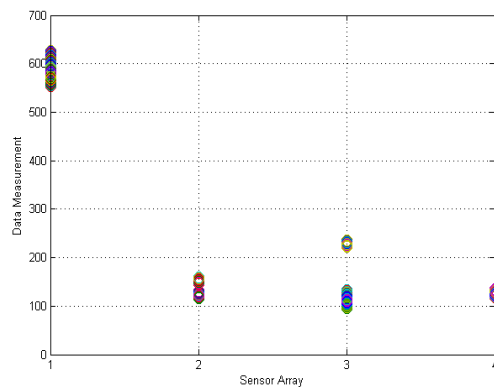
Figure 4. Raw data for sample A.
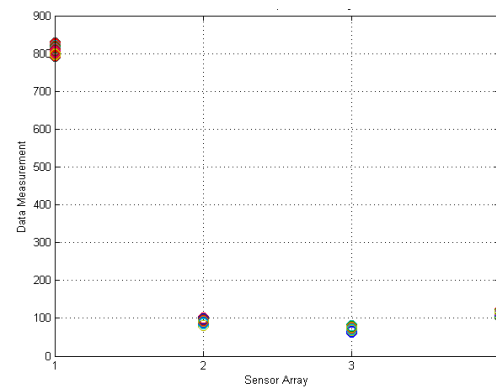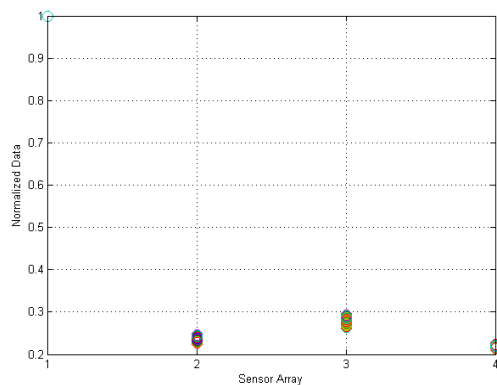

Figure 5. Raw data for sample B.
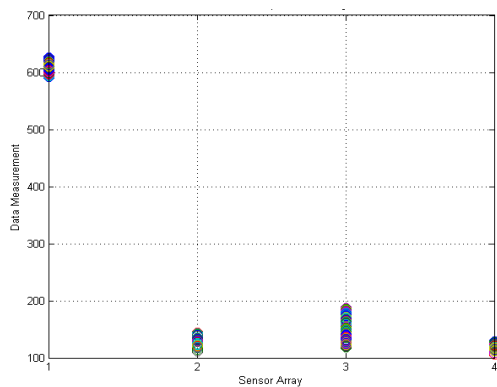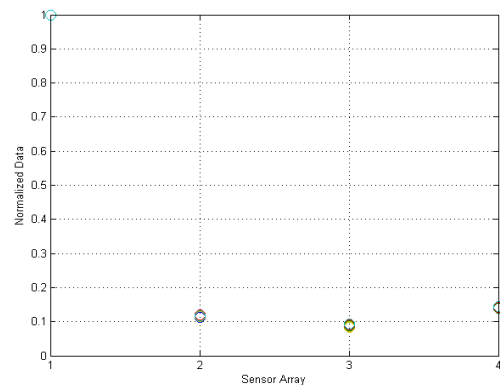
Figure 6. Raw data for sample C.
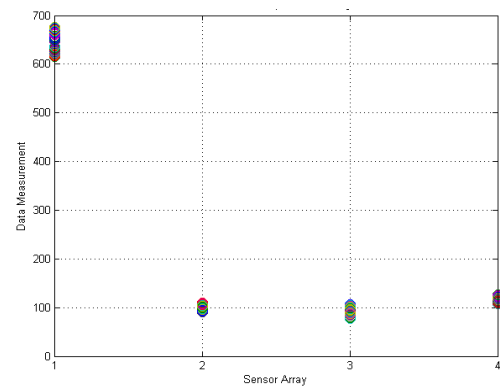
Figure 7. Raw data for sample D.

3.2. Data Pre-Processing

In the data processing stage, all of the 16,000 data will be divided with the biggest value in their own row (refer to Table 1). Figure 8-11 shows the graph of normalized data for each type of waxy crude oil sample. Y-axis indicates the normalized data while the x-axis indicates the E-Nose sensor array. As mentioned in the previous section, the normalization phase will rescale all the data in the range of 0 to 1 interval. The figures show that all the data was successfully normalized.
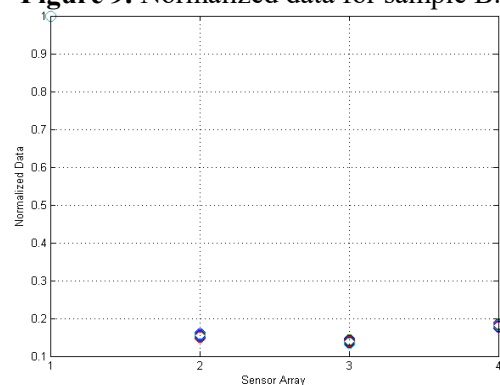

Figure 8. Normalized data for sample A.

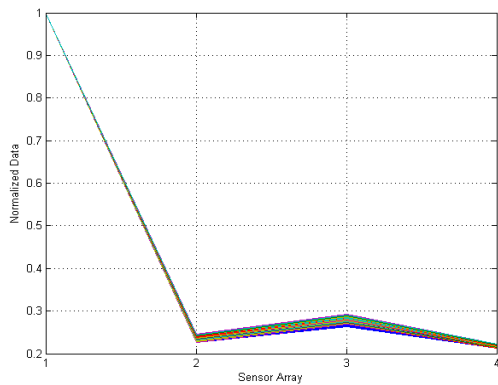**Figure 9.** Normalized data for sample B.
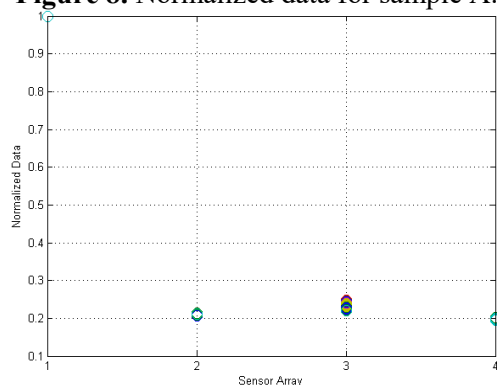
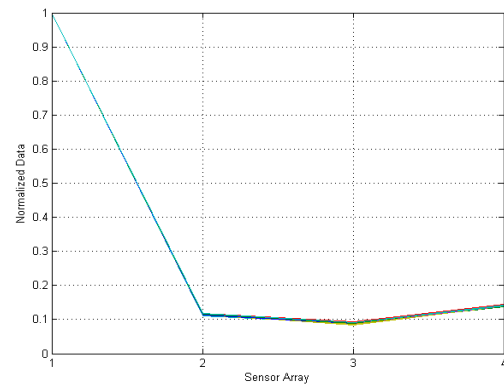**Figure 10.** Normalized data for sample C.

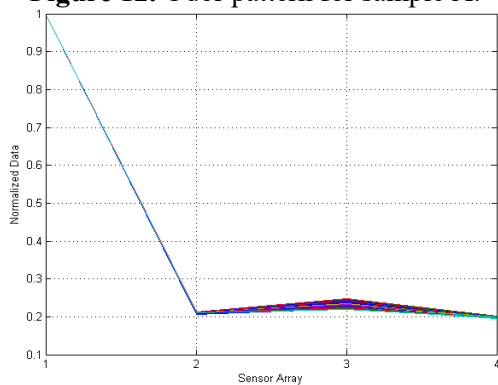**Figure 11.** Normalized data for sample D.

After the normalization stage, the odor-pattern will be extracted as in Figure 12-15 on page 6. It is clearly shown that the odor pattern of each type of waxy crude oil sample is very unique and different between each other. It also shows that each type of waxy crude oil have their own properties and release different gasses at 30ºC.
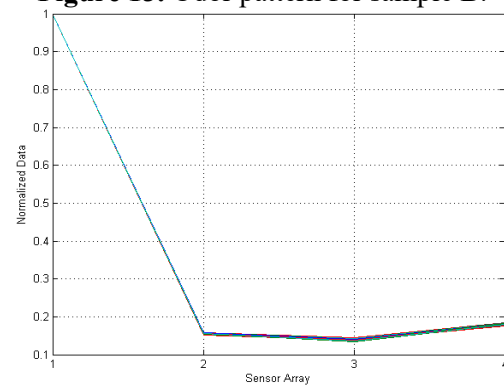
Figure 12. Odor pattern for sample A.

**Figure 13.** Odor pattern for sample B.


Figure 14. Odor pattern for sample C.

Figure 15. Odor pattern for sample D.

3.3. Boxplot Analysis

Before the classification stage, the normalized data will be analyzed using boxplot analysis. Boxplot is a well-known statistical method and very useful in identifying and comparing the data distributions. The results of boxplot analysis are shown in Figure 16-17 in page 6 and Figure 18-19 in page 7. Y-axis shows the normalized data and x-axis shows the number of associated sensors. For this type of analysis, there is a median, first quartile, third quartile, the maximum and minimum value of the data at each boxplot for each sensor. Sensor 1 shows the best data result due to the constant reading in all waxy crude sample.
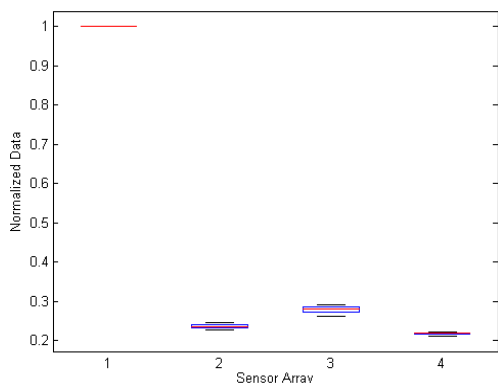
Figure 16. Boxplot analysis for sample A.
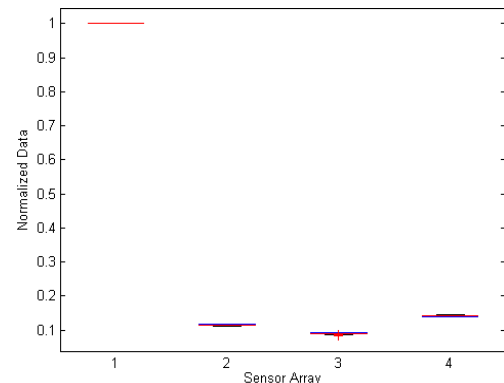
Figure 17. Boxplot analysis for sample B.
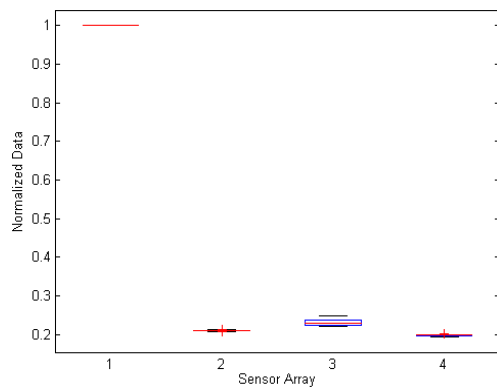
Figure 18. Boxplot analysis for sample C.

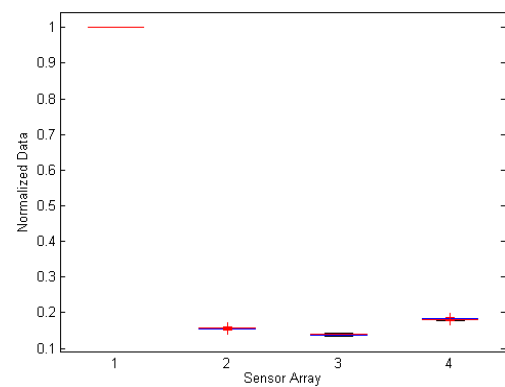
Figure 19. Boxplot analysis for sample D.

3.4. K-Nearest Neighbour (K-NN)

K-NN is an excellent classification technique and frequently used among researchers. There are nine data splitting ratios for training and testing data in K-NN classification which are 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20 and 90:10. Table 3 below shows the outstanding classification results at 60:40 splitting data ratio (60% data training and 40% data testing). All four distances (Euclidean, Cityblock, Cosine, and Correlation), three rules (Nearest, Random and Consensus) and three types ok K (K=1, K=2 and K=3) in parameter optimization table were able to classify all the four types of waxy crude oil with 100% accuracy.

Table 3. Parameter optimization table for waxy crude oil classification using K-NN at 60:40 splitting data.

Distance	Rule	Ratio 60:40		
		Percentage similarity (%)		
		K=1	K=1	K=1
Euclidean	Nearest	100	100	100
	Random	100	100	100
	Consensus	100	100	100
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	100
Correlation	Nearest	100	100	100
	Random	100	100	100
	Consensus	100	100	100

3.5. Linear Regression

The training and testing data in the previous subsection were validated using linear regression. Regression plot for a type of waxy crude oil was shown in Figure 20 in page 8 for sample A, B, C, and D respectively. Y-axis is the output of crude oil sample and x-axis is the target. The regression graph was plotted to measure the percentage of error in the classification stage. In Figure 8, the testing data is in the blue line with '+' sign while the training data is in the red line with 'O' sign. The data was successfully divided into four classes which are sample A, B, C and D respectively. It obviously shows that the classification results for all types of waxy crude oil are 100% accurate with zero percentage of error.
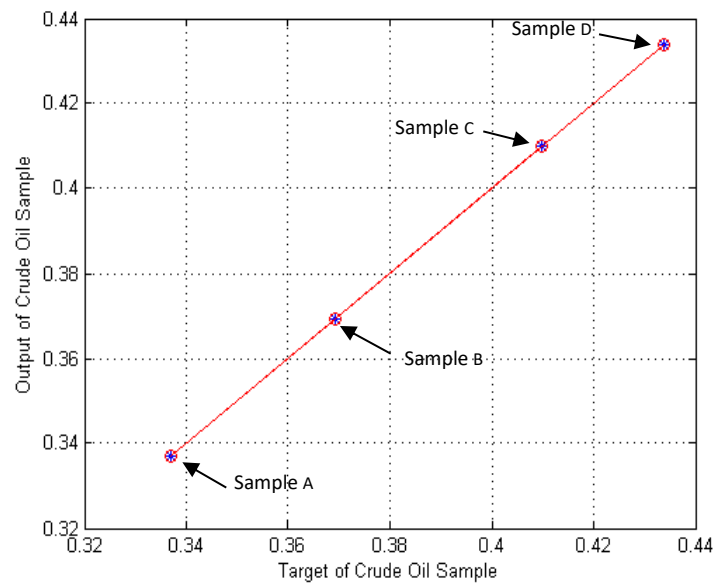
Figure 20. Regression plot for type of waxy crude oil.

4. Conclusion
This paper has presented the classification of four types of waxy crude oil based on their odor. The study shows that for each type of waxy crude oil, they have their own odor-profile at 30°C. The distinctive and different odor-profiles were influenced by the sample properties which releases different gases and aroma. The data pre-processing and boxplot analysis was done successfully to extract the odor pattern for each type of waxy crude oil. The k-NN classification technique used has achieved 100% accuracy with zero percentage of error. Correlation between odor and waxy crude oil properties could be done later for future study.

5. References
[1] N. Ridzuan, "Experimental Investigation and Modelling of Wax Deposition Inhibition in Pipeline Transportation of Crude Oil," Universiti Malaysia Pahang, 2016.
[2] A. Mohammad and U. M. Maung, "The Drilling And Completion Of The First Monobore Well In Peninsular Malaysia," 2000.
[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. 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.
[5] 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.
[6] 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.
[7] F. A. Halim, "Classification of Ammonia from Water Based on Odor-Profile using K-NN and CBR," Universiti Malaysia Pahang, 2016.
[8] 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.
[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] N. Zahed, "Intelligent Odour Profile Classification of Honey using Case Based Reasoning and K Nearest Neighbors," 2018.