

Research Article

Predicting the Viscosity of Petroleum Emulsions Using Gene Expression Programming (GEP) and Response Surface Methodology (RSM)

Abubakar A. Umar , Ismail M. Saaid , Aliyu A. Sulaimon , and Rashidah M. Pilus

Department of Petroleum Engineering, Universiti Teknologi PETRONAS, 32610 Bandar Seri Iskandar, Perak, Malaysia

Correspondence should be addressed to Aliyu A. Sulaimon; aliyu.adebayor@utp.edu.my

Received 4 April 2019; Revised 5 August 2019; Accepted 28 August 2019; Published 6 January 2020

Academic Editor: Syed Abdul Mohiuddine

Copyright © 2020 Abubakar A. Umar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper summarizes an investigation of certain operating parameters on the viscosity of petroleum emulsions. The production of crude oil is accompanied by emulsified water production, which comes along with various challenges like corroding the transport systems and catalysts poisoning during petroleum refining in the downstream. Several process variables are believed to affect the ease with which emulsified water can be separated from emulsions. Some of the issues have not been extensively examined in the literature. The simplicity with which water is separated from petroleum changes with age (after formation) of the emulsion; notwithstanding, this subject has not been investigated broadly in literature. This study tries to assess the correlation between aging time, water cut, crude oil viscosity, water viscosity and amount of solids and viscosity of petroleum emulsions. To achieve that, a response surface methodology (RSM) based on Box-Behnken design (BBD) was used to design the experiment. Synthetic emulsions were prepared from an Offshore Malaysian Crude oil based on the DoE design and were aged for 7 days. The emulsions viscosities were measured at 60-degree Celsius using an electromagnetic viscometer (EV100). The broad pressure and temperature range of the HPHT viscometer permit the imitation of acute conditions under which such emulsions may form. The data obtained from the RSM analysis was used to develop a prediction model using gene expression programming (GEP). It was discovered that the viscosity of water has no effect on the viscosities of the studied emulsions, as does the water cut and amount of solids. The most significant factor that affects emulsion viscosity is the aging time, with the emulsion becoming more viscous over time. This is believed to be imminent because of variations in the interfacial film structure. This is followed by the amount of solids, also believed to be as a result of increasing coverage at the interface of the water droplets, limiting the movements of the dispersed droplets (reduced coalescence), thereby increasing the viscosity of the emulsions.

1. Introduction

An emulsion has been defined in numerous ways, to mean a system that contains two (immiscible) aqueous phases, one of the phases dispersed as droplets (internal phase) in another (the continuous phase). In the presence of oil and brine, at least two types of emulsions can form. These emulsions are, either one in which the oil is dispersed in the brine (also called oil-in-water) and the second in which brine is dispersed in the oil (also called water-in-oil) [1–2]. Almost all emulsions found in the oil fields are those in which brine is dispersed in the oil; the water-in-oil type. In this paper, all emulsions discussed refer to this type of emulsion [3–5]. Other definitions provided

described emulsion as a blend of two unmixable liquids when blended, one of which is spread (dispersed) as very small globules in the other, and is made stable by an emulsifying agent [5–7]. Something basic to all the descriptions provided in the literature by several authors is evident in the thermodynamic instability of these emulsions, which makes them separate into water and oil phases upon leaving them to settle over time [6]. The thermodynamic instability of these emulsions, herein defined as macro-emulsions (with dispersed phase greater than $0.1 \mu\text{m}$), is assumed so because the contact between the oil and water molecules does not support the mix-up, hence, they will invariably break into separate phases with time [8, 9, 10]. This aspect of emulsions has received a lot of

attention from researchers, leading to the availability of comprehensive literature about its formation and stabilization [11]. A more extensive understanding of w/o emulsions requires a deeper knowledge of the materials responsible for their formation and stabilization, especially so, since the materials responsible for their formation vary across areas of application and study. Also, extensive knowledge of their physical properties like viscosity is necessary for its handling and treatment.

The viscosity of a liquid is described as a measure of the internal fluid friction in the liquid. In other words, it is the opposition to fluid flow, shear, and agitation. In addition to the viscosity of the continuous phase, the viscosity of an emulsion is complicated by the presence of certain solid particles it contains, as well as the aging period [12]. The flow behaviours of undiluted emulsions are of concern in several applications. For instance, the selection of blending equipment and the energy specifications for creating an emulsion is subject to the emulsion rheology. Predicting the transport function of emulsions within pipelines requires familiarity with likely shifts of viscosity with the shear rate besides concentration. Several influential industry operations such as oil-well drilling, fracturing, and acidizing methods, together with other operations that involve handling of large volumes of emulsions require understanding and control of the viscous properties of these emulsions. Despite the importance of this problem, there has been limited theoretical or empirical expressions that correlate the viscosity of petroleum emulsions with the amount of solids present in the continuous phase, the viscosity of the phases and variation in aging time [13, 14].

2. Materials and Methods

2.1. Design of Experiment-Response Surface Methodology. Both DoE and RSM are typically used to statistically identify the significant factors (independent variables) in a model and to generate response surface model that can be used to predict the outcome by providing a mathematical equation. In this study, the experiment was designed using Stratigraphic Centurion version VII. In the design of experiments, five independent variables were considered, which are: aging time, water viscosity, oil viscosity, water cut and amount of solids. The dependent variable (response) is the emulsion viscosity (μ_E). Each of the independent variable (x_1, x_2, x_3, x_4 , and x_5) was varied numerically over 3 levels and coded as -1, 0, and 1. Analysis of variance (ANOVA) and regression analysis was performed to ascertain how statistically consequential the model terms are, and as well fit a regression correlation correlating the experimental data to the independent variables. Assuming the variations of Y , (emulsion viscosity) obey an eight parameter, second order equation of the form shown below:

$$\text{Viscosity } (Y) = \beta_0 + \sum \beta_i x_i + \sum \beta_{ii} x_i^2 + \sum \beta_{ij} x_i x_j, \quad (1)$$

where Y , the response value estimated by the model; β_0 is constant, β_1 an offset value is linear, β_{ii} is quadratic and β_{ij} is inter-relationship coefficients of regression.

The model competency was defined using model summary; lack-of fit test and coefficient of determination (R^2) analysis. Joglekar, [15] suggested that for R^2 to be regarded for an acceptable fitness of a response model, the value should be at least 0.80. For all statistically nonsignificant terms ($P > 0.05$), they were discarded from the initial model and the empirical

TABLE 1: Physicochemical properties of mineral oil.

Properties of oil phase			
	Property	Value	Units
	Density @ 25°C and 60°C	0.8159 and 0.7904	g/cc
Offshore Malaysian crude oil	Refractive index (RI) @ 25°C and 60°C	1.4617 and 1.4448	—
	Specific gravity @ 25°C and 60°C	0.8183 and 0.8061	—
	Viscosity @ 25°C	26.5	cP

data refitted only to significant ($P < 0.05$) variable effects to obtain the final reduced model.

2.2. Preparation of Synthetic Emulsions. The synthetic emulsions used in this study were prepared using different water-to-oil ratios based on the design suggested by the DoE. Distilled water was used as the internal phase and a mineral oil (with properties shown in Table 1) as the continuous phase.

Before mixing the oil and water, the emulsifiers used in this study, a nonionic surfactant, Sorbitan monooleate (Span 80) and various amount of solids (with size $0.05\text{ }\mu\text{m}$) were combined with the oil phase and mixed with a homogenizer, for around 3 minutes. The concentration of surfactant Sorbitan Monooleate (Span 80), was 2 mg/L. Span 80 was chosen because it is a nonionic surfactant with Hydrophilic-lipophilic balance (HLB) of 4.3, used in the formation of water-in-oil emulsions. Surfactants with HLB values between 3 and 6, the oil-loving group (lipophilic) predominates, and they form w/o emulsions. The solids used in the study are inorganic solid extracts from Malaysian oilfield emulsions, majorly calcite materials. The characteristics of the solids have been reported in our previous study, Abubakar et al. [17].

This method is known as the surfactant-in-oil method. Furthermore, the oil that now contains the solids and surfactants is mixed with the deionized water employing a Homogenizer. The homogenization was made at 10 000 rpm for a period of 5 min. The prepared emulsions are then stored in graduated plastic centrifuge bottles for further analysis.

2.3. Gene Expression Programming (GEP). Gene expression programming (GEP) is a typically comprehensive model/phenotype system in which the chromosomes (genetic code)/expression trees form a truly operative, inseparable entity [16]. Indeed, in GEP things like invalid expression tree or program do not exist. Unmistakably, the interaction of GEP chromosomes and expression trees needs a clear-cut interpretation system to allot the sign of chromosomes into the language of expression trees. The benefits of a system like GEP are distinct from nature, but the most significant should be emphasized. First, the chromosomes are simple entities: linear, compact, comparatively small, simple to manipulate genetically (replicate, mutate, recombine, transpose, etc.). Secondly, the expression trees (ETs) are solely the expression of their individual chromosomes; they are the objects upon which assortment works and, according to fitness, they are chosen to generate with modification. During replication, it is the chromosomes of the individuals that are replicated with modification and forwarded to the next generation, not the ETs. Because of these characteristics, GEP is remarkably versatile and exceedingly outperforms the current evolutionary techniques [18, 19]. The flow diagram of a gene expression

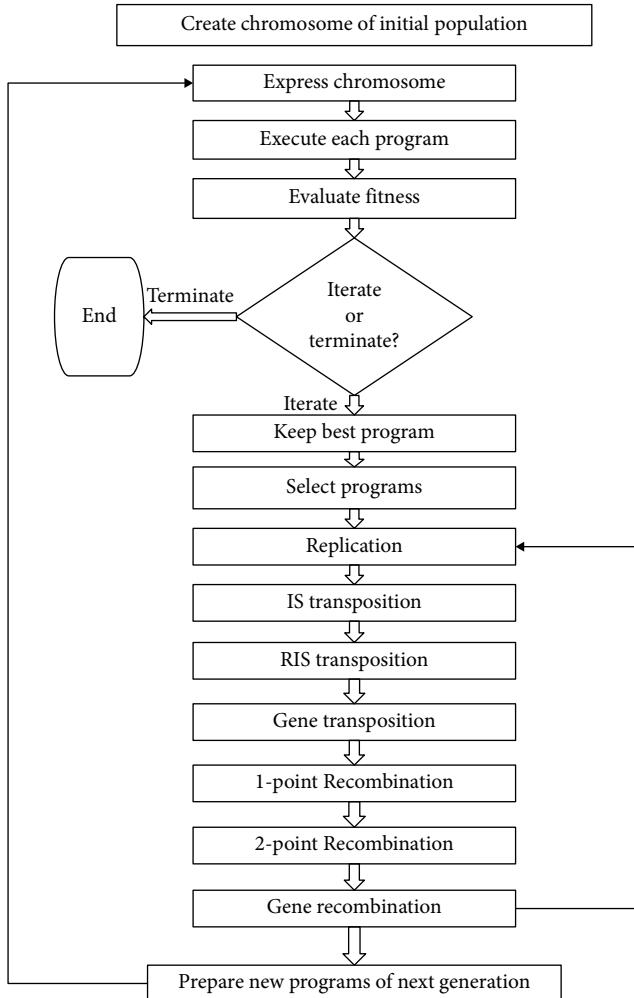


FIGURE 1: The flowchart of gene expression algorithm [16].

3. Results and Discussions

$$\begin{aligned}
 \text{Emulsion Viscosity(cP)} = & -744.412 + 11.4208 * \text{Time(Days)} + 1201.75 * \text{Oil viscosity(cP)} + 1.42624 * \text{Water viscosity(cP)} \\
 & + 5.81993 * \text{Water cut(v/v\%)} + 507.859 * \text{Amount of solids(wt\%)} - 0.166667 * \text{Time(Days)}^2 \\
 & - 8.49567 * \text{Time(Days)} * \text{Oil viscosity(cP)} + 0.00595238 * \text{Time(Days)} * \text{Water viscosity(cP)} \\
 & + 0.0704365 * \text{Time(Days)} * \text{Water cut(v/v\%)} + 7.39087 * \text{Time(Days)} * \text{Amount of solids(wt\%)} \\
 & - 496.36 * \text{Oil viscosity(cP)}^2 - 1.69977 * \text{Oil viscosity(cP)} * \text{Water viscosity(cP)} \\
 & - 4.68615 * \text{Oil viscosity(cP)} * \text{Water cut(v/v\%)} + 127.165 * \text{Oil viscosity(cP)} * \text{Amount of solids(wt\%)} \\
 & + 0.0046754 * \text{Water viscosity(cP)}^2 + 0.0413866 * \text{Water viscosity(cP)} * \text{Water cut(v/v\%)} \\
 & - 5.26961 * \text{Water viscosity(cP)} * \text{Amount of solids(wt\%)} - 0.010086 * \text{Water cut(v/v\%)}^2 \\
 & - 4.81151 * \text{Water cut(v/v\%)} * \text{Amount of solids(wt\%)} - 268.271 * \text{Amount of solids(wt\%)}^2.
 \end{aligned} \tag{2}$$

$$\begin{aligned}
 \text{Emulsion Viscosity(cP)} = & -645.828 + 4.36012 * \text{Time(Days)} + 1061.51 * \text{Oil viscosity(cP)} \\
 & + 0.0796569 * \text{Water viscosity(cP)} + 5.77108 * \text{Water cut(v/v\%)} \\
 & + 501.996 * \text{Amount of solids(wt\%)} + 0.0704365 * \text{Time(Days)} * \text{Water cut(v/v\%)} \\
 & - 444.461 * \text{Oil viscosity(cP)}^2 - 4.68615 * \text{Oil viscosity(cP)} * \text{Water cut(v/v\%)} \\
 & + 0.0413866 * \text{Water viscosity(cP)} * \text{Water cut(v/v\%)} \\
 & - 5.26961 * \text{Water viscosity(cP)} * \text{Amount of solids(wt\%)} - 0.00938823 * \text{Water cut(v/v\%)}^2 \\
 & - 4.81151 * \text{Water cut(v/v\%)} * \text{Amount of solids(wt\%)}.
 \end{aligned} \tag{3}$$

algorithm (GEA) shown in Figure 1 starts with the stochastic creation of the chromosomes of the original population. This is followed by displaying the chromosomes, and estimation of the fitness of each chromosome displayed. This is followed by the selection of individuals according to appropriateness to generate with adjustment, transmitting new traits to the progeny (descendants). The individuals of the new generation are constrained by the same exploratory process which involves the creation of the genomes, their contact with the collection environment and finally reproduction with adjustment. This process is duplicated for a specific number of generations or at least continuously till a solution has been found.

2.4. Viscosity Measurement – Electromagnetic Viscometer. The principle of operation of the electromagnetic viscometer is an effortless and definitive electromagnetic concept. Two coils push the piston backward and forward through a fixed magnetic induction force. Established connections analyze the piston's two-way movement time to estimate absolute viscosity. A built-in temperature detector (RTD) recognizes (by sensing) the actual temperature in the sampling container. The viscometer is made up of a Cambridge Electromagnetic Viscometer SPSL 440 and a set of six calibrated pistons to cover viscosities ranging from 0.2 cP to 10,000 cP. An external heating and recirculating bath provide heat. It is an equipment that gives a very accurate measurement of large viscosity ranges at reservoir conditions.

The equipment was ensured to be in good condition before use. One day before each test, the equipment was warmed while observing all the safety procedure. To take a measurement, the equipment, cooling bath and the desktop computer were switched on and all connections were checked properly. The cell was vacuumed and the range of viscosity (in centipoise) was chosen, the temperature and pressure desired were chosen and the sample cell was filled. The “run measurement” button was pressed, and measurement allowed to be taken. The result was taken, and data saved in the computer.

TABLE 2: Experimental design parameters.

Parameters	(-1)	0	1
X1, aging time (days)	1.0	4.0	7.0
X2, water viscosity (cP)	0.890	0.945	1.000
X3, oil viscosity (cP)	18	26.5	35
X4, water cut (v/v%)	20	35	50
X5, amount of solids (wt%)	0.17	0.29	0.41

Table 2 shows the lower, middle and upper experimental design parameters defined by -1, 0, and +1 respectively. Table 3 presents the predicted regression coefficients for the response variable, simultaneously with the equivalent R^2 , F -value, and the p -value of lack of fit. The ANOVA table distributes the variability in Emulsion Viscosity (cP) into discrete segments for the individual parameters. In addition, it examines the statistical significance (consequence) of the individual parameters by analyzing the mean square upon an estimation of the experimental error. Here, 11 parameters have P -values less than 0.05, showing that they are significantly different from zero at the 95.0% confidence level. The r -squared statistic proves that the model as fitted describes 87.164% of the unevenness in Emulsion Viscosity (cP). The adjusted r -squared statistic, which is more fitting when relating models with different numbers of independent variables, is 85.6359%. The standard error of the estimation shows the standard deviation of the residuals to be 8.71296. The mean absolute error (MAE) of 6.54856 is the average value of the residuals. The Durbin-Watson (DW) statistic examines the residuals to determine if there is any significant relationship with regards to the order in which they occur. Since the p -value is less than 0.05, there is an indication of possible serial correlation at the 5% significance level. From Table 3, it is evident that 11 parameters show significance in the viscosity of emulsions studied. Also, from Figures 2 and 3, Equations 2 and 3, we can observe that the length of each bar in the Pareto charts is proportional to the value of a t -statistic calculated for the corresponding effect. Any bars beyond the vertical line are statistically significant at the selected significance level, set by default at 5%. In this case, there are 11 significant main effects on emulsion viscosity. Equation 3 is therefore the refitted model equation that excludes all insignificant parameters.

Figure 4 is the main effect plot, and it shows how each factor affects the viscosity of the emulsion. The lines specify the predictable change in emulsion viscosity as each factor is moved from its low level to its high. The four factors with significant main effects are shown. These factors have higher influence on the response, i.e. emulsion viscosity. For example, the average emulsion viscosity at the 1st day of the emulsion's life is around 97, but averages to around 141 at the 7th day of observation. The difference of 44 is called the "main effect" of aging time. The main effect of oil viscosity is around 13, that of water cut is around 44 and that of the solids content is up to 49, indicating very significant effect. According to [20], the W/O emulsion would exhibit Newtonian characteristic when water content is no more than 20% and exhibit nonNewtonian characteristic if otherwise.

TABLE 3: Analysis of variance (ANOVA) for emulsion viscosity.

Source	Sum of squares	Df	Mean square	F-ratio	P-value
A:time (days)	35219.0	1	35219.0	463.92	0.0000
B:oil viscosity (cP)	1678.34	1	1678.34	22.11	0.0000
C:water viscosity (cP)	1.72024	1	1.72024	0.02	0.8805
D:water cut (v/v%)	16900.1	1	16900.1	222.62	0.0000
E:amount of solids (wt%)	22750.1	1	22750.1	299.68	0.0000
AA	63.0	1	63.0	0.83	0.3636
AB	146.72	1	146.72	1.93	0.1663
AC	1.72024	1	1.72024	0.02	0.8805
AD	750.149	1	750.149	9.88	0.0020
AE	132.149	1	132.149	1.74	0.1888
BB	336.667	1	336.667	4.43	0.0367
BC	70.7232	1	70.7232	0.93	0.3358
BD	1674.01	1	1674.01	22.05	0.0000
BE	19.7232	1	19.7232	0.26	0.6109
CC	17.04	1	17.04	0.22	0.6363
CD	3118.58	1	3118.58	41.08	0.0000
CE	808.938	1	808.938	10.66	0.0013
DD	769.056	1	769.056	10.13	0.0017
DE	2100.22	1	2100.22	27.67	0.0000
EE	139.286	1	139.286	1.83	0.1774
Total error	12753.8	168	75.9156		
Total (corr.)	99359.9	188			

The insignificance of water viscosity in the overall emulsion viscosity can be ascribed to the fact that water, which has lower viscosity than oil, is the dispersed phase whose molecules are smaller compared to the crude oil molecules that exist as chains which can get tangled up with each other. Due to droplet crowding, emulsions indicate nonNewtonian behavior, which is why their viscosity is higher than the viscosity of oil and water [21]. Another interesting observation that can be made on the main effects plot is the nature of variation of the effects. As seen, both aging time and amount of solids show parallel lines that continues, indicating that as the aging time and amount of solids increase, the viscosity of the emulsion would keep increasing. However, the water cut show a sign of curving upon reaching a maximum, indicating that there is a maximum value to which the effect of both oil viscosity and water cut would either cease or would be in the reverse effect.

As observed in the study, increase in solids concentration leads to a corresponding increase in emulsion viscosity, believed to have affected the dispersed droplets sizes (making them smaller). This could be as a result of numerous likely reasons: (a) with the decrease in droplet size, the average distance between the droplets decreases, leading to an increase in hydrodynamic interaction and viscosity; (b) the thickness of the adsorbed surfactant layer at the oil/water interface, and (c) the increase in viscosity, upon reduction of droplet size, could be partially due to an increase in the effective dispersed-phase concentration. It is difficult to determine which mechanism leads here, but all three are possible [22, 23].

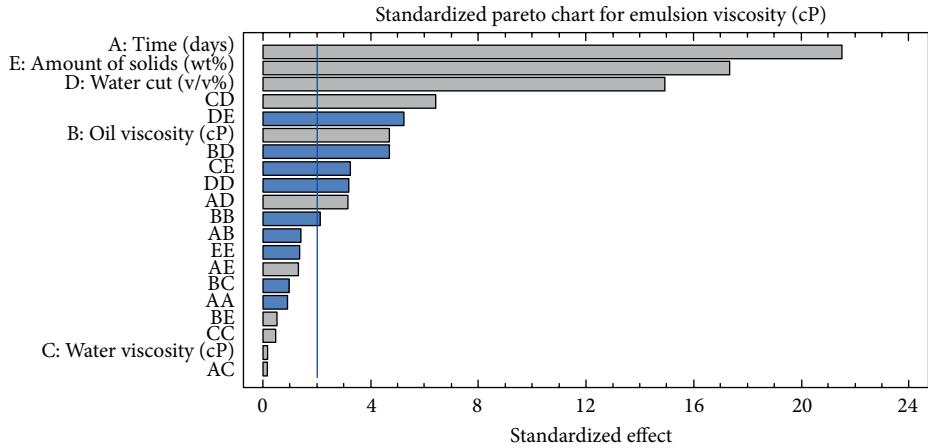


FIGURE 2: Pareto chart for emulsion viscosity (original model).

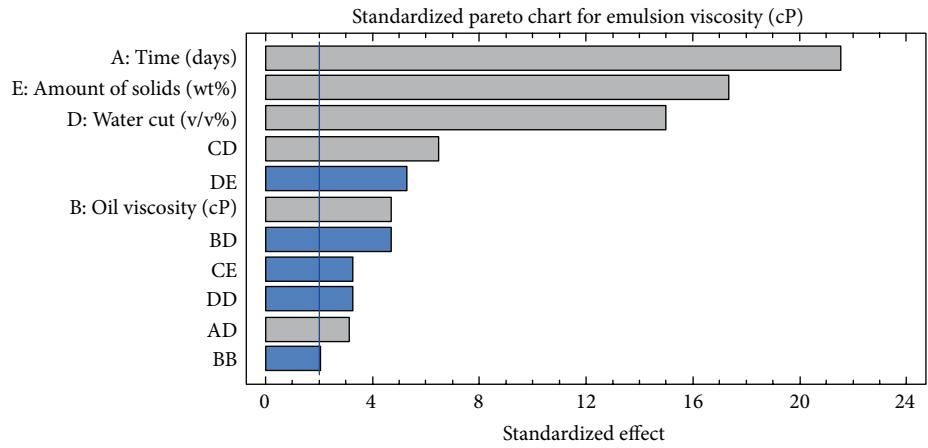


FIGURE 3: Pareto chart for emulsion viscosity (refitted model).

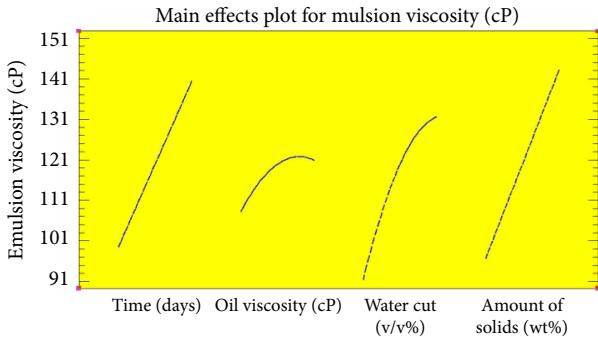


FIGURE 4: Main effects plot for emulsion viscosity.

To further understand the fitted model, 3-D surface plots shown in Figure 5 were plotted. On the Y-axis of all the plots shown is the emulsion viscosity, while the X and Z axes carry the variables that are compared. These graphs provide a clearer view of the effects of the variables as explained thus far.

To explore the accuracy of the models, a comparison between the predicted stabilities from all the three models was made with the corresponding results observed from the experiments, as depicted in Figure 6. The point collection around the 45-degree line illustrate that the deviation between the

predicted stabilities in the GEP model was slightly greater than of the RSM model, corresponding to the values of their R^2 of 0.8716 and 0.8318 respectively. This indicates that both models can be used to predict the viscosity of petroleum emulsions stabilized by the solids, water cuts and oil/water viscosities.

The gene expression programming (GEP) is a learning algorithm that combines both the simple, linear chromosomes of fixed length like the ones used in genetic algorithm (GAs) and the branched structures of different sizes and shapes like the parse trees of genetic programming (GP). This evolutionary technique was invented by Ferreira [16] in 1999. GEP algorithm performs excellently in providing solution to problems related to symbolic regression, optimization, time series analysis, classification and logic synthesis. The output from the GEP model is usually presented in form of mathematical equations, logical expressions, decision trees and neural networks.

Similarly, using the data obtained from visual observations and following the procedure for prediction using GEP highlighted in the methodology section, a GEP equation for the prediction of emulsion viscosity was obtained as:

$$ET_1 = \left[\left(\frac{C_6 + d_2}{d_4} \right) \times \left(\frac{d_3}{C_3} \right) \right] - \left(\frac{d_0}{d_4} \right) - \left(\frac{C_9}{d_4} \right). \quad (4)$$

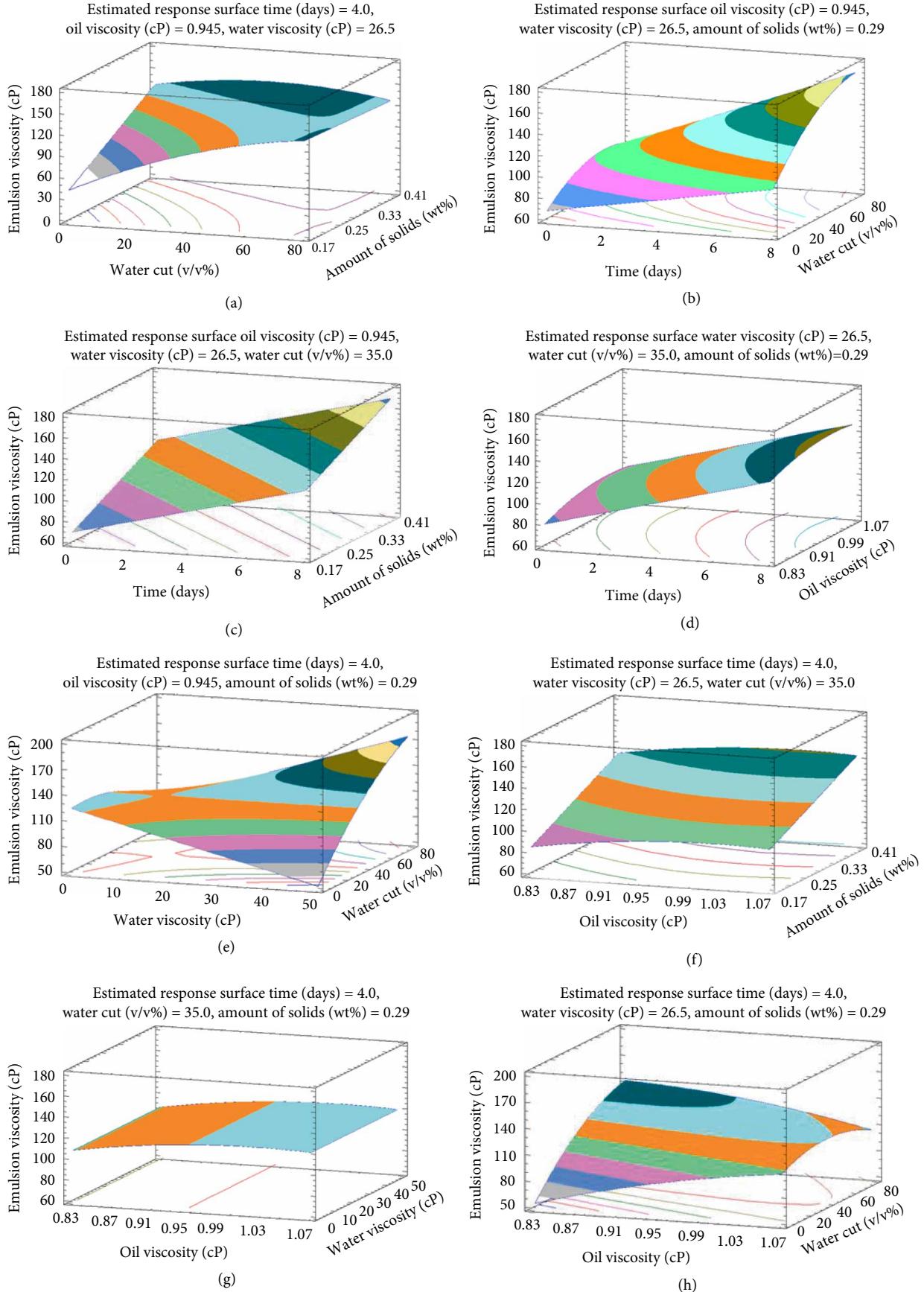


FIGURE 5: 3-D response surface plots of the combined effects of different parameters.

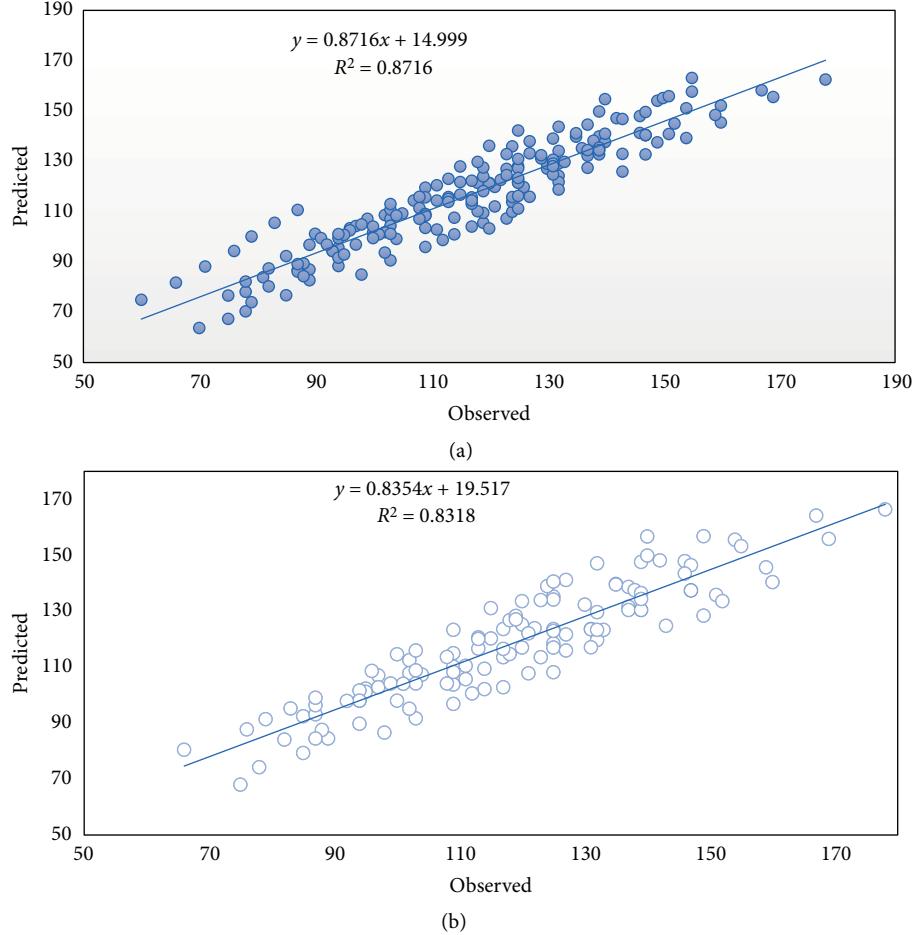


FIGURE 6: Cross plot of predicted viscosity versus measured viscosity for model developed with (a) RSM, and (b) GEP using experimental data.

$$ET_2 = \left[(C_2)^2 + \left(\frac{C_8}{d_2} \right) - \left(\frac{d_2 * C_0}{d_3} \right) \right] \div d_1. \quad (5)$$

$$ET_3 = \left[(C_3 + (d_4 * d_0)) + \left(\frac{C_6}{d_3} \right) \times (d_1)^3 \right]^2. \quad (6)$$

The overall viscosity is $V_E = ET_1 + ET_2 + ET_3$, where d_0, d_1, d_2, d_3 , and d_4 corresponds to the model input variables; aging time, water viscosity, oil viscosity, water cut and amount of solids respectively. The numerical constants for expression trees ET_1, ET_2 , and ET_3 are as shown below:

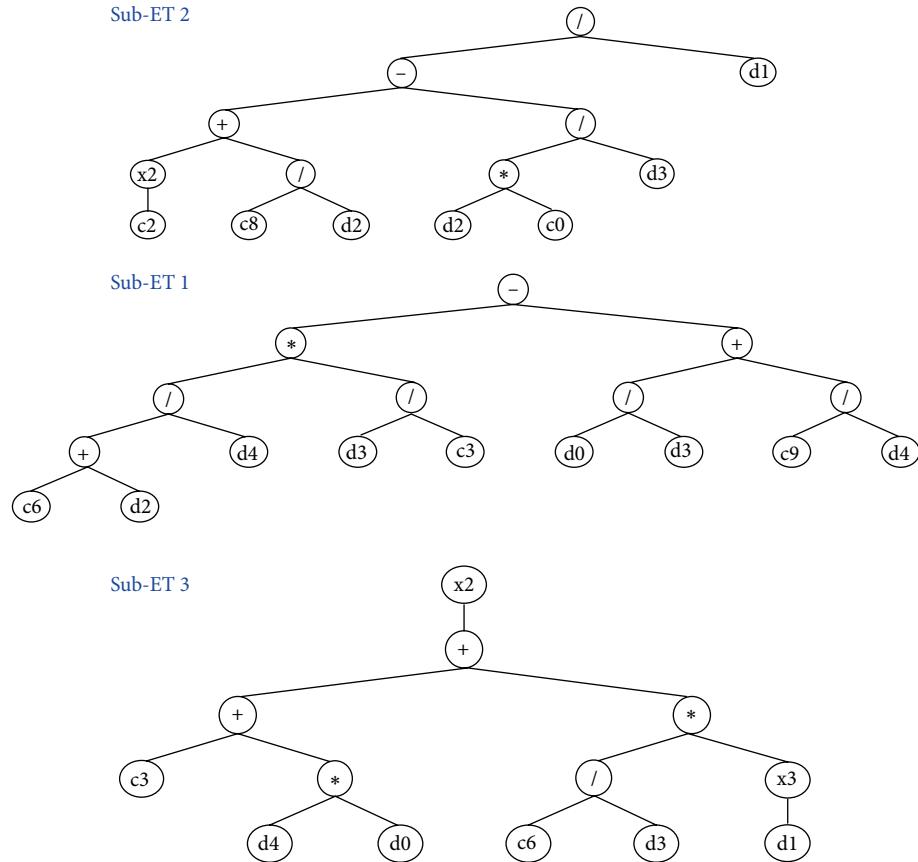
3.1. Numerical Constants. ET_1 :

$$\begin{aligned} c0 &= 0.353666096987823 \\ c1 &= 0.197454756309702 \\ c2 &= -1.6741522001709 \\ c3 &= 131.464361463475 \\ c4 &= -1.0879848628193 \\ c5 &= -1.22901089510788 \\ c6 &= -8.04498428296762 \\ c7 &= 3.79131443220313 \\ c8 &= 97.7135468611713 \\ c9 &= 12.523634793438. \end{aligned} \quad (7)$$

$$\begin{aligned} c0 &= 19.2643691418623 \\ c1 &= -0.258491775261696 \\ c2 &= -4.67818231757561 \\ c3 &= -7.49235365764336 \\ c4 &= -1.87968169194617 \\ c5 &= -1.56590472121342 \\ c6 &= -4.71236304818873 \\ c7 &= 3.50215552232429 \\ c8 &= 163.732440983008 \\ c9 &= 0.207415595263527 \end{aligned} \quad (8)$$

ET_2 :

$$\begin{aligned} c0 &= 7.19873653370769 \\ c1 &= -0.947325357829524 \\ c2 &= 3.61120059816279 \\ c3 &= 9.66429639576403 \\ c4 &= -64.6451363405255 \\ c5 &= -7.35761899380474 \\ c6 &= 22.1583136813391 \\ c7 &= -8.37128428758507 \\ c8 &= 45.5888851588488 \\ c9 &= 6.37928403576769 \end{aligned} \quad (9)$$

FIGURE 7: Sub-expressions trees ET_1 , ET_2 , and ET_3 .

Substituting the constants into Equations 4–6 gives:

$$\begin{aligned}
 V_E = & \left(\frac{d_2 - 8.04}{d_4} \right) \times \left(\frac{d_3}{131} \right) - \left(\frac{d_0}{d_4} \right) - \left(\frac{12.524}{d_4} \right) \\
 & + \left[21.81 + \left(\frac{163}{d_2} \right) - \frac{19.26d_2}{d_3} \right] \div d_1 \\
 & + \left[9.66 + (d_4 * d_0) + \left(\frac{22.16}{d_3} \right) \times (d_1)^3 \right]^2.
 \end{aligned} \quad (10)$$

Equation 10 is the finalized viscosity emulsion equation derived from the expression trees shown in Figure 7.

4. Conclusions

Understanding the behaviors of petroleum emulsions remain an important routine operation in the oil industry. It is believed that two third of all crude oils in the world are produced in form of emulsions. While some of these emulsions are easy to be separated, some separate with varying levels of difficulties. Understanding the viscosity of these water-in-oil emulsions are significant in designing field gathering systems and pipelines that handle wet oil (crude oil that contains water). Many wells produce crude oil in form of emulsions, and the viscosity of the emulsion will directly affect pumping pressures and tubing sizes. In this study, we have investigated the effects of

certain variables on the viscosity of emulsions. We have found that aging time of an emulsion has the greatest effect on its viscosity, where the emulsion viscosity increases with time. The amount of solids that accompany these emulsions, either from the continuous phase or from the surrounding environments also play significant importance, especially since it strongly affects the interfacial film that surrounds the dispersed phase. Among all the parameters investigated, this study shows that the viscosity of the water plays no significant role in the viscosity of the emulsion.

A statistical approach was adopted to assess the effect of the variables singly and in combination and resulted into a model equation that can be used to predict the viscosity of the emulsion. This is followed by a GEP model which also gives an equation that can be used to predict the emulsion viscosity.

Data Availability

The experimental (stability) data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors wish to extend maximum appreciation to the Institute of Contaminants, Universiti Teknologi PETRONAS (UTP) and yayasan utp grant (015LCO-076) for sponsoring this study.

Supplementary Materials

Supplementary 1: Data: the excel sheet contains experimental data from the viscosity measurements conducted for this study. This data is extracted from the data pane from the Design of Experiment (DoE) software. It contains four input parameters (Time, water viscosity, water cut, and amount of solids) as the parameters whose effects on emulsion viscosity we are studying. Different ratios based on the DoE suggestion were used to study each emulsion and the data is as presented in the “Data” sheet. RSM Plot: this is the second sheet on the excel sheet. It is extracted from the ANOVA pane on the DoE. It shows the observed emulsion viscosity (obtained from the experiments conducted) and the fitted viscosity (based on the prediction of the software). A plot of confidence level 5% and -5% was plotted to show the 95% confidence of the fitness. The plot is presented and the R² value with the equation is shown. GEP Plot: the GEP plot like the RSM plot shows the relationship between the observed viscosity values and the predicted by the GEP algorithm. The equation as well as the R² are shown. Viscosity Data: the last sheet is of the viscosity data, which shows the recorded values of different runs in the viscosity measurements conducted. (*Supplementary Materials*)

References

- [1] S. L. Kokal and M. Al-Dokhi, “Case studies of emulsion behavior at reservoir conditions,” *SPE Production & Operations*, vol. 23, no. 03, pp. 312–317, 2008.
- [2] P. L. Bansbach, “Treating emulsions produced by thermal recovery operations,” in *SPE California Regional Meeting*, Bakersfield, CA, USA, 1965.
- [3] C. Dalmazzone, C. Noik, and L. Komunjer, “Mechanism of crude-oil/water interface destabilization by silicone demulsifiers,” *SPE Journal*, vol. 10, no. 01, pp. 44–53, 2005.
- [4] D. J. McClements, *Food Emulsions: Principles, Practices, and Techniques*, CRC Press, Boca Raton, FL, USA, 2015.
- [5] H. M. A. Aziz, S. F. Darwish, and F. M. Abdeen, “Downhole emulsion problem, the causes and remedy, ras budran field,” in *SPE Asia Pacific Oil and Gas Conference and Exhibition*, Society of Petroleum Engineers, 2002.
- [6] P. Singh, W. H. Thomason, S. Gharfeh, L. D. Nathanson, and D. J. Blumer, “Flow properties of Alaskan heavy-oil emulsions,” in *SPE Annual Technical Conference and Exhibition*, Society of Petroleum Engineers, 2004.
- [7] S. Kokal, *Crude Oil Emulsions: Everything You Wanted to Know but Were Afraid to Ask*, Society of Petroleum Engineers, 2008.
- [8] N. Shahidzadeh, D. Bonn, O. Aguerre-Chariol, and J. Meunier, “Spontaneous emulsification: relation to microemulsion phase behaviour,” *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, vol. 147, no. 3, pp. 375–380, 1999.
- [9] D. Craig, S. Barker, D. Banning, and S. Booth, “An investigation into the mechanisms of self-emulsification using particle size analysis and low frequency dielectric spectroscopy,” *International Journal of Pharmaceutics*, vol. 114, no. 1, pp. 103–110, 1995.
- [10] L. M. Bodahl, *Rheology and flow properties of water-in-oil emulsions-comparing measured flow rates to predicted flow rates based on the rheological results from the laboratory [M.S. thesis]*, Norwegian University of Science and Technology, Trondheim, Norway, 2017.
- [11] R. Oliveira and M. Goncalves, “Emulsion rheology-theory vs. field observation,” in *Offshore Technology Conference*, Offshore Technology ConferenceHouston, TX, USA, 2005.
- [12] W. Woelflin, *The Viscosity of Crude-Oil Emulsions: Drilling and Production Practice*, American Petroleum Institute, Washington, DC, USA, 1942.
- [13] D. Dan and G. Jing, “Apparent viscosity prediction of nonNewtonian water-in-crude oil emulsions,” *Journal of Petroleum Science and Engineering*, vol. 53, no. 1-2, pp. 113–122, 2006.
- [14] W. Wei, P. Wang, K. Li, J. Duan, K. Wu, and J. Gong, “Prediction of the apparent viscosity of nonNewtonian water-in-crude oil emulsions,” *Petroleum Exploration and Development*, vol. 40, no. 1, pp. 130–133, 2013.
- [15] A. M. Joglekar and A. T. May, “Product excellence through experimental design,” In *Food Product Development from Concept to the Market Place*, E. Graf and I. S. Saguy, Eds., pp. 211–230, The AVI Publications, New York, NY, USA, 1987.
- [16] C. Ferreira, “Gene expression programming: a new adaptive algorithm for solving problems,” *Complex Systems*, vol. 13, no. 2, pp. 87–129, 2001.
- [17] A. A. Umar, I. B. M. Saaid, and A. A. Sulaimon, “Characteristics of native colloids and stability prediction of emulsions stabilized by asphaltenes, wax, silicates and calcites,” in *SPE Middle East Oil and Gas Show and Conference*, Society of Petroleum Engineers, 2019.
- [18] L. Teodorescu, “Gene expression programming software application for high energy physics data analysis,” in *Nuclear Science Symposium Conference Record*, 2006, IEEE, 2006.
- [19] S. Samadianfard, “Gene expression programming analysis of implicit Colebrook–White equation in turbulent flow friction factor calculation,” *Journal of Petroleum Science and Engineering*, vol. 92, pp. 48–55, 2012.
- [20] C. Li, Q. Huang, S. Ma, and C. Ji, “An experimental study on the viscosity of water-in-oil emulsions,” *Journal of Dispersion Science and Technology*, vol. 37, no. 3, pp. 305–316, 2016.
- [21] S. Nasery, S. Hoseinpour, L. T. K. Phung, and A. Bahadori, “Prediction of the viscosity of water-in-oil emulsions,” *Petroleum Science and Technology*, vol. 34, no. 24, pp. 1972–1977, 2016.
- [22] H. A. Barnes, “Rheology of emulsions—a review,” *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, vol. 91, pp. 89–95, 1994.
- [23] M. B. Lunde, *Rheology and Flow Properties of Water-in-Oil Emulsions-Comparing Measured Flow Rates to Predicted Flow Rates Based on the Rheological Results from the Laboratory*, [M. S. thesis], NTNU, Norway, 2017.

