Science & Technology Development Journal – Engineering and Technology 2025, ():1-21

Open Access Full Text Article

Research Article

Study of the variation of drilling mud density with temperature, pressure, and circulation rate using artificial neural networks, statistical models, and empirical correlations

Phạm Sơn Tùng[1,2,*], Phạm Thanh Nhân[1,2]

[1]*Department of Drilling and Production Engineering, Faculty of Geology and Petroleum Engineering, Ho Chi Minh University of Technology (HCMUT), 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Vietnam*

[2]*Vietnam National University Ho Chi Minh City, Linh Trung Ward, Ho Chi Minh City, Vietnam*

Correspondence

Phạm Sơn Tùng, Department of Drilling and Production Engineering, Faculty of Geology and Petroleum Engineering, Ho Chi Minh University of Technology (HCMUT), 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Vietnam

Vietnam National University Ho Chi Minh City, Linh Trung Ward, Ho Chi Minh City, Vietnam

Email: phamsontung@hcmut.edu.vn

History

- Received: 11-01-2024
- Revised: 30-4-2024
- Accepted: 13-11-2024
- Published Online:

DOI :

ABSTRACT

Well control is an important aspect of drilling operations because improper well control can result in kicks and blowouts with grave consequences. A successful well control requires a good understanding of the relationships between drilling mud pressure and formation pressure, as well as the variation of bottom hole pressure during drilling operations. As the hydrostatic pressure of the drilling mud column accounts for most of the pressure, a more accurate control of the changes of mud density will contribute to a more accurate bottom hole pressure modeling. Regarding the control of the mud density, a practical problem has existed so far in petroleum drilling: the mud density is determined at the surface condition, and its values vary along the depth of the well because of the changes of temperature and pressure, which consequently leads to an inaccuracy in mud density control in reality.

In order to reduce the inaccuracy in mud density control, this research aims to provide a reliable method to correctly predict the drilling mud's density under specific conditions. Different artificial neural networks (ANN) were proposed to predict drilling mud density based on the value of mud density at surface conditions, circulation rate, bottomhole pressure, and temperature. This study then used statistical methods to compare the predicted results with results obtained from existing empirical correlations and from other researchers' works to find out the most optimal artificial neural network which should consist of only one hidden layer. The main contributions of this research in comparison with existing papers are that: 1) Existing methods did not take into account the influence of circulation rate, therefore the real working conditions of the drilling mud were not represented entirely. Our research included the circulating rate in the ANN modeling and in the study of relative importance. The results indicated that the value of mud density at surface conditions had the greatest effect on the prediction results, and the influence of the circulating pump flow rate is small but should not be ignored; 2) Our research used different methods (ANN, Generalized Additive, Nonlinear Function) to predict the mud density in variation with temperature and pressure, which has never been approached in existing literature; 3) The sufficiency in the number of data was studied in this research, which has never been treated in previous studies. The Bootstrap method was used in this regard; 4) We remarked that the overfitting has not been treated properly in the existing literature review in this field, hence we included a thorough analysis of the overfitting in this paper. Finally, the results of this paper can be useful in real life because it can help drillers to accurately predict the mud density under varied conditions of pressure and temperature, and therefore to increase the safety of the drilling operations.

Key words: mud weight, machine learning, artificial neural networks, empirical correlations

INTRODUCTION

Ensuring safety is always the top priority in the oil and gas industry because accidents related to the petroleum sector often lead to loss of time, infrastructure, finance, and manpower. One of the accidents causing severe consequences is the loss of well control during the drilling process, specifically when the pressure in the wellbore is lower than the formation pressure. This scenario can happen if the mud density is not controlled adequately during the drilling operation due to the variation of pressure and temperature inside the wellbore, and consequently the mud density may be too low to maintain bottomhole pressure equal to formation pressure Cormack, 2017[1]. Therefore, being able to accurately calculate the mud density will help to assure a successful drilling operation. In order to achieve this objective, studying the influence of different factors affecting the density of the drilling fluid is extremely necessary.

In literature, there have been various studies relating to the prediction of drilling mud density at different conditions. It is well known that when bottom-

Cite this article : Tùng P S, Nhân P T. **Study of the variation of drilling mud density with temperature, pressure, and circulation rate using artificial neural networks, statistical models, and empirical correlations** . *Sci. Tech. Dev. J. – Engineering and Technology* 2025; ():1-21.

hole pressure increases, drilling mud density will increase since the drilling fluid volume is compressed, and conversely, when the bottom hole temperature increases, the drilling fluid volume expands leading to a decrease in its density, which is mentioned in Babu, 1996[2]; Hussein and Amin, 2010[3]; An et al., 2015[4]. McMordie et al., 1982[5] conducted an experimental research about the changes of drilling mud density with temperature (70-400 ^{o}F) and pressure (0-14000 psi). Similarly, Demirdal & Cunha, 2009[6] conducted experiments to study the variation of drilling mud density with the same range of pressure (0-14000 psi) but with a different range of temperature (25-175 ^{o}C). Zamora et al., 2013[7] also conducted experiments to study the volumetric behavior and the variation of density of base oils, brines, and drilling fluids with the range of temperature (36^{o}F–600^{o}F) and pressure (0-30000 psi). Some studies provided empirical correlations between mud density and pressure and temperature, such as Kemp, 1989[8]; Peters et al., 1990[9]; Isambourg et al., 1996[10]; Zamora et al., 2000[11]; Hemphill and Isambourg, 2005[12]; and Peng et al., 2016[13]. egarding the application of machine learning in this field, some authors used Artificial Neural Network (Osman et al., 2003[14]; Adesina 2015[15], Okorie E. Agwu et al., 2020[16]), while some others used different methods such as Fuzzy logic (Ahmadi et al., 2018[17]), Support Vector Machine (Xu et al., 2014[18]; Ahmadi, 2016[19]; Kamari et al., 2017[20]), Radial Basis Function Artificial Neural Network (Rahmati & Tatar, 2019[21]), and Particle Swarm Optimization Artificial Neural Network (Ahmadi et al., 2018[17]; Zhou et al., 2016[22]). It is also worth mentioning the hydraulic model proposed by Charlez et al., 1998[23] to calculate downhole pressure and then to predict fluid downhole density. In brief, the common point of these studies is to predict drilling mud density at different bottomhole pressures and temperatures.

However, besides temperature and pressure, some other factors also affect the density of drilling fluid, such as the inclination angle of the well which was highlighted in the study of Tian et al., 2013[24]; or the type of drilling fluid which was mentioned in the studies of Demirdal et al., 2007[25] and Demirdal & Cunha, 20096; and finally the circulation rate which was mentioned in the studies of Kårstad & Aadnøy, 1998[26] and Harris & Osisanya, 2005[27]. The study of Hemphill, 1996[28] investigated the effect of inclination angle and of cuttings on drilling fluid properties. Boatman, 1967[29] studied the influence of shale on drilling fluid density.

In reality, it is challenging to observe the changes in drilling fluid density because of costly specialized measuring equipment which must comply with well design requirements. Ombe et al., 2020[30] developed a specific measurement to achieve this task. Hoseinpour et al., 2022[31] combined well logging and geomechanical parameters to determine the mud window, but the authors could not predict the variation of the mud density in function of pressure, temperature, and some other factors.

In brief, the above literature review showed that developing a new method to accurately predict drilling mud density in the well under influence of various factors is necessary, which is the objective of our study. In this study, we resorted to not only machine learning methods but also empirical correlations as well as mathematical, and statistical methods.

Regarding the empirical correlations, Furbish, 1997[32] provided the following equation of state for liquid density:

$$\rho = \rho_0 \left[1 - \alpha (T - T_0) + \beta (P - P_0)\right] \tag{1}$$

ρ (ppg) is predicted drilling mud density, ρ_0 is value of mud density at surface conditions, T and P are final temperature (^{o}F) and pressure (psi), T_0 and P_0 are standard temperature (^{o}F) and pressure (psi), α ($^o\mathrm{F}^{-1}$) is isobaric coefficient and β($^o\mathrm{F}^{-1}$) is isothermal compressibility. These coefficients were taken from the work of Zamora et al. (2000) wherein they used 0.0002546 và 2.823×10^{-6} for α and β respectively for oil-based mud.

Another empirical correlation given by Hoberock et al., 1982[33] predicted oil-based mud density and water-based mud through the law of conservation of mass as detailed in the following:

$$\rho\left(P_2, T_2\right) = \frac{\rho_1}{1 + f_0 \left(\dfrac{\rho_{01}}{\rho_{02}} - 1\right) + f_w \left(\dfrac{\rho_{w1}}{\rho_{w2}} - 1\right)} \tag{2}$$

$\rho\left(P_2, T_2\right)$ is predicted drilling mud density, ρ_1 (ppg) is value of mud density at surface conditions, ρ_{01} (ppg) is initial oil density, ρ_{02} (ppg) is oil density in predicted drilling mud, ρ_{w1} (ppg) is water density in initial drilling mud, ρ_{w2} (ppg) is water density in predicted drilling mud, f_0 (%) is the percentage of oil volume in the drilling fluid, f_w (%) is the percentage of water volume in the drilling fluid.

Kutasov, 1988[34] presented an empirical correlation to calculate drilling mud density:

$$\rho_m = \rho_{mo} e^{[\alpha(P-P_0)-\beta(T-T_0)-\gamma(T-T_0)]} \tag{3}$$

ρ_m (ppg) is the predicted drilling mud density, ρ_{mo} (ppg) is the drilling mud density at standard conditions. P_0(psi) and T_0(^{o}F) are standard pressure and temperature. P(psi) and T(^{o}F) are the pressure and temperature at the predicted position. Kutasov evaluated α, β, γ, and ρ_{mo} with 5 drilling mud examples from McMordie et al., 1982[5]. Besides, Kutasov's correlation can be applied to oil-based mud and water-based mud. In our paper, the values of , and , which were taken from the work of Micah, 2011[35], were 3.0997×10^{-6}, 2.2139×10^{-4}, and 5.0123×10^{-7}, respectively.

Sorelle et al., 1982[36] focused on the changes in the volume of the components in drilling fluid caused by temperature and pressure, as being shown in the following formula:

$$\rho_f = \frac{\rho_i}{1 + \frac{\triangle V_o}{V} + \frac{\triangle V_w}{V}} \qquad (4)$$

ρ_f (ppg) is the predicted drilling mud density, ρ_i (ppg) is the value of mud density at surface conditions, $\triangle V_o$ (gal) is the change in oil volume, $\triangle V_w$ (gal) is the change in water volume, V(gal) is the total volume.

The literature review allowed us to see some possible contributions that we can bring to the research in this domain. Firstly, the study developed an artificial neural network modeling to predict drilling fluid density, combined with various mathematical, statistical (generalized additive model) and experimental models on the same dataset to provide a comprehensive and multidimensional understanding of the changes in drilling fluid density inside the wellbore. The simulation results were compared with actual data to verify the accuracy of the model.

Secondly, the number of features that our research used for the artificial neural network was greater than in previous studies. As mentioned above, previous papers considered mostly temperature and pressure as input features, while this study presented an artificial neural network modeling with inputs consisting of not only bottomhole temperature and pressure but initial drilling fluid density and circulation rate as well. Consequently, this paper conducted a study about the effect of various influencing factors mentioned above, besides pressure and temperature, on the drilling mud density.

Thirdy, this paper took into account the possible influence of the low number of input data used for ANN modeling. It is difficult to answer the question if a data set is enough for neural networks modeling because the conclusion depends on each particular case. Hence, in this study, we tried to answer this question by using the Bootstrap method to resample the data.

Finally, we remarked that the overfitting analysis was neglected in many previous researches as shown in the above literature review, we therefore included in this paper a thorough solution for the overfitting problem. The findings of this study have the potential to be applied in real life because they help to improve the accuracy of the mud density's determination, which in turn will improve the safety of the operations.

METHODOLOGY

Mathematical models

Regarding the mathematical models, we initially intended to use a linear function, which is easy to implement, to calculate the drilling mud density based on bottomhole pressure, bottomhole temperature and value of mud density at surface conditions. However, there are some assumptions that we must comply with which can be found in Dahraj & Bhutto, 2014[37] and Molnar, 2021[38]. The input data was collected from the works of McMordie et al., 1982[5] and Demirdal & Cunha, 2009[6], which were summarized in Figure 1.

Figure 1 illustrates the variation of drilling mud density in function of temperature and pressure. The blue graph represents the temperature, the orange graph shows the pressure and the green one describes the value of mud density at surface conditions.

Figure 2 to Figure 4 showed that all the histograms of variables are not bell-shaped. Moreover, we also analyzed the distribution of residuals in Figure 5. We observed that the distribution of residuals was not in shape with the red curve, which presented the normal distribution. Instead, the distribution was likely the fat-tailed distribution, which was not normal distribution, so the linear function was not suitable in this case. Consequently, we had to think about another method, which is the nonlinear function, to deal with the problem. This nonlinear function will also be used later to verify the results given by the artificial neural network modeling.

For the nonlinear function, the quadratic and cubic functions were tested, and we obtained that the correlation coefficient of the cubic function (0.9997) was higher than the one of the quadratic functions (0.9994). In reality, there may be other nonlinear functions with higher correlation coefficients, however, the more complex the function, the higher the risk of overfitting. The cubic function was therefore chosen for this study.

The nonlinear model was constructed by solving the linear least squares problems while using QR factorization which can be referred to the work of Golub & Loan, 1996[39]. The cubic function has the following form:

$$\begin{aligned} \rho = &\left(A\rho_i^2 + B\rho_i^2 + C\rho_i^3\right) + \\ &\left(D \times P^3 + E \times P^2 + F \times P\right) \\ &+ \left(G \times T^3 + H \times T^2 + I \times T\right) + J \end{aligned} \qquad (5)$$

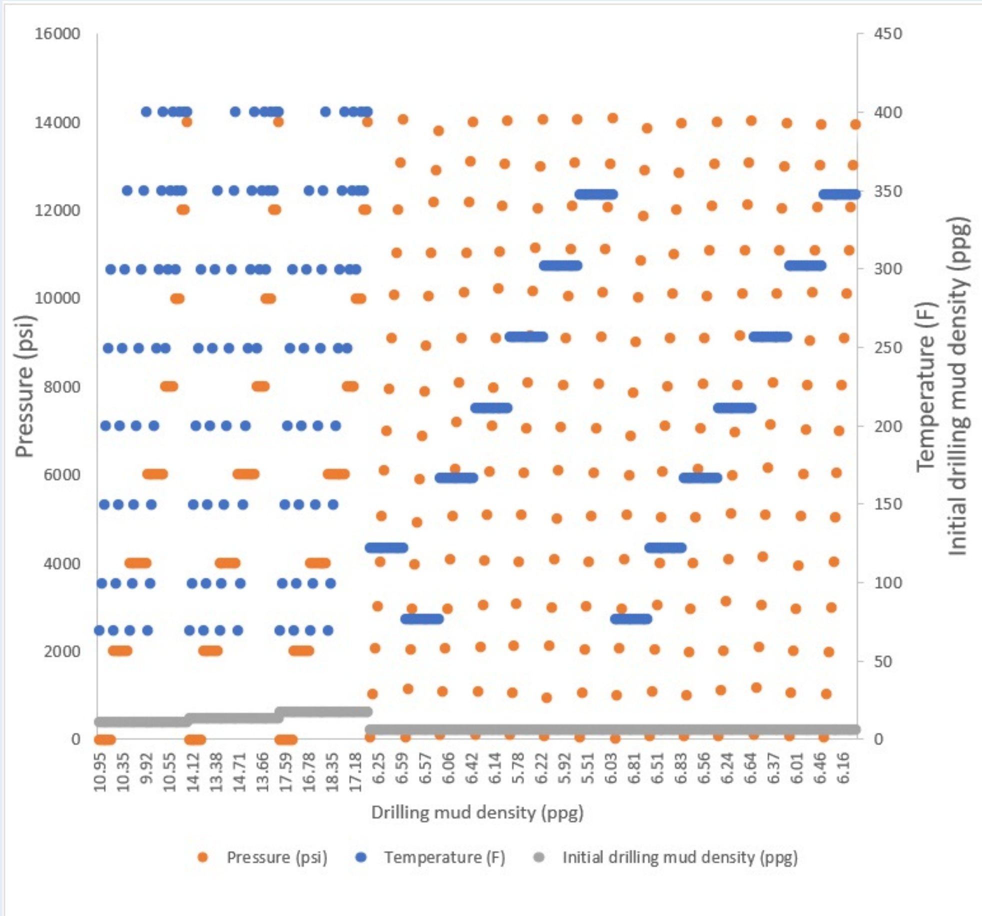

Figure 1: The data collected from the works of McMordie et al. (1982) and Demirdal & Cunha (2009) were used in this research for the nonlinear function

ρ_I (ppg) the value of mud density at surface conditions, P and T are pressure (psi) and temperature (^{o}F) at the location of interest. The values from A to J were determined and listed in the following:

$$
\begin{aligned}
&A = -2.761 \times 10^{-3} \\
&B = 9.753 \times 10^{-2} \\
&C = -5.393 \times 10^{-2} \\
&D = 1.412 \times 10^{-13} \\
&E = -4.372 \times 10^{-9} \\
&F = 8.317 \times 10^{-5} \\
&G = -2.058 \times 10^{-8} \\
&H = 1.532 \times 10^{-5} \\
&I = -6.673 \times 10^{-3} \\
&J = 3.785
\end{aligned}
$$

The nonlinear function presented a high coefficient of determination $R^2 = 0.9994$. Moreover, the value of mean square error was also accepted, with the MSE = 0.00971 for the nonlinear function (the caluclation of MSE is described in detail in the Appendix section). With input taken from Table 1, the calculated values of drilling mud density (ppg) from empirical correlations and nonlinear function are presented in Table 2.

Table 2 showed that results obtained from the empirical correlations are close to results obtained from the nonlinear function. Hence, the nonlinear function can be used as an alternative method to predict the drilling fluid density in function of pressure and temperature. However, these methods do not take into account the influence of other factors such as the circulation rate. Hence, in the next section, an artificial neural network modeling will be presented.

Machine learning model

Overview of artificial neural network

Artificial Neural Network (ANN) is an artificial intelligence information processing system inspired by the

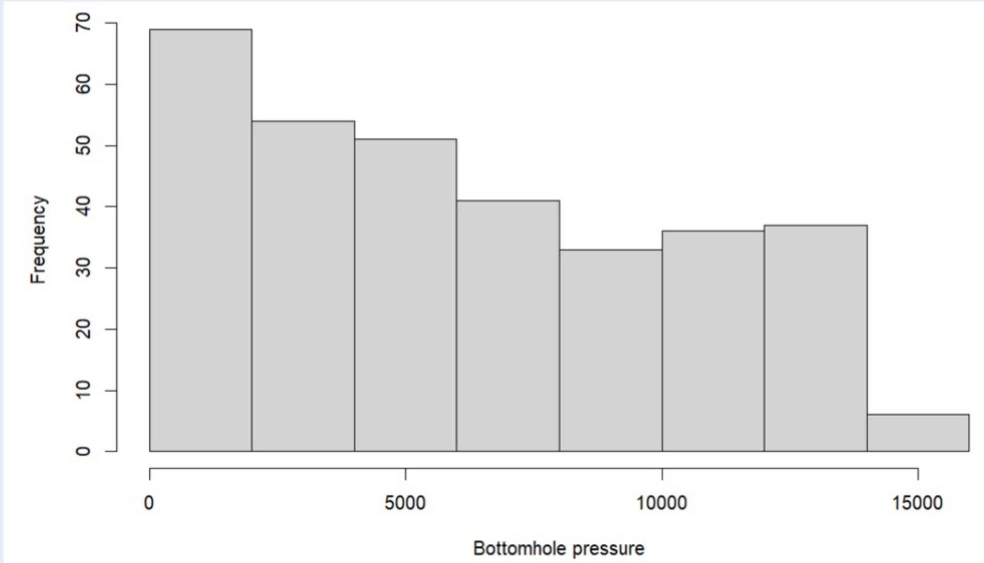

Figure 2: Histogram of bottomhole pressure values (psi)

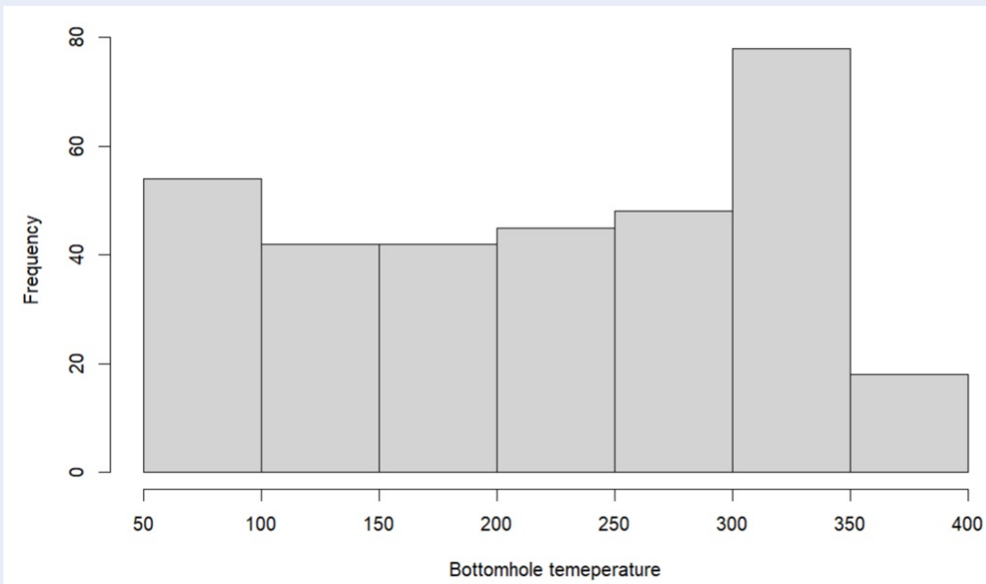

Figure 3: Histogram of bottomhole temperature values (oF)

operation of biological neural networks in the human brain. One of the notable features of artificial neural networks is their limited learning ability.

An artificial neural network usually consists of 3 layers and each layer will have a different number of neurons:

- Input layer: the main function is providing necessary information. A number of neurons in input layer are corresponding to a number of factors and these factors are assumed in the form of vectors
- Hidden layers contain hidden neurons helping the inputs connect and outputs. A neural network may have one or multiple hidden layers, and in some cases, there is no hidden layer.
- Output layer includes the neurons which hold output information. A neural network can have many output factors.

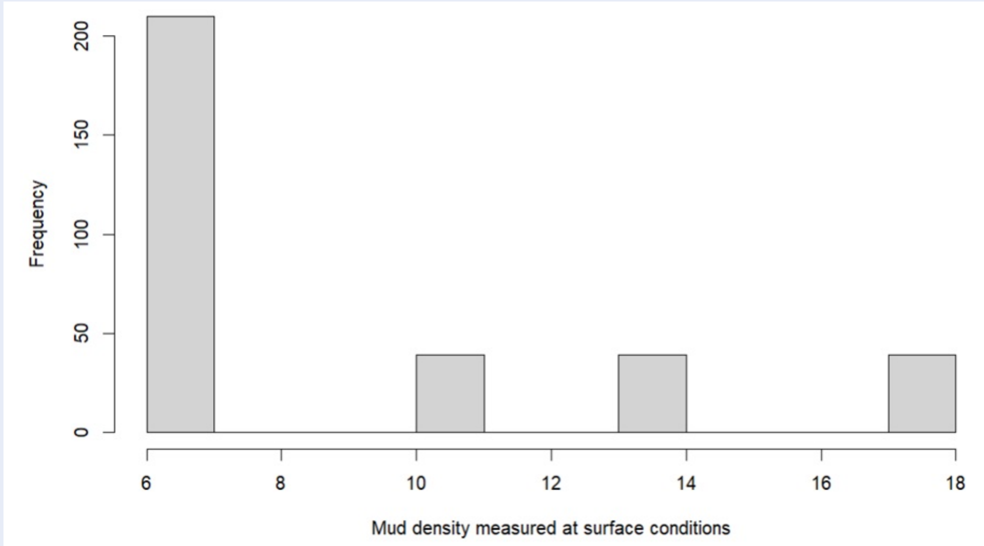

Figure 4: Histogram of mud density at surface conditions (ppg)

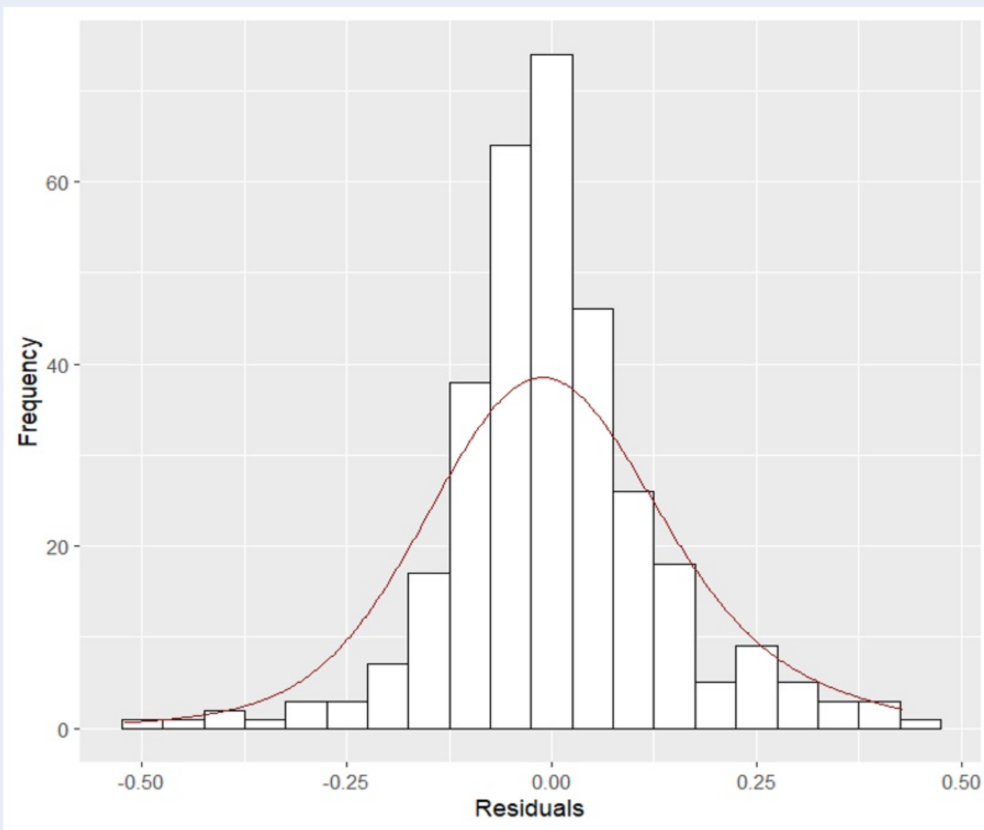

Figure 5: Histogram of residuals

Table 1: The input data which was used in this study for empirical correlations and nonlinear function

Bottomhole pressure (psi)	Bottomhole temperature (^{o}F)	Circulation rate (gal/min)	Oil volume fraction	Water volume fraction	Oil density (ppg)	Water density (ppg)
8562.279	128.55	0	0.67711	0.17524	6.6619	8.3641
8671.2	128.63	0	0.67715	0.17523	6.6603	8.3632
8942.688	127.15	0	0.67684	0.17532	6.6735	8.3717
8945.111	90.74	126.8	0.67572	0.17559	6.7226	8.4065
8945.056	90.28	126.8	0.67574	0.17558	6.7217	8.4059
8946.967	90.07	126.8	0.67574	0.17558	6.7217	8.4059
8944.121	90.02	126.8	0.67574	0.17558	6.7216	8.4059
8945.887	89.91	126.9	0.67574	0.17558	6.7217	8.4059

Table 2: Drilling mud density (ppg) obtained from empirical correlations and the nonlinear functions using input data in Table 1

Furbish	Hoberock	Kutasov	Sorelle	Nonlinear function
10.8347	11.0440	10.8599	12.63049	10.8949
10.8378	11.0416	10.8633	12.63005	10.8989
10.8501	11.0613	10.8771	12.63331	10.9150
10.9500	11.1354	10.9852	12.64592	11.0635
10.9513	11.1341	10.9865	12.64577	11.0655
10.9519	11.1341	10.9871	12.64577	11.0665
10.9520	11.1340	10.9872	12.64577	11.0666
10.9523	11.1341	10.9876	12.64577	11.0671

Determining the number of hidden layers and the number of neurons is a relatively complex task, there is no rule that finds out the optimal number of hidden layers and hidden neurons. The method of selecting the number of neurons and layers is a trial-and-error approach. The connections between neurons in different layers contain their own individual weights. The number of weights depends on network configuration.

The general relationship between the input data and output data is described below:

$$y_k = f_o\left[\sum_j w_{kj} \times f_h\left(\sum_j w_{ji} x_i + b_j\right) + b_k\right] \quad (6)$$

x_i is an input vector, w_{ji} denotes the connection weight from the *i*th neuron in the input layer to the *j*th neuron in the hidden layer, b_j represents the threshold value or bias of *j*th hidden neuron, w_{kj} stands for the connection weight from the *j*th neuron in the hidden layer to the *k*th neuron in the output layer, b_k refers to the bias of the *k*th output neuron, f_h and f_o are the activation functions for the hidden and output neuron, respectively.

The Transfer Function is responsible for transforming the input variable into a different range of values. Some commonly used transfer functions include the logistic sigmoid function, the tangent sigmoid function, and the linear function. Each type of function used has a different purpose for each layer and different types of problems. Nonlinear functions are often used for pattern recognition and discrimination problems and are typically used in the hidden layer. The linear function is used in matching and prediction problems and is usually used in the output layer. This study only covers basic knowledge of machine learning, and readers can refer to additional sources for more information, such as Ghaffari et al., 2006 [40], F. Parrella, 2007 [41], and Mohaghegh, 2000 [42].

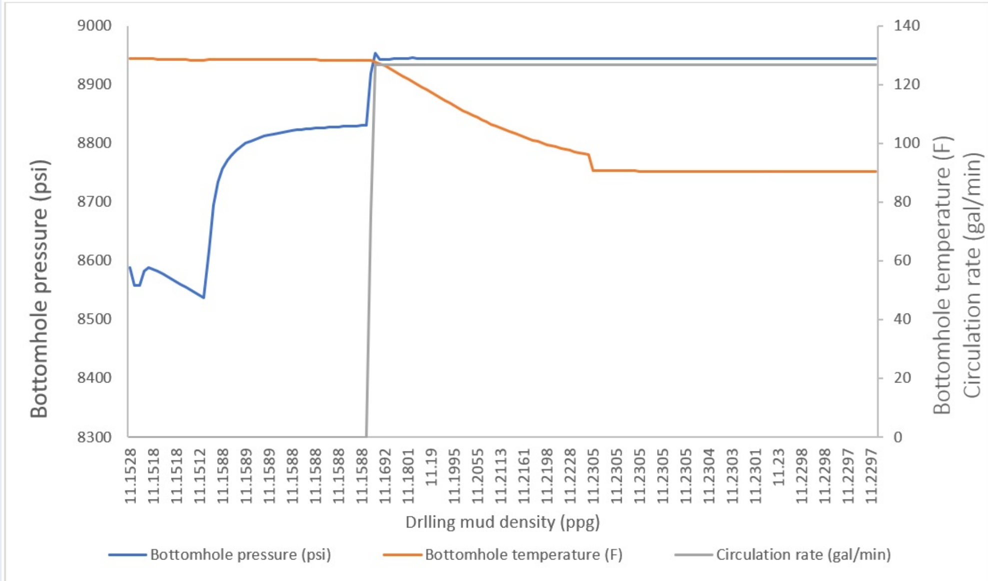

Figure 6: Input data used in this study to build the artificial neural network was provided by Schlumberger in Drillbench software's tutorial

Input data for ANN modeling

Input data, which was used to build and calibrate ANN in this study, covered 162 positions of a well at different conditions. The value of drilling mud density at standard conditions is 10.7656 ppg. The data can be viewed in Figure 6.

Artificial neural networks optimization before analysis of overfitting

According to Kårstad & Aadnøy, 1998 [26] and Harris & Osisanya, 2005 [27], the circulation rate also had an effect on drilling mud density. With the desire to contribute a small part to research on predicting drilling mud density, our paper would like to introduce an artificial neural network for predicting drilling fluid density in the function of 4 input factors: surface drilling fluid density, bottom hole pressure, bottom hole temperature, and circulation rate. In the first hidden layer, the transfer function used is the logistic sigmoid function. In the second hidden layer, the transfer function used is the tangent sigmoid function, and in the output layer, the transfer function used is the linear function. This paper used the trial-and-error method to build the networks. Each network structure was run 10 times to avoid random distribution and was selected based on the smallest mean square error in the 10 training runs. In Table 3, with the lowest mean square error MSE , the optimized network consisted of 4 neurons in the input layer, 6 neurons in the first hidden layer, 10 neurons in the second hidden layer, and 1 neuron in the output layer (Figure 7).

However, the solution is not as simple as it seems. We remarked here that the MSE values were anomaly small, which manifested the overfitting problem. Hence, the model can not be used in real life. Therefore, in the following section, we will solve the overfitting problem.

Solving the overfitting problem and optimizing the artificial neural network.

a. Data pre-processing

The authors knew that the data sets are very important in ANN, that's why we tried to collect as much data as possible. In this research, we had 327 observations for the non-linear analysis and 162 data for the neural network modeling. Understanding the number of data might be low, hence we referred Horowitzto's paper in 2008 [43] and conducted the Bootstrap method to resample the data set and obtained a new one with the same statistical characteristics for 400 data points. After that, we divided the data into training set, validation set and test set with proportions of 70%, 15%, 15%, respectively, and used the same ANN models for both original and Bootstrap datasets.

For the targets in neural network training, we used the difference between the density of initial drilling mud and density at bottomhole condition. The input and

Table 3: Mean square errors of different networks

Layer 1 Layer 2	1 neuron	2 neurons	3 neurons	4 neurons	5 neurons	6 neurons	7 neurons	8 neurons	9 neurons	10 neurons
1 neuron	3.87E-7	1.6E-7	1.65E-7	2.18E-7	1.16E-7	1.28E-7	1.15E-7	1.23E-07	1.30E-07	1.59E-07
2 neurons	4.63E-7	1.99E-7	1.92E-7	2.63E-7	1.4E-7	1.24E-7	6.31E-8	1.03E-07	1.51E-07	1.21E-07
3 neurons	4.62E-7	4.58E-7	1.46E-7	2.79E-7	1.39E-7	1.13E-7	2.17E-7	1.32E-07	1.21E-07	1.12E-07
4 neurons	2.76E-7	2.15E-7	2.32E-7	8.48E-8	1.08E-7	1.62E-7	1.35E-7	1.40E-07	1.27E-07	1.26E-07
5 neurons	3.94E-7	3.08E-7	7.83E-8	1.14E-7	1.07E-7	1.1E-7	1.19E-7	1.30E-07	1.08E-07	1.26E-07
6 neurons	2.51E-7	2.98E-7	1.09E-7	1.67E-7	1.38E-7	6.11E-8	8.13E-8	1.30E-07	1.20E-07	1.11E-07
7 neurons	1.57E-7	1.14E-7	2.34E-7	1.36E-7	1.75E-7	9.06E-8	8.95E-8	1.21E-07	1.05E-07	1.16E-07
8 neurons	1.47E-7	1.03E-7	1.47E-7	7.69E-8	1.08E-7	1.39E-7	7.03E-8	1.41E-07	1.12E-07	1.04E-07
9 neurons	9.06E-6	1.3E-7	1.12E-7	1.18E-7	1E-7	1.77E-7	6.02E-6	1.37E-07	1.01E-07	1.17E-07
10 neurons	1.99E-7	8.9E-8	1.25E-7	1.08E-7	1.13E-7	5.43E-8	1.22E-7	1.16E-07	1.01E-07	1.18E-07

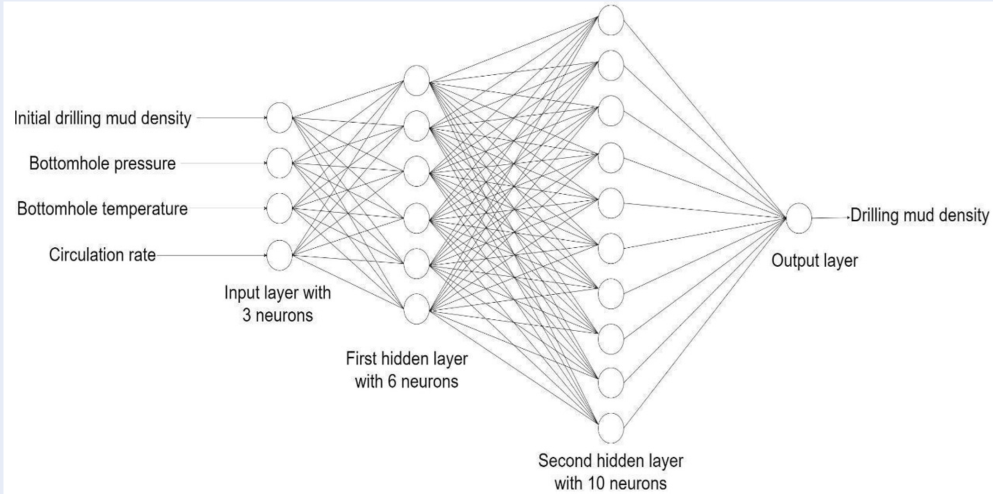

Figure 7: The architecture of the optimized artificial neural network resulted from this study before analysis of overfitting

target data were normalized as shown in the following formulas:

$$x_j = \frac{x_i - x_i^{min}}{x_i^{max} - x_i^{min}} \tag{7}$$

$$y_j = \log(1 + y_i) \tag{8}$$

x_j is a dimensionless value of input data, x_i is a true value of input data, y_j is a dimensionless value of target data and y_i is a true value of target data.

b. Artificial neural networks modeling using Bootstrap data added to original data

Using both original and Bootstrap datasets for the same neural network (3-6-10-1), we observed that the overfitting was decreased (Figure 8) because the MSE values of the test set and validation set were similar. However, in Figure 9, we observed that although the R values obtained from using Bootstrap data was reasonably high (the overfitting problem did not occur), but we also observed that the regression graphs were anormal: many output values fluctuated only around the value of 0.36, which is unusual because in reality the values of target data were more varied. In conclusion, the utilization of Bootstrap method could reduce the overfitting problem, but did not provide satisfactory results. Consequently, in the next section, we will use only original data.

c. Artificial neural networks modeling using only original data with analysis of overfitting

After realizing that the Bootstrap method did not improve the results, the author went back to the normalized original datasets. We then used the same neural network (3-6-10-1), and overfitting was observed in the results: firstly, because the R values were abnormally high (Figure 10); secondly, the MSE of testing set is larger than the one of the training set (Figure 11). Therefore, we trained different models which consisted of two hidden layers, and the number of neurons varied from 1 to 10 for each hidden layer. However, the overfitting still existed, so we had to go back to the model with one hidden layer. The results in Figure 12 showed the validation and test curves were very similar, and the MSE of the test set and of the validation set were lower than the one of the train set, which indicated that the overfitting had been excluded. Figure 13 showed that R values and the regression graphs were reasonable without abnormal distribution. In literature, the research of Okorie E. Agwu et al., 2020 [16] possibly had an overfitting problem with very high R value and the predicted values were exactly the same as experimental values. The thorough analysis of overfitting in our research helped to avoid this same problem.

In brief, the results indicated that the optimized network with the best performance without encountering overfitting consisted of one hidden layer with 5 neurons, and the transfer function was tangent sigmoid.

The results in previous sections showed that the number of input data is not a problem for ANN modeling as we were afraid at first. There is no simple answer to the question if a data set is enough for neural networks modeling. It really depends on each particular case. The 327 observations for the non-linear

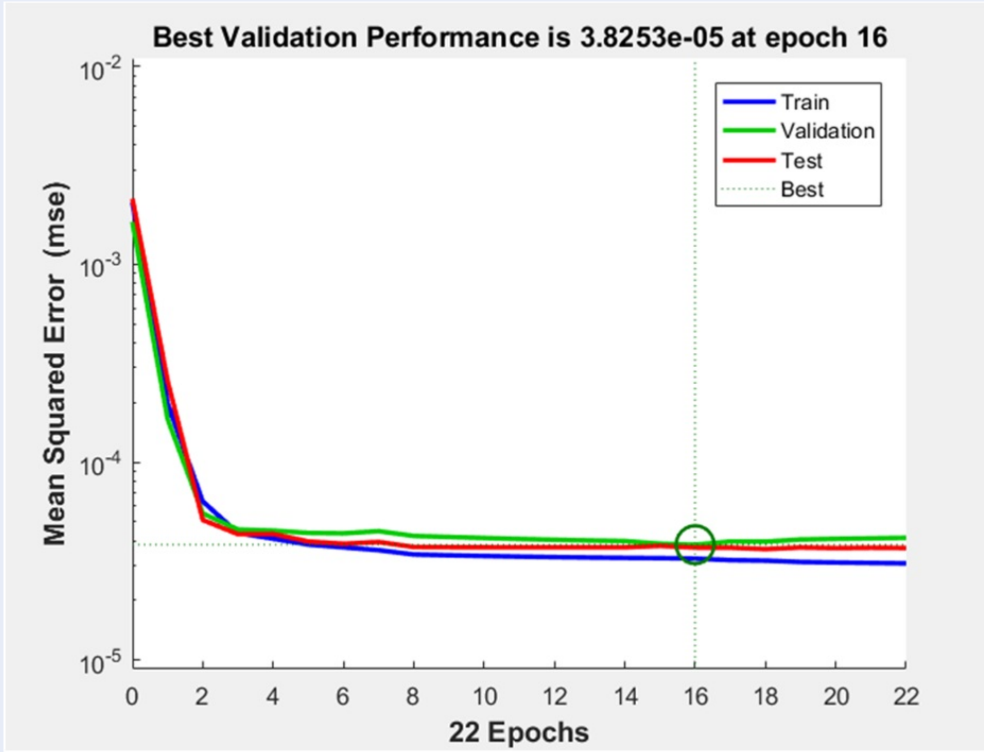

Figure 8: Graph shows the MSE values when using Bootstrap datasets to train the same ANN (3-6-10-1)

analysis and 162 data for the neural network modeling used in this research are therefore enough for a proper analysis. The results showed that we must choose the right neural network model with optimized layers and nodes to have a high accuracy without encountering an overfitting problem. For this case, using one hidden layer is optimized for the ANN modeling. This can be explained by the fact that the more complicated a neural network is, the more data it requires in order to not be overfitted (Muhammad Uzair and Noreen Jamil 2020[44]). Hence, in this study, with our available data, the number of hidden layers must be one, so that no overfitting can occur.

RESULTS AND DISCUSSION

Since the authors wanted to present various models to predict drilling mud density, a generalized additive model (GAM) was built based on the input data in Figure 6 and evaluated using the same data from Table 1. A generalized additive model is a generalized linear model with a linear predictor involving a sum of smooth functions of covariates (Hastie and Tibshirani 1990[45]). The GAMs can model non-Gaussian outcome variables, in terms of several predictor variables. The requirement of the generalized linear models that the relationships between the outcome and the predictors be linear was relinquished by Vanhove, 2014[46]. Instead, non-linear relationships can also be modeled with the form estimated from the data. This can be accomplished by fitting higher-order polynomial regressions on subsets of the data and adding the pieces together. The more subset regressions are fitted and connected together, the more wiggly the overall curve will be. Fitting too many subset regressions results in overwiggly curves that fit disproportionally much noise in the data ('oversmoothing'). In order to prevent this, the algorithm can be furnished with a cross-validation procedure or a generalized (algebraic) approximation (Wood, 2006[47]).

Whereas the additive model was estimated by penalized least squares, the GAM will be fitted by penalized likelihood maximization, and in practice this will be achieved by penalized iterative least squares. More specific details can be viewed in the paper of Wood, 2006[47]; Zuur et al., 2009[48]; Vanhove, 2014[46]. Table 4 will show the specific results of drilling mud density obtained from the generalized additive model.

To confirm the effect of circulation rate on the mud density and prove that the network obtained from this study can be applied, the results of drilling mud density obtained from ANN model and generalized additive model were compared with the results from the

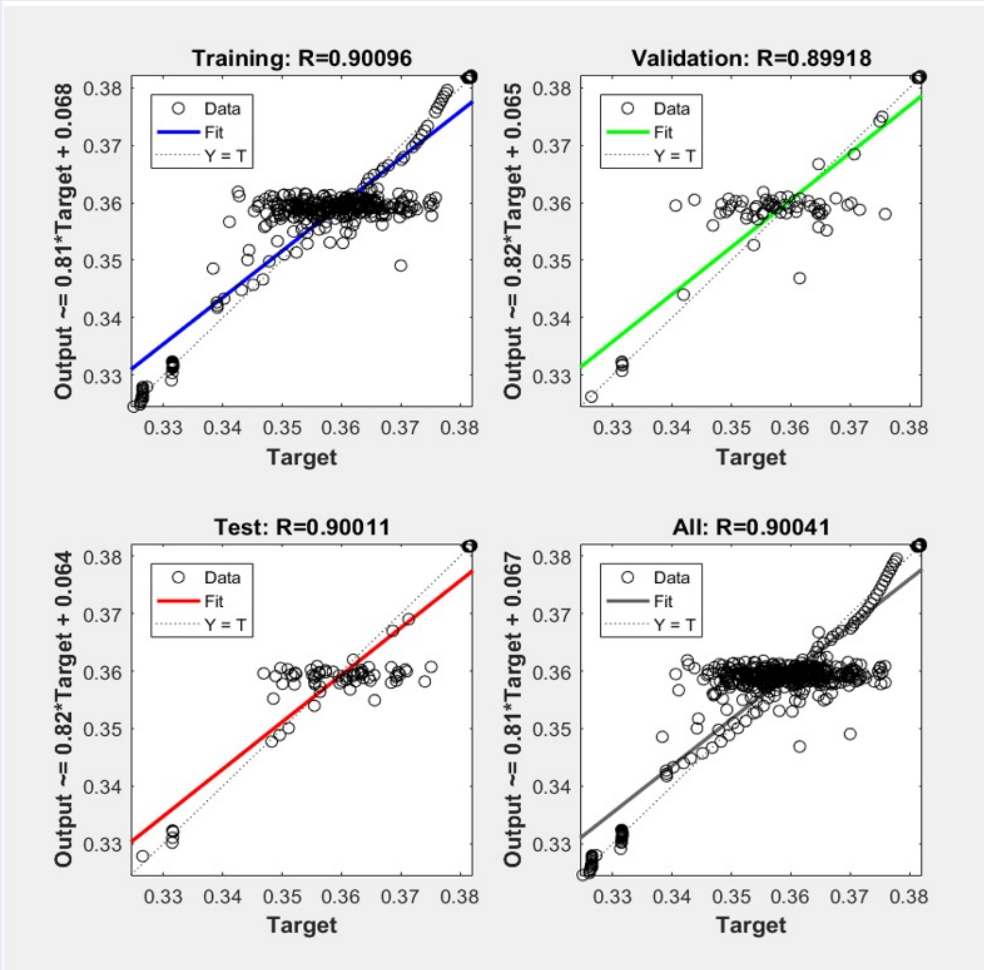

Figure 9: Graph shows the R values when using Bootstrap datasets to train the same artificial neural network

ANN model of Okorie E. Agwu et al., 2020 [16] in Table 4.

The determination coefficient between the results obtained from our ANN model and input data is 0.9972, which is rather similar to the determination coefficient (0.9970) obtained from the ANN model of Okorie E. Agwu et al., 2020 [16]. However, the mean square error of our network (0.01321) is lower than Okorie E. Agwu's (0.04754). In addition, as mentioned above, the overfitting problem was included in our analysis, which was not done in Agwu et al., 2020 [16]. We concluded that our ANN model provided a high value of coefficient of determination without encountering the overfitting problem. Moreover, the determination coefficient given by the generalized additive model is high (R^2 = 0.99865) while the mean square error is low (3.65×10^{-6}). Hence, our ANN model and generalized additive model can be used in real life applications.

Eventually, Table 6 showed that almost all of the methods were reliable. Only the calculated results given by Sorelle et al., 1982 [36] gave a significant deviation compared to the input data, hence using the model of Sorelle is not highly recommended. Although the determination coefficient of our ANN model is lower than the one given by the generalized additive model, the ANN method can still be accepted because of its small mean square error (Table 5), and because it can include more influence factors in the input data than the other methods.

Figure 14 shows the predicted results obtained from different methods that were used in this study. The measured data in Figure 6 were the same data as the input data used in ANN modeling. Figure 14 allowed us to draw the same conclusions as mentioned in the previous paragraph.

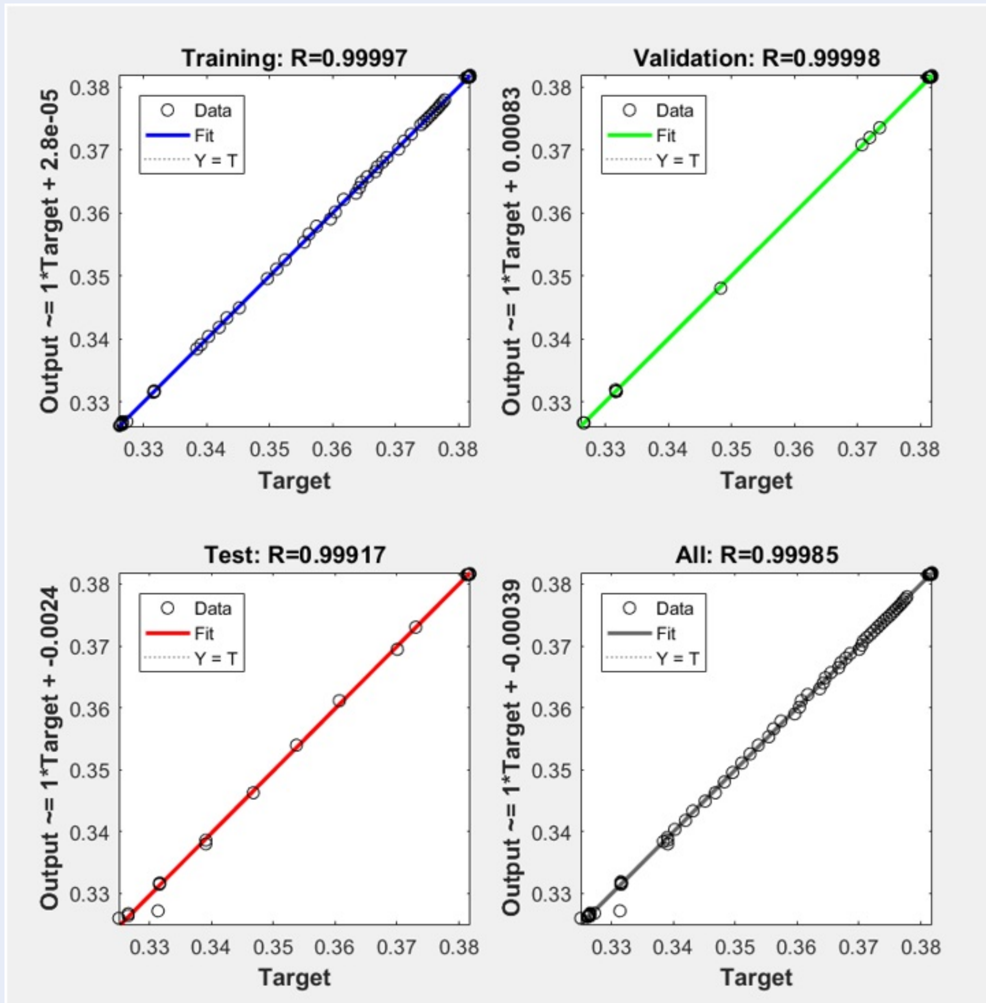

Figure 10: Graph shows the R values when applying the original normalized datasets to the same artificial neural network

Table 4: Drilling mud density (ppg) obtained from different artificial neural networks and the generalized additive model using the same input data in Table 1

The optimized ANN obtained from this study	Generalized additive model	ANN model of Okorie E. Agwu et al. (2020)
11.2358	11.1515	10.9003
11.2439	11.1603	10.9044
11.2536	11.1732	10.9199
11.3584	11.2300	11.0323
11.3587	11.2302	11.0337
11.3594	11.2305	11.0345
11.3585	11.2303	11.0345
11.3592	11.2305	11.0349

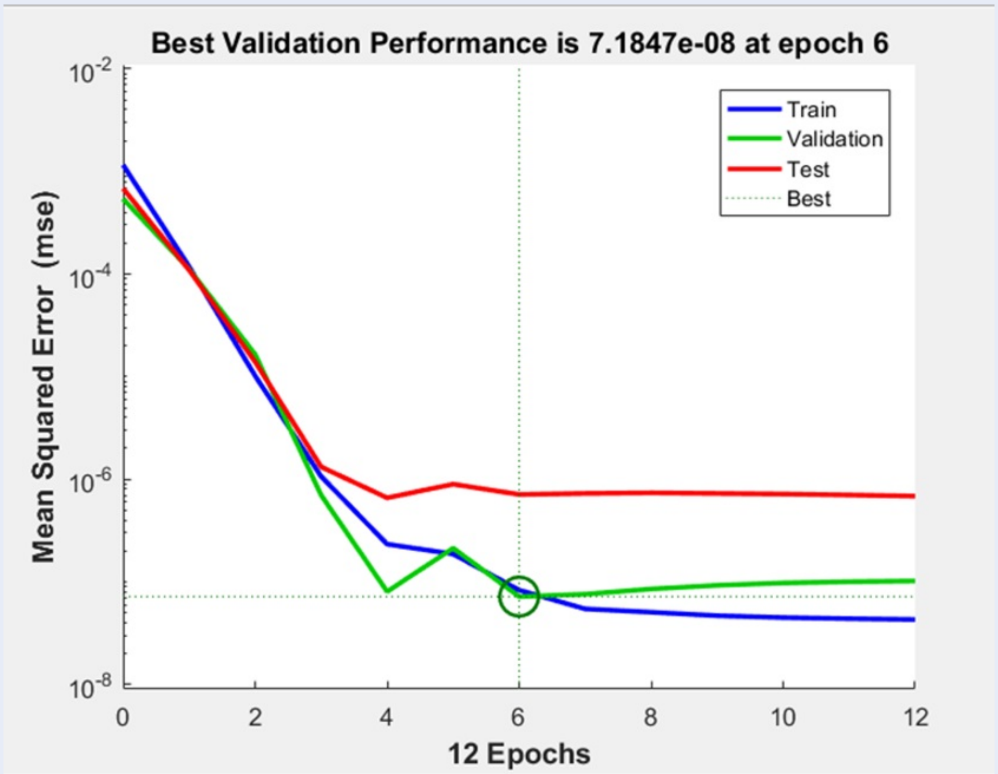

Figure 11: Graph shows the MSE values when applying the original normalized datasets to the same artificial neural network

Table 5: The results of drilling mud density (ppg) obtained from the optimized ANN, generalized additive model, and empirical correlations for the same input data from Table 1

Nonlinear function	Generalized additive model	Furbish	Hoberock	Kutasov	Sorelle	The optimized ANN obtained from this study
10.8949	11.1515	10.8347	11.0440	10.8599	12.63049	11.2358
10.8989	11.1603	10.8378	11.0416	10.8633	12.63005	11.2439
10.9150	11.1732	10.8501	11.0613	10.8771	12.63331	11.2536
11.0635	11.2300	10.9500	11.1354	10.9852	12.64592	11.3584
11.0655	11.2302	10.9513	11.1341	10.9865	12.64577	11.3587
11.0665	11.2305	10.9519	11.1341	10.9871	12.64577	11.3594
11.0666	11.2303	10.9520	11.1340	10.9872	12.64577	11.3585
11.0671	11.2305	10.9523	11.1341	10.9876	12.64577	11.3592

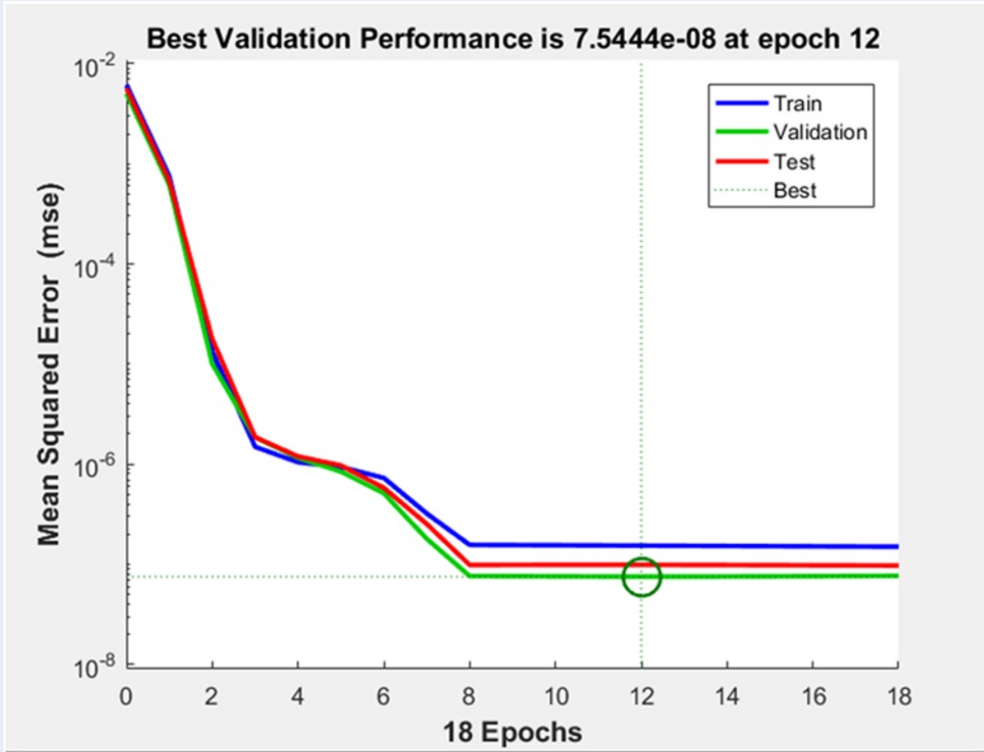

Figure 12: Graph shows the MSE values when applying the original normalized datasets to the artificial neural network with one hidden layer

Table 6: **Comparison of correlation coefficients and errors given by different methods**

Statistical parameters	MSE	RMSE
Methods		
The optimized ANN obtained from this study	0.01321	0.1149
ANN model of Okorie E. Agwu et al. (2020)	0.04754	0.2180
Furbish	0.08631	0.2938
Hoberock	0.01023	0.1011
Nonlinear function	0.04083	0.2021
Kutasov	0.06892	0.2625
Sorell	2.06639	1.4375
Generalized additive model	3.65E-06	0.0019

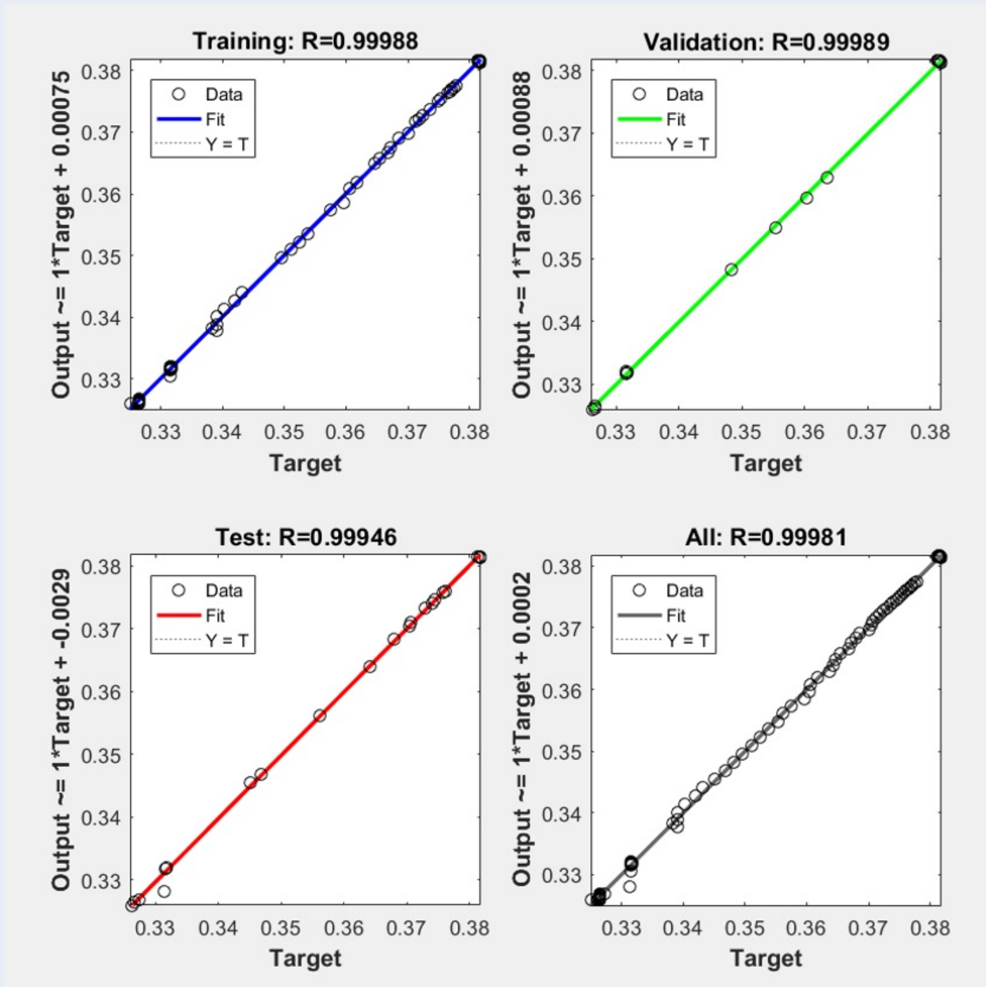

Figure 13: Graph shows the R values when applying the original normalized datasets to the artificial neural network with one hidden layer

Study of relative importance of different input parameters

Analyzing the impact of the three factors (pressure, temperature, and surface density) using the nonlinear mathematical model

To analyze the impact of the three factors, which are pressure, temperature, and surface density, input data in Figure 1 was used for the evaluation with help of the Equation (5). The effect of these three factors are illustrated in Figure 15. We observed that if the value of mud density at surface conditions is reduced to 60%, the drilling mud density at the wellbore conditions will decrease to approximately 55%. Another remark is that if the bottomhole temperature is reduced to 60%, the drilling mud density will increase by approximately 1.05 times. Besides, if the bottomhole pressure is reduced to 60%, the drilling fluid density will be 0.95 compared to the initial value.

These above observations are similar to the ones discussed in the works of Agwu et al. 2020 [16] and Osman et al., 2003 [14]. Both of these two papers concluded that surface density had the biggest impact, followed by bottomhole temperature and bottomhole pressure.

Analyzing the impact of the four factors (pressure, temperature, surface density, and circulation rate) using generalized additive model

Since the value of mud density at surface conditions is constant during the operation, it may not be wise to include it in the study. Therefore, instead of considering the impact of surface drilling fluid density on the bottomhole drilling mud density, we evaluated

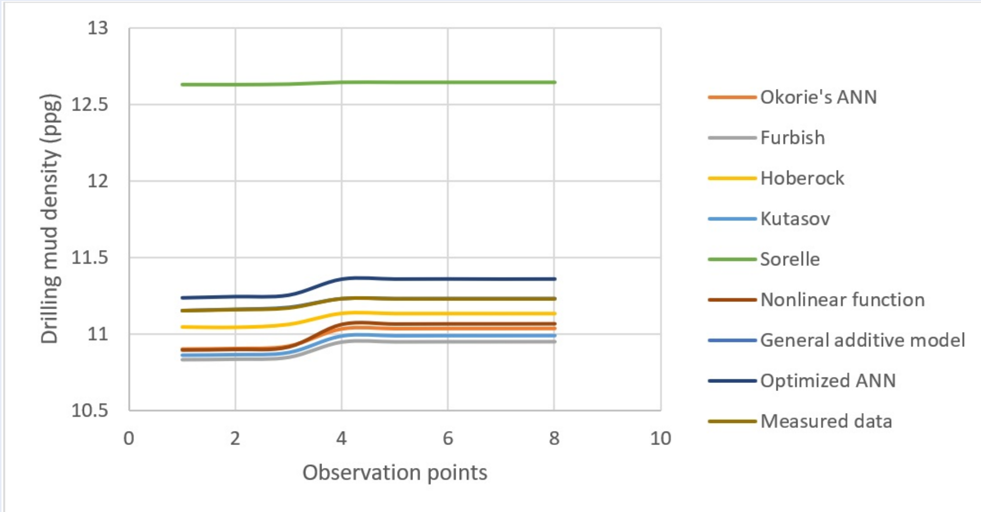

Figure 14: Graph shows results of drilling mud density (ppg) obtained from empirical correlations, nonlinear function, generalized additive model, and machine learning models

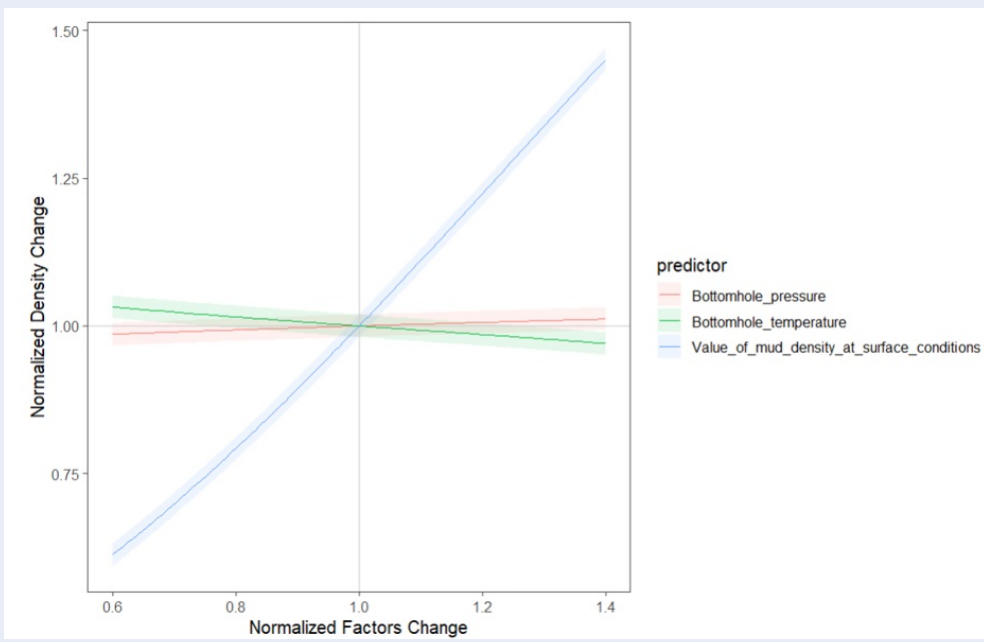

Figure 15: Relative importance of bottomhole pressure, bottomhole temperature, and the value of mud density at surface conditions to the drilling mud density

another factor which was the circulation rate. Harris and Osisanys, 2005[27] mentioned that the circulation rate was proportional to the drilling fluid density at the bottomhole condition because higher flow rates would cause the bottomhole pressure to increase and the bottomhole temperature to decrease. Besides that, no study about the influence of circulation rate has ever been realized so far.

Our generalized additive model was used to study the level of variables' importance with help of the data presented in Figure 6. Figure 16 showed that the effect of the circulation rate on the drilling mud density was quite low. Combined with the results shown in the previous section, it can be concluded that the level of influence of different factors on the drilling fluid density is in the following order: value of mud density at surface conditions, bottomhole pressure, bottomhole temperature, and circulation rate.

CONCLUSIONS

This paper presented various methods (artificial neural network, generalized additive model, nonlinear function, empirical correlations) to predict drilling mud density in function of temperature, pressure, surface value of the drilling fluid, and circulation rate. The results lead to the following conclusions:

- The Generalized Additive model and Artificial Neural Network have higher coefficient of determination R^2 and lower MSE than the other methods. However, it is recommended to use our optimized ANN method because we demonstrated that it did not have a problem of overfitting, while the Generalized Additive model presented a very low MSE, which should be used with caution.
- The optimized ANN model consisted of only one hidden layer. In addition, the answer to the question if a data set is enough for neural networks modeling is not simple because it depends on each particular case. In this study, the Bootstrap method was used to resample the data and the conclusion was that the number of input data was enough to avoid the overfitting problem. Moreover, it is worthy to note that since there was often a lack of overfitting analysis in previous studies in literature review regarding this specific case, we solved this problem by conducting a thorough analysis of overfitting in this paper.
- The nonlinear model is more appropriate than the linear model in this case based on the analysis of the histograms of different variables.
- The empirical correlations presented higher deviation between predicted results and measured data, especially the correlation given by Sorelle et al. (1982).
- The level of impact on drilling mud density is in the following order: value of mud density at surface conditions, bottomhole pressure, bottom hole temperature, and circulation rate.

ABBREVIATIONS

ANN: Artificial neural network

f_0: Percentage of oil volume in the drilling fluid

f_w: Percentage of water volume in the drilling fluid

GAM: Generalized additive model

MSE: Mean Squared Error

P_0 (psi): Standard pressure

P, P_2 (psi): Pressure at the predicted position

RMSE: Root Mean Squared Error

T_0 (^{o}F): Standard temperature

T, T_2 (^{o}F): Temperature at the predicted position

V (gal): Total volume

$\triangle V_0$ (gal): Difference in oil volume

$\triangle V_w$ (gal): Difference in water volume

x_i: A true value of input data

x_i^{max}: A maximum value of input data

x_i^{min}: A minimum value of input data

x_j: A dimensionless value of input data

y_i: A true value of target data

y_j: A dimensionless value of target data

$\rho_i, \rho_{mo}, \rho_1$ (ppg): Value of mud density at surface conditions

ρ, ρ_f, ρ_m (ppg): Predicted drilling mud density

ρ_{o1} (ppg): Initial oil density

ρ_{o2} (ppg): Oil density in predicted drilling mud

ρ_{w1} (ppg): Initial water density

ρ_{w2} (ppg): Water density in predicted drilling mud

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest in the publication of this article.

AUTHORS' CONTRIBUTION

Pham Son Tung directed and supervised the development and completion of the research, as well as reviewed and revised the article.

Pham Thanh Nha collected data, built the models, and drafted the manuscript.

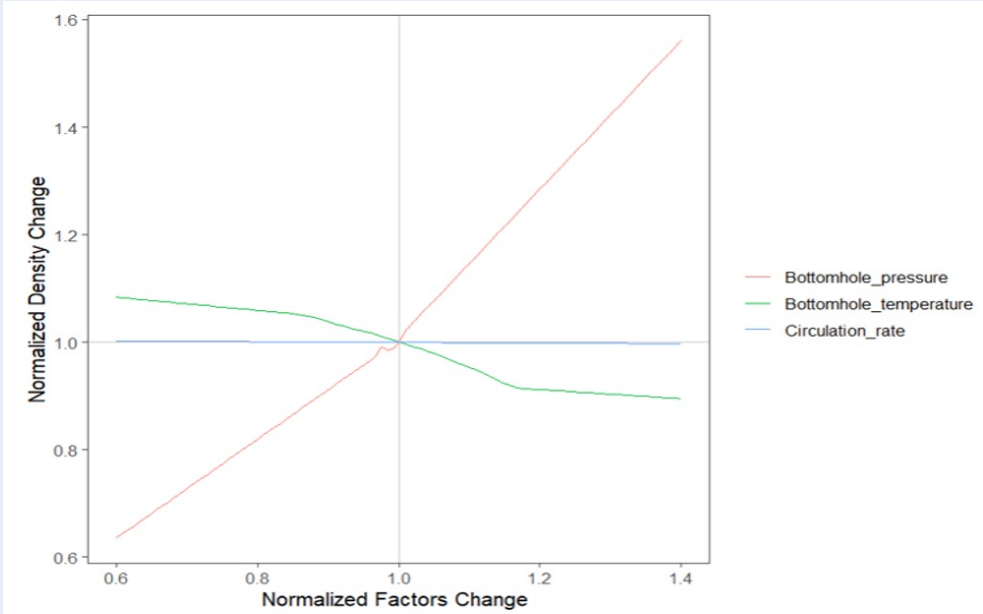

Figure 16: Relative importance of bottomhole pressure, bottomhole temperature, and circulation rate to the drilling mud density

APPENDIX

Mean Squared Error (MSE) is a formula for estimating the squared value of an error. The smaller the value of MSE, the more accurate the prediction is.

$$MSE = \frac{1}{N} * \sum_{i=1}^{N} (X_i^* - X_i)^2$$

Root Mean Square Error (RMSE) is used to evaluate how well a model fits the data. When the value of RMSE is near 0, the model will be more accurate.

$$RMSE = \left[\frac{\sum_{i=1}^{N} \left(X_i^* - X_i\right)^2}{N} \right]^{\frac{1}{2}}$$

T-value is a measure that indicates the degree of influence of input factors on the results. The absolute value of the t-value indicates the greater the degree of influence. A negative t-value indicates an inverse relationship between the input factor and the result, and vice versa.

The correlation coefficient is a statistical parameter that measures the degree of fit between predicted and actual data of drilling fluid density.

$$R^2 = 1 - \frac{\sum_{i=1}^{N} \left(X_i^* - X_i\right)^2}{\sum_{i=1}^{N} \left(X_i^* - \frac{1}{N} \sum_{i=1}^{N} X_i\right)^2}$$

N is the total number of observations, I is the index of I observation; X_i* is the value of drilling mud density which is predicted from empirical correlations or machine learning models.

Pr (>|t|) is the p-value corresponding to the t-value. If the p-value is less than the statistical significance level α (usually 0.05), the factors associated with it will be statistically significant in the results, otherwise, it will be a random factor.

REFERENCES

1. Cormack D. An Introduction to Well Control Calculations for Drilling Operations. New York, USA; 2017;Available from: https://doi.org/10.1007/978-3-319-63190-5.
2. Babu DR. Effects of P-ρ-T behavior of muds on static pressures during deep well drilling — part 2: static pressures. In: Paper (SPE27419) SPE drilling and completion, 1996; vol 11, no 2; 1996;Available from: https://doi.org/10.2118/27419-PA.
3. Hussein AMO, Amin RAM. Density measurement of vegetable and mineral based oil used in drilling fluids. In: Paper SPE 136974 presented at the 34th annual SPE international conference and exhibition held in Tinapa — Calabar, Nigeria;Available from: http://dx.doi.org/10.2118/136974-MS.
4. An J, Lee K, Choe J. Well control simulation model of oil based muds for HPHT wells. In: Paper SPE 176093 presented at the SPE/IATMI Asia Pacific oil and gas conference and exhibition held in Nusa Dua, Bali, Indonesia; 2015;Available from: https://doi.org/10.2118/176093-MS.
5. McMordie Jr WC, Bland RG, Hauser JM. Effect of temperature and pressure on the density of drilling fluids. In: Paper SPE 11114 presented at the 57th annual Fall Technical Conference and Exhibition of the Society of Petroleum engineers of AIME held in New Orleans; 1982;Available from: http://dx.doi.org/10.2118/11114-MS.

6. Demirdal B and Cunha JC. Olefin-based synthetic-drilling-fluids volumetric behavior under downhole conditions. SPE Drilling and Completion 24(2): 239–248; 2009;Available from: https://doi.org/10.2118/108159-PA.
7. Zamora M, Roy D, Slater K, Troncoso J. Study on the Volumetric Behavior of Base Oils, Brines, and Drilling Fluids Under Extreme Temperatures and Pressures. SPE Drilling & Completion 28 (3): 278–288; 2013;.
8. Kemp NP, Thomas DC, Atkinso G, Atkinson BL. Density Modeling for Brines as a Function of Composition, Temperature, and Pressure. SPE Prod Eng 4: 394–400; 1989;Available from: https://doi.org/10.2118/16079-PA.
9. Peters EJ, Chenevert ME, Zhang C. A model for predicting the density of oil-based muds at high pressures and temperatures. In: Paper SPE-18036-PA, SPE Drilling Engineering, vol 5, no 2; 1990;Available from: https://doi.org/10.2118/18036-PA.
10. Isambourg P, Anfinsen BT, Marken C. Volumetric Behavior of Drilling Muds at High Pressure and High Temperature. Paper presented at the European Petroleum Conference, Milan, Italy; 1996;Available from: https://doi.org/10.2118/36830-MS.
11. Zamora M, Broussard PN, Stephens MP. The top 10 mud related concerns in deepwater drilling operations. In: Paper SPE 59019 presented at the SPE international petroleum conference and exhibition in Mexico, Villahermosa, Mexico; 2000;Available from: https://doi.org/10.2118/59019-MS.
12. Hemphill T, Isambourg P. New Model Predicts Oil, Synthetic Mud Densities. Oil & Gas Journal 103 (16): 56–58; 2005;.
13. Peng Q, Fan H, Zhou H, Liu J, Kang B, Jiang W, Gao Y, Fu S. Drilling fluid density calculation model at high temperature high pressure. In: Paper OTC-26620-MS presented at the offshore technology conference Asia, Kuala Lumpur, Malaysia; 2016;Available from: https://doi.org/10.4043/26620-MS.
14. Osman EA, Aggour MA. Determination of Drilling Mud Density Change with Pressure and Temperature Made Simple and Accurate by ANN. In: Paper SPE 81422 presented at the SPE 13th middle east oil show and conference, Bahrain; 2003;Available from: https://doi.org/10.2118/81422-MS.
15. Adesina FAS, Abiodun A, Anthony A, Olugbenga F. Modelling the effect of temperature on environmentally safe oil based drilling mud using artificial neural network algorithm. Pet Coal J 57(1): 60–70; 2015;.
16. Agwu Okorie E., Akpabio JU, Dosunmu A. Artificial neural network model for predicting the density of oil-based muds in high-temperature, high-pressure wells. J Petrol Explor Prod Technol 10: 1081–1095; 2020;Available from: https://doi.org/10.1007/s13202-019-00802-6.
17. Ahmadi MA, Shadizadeh SR, Shah K, Bahadori A. An accurate model to predict drilling fluid density at wellbore conditions. Egypt J Pet 27(1): 1–10; 2018;Available from: https://doi.org/10.1016/j.ejpe.2016.12.002.
18. Xu S, Li J, Wu J, Rong K, Wang G. HTHP static mud density prediction model based on support vector machine. Drill Fluid Complet Fluid 31(3):28–31; 2014;.
19. Ahmadi MA. Toward reliable model for prediction of drilling fluid density at wellbore conditions: a LSSVM model. Neurocomput J 211: 143–149; 2016;Available from: https://doi.org/10.1016/j.neucom.2016.01.106.
20. Kamari A, Gharagheizi F, Shokrollahi A, Arabloo M, Mohammadi AH. Estimating the drilling fluid density in the mud technology: application in high temperature and high pressure petroleum wells. In: Mohammadi AH (ed) Heavy oil. New York, pp 285–295; 2017;.
21. Rahmati AS, Tatar A. Application of radial basis function (RBF) neural networks to estimate oil field drilling fluid density at elevated pressures and temperatures. Oil Gas Sci Technol Rev IFP Energies Nouvelles; 2019;Available from: https://doi.org/10.2516/ogst/2019021.
22. Zhou H, Niu X, Fan H, Wang G. Effective calculation model of drilling fluids density and ESD for HTHP well while drilling. In: Paper IADC/SPE-180573-MS presented at the 2016 IADC/SPE Asia Pacific Drilling Technology Conference, Singapore; 2016;Available from: https://doi.org/10.2118/180573-MS.
23. Charlez PhA, Easton M, Morrice G, and P. Tardy. Validation of Advanced Hydraulic Modeling using PWD Data. Paper presented at the Offshore Technology Conference, Houston, Texas; 1998;Available from: https://doi.org/10.4043/8804-MS.
24. Li TT, Zhang M, Zhang HL, Tang LP. A new method to determine mud density in deflecting well. Advanced Materials Research (734–737): 1338–1342; 2013;Available from: https://doi.org/10.4028/www.scientific.net/AMR.734-737.1338.
25. Demirdal B, Miska S, Takach N, Cunha JC (2007). Drilling Fluids Rheological and Volumetric Characterization Under Downhole Conditions. In: Paper SPE 108111 presented at the Latin American and Caribbean Petroleum Engineering Conference, Buenos Aires, Argentina; 2007;Available from: https://doi.org/10.2118/108111-MS.
26. Kårstad E and Aadnøy BS. Density Behavior of Drilling Fluids During High Pressure High Temperature Drilling Operations. Paper presented at the IADC/SPE Asia Pacific Drilling Technology, Jakarta, Indonesia; 1998;Available from: https://doi.org/10.2118/47806-MS.
27. Harris OO and Osisanya SO. Evaluation of Equivalent Circulating Density of Drilling Fluids under High-Pressure/High-Temperature Conditions. Paper presented at the SPE Annual Technical Conference and Exhibition, Dallas, Texas; 2005;Available from: https://doi.org/10.2118/97018-MS.
28. Hemphill T. Prediction of Rheological Behavior of Ester-Based Drilling Fluids under Downhole Conditions. Paper presented at the International Petroleum Conference and Exhibition of Mexico, Villahermosa, Mexico; 1996;Available from: https://doi.org/10.2118/35330-MS.
29. Boatman WA. Measuring and Using Shale Density to Aid in Drilling Wells in High-Pressure Areas. J Pet Technol 19: 1423–1429; 1967;Available from: https://doi.org/10.2118/1799-PA.
30. Ombe EM, Elyas, OA., Qader TA, Mehdi M. Application of a real time mud density and rheology monitoring system to enhance drilling in high pressure high temperature gas wells with MPD systems. Paper presented at the International Petroleum Technology Conference, Dhahran, Kingdom of Saudi Arabia; 2020;Available from: https://doi.org/10.2523/iptc-19909-abstract.
31. Hoseinpour M, Riahi MA. Determination of the mud weight window, optimum drilling trajectory, and wellbore stability using geomechanical parameters in one of the Iranian hydrocarbon reservoirs. J Petrol Explor Prod Technol 12: 63–82; 2022;Available from: https://doi.org/10.1007/s13202-021-01399-5.
32. Furbish DJ. Fluid physics in geology: an introduction to fluid motions on earth's surface and within its crusts. Oxford University Press, New York; 1997;Available from: https://doi.org/10.1093/oso/9780195077018.001.0001.
33. Hoberock LL, Thomas DC, Nickens HV. Here's how compressibility and temperature affect bottom-hole mud pressure. Oil Gas J 80(12):159–164; 1982;.
34. Kutasov IM. Empirical correlation determines downhole mud density. Oil Gas J 86:61–63; 1988;.
35. Micah EM. Combined Effect of Rheological Model and Equivalent Diameter Definitions on Pressure Losses/Equivalent Circulating Density Estimation. Master Thesis, African University Of Science And Technology Abuja; 2011;.
36. Sorelle RR, Jardiolin RA, Buckley P, Barrrios JR. Mathematical field model predicts downhole density changes in static drilling fluids. In: Paper SPE 11118 presented at SPE annual fall technical conference and exhibition, New Orleans, LA; 1982;Available from: http://dx.doi.org/10.2118/11118-MS.
37. Dahraj NUH and Bhutto AA. Linear mathematical model developed using statistical methods to predict permeability from porosity. Society of Petroleum Engineers - PAPG/SPE Pakistan Section Annual Technical Conference; 2014;Available from: https://doi.org/10.2118/174716-ms.
38. Molnar C. Interpretable Machine Learning - A Guide for

Making Black Box Models Explainable. Independently published; 2021;Available from: https://christophm.github.io/interpretable-ml-book/index.html.

39. Golub GH, Loan CFV. Matrix computation, 3rd edn. Baltimore, USA; 1996;.
40. Ghaffari A, Abdollahi H, Khoshayand MR, Bozchalooi IS., Dadgar A, Rafiee-Tehrani M. Performance comparison of neural network training algorithms in modeling of bimodal drug delivery. International Journal of Pharmaceutics, 327(1–2): 126–138; 2006;Available from: https://doi.org/10.1016/j.ijpharm.2006.07.056.
41. Parrella F. Online support vector regression. A Thesis Presented for the Degree of Information Science, University of Genoa; 2007;.
42. Mohaghegh S. Virtual-Intelligence Applications in Petroleum Engineering: Part 1 - Artificial Neural Networks. J Pet Technol 52: 64–73; 2000. 2000;Available from: https://doi.org/10.2118/58046-JPT.
43. Horowitz J. Bootstrap. In: The New Palgrave Dictionary of Economics, 2nd edn. London, pp 1–7; 2020;Available from: https://doi.org/10.1057/978-1-349-95121-5.
44. Uzair M, Jamil N. Effects of Hidden Layers on the Efficiency of Neural networks. Proceedings - 2020 23rd IEEE International Multi-Topic Conference, INMIC 2020, 1–6; 2020;Available from: https://doi.org/10.1109/INMIC50486.2020.9318195.
45. Hastie T, Tibshirani R. Generalized Additive Models, 1st edn. Suffolk, Great Britain; 1990;Available from: https://doi.org/10.1201/9780203753781.
46. Vanhove J. Receptive multilingualism across the lifespan Cognitive and linguistic factors in cognate guessing. Languages and Literatures/Multilingualism Research. PhD thesis. University of Fribourg; 2014;.
47. Wood SN. Generalized additive models: An introduction with R, 2nd edn. Boca Raton, USA; 2006;.
48. Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM. Mixed effects models and extensions in ecology with R. New York; 2009;.

Bài nghiên cứu

Nghiên cứu về sự thay đổi tỷ trọng dung dịch khoan do nhiệt độ, áp suất và lưu lượng bơm tuần hoàn bằng cách sử dụng mạng nơron nhân tạo, các mô hình thống kê và các tương quan thực nghiệm

Phạm Sơn Tùng[1,2,*], Phạm Thanh Nhân[1,2]

TÓM TẮT
Bài báo sử dụng một số phương pháp thống kê và học máy nhằm xác định tỷ trọng dung dịch khoan trong các điều kiện áp suất và nhiệt độ khác nhau. Bên cạnh đó, sự ảnh hưởng của các thông số vận hành như là tỷ trọng dung dịch khoan ở điều kiện tiêu chuẩn và lưu lượng bơm tuần hoàn cũng được đề cập tới trong nghiên cứu này. Các loại mô hình khác nhau (mô hình thực nghiệm, mạng nơron nhân tạo, mô hình Generalized Additive, mô hình tuyến tính) đã được xây dựng và so sánh kết quả trên cùng các bộ số liệu đầu vào. Kết quả nghiên cứu cho thấy việc xác định chính xác tỷ trọng dung dịch khoan ở điều kiện bề mặt có ảnh hưởng lớn nhất tới độ chính xác của giá trị dung dịch khoan tại các độ sâu khác nhau. Ngoài ra, mức độ ảnh hưởng của lưu lượng bơm tuần hoàn dù không lớn nhưng cũng không nên bỏ qua nếu muốn tăng tính chính xác trong dự đoán. Phương pháp Bootstrap cũng được dùng trong nghiên cứu này nhằm giải quyết vấn đề số lượng số liệu đầu vào bị hạn chế. Hiện tượng overfitting (quá khớp) cũng đã được nghiên cứu kỹ lưỡng trong bài báo này, nhằm giải quyết một vấn đề thường rất hay gặp trong các nghiên cứu sử dụng học máy ngày nay, khi mà các mô hình cho kết quả dự báo rất chính xác trên bộ số liệu đầu vào, nhưng khi áp dụng cho số liệu thực tế thì lại không thể sử dụng được.
Từ khoá: tỷ trọng dung dịch khoan, học máy, mạng nơron nhân tạo, tương quan thực nghiệm

[1]*Bộ môn Khoan và Khai thác Dầu khí, Khoa Kỹ thuật Địa chất và Dầu khí, Trường Đại học Bách Khoa TP. HCM*

[2]*Đại học Quốc gia Thành phố Hồ Chí Minh*

Liên hệ

Phạm Sơn Tùng, Bộ môn Khoan và Khai thác Dầu khí, Khoa Kỹ thuật Địa chất và Dầu khí, Trường Đại học Bách Khoa TP. HCM

Đại học Quốc gia Thành phố Hồ Chí Minh

Email: phamsontung@hcmut.edu.vn

Lịch sử
- Ngày nhận: 11-01-2024
- Ngày sửa đổi: 30-4-2024
- Ngày chấp nhận: 13-11-2024
- Ngày đăng:

DOI :

Trích dẫn bài báo này: Tùng P S, Nhân P T. **Nghiên cứu về sự thay đổi tỷ trọng dung dịch khoan do nhiệt độ, áp suất và lưu lượng bơm tuần hoàn bằng cách sử dụng mạng nơron nhân tạo, các mô hình thống kê và các tương quan thực nghiệm**. *Sci. Tech. Dev. J. - Eng. Tech.* 2025; ():1-1.