Deep Learning Approach for Robust Prediction of Reservoir Bubble Point Pressure

Fahd Saeed Alakbari,* Mysara Eissa Mohyaldinn,* Mohammed Abdalla Ayoub, and Ali Samer Muhsan

ABSTRACT: The bubble point pressure \( (P_b) \) is a crucial pressure–volume–temperature (PVT) property and a primary input needed for performing many petroleum engineering calculations, such as reservoir simulation. The industrial practice of determining \( P_b \) is by direct measurement from PVT tests or prediction using empirical correlations. The main problems encountered with the published empirical correlations are their lack of accuracy and the noncomprehensive data set used to develop the model. In addition, most of the published correlations have not proven the relationships between the inputs and outputs as part of the validation process (i.e., no trend analysis was conducted). Nowadays, deep learning techniques such as long short-term memory (LSTM) networks have begun to replace the empirical correlations as they generate high accuracy. This study, therefore, presents a robust LSTM-based model for predicting \( P_b \) using a global data set of 760 collected data points from different fields worldwide to build the model. The developed model was then validated by applying trend analysis to ensure that the model follows the correct relationships between the inputs and outputs and performing statistical analysis after comparing the most published correlations. The robustness and accuracy of the model have been verified by performing various statistical analyses and using additional data that was not part of the data set used to develop the model. The trend analysis results have proven that the proposed LSTM-based model follows the correct relationships between the inputs and outputs and performing statistical analysis after comparing the most accurate models in the literature. These values are much better than those given by the most accurate models in the literature.

1. INTRODUCTION

Bubble point pressure \( (P_b) \) is known as the pressure at which the gas’s initial bubble emerges from the liquid oil solution.\(^1\) Obtaining accurate \( P_b \) is essential for performing proper reservoir simulations, production optimization, and oil recovery calculations.\(^2\)\(^–\)\(^5\) Thus, accurate calculation of \( P_b \) is hugely significant in the petroleum industry. Therefore, enhancing \( P_b \) correlations and models for robust and accurate determination of reservoir \( P_b \) has been one of the hot topics of petroleum engineering research.\(^6\)

In this regard, substantial research efforts have been devoted by many researchers leading to the availability of valuable prediction correlation and reported procedures for determining \( P_b \). Standing\(^7\) used 105 data points from California crude oil, USA, to obtain the \( P_b \) correlation. Lasater\(^8\) published a correlation to obtain \( P_b \) based on 158 experimental data points from the USA and Canada. Glaso\(^9\) applied linear and nonlinear regressions to obtain a \( P_b \) correlation utilizing 41 data points from the North Sea to estimate the \( P_b \). Vasquez and Beggs\(^10\) implemented regression techniques to find correlations for the \( P_b \) prediction utilizing 6004 data points worldwide. Al-Marhoun\(^11\) used nonlinear multiple regression analysis and introduced a \( P_b \) correlation using 160 data points from the Middle East. Kartoatmodjo and Schmidt\(^12\) applied nonlinear regression analysis to develop a bubble point prediction correlation using 5392 data points from the USA and South-East Asia. They formed a correlation to determine \( P_b \). Dokal and Osman\(^13\) utilized a regression method utilizing 51 data points from the UAE to build a \( P_b \) prediction correlation. Petrosky and Farshad\(^14\) proposed a bubble point pressure correlation utilizing 90 data points from the Gulf of Mexico. They applied regression methods using the statistical analysis system (SAS) software. Macary and El-Batanoney\(^15\) applied 90 data points from the Gulf of Suez to build a \( P_b \) correlation. Omar and Todd\(^16\) utilized a nonlinear regression method to obtain \( P_b \) utilizing 93 data points from Malaysia. De Ghetto et al.\(^17\) applied multiple, linear, and nonlinear regressions using the statistical analysis system (SAS) program to develop a technique to obtain \( P_b \).
bubble point pressure correlation utilizing 3700 data points from the Mediterranean Basin, Africa, Persian Gulf, and the North Sea. Frashad et al. applied multiple nonlinear regression techniques to develop a \( P_b \) correlation utilizing 43 data points from Colombia. Almehaideb proposed a \( P_b \) correlation utilizing 62 data points from the United Arab Emirates (UAE). He used regression analysis to obtain the correlation. Hanafy et al. applied regression analysis to develop a \( P_b \) prediction using 324 data points from Egypt. Velarde et al. applied nonlinear regression methods utilizing 2097 data points worldwide to create a \( P_b \) correlation. Gharibi et al. utilized a neural network to find \( P_b \). Al-Shammasi used linear and nonlinear regression to propose a \( P_b \) correlation utilizing 1661 data points worldwide. Dindoruk and Christman developed the equation to estimate \( P_b \) utilizing Microsoft Excel. The correlation was based on 100 Gulf of Mexico pressure−volume−temperature (PVT) laboratory. Mehran et al. and Bolondarzadeh et al. proposed a \( P_b \) correlation using 387 data points and 166 data points, respectively, from Iran fields. Malallah et al. used graphical alternating conditional expectation to calculate a \( P_b \). Hemmati and Kharrat proposed a \( P_b \) correlation using 287 data points from Iranian crude oils. They used the nonlinear multiple regression techniques to build the \( P_b \) correlation. Mazandarani and Asghari applied multiple regression analysis to develop a \( P_b \) correlation utilizing 55 data points from different locations in Iran. Kamehchi et al. applied a stepwise multiple linear regression analysis using a statistical software package Statistical Package for the Social Sciences (SPSS) for Windows to develop a \( P_b \) correlation, utilizing 94 crude oil data points. Dutta and Gupta used support vector regression to find a \( P_b \) calculation. Arabloo et al. introduced a \( P_b \) correlation by applying a multistart feature of LINGO software and using 750 data points worldwide. Patil et al. applied the multiple regression methods and utilized 23 data points from Niger Delta to calculate \( P_b \). Gomaa used support vector regression to obtain \( P_b \) models. Sharrad and Abd-Alrahman presented a \( P_b \) correlation applying 35 data points from Libya crude oil, using linear regression analysis through EVViews software. Seyyedattar et al. used 569 data points from different places applying extra trees, a least square support vector machine, and an adaptive network-based fuzzy inference system to predict \( P_b \). They showed that the extra trees model is a robust tool to forecast \( P_b \). Talekeikhah et al. utilized a radial basis function neural network, multilayer perceptron, support vector regression, adaptive neuro-fuzzy inference system, decision trees, and random forest to obtain \( P_b \). They indicated that the decision trees model is the best model with an average absolute relative error of 3.379%. Tariq et al. used a functional network coupled with particle swarm optimization to determine \( P_b \). However, the previous models lack accuracy and have not applied trend analysis to identify the correct relationships between the inputs and outputs as part of the validation process. To our knowledge, this paper is the first study to apply the long short-term memory (LSTM) network and trend analysis for developing a bubble point pressure model.

In this work, we developed an accurate and reliable LSTM-based model by implementing trend analysis to predict \( P_b \). LSTM was applied in this study because it has surpassed other deep learning techniques such as recurrent neural network (RNN) and can capture the nonlinear trends in data and remember previous information for a long time. LSTM has self-connecting hidden layers and a gating structure to control long-sequence time-series data efficiently. Therefore, LSTM has been successfully used to solve many problems. Furthermore, the trend analysis is used to present the relationships between inputs (gas solubility (\( R_s \)), gas-specific gravity (\( g \)), oil specific gravity (API), and reservoir temperature (\( T_r \)) and output (bubble point pressure (\( P_b \))) in the existing and LSTM models. The trend analysis defines errors in the existing and LSTM models to present unexpected relationships between inputs and output, highlighting the need to show the model’s reliability. Moreover, several statistical analyses such as average absolute percentage relative error (AAPRE), maximum absolute percent relative error (\( E_{\text{max}} \)), minimum absolute percent relative error (\( E_{\text{min}} \)), root mean square error (RMSE), standard deviation (SD), and correlation coefficient (\( R \)) have been performed to describe, validate, and compare the LSTM model with the previously published models.
2. METHODOLOGY

Figure 1 presents the workflow diagram of the previous and proposed LSTM models. The workflow mainly included six phases: data collection, testing of previous models, training the LSTM model, trend analysis, testing the LSTM model, and comparing the models. **2.1. Data Collection and Preprocessing.** Seven hundred and sixty data points, collected from different published sources relevant to this model, were utilized in this study.

**Table 1. Statistical Description of the Collected Data (760 Data Points)**

| Parameters | training data | | testing data | |
|---|---|---|---|---|
| $P_b$ (psi) | $R_s$ (SCF/STB) | $\gamma_g$ API (°API) | $T_f$ (°F) | $P_b$ (psi) | $R_s$ (SCF/STB) | $\gamma_g$ API (°API) | $T_f$ (°F) |
| minimum | 126 | 9 | 0.589 | 15.3 | 74 | 130 | 26 | 0.589 | 19.4 | 74 |
| maximum | 7127 | 2637 | 1.367 | 59.5 | 19.4 | 4432 | 1850 | 1.367 | 51.7 | 271 |
| standard deviation (SD) | 1790 | 536 | 0.905 | 36.05 | 171 | 1135 | 424 | 0.160 | 6.4 | 45.3 |
| mean | 1697 | 433 | 0.876 | 36.7 | 169 | 1912 | 560 | 0.930 | 32.6 | 160 |
| median | 500 | 585 | 0.802 | 33.3 | 100 | 4005 | 61 | 0.802 | 31.2 | 100 |
| Kurtosis | 1.447 | 1.959 | 0.325 | 7.32 | 49.4 | 1.134 | 1.188 | 0.260 | 7.32 | 51.7 |
| Skewness | 0.881 | 1.230 | 0.834 | 36.05 | 171 | 0.107 | 0.107 | 0.766 | 0.734 | 0.182 |
| Range | 7001 | 2628 | 0.778 | 44.2 | 220 | 4302 | 1824 | 0.778 | 32.3 | 197 |
| Interquartile range | 1648 | 559 | 0.201 | 10.35 | 74 | 2086 | 731 | 0.230 | 9.9 | 67 |

**Table 2. Comparison of the Used Data and Accuracy for the Previous and the Proposed LSTM Models**

| No. | Model | Location | no. of data points | $R_b$ (bbl/STB) | $R_s$ (SCF/STB) | $\gamma_g$ API (°API) | $T_f$ (°F) | Average absolute error % | Standard deviation |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Standing | California USA | 105 | | | | | | |
| 2 | Lasater | Canada West and Mid-cont U.S. and SA | 158 | | | | | | |
| 3 | Glaso | North Sea | 41 | | | | | | |
| 4 | Vazquez and Beggs | Worldwide | 6004 | | | | | | |
| 5 | Al-Marhoun | Middle East | 160 | | | | | | |
| 6 | Kartosmatmojo and Schmidt | Worldwide | 5392 | | | | | | |
| 7 | Dokla and Osman | UAE | 51 | | | | | | |
| 8 | Petrovsky and Farshed | Gulf of Mexico | 90 | | | | | | |
| 9 | Macary and El-Batanoney | Gulf of Suez | 90 | | | | | | |
| 10 | Omar and Todd | Malaysia | 93 | | | | | | |
| 11 | De Ghetto et al. | Mediterranean Basin, Africa, Persian Gulf, and the North Sea | 195 | | | | | | |
| 12 | Frashad et al. | Colombia | 43 | | | | | | |
| 13 | Almehaideb | UAE | 62 | | | | | | |
| 14 | Hanafy et al. | Egypt | 90 | | | | | | |
| 15 | Velarde et al. | USA | 728 | | | | | | |
| 16 | Al-Shammasi | Worldwide | 1709 | | | | | | |
| 17 | Dindoruk and Christman | Gulf of Mexico | 100 | | | | | | |
| 18 | Mehran et al. | Iran | 387 | | | | | | |
| 19 | Bologjadeh et al. | Iran | 166 | | | | | | |
| 20 | Hemati and Kharrat | Iran | 287 | | | | | | |
| 21 | Mazandarani and Aghar | Iran | 55 | | | | | | |
| 22 | Khamchi et al. | Iran | 106 | | | | | | |
| 23 | Arablobo et al. | Worldwide | 756 | | | | | | |
| 24 | Patil et al. | Niger Delta | 23 | | | | | | |
| 25 | Gomaa | Middle East | 441 | | | | | | |
| 26 | Sharrad and Abd-Alrahman | Libya | 35 | | | | | | |
| 27 | Tariq et al. | Worldwide | 760 | | | | | | |
| 28 | proposed LSTM | Worldwide | 760 | | | | | | |

ACS Omega http://pubs.acs.org/journal/acsodf

ACS Omega 2021, 6, 21499−21513

ACS Omega Article

21501 https://doi.org/10.1021/acsomega.1c02376

2.2. LSTM Model. Trend analysis, testing the LSTM model, and comparing the models.
work. The preprocessing step involved data cleaning for outliers in the collected data. The collected data set was split into two subparts; 70% (532 data points) of the data was allocated for training. In contrast, 30% (228 data points) of the collected data was allocated for testing the model. Table 1 lists the statistical description of the collected training and testing data. The box and whisker plot was used to detect and remove outliers in the collected data to obtain normal data. The box and whisker plot was introduced by Tukey.40 Figure S1 (Supporting Information) presents the box and whisker plot that displays lower quartile Q1, upper quartile Q3, interquartile range (IQR), lowest observation (Q1 − 1.5 × IQR), and the highest observation (Q3 + 1.5 IQR). Outliers are any value less than (Q1 − 1.5 × IQR) or more than (Q3 + 1.5 IQR), Figure S1. The box and whisker plots for the inputs and output parameters for the collected and normal training data are presented in Figures S2 and S3, respectively. The testing data does not have any outlier to make the box and whisker plots of the inputs and output parameters for the collected and normal testing data the same (refer to Figure S4). The statistical description of the normal data is shown in Table S1 (Supporting Information). Figures S5 and S6 show histograms of the inputs and output parameters for the collected and the normal training data are presented in Figures S2 and S3, respectively. The histograms of the inputs and output parameters for the collected and the normal testing data are shown in Figure S7.

We considered what is reported in the literature regarding the dependency of $P_{b}$ on other fluid properties. We found that $P_{b}$ highly depends on $R_{c}$, $\gamma_{f}$, API, and $T_{p}$ as reported by most previous studies; Table 2. Randomization of data was applied using the shuffle method that changes the original list of the data, and it does not return a new list of the data. Randomization can be used to assure that each data set does not memorize the pattern to avoid generalization and model overfitting. The correlation coefficient ($R$) was calculated to assess each feature’s significance ($R_{c}$, $\gamma_{f}$, API, and $T_{p}$) to the output $P_{b}$. It is found that $P_{b}$ is a strong function of $R_{c}$ (the correlation coefficient of 0.876 and 0.851 for collected and normal data, respectively) but $\gamma_{f}$ is moderate (the correlation coefficient of −0.513 and −0.544 for collected and normal data, respectively). $P_{b}$ is a weak function of API (correlation coefficient of 0.383 and 0.410 for collected and normal data, respectively). $P_{b}$ is also a weak function of $T_{p}$ (correlation coefficient of 0.315 and 0.265 for collected and normal data, respectively); Figure 2.

Before the proposed LSTM model was applied, all of the parameters (inputs and outputs) are normalized between −1 and 1 to scale the parameters within a specified range using the following equation

$$Y = \frac{Y_{max} - Y_{min}}{X_{max} - X_{min}} (X - X_{min}) + Y_{min}$$  \hspace{1cm} (1)

$$Y = \frac{X - X_{min}}{X_{max} - X_{min}}$$  \hspace{1cm} (2)

where $Y$ is the parameter in the normalized form, $Y_{max}$ is the maximum value of the normalized form (1), $Y_{min}$ is the minimum value of the normalized form (−1), $X$ is the parameter to be normalized, $X_{min}$ is the minimum value of the parameter, and $X_{max}$ is the maximum value of the parameter.

2.2. Proposed LSTM Model. LSTM is a special type of recurrent neural network (RNN) mostly applied to define the relationship between the current and previous input data.41 Its memory is utilized to save the internal information of the previously input data. The long short-term memory hidden state can contain some memory cells (recurrently connected cells). Each LSTM memory cell involves three gates, the forget gate, the input gate, and the output gate that maintains information movement, which moves through the given layer; Figure 3. A memory cell, identified as a cell state, aims to store all significant information that is to be passed in the next layer. Eq 3 is used to update the cell state to a new value relevant to the new cell state. The inputs are shown as x at the bottom, and the outputs are shown as $h_{i}$; Figure 4.42 The gates can learn and remember the input sequence data, which is significant:

$$C_{t} = f_{t} * C_{t-1} + i_{t} * \tilde{C}_{t}$$  \hspace{1cm} (3)
On the other hand, the method can discard unnecessary information. There are some multiple activation functions applied in each gate, namely, tanh and sigmoid activation. The activation functions can be used to determine which information passes through the layer.

2.2.1. tanh Activation. tanh activation can control the data passing through the layer and squish data to ensure that the range of data falls between −1 and 1. Moreover, it can regulate the output to ensure that the value is between −1 and 1 or otherwise make the other insignificant value.

\[ \phi(z) = \frac{1 - e^{-2z}}{1 + e^{-2z}} \]  

(4)

2.2.2. Sigmoid Activation. Sigmoid activation squishes values between one and zero and can update or forget the information. When it is zero, the information disappears or is forgotten, whereas the information is remembered when it is one. Therefore, the network can learn when the information can be significant and can be maintained. When the information is not significant, it can be forgotten.

\[ \sigma(z) = \frac{1}{1 + e^{-z}} \]  

(5)

2.2.3. Forget Gate. Forget gate can be used to decide on retaining or discarding the information. It can accept both the present sequence and the previous state information and decide to discard irrelevant data. The insignificant data is extracted, applying the multiplication of a filter. Sigmoid activation can be required here to make the information be between one and zero. The operation in the forget gate can be obtained using

\[ f_i = \sigma(W_f[h_{t-1}, x_t] + b_f) \]  

(6)

The forget gate can accept data from the prior hidden state \( h_{t-1} \) and the sequence input \( x_t \). Then, the data can be added with a bias and multiplied by the weights. Next, the values are followed through the sigmoid activation to be kept in the cell state or removed.

2.2.4. Input Gate. Input gate obtains information from \( h_{t-1} \) and \( x_t \) to update information in the cell state. The data is passed into two functions: sigmoid and tanh. Sigmoid activation can be used to decide significant information and can be updated. The values of \( x_t \) and \( h_{t-1} \) are added with a bias and multiplied by the weight to obtain a vector from zero or one. The value of the sigmoid is utilized to determine significant data applying the following equation

\[ i_i = \sigma(W_i[h_{t-1}, x_t] + b_i) \]  

(7)

Then, eq 8 can be used to obtain the tanh function used to convert the data between −1 and 1 using the following equation. This was to create a new vector that needed to be applied to the cell.

\[ \tilde{C}_i = \tanh(W_C[h_{t-1}, x_t] + b_C) \]  

(8)

Before being applied to the present cell state, the sigmoid and tanh function values were multiplied together. This step is significant to ensure that the data added to the cell is essential and not redundant.

2.2.5. Output Gate. It is the last gate that relies on current inputs to obtain the next hidden state. Two steps are included in the output gate to complete the process.

The first step can be to obtain the current input \( x_t \) and the prior hidden state \( h_{t-1} \). It then provides values into the sigmoid function. The sigmoid activation can be similar to the other two gates, where they can be multiplied by the weight before being added with bias.

\[ o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \]  

(9)

The next step can be to obtain the cell state’s value and bring it into the tanh activation. Then, the tanh function’s value is multiplied with the sigmoid’s output to determine the information for the hidden state: \(^{43}\) Our study focuses on applying LSTM to predict \( P_b \) (more about LSTM is found in different references, such as \(^{44−46}\)).

\[ h_t = o_t \times \tanh(C_t) \]  

(10)

where \( f_i \) is the output of the forget gate, \( i_i \) is the output of the input gate, \( C_t \) is the temporary memory cell in the memory block that is produced by a tanh layer, \( C_i \) is the temporary memory cell in the memory block for the previous cell state, \( C_{t-1} \) is the temporary memory cell in the memory block for the current cell state, \( C_{t+1} \) is the temporary memory cell in the memory block for the previous cell state, and \( o_t \) is the output of the output gate. \( W_\rho \), \( W_\sigma \), \( W_o \) are the recurrent weight matrices; \( b_\rho \), \( b_\sigma \), \( b_o \) are the related bias vectors; \( h_{t-1} \) represents the previous results in each LSTM cell; \( h_t \) represents the results in each LSTM cell; \( x_t \) is the current input in each LSTM cell; and \( \sigma \) is the sigmoid function that determines the propagation of the information.

MATLAB software was used to develop the proposed LSTM model. The pseudocode of the proposed LSTM model is presented in Algorithm 1 (see the Supporting Information). Table 3 shows the hyperparameters of the LSTM MATLAB code applied to find \( P_b \).
Table 3. Hyperparameters of the LSTM Model

| parameter                  | description/value |
|----------------------------|-------------------|
| features (inputs)          | 4                 |
| responses (outputs)        | 1                 |
| hidden units               | 50                |
| max epochs                 | 256               |
| mini-batch size            | 128               |
| gradient threshold         | 0.2000            |
| initial learn rate         | 0.3000            |
| verbose                    | 1                 |
| training options           | sgd (stochastic gradient descent with momentum) |
| momentum                   | 0.9000            |
| L2 Regularization (factor for L2 regularization) | 0.0001 |
| OutputMode (format of output) | sequence |
| StateActivationFunction (activation function to update the cell and hidden state) | tanh |
| GateActivationFunction (activation function to apply to the gates) | sigmoid |
| shuffle                    | once              |
| InputWeightsInitializer (function to initialize input weights) | Glorot (Glorot initializer) |
| LearnRateSchedule (option for dropping the learning rate during training) | none (the learning rate remains constant throughout the training) |
| RecurrentWeightsInitializer (function to initialize recurrent weights) | orthogonal |
| BiasInitializer (function to initialize bias) | unit-forget-gate |
| GradientThresholdMethod (gradient threshold method) | l2norm |
| SequenceLength (option to pad, truncate, or split input sequences) | longest |
| SequencePaddingDirection (direction of padding or truncation) | right |
| ExecutionEnvironment (hardware resource for the training network) | auto |

3. RESULTS AND DISCUSSION

3.1. \( P_b \) Prediction Model Using LSTM. The proposed LSTM model was validated using 228 unseen data points (30% testing data set) after being studied by applying trend analysis to ensure that the model was stable and had sound relationships between the inputs and the output (see Section 3.3). In addition, statistical error analysis was performed to evaluate the LSTM model, as seen in Table 4. As a result, the LSTM model has the lowest average absolute percentage relative error (AAPRE), maximum absolute percent relative error (\( E_{\text{max}} \)), minimum absolute percent relative error (\( E_{\text{min}} \)), root mean square error (RMSE), and standard deviation (SD) and the highest correlation coefficient (\( R \)). As per the tabulated results in Table 4, the accuracy of the LSTM model is outstanding.

The cross-plotting analysis was applied to assist in visualizing the model’s accuracy. First, the measured \( P_b \) values are plotted vs the predicted \( P_b \) values. Then, a 45° straight line is plotted on the measured \( P_b \) values’ cross plot and the expected \( P_b \) values. The closer the plotted data points to the straight line, the higher the correlation or model’s accuracy.

Table 4. Statistical Error Analysis of the LSTM Model

| AAPRE (%) | \( E_{\text{max}} \) (%) | \( E_{\text{min}} \) (%) | RMSE | SD  | \( R \) |
|-----------|------------------------|------------------------|------|-----|-------|
| 8.422     | 45.409                 | 0.040                  | 12.20| 0.088| 0.990 |

Figures 5 and 6 demonstrate the cross-plotting of the training and testing data set of the proposed LSTM model with the coefficient of determination \( (R^2) \) of 0.9837 and 9843 for the training and testing data sets, respectively.

3.2. Proposed LSTM Model and Current Models’ Comparison. Some statistical analyses have been performed to describe, validate, and compare the proposed LSTM model with the previously published models that follow the correct relationships between the inputs and the output (see Section 3.3). The statistical error analysis used for the LSTM model and the previous models included cross plot, correlation coefficient, AAPRE, RMSE, SD, \( E_{\text{max}} \) and \( E_{\text{min}} \). In this study, AAPRE and \( R \) are used as key indicators.

3.2.1. Cross-Plotting Analysis. Figure 7 displays the cross-plotting and coefficient of determination \( (R^2) \) of the LSTM model compared to the existing correlations. As presented in the figure, the LSTM model shows the highest accuracy and can predict \( P_b \) with the coefficient of determination \( (R^2) \) of 0.9843.

3.2.2. Statistical Error Analysis. Figure 8 shows the AAPRE and correlation coefficient (\( R \)) comparison of the LSTM model with the previously published models. As shown, the LSTM model indicates the lowest AAPRE of 8.422% and the highest correlation coefficient (\( R \)) of 0.99. On the other hand, the Velarde et al.\(^2\) and Mehran et al.\(^24\) models display AAPRE (%) of 9.0 and 9.75% and correlation coefficient (\( R \)) of 0.972 and 0.969, respectively. The Petrosky and Farshad\(^14\) model shows the highest AAPRE of 76.59%.

The prediction performance of the published correlations was compared with the LSTM model. Table 5 shows the statistical parameters for the comparative evaluation of the existing correlations and the LSTM model. Statistical error analysis has been performed to describe the robustness of the LSTM model. AAPRE and correlation coefficient (\( R \)) were used as primary indicators of the model’s accuracy. As seen in the table, the LSTM model and the existing correlations are ranked based on AAPRE and the correlation coefficient (\( R \)). The LSTM model also has the lowest RMSE and SD compared to other models. This useful comparison of all correlations and the LSTM model forms an essential means of evaluating correlations and the LSTM performance. Investigation of these statistical error analyses proves that the LSTM model exceeds all of the current correlations.

From Table 5, it can also be observed that the Velarde et al.\(^2\) correlation is the second-ranked correlation with AAPRE of 9% and the correlation coefficient (\( R \)) of 0.972. The Petrosky and Farshad\(^14\) model shows AAPRE of 76.59% and the correlation coefficient (\( R \)) of 0.970 in predicting \( P_b \).

3.3. Sensitivity Analysis. Sensitivity analysis or trend analysis is the study of the uncertainty in the output of a model.\(^47\) The trend analysis is utilized to determine the model’s reliability. Trend analysis may define the relationships in a model between output and input variables and look for model errors. Additionally, trend analysis can also identify and remove unnecessary parts of the model structure.\(^48\) An input parameter can be changed between the minimum value and the maximum value, while the others are fixed at their constant mean values.\(^22,49–51\) Graphs are plotted for the input parameter values as the x-axis against \( P_b \) as the y-axis for the previous models and the LSTM model.

Four input parameters (\( R_w \), \( \phi \), \( API \), and \( T_i \)) have been selected for the trend analysis. Oil formation volume factor at the bubble point pressure (\( B_{vb} \) (bbl/STB)) has been discarded from the model’s inputs as Omar and Todd,\(^19\) Almeheidab,\(^19\)
and Hemati and Kharrat\textsuperscript{27} only apply it. No apparent trend of \(P_b\) has been reported in the literature.

The trend of \(R_s\) is displayed in Figure 9. Petrosky and Farshad\textsuperscript{14} and Almehaideb\textsuperscript{19} developed correlations based on the \(R_s\) ranges of 217–1406 and 128–3871 SCF/STB, respectively. The Petrosky and Farshad\textsuperscript{14} and Almehaideb\textsuperscript{19} correlations show that \(P_b\) was negative, \(-812.6\) and \(-207.5\) psi, respectively, at \(R_s\) 26 SCF/STB (Figure 9). The LSTM model demonstrates that the trend obeys the trend of the correlations in the literature where \(R_s\) is directly proportional to \(P_b\); Figure 10. Therefore, the LSTM model is successful in following the models’ trends.

The trend of gas-specific gravity (\(\gamma_g\)) is demonstrated in Figure 11. As presented by the previous models, the gas-specific gravity (\(\gamma_g\)) is inversely proportional to \(P_b\). Nevertheless, the Hanafy et al.\textsuperscript{20} correlation indicates that \(P_b\) is constant as they did not include the gas-specific gravity in their correlation. The Goma\textsuperscript{32} correlation has shown that a slight increase in gas-specific gravity can increase \(P_b\), contradicting the previous correlation behavior. Omar and Todd’s\textsuperscript{16} correlation showed different trends in which the gas-specific gravity shows a decreasing trend and then increases. Consequently, Hanafy et al.,\textsuperscript{20} Omar and Todd,\textsuperscript{16} and Goma\textsuperscript{32} failed to represent the behavior accurately. The LSTM model trend followed the most current correlation trend; Figure 12 shows that the LSTM model follows the proper gas-specific gravity trend.

The trend of oil API gravity (API) is displayed in Figure 13. The oil API gravity (API) is inversely proportional to \(P_b\). Hanafy et al.\textsuperscript{20} showed a horizontal line behavior, which indicates that their correlation does not consider the oil API gravity. Dokla and Osman\textsuperscript{13} and Gomaa\textsuperscript{32} correlations followed the trend of existing correlations, but increasing the oil API gravity slightly leads to a decrease in \(P_b\). The Tariq et al.\textsuperscript{37} model shows that increasing the oil API gravity leads to an increase of \(P_b\).

Consequently, Hanafy et al.,\textsuperscript{20} Dokla and Osman,\textsuperscript{13} Gomaa,\textsuperscript{32} and Tariq et al.\textsuperscript{37} have not proven a proper trend for the \(P_b\) correlation. Petrosky and Farshad\textsuperscript{14} established their
Figure 7. Cross plot comparison of the LSTM model with the previously published models.

Figure 8. Correlation coefficient (R) and AAPRE (%) comparison of the LSTM model with previously published models.
correlation based on the API oil gravity range of 16.3–45 °API. Thus, the Petrosky and Farshad14 correlation shows that $P_b$ is negative, −37.35 and −145.91 psi at API oil gravity of 48.11 and 51.7 °API, respectively; Figure 13. Figure 14 displays that the oil API gravity trend of the LSTM model obeys the trend of the existing correlations.

Table 5. Statistical Error Analysis Comparison of the LSTM Model with the Previously Published Models

| rank | model                        | AAPRE (%) | $E_{max}$ (%) | $E_{min}$ (%) | RMSE  | SD   | R     |
|------|------------------------------|-----------|---------------|---------------|-------|------|-------|
| 1    | The proposed LSTM (2021)     | 8.422     | 45.409        | 0.040         | 12.20 | 0.088| 0.990 |
| 2    | Velarde et al. (1997)        | 9.00      | 62.47         | 0.039         | 13.04 | 0.095| 0.9724|
| 3    | Mehran et al. (2006)         | 9.75      | 63.86         | 0.035         | 13.60 | 0.095| 0.9699|
| 4    | Lasater (1958)               | 11.07     | 66.08         | 0.016         | 15.31 | 0.106| 0.9742|
| 5    | Standing (1947)              | 12.35     | 69.28         | 0.032         | 16.26 | 0.106| 0.9753|
| 6    | Arabloo et al. (2014)        | 12.66     | 72.98         | 0.000         | 17.12 | 0.116| 0.9589|
| 7    | Hemati and Kharrat (2007)    | 13.76     | 85.01         | 0.026         | 22.13 | 0.174| 0.9741|
| 8    | Vazquez and Beggs (1980)     | 16.88     | 74.79         | 0.493         | 21.65 | 0.136| 0.9767|
| 9    | Kartosmos and Schmit (1991)  | 16.94     | 78.37         | 0.085         | 22.74 | 0.152| 0.9722|
| 10   | Al-Shammasi (1999)           | 17.33     | 62.95         | 0.205         | 22.60 | 0.145| 0.9663|
| 11   | Frashad et al. (1996)        | 18.23     | 74.23         | 0.042         | 24.30 | 0.161| 0.9621|
| 12   | De Ghetto et al. (1994)      | 18.37     | 73.97         | 0.007         | 24.83 | 0.167| 0.9720|
| 13   | Dindoruk and Christman (2001)| 20.89     | 77.83         | 0.432         | 25.81 | 0.152| 0.9369|
| 14   | Glaso (1980)                 | 23.02     | 79.52         | 0.281         | 27.70 | 0.154| 0.9701|
| 15   | Mazandarani and Aghari (2007)| 23.91     | 120.93        | 0.127         | 34.19 | 0.245| 0.9462|
| 16   | Almeidaheb (1997)            | 26.15     | 234.92        | 0.037         | 44.18 | 0.357| 0.9482|
| 17   | Macary and El-Batanoney (1993)| 31.20  | 149.75        | 0.111         | 42.62 | 0.291| 0.9499|
| 18   | Khami et al. (2009)          | 31.24     | 97.52         | 0.059         | 37.27 | 0.204| 0.9652|
| 19   | Bolozadazadeh et al. (2006)  | 40.42     | 434.20        | 0.175         | 84.69 | 0.746| 0.9694|
| 20   | Sharrad and Abd-Alrahman (2019)| 45.93    | 72.46         | 0.346         | 47.96 | 0.139| 0.8929|
| 21   | Al-Marhoun (1988)            | 54.06     | 79.22         | 27.176        | 54.40 | 0.148| 0.9538|
| 22   | Petrosky and Farshad (1993)  | 76.59     | 784.59        | 0.295         | 159.87| 1.406| 0.9703|

Figure 9. Gas solubility trend analysis of the LSTM model and previously published models.

Figure 15 reveals the trend of the reservoir temperature. $P_b$ is an increasing function regarding reservoir temperature. However, the Hanafy et al.20 correlation shows a horizontal line, which indicates that their correlation did not include the reservoir temperature parameter. Dindoruk and Christman23 and Arabloo et al.6 correlations indicate that $P_b$ is almost
constant, including the reservoir temperature in their correlations. Dokla and Osman’s correlation suggests that $P_b$ decreases with the increase of the reservoir temperature. Therefore, the Dokla and Osman’s correlation shows that the trend fails to show dependency adequately. The trend demonstrated by the LSTM model also follows the correct relationships between inputs and output, Figure 16, indicating that it represents the proper trend for the reservoir temperature.

In conclusion, all of the input parameters ($R_s$, $\gamma_w$, API, and $T_r$) of the LSTM model follow the proper trends. Consequently,
the trend analysis represented the reliability of the proposed LSTM model. In contrast, Hanafy et al.\textsuperscript{20} Goma,\textsuperscript{32} Omar and Todd,\textsuperscript{16} Dokla and Osman,\textsuperscript{13} and Tariq et al.\textsuperscript{37} correlation trends fail to present the behavior correctly. Finally, Petrosky and Farshad\textsuperscript{14} and Almehaideb\textsuperscript{19} correlations show the predicted $P_b$’s negative values because the used values of $R_s$ and oil API are out of their considered ranges.

4. CONCLUSIONS

The LSTM-based prediction model was developed, validated, and compared against the previous models available in the literature to predict $P_b$ using 760 data points. As a result, the following conclusions can be drawn.

The model’s trend analysis results indicate that the model can accurately describe the profile $P_b$ as a function of all of the considered independent variables following the actual trend expected from the physical relationship.
The LSTM model generates a substantially higher accuracy of $P_b$ compared to all published models. This accuracy is proven by the highest correlation coefficient of 0.99, the lowest AAPRE of 8.422%, the smallest root mean squared error of 12.20, and the lowest standard error deviation (SD) of 0.088 compared to 27 published correlations. The model accuracy and reliability were further enhanced by data randomization to ensure that each data set does not memorize the pattern and avoid generalization and model overfitting.

5. LIMITATIONS AND RECOMMENDATIONS

The LSTM model was built based on bulk properties (gas solubility ($R_s$), gas-specific gravity ($\gamma_g$), oil specific gravity (API), and reservoir temperature ($T_r$)) without considering the compositional data of the crude oil because of the data limitations. The proposed model provides high-accuracy results with the four input parameters considered ($R_s$, $\gamma_g$, API, and $T_r$), especially if the input values are within the data ranges used to
develop the model. However, the proposed model’s benefits outweigh the limitations. The results presented here are self-explanatory. The proposed LSTM model was the most robust and accurate model compared to the published models. We recommend using more data to enhance the performance of LSTM for other applications with optimal hyperparameters. For example, deep learning (LSTM) can be applied for many petroleum industry calculations, such as production engineering and drilling engineering calculations.

ASSOCIATED CONTENT
Supporting Information
The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.1c02376.

Appendix for statistical error analysis; pseudocode of the proposed LSTM model; statistical description of the normal data; box and whisker plot; box and whisker plots for the collected training data; box and whisker plots for the normal training data; box and whisker plots for the collected and normal testing data; histograms of the inputs and output parameters for the collected training data; histograms of the inputs and output parameters for the normal training data; and histograms of the inputs and output parameters for the collected and normal testing data (PDF)

AUTHOR INFORMATION

Corresponding Authors
Fahd Saeed Alakbari – Petroleum Engineering Department, Universiti Teknologi PETRONAS, 32610 Bandar Seri Iskandar, Perak, Malaysia; orcid.org/0000-0002-3227-698X; Email: fahd_19001032@utp.edu.my, alakbarifahd@gmail.com
Mysara Eissa Mohyaldinn – Petroleum Engineering Department, Universiti Teknologi PETRONAS, 32610 Bandar Seri Iskandar, Perak, Malaysia; Email: mysara.eissa@utp.edu.my

Authors
Mohammed Abdalla Ayoub – Petroleum Engineering Department, Universiti Teknologi PETRONAS, 32610 Bandar Seri Iskandar, Perak, Malaysia
Ali Samer Mul san – Mechanical Engineering Department, Universiti Teknologi PETRONAS, 32610 Bandar Seri Iskandar, Perak, Malaysia

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.1c02376

Author Contributions
F.S.A. implemented the writing: original draft preparation, review and editing, methodology, software, and data collection. M.E.M. implemented the problem conceptualization, research supervision and funding, and drafting. M.A.A. implemented model development, result verification, and review and editing. A.S.M. implemented the review and editing.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS
The authors express their appreciation to the Universiti Teknologi PETRONAS for supporting this work under the YUTP-Grant cost center #01SLC0-098. They also especially thank COREOR, Petroleum Engineering, Universiti Teknologi PETRONAS, for supporting this work. The first author, in particular, is grateful to Universiti Teknologi PETRONAS for supporting his Ph.D. study under the Graduate Assistance (GA) scheme.

REFERENCES
(1) McCain, W. D., Jr. The Properties of Petroleum Fluids, 2nd ed.; PennWell Publ. Co.: Tulsa, Oklahoma, 1990.
(2) Velarde, J.; Blasingame, T. A.; McCain, W. D., Jr. In Correlation of Black Oil Properties at Pressures Below Bubble Point Pressure–A New Approach, Annual Technical Meeting: Petroleum Society of Canada, 1997.
(3) Valkó, P. P.; McCain, W. D., Jr. Reservoir Oil Bubblepoint Pressures Revisited; Solution Gas–Oil Ratios and Surface Gas Specific Gravities. J. Pet. Sci. Eng. 2003, 37, 153–169.
(4) Bandyopadhyay, P.; Sharma, A. Development of a New Semi Analytical Model for Prediction of Bubble Point Pressure of Crude Oils. J. Pet. Sci. Eng. 2011, 78, 719–731.

(5) Farasat, A.; Shokrollahi, A.; Arabloo, M.; Gharagheizi, F.; Mohammadi, A. H. Toward an Intelligent Approach for Determination of Saturation Pressure of Crude Oil. Fuel Process. Technol. 2013, 115, 201–214.

(6) Arabloo, M.; Amooie, M.-A.; Hemmati-Saraparodeh, A.; Ghazanfari, M.-H.; Mohammadi, A. H. Application of Constrained Multi-Variable Search Methods for Prediction of PVT Properties of Crude Oil Systems. Fluid Phase Equilib. 2014, 363, 121–130.

(7) Standing, M. B. A Pressure-Volume-Temperature Correlation for Mixtures of California Oils and Gases. In Drilling and Production Practice; American Petroleum Institute, 1947.

(8) Lasater, J. A. Bubble Point Pressure Correlation. J. Pet. Technol. 1958, 10, 65–67.

(9) Glaso, O. Generalized Pressure-Volume-Temperature Correlations. J. Pet. Technol. 1980, 32, 785–795.

(10) Vasquez, M.; Begg, H. D. Correlations for Fluid Physical Property Prediction. J. Pet. Technol. 1980, 32, 968–970.

(11) Al-Marhoun, M. A. PVT Correlations for Middle East Crude Oils. J. Pet. Technol. 1988, 40, 650–666.

(12) Kartoatmodjo, R. S. T.; Schmidt, Z. New Correlations for Crude Oil Physical Properties. 1991.

(13) Dokla, M.; Osman, M. Correlation of PVT Properties for UAE Crudes (Includes Associated Papers 26135 and 26316). SPE Form. Eval. 1992, 7, 41–46.

(14) Petrovsky, G. E.; Farshad, F. F. In Pressure-Volume-Temperature Correlations for Gulf of Mexico Crude Oils, SPE Annual Technical Conference and Exhibition; Society of Petroleum Engineers, 1993.

(15) Macary, S. M.; El-Batanoney, M. H. Derivation of PVT Correlations for the Gulf of Suez Crude Oils. Sekiyu Gakkaisi 1993, 36, 472–478.

(16) Omar, M. I.; Todd, A. C. In Development of New Modified Black Oil Correlations for Malaysian Crudes, SPE Asia Pacific Oil and Gas Conference; Society of Petroleum Engineers, 1993.

(17) De Ghette, G.; Paone, F.; Villa, M. Reliability Analysis on PVT Correlations. Presented at the European Petroleum Conference Held in London, U.K., October 25–27, 1994; Paper SPE-28904.

(18) Frashad, F.; LeBlanc, J. L.; Garber, J. D.; Osorio, J. G. In Empirical PVT Correlations for Colombian Crude Oils, SPE Latin America/Caribbean Petroleum Engineering Conference; Society of Petroleum Engineers, 1996.

(19) Almeiahed, R. A. In Improved PVT Correlations for UAE Crude Oils, Middle East Oil Show and Conference; Society of Petroleum Engineers, 1997.

(20) Hanafi, H. H.; Macary, S. M.; EI Nady, Y. M.; Bayomi, A. A.; El Batanoney, M. H. In Empirical PVT Correlations Applied to Egyptian Crude Oils Exemplify Significance of Using Regional Correlations, International Symposium on Oilfield Chemistry; Society of Petroleum Engineers, 1997.

(21) Gharbi, R. B.; Elsharkawy, A. M.; Karkoub, M. Universal Neural-Network-Based Model for Estimating the PVT Properties of Crude Oil Systems. Energy Fuels 1999, 13, 454–458.

(22) Al-Shammai, A. A. In Bubble Point Pressure and Oil Formation Volume Factor Correlations, Middle East Oil Show and Conference; Society of Petroleum Engineers: Bahrain, 1999; p 17.

(23) Dindoruk, B.; Christman, P. G. In PVT Properties and Viscosity Correlations for Gulf of Mexico Oils, SPE Annual Technical Conference and Exhibition; Society of Petroleum Engineers: New Orleans, Louisiana, 2001; p 14.

(24) Mehram, T.; Movagharnejad, K.; Didanloo, A. In New Correlation for Estimation of Formation Volume Factor and Bubblepoint Pressure for Iranian Oil Fields, 1st Iranian Petroleum Engineering Conference, 2006.

(25) Bolandrazadeh, A.; Hashemi, S.; Solgani, B. In The New PVT Generated Correlations of Iranian Oil Properties, 4th Iranian Petroleum Engineering Student Conference, 2006.

(26) Malallah, A.; Gharbi, R.; Algharaib, M. Accurate Estimation of the World Crude Oil PVT Properties Using Graphical Alternating Conditional Expectation. Energy Fuels 2006, 20, 688–698.

(27) Hemmati, M. N.; Kharrat, R. In A Correlation Approach for Prediction of Crude Oil PVT Properties, SPE Middle East Oil and Gas Show and Conference; Society of Petroleum Engineers, 2007.

(28) Mazandarani, M. T.; Agha, M. S. In Correlations for Predicting Solution Gas-Oil Ratio, Bubble Point Pressure and Oil Formation Volume Factor at Bubblepoint of Iran Crude Oils, European Congress of Chemical Engineering, Copenhagen, 2007.

(29) Khamechini, E.; Rashidi, F.; Rasouli, H.; Ehramian, A. Novel Empirical Correlations for Estimation of Bubble Point Pressure, Saturated Viscosity and Gas Solubility of Crude Oils. Pet. Sci. 2009, 6, 86–90.

(30) Dutta, S.; Gupta, J. P. PVT Correlations of Indian Crude Using Support Vector Regression. Energy Fuels 2009, 23, 5483–5490.

(31) Patil, P. A.; Bai, M. X.; Teodoriu, C.; Reinicke, K. M. Development of PVT Correlations According to Geography. Pet. Sci. Technol. 2014, 32, 991–999.

(32) Gomaa, S. New Bubble Point Pressure Correlation for Middle East Crude Oils. Int. Adv. Res. J. Sci. Eng. Technol. 2016, 3, 1–9.

(33) Alakbari, F. S.; Elkhatatny, S.; Baarimah, S. O. In Prediction of Bubble Point Pressure Using Artificial Intelligence AI Techniques, SPE Middle East Artificial Lift Conference and Exhibition; Society of Petroleum Engineers: Manama, Kingdom of Bahrain, 2016; p 9.

(34) Sharrad, M.; Abd-Alrahman, H. H. New Derived Correlations for Libyan Crude Oil to Estimate Bubble-Point Pressure. Sci. J. Appl. Sci. Sabratha Univ. 2019, 2, 1–13.

(35) Seyyedattar, M.; Ghasi, M. M.; Zendebhoudi, S.; Butt, S. Determination of Bubble Point Pressure and Oil Formation Volume Factor: Extra Trees Compared with LSSVM-CSA Hybrid and ANFIS Models. Fuel 2020, 269, No. 116834.

(36) Talekhiakh, M.; Amar, M. N.; Nasiri, A.; Humand, M.; Hemmati-Saraparodeh, A.; Dabir, B.; Seghier, M. E. A. Ben. Experimental Measurement and Compositional Modeling of Crude Oil Viscosity at Reservoir Conditions. J. Taiwan Inst. Chem. Eng. 2020, 109, 35–50.

(37) Tariq, Z.; Mahmoud, M.; Abdulameer, A. Machine Learning-Based Improved Pressure–Volume–Temperature Correlations for Black Oil Reservoirs. J. Energy Resour. Technol. 2021, 143, No. 113003.

(38) Fan, D.; Sun, H.; Yao, J.; Zhang, K.; Yan, X.; Sun, Z. Well Production Forecasting Based on ARIMA-LSTM Model Considering Manual Operations. Energy 2021, 220, No. 119708.

(39) Iqbal, Z.; Mahmoud, M.; Abdulraheem, A. Machine Learning-Based Correlation for Prediction of Bubble Point Pressure Using Artificial Intelligence. Energies 2020, 13, 35.

(40) Tukey, J. W. Exploratory Data Analysis; Addison-Wesley Publishing Company: Reading, MA, 1977; Vol. 2.

(41) Karim, F.; Majumdar, S.; Darabi, H.; Chen, S. LSTM Fully Convolutional Networks for Time Series Classification. IEEE Access 2018, 6, 1662–1669.

(42) Xiang, Z.; Yan, J.; Demir, I. A Rainfall-runoff Model with LSTM-based Sequence-to-sequence Learning. Water Resour. Res. 2020, 56, e2019WR025326.

(43) Pysal, D.; Abdulkadir, S. J. H.; Shukri, S. R. M.; Alhussain, H. Classification of Children’s Drawing Strategies on Touch-Screen of Seriesation Objects Using a Novel Deep Learning Hybrid Model. Alexandria Eng. J. 2020, 115–129.

(44) Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. Neural Comput. 1997, 9, 1735–1780.

(45) Schmidhuber, J. Deep Learning in Neural Networks: An Overview. Neural Networks 2015, 61, 85–117.

(46) LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. Nature 2015, 521, 436–444.

(47) Saltelli, A. Sensitivity Analysis for Importance Assessment. Risk Anal. 2002, 22, 579–590.
(48) Bahremand, A.; De Smedt, F. Distributed Hydrological Modeling and Sensitivity Analysis in Torysa Watershed, Slovakia. *Water Resour. Manage*. 2008, 22, 393−408.

(49) Osman, E. S. A.; Ayoub, M. A.; Aggour, M. A. *Artificial Neural Network Model for Predicting Bottomhole Flowing Pressure in Vertical Multiphase Flow*; Society of Petroleum Engineers, 2005.

(50) Ayoub, M. A.; Zainal, S. N.; Elhaj, M. E.; Ku Ishak, K. E. H.; Ahmed, Q. In *Revisiting the Coefficient of Isothermal Oil Compressibility Below Bubble Point Pressure and Formulation of a New Model Using Adaptive Neuro-Fuzzy Inference System Technique*, International Petroleum Technology Conference; International Petroleum Technology Conference, 2020.

(51) Alakbari, F. S.; Mohyaldinn, M. E.; Ayoub, M. A.; Muhsan, A. S.; Hussein, I. A. A Robust Fuzzy Logic-Based Model for Predicting the Critical Total Drawdown in Sand Production in Oil and Gas Wells. *PLoS One* 2021, 16, No. e0250466.