Applications of Artificial Intelligence to Predict Oil Rate for High Gas–Oil Ratio and Water-Cut Wells

Ahmed Farid Ibrahim, Redha Al-Dhaif, Salaheldin Elkatatny,* and Dhafer Al Shehri

ABSTRACT: Measuring oil production rates of individual wells is important to evaluate a well’s performance. Multiphase flow meters (MPFMs) and test separators have been used to estimate well production rates. Due to economic and technical issues with MPFMs, especially for high gas–oil ratio (GOR) reservoirs, the use of a choke formula for estimating well production rate is still popular. The objective of this study is to implement different artificial intelligence (AI) techniques to predict the oil rate through wellhead chokes. Support-vector machine (SVM) and random forests (RF) were used to generate different models to predict the production rates for high GOR and WC wells. A set of data (548 wells) was obtained from oil fields in the Middle East. GOR varied from 1000 to 9351 scf/stb, and WC ranged from 1 to 60%. Around 300 wells were flowing under critical flow conditions, while the rest were subcritical. Hence, two cases were studied using each AI model. Seventy percent of the data was used to train both RF and SVM models, while 30% of the data was used to test and validate these models. The developed RF and SVM models were then compared against the previous empirical formulas. The RF model in both critical and subcritical flow conditions was able to perfectly match the actual oil rates. SVM was able to predict the general trend for the oil rates but missed some of the sharp changes in the oil rate trend. The average absolute percent error (AAPE) values in the subcritical flow were 0.7%, while in the critical flow, the AAPE values were 1.4 and 0.75% for SVM and RF models, respectively. SVM and RF models outperform the published formulas by 34%. The results from this study will help to estimate the real-time oil and gas rates based on the available data from wellhead chokes without the need for field intervention.

1. INTRODUCTION

In oil and gas fields, a well produces three phases of oil, gas, and water. Measuring the different rates in this mixture is difficult, especially for high gas–oil ratio (GOR) and WC conditions. The measurements of oil rates are important for well performance evaluation and designing field operations such as oil recovery calculations, reservoir performance monitoring, and piping design of surface facilities.\(^5\)\(^6\)

Test separators are the most conventional and reliable method for well testing. Test separators can be two phase or three phase or horizontal, vertical, or spherical and equipped with different meters to measure the oil, gas, and water rates. As the test separators are usually portable, the transportation issue is one of the main limitations for the use of test separators. Besides, their big size that requires a large footprint on-site and longer retention time before measurements also limit the continuous use of test separators to measure the oil rates.\(^5\)\(^6\)\(^7\)

Multiphase flow meters (MPFMs) were introduced to overcome the test separator limitations. MPFMs depend on inline measuring of oil, gas, and water flow rates without any physical separation of each phase.\(^5\)\(^6\)\(^7\) Their operating principle relies on using the mass flow and volume fractions to determine the mass flow rate of oil, gas, and water. MPFMs have gone through different design improvements to enhance their measurement accuracy. However, the challenge associated with MPFMs is the accuracy of the reading for fluids with a high gas–oil ratio (GOR) or high gas volume fraction (GVF). The MPFM accuracy decreases as the GVF increases more than 85%.\(^8\) In addition, Ganat et al. compared MPFM measurements vs testing separators, and they found that MPFMs were unreliable in a high water-cut environment with an error that can reach 50% at a water cut of 61%.\(^9\) Moreover, accurate pressure volume temperature (PVT) properties of each fluid from each reservoir are needed for MPFM calibration. Furthermore, it is not feasible sometimes to install MPFMs in a location (onshore or offshore) due to logistical constraints and financial priorities.\(^5\)

Due to the economic and technical issues for test separators and MPFMs, the use of a choke formula for estimating well...
production rates is still popular. Wellhead chokes are regularly installed in almost all wells to control flow rates to the desired levels and control production drawdown to avoid formation damage and surface equipment's slugging.\(^\text{10}\) Two-phase flow through the choke can be classified as sonic (critical) or subsonic (subcritical) flow conditions based on the downstream pressure to an upstream pressure ratio.\(^\text{11}\) At critical flow conditions, downstream pressure changes do not affect the production rate, which is a function of only GOR, upstream pressure, and choke size.\(^\text{12,13}\) Gilbert\(^\text{14}\) proposed an empirical correlation based on production well-test data and analyzed 260 test data sets for different choke sizes to predict production rates at critical flow conditions.\(^\text{14}\) Several studies were developed to provide different correlations with global purposes or targeting specific fields. Most of the available formulas are unpretentious based on Gilbert-type formulas while neglecting the differential pressure across the choke.\(^\text{15−18}\) Espinoza came up with a modified empirical correlation to calculate and predict the liquid rate using choke size, upstream wellhead pressure, and gas−oil ratio only in oil fields with stable water cut and naturally flowing wells. He used Gilbert\(^\text{14}\) and Ros correlations and modified their forms. A new empirical coefficient to the equation was introduced to match the historical production rate data of the studied field in the Caspian Sea. However, this coefficient needs to be recalculated every time when a new test is available.\(^\text{19}\) Therefore, an alternative solution needs to be adopted and applied to resolve this challenge. Artificial intelligence (AI) is
one of those top solutions. Artificial intelligence (AI) and machine learning have been used for different applications in oil and gas fields.21–23

AlAjmi et al.24 applied the artificial neural network (ANN) technique on 704 test data sets from 31 wells to predict the flow rate as a function of upstream wellhead pressure, upstream wellhead temperature, gas–oil ratio, water cut, and choke size. The mean absolute percentage error (MAPE) was 13.92 and 16.39% for critical and subcritical flow models, respectively.24 Khan et al.25 applied support-vector machine (SVM) and ANN to estimate oil flow rates in artificially gas-lift wells using 1400 separator production test data points. The SVM model had a $R^2$ of 0.96 and an average absolute percent error (AAPE) of about 4%. However, the AI models predicted oil rates with an inconsistency in the range of 60–70 bpd stb/d.25 Elhaj et al.26 applied the SVM technique to estimate the gas flow rate through chokes using 162 data points from a gas field in Sudan. The optimum model with a training/testing data ratio of 70/30 had an AAPE of 1.7%.26 Moreover, different studies applied AI techniques to predict the oil rates; however, the GOR was low (175 scf/stb) or was not specified.27,28

Measuring the oil rate using test separators and MPFM is expensive and includes field interventions. The available empirical correlations are either simple and based on the Gilbert14 form and not applicable for subcritical flow conditions. The literature provides the application of AI to predict the oil rate from the choke parameters. However, to the best of the authors’ knowledge, these models were applied for a low GOR flow system with some inconsistency in the predicted oil rates. Hence, the current study presents the application of two AI techniques (random forests (RF) and SVM) to predict the oil rate from the choke parameters. The models are applicable for both critical and subcritical flow conditions and at high WC and GOR flow systems. In this study, the oil production rate was predicted as a function of GOR, WC, upstream and downstream pressures, and choke size. A total of 548 data points from an oil field in the Middle East were used to build and validate the AI models with GOR up to 10 000 scf/stb and WC up to 60%.

2. METHODOLOGY

2.1. Data Description. A data set of 548 real field-portable testing separator data points was collected from one field. GOR ranged from 1000 to 9351 scf/stb, and WC was up to 60%. The available parameters included the total gas–oil ratio (GOR), water cut (WC), flowing wellhead pressure (upstream pressure–$P_{U}$), choke size, test separator pressure (downstream pressure–$P_{D}$), and oil flow rate ($Q_{o}$). To meet the objective of this study, rate tests with GOR less than 1000 scf/stb and data with zero WC were excluded.

The data were classified into two data sets. The subcritical flow condition data set with 248 data points had the downstream pressure to upstream pressure ratio higher than or equal to 0.5. The critical flow condition data set with 300 points had the downstream pressure to upstream pressure ratio less than 0.5. In the case of critical flow, the flow rate was independent of the downstream pressure.

Figure 1 shows a scatter matrix plot for both critical data set in orange color and subcritical data set in blue color to visualize the connections and correlations between the features. The diagonal of Figure 1 presents the data frequency for each set. Both data sets showed almost the same range for most of the variables, except GOR with a range up to 9351 scf/stb for the subcritical data set compared to 5870 scf/stb for the critical data set. Most of the parameters showed a nearly normal distribution, while GOR and WC showed a lognormal distribution with a high skewness factor. Tables 1 and 2 present statistical analysis for subcritical and critical data sets. Statistical analysis includes minimum, maximum, mean, median, mode, skewness, kurtosis, and sample standard deviation. The minimum GOR in the two sets was 1000 scf/stb, while the maximum GOR was 5870 and 9351 scf/stb for subcritical and critical flow conditions, respectively. High values of the skewness and kurtosis coefficients for GOR and WC represent an asymmetric distribution for these two variables with most of the data shifted to the lower end.

To examine the relative influence of the input variables on the dependent variable (oil rate), the nonparametric Spearman’s correlation coefficient ($R$) was calculated using the following equation.

$$R = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n - 1 \sigma_x \sigma_y} \quad (1)$$

where $R$ is the correlation coefficient between the dependent (output = oil rate) and independent (input parameters) parameters; $x_i$ is the independent parameter, which includes GOR, WC, $P_{U}$, choke size, and $P_{D}$; $y_i$ is the dependent parameter (oil rate); $\mu_x$ and $\mu_y$ are the mean for the independent and dependent parameters, respectively; and $\sigma_x$ and $\sigma_y$ are the standard deviation for the independent and dependent parameters, respectively. Figure 2 presents the
correlation coefficients between all of the parameters with the oil rate for both subcritical and critical data sets. It ranged from −1 for the strong inverse relationship between the parameter and the oil rate to 1 for a strong direct relationship. Both critical and subcritical data sets show similar correlation sign with the different features, with a −ve correlation coefficient for WC and GOR and +ve for the rest of the parameters. Choke size showed the highest correlation coefficient, indicating a strong direct relationship between choke size and oil rate. $P_U$ and $P_D$ showed a low correlation with oil rate.

2.2. Model Development. MATLAB software was used to develop the SVM model, while the RF model was built using scikit-learn in Python.

The RF model is made up of decision trees that achieve high performance in a low-dimensional data set. RF overcomes the large variance and the overfitting in the decision tree method by building hundreds or thousands of decision trees on different bootstrapped data to effectively reduce the variance and improve the power capability of the RF algorithm. In addition, at each step in the tree building process, a limited number of features are arbitrarily selected using cross-validation. This helps to decorrelate the input trees, which enhances the model accuracy. RF is an algorithm where different bootstrap samples (i.e., random samples of the training sample) are used to train different trees. In addition to bootstrap samples, RF uses subsets of features for training the individual trees. This will lead to faster training and better predictive performance due to a better variance-bias trade-off. The RF technique was applied on the data to develop

![Figure 2. Correlation coefficients between the oil rates and the different independent parameters for both critical and subcritical data sets.](image)

![Figure 3. Flowchart for building the different RF and SVM models for critical and subcritical flow conditions.](image)

**Table 3. $R^2$ as a Function of the Input Parameters for the Subcritical Case**

| input parameters | $R^2$ training |
|------------------|----------------|
| GOR, $P_U$, choke size, $P_D$ | 0.91 |
| GOR, WC, $P_U$, choke size | 0.89 |
| GOR, $P_U$, choke size, $P_D$ | 0.77 |
| GOR, WC, $P_U$, $P_D$ | 0.23 |
| GOR, WC, choke size, $P_D$ | 0.85 |
| WC, $P_U$, choke size, $P_D$ | 0.87 |

![Figure 4. Cross-plots of actual and predicted oil rates for the SVM model in the case of subcritical conditions: training (a) and testing (b).](image)
a model to predict the oil rate. Different combinations of RF parameters were used. These parameters include the maximum depth of the tree, the maximum features to be considered when splitting the node in each tree, and the number of trees in the forest (N of estimators). N of estimators varied from 3 to 150, and the maximum depth had values varied from [3, 4, 5, ..., 30]. The maximum feature varied among three different features ["auto", "sqrt", "log 2"].

The support-vector machine (SVM) technique was developed by Vapnik in 1995. SVM is a supervised learning method with an associated learning algorithm that analyzes data and recognizes patterns of input/output data. It is a tool used for classification and regression tasks. The SVM technique builds the input prototypes in a space with greater dimensions by employing a nonlinear mapping method. The SVM finds specific linear models between two different classes and orientates them in a way that maximizes the margin of the separator hyperplane. The nearest training data points, used to define the margin, are called support vectors. Different SVM parameters were optimized such as regularization parameter C, γ, and kernel type to improve the performance of the SVM.

Figure 3 shows the flowchart used to develop the RF and SVM models. For each model, two cases were created: subcritical and critical with 248 and 300 data points, respectively. The data were classified into training and testing (includes validation). The percentage distribution of each group was determined based on trials and errors. The optimized percentages were 70 and 30% for training and testing, respectively. For critical condition models, the input parameters to the model were upstream pressure, GOR, WC, and choke size as independent variables. Downstream pressure was added in the cases of subcritical conditions. The output

|      | R^2 training | R^2 testing | AAPE training (%) | AAPE testing (%) |
|------|--------------|-------------|-------------------|------------------|
| RF   | 0.98         | 0.91        | 0.6               | 1.3              |
| SVM  | 0.90         | 0.88        | 1.7               | 1.8              |

Figure 5. Cross-plots of the actual and predicted oil rates for the SVM model in the case of subcritical conditions: training (a) and testing (b).

Table 4. R^2 and AAPE Summary for the Subcritical Case

Figure 6. Actual vs predicted oil flow rate under subcritical conditions for RF (a) and SVM (b) models for all data set.
parameter is the oil rate from each model as the dependent variable.

The assumptions for the model development are as follows: the data are classified into critical and subcritical flow conditions based on the downstream to the upstream pressure ratio of 0.5, the downstream pressure is neglected for critical flow conditions, and the model is applicable for high GOR flow streams with GOR from 1000 to 9351 scf/stb and WC from 1 to 60%.

The average absolute percentage error (AAPE) and coefficient of determination ($R^2$) were used to evaluate the developed model for each case. The average absolute percentage error (AAPE) was used for the error analysis that can be calculated as follows:

$$\text{AAPE} = \frac{\sum_{i=1}^{N} |Q_{o,\text{Actual}} - Q_{o,\text{Predicted}}|}{N} \times 100\%$$

where $Q_{o,\text{Actual}}$ and $Q_{o,\text{Predicted}}$ are the actual and predicted oil rates, respectively, and $N$ is the number of data points.

$R^2$ was used to show the goodness of fit, which can be calculated from eq 3.

$$R^2 = 1 - \frac{\left[ \sum_{i=1}^{N} (Q_{o,\text{Actual}} - Q_{o,\text{Predicted}})^2 \right]}{\left[ \sum_{i=1}^{N} (Q_{o,\text{Actual}} - \bar{Q})^2 \right]^{1/2}}$$

where $Q_{o,\text{Actual}}$ and $Q_{o,\text{Predicted}}$ are the actual and predicted oil rates, respectively, and $N$ is the number of data points.

AAPE and $R^2$ were reported for each SVM and RF model in both critical and subcritical cases and then were compared with existing correlations, namely, Gilbert, Baxendell, Ros, and Achong.14,16,18,20 These empirical correlations are explained in detail in Appendix A.

3. RESULTS AND DISCUSSION

3.1. Case 1: Subcritical Flow. The input parameters for this model were GOR, WC, $P_{U}$, choke size, and $P_{D}$, while $Q_{Oil}$ was the output parameter, and the number of data points was 248 points.

The RF technique was applied on the data to develop a model to predict the oil rate under subcritical conditions. Different combinations of RF parameters were used. These parameters include the maximum depth of the tree, the maximum features to be considered when splitting the node in each tree, and the number of trees in the forest ($N$ of estimators). $N$ of estimators varied from 3 to 150, and the optimum value used in this case was 100. The maximum depth had values that varied from [3, 4, 5, ..., 30], and the optimum maximum depth was found to be 25. The optimum maximum feature was $(\sqrt{2})$ out of three different features [auto, sqrt, log2].

To investigate the sensitivity of the model output to the different input parameters, the model was run by excluding the input parameters one by one from the model. Initially, GOR, WC, $P_{U}$, choke size, and $P_{D}$ were used as inputs, and then one parameter was removed each time. Table 3 summarizes the change in $R^2$ on removing one of the input parameters. The results from the sensitivity analysis are in agreement with the correlation coefficient calculated in Figure 2. The choke size showed the most effective input in the oil rate, and the downstream pressure still had an effect on the results but with the lowest degree.

Figure 4 shows a cross-plot between the actual and predicted oil rates in the case of training and testing data sets. The RF model was able to accurately predict the oil rate from the choke parameters, where most of the data points aligned to the 45° line for the training data set. However, its accuracy decreased for the testing data set with an $R^2$ of 0.91. The RF model had an $R^2$ of 0.98 and 0.91 for training and testing data sets, respectively. AAPE for the RF model was estimated to be 0.6% and 1.3% for training and testing data sets, respectively, with an average of 0.7% for all data.

Similarly, the SVM technique was applied on the data. Different SVM parameters were used such as the regularization parameter $C$, $\gamma$, and kernel type to calculate the performance of the SVM. The optimum results in this model were obtained...
using lambda = 1 × 10^{-2}, epsilon = 0.01, kerneloption = 2, verbose = 1, and C = 500 with Kernel “Gaussian”. Figure 5 shows a cross-plot between the actual and the predicted data from the SVM model for the training and testing data sets. The developed SVM model had an $R^2$ of 0.90 and 0.88 for training and testing data sets, respectively. AAPE for the SVM model was estimated to be 1.7 and 1.8% for training and testing data sets, respectively, with an average of 1.7% for all data. The SVM model was slightly less performing than the RF model as shown in Table 4.

Figure 6 presents the oil flow rate obtained from the SVM and RF models plotted against the actual flow rate from well testing. The two graphs are almost identical in the case of the RF model, which confirms the reliability of the RF model in oil flow rate estimation for production under subcritical flow conditions with a high GOR up to 5870 scf/stb. The SVM model was able to capture the general trend of the oil rates but slightly missed some of the sharp changes in the oil rate.

### 3.2. Case 2: Critical Flow

The input parameters for the critical flow case exclude the downstream pressure effect, and the main inputs were GOR, $P_U$, and choke size while $Q_{OS}$ was the output parameter. The total number of points used in this

---

**Figure 8.** Cross-plots of actual and predicted oil rates for the RF model in the case of critical flow conditions: training (a) and testing (b).

**Table 6. $R^2$ and AAPE Summary for the Critical Case**

|       | $R^2$ training | $R^2$ testing | AAPE training (%) | AAPE testing (%) |
|-------|----------------|---------------|------------------|------------------|
| RF    | 0.98           | 0.91          | 0.6              | 1.3              |
| SVM   | 0.94           | 0.93          | 1.0              | 1.8              |

**Figure 9.** Actual vs estimated oil flow rate under critical flow conditions for the RF (a) and SVM (b) models for all data.

**Figure 10.** SVM and RF models vs existing correlations for critical flow.
case was 300 data points. The maximum GOR for this set was 9351 scf/stb and water cut was up to 42%.

The RF technique was applied on the choke data to estimate the oil rate under critical conditions. The ratio of the number of data points in the training and testing data sets was 70:30%. The optimum RF parameters for this case were N of estimators of 150, the maximum depth of 17, and the optimum maximum feature was (auto).

Similar to the subcritical condition, the sensitivity of input parameters was studied for the critical flow conditions. Initially, GOR, WC, \( P_U \), choke size, and \( P_D \) were used as inputs, and then one parameter was removed at each time. Table 5 summarizes the change in \( R^2 \) on removing one of the input parameters. The results from the sensitivity analysis are in agreement with the correlation coefficient calculated in Figure 2. The downstream pressure has no effect on the model development, so it was removed from the model development.

The choke size showed the most effective input in the oil rate. Figure 7 shows a cross-plot between the actual and predicted oil rates from RF model. Most of the data aligned around the 45° line. \( R^2 \) between the developed RF model and the actual data was found to be 0.98 and 0.94 for training and testing data sets, respectively. AAPE for the RF model was estimated to be 0.6 and 1.3% for training and testing data sets, respectively, with an average of 0.75% for all data.

Similarly, the SVM model was applied on the choke data with flow under critical conditions. The optimum results in this model were obtained using \( \text{lambda} = 1 \times 10^{-3} \), \( \epsilon = 0.01 \), \( \text{kerneloption} = \text{S} \), \( \text{verbose} = 10 \), and \( C = 800 \) with Kernel “JCB”. Figure 8 shows the cross-plot and model coefficient of determination \( (R^2) \) values for the training and testing data. \( R^2 \) between the developed SVM model and the actual oil data was estimated to be 0.94 and 0.93 for training and testing data sets, respectively. AAPE for the SVM model was 1.0 and 1.8% for training and testing data sets, respectively, with an average of 1.4% for all data. Table 6 summarizes the performance evaluation parameters, where the SVM model shows lower accuracy compared to the RF model, with an average AAPE for all data of 1.4% compared to 0.7% for the RF model.

The oil flow rates obtained from the developed SVM and RF models were plotted against the actual flow rate from well testing as shown in Figure 9. The RF model was able to accurately predict the oil rates at different conditions with the actual and predicted oil graphs almost identical, which confirms the reliability of the RF model in oil flow rate estimation in the case of critical fluid flow conditions with high GOR between 1000 and 9351 scf/stb and WC up to 42%. The SVM model missed some of the actual oil rates with the ability to predict the general trend of oil rate values.

SVM and RF models were compared with empirical correlations: Gilbert, Baxendell, Ros, and Achong.\(^{14,16,18,20}\) As demonstrated in Figure 10, the SVM and RF models have shown higher performance as compared with all of these correlations. The AAPE values for the SVM and RF models are about 1.4—0.7%, while the correlation AAPEs are 43.19, 34.67, 35.14, and 37.02% for Gilbert, Baxendell, Ros, and Achong, respectively. The possible reasons for the low behavior of the empirical correlations are that these correlations are based on curve fitting that depends on a theory of cause and effect underlying the data, which usually include random or systematic errors, while the current study is based on the machine learning technique, where the ML technique discovers the relations through data mining. Moreover, the published empirical correlations were developed from data sets with lower data points and narrow input ranges compared with the current study (details in Table A2 in Appendix A).

The results showed that in both critical and subcritical conditions, the chock size was the more dominant parameter affecting the predicted oil rate followed by the WC and GOR. Removing the downstream pressure from the input data for the critical flow did not affect the prediction of the oil rate from the chock parameters.

The developed models were able to predict the surface oil rate from the choke parameters. It should be highlighted that it is recommended to employ the developed models using choke parameters within the same model’s input ranges to ensure trustworthy results, and a validation step should be conducted first. Moreover, further validation for the developed models with different field data will be conducted in future work.

### 4. CONCLUSIONS

This study presents the application of the SVM and RF techniques to predict oil rates from the wellhead choke performance as a function of upstream and downstream pressures, GOR, and chock size in the case of high GOR and high WC reservoirs. The following are the main conclusions:

1. RF models were able to accurately predict oil rates in critical and subcritical flow cases with AAPE less than 1%.
2. AAPE from SVM in the case of critical and subcritical flows with GOR up to 9000 scf/stb was less than 1.4%.
3. SVM and RF models outperform the published empirical correlations.
4. RF models perform slightly better than SVM models in both critical and subcritical flow conditions.

This study showed the capability and robustness of AI to estimate surface oil flow rates with about ±1.4% accuracy error and goodness of fit \( (R^2) \) of 0.93—0.98.

### APPENDIX A

Gilbert in 1954 developed the general correlation form to describe the critical flow as shown in eq A1.

\[
R_U = \frac{C \text{GOR}^m Q_{oil}^n}{\text{Chk}^m}
\]

(A1)

where GOR is the gas—oil ratio, \( Q_{oil} \) is the oil rate, Chk is the choke size, \( R_U \) is the upstream pressure, and the constants \( C, m, \) and \( n \) are correlation constants. Different authors have then modified the Gilbert constant based on their oil field properties. Table A1 shows the correlation constant for each correlation. Moreover, Table A2 summarizes some details about each correlation including the number of data points used to generate the correlations, input parameters, and some of the parameter ranges.

A modified correlation can be used for subcritical flow conditions by changing the upstream pressure to pressure difference as follows

\[
\Delta P = \frac{C \text{GOR}^m Q_{oil}^n}{\text{Chk}^m}
\]

(A2)

where \( \Delta P = P_U - P_D \) and \( P_D \) is the downstream pressure. 
Table A1. Empirical Correlation Constants for Gilbert, Baxendell, Ros, and Achong Correlations

| correlation | C   | m   | n   |
|------------|-----|-----|-----|
| Gilbert    | 10  | 0.546 | 1.84 |
| Baxendell  | 9.56 | 0.546 | 1.93 |
| Ros        | 17.40 | 0.500 | 2   |
| Achong     | 3.82  | 0.650 | 1.88 |

Table A2. Summary of Empirical Correlation Input Parameters and Some of the Parameter Ranges for Gilbert, Baxendell, Ros, and Achong Correlations

| correlation | no. of data points | input parameters | parameter range |
|------------|--------------------|-----------------|----------------|
| Gilbert    | 260                | GLR, $P_{upr}$  | $D_{upr} = 6/64$ to 18/64 in. |
| Baxendell  |                    | GLR, $P_{upr}$  |                  |
| Ros        | 104                | GLR, $P_{upr}$  | $D_{upr} = D$ = 1/2 − 4 in. |
| Achong     |                    | GLR, $P_{upr}$  |                  |

REFERENCES

1. Farid, A. M. M.; El-Banbi, A. H.; Abdelwaly, A. A. A. An Integrated Model for History Matching and Predicting Reservoir Performance of Gas/Condensate Wells. SPE Reservoir Eval. Eng. 2013, 16, 412–422.
2. Ghorbani, H.; Wood, D. A.; Choubineh, A.; Mohamadian, N.; Tatar, A.; Farhangian, H.; Nikooye, A. Performance comparison of bubble point pressure from oil PVT data: Several neurocomputing techniques compared. Exp. Comput. Multiphase Flows 2020, 2, 225–246.
3. Ibrahim, A. F.; Assem, A.; Ibrahim, M. A novel workflow for water flowback RTA analysis to rank the shale quality and estimate fracture geometry. J. Nat. Gas Sci. Eng. 2020, 81, No. 103387.
4. Selim, I. E.-S.; Shokir, E. M. In Tracking Subsea Gas Wells Performance without Periodic Production Testing on Test Separator, North Africa Technical Conference and Exhibition, SPE-152768-MS, Cairo, Egypt, February, 2012.
5. Coimbra, A.; Punet, E. In Flow Rate Measurement Using Test Separator and PDG Data Allows Individual and Commingled Production Zone Flow Rate History Calculation, OTC-27963-MS, OTC Brasil, Rio de Janeiro, Brazil, October, 2017.
6. Alsalman, A.; Almutairi, A.; Alyed, S.; Kumar, V. In Firsttime Utilization of Portable Multiphase Flow Meter for Testing Offshore Wells in Saudi Arabia, SPE Middle East Oil & Gas and Show Conference, SPE-172696-MS, Manama, Bahrain, March, 2015.
7. Falcone, G.; Hewitt, G. F.; Alimonti, C.; Harrison, B. In Multiphase Flow Metering: Current Trends and Future Developments, SPE Annual Technical Conference and Exhibition, SPE-71474-MS, New Orleans, Louisiana, September, 2001.
8. Nasri, A.; Al-Anizi, A.; Al-Amri, M. A.; Al-Khelaib, F. T.; Al-Anazi, A. In MultiphaseFlow Meters Trial Testing in High GOR/GVF Environment, International Petroleum Technology Conference, IPTC-17422-MS, Doha, Qatar, January, 2014.
9. Ghorbani, H.; Wood, D. A.; Choubineh, A.; Mohamadian, N.; Tatar, A.; Farhangian, H.; Nikooye, A. Performance comparison of bubble point pressure from oil PVT data: Several neurocomputing techniques compared. Exp. Comput. Multiphase Flows 2020, 2, 225–246.
(17) Fortunati, F. In *Two-Phase Flow through Wellhead Chokes*, SPE European Spring Meeting, SPE-3742-MS, Amsterdam, Netherlands, May, 1972.

(18) Baxendell, P. B. Producing Wells on Casing Flow - An Analysis of Flowing Pressure Gradients. *Trans. AIME* 1958, 213, 202–206.

(19) Espinoza, R. In *Digital Oil Field Powered with New Empirical Equations for Oil Rate Prediction*, SPE Middle East Intelligent Oil and Gas Conference and Exhibition, SPE-176750-MS, Abu Dhabi, UAE, September, 2015.

(20) Achong, I. *Revised Bean Performance Formula for Lake Maracaibo Wells*, Internal Company Report; Shell Oil Co.: Houston, TX, USA, 1961.

(21) Hagan, M. T.; Menhaj, M. B. Training feedforward networks with the Marquardt algorithm. *IEEE Trans. Neural Networks* 1994, 5, 989–993.

(22) Jang, J. R.; Chuen-Tsai, S. Neuro-fuzzy modeling and control. *Proc. IEEE* 1995, 83, 378–406.

(23) Ahmed, A. A.; Elkhatemy, S.; Abdulraheem, A.; Mahmoud, M. In Application of Artificial Intelligence Techniques in Estimating Oil Recovery Factor for Water Derive Sandy Reservoirs, SPE Kuwait Oil & Gas Show and Conference, Kuwait City, Kuwait, 2017.

(24) AlAjmi, M. D.; Alarifi, S. A.; Mahsoon, A. H. In *Improving Multiphase Choke Performance Prediction and Well Production Test Validation Using Artificial Intelligence: A New Milestone*, SPE Digital Energy Conference and Exhibition, SPE-173394-MS, The Woodlands, Texas, USA, March, 2015.

(25) Khan, M. R.; Tariq, Z.; Abdulraheem, A. In *Utilizing State of the Art Computational Intelligence to Estimate Oil Flow Rate in Artificial Lift Wells*, SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition, Dammam, Saudi Arabia, April, 2018.

(26) Elhaj, M. A.; Anifowose, F.; Abdulraheem, A. In *Single Gas Flow Prediction through Chokes Using Artificial Intelligence Techniques*, SPE Saudi Arabia Section Annual Technical Symposium and Exhibition, Al-Khobar, Saudi Arabia, April, 2015.

(27) Al Kadem, M.; Al Dabbous, M.; Al Mashhad, A.; Al Sadah, H. In *Utilization of Artificial Neural Networking for Real-Time Oil Production Rate Estimation*, Abu Dhabi International Petroleum Exhibition & Conference, SPE-197879-MS, Abu Dhabi, UAE, November, 2019.

(28) Rahmani, B.; Mohsenpour, S.; Shamshir Noghabi, H. R.; Hemmati, N.; Tabzar, A. Estimation of flow rates of individual phases in an oil-gas-water multiphase flow system using neural network approach and pressure signal analysis. *Flow Meas. Instrum.* 2019, 66, 28–36.

(29) Hegde, C.; Gray, K. Use of machine learning and data analytics to increase drilling efficiency for nearby wells. *J. Nat. Gas Sci. Eng.* 2017, 40, 327–335.

(30) Yarveicy, H.; Saghafi, H.; Ghiasi, M. M.; Mohammadi, A. H. Decision tree-based modeling of CO2 equilibrium absorption in different aqueous solutions of absorbents. *Environ. Prog. Sustainable Energy* 2019, 38, 5441–5448.

(31) Efron, B. *The Jackknife, the Bootstrap and other Resampling Plans*; SIAM, 1982.

(32) Hedge, C.; Wallace, S.; Gray, K. In “Using Trees, Bagging, and Random Forest to Predict Rate of Penetration During Drilling”, Proceeding of the SPE Middle East Intelligence Oil and Gas Conference and Exhibition held in Dubai, UAE, SPE-176792-MS, 15–16 September, 2015.

(33) Abbas, A. K.; Almubarak, H.; Abbas, H.; Dawood, J. In Application of Machine Learning Approach for Intelligent Prediction of Pipe Sticking, Abu Dhabi International Petroleum Exhibition & Conference, SPE-197396-MS, Abu Dhabi, UAE, November, 2019.

(34) Tao, G. L.; Yao, Z. S.; Tan, B. Z.; Gao, C. C.; Yao, Y. W. Application of Support Vector Machine for Prediction of Slope Stability Coefficient Considering the Influence of Rainfall and Water Level. *Appl. Mech. Mater.* 2016, 851, 840–845.