Research Article

Application of Various Machine Learning Techniques in Predicting Total Organic Carbon from Well Logs

Osama Siddig, Ahmed Farid Ibrahim, and Salaheldin Elkatatny

Department of Petroleum Engineering, King Fahd University of Petroleum and Minerals, Dhahran 31261, Box: 5049, Saudi Arabia

Correspondence should be addressed to Ahmed Farid Ibrahim; ahmed.ibrahim@kfupm.edu.sa and Salaheldin Elkatatny; elkatatny@kfupm.edu.sa

Received 7 July 2021; Revised 14 August 2021; Accepted 23 August 2021; Published 31 August 2021

Academic Editor: Gaurav Singal

Copyright © 2021 Osama Siddig et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Unconventional resources have recently gained a lot of attention, and as a consequence, there has been an increase in research interest in predicting total organic carbon (TOC) as a crucial quality indicator. TOC is commonly measured experimentally; however, due to sampling restrictions, obtaining continuous data on TOC is difficult. Therefore, different empirical correlations for TOC have been presented. However, there are concerns about the generalization and accuracy of these correlations. In this paper, different machine learning (ML) techniques were utilized to develop models that predict TOC from well logs, including formation resistivity (FR), spontaneous potential (SP), sonic transit time (Δt), bulk density (RHOB), neutron porosity (CNP), gamma ray (GR), and spectrum logs of thorium (Th), uranium (Ur), and potassium (K). Over 1250 data points from the Devonian Duvernay shale were utilized to create and validate the model. These datasets were obtained from three wells; the first was used to train the models, while the data sets from the other two wells were utilized to test and validate them. Support vector machine (SVM), random forest (RF), and decision tree (DT) were the ML approaches tested, and their predictions were contrasted with three empirical correlations. Various AI methods’ parameters were tested to assure the best possible accuracy in terms of correlation coefficient (R) and average absolute percentage error (AAPE) between the actual and predicted TOC. The three ML methods yielded good matches; however, the RF-based model has the best performance. The RF model was able to predict the TOC for the different datasets with R values range between 0.93 and 0.99 and AAPE values less than 14%. In terms of average error, the ML-based models outperformed the other three empirical correlations. This study shows the capability and robustness of ML models to predict the total organic carbon from readily available logging data without the need for core analysis or additional well interventions.

1. Introduction

Due to the continuous oil and gas exploitation, conventional hydrocarbon reserves are gradually depleted, and the production rates of the current reservoirs are significantly declining.

Conventional hydrocarbon reserves are gradually depleting due to the continuous oil and gas exploitation and the production rates of the current reservoirs dramatically dropped [1, 2]. Source rock and unconventional reservoirs have recently attracted interest as a result [3, 4]. Since unconventional resources are more complex, tight, and less permeable, unconventional reservoirs exploration is more challenging and demanding in contrast to conventional reservoirs [5]. However, considerable discoveries of unconventional resources have been announced around the globe, namely, in North and South America, Middle East, and North Africa, which represents a significant addition to the total world oil reserves [6, 7]. Unconventional resources, in contrast to conventional reservoirs, are self-storing and self-generating reservoirs; consequently, evaluating their hydrocarbon generation potential is critical. Characterization, development, and hydrocarbon extraction of unconventional resources are sophisticated and costly operations, all of which underscores the importance of evaluating the unconventional resources’ ability to generate hydrocarbons in a cost-effective and precise manner [4, 6].
Total organic carbon (TOC), which has been widely considered as a quantification of the hydrocarbon generation potentials [8–10], is one of the most efficient parameters that evaluate the quality of unconventional resources [11]. In general, the rock pyrolysis experiment is used to determine TOC content in the laboratory [12, 13]. Due to the high cost of the experiments, there is a limitation on the number of laboratory tests to measure TOC. Consequently, it is very difficult to obtain a complete TOC profile for the formations of the interest, which severely affects the reservoir evaluation [14].

Several authors developed empirical correlations to determine TOC, and these models were developed based on the experimental TOC measures on drilling cuttings or core samples and the corresponding well logs. Then, these developed correlations are proposed to be applied to determine the TOC for different wells [15–19]. These correlations are summarized in Table 1 [20–24]. One concern about these empirical correlations is the low accuracy of the predictions when used with different datasets.

Artificial intelligence (AI) has been used in different industries [25]. AI techniques are known to have the capability to generate high accuracy models; therefore, several studies utilized them in TOC prediction [26, 27]. In the appendix, Table 2 summarizes the different research studies that utilized AI techniques to estimate the TOC from well logs [8, 9, 14, 17, 28–42]. The used well logs include formation resistivity (FR), spontaneous potential (SP), sonic transit time (Δt), bulk density (RHOB), neutron porosity (CNP), gamma ray (GR), and spectral logs of thorium (Th), potassium (K), and uranium (Ur).

TOC is a vital parameter to characterize the unconventional resources. Experimental analysis can be used to measure the TOC, but it is expensive, time-consuming, and does not give a continuous profile for the total organic carbon with depth. Empirical correlations can be used to estimate TOC. However, there are concerns about the generalization and accuracy of these correlations. In this paper, the application of different AI techniques in TOC prediction in Devonian shale formation from the well logs will be tested. These well logs include sonic transient time, resistivity, bulk gamma ray, bulk density, and spectral GR logs of Th, Ur, and K and neutron log porosity. The next sections in the paper describe the methodology that was used to build the ML models, followed by building the model. The models were then tested and validated with different data sets. A sensitivity analysis was conducted to investigate the importance of input logging parameters in the predicted TOC values.

2. Methodology

In this study, three machine learning tools were used to estimate the TOC as a function of eight well logs records. Figure 1 summarizes the methodology applied for optimizing ML models and validation of these developed models.

2.1. Data Description. Experimental data for TOC from three different wells have been collected together with their corresponding well logs. 891, 291, and 82 data points from Well-A, Well-B, and Well-C were used to train, test, and validate the AI models all, respectively. All wells are in Devonian Duvernay shale, which is an organic liquid-rich source rock. The sedimentary basin is located in Alberta, Canada, with 61.7 billion barrels and 443 trillion cubic feet of oil and gas reserves, respectively [43, 44].

2.2. Core Samples Testing. Rock-Eval 6 was employed to estimate the actual TOC values of drilling cuttings from different wells. Tests procedures are shown in Figure 2. More detailed discussion about TOC experimental procedure is presented by Chen et al. [44].

2.3. Data Preprocessing. Prior to the AI model’s training, outliers, incomplete, or unrealistic data points were eliminated from the data used to construct the model. Data points that contain any value that is away from the mean of the data with three times the standard deviation were considered as an outlier. The statistical characteristics of Well-A’s dataset are illustrated in Table 3.

2.4. Model Development. The AI model was trained and optimized in this work using Well-A dataset that contains 891 data points with wide ranges of TOC and well logs values. The effect of various parameters inside the AI algorithms was tested to optimize the models, by running the AI tools inside a mutable for loops in MATLAB.

In SVM models, different kernel functions, values for kernel options, epsilon, and regularization were tested while in RF, different sets of number of trees, maximum number of level in each tree, and maximum number of features were used. In DT, various values for maximum tree depth, minimum sample split, and maximum number of features were tested. The correlation coefficient between the known and predicted TOC and the average absolute percentage error (AAPE) were used as evaluation criteria.

In addition, different sets of inputs were used to evaluate the significance of each well log parameter in TOC prediction. Seven sets were considered, the least one includes four parameters, and the most comprehensive consists of all eight parameters, as shown in Table 4.

The accuracies of the generated models were tested and validated using 291 and 82 data points from Well-B and Well-C, respectively. The two wells are in the same field as Well-A. The performance of AI models was also compared with that of currently existing correlations, such as the Schmoker model, the logR approach, and the Zhao et al. model.

3. Results and Discussion

The AI models were trained for TOC estimation based on eight well log data of RHOB, Δt, CNL, FR, GR, and spectral GR. The training dataset consisted of 891 data points from Well-A, while the testing dataset contains 291 data points from Well-C. This section presents the results obtained using each method.
3.1. Support Vector Machine. Using data set from Well-A and Well-B, different trials have been applied using support SVM with changing some tuning parameters inside the algorithm, such as kernel function and regularization. The best results were achieved using the Gaussian kernel function. Good results have been achieved in the training dataset with a 7.1% average error; however, the accuracy in the testing datasets was relatively low with an average error of 19.7%. The correlation coefficients were 0.974 and 0.856 for training and testing, respectively. Figure 3 presents the cross-plots of the actual and SVM-predicted TOC values for the training and testing data sets.

3.2. Decision Tree. This technique resulted in perfect fitting in the training dataset with a 0.994 correlation coefficient as shown in Figure 4(a). However, the prediction performance

### Table 1: Summary of different empirical correlations.

| Authors          | Model                                                                 | Remarks                                                                 |
|------------------|----------------------------------------------------------------------|------------------------------------------------------------------------|
| Schmoker         | $\text{TOC} = \frac{(\rho_B - \rho)}{1.378}$                      | Predicted TOC in volume percentage and used data from Devonian shale.  |
|                  | where densities are in g/cm$^3$                                      | $\rho$ is the formation bulk density, and $\rho_B$ represents the     |
|                  |                                                                     | organic matter free rock density.                                      |
| Schmoker         | $\text{TOC} = \frac{(100\rho_B) - (0.9922\rho_{mi} - 0.039)}{135\rho_{mi} - 0.675}$ | A revised model used Bakken shale’s data and predicted TOC in weight   |
|                  |                                                                     | percentage. $R$ is ratio between the organic material and organic      |
|                  |                                                                     | carbon. $\rho_B$ is the density of the organic organic matter,         |
|                  |                                                                     | and $\rho_{mi}$ is the average bulk density.                           |
| Passy et al.     | $\Delta \log R = \log \left( \frac{\text{BR/FR}_{\text{baseline}}}{0.02 \times (\Delta t - \Delta t_{\text{baseline}})} \right)$ | Widespread model and known as the $\Delta \log R$ model. First, the     |
|                  | $\text{TOC} = \Delta \log R \times 10^{1.297 - 0.1688 \times \text{LOM}}$ | logs separation ($\Delta \log R$) is calculated from acoustic transit  |
|                  | where resistivity is in ohm-m and transit time is in $\mu$s/ft       | time ($\Delta t$) and formation resistivity (FR). Then, TOC is        |
|                  |                                                                     | estimated from $\Delta \log R$ and the level of maturity (LOM).       |
| Wang et al.      | $\text{TOC} = [a \Delta \log R + b (\text{GR} - \text{GR}_{\text{baseline}})] \times 10^{(0.67 + \eta \times T_{\text{max}})}$ | Revised $\Delta \log R$ models to estimate the TOC developed for       |
|                  | where $a$, $b$, $\delta$, and $\eta$ are constants                 | Devonian shale using $\Delta \log R$, gamma-ray (GR), and the indicator |
|                  |                                                                     | of maturity ($T_{\text{max}}$).                                       |
| Zhao et al.      | $\text{TOC} = a \Delta \log R + b (\text{GR} - \text{GR}_{\text{baseline}}) + c$ | Revised Wang’s model. Do not depend on the level of maturity, vitrinite  |
|                  | where $a$, $b$, and $c$ are constants                               | reflectance ($R_v$), or $T_{\text{max}}$.                             |

### Table 2: Summary of different research studies that employed AI techniques to predict the TOC.

| Ref | Data sources                  | Data points | AI tools | Inputs | Accuracy ($R^2$) |
|-----|-------------------------------|-------------|----------|--------|------------------|
| [8] | Barnett and Devonian shale    | 442         | ANN      | RHOB, GR, $\Delta t$, FR | 0.89–0.93 |
| [9] | Shahejie formation            | 125         | CNN      | RHOB, GR, $\Delta t$, FR | 0.83   |
| [10] | Kazhdumi formation            | 31          | FL       | C1, K, Th | 0.94  |
| [11] | Devonian and Barnett shales   | 500         | FNN, SVM | RHOB, GR, $\Delta t$, FR | 0.74–0.77 |
| [12] | Barnett shale                 | 645         | FL       | RHOB, GR, $\Delta t$, FR | 0.91  |
| [13] | Jiunenchiung formation        | 31          | SVM      | C1, K, Th | 0.69  |
| [14] | Kangan-Dalan formation        | 124         | FL, NF, NN | RHOB, GR, $\Delta t$, FR | 0.85  |
| [15] | Kazhdumi and Kangan-Dalan formations | 78     | ANN      | GR, $\Delta t$, K, Th | 0.89  |
| [16] | Ordos basin and Canning basin | NA          | GPR      | C1, K, Th | NA   |
| [17] | Gadvan formation              | 2875        | ANN, FL  | RHOB, GR, $\Delta t$, FR | 0.78–0.99 |
| [18] | Kazhdumi and Kangan-Dalan formations | 200 | ANN | C1, K, Th | NA |
| [19] | NA                            | 70          | ANN      | $\Delta t$, K, Th | 0.98  |
| [20] | Khatata and Ras Qattara formations | 54       | ANN      | RHOB, GR, $\Delta t$, FR | 0.96  |
| [21] | Beibu Gulf basin              | 18          | SVM      | RHOB, GR, $\Delta t$, SP, FR | 0.75  |
| [22] | Sichuan basin                 | 185         | ELM, ANN | C1, K, Th | 0.87–0.91 |
| [23] | Tonghua basin                 | 215         | ANN, SVM | RHOB, GR, $\Delta t$, SP, FR, K, Th, Ur | 0.9–0.93 |
| [24] | Barnett and Duvernay shales   | 460         | ANN      | RHOB, GR, $\Delta t$, FR | 0.98  |
| [25] | Barnett shale                 | 442         | ANN      | RHOB, GR, $\Delta t$, FR | 0.93  |
| [26] | Barnett shale                 | +800        | ANFIS, FNN, SVM | RHOB, GR, $\Delta t$, FR | 0.82–0.87 |
| [27] | Qaidam basin                  | 19          | ANN      | RHOB, GR, $\Delta t$, FR | NA   |
| [28] | Bohai Bay, Sichuan, Ordos and western Canada sedimentary basins | 353 | ANN, SVM | RHOB, GR, $\Delta t$, FR | 0.89  |
Data gathering and preprocessing

Statistical analysis

Randomly divide the dataset into training and testing with 70/30 ratio

Selecting the model’s parameters for AI algorithms

Run the AI code

Calculate R and AAPE

Best outcomes?

No

Yes

Test the model with the validation dataset

Results analysis

Figure 1: Methodology flow chart.

Grinding:
samples were grinded to less than 63 μm

Thermal decomposition:
Using a pyrolysis oven

Samples burning:
Using oxidization oven at 300 °C for 30 seconds

Figure 2: TOC test procedures.

Table 3: The statistical description of the training data.

| Statistical parameter | FR (Ω·m) | $\Delta t$ (μsec/ft) | RHOB (g/cm³) | CNP | GR (°API) | Ur (wt%) | TH (ppm) | K (ppm) | TOC (wt%) |
|-----------------------|----------|----------------------|---------------|-----|-----------|----------|----------|---------|-----------|
| Minimum               | 3.71     | 51.0                 | 2.39          | 0.019 | 22.9      | 1.39     | 1.97     | 0.130   | 0.76      |
| Maximum               | 1675     | 96.6                 | 2.77          | 0.346 | 298       | 22.6     | 17.0     | 4.06    | 5.66      |
| Mean                  | 110      | 77.9                 | 2.545         | 0.174 | 95.5      | 6.16     | 9.01     | 1.51    | 2.78      |
| Standard deviation    | 176      | 8.56                 | 0.075         | 0.052 | 38.9      | 3.16     | 2.517    | 0.607   | 1.30      |
| Kurtosis              | 21.8     | 0.227                | −0.465        | 0.984 | 3.43      | 6.81     | 0.315    | 1.14    | −1.08     |
| Skewness              | 3.98     | −0.630               | 0.436         | −0.127 | 0.837    | 2.13     | −0.135   | 0.554   | 0.181     |

Table 4: Different sets of inputs.

| No. | Scenarios                                      | Inputs                                      |
|-----|------------------------------------------------|---------------------------------------------|
| 1   | Without GR and spectral GR                     | CNP, RHOB, $\Delta t$, FR                  |
| 2   | Without spectral GR                             | CNP, RHOB, GR, $\Delta t$, FR              |
| 3   | Without density                                 | CNP, GR, $\Delta t$, FR, K, TH, Ur         |
| 4   | Without porosity                                | RHOB, GR, $\Delta t$, FR, K, TH, Ur        |
| 5   | Without resistivity                             | CNP, RHOB, GR, $\Delta t$, K, TH, Ur       |
| 6   | Without sonic transient time                   | CNP, RHOB, GR, FR, K, TH, Ur               |
| 7   | All logs                                       | CNP, RHOB, GR, $\Delta t$, FR, K, TH, Ur   |
Figure 3: Given and SVM-predicted TOC cross-plots for (a) the training and (b) the testing datasets.

Figure 4: Given and DT-predicted TOC cross-plots for (a) the training and (b) the testing datasets.

Figure 5: Given and RF-predicted TOC cross-plots for (a) the training and (b) the testing datasets.
Table 5: Performance of different sets of inputs.

| No. | Inputs                        | Training $R$ | AAPE (%) | Testing $R$ | AAPE (%) |
|-----|-------------------------------|--------------|----------|-------------|----------|
| 1   | CNP, RHOB, $\Delta t$          | 0.978        | 8.5      | 0.811       | 25.7     |
| 2   | CNP, RHOB, GR, $\Delta t$, FR | 0.987        | 5.8      | 0.900       | 16.5     |
| 3   | CNP, GR, $\Delta t$, FR, K, Th, Ur | 0.988 | 5.5     | 0.924       | 14.6     |
| 4   | RHOB, GR, $\Delta t$, FR, K, Th, Ur | 0.989 | 5.2     | 0.921       | 14.8     |
| 5   | CNP, RHOB, GR, $\Delta t$, K, Th, Ur | 0.985 | 6.5     | 0.913       | 16.6     |
| 6   | CNP, RHOB, GR, FR, K, Th, Ur   | 0.988        | 5.5      | 0.918       | 14.6     |
| 7   | CNP, RHOB, GR, $\Delta t$, FR, K, Th, Ur | 0.989 | 5.3     | 0.931       | 13.8     |

Figure 6: TOC prediction of the AI models and different correlation for the validation dataset of Well-C.
was significantly less accurate in testing with $R$ value which equals 0.864 and which reduces the ability to generalize the model. In Figure 4(b), it is noticeable that some points fall relatively away from the 45° line.

3.3. Random Forest. In comparison with DT, RF performed better in testing with a 0.931 correlation coefficient, and its performance with the training data set was very close with $R$ which equals 0.989. Figure 5 indicates the deviation between the actual and TOC prediction visually. In comparison with testing results of DT shown in Figure 4(b), it is noticeable in Figure 5(b) that the points fall closer to the 45° line.

4. Input Parameters Sensitivity

From the previous analysis, RF was chosen to perform the sensitivity on the inputs. Seven different sets of the well logs information are tested and reported in Table 5. The best performance was achieved when all eight parameters were incorporated (case 7) while the worst fitting accuracy happened when the gamma ray and spectral gamma ray were excluded (case 1). In the cases that exclude density, porosity, and sonic transient time each alone (cases 3, 4, and 6), the least effect in TOC prediction was noticed. It is noteworthy to mention that the effect of changing the inputs' set is more obvious in the testing dataset in contrast to the training dataset.

5. Models Validation

As shown in the previous results, good matching accuracies for Devonian shale's TOC have been achieved by the two methods in training and testing datasets. As an additional validation, eighty-two data points from Well-C were kept hidden from the AI tool during model construction utilized later to ensure the generalization of the new models. This well is located in the vicinity of the Well-I and Well-II that has been used in model building. The validation data points fall in 140 ft depth range. The accuracy of the two AI techniques was compared with that of three of the available models for TOC estimation, namely, the Schmoker, $\Delta$log$R$, and Zhao et al. correlations.

The TOC values obtained from different empirical correlations and AI-based models against the actual values of Well-C dataset are presented in Figure 6. This figure shows that the RF-based model outmatched all other models with AAPE of only 14% and a high $R$ of 0.94. DT and SVM models were less accurate than RF; however, in terms of AAPE, both were better than the other three correlations with AAPE values less than 16.4%. Zhao et al. [24] correlation results were the closest to AI-based models and not far from $\Delta$log$R$ predictions with AAPE values range between 20% and 25% and $R$ values between 0.83 and 0.83 while the least favorable results in the validation dataset were from the Schmoker model with AAPE above 48%.

The TOC prediction in the different datasets and their comparison with the existing correlations show the accuracy of the developed models for Devonian shale. These models determine the TOC using the conventional well logs and spectral GR logs. It outperformed the existing models that calculate the TOC based on RHOB log only (Schmoker model) or a combination of FR and $\Delta t$ logs and level of maturity (LOM) ($\Delta$log$R$ method) or bulk gamma ray, FR, and $\Delta t$ or RHOB logs (Zhao et al. [24] correlation). This result and the previous results in this study demonstrate the reliability of the AI models for TOC estimation in Devonian shale.

The developed models were able to accurately predict the TOC from the conventional well log including CNP, RHOB, GR, $\Delta t$, FR, K, Th, and Ur logs, which helps to obtain a continuous profile for TOC with depth without the need for core analysis or additional well interventions. Similar to any developed model, we recommend employing the developed models using input parameters within the same model’s inputs ranges to ensure reliable results. For future work, more data will be collected to validate the developed models and other ML techniques will be applied.

6. Conclusions

This study established three models for TOC prediction in Devonian shale from conventional well logs and spectral GR logs using different machine learning techniques and approximately 1250 data points from three wells. The employed ML techniques were support vector machine (SVM), decision tree (DT), and random forest (RF). A summary of the findings reported in this paper is as follows [45]:

(i) In training and testing datasets, the three AI algorithms produced good matches; however, the RF-based model has the best accuracy. The RF model was able to predict the TOC for the training and testing datasets, with $R$ values of 0.99 and 0.93, respectively, and AAPE values of 5.3% and 13.8% in the same order.

(ii) Data from a different well were hidden entirely from the AI tools and used to validate the built model. In this dataset, the RF model produced a 0.94 correlation coefficient and a 14% AAPE.

(iii) The AI-based models’ predictions were compared with three other empirical correlations. The AI models yielded more accurate results contrasted to the other models which resulted in AAPEs greater than 20%.

Abbreviations

AAPE: Average absolute percentage error
ANFIS: Adaptive neuro-fuzzy interference system
ANN: Artificial neural network
CNN: Convolutional neural network
CNP: Neutron porosity
DT: Decision tree
ELM: Extreme learning machine
FL: Fuzzy logic
FN: Functional network
FR: Formation resistivity
GPR: Gaussian process regression
GR: Gamma ray
K: Spectrum logs of potassium
ML: Machine learning
NF: Neuro-fuzzy
NN: Neural network
R: Correlation coefficient
RF: Random forest
RHOB: Bulk density
SP: Spontaneous potential
SVM: Support vector machine
Th: Spectrum logs of thorium
TOC: Total organic carbon
Ur: Spectrum logs of uranium
\( \rho \): Density
\( \Delta t \): Sonic transit time.

Data Availability
Most of the data are included in the manuscript. A detailed data sample will be provided upon request.

Conflicts of Interest
The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments
The authors would like to thank KFUPM for giving permission to publish this work.

References
[1] H. Tang, Z. Sun, Y. He, Z. Chai, A. R. Hasan, and J. Killough, “Investigating the pressure characteristics and production performance of liquid-loaded horizontal wells in unconventional gas reservoirs,” Journal of Petroleum Science and Engineering, vol. 176, pp. 456–465, 2019.
[2] P. Zhao, M. Ostadhassan, B. Shen et al., “Estimating thermal maturity of organic-rich shale from well logs: case studies of two shale plays,” Fuel, vol. 235, pp. 1195–1206, 2019.
[3] Y. Wu, P. Tahmasibi, H. Yu, C. Lin, H. Wu, and C. Dong, “Pore-scale 3D dynamic modeling and characterization of shale samples: considering the effects of thermal maturation,” Journal of Geophysical Research: Solid Earth, vol. 125, 2020.
[4] L. Zhu, C. Zhang, C. Zhang et al., “A new and reliable dual model- and data-driven TOC prediction concept: a TOC logging evaluation method using multiple overlapping methods integrated with semi-supervised deep learning,” Journal of Petroleum Science and Engineering, vol. 188, Article ID 106944, 2020.
[5] C. N. Zou, S. Z. Tao, B. Bai, and Z. Yang, “Differences and relations between unconventional and conventional oil and gas,” China Petroleum Exploration, vol. 20, pp. 1–16, 2015.
[6] S. Kumar, S. Das, R. Bastia, and K. Ojha, “Mineralogical and morphological characterization of older cambay shale from North cambay basin, India: implication for shale oil/gas development,” Marine and Petroleum Geology, vol. 97, pp. 339–354, 2018.
[7] S. Rani, E. Padmanabhan, and B. K. Prusty, “Review of gas adsorption in shales for enhanced methane recovery and CO2 storage,” Journal of Petroleum Science and Engineering, vol. 175, pp. 634–643, 2019.
[8] A. A. Mahmoud, S. Elkatatny, M. Mahmoud, M. Abouelresh, A. Abdulraheem, and A. Ali, “Determination of the total organic carbon (TOC) based on conventional well logs using artificial neural network,” International Journal of Coal Geology, vol. 179, pp. 72–80, 2017.
[9] H. Wang, W. Wu, T. Chen, X. Dong, and G. Wang, “An improved neural network for TOC, S1 and S2 estimation based on conventional well logs,” Journal of Petroleum Science and Engineering, vol. 176, pp. 664–678, 2019.
[10] S.-C. Yang, N. Wang, M.-R. Li, and J. Yu, “The logging evaluation of source rocks of triassic Yanchang formation in Chongxin area, Ordos basin,” Journal of Natural Gas Science, vol. 24, pp. 470–476, 2013.
[11] L. Ma, K. G. Taylor, P. J. Dowey, L. Courtois, A. Gholinia, and P. D. Lee, “Multi-scale 3D characterisation of porosity and organic matter in shales with variable TOC content and thermal maturity: examples from the Lublin and Baltic Basins, Poland and Lithuania,” International Journal of Coal Geology, vol. 180, pp. 100–112, 2017.
[12] H. Carvajal-Ortiz and T. Gentzis, “Critical considerations when assessing hydrocarbon plays using Rock-Eval pyrolysis and organic petrology data: data quality revisited,” International Journal of Coal Geology, vol. 152, pp. 113–122, 2015.
[13] B. Hazra, S. Dutta, and S. Kumar, “TOC calculation of organic matter rich sediments using Rock-Eval pyrolysis: critical consideration and insights,” International Journal of Coal Geology, vol. 169, pp. 106–115, 2017.
[14] V. Bolandi, A. Kadkhodaie, and R. Farzi, “Analyzing organic richness of source rocks from well log data by using SVM and ANN classifiers: a case study from the Kazhdumi formation, the Persian Gulf basin, offshore Iran,” Journal of Petroleum Science and Engineering, vol. 151, pp. 224–234, 2017.
[15] Y. Chen, S. Jiang, D. Zhang, and C. Liu, “An adsorbed gas estimation model for shale gas reservoirs via statistical learning,” Applied Energy, vol. 197, pp. 327–341, 2017.
[16] H. Daigle, N. W. Hayman, E. D. Kelly, K. L. Milliken, and H. Jiang, “Fracture capture of organic pores in shales,” Geophysical Research Letters, vol. 44, pp. 2167–2176, 2017.
[17] A. A. Mahmoud, S. Elkatatny, A. Ali, A. Abdulraheem, and M. Abouelresh, “Estimation of the total organic carbon using functional neural networks and support vector machine,” in Proceedings of the Day 3 Wed, Dhahran, Saudi Arabia, January 2020.
[18] A. A. Mahmoud, S. Elkatatny, A. Ali, M. Abouelresh, and A. Abdulraheem, “New robust model to evaluate the total organic carbon using fuzzy logic,” in Proceedings of the Day 4 Wed, Mishref, Kuwait, October 2019.
[19] E. J. Mathia, T. F. T. Rexer, K. M. Thomas, L. Bowen, and A. C. Aplin, “Influence of clay, calcareous microfossils, and organic matter on the nature and diagenetic evolution of pore systems in mudstones,” Journal of Geophysical Research: Solid Earth, vol. 124, pp. 149–174, 2019.
[20] J. W. Schmoker, “Determination of organic content of Appalachian Devonian shales from formation-density logs: geologic notes,” The American Association of Petroleum Geologists Bulletin, vol. 63, 1979.
[21] J. W. Schmoker, “Organic content of Devonian shale in western Appalachian basin,” The American Association of Petroleum Geologists Bulletin, vol. 64, pp. 2156–2165, 1980.
[22] Q. R. Passey, S. Creaney, J. B. Kulla, F. J. Morett, and J. D. Stroud, “A practical model for organic richness from porosity and resistivity logs,” The American Association of Petroleum Geologists Bulletin, vol. 74, pp. 1777–1794, 1990.

[23] P. Wang, Z. Chen, X. Pang, K. Hu, M. Sun, and X. Chen, “Revised models for determining TOC in shale play: example from devonian Duvernay shale, western Canada sedimentary basin,” Marine and Petroleum Geology, vol. 70, pp. 304–319, 2016.

[24] P. Zhao, H. Ma, V. Rasouli, W. Liu, J. Cai, and Z. Huang, “An improved model for estimating the TOC in shale formations,” Marine and Petroleum Geology, vol. 83, pp. 174–183, 2017.

[25] I. Kumar, J. Virmani, H. S. Bhadauria, and M. K. Panda, “Chapter 13-classification of breast density patterns using PNN, NFC, and SVM classifiers,” in Soft Computing Based Medical Image Analysis, N. Dey, A. S. Ashour, F. Shi, and V. E. Balas, Eds., pp. 223–243, Academic Press, Cambridge, MA, USA, 2018.

[26] M. Tan, X. Song, X. Yang, and Q. Wu, “Support-vector-regression machine technology for total organic carbon content prediction from wireline logs in organic shale: a comparative study,” Journal of Natural Gas Science and Engineering, vol. 26, pp. 792–802, 2015.

[27] L. Zhu, C. Zhang, C. Zhang et al., “Forming a new small sample deep learning model to predict total organic carbon content by combining unsupervised learning with semi-supervised learning,” Applied Soft Computing, vol. 83, Article ID 105596, 2019.

[28] A. Kadkhodaie-Ilkhchi, H. Rahimpour-Bonab, and M. Rezaee, “A committee machine with intelligent systems for estimation of total organic carbon content from petrophysical data: an example from Kangan and Dalan reservoirs in South Pars Gas Field,” Computers & Geosciences, vol. 35, pp. 459–474, 2009.

[29] B. Alizadeh, S. Najjari, and A. Kadkhodaie-Ilkhchi, “Artificial neural network modeling and cluster analysis for organic facies and burial history estimation using well log data: a case study of the South Pars,” Computers & Geosciences, vol. 45, pp. 261–269, 2012.

[30] H. Yu, R. Rezaee, Z. Wang et al., “A new method for TOC estimation in tight shale gas reservoirs,” International Journal of Coal Geology, vol. 179, pp. 269–277, 2017.

[31] M. Khoshnoodka, H. Mohseni, O. Rahmani, and A. Mohammadi, “TOC determination of Gadvan Formation in South Pars Gas field, using artificial intelligent systems and geochemical data,” Journal of Petroleum Science and Engineering, vol. 78, pp. 119–130, 2011.

[32] E. Sfidari, A. Kadkhodaie-Ilkhchi, and S. Najjari, “Comparison of intelligent and statistical clustering approaches to predicting total organic carbon using intelligent systems,” Journal of Petroleum Science and Engineering, vol. 86–87, pp. 190–205, 2012.

[33] B. Alizadeh, K. Maroufi, and M. H. Heidarifard, “Estimating source rock parameters using wireline data: an example from Dezful Embayment, South West of Iran,” Journal of Petroleum Science and Engineering, vol. 167, pp. 857–868, 2018.

[34] M. R. Shalaby, N. Jumat, D. Lai, and O. Malik, “Integrated TOC prediction and source rock characterization using machine learning, well logs and geochemical analysis: case study from the Jurassic source rocks in Shams Field, NW Desert, Egypt,” Journal of Petroleum Science and Engineering, vol. 176, pp. 369–380, 2019.

[35] J. Rui, H. Zhang, D. Zhang, F. Han, and Q. Guo, “Total organic carbon content prediction based on support-vector-regression machine with particle swarm optimization,” Journal of Petroleum Science and Engineering, vol. 180, pp. 699–706, 2019.

[36] X. Shi, J. Wang, G. Liu, L. Yang, X. Ge, and S. Jiang, “Application of extreme learning machine and neural networks in total organic carbon content prediction in organic shale with wire line logs,” Journal of Natural Gas Science and Engineering, vol. 33, pp. 687–702, 2016.

[37] P. Wang, S. Peng, and T. He, “A novel approach to total organic carbon content prediction in shale gas reservoirs with well logs data. Tonghua Basin, China,” Journal of Natural Gas Science and Engineering, vol. 55, pp. 1–15, 2018.

[38] S. Elkattatny, “A self-adaptive artificial neural network technique to predict total organic carbon (TOC) based on well logs,” Arabian Journal for Science and Engineering, vol. 44, pp. 6127–6137, 2019.

[39] A. A. Mahmoud, S. ElKatatny, A. Abdulraheem, M. Mahmoud, M. Omar Ibrahim, and A. Ali, “New technique to determine the total organic carbon based on well logs using artificial neural network (white box),” in Proceedings of the Day 3 Wed, Dammam, Saudi Arabia, April 2017.

[40] A. A. Mahmoud, S. Elkattatny, A. Z. Ali, M. Abouelresh, and A. Abdulraheem, “Evaluation of the total organic carbon (TOC) using different artificial intelligence techniques,” Sustainability, vol. 11, p. 5643, 2019.

[41] D. Liu, “Prediction method of TOC content in mudstone based on artificial neural network,” IOP Conference Series: Earth and Environmental Science, vol. 781, no. 2, Article ID 022087, 2021.

[42] D. Z heng, S. Wu, and M. Hou, “Fully connected deep network: an improved method to predict TOC of shale reservoirs from well logs,” Marine and Petroleum Geology, vol. 132, Article ID 105205, 2021.

[43] S. Creaney, J. Allan, K. S. Cole et al., ”Petroleum generation and migration in the western Canada Sedimentary Basin,” in Geological Atlas of the Western Canada Sedimentary Basin, pp. 455–468, Canadian Society of Petroleum Geologists and Alberta Research Council, Alberta, Canada, 1994.

[44] C. D. Rokosh, S. Lyster, S. D. A. Anderson et al., “Summary of Alberta’s Shale-and siltstone-hosted hydrocarbon resource potential,” ERCB/AGS Open File Report, 2012.

[45] Z. Chen, C. Jiang, D. Lavoie, and J. Reyes, “Model-assisted Rock-Eval data interpretation for source rock evaluation: examples from producing and potential shale gas reservoir plays,” International Journal of Coal Geology, vol. 165, pp. 290–302, 2016.