ACS OMEGA

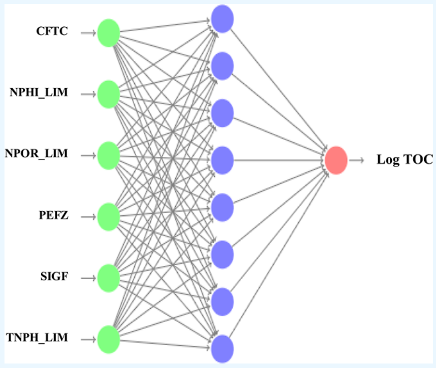
http://pubs.acs.org/journal/acsodf Article

Prediction of Total Organic Carbon in Organic-Rich Shale Rocks Using Thermal Neutron Parameters

Amjed Hassan,* Emad Mohammed, Ali Oshaish, Dhafer Badhafere, Korhan Ayranci, Tian Dong, Umair bin Waheed, Ammar El-Husseiny, and Mohamed Mahmoud*

Cite This: *ACS Omega* 2023, 8, 4790−4801 Read Online

ABSTRACT: Total organic carbon (TOC) content is one of the crucial parameters that determine the value of the source rock. The TOC content gives important indications about the source rocks and hydrocarbon volume. Various techniques have been utilized for TOC quantification, either by geochemical analysis of source rocks in laboratories or using well logs to develop mathematical correlations and advanced machine learning models. Laboratory methods require intense sampling intervals to have an accurate understanding of the reservoir, and depending on the thickness of the interested formation, it can be time-consuming and costly. Empirical correlations based on well logs (e.g., density, sonic, gamma ray, and resistivity) showed fast predictions and very reasonable accuracies. However, other important parameters such as thermal neutron logs have not been studied yet as a potential input for providing reliable TOC predictions. Also, different studies estimate the TOC based on the well-logging data for various formations; however, limited studies were reported to predict the TOC for the Horn River Formation. Therefore, the objective of this study is to estimate the TOC variations based on the thermal neutron logs using one of the largest source rocks in Canada: The Horn River Formation. More than 150 data sets were collected and used in this work. The parameters of the artificial neural network (ANN) model were fine-tuned in order to improve the model's prediction performance. Furthermore, an empirical correlation was developed utilizing the optimized ANN model to allow fast and direct application for the developed model. The developed correlation can predict the TOC with an average absolute error of 0.52 wt %. The proposed TOC model was able to outperform the previous models, and the coefficient of determination was increased from 0.28 to 0.73. Overall, the proposed TOC model can provide high accuracy for TOC ranges from 0.3 to 6.44 wt %. The developed model can provide a real-time quantification for the organic matter maturity, helping to allocate the zones of mature organic matter within the drilled formations.



1. INTRODUCTION

In recent years, unconventional hydrocarbon production from shale source rocks has gained a vast interest in the oil and gas industry. Such interest is attributed to the depletion of conventional resources and the advancements in directional drilling and hydraulic fracturing.[1−4] The total organic carbon (TOC) content is one of the very important petrophysical properties that indicate the quality of source rocks. TOC represents the amount of organic matter deposited in the rock and is one of the most critical parameters to investigate prior to drilling for its significance in hydrocarbon quantification and quality measurement of the resource.[2,4−9] Organic matter depositions are controlled by the primary production, destruction, and dilution of organic matter where it maximizes when the first is fairly greater than the latter two factors.[10] Many researchers proved the strong relation of TOC with many source rock parameters, such as gas adsorption capacity[11] and porosity[10] which emphasizes the importance of TOC as a rock petrophysical feature.[12,13]

For TOC quantification, various techniques have been utilized either by a geochemical analysis of source rock cores in laboratories or using well logs to develop mathematical correlations and advanced machine learning models. Laboratory TOC quantification methods where core samples are tested for organic material are considered to be reliable and standard methods, such as the Rock-Eval 6 pyrolysis method.[12−15] In the Rock-Eval 6 method, a small amount of the targeted sample (up to 100 mg) is thermally analyzed at high temperatures (up to 850 °C) in two stages: pyrolysis and oxidation. The pyrolysis stage gives the amount of free hydrocarbons, released hydrocarbons as a result of thermal cracking, and the amount of CO

Received: October 27, 2022
Accepted: January 17, 2023
Published: January 24, 2023

Table 1. Summary for Predicting the TOC Using Different AI Methods

reference	objective	technique	input parameters	geological information
Alizadeh et al.[51]	TOC and S_2 estimations	ANN	RT and DT logs	Dezful Embayment basin in Iran
Tan et al.[52]	TOC prediction	SVM and RBF	GR, AC, CNL, TH, U, K, PE, RT, and RHOB	Huangping syncline basin in China
Wang et al.[53]	TOC prediction	nonlinear regression	RT and DT logs	Sichuan Basin in China
Mahmoud et al.[15]	TOC prediction	ANN	DT, GR, RHOB, and RT	Devonian Shale in America
Zhao et al.[54]	TOC prediction	nonlinear regression	CNL	Bakken formations in the USA
Rui et al.[55]	TOC prediction	SVM	GR, RHOB, SP, DT, and RT	Beibu Gulf formations in China
Lawal et al.[56]	TOC prediction	ANN	XRD such as Al_2O_3, SiO_2, CaO, and MgO	Devonian formations in America
Sultan[57]	TOC prediction	ANN and differential evolution	DT, GR, RHOB, and RT	Devonian formations in America
Wang et al.[46]	TOC and S_2 estimations	ANN	NPHI, RHOB, DT, and RT	Bohai Bay formations in China
Handhal et al.[14]	prediction of TOC	SVM, ANN, and random forest	RHOB, GR, RLLD, DT, and NPHI	Rumaila formations in Iran

and CO_2 released. On the other hand, the oxidation stage determines the amount of released CO and CO_2 when oxidizing the residual organic matter from the pyrolysis stage.[16,17] Eventually, TOC is determined by adding the amount of produced organic carbon in the pyrolysis stage and the residual amount of organic carbon from the oxidation stage.[17,18] Peters et al.[19] stated many other laboratory techniques for TOC quantification such as filter acidification,[19] nonfilter acidification,[20,21] total minus coulometric,[22] laser-induced pyrolysis,[23] and diffuse reflectance infrared Fourier transform spectroscopy.[24,25] Although these techniques can be reliable, in order to achieve accurate results, researchers need to analyze cores and outcrops in short intervals (e.g., every 25 cm). This method can therefore be time-consuming and costly particularly if the reservoir is thick and if the lateral TOC variations are also needed to be taken into account. Moreover, core preservation may be limited which is another drawback of laboratory-based TOC quantitation tools.[17] However, it is still crucial to determine TOC even in the absence of core data. Hence, other approaches that employ well logs have been established to cover the shortage in TOC quantification when laboratory methods are not applicable. Throughout the past decades, well-log data have been used to assess numerous parameters starting from the identification of source rocks and organic materials since the 1940s[26] and reaching the calculation of petrophysical parameters.[27−30] Additionally, the TOC content has been assessed *via* well logs by approaches like the gross gamma ray that utilizes gamma ray log,[30] using hyperspectral imagery,[31] utilizing bulk density log,[26] and using transit time and resistivity logs such as the widely used ΔlogR method.[32] However, most of the available methods may fail to provide accurate estimations for the TOC due to the involvement of different parameters. For example, the gross gamma ray is affected more by uranium, while the bulk density log overestimates the TOC in clay-rich formations. Also, the approach based on Δlog*R* requires multiple variables such as the level of maturation and baseline resistivity which are not uniform especially in the frequently interbedded shale rocks due to their extreme heterogeneity.[13,15,19,33]

New logging techniques such as nuclear magnetic resonance and lithology scanning tools offer developments in TOC quantification accuracy.[34] However, they incur prohibitively high costs and cannot be used conventionally, hence requiring new robust and more convenient methods for TOC quantification.[7] Artificial intelligence (AI) has been considerably used in the last few years in oil and gas research, and much work has been made on the prediction of TOC based on core and well log data.[15,35−41] In most cases, AI methods are not globally applicable due to the heterogeneity of shales, which indicates the cruciality of studying the nature of the targeted fields and picking proper logs for the model.[1,33] Huang et al.[39] demonstrated one of the early applications of AI in predicting TOC using only three conventional (gamma ray, resistivity, and sonic) logs as an input and a pseudo-TOC log calculated from an empirical correlation following the Passey et al.[32] approaches.[39] Subsequently, conventional logs have been fed as inputs, that is, gamma ray log, density log, acoustic log, deep and medium resistivity logs, and porosity log in addition to uranium (U), thorium (Th), and potassium (K) contents.[42−44] X-ray fluorescence elements' data (such as copper and nickel) and thermal neutron porosity[1] as well as conventional log combinations displayed correlation with the TOC in the investigated environment[45−50] and were proofing evidence of AI reliability in predicting TOC. Table 1 provides a summary of the studies conducted for predicting the organic matter in shale formations.

Different approaches have been used to estimate the TOC including experimental measurements and empirical correlations. Laboratory TOC determinations are expensive and time-consuming methods, as well as require a large number of core samples to construct the TOC profiles. Empirical correlations based on regression and AI techniques showed a fast prediction approach with very reasonable accuracy. Different algorithms have been utilized to determine the TOC based on a wide range of inputs; however, thermal neutron logs have never been studied yet as a potential parameter that can be used to provide reliable TOC estimations. The TOC content can be estimated based on the response of thermal neutron logs due to the relationship between the thermal neutron and TOC. The major constituents of organic matter are carbon and hydrogen elements which can significantly affect the movement of the neutrons ejected during the thermal neutron logging. The neutron tool responds primarily to the presence of hydrogen; increasing the amount of hydrogen will lead to an increase in the neutrons' slowness.[58] Also, hydrogen has a thermal capture unit of 200 per gm/cm^3, which is very high compared to other elements.[59] Hence, the theoretical relationship between neutron count rates (thermal logs) and TOC is always feasible. Therefore, in this study, we used the thermal neutron logs to estimate the TOC variations. Different studies tried to estimate the TOC based on the well-logging data for various formations; however, limited studies were reported to predict the TOC for the Horn River Formation.[60,61] Therefore, the major objective of this work is to develop a model to predict the TOC based on the thermal neutron logs for the Horn River Formation. More than 150 data sets were collected and used in this work. We used the artificial neural network (ANN) method to propose a new TOC model. The developed model was optimized to improve

https://doi.org/10.1021/acsomega.2c06918
ACS Omega 2023, 8, 4790−4801

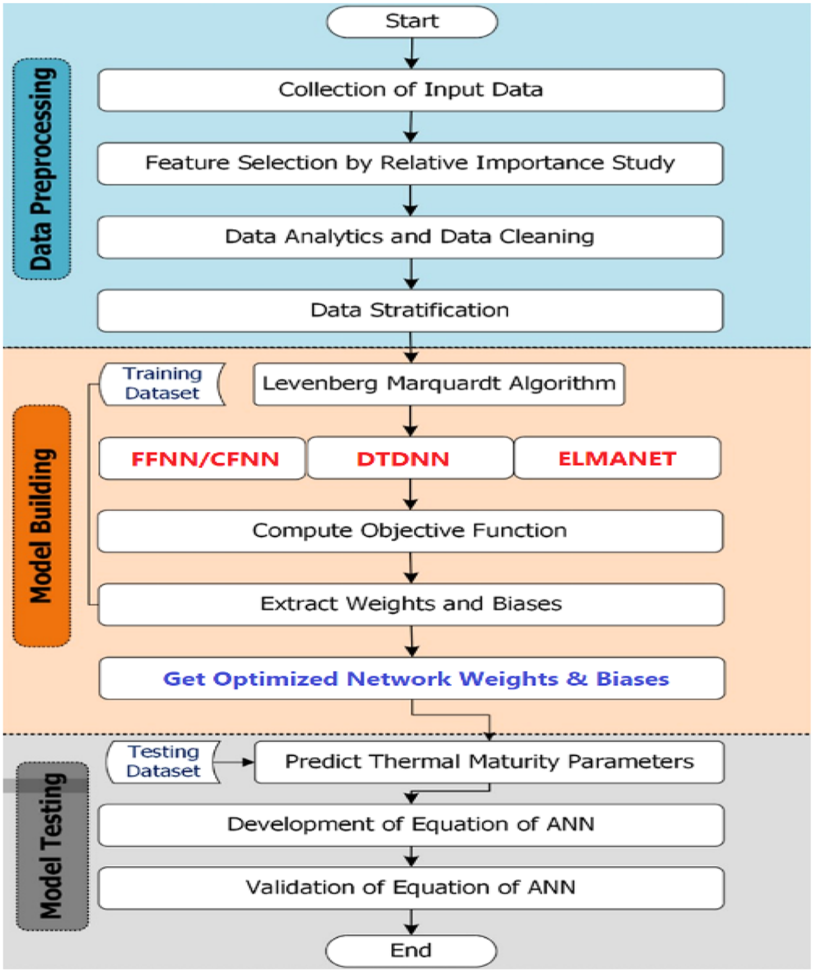


Figure 1. Flow chart describing the algorithm used for optimizing the artificial neural network (ANN) model for determining the total organic carbon (TOC) profile.

the prediction performance. Finally, an empirical correlation was proposed using the optimized ANN program to allow fast and direct application for the developed model.

2. METHODS AND DATA DESCRIPTION

2.1. Methods. In this work, ANN models were developed and trained using the petrophysical parameters and the measured TOC. More than 150 data sets including TOC and well-logging data were collected and used in this study. The petrophysical logs are corrected far thermal counting (CFTC) rate, thermal neutron porosity (NPHI), standard resolution formation photoelectric factor (PEFZ), enhanced thermal neutron porosity (NPOR), and formation capture cross-section (SIGF) logs. Basically, in this work, the thermal logs (CFTC, NPHI, and NPOR) were selected since they were not tested before in predicting the TOC, while the formation lithology logs (PEFZ and SIGF) were used to give more information about the layers. The PEFZ log provides very useful information about rock mineralogy based on measuring the photoelectric absorption factor (Pe). For example, high Pe values can indicate high percentages of clays. The TOC values were obtained for around 150 core samples. The TOC contents were measured by Weatherford Laboratories using LECO (Laboratory Equipment Corporation) combustion; more details about the TOC measurements can be found in Dong et al.[12] Elman-type neural networks were developed and trained using a Bayesian regularization backpropagation algorithm. Different types of transfer functions were utilized to propose the best prediction model. In addition, the model parameters were optimized by examining different numbers of neurons and training function types, till achieving the best prediction performance. The ANN model was assessed using different approaches such as visual check and various evaluation indexes including correlation coefficient (CorrCoef), root-mean-square error (RMSE), average absolute error (AAE), and coefficient of determination (R^2). Moreover, an explicit empirical equation was developed based on the weights and biases of the best-performing ANN model. The proposed model can provide fast and reliable TOC estimations based on the logging parameters.

Figure 1 shows a flow chart describing the main steps used in this work which are data preprocessing, model development, and model testing. The preprocessing stage includes data collection and cleaning in order to remove the noisy and outlier points. We detected and removed the outliers manually based on the standard deviation values. Also, zeros and negative readings were removed manually. Data stratification was applied which indicates the separation of all data into different groups: training and testing. In the second phase, we developed and optimized the ANN model by examining different ANN types and structures. The examined types include feedforward neural networks, correlation filter neural networks, densely connected time delay neural networks, and Elman-type neural networks.

https://doi.org/10.1021/acsomega.2c06918
ACS Omega 2023, 8, 4790−4801

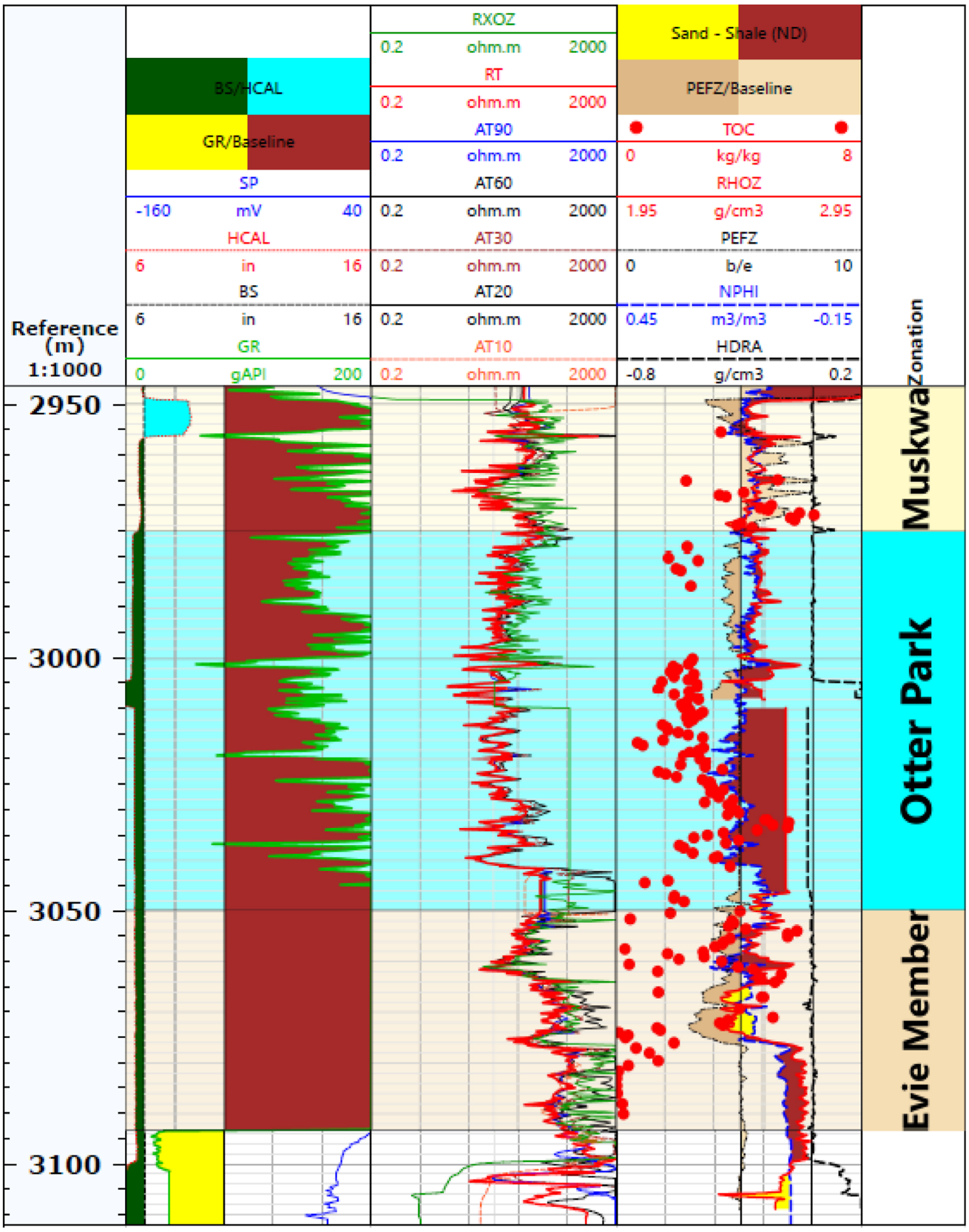


Figure 2. Log profiles used in this work to estimate the total organic carbon variation.

Also, we tested the performance of the developed model using the testing data set, which was kept unseen by the model during the development/training stage. Finally, we extracted the ANN parameters to be used for developing a new ANN-based equation.

2.2. Data Description. The TOC values used in this study come from the Maxhamish core of the Horn River Formation, Canada, and are adapted from Dong et al.[12,13] The Horn River Formation shows significant reservoir heterogeneity issues due to subtle changes in grain size within the shale formation.[59] Figure 2 shows the log profiles used to estimate the TOC; TechLog software was used to plot the logs. The logs include CFTC, NPHI, NPOR, PEFZ, SIGF, and TNPH. NPOR measures the formation porosity based on the emission of the fast neutron, and the tool can detect the hydrogen atoms present in the pores, since hydrogen has the biggest effect in slowing down and capturing neutrons. PEFZ is a supplementary measurement to the bulk density, and the tool is used as an additional input to resolve mixtures of minerals such as complex carbonates, dolomitic limestones, and anhydritic dolomites. SIGF is the atomic capture section for neutrons which can detect the effective area required to pass and capture the neutrons. TNPH can measure the slowing down and capture of neutrons between a source and one or more thermal neutron detectors. It should be highlighted that different inputs were examined in order to improve the TOC prediction. We have found out that using logs such as the formation lithology (PEFZ and SIGF) can improve the TOC predictions. Also, relatively high values for

https://doi.org/10.1021/acsomega.2c06918
ACS Omega 2023, 8, 4790−4801

Table 2. Statistical Analysis of the Data Used in This Study

parameters	CFTC (HZ)	NPHI (V/V)	NPOR (V/V)	PEFZ (B/E)	SIGF (M-1)	TNPH (V/V)	TOC (wt %)
max	15968.94	0.18	0.25	8.13	3.43	0.16	6.44
min	2217.21	0.001	0.01	1.81	0.81	0.01	0.03
mean	4649.30	0.09	0.09	4.95	2.09	0.08	2.83
range	13751.72	0.17	0.27	5.13	2.62	0.16	6.41
st. deviation	2295.38	0.04	0.04	0.81	0.61	0.03	1.45

CorrCoef were observed between the PEFZ and TOC, as will be discussed in the Results and Discussion section.

In addition, statistical analysis was carried out to estimate the minimum, maximum, mean, and standard deviation, as provided in Table 2. The studied region showed a wide porosity range between 1 and 25%, while TOC varies between 0.03 and 6.44 wt % with a mean value of 2.83%. Finally, the histograms for all input and output are provided in Figure 3. It can be seen that the data showed different distributions; however, this will not affect the ANN performance, since the ANN tool automatically normalizes the data between the minimum and maximum values for each parameter, which will minimize the impact of different data distributions on the prediction process.

3. RESULTS AND DISCUSSION

Thermal neutron logs can be positively related to the total carbon content because the ejected neutrons from the logging tool can be significantly influenced by the high percentages of carbon and hydrogen elements in the formation. Therefore, a continuous TOC profile can be constructed utilizing the thermal neutron tool. In this work, an ANN technique was employed to develop a new model for estimating the TOC. This section presents the data analytics and the machine learning approach used in this work to provide reliable TOC predictions. Also, a new correlation for estimating the TOC is presented and validated.

3.1. Data Analytics. The results of the statistical analysis of the data set used in this work are given in Tables 3 and 4, for the training and testing data sets, respectively. The input parameters are CFTC rate, thermal neutron porosity based on ratio method (NPHI), enhanced thermal neutron porosity (NPOR), standard resolution formation photoelectric factor (PEFZ), formation capture cross-section (SIGF), and thermal neutron porosity (TNPH). In this work, the TOC values were transferred into different domains, and we found out that the logarithmic domain will lead to the best TOC predictions; similar observations were reported by many studies.[1,9,15] Also, the ANN technique usually uses the L2 norm regularization for its cost function. The L2 regularization is the most common type of all regularization techniques and is also commonly known as weight decay regression. L2 norm works well with normal Gaussian distribution data; this could be the reason for getting better results once the TOC data were transformed to the log domain. The values of the Log (TOC) are given in the last columns in Tables 3 and 4. Statistical parameters such as skewness and kurtosis were determined to indicate the distribution of each parameter. Most of the variables showed non-normal distributions. It is worth mentioning that the whole data were subdivided into two groups: training and validation data. We used a randomized function to select the training and validation data sets. Also, different training−validation ratios were examined. The most suitable ratio was selected based on the minimum error profiles. The training data represent 70% of the whole data and were mainly used to train the ANN model to determine the relationship between the input parameters and the TOC. The validation data were not used during the training stage and were used only to test the model's reliability.

Figure 4 presents the relationship between the input and output parameters where a high value indicates a strong relationship. Also, positive values reveal direct proportion, while negative values indicate inverse proportion. The TOC showed a strong relationship with all of the thermal counting rates, formation photoelectric factor, and thermal neutron porosity. All parameters (except CFTC) showed a positive impact. Increasing any of these parameters can lead to an increase in the organic content. Also, all porosity logs (NPHI, NPOR, and TNPH) showed positive relationships with the TOC, which is very reasonable. In general, increasing the volume of organic matter will result in reducing the matrix density and hence increase the formation porosity. Moreover, higher TOC can lead to higher PEFZ which can be attributed to the higher absorption factor by the organic matter. However, the thermal counting rate showed a negative relationship with the TOC value. The formation of high carbon content can lead to reducing the readings of the CFTC logs.

In addition, Figure 4 shows the relative importance of the input variables with TOC after using the logarithmic transformation. Transferring the TOC values into the logarithmic domain showed considerable improvements in the input−output relationship. All parameters showed an increase in relative importance without changing the signs of the values, indicating that transferring the TOC into logarithmic values can improve the prediction performance as well as persevering the petrophysical relationship between the input parameters and the TOC values.

3.2. Building the ANN Model. In this work, the ANN technique was used to propose a new TOC model; ANN was selected because it has shown strong prediction performance in different applications. Various training and transferring functions can be used in the ANN model till achieving an acceptable prediction performance. In the ANN model, training functions are used to train the model to capture the relationship between the inputs and output, while the transfer function translates the input signals to output signals. In this work, a new ANN model was built to estimate the TOC values based on the well logs. All ANN parameters are carefully selected in order to minimize the prediction error. RMSE, AAE, and the R^2 were utilized as error indexes. The ultimate objective is to find the best ANN parameters that can provide the minimum RMSE and AAE and the maximum R^2.

In general, the TOC ANN model is composed of six input neurons which are the thermal counting rate (CFTC), thermal neutron porosity (NPHI), neutron porosity (NPOR), photoelectric factor (PEFZ), formation capture (SIGF), and thermal neutron porosity (TNPH) logs. It should be highlighted that the data analysis showed a weak relationship between the TOC and SIGF and PEFZ logs, but we included these logs because they can give important information about the formations.

https://doi.org/10.1021/acsomega.2c06918
ACS Omega 2023, 8, 4790−4801

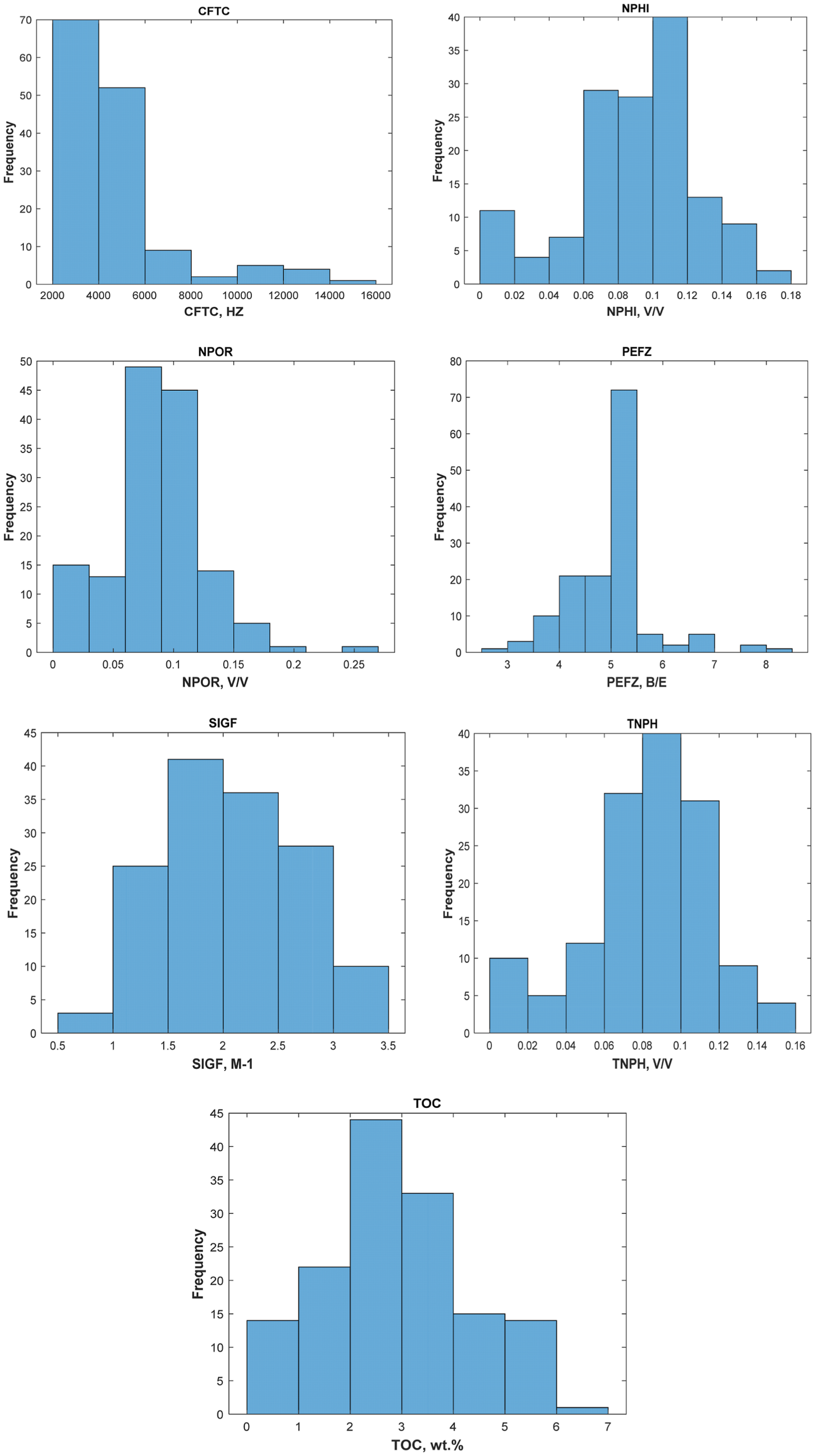


Figure 3. Histograms for the input and the output parameters.

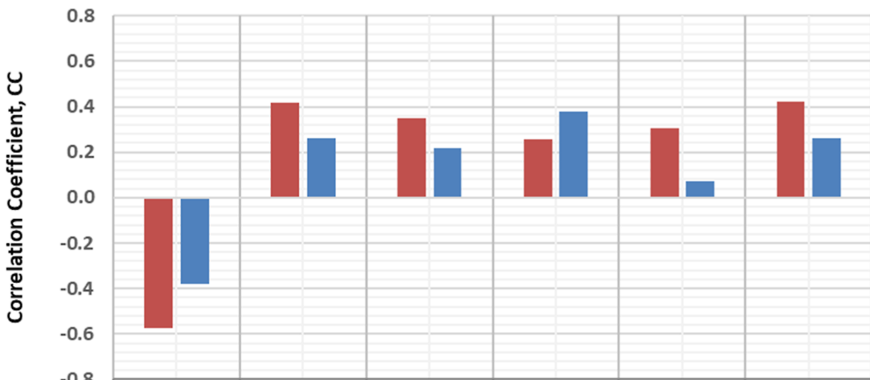
https://doi.org/10.1021/acsomega.2c06918
ACS Omega 2023, 8, 4790−4801

Table 3. Statistical Analysis of the Training Data Used for TOC Prediction

parameters	CFTC (HZ)	NPHI (V/V)	NPOR (V/V)	PEFZ (B/E)	SIGF (M-1)	TNPH (V/V)	TOC (wt %)	log (TOC)
max	15968.94	0.18	0.25	8.13	3.43	0.16	6.44	0.81
min	2217.21	0.00	0.02	3.00	0.81	0.01	0.03	−1.55
mean	4649.30	0.09	0.09	4.95	2.09	0.08	2.83	0.35
range	13751.72	0.17	0.27	5.13	2.62	0.16	6.41	2.36
st. deviation	2295.38	0.04	0.04	0.81	0.61	0.03	1.45	0.40
skewness	2.52	−0.34	0.39	1.25	0.11	−0.51	0.20	−2.50
kurtosis	7.49	0.06	2.39	4.09	−0.79	0.39	−0.31	7.57

Table 4. Statistical Analysis of the Testing Data Used for TOC Prediction

parameters	CFTC (HZ)	NPHI (V/V)	NPOR (V/V)	PEFZ (B/E)	SIGF (M-1)	TNPH (V/V)	TOC (wt %)	log (TOC)
max	12238.08	0.15	0.15	6.46	3.21	0.14	5.87	0.77
min	2492.29	0.01	0.01	1.81	1.03	0.01	0.03	−1.57
mean	4854.40	0.09	0.08	4.46	1.98	0.08	2.79	0.32
range	9745.78	0.14	0.14	4.65	2.17	0.13	5.85	2.34
st. deviation	2492.00	0.03	0.03	0.93	0.53	0.03	1.49	0.47
skewness	2.35	−0.64	−0.39	−0.67	0.08	−0.90	0.14	−2.69
kurtosis	4.79	0.79	0.56	1.28	−0.41	1.16	−0.30	8.74



	CFTC	NPHI_LIM	NPOR_LIM	PEFZ	SIGF	TNPH_LIM
logTOC	-0.575	0.415	0.348	0.258	0.304	0.420
TOC	-0.380	0.261	0.216	0.377	0.074	0.263

Figure 4. Relative importance of the input variables with total organic carbon (TOC), before and after using the logarithmic transformation.

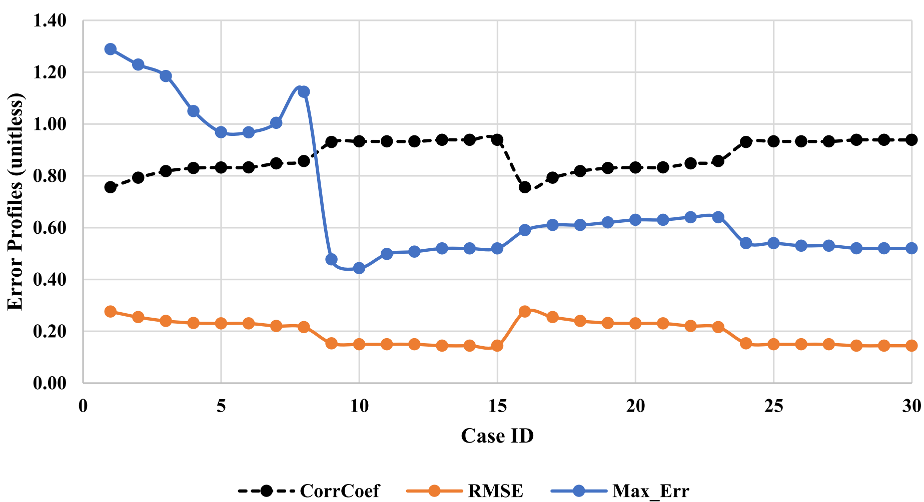


Figure 5. Correlation coefficient (CorrCoef), root-mean-square error (RMSE), and maximum error (Max_Err) performance with different cases.

The best ANN model is defined with three layers with eight neurons in the middle ANN layer. The number of ANN neurons was examined till acquiring the best prediction performance. Different combinations of neurons' number, hidden layers, training, and transfer functions were used, and around 30 combinations (cases) were examined. Figure 5 shows the error

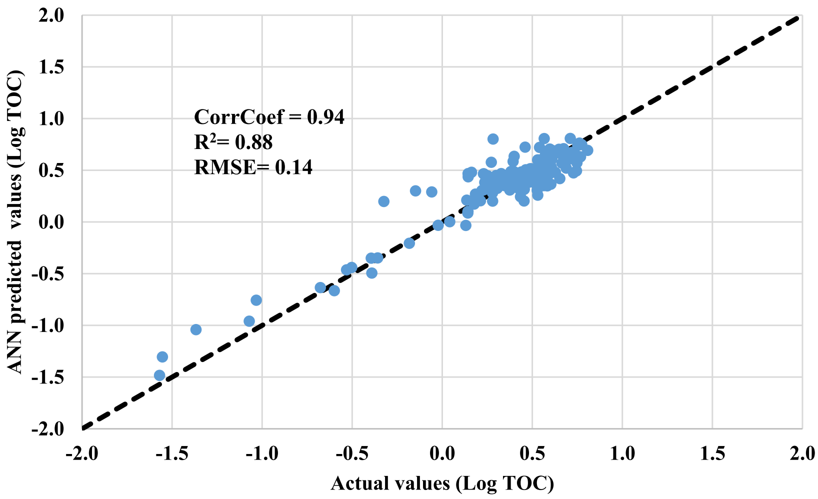


Figure 6. Regression plot of the actual measured total organic carbon (TOC) and the predicted values from the artificial neural network (ANN) model, in the logarithmic domain.

profiles for different combinations or cases. In addition, Figure 6 shows the measured and predicted TOC values based on the developed ANN model. The values showed good alignment along with the 45°-line, indicating effective prediction performance from the ANN model. The developed model can predict the TOC values with an RMSE of 0.14 and an R^2 value of 0.88. It should be noted that the ANN model will predict the values of log(TOC), and then the values should be transformed from the logarithmic domain into normal TOC values. It could be observed that the values presented in Figure 6 are overlapped; hence, to provide clearer visualization, the distribution of TOC values along depth is provided in Figure 7. The AAE between the laboratory-measured TOC and that predicted by ANN is 0.55 wt %. The laboratory-measured TOC is labeled by (Core), and the predicted values are labeled by (ANN). A good match between the measured and predicted values can be observed.

3.3. New Correlation. In this work, the optimized ANN model was used to propose an easy and fast-applicable model that can be used to compute the TOC variations based on the thermal logs. Figure 8 shows the schematic of the ANN model proposed in this work. A new TOC mathematical model is presented by utilizing the weights and biases of the ANN program. The proposed equation can help in estimating the TOC without the need of using the ANN code. The required inputs to estimate the TOC are the corrected thermal counting rate (CFTC), thermal neutron porosity (NPHI), enhanced thermal neutron porosity (NPOR), formation photoelectric factor (PEFZ), formation capture cross-section (SIGF), and thermal neutron porosity (TNPH) logs. Equation 1 represents the proposed TOC equation based on the optimized ANN model.

$$\mathrm{TOC} = \left[\sum_{i=1}^{N} w_{2i}\left(\frac{2}{1 + e^{-2(w_{1i,1}X_{j_n} + b_{1i})}}\right)\right] + b_2 \tag{1}$$

where TOC indicates the amount of organic matter, N equals 8 which is the number of neurons used in this work, w_1 and w_2 are the weights of the input and last ANN layers, respectively, X_{j_n} represents the normalized model inputs, and b_1 and b_2 are the biases for the input and target layers, respectively.

Importantly, the range of applicability of this correlation is provided in Table 3. Also, all inputs used in eq 1 should be used

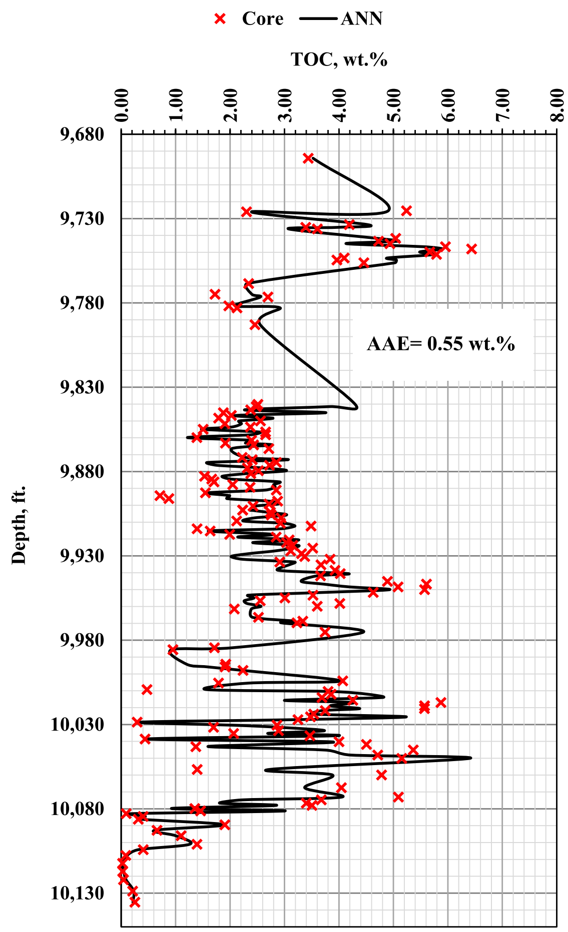


Figure 7. Variations of core measured total organic carbon (TOC) values and artificial neural network (ANN) predicted ones along with the depth.

as normalized values (X_{j_n}). The minimum ($X_{j_{\min}}$) and maximum ($X_{j_{\max}}$) can be used to determine the normalized values for each variable, as given by eq 2. The maximum ($X_{j_{\max}}$) and minimum ($X_{j_{\min}}$) values for all variables are given previously in Table 3, and

https://doi.org/10.1021/acsomega.2c06918
ACS Omega 2023, 8, 4790−4801

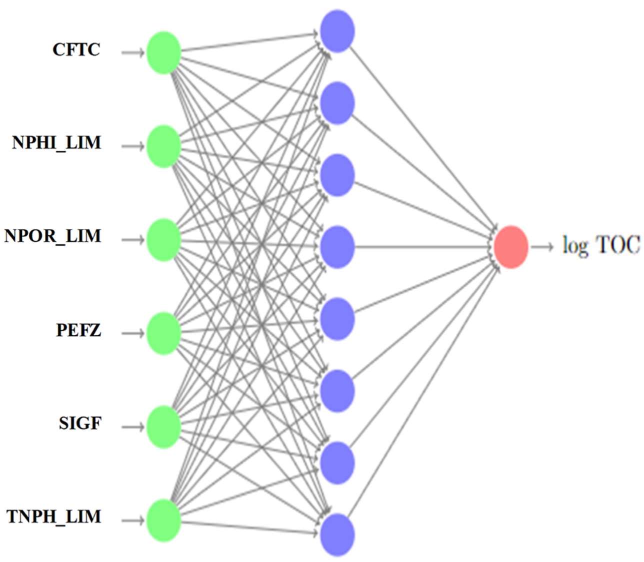


Figure 8. Neural network schematic showing the input variables and the output variable total organic carbon (TOC).

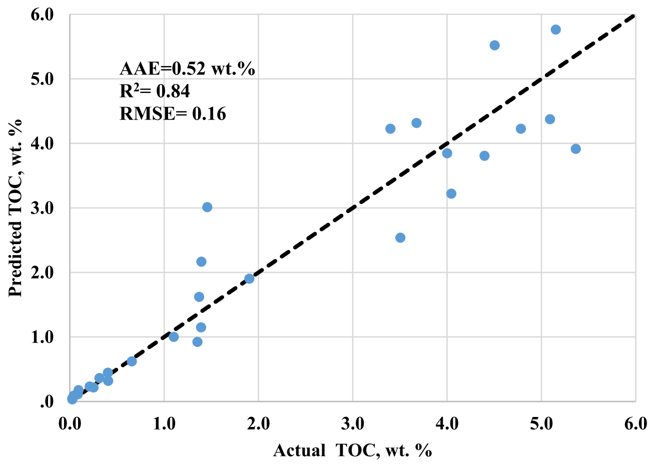


Figure 9. Cross plot between the actual measured total organic carbon (TOC) and the predicted values based on the artificial neural network (ANN) program, for the validation data.

weights and biases of the ANN-based model for determining the total organic content are given in Table 5

$$X_{j_n} = 2 \times \left(\frac{X_j - X_{j_{\min}}}{X_{j_{\max}} - X_{j_{\min}}} \right) - 1 \tag{2}$$

3.4. Model Validation. The developed TOC model was validated using hidden data that were not used previously. We have chosen the validation data randomly; however, it is ensured that a reasonable TOC range was covered. Figure 9 shows a cross plot between the measured TOC values and the predicted one using the proposed ANN correlation (eq 1). The TOC values were predicted with an AAE of 0.52 wt % and a CorrCoef of 0.84. Also, the variation of the predicted and actual TOC values along with the depth, for the validation data, is given in Figure 10.

In addition, the performance of the developed model in predicting the TOC was compared with different models available in the literature. The conventional logs such as density, gamma ray, resistivity, and sonic logs were used to estimate the TOC utilizing Passey's model, Schmoker's model, and the ANN model developed by Mahmoud et al.[15] In this work, we compared and presented our predictions with the Mahmoud et al.'s[15] model because it is one of the effective and most recent methods developed for TOC determinations. Also, their developed model was able to outperform all previous TOC models.[15] It should be mentioned that ANN models are empirical models that have been built using a certain set(s) of data, and the weights and biases of Mahmoud et al.'s model were determined using a special set and values of logging tools. However, we are comparing our developed model with Mahmoud et al.'s model since the formation studied in this work is similar to Mahmoud et al.'s formation, where the TOC is varying between 0.1 and 6 wt %, the gamma ray is changing between 20 and 200 API, and the bulk formation density is varying between 2.4 and 2.8 g/cm^3, with very few points are outside these ranges. Figure 11 shows the actual against the predicted TOC using the ANN model developed in this work and the model proposed by Mahmoud et al.[15] The correlation of determination was increased from 0.28 to 0.73, and the AAE was reduced from 1.56 to 0.55 wt % using the ANN model developed in this work, indicating the good performance of the newly developed TOC model.

4. CONCLUSIONS

This work examines the use of thermal neutron logs to predict the TOC using the ANN approach. More than 150 data sets were used to build and validate the proposed TOC model. Based on the current study, the following conclusions can be drawn:

- Thermal logs can be effectively used to predict the TOC due to the relationship between neutron count rates and TOC.
- The CorrCoef analysis confirmed the strong relationship between the TOC and CFTC rate.

Table 5. Weights and Biases of ANN-Based Model for Determining the Total Organic Content

	weights between input and middle layer (w_1)							
hidden layer neurons (N)	CFTC (HZ)	NPHI (V/V)	NPOR (V/V)	PEFZ (B/E)	SIGF (M-1)	weights between hidden and target layer (w_2)	hidden layer bias (b_1)	target layer bias (b_2)
1	0.711	1.838	−2.525	2.376	0.425	−2.173	−0.936	1.111
2	−2.967	−1.052	0.073	0.016	2.360	−0.297	−1.176	
3	−1.693	2.912	−1.767	0.443	5.244	−1.656	1.873	
4	−0.120	−0.601	0.248	−0.471	−2.955	1.187	1.150	
5	0.422	1.281	2.635	0.804	3.634	3.282	1.781	
6	3.285	3.801	−0.274	1.240	−2.803	0.121	−0.661	
7	−1.729	−0.319	−3.165	0.753	−1.208	−0.740	−2.883	
8	2.150	−0.764	1.627	−3.518	−0.423	−0.365	−1.722	

https://doi.org/10.1021/acsomega.2c06918
ACS Omega 2023, 8, 4790−4801

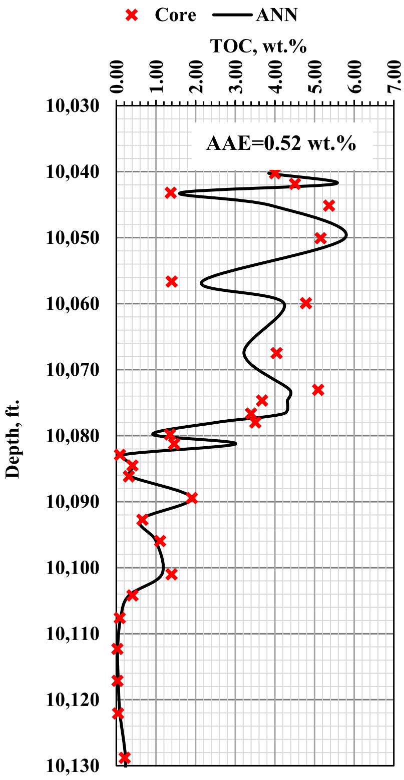


Figure 10. Variation of the predicted and actual total organic carbon (TOC) values along with the depth, for the validation data.

- Transferring the TOC data into the logarithmic domain showed an improvement in the CorrCoef values, leading to a significant increase in the prediction accuracy.
- Also, a new correlation was developed based on the optimized ANN model. The new correlation can predict the TOC values with high accuracy; the AAE is 0.52 wt % on average.
- The proposed TOC model was able to outperform the previous model, and R^2 was increased from 0.28 to 0.73.
- Overall, the proposed TOC model can give a real-time determination of the organic matter, which will help in identifying the zones of mature organic matter.

■ AUTHOR INFORMATION

Corresponding Authors

Amjed Hassan − *College of Petroleum Engineering & Geosciences, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* orcid.org/0000-0002-4487-6982; Email: amjed.mohammed@kfupm.edu.sa

Mohamed Mahmoud − *College of Petroleum Engineering & Geosciences, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* orcid.org/0000-0002-4395-9567; Email: mmahmoud@kfupm.edu.sa

Authors

Emad Mohammed − *College of Petroleum Engineering & Geosciences, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia*

Ali Oshaish − *College of Petroleum Engineering & Geosciences, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia*

Dhafer Badhafere − *College of Petroleum Engineering & Geosciences, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* orcid.org/0000-0002-2007-1469

Korhan Ayranci − *College of Petroleum Engineering & Geosciences, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia*

Tian Dong − *Key Laboratory of Tectonics and Petroleum Resources of Ministry of Education, China University of Geosciences, Wuhan 430074, China*

Umair bin Waheed − *College of Petroleum Engineering & Geosciences, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia*

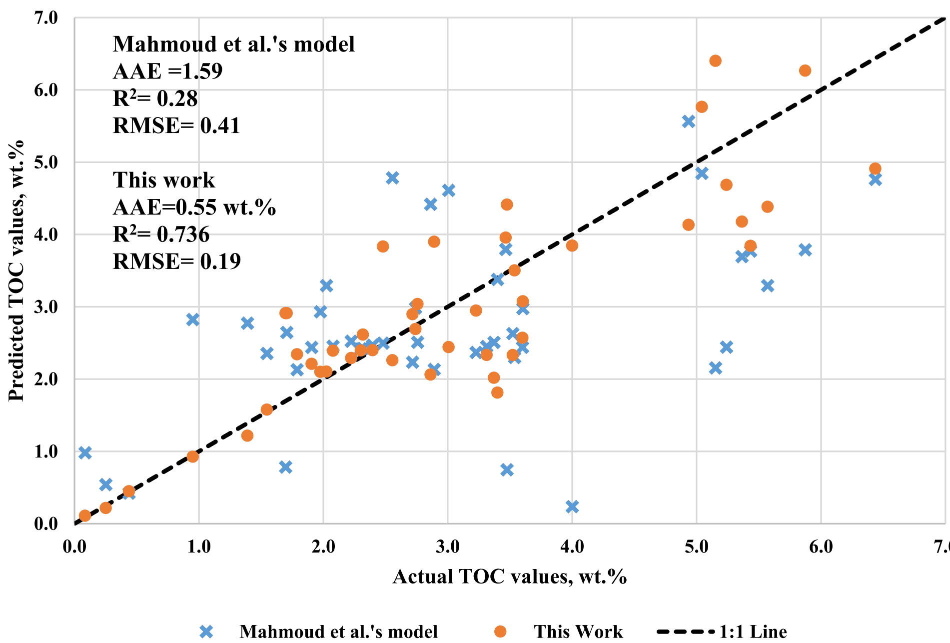


Figure 11. Actual total organic carbon (TOC) plotted against the predicted values using the model developed in this work and Mahmoud et al.'s model.

https://doi.org/10.1021/acsomega.2c06918
ACS Omega 2023, 8, 4790−4801

Ammar El-Husseiny − *College of Petroleum Engineering & Geosciences, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia;* orcid.org/0000-0001-5762-6109

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsomega.2c06918

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The College of Petroleum and Geoscience at King Fahd University of Petroleum & Minerals is acknowledged for the support and permission to publish this work.

■ ABBREVIATION

AC, acoustic log
CNL, neutron porosity
DEN, lithology density
DT, decision tree
RHOB, bulk density log
RLLD, deep lateral resistivity log
RT, resistivity of uninvaded zone
SP, spontaneous log
TH, thorium

■ REFERENCES

(1) Chan, S. A.; Hassan, A. M.; Usman, M.; Humphrey, J. D.; Alzayer, Y.; Duque, F. Total organic carbon (TOC) quantification using artificial neural networks: Improved prediction by leveraging XRF data. *J. Pet. Sci. Eng.* **2022**, *208*, 109302.

(2) Ahangari, D.; Daneshfar, R.; Zakeri, M.; Ashoori, S.; Soulgani, B. S.*On the prediction of geochemical parameters (TOC, S1 and S2) by considering well log parameters using ANFIS and LSSVM strategies*; Petroleum, 2021.

(3) Curtis, J. B. Fractured shale-gas systems. *AAPG Bull.* **2002**, *86*, 1921−1938.

(4) Iqbal, O.; Ahmad, M.; Abd Kadir, A. A. Geomechanical characterization of potential Roseneath shale gas, Cooper basin, Australia. *ARPN J. Eng. Appl. Sci.* **2017**, *12*, 5221−5231.

(5) Sfidari, E.; Kadkhodaie-Ilkhchi, A.; Najjari, S. Comparison of intelligent and statistical clustering approaches to predicting total organic carbon using intelligent systems. *J. Pet. Sci. Eng.* **2012**, *86-87*, 190−205.

(6) Yoon, G.; Park, S. M.; Yang, H.; Tsang, D. C.; Alessi, D. S.; Baek, K. Selection criteria for oxidation method in total organic carbon measurement. *Chemosphere* **2018**, *199*, 453−458.

(7) Bai, Y.; Tan, M. Dynamic committee machine with fuzzy-c-means clustering for total organic carbon content prediction from wireline logs. *Comput. Geol.* **2021**, *146*, 104626.

(8) Cheng, S.; Gao, X.; Cao, L.; Wang, Q.; Qiao, Y. Quantification of Total Organic Carbon in Ashes from Smoldering Combustion of Sewage Sludge via a Thermal Treatment—TGA Method. *ACS Omega* **2020**, *5*, 33445−33454.

(9) Tariq, Z.; Mahmoud, M.; Abouelresh, M.; Abdulraheem, A. Data-driven approaches to predict thermal maturity indices of organic matter using artificial neural networks. *ACS Omega* **2020**, *5*, 26169−26181.

(10) Passey, Q. R.; Bohacs, K. M.; Esch, W. L.; Klimentidis, R.; Sinha, S.From oil-prone source rock to gas-producing shale reservoir−geologic and petrophysical characterization of unconventional shale-gas reservoirs. In *International Oil and Gas Conference and Exhibition in China*; OnePetro, 2010, June.

(11) Kim, J.; Kim, D.; Lee, W.; Lee, Y.; Kim, H. Impact of total organic carbon and specific surface area on the adsorption capacity in Horn River shale. *J. Pet. Sci. Eng.* **2017**, *149*, 331−339.

(12) Dong, T.; Harris, N. B.; Ayranci, K. Relative sea-level cycles and organic matter accumulation in shales of the Middle and Upper Devonian Horn River Group, northeastern British Columbia, Canada: Insights into sediment flux, redox conditions, and bioproductivity. *GSA Bull.* **2017**, *130*, 859−880.

(13) Dong, T.; Harris, N. B.; Ayranci, K.; Yang, S. The impact of rock composition on geomechanical properties of a shale formation: Middle and Upper Devonian Horn River Group shale, Northeast British Columbia, Canada. *Am. Assoc. Pet. Geol. Bull.* **2017**, *101*, 177−204.

(14) Handhal, A. M.; Al-Abadi, A. M.; Chafeet, H. E.; Ismail, M. J. Prediction of total organic carbon at Rumaila oil field, Southern Iraq using conventional well logs and machine learning algorithms. *Mar. Pet. Geol.* **2020**, *116*, 104347.

(15) Mahmoud, A. A. A.; Elkatatny, S.; Mahmoud, M.; Abouelresh, M.; Abdulraheem, A.; Ali, A. Determination of the total organic carbon (TOC) based on conventional well logs using artificial neural network. *Int. J. Coal Geol.* **2017**, *179*, 72−80.

(16) Behar, F.; Beaumont, V.; Penteado, H. D. B. Rock-Eval 6 technology: performances and developments. *Oil Gas Sci. Technol.* **2001**, *56*, 111−134.

(17) Lafargue, E.; Marquis, F.; Pillot, D. Rock-Eval 6 applications in hydrocarbon exploration, production, and soil contamination studies. *Rev. Inst. Fr. Pet.* **1998**, *53*, 421−437.

(18) Hazra, B.; Dutta, S.; Kumar, S. TOC calculation of organic matter rich sediments using Rock-Eval pyrolysis: Critical consideration and insights. *Int. J. Coal Geol.* **2017**, *169*, 106−115.

(19) Peters, K. E.; Xia, X.; Pomerantz, A. E.; Mullins, O. C.Geochemistry applied to evaluation of unconventional resources. In *Unconventional oil and gas resources handbook (71-126)*; Gulf Professional Publishing, 2016.

(20) Peters, K. E.; Simoneit Bernd, R. T.*Rock-Eval pyrolysis of quaternary sediments from Leg 64, sites 479 and 480*; Gulf of California, 1982.

(21) Wimberley, J. W. A rapid method for the analysis of total and organic carbon in shale with a high-frequency combustion furnace. *Anal. Chim. Acta* **1969**, *48*, 419−423.

(22) Engleman, E. E.; Jackson, L. L.; Norton, D. R. Determination of carbonate carbon in geological materials by coulometric titration. *Chem. Geol.* **1985**, *53*, 125−128.

(23) Elias, R.; Duclerc, D.; Le-Van-Loi, R.; Gelin, F.; Dessort, D.A new geochemical tool for the assessment of organic-rich shales. In *SPE/AAPG/SEG Unconventional Resources Technology Conference*; OnePetro, 2013, August.

(24) Herron, M. M.; Loan, M.; Charsky, A. M.; Herron, S. L.; Pomerantz, A. E.; Polyakov, M. Kerogen content and maturity, mineralogy and clay typing from DRIFTS analysis of cuttings or core. *Petrophysics* **2014**, *55*, 435−446.

(25) Mao, J. D.; Hundal, L. S.; Schmidt-Rohr, K.; Thompson, M. L. Nuclear magnetic resonance and diffuse-reflectance infrared Fourier transform spectroscopy of biosolids-derived biocolloidal organic matter. *Environ. Sci. Technol.* **2003**, *37*, 1751−1757.

(26) Russell, W. L. Relation of radioactivity, organic content, and sedimentation. *AAPG Bull.* **1945**, *29*, 1470−1493.

(27) Patterson, C. D.; Quarles, C. A.; Breyer, J. A. Possible new well-logging tool using positron Doppler broadening to detect total organic carbon (TOC) in hydrocarbon source rocks. *Radiat. Phys. Chem.* **2003**, *68*, 523−526.

(28) Jumat, N.; Shalaby, M. R.; Islam, M. A. An integrated source rock characterization using geochemical analysis and well logs: a case study of Taranaki Basin, New Zealand. *Pet. Coal* **2017**, *59*, 884−910.

(29) Wang, P.; Peng, S.; He, T. A novel approach to total organic carbon content prediction in shale gas reservoirs with well logs data, Tonghua Basin, China. *J. Nat. Gas Sci. Eng.* **2018**, *55*, 1−15.

(30) Schmoker, J. W. Determination of organic-matter content of Appalachian Devonian shales from gamma-ray logs. *AAPG Bull.* **1981**, *65*, 1285−1298.

(31) Rivard, B.; Harris, N. B.; Feng, J.; Dong, T. Inferring total organic carbon and major element geochemical and mineralogical character-

istics of shale core from hyperspectral imagery. *AAPG Bull.* **2018**, *102*, 2101−2121.
(32) Passey, Q. R.; Creaney, S.; Kulla, J. B.; Moretti, F. J.; Stroud, J. D. A practical model for organic richness from porosity and resistivity logs. *AAPG Bull.* **1990**, *74*, 1777−1794.
(33) Zeng, B.; Li, M.; Zhu, J.; Wang, X.; Shi, Y.; Zhu, Z.; Guo, H.; Wang, F. Selective methods of TOC content estimation for organic-rich interbedded mudstone source rocks. *J. Nat. Gas Sci. Eng.* **2021**, *93*, 104064.
(34) Hou, H.; Qin, Q.; Shao, L.; Liang, G.; Tang, Y.; Zhang, H.; Li, Q.; Liu, S. Study on the Applicability of Reservoir Fractal Characterization in Middle−High Rank Coals with NMR: Implications for Pore-Fracture Structure Evolution within the Coalification Process. *ACS Omega* **2021**, *6*, 32495−32507.
(35) Al-AbdulJabbar, A.; Elkatatny, S.; Mahmoud, M.; Abdulraheem, A.Predicting formation tops while drilling using artificial intelligence. In *SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition*; OnePetro, 2018, April.
(36) Carpenter, C. Artificial neural network models and predicts reservoir parameters. *J. Pet. Technol.* **2021**, *73*, 44−45.
(37) Ahmed S, A.; Elkatatny, S.; Ali, A. Z.; Abdulraheem, A.; Mahmoud, M.Artificial neural network ANN approach to predict fracture pressure. In *SPE middle east oil and gas show and conference*; OnePetro, 2019, March.
(38) Al-Khdheeawi, E. A.; Mahdi, D. S.; Feng, R.Lithology determination from drilling data using artificial neural network. In *53rd US Rock Mechanics/Geomechanics Symposium*; OnePetro, 2019, June.
(39) Huang, Z.; Williamson, M. A. Artificial neural network modelling as an aid to source rock characterization. *Mar. Pet. Geol.* **1996**, *13*, 277−290.
(40) Salem, K. G.; Gad, K.; Abdulaziz, A. M.; Aziz, A.; Abdel Sattar, A., Dahab, A. S.Prediction of hydraulic properties in carbonate reservoirs using artificial neural network. In *Abu Dhabi international Petroleum Exhibition & Conference.*; OnePetro, 2018, November.
(41) Xie, L.; Chen, Z.; Chaudhary, S.; Alegria, A. E.A Novel Rheological Hierarchy Optimization Methodology with Artificial Intelligence and Inverse Technique Improves Spacer Design. In *SPE/IADC Drilling Conference and Exhibition*; OnePetro, 2017, March.
(42) Nie, X.; Wan, Y.; Bie, F. Dual-shale-content method for total organic carbon content evaluation from wireline logs in organic shale. *Open Geosci.* **2017**, *9*, 133−137.
(43) Zhu, L.; Zhang, C.; Zhang, C.; Zhang, Z.; Zhou, X.; Liu, W.; Zhu, B. 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. *J. Pet. Sci. Eng.* **2020**, *188*, 106944.
(44) Shi, X.; Wang, J.; Liu, G.; Yang, L.; Ge, X.; Jiang, S. Application of extreme learning machine and neural networks in total organic carbon content prediction in organic shale with wire line logs. *J. Nat. Gas Sci. Eng.* **2016**, *33*, 687−702.
(45) Amosu, A.; Imsalem, M.; Sun, Y. Effective machine learning identification of TOC-rich zones in the Eagle Ford Shale. *J. Appl. Geophys.* **2021**, *188*, 104311.
(46) Wang, H.; Wu, W.; Chen, T.; Dong, X.; Wang, G. An improved neural network for TOC, S1 and S2 estimation based on conventional well logs. *J. Pet. Sci. Eng.* **2019**, *176*, 664−678.
(47) Zheng, D.; Wu, S.; Hou, M. Fully connected deep network: An improved method to predict TOC of shale reservoirs from well logs. *Mar. Pet. Geol.* **2021**, *132*, 105205.
(48) Khoshnoodkia, M.; Mohseni, H.; Rahmani, O.; Mohammadi, A. TOC determination of Gadvan Formation in South Pars Gas field, using artificial intelligent systems and geochemical data. *J. Pet. Sci. Eng.* **2011**, *78*, 119−130.
(49) Shalaby, M. R.; Malik, O. A.; Lai, D.; Jumat, N.; Islam, M. A. Thermal maturity and TOC prediction using machine learning techniques: case study from the Cretaceous−Paleocene source rock, Taranaki Basin, New Zealand. *J. Pet. Explor. Prod. Technol.* **2020**, *10*, 2175−2193.
(50) Wang, P.; Peng, S. A new scheme to improve the performance of artificial intelligence techniques for estimating total organic carbon from well logs. *Energies* **2018**, *11*, 747.
(51) Alizadeh, B.; Najjari, S.; Kadkhodaie-Ilkhchi, A. 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 Gas Field, Persian Gulf, Iran. *Comput. Geosci.* **2012**, *45*, 261−269.
(52) Tan, M.; Song, X.; Yang, X.; Wu, Q. Support-vector-regression machine technology for total organic carbon content prediction from wireline logs in organic shale: A comparative study. *J. Nat. Gas Sci. Eng.* **2015**, *26*, 792−802.
(53) Wang, P.; Chen, Z.; Pang, X.; Hu, K.; Sun, M.; Chen, X. Revised models for determining TOC in shale play: Example from Devonian Duvernay shale, Western Canada sedimentary basin. *Mar. Pet. Geol.* **2016**, *70*, 304−319.
(54) Zhao, P.; Ma, H.; Rasouli, V.; Liu, W.; Cai, J.; Huang, Z. An improved model for estimating the TOC in shale formations. *Mar. Pet. Geol.* **2017**, *83*, 174−183.
(55) Rui, J.; Zhang, H.; Zhang, D.; Han, F.; Guo, Q. Total organic carbon content prediction based on support-vector-regression machine with particle swarm optimization. *J. Pet. Sci. Eng.* **2019**, *180*, 699−706.
(56) Lawal, L. O.; Mahmoud, M.; Alade, O. S.; Abdulraheem, A. Total organic carbon characterization using neural-network analysis of XRF data. *Petrophysics* **2019**, *60*, 480−493.
(57) Sultan, A.New artificial neural network model for predicting the TOC from well logs. In *SPE Middle East Oil and Gas Show and Conference*; OnePetro, 2019, March.
(58) Broglia, C.; Ellis, D. Effect of alteration, formation absorption, and standoff on the response of the thermal neutron porosity log in gabbros and basalts: Examples from Deep Sea Drilling Project-Ocean Drilling Program Sites. *J. Geophys. Res.: Solid Earth* **1990**, *95*, 9171−9188.
(59) Clavier, C.; Hoyle, W.; Meunier, D. Quantitative interpretation of thermal neutron decay time logs: part I. Fundamentals and techniques. *J. Pet. Technol.* **1971**, *23*, 743−755.
(60) Harris, N. B.; Moghadam, A.; Dong, T. The effects of organic carbon content and thermal maturity on acoustic parameters in a black shale: Woodford Shale, Permian Basin, West Texas. *Geophysics* **2019**, *84*, 231−248.
(61) Ayranci, K.; Harris, N.; Dong, T. Sedimentological and ichnological characterization of the Middle to Upper Devonian Horn River Group, British Columbia, Canada: Insights into mudstone depositional conditions and processes below storm wave base. *J. Sediment. Res.* **2018**, *88*, 1−23.

https://doi.org/10.1021/acsomega.2c06918
ACS Omega 2023, 8, 4790−4801