Arabian Journal for Science and Engineering (2019) 44:6127–6137
https://doi.org/10.1007/s13369-018-3672-6

RESEARCH ARTICLE - PETROLEUM ENGINEERING

A Self-Adaptive Artificial Neural Network Technique to Predict Total Organic Carbon (TOC) Based on Well Logs

Salaheldin Elkatatny[1]

Received: 9 August 2018 / Accepted: 28 November 2018 / Published online: 4 December 2018

Abstract

Determination of total organic carbon (TOC) is a key method of characterizing shale reservoirs. The conventional method for TOC determination using cores from shale reservoirs is time-consuming and costly. TOC can be estimated by an indirect method using petrophysical well logs. The existing models assume a linear relationship between the well logs and TOC and have a high error and low correlation coefficients (CC) between the actual and predicted TOC. The first goal of this study is to apply a self-adaptive differential evolution (SaDE) optimizing method to determine the best combination of artificial neural network (ANN) parameters (number of hidden layers, number of neurons in each layer, training function, transferring function, and the training over testing ratio). The second goal is to develop a new empirical correlation that can be used to estimate TOC using well logs based on the optimized SaDE-ANN model. Four-hundred and sixty data points from Barnett shale were used for training and testing the developed SaDE-ANN model. Another set of data (29 data points) of Duvernay shale was used to compare the developed TOC correlation with the previous models. The obtained results show that the developed SaDE-ANN model predicted the TOC using only well logs: gamma ray (GR), compressional time (DT), deep resistivity (RD), and bulk density (RHOB) with a high accuracy (a CC of 0.99 and the average absolute percentage error (AAPE) of 6%). The developed TOC correlation outperformed the models proposed by Wang et al. (Mar Pet Geol 70:304–319, 2016. https://doi.org/10.1016/j.marpetgeo.2015.11.023) and Abdulhamid et al. (Int J Coal Geol 179(15):72–80, 2017). The new empirical correlation for TOC estimation reduced AAPE by 67% as compared with the ANN model developed by Abdulhamid et al. (2017) for the Duvernay formation. The developed TOC correlation is simple and can be applied using any computer without the need for the ANN model or special software. The developed technique will help reservoir engineers and geologists to estimate the TOC values using only the well logs with a high accuracy.

Keywords Self-Adaptive · Artificial neural network · Total organic carbon · Well logs · Barnett shale · Duvernay formation

Abbreviations

SaDE	Self-adaptive differential evolution
ANN	Artificial neural network
AI	Artificial intelligence
CC	Correlation coefficient
AAPE	Average absolute percentage error
TOC	Total organic carbon, wt%
DT	Compressional time, us/ft
RD	Deep resistivity, ohm m
RHOB	Bulk density, g/cm^3
GR	Gamma ray, API
N	Number of neurons
w	Weight associated with a layer and a hidden layer
i	The index of each neuron in a hidden layer
b	Bias associated with a hidden layer

✉ Salaheldin Elkatatny
elkatatny@kfupm.edu.sa

[1] Department of Petroleum Engineering, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

1 Introduction

The ability to recover hydrocarbons from unconventional resources (i.e., low porosity–permeability reservoirs) such as tight oil and shale gas was recently extended by the advances in horizontal drilling and multistage fracturing. TOC is one of the essential parameters that affect reserve estimation and hydraulic fracture design [1,2]. In addition, organic matter, similar to all other rock matrix components, has a considerable effect on the geomechanical properties of a shale

formation. These components are essential for developing such reservoirs [3]. Moreover, the maturity and carbon content control the organic porosity of a reservoir and, hence, affect the amount of gas adsorbed by the organic matter [4,5].

The microstructure, texture, permeability, and reservoir wettability are also controlled by the TOC present in shales [2,3,6,7]. Thus, hydrocarbon exploration and production require a reliable method for characterizing the organic matter present in a shale formation [8,9].

Currently, TOC is estimated using well log data through the application of empirical correlations. These correlations are developed for specific formations and conditions and are based on certain assumptions.

Schmoker [10] developed the first empirical equation for TOC based on data from the Devonian formations. Equation 1 can be used to determine the vol% of TOC which can be converted to wt% as explained by Schmoker [10].

$$\text{TOC (vol\%)} = \frac{(\rho_B - \rho)}{1.378} \tag{1}$$

where ρ_B is the rock density without including the organic matter in g/cm^3 and ρ denotes the rock bulk density in g/cm^3.

Schmoker [11] modified the above model to be used for the Bakken shale formation by assuming the same pyrite–organic matter relationships used for developing the Devonian correlation, Eq. 2:

$$\text{TOC (wt\%)} = \frac{[(100\,\rho_o)\,(\rho - 0.9922\rho_{mi} - 0.039)]}{[(R\rho)\,(\rho_o - 1.135\rho_{mi} - 0.675)]} \tag{2}$$

where ρ_o represents the organic matter density in g/cm^3, R is the weight percent ratio of the organic matter to organic carbon, ρ_{mi} is the volume-weighted average density of the pore fluid and grain in g/cm^3.

Schmoker and Hester [12] developed a more generalized model for TOC of the Bakken formation represented by a simplified equation (Eq. 3), assuming that R is 1.3, ρ_o is 1.01 g/cm^3, and ρ_{mi} is 2.68 g/cm^3.

$$\text{TOC (wt\%)} = \frac{77.44}{\rho} - 28.7 \tag{3}$$

Passey et al. [13] developed a simplified method for the estimation of TOC (Eqs. 4 and 5), which found wide application in the field of unconventional resource evaluation.

$$\Delta \log R = \log_{10}\left(R/R_{\text{baseline}}\right) + 0.02 \times \left(\Delta t - \Delta t_{\text{baseline}}\right) \tag{4}$$

$$\text{TOC} = \Delta \log R \times 10^{(2.297 - 0.1688 \times \text{LOM})} \tag{5}$$

where $\Delta \log R$ is the logs separation, R is the resistivity in ohm m, Δt is the sonic transient time in μs/ft, R_{baseline} is the resistivity of the base formation in ohm m, $\Delta t_{\text{baseline}}$ is the base sonic transient time in μs/ft, and LOM is the level of maturity.

In many cases, these empirical equations failed in the evaluation of the TOC content in formations other than the formations they were developed for [8,14]. Abdulhamid et al. [14] developed an empirical correlation to estimate the TOC content using well log data. The optimized model based on artificial neural networks (ANN) they developed contains only 1 hidden layer with 5 neurons, and they tested it using data from the Duvernay formation. The correlation coefficient (CC) is 0.90, and average absolute percentage error (AAPE) is high (19%) when the predicted TOC data are compared to the actual data.

In this study, the self-adaptive differential evolution (SaDE) technique will be applied to optimize the ANN architecture to reduce the AAPE and increase the CC between the predicted and actual TOC data. The overall objective of the study is to develop a new empirical correlation to determine TOC with a high accuracy based on the optimized SaDE-ANN model.

1.1 Artificial Neural Networks

ANN is a computational method derived from the biological neural networks construction features [15]. A neural network is first trained on training data containing input–output mapping. After training the network using the training dataset, the testing data output could be predicted using the weighted average of the outputs of the training dataset. The weights are calculated using the Euclidean distance between the training and testing data [16,17]. Based on the interconnection between the different layers and neurons, the ANN could be classified into several types [18]. ANN has the ability to detect all possible integrations between predicted parameters, and it requires only nominal statistical training [19,20]

Artificial intelligence (AI) techniques have been extensively applied recently in the petroleum industry, especially for predicting the well or field performance. Alajmi et al. [21] predicted the choke performance using an ANN. Elkatatny et al. [22,23] applied ANN to comprehensively determine the permeability of heterogeneous reservoirs and to estimate the rheological properties of drilling fluids based on real-time measurements. They developed robust models that would help petroleum engineers make predictions with a high accuracy in a short time. Van and Chon [24,25] evaluated the performance of CO_2 flooding using ANN techniques. They developed ANN models for determining the oil production rate, CO_2 production, and the gas–oil ratio (GOR).

2 Methodology

This study is aimed to use a SaDE optimizing technique to accurately optimize the ANN-related parameters to precisely identify the TOC using well logs: DT, RD, RHOB, and GR. The ANN-related parameters to be optimized are the number of hidden layers, the number of neurons in each layer, training function, and transferring function. The advantage of the proposed model is its ability to automatically determine the training over testing ratio, which requires a very long time if determined using ANN without SaDE.

The optimization process was continued until the outcome of the proposed SaDE-ANN model is acceptable. (The values of estimated parameters from the model are very close to the corresponding experimental data.) The accuracy of the proposed model is determined using AAPE and CC.

After training the SaDE-ANN model on randomly selected training data, the model is validated on unseen testing data followed by extracting a new empirical correlation to estimate TOC using GR, RHOB, DT, and RD logs.

3 Data Description

Barnett shale is the major source rock in the Mississippian age, and it is considered the second source of productive unconventional shale gas in the USA with proven reserves up to 43.4 trillion cubic feet of natural gas [4,26–28]. Abouelresh and Slatt [29,30] found that Barnett shale is mainly organic-rich shale with variable ratios of siliceous, calcareous, and phosphatic content. Sedimentologically, Barnett shale is primarily a result of suspension settling and density currents as well as hyperpycnal flows from the adjacent shelf [29,30]. Loucks and Ruppel [31] states that Barnett shale is deposited in deep, long, and narrow sub-basins shielded from the open ocean. The resulting anoxic condition is the key factor which preserves the organic content of shale. The thickness of the Barnett shale reservoir ranges between 300 and 700 ft [32].

Bowker [33] has stated that the measured TOC in the Barnett shale ranges from less than 0.5 to more than 6 wt%, with an average of about 4.5 wt%. Type II oil-prone marine kerogens are the main hydrocarbon type in Barnett shale [26]. The gas is thermogenic, most probably generated from kerogen cracking and secondary cracking of previously formed non-migrated hydrocarbons. The maturity level of Barnett shale lies within the thermal gas generation window, with around greater than 140 BCFG/mi^2 (54.05 BCFG/km^2) gas in place [27]. Bowker [33] found that the average porosity of Barnett shale is 5.5% and the average water saturation is 25%. Fu et al. [34] found that the unconventional Barnett shale reservoir has elevated levels of gamma rays (generally greater than 90 API units).

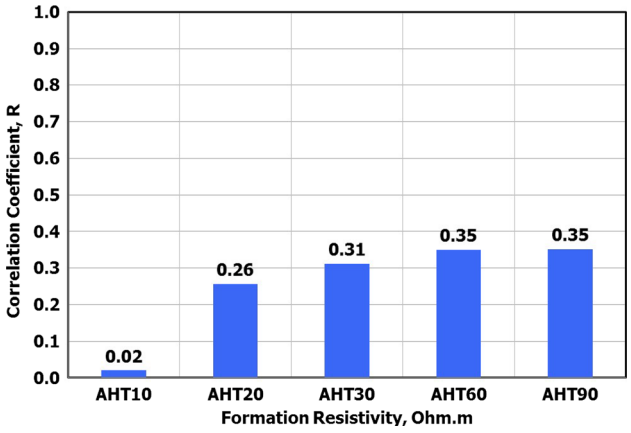

Fig. 1 Correlation coefficient of TOC with the formation resistivity

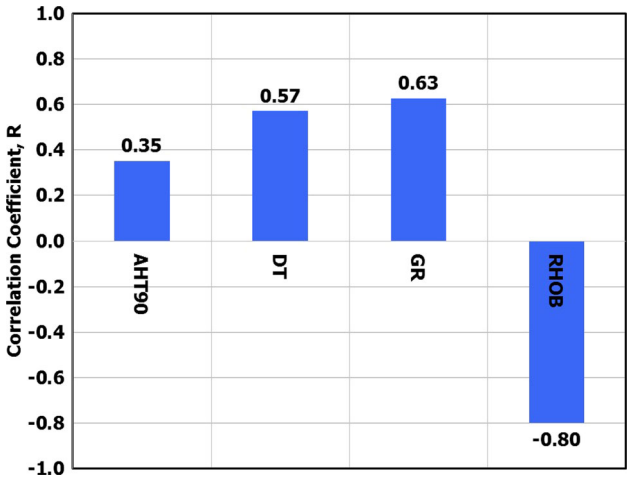

Fig. 2 Correlation coefficient of TOC with different logs

Actual TOC measurements (377 data points) and log data were collected to train and optimize the SaDE-ANN model. Figure 1 shows the relative importance between TOC and the resistivity logs. The correlation coefficient between TOC and the resistivity increases when the depth of invasion is increased. The relative importance increases from 0.02 to 0.35 when the depth of the resistivity log increases from 10″ to 60″ and remains constant when the resistivity depth increases to 90″. Based on this result, deep resistivity (AHT90) was used as an input for training the SaDE-ANN model.

Figure 2 shows that TOC has a strong function with the density, where the relative importance is −0.8, and a moderate function with GR and the compressional time. The relative importance of the TOC with GR and DT is 0.63 and 0.57, respectively.

Many researchers have reiterated the importance of including the aforementioned four parameters in predicting and modeling TOC. Passey et al. [13] and Heslop [35] stated that kerogen presence in the formation altered the deep

Table 1 Statistical parameters for the input data (377 training data points from Barnett shale)

	Gamma ray (API)	Resistivity (Ohm m)	Compressional time (us/ft)	Density (g/cm^3)	TOC (%)
Minimum	26	4	52	2.40	0.76
Maximum	180	172	97	2.77	5.55
Mean	92.74	47.46	76.22	2.57	2.60
Range	153.57	168.25	44.81	0.37	4.79
Standard deviation	32.04	43.34	10.98	0.08	1.18
Kurtosis	− 0.31	− 0.05	− 0.56	− 0.80	− 0.99
Skewness	− 0.35	1.06	− 0.51	0.34	0.23

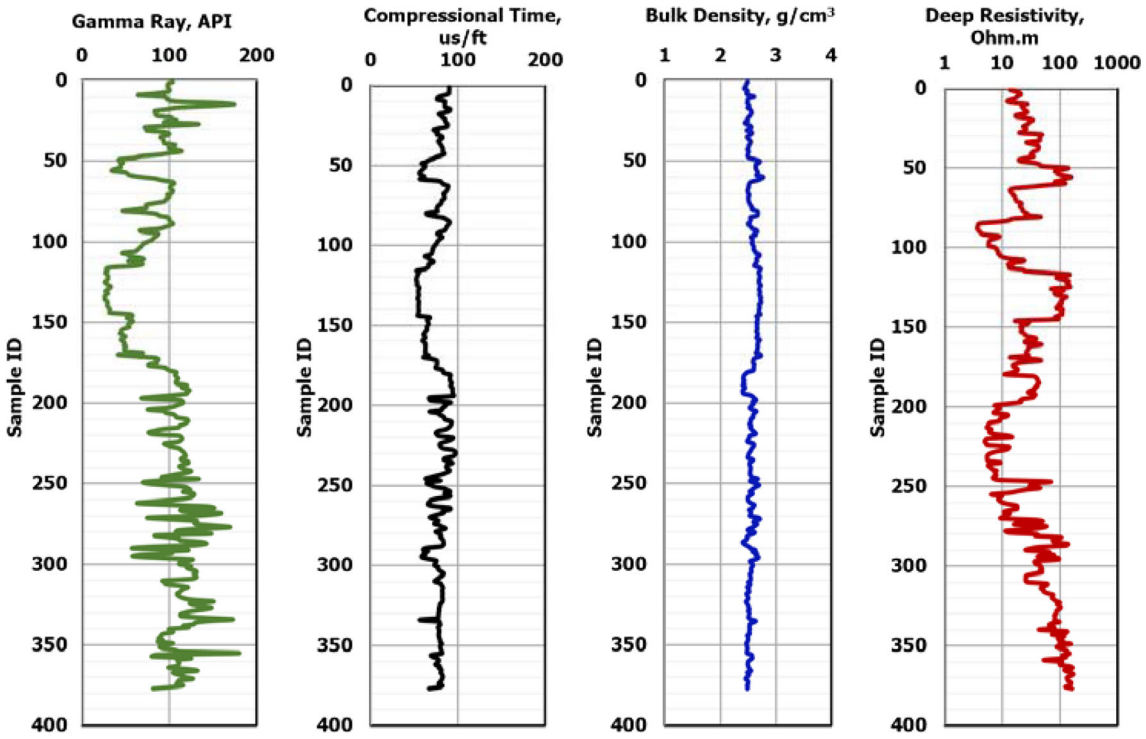

Fig. 3 Log data for Barnett shale formation (377 training data points)

resistivity measurement, since the increase in the kerogen aromaticity with the increase in the thermal maturity leads to the kerogen conductive behavior. Liu et al. [36] concluded that compressional time decreased with increasing the TOC in the formation; this is attributed to the presence of the kerogen which is characterized by the low transit velocity. Schmoker [10] stated that as the concentration of the kerogen which has considerably low density increased in a formation, the formation bulk density will be decreased. Fu et al. [34] mentioned that shale gas formation can be identified using gamma ray because of the high radioactivity of the organic content. The relationship between TOC and GR is not well understood [37–39]. Including GR in TOC models enhances the prediction performance of these models [8,35,40].

Table 1 lists the statistical parameters of the training data (377 data points) for Barnett shale. TOC ranges from 0.76 to 5.55 wt%; formation density ranges from 2.40 to 2.77 g/cm^3; compressional time ranges from 52 to 97 us/ft; deep resistivity ranges from 4 to 172 ohm m; and GR ranges from 26 to 180 API degrees.

4 Building the SaDE-ANN Model

The first step in building the SaDE-ANN model is the training process, where the log data such as GR, RT, DT, and RHOB and the TOC data are used to train the model. Figure 3 shows the plots of the aforementioned logs. The input data should be normalized to be used for training the SaDE-ANN model. The following equations were extracted from the SaDE-ANN model:

$$GR_n = 0.01218\,(GR - 15.66) - 1 \quad (6)$$

$$RD_n = 0.0119\,(RD - 3.653) - 1 \quad (7)$$

$$DT_n = 0.0446\,(DT - 52.288) - 1 \quad (8)$$

$$RHOB_n = 5.1546\,(RHOB - 2.402) - 1 \quad (9)$$

$$TOC_n = 0.4127\,(TOC - 0.7) - 1 \quad (10)$$

where RD_n, GR_n, $RHOB_n$, DTn, and TOC_n are the logs of normalized deep induction resistivity, gamma rays, bulk formation density, compressional transient time, and total organic carbon, respectively.

The main reason for using the SaDE technique is to optimize the variable parameters for the ANN model. The results of the optimization process in terms of the lowest AAPE and the highest CC show that the optimum ANN design parameters which will generate the best results are the following:

- The ANN structure of 4–16–16–1 which indicates 4 neurons, representing the input parameters, in the input, 16 neurons in the first hidden layer, 16 neurons in the second hidden layer, and only one neuron, representing TOC, in the output layer.
- Training was conducted using 82% of the data points, while testing was performed using 18% of the data points.
- Levenberg–Marquardt back propagation (trainlm) as a function for training.
- Logarithmic sigmoid (logsig) as a function of transfer.

Figure 4 shows that the developed SaDE-ANN model predicted the TOC content with a high accuracy and a CC of 0.98 and an AAPE of 6.0%. Figure 5 shows that R^2 between the predicted and actual TOC using the optimized SaDE-ANN model is 0.97.

Another set of data from the same field (83 data points) was used for testing the developed SaDE-ANN model using unseen data. Table 2 lists the statistical parameters of the testing data; these data represents Barnet shale formation. The TOC ranges from 0.8 to 5.5 wt%; RHOB ranges from 2.4 to 2.7 g/cm^3; DT ranges from 52 to 94 us/ft; RD ranges from 4 to 163 ohm m; and GR ranges from 27 to 166 API degrees. The range of the testing data lies within the same range as the training data.

Figure 6 confirms the high accuracy of the developed SaDE-ANN model in predicting TOC, where the CC is 0.98 and the AAPE is 7.0%. Based on these results, it can be concluded that the SaDE-ANN model can be used to predict TOC as a function of well logs GR, DT, RHOB, and RD.

To avoid overfitting, 15% of the data points are randomly used to assess the generalization of the developed network. Figure 7 shows that the optimum number of iterations, which yields the lowest mean square error (MSE), is 24. The network performance on the testing dataset is also considered (where the MSE is calculated for each evaluation). The performance of both training and testing of the network must be good. If training alone is good, it will be considered as overfitting, and the network is not selected.

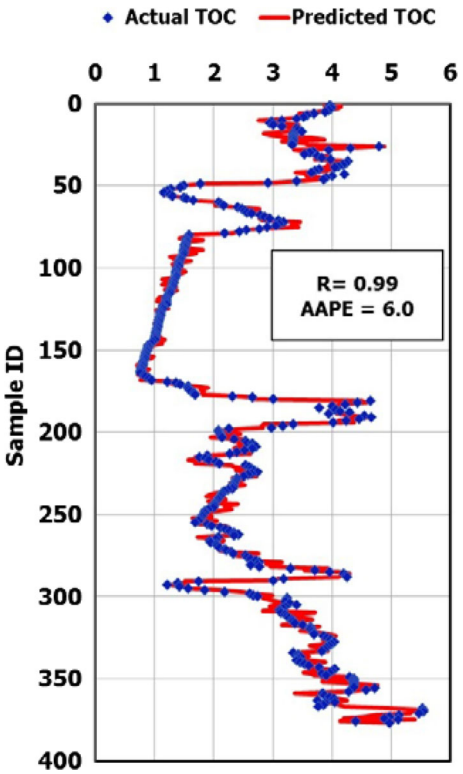

Fig. 4 Prediction of TOC using the SaDE-ANN model for training data (377 data points)

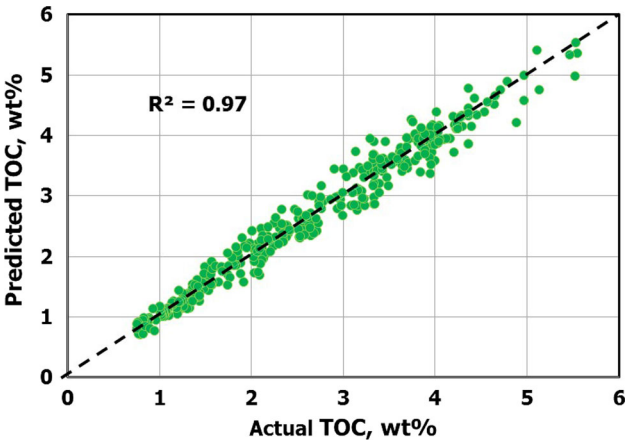

Fig. 5 Coefficient of determination for TOC prediction for training data (377 data points) using the SaDE-ANN model

5 A New Correlation for TOC Estimation

To change the developed SaDE-ANN model to a white box, a new empirical equation for estimating the TOC using only well log data was developed based on the optimized parameters (weights and biases) of the SaDE-ANN model as follows:

Table 2 Statistical parameters for the unseen data (83 testing data points from Barnett shale)

	Gamma ray (API)	Resistivity (Ohm m)	Compressional time (us/ft)	Density (g/cm^3)	TOC (%)
Minimum	27	4	52	2.4	0.8
Maximum	166	163	94	2.7	5.5
Mean	91.92	53.66	74.92	2.58	2.51
Range	139.05	158.70	41.85	0.31	4.78
Standard deviation	33.41	43.65	10.40	0.08	1.22
Kurtosis	−0.70	−0.41	−0.47	−0.84	−0.83
Skewness	−0.30	0.79	−0.55	0.02	0.40

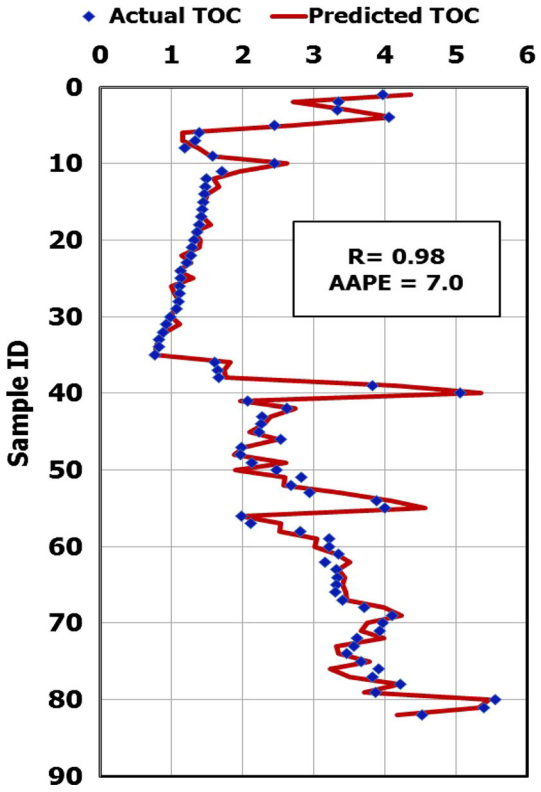

Fig. 6 Prediction of TOC using the SaDE-ANN model for unseen data (83 data points)

$$X_i = \frac{1}{1 + e^{-\left(w_{1_{i,1}}\mathrm{GR_n} + w_{1_{i,2}}\mathrm{RD_n} + w_{1_{i,3}}\mathrm{DT_n} + w_{1_{i,4}}\mathrm{RHOB_n} + b_{1i}\right)}} \tag{11}$$

Table 3 lists the different values of the weight (w_{1i}) and the biases (b_{1i}) for each input parameter in the first hidden layer. The optimization process shows that the optimum number of neurons in the first hidden layer (N_1) is 16, and that in the second layer (N_2) is also 16. Table 4 lists the weights (W_{2ji}) and biases (b_{2j}) associated with the hidden layers N_1 and N_2. Table 4 also lists the weights (w_{3j}) and the bias ($b_3 = -2.0042$) associated with the hidden layer N_2 and the output layer.

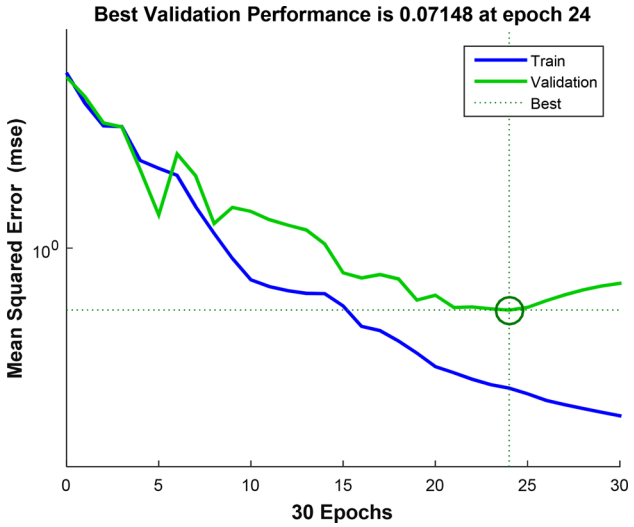

Fig. 7 Training and validation errors

Table 3 Weight and biases for the first hidden layer of the ANN-based TOC model

i	$w_{1_{i,1}}$	$w_{1_{i,2}}$	$w_{1_{i,3}}$	$w_{1_{i,4}}$	b_{1_i}
1	−4.29	3.57	−0.06	−0.40	6.33
2	−6.50	2.74	−0.13	1.93	6.18
3	−3.41	0.86	1.92	4.72	5.18
4	−4.30	4.26	−0.13	−5.47	1.57
5	−2.25	1.19	−4.83	−5.18	−3.94
6	0.74	2.59	4.82	−8.09	−2.27
7	−2.52	−7.56	3.51	−0.65	−5.39
8	6.50	0.24	3.70	1.35	−1.31
9	−4.31	1.46	−6.65	3.37	−1.82
10	−2.26	−2.27	5.21	3.40	−0.86
11	−5.55	1.94	−1.15	1.71	−3.63
12	−8.31	−1.76	−1.08	1.14	−2.70
13	−3.30	5.80	−2.15	−3.49	−2.64
14	3.11	−4.31	2.77	0.96	0.98
15	6.08	1.81	−2.77	3.44	4.64
16	4.58	3.67	−5.73	2.00	7.25

Table 4 Weight and biases for the second hidden layer of the ANN-based TOC model

j	$w_{2_{j,i}}$																b_{2_j}	w_{3_j}
	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$	$i=9$	$i=10$	$i=11$	$i=12$	$i=13$	$i=14$	$i=15$	$i=16$		
1	2.37	− 1.10	0.87	1.27	1.13	2.81	1.35	0.68	− 0.26	1.31	4.23	3.43	− 2.39	1.48	− 0.16	− 3.41	− 8.76	2.31
2	2.81	− 1.27	1.98	0.54	− 2.88	− 1.99	2.22	− 2.00	− 1.21	0.26	0.83	− 1.24	− 0.08	1.19	1.34	0.51	− 6.40	4.32
3	− 1.45	0.58	0.58	1.69	0.03	− 3.74	1.86	1.77	0.67	1.91	0.31	− 1.68	− 0.39	− 2.30	2.35	0.49	0.31	− 3.83
4	1.67	0.81	0.06	2.22	− 0.16	0.87	− 1.95	− 1.33	− 3.77	− 1.64	− 3.52	0.74	− 2.68	− 1.50	4.16	− 1.63	0.89	2.35
5	0.16	− 1.17	− 0.07	− 1.97	1.26	− 0.69	0.84	1.43	2.64	0.91	2.75	4.47	3.26	0.80	1.31	3.78	− 9.40	1.95
6	− 2.86	0.35	2.36	− 3.18	0.26	− 2.94	0.87	− 0.10	− 1.88	− 2.39	1.83	− 4.16	0.37	4.13	3.43	1.01	1.59	3.61
7	− 2.05	1.00	− 1.07	− 2.33	− 1.92	− 0.29	− 2.17	− 0.93	0.58	− 1.56	− 1.65	2.08	− 0.40	0.61	0.53	3.38	3.45	− 3.58
8	− 1.82	− 1.57	3.64	1.28	3.47	− 4.01	0.49	3.10	1.30	1.87	2.77	− 2.46	2.32	− 1.49	− 1.90	2.53	− 1.51	2.53
9	− 2.46	− 1.59	2.01	− 0.08	− 3.34	0.50	− 3.20	3.03	1.24	− 0.69	2.36	0.15	1.64	− 2.86	− 1.33	− 0.11	− 0.14	2.21
10	− 2.99	− 0.93	− 1.97	− 2.34	− 1.76	1.85	− 1.61	0.25	− 1.61	0.99	1.69	1.79	1.45	1.56	− 0.92	1.32	− 0.15	− 0.31
11	− 1.86	− 0.25	− 0.63	5.19	0.29	2.23	− 2.31	− 2.25	2.24	5.27	0.20	3.14	− 0.20	1.71	− 3.85	− 4.97	− 1.74	3.55
12	− 0.67	3.72	− 0.73	− 2.75	2.07	− 0.57	4.55	− 0.92	− 5.52	2.18	0.38	0.59	3.39	0.27	− 1.85	− 0.76	− 2.61	− 1.69
13	− 2.23	− 1.53	0.26	0.76	2.23	2.07	2.57	0.64	− 2.49	1.94	1.61	0.34	− 0.53	0.24	− 1.48	− 2.07	− 3.99	− 0.91
14	− 0.01	0.81	− 1.80	1.41	3.00	5.28	− 0.55	2.73	− 0.50	− 1.00	− 2.46	− 0.22	− 0.28	− 3.97	− 0.52	2.33	− 0.15	1.56
15	4.03	3.90	2.48	2.62	− 2.74	− 0.15	− 0.43	0.48	0.17	− 2.79	1.62	− 0.50	0.84	0.34	0.69	− 2.54	0.68	− 3.86
16	− 0.05	1.73	− 1.22	− 2.59	3.15	− 3.76	1.24	2.65	− 0.44	− 2.30	0.45	1.90	0.93	− 2.49	4.12	2.64	2.26	3.79

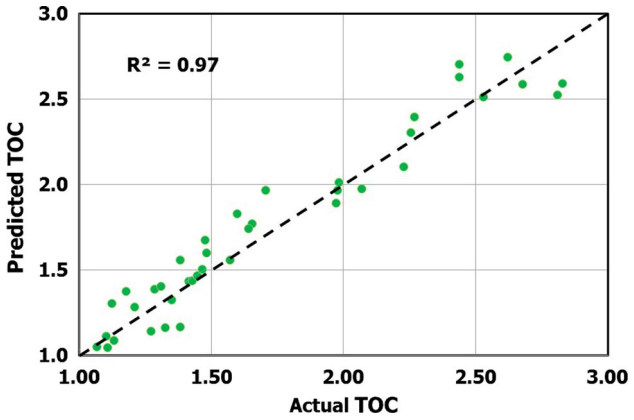

Fig. 8 Actual and predicted TOC for unseen data from Barnett shale using the developed correlation

$$Y_j = \sum_{i=1}^{N_1} w_{2_{j,i}} \frac{1}{1+e^{-X_i}} + b_{2_j} \tag{12}$$

$$\mathrm{TOC_n} = \sum_{j=1}^{N_2} w_{3_j} \frac{1}{1+e^{-Y_j}} + b_3 \tag{13}$$

The de-normalized value of TOC can be calculated by rewriting Eq. 10 as follows:

$$\mathrm{TOC} = \frac{\mathrm{TOC_n} + 1}{0.41217} + 0.7 \tag{14}$$

The developed correlation was tested using unseen data from Barnett shale. Figure 8 shows that the developed correlation (Eqs. 13 and 14) is able to predict TOC with a high accu-

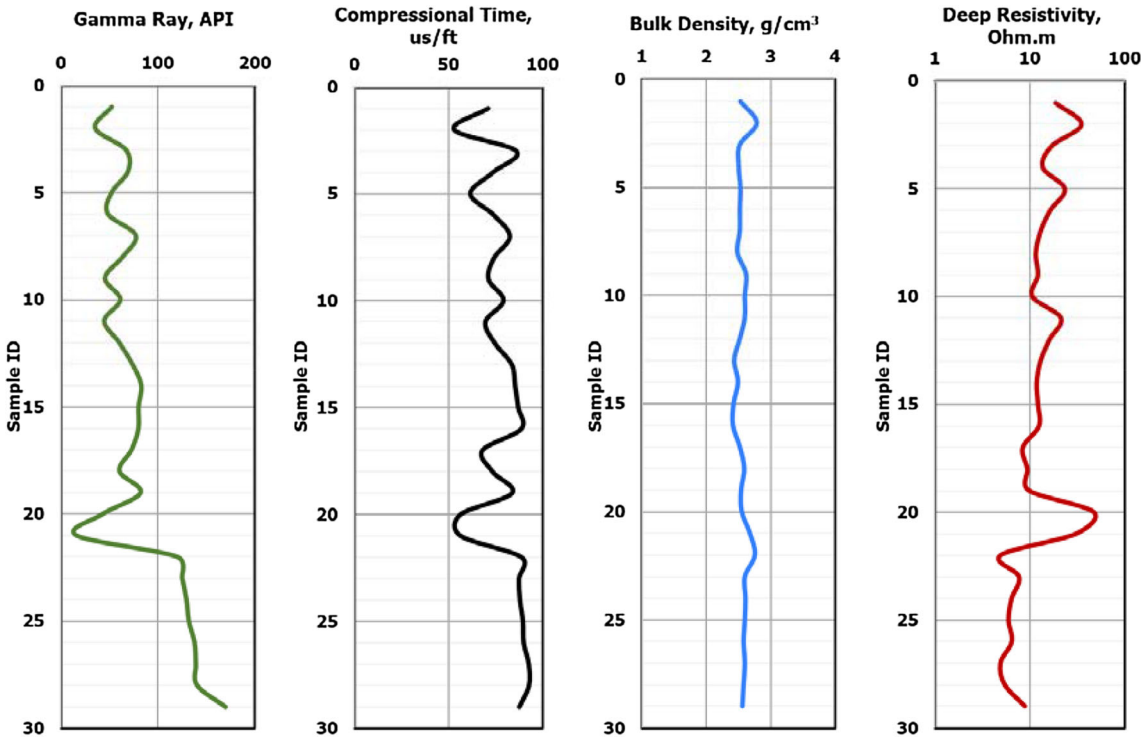

Fig. 9 Log data of Duvernay shale (29 data points)

Table 5 Statistical parameters for the validation data (29 data points from Duvernay shale)

	Gamma ray (API)	Resistivity (Ohm m)	Compressional time (us/ft)	Density (g/cm^3)	TOC (%)
Minimum	16	5	53	2.41	0.7
Maximum	169	47	93	2.77	5
Mean	81.02	14.22	78.24	2.56	2.77
Range	153.56	41.99	39.71	0.38	4.3
Standard deviation	38.46	9.55	11.45	0.09	1.36
Kurtosis	− 0.45	4.30	− 0.40	1.40	− 1.07
Skewness	0.67	1.95	− 0.74	0.76	0.24

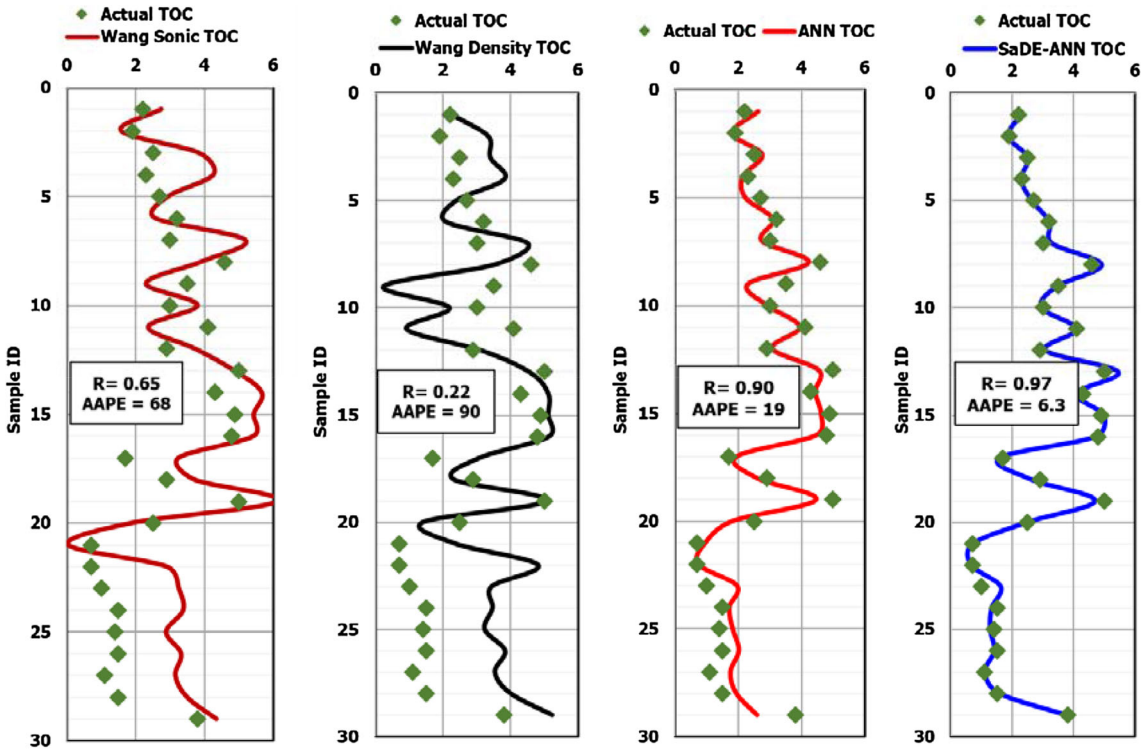

Fig. 10 Comparison of the developed SaDE-ANN model with other methods

racy. The coefficient of determination is 0.97 between the predicted and actual TOC data.

Another set of data from Duvernay formation was used for the comparison of the developed correlation for TOC with previous methods. The input log data used for the comparison are shown in Fig. 9. Table 5 lists the statistical parameters for the Duvernay shale data (29 data points). The TOC ranges from 0.7 to 5.0 wt%; RHOB ranges from 2.41 to 2.77 g/cm^3; DT ranges from 53 to 93 us/ft; RD ranges from 5 to 47 ohm m, and GR ranges from 16 to 169 API degrees. The range of the Duvernay formation data is within the range of the training data for the developed SaDE-ANN model.

Figure 10 shows that the Wang sonic model [8] overestimated TOC in the lower part of the data points (from 20 to 29) and the AAPE is very high (68%), in addition to the low CC (0.65). The density model proposed by Wang et al. [8] for TOC prediction yielded the worst results when compared with the sonic model, with a CC of 0.22 and an AAPE of 90%. The correlation developed based on the optimized ANN model [14] yielded better results as compared with the models proposed by Wang et al. [8], with a CC of 0.90 and an AAPE of 19%. The ANN-TOC correlation also overestimated TOC in the data from 20 to 29 as shown in Fig. 10.

The correlation developed based on the SaDE-ANN optimized model provides an accurate match between the actual and predicted TOC data. The AAPE is 6.3, and the CC is 0.97. Thus, using the SaDE-ANN model reduces the AAPE by 67% as compared with the ANN model. Figure 11 shows a comparison between the coefficient of the determination of the ANN model and the SaDE-ANN model. Applying the SaDE for ANN optimization yields a higher coefficient of determination of 0.98 as compared with 0.88 for ANN alone. The results of this study confirm the importance of applying the SaDE technique to optimize the ANN architecture in order to obtain the best results.

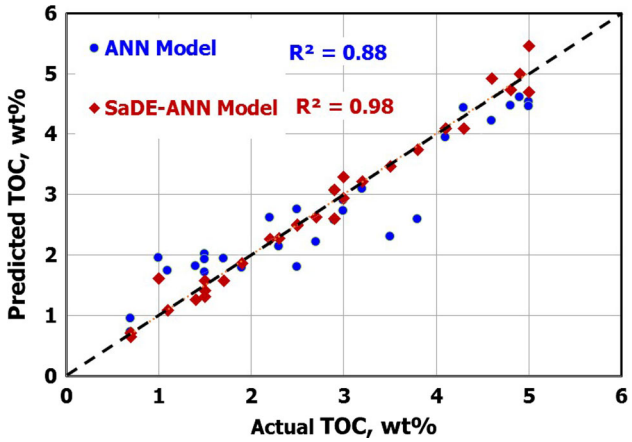

Fig. 11 Coefficient of determination of the actual and predicted TOC using the ANN and SaDE-ANN models for the validation data

6 Conclusions

A robust ANN model was built to accurately estimate the TOC. The following conclusions could be drawn out of the results obtained in this study:

1. The developed SaDE-ANN model predicted the TOC based on well logs: GR, DT, RD, and RHOB with a high accuracy (a CC of 0.99 and an AAPE of 6%).
2. The developed correlation for TOC estimation based on the SaDE-ANN model outperformed models proposed by Wang et al. [8] and Abdulhamid et al. [14].
3. The developed TOC correlation based on the SaDE-ANN model reduces the AAPE by 67% as compared to the ANN model proposed by Abdulhamid et al. [14] for the Duvernay formation.

Compliance with Ethical Standards

Conflict of interest The author has no conflicts of interest to declare.

References

1. Passey, Q.R.; Bohacs, K.; 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. Society of Petroleum Engineers (2010)
2. Sondergeld, C.H.; Ambrose, R.J.; Rai, C.S.; Moncrieff, J.: Micro-structural studies of gas shales. In: SPE Unconventional Gas Conference. Society of Petroleum Engineers (2010)
3. Altowairqi, Y.; Rezaee, R.; Evans, B.; Urosevic, M.: Shale elastic property relationships as a function of total organic carbon content using synthetic samples. J. Pet. Sci. Eng. **133**, 392–400 (2015)
4. Montgomery, S.L.; Jarvie, D.M.; Bowker, K.A.; Pollastro, R.M.: Mississippian Barnett Shale, Fort Worth basin, north-central Texas: gas-shale play with multi-trillion cubic foot potential. AAPG Bull. **89**(2), 155–175 (2005)
5. Ross, D.J.; Bustin, R.M.: Impact of mass balance calculations on adsorption capacities in microporous shale gas reservoirs. Fuel **86**(17), 2696–2706 (2007)
6. Sone, H.; Zoback, M.D.: Mechanical properties of shale-gas reservoir rocks—part 1: static and dynamic elastic properties and anisotropy. Geophysics **78**(5), D381–D392 (2013)
7. Zhang, T.; Ellis, G.S.; Ruppel, S.C.; Milliken, K.; Yang, R.: Effect of organic-matter type and thermal maturity on methane adsorption in shale-gas systems. Org. Geochem. **47**, 120–131 (2012)
8. 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. **70**, 304–319 (2016). https://doi.org/10.1016/j.marpetgeo.2015.11.023
9. Ding, J.; Xiaozhi, C.; Xiudi, J.; Bin, W.; Jinmiao, Z.: Application of AVF inversion on shale gas reservoir TOC prediction. In: SEG Annual Meeting: Society of Exploration Geophysicists (2015)
10. Schmoker, J.W.: Determination of organic content of Appalachian Devonian shales from formation-density logs. Am. Assoc. Pet. Geol. Bull. **63**, 1504–1509 (1979). https://doi.org/10.1306/2F9185D1-16CE-11D7-8645000102C1865D
11. Schmoker, J.W.: Organic content of Devonian shale in Western Appalachian Basin. AAPG Bull. **64**(12), 2156–2165 (1980)
12. Schmoker, J.W.; Hester, T.C.: Organic carbon in Bakken formation, United States portion of Williston basin. AAPG Bull. **67**(12), 2165–2174 (1983)
13. 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. **74**(12), 1777–1794 (1990)
14. Abdulhamid, A.; Elkatatny, S.M.; Mahmoud, M.A.; Aburesh, 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. **179**(15), 72–80 (2017)
15. Arabjamaloei,; Shadizadeh, S.: Modeling and optimizing rate of penetration using intelligent systems in an Iranian Southern Oil Field (Ahwaz Oil Field). Pet. Sci.Technol. **29**(16), 1637–1648 (2011). https://doi.org/10.1080/10916460902882818
16. Lippmann, R.: An introduction to computing with neural nets. IEEE ASSP Mag. **4**(2), 4–22 (1987)
17. Jain, A.K.; Mao, J.; Mohiuddin, K.M.: Artificial neural networks: a tutorial. Computer **29**(3), 31–44 (1996)
18. Goyal, S.; Goyal, G.K.: Cascade and feedforward backpropagation artificial neural network models for prediction of sensory quality of instant coffee flavoured sterilized drink. Can. J. Artif. Intell. Mach. Learn. Pattern Recognit. **2**(6), 78–82 (2011)
19. Vineis, P.; Rainoldi, A.: Neural networks and logistic regression: analysis of a case-control study on myocardial infarction. J. Clin. Epidemiol. **50**, 1309–1310 (1997). https://doi.org/10.1016/S0895-4356(97)00163-7
20. Burbidge, R.; Trotter, M.; Buxton, B.; Holden, S.: Drug design by machine learning: support vector machines for pharmaceutical data analysis. Comput. Chem. **26**, 5–14 (2001)
21. AlAjmi, M.D.; Alarifi, S.A.; Mahsoon, A.H.: Improving multiphase choke performance prediction and well production test validation using artificial intelligence: a new milestone. SPE-173394-MS, presented at the SPE Digital Energy Conference and Exhibition, held in the Woodlands, Texas, USA, 3–5 March 2015
22. Elkatatny, S.; Mahmoud, M.; Tariq, Z.; Abdulraheem, A.: New insights into the prediction of heterogeneous carbonate reservoir permeability from well logs using artificial intelligence network. Neural Comput. Appl. 1–11 (2017)
23. Elkatatny, S.; Tariq, Z.; Mahmoud, M.: Real time prediction of drilling fluid rheological properties using artificial neural networks visible mathematical model (white box). J. Pet. Sci. Eng. **146**, 1202–1210 (2016)
24. Van, S.L.; Chon, B.H.: Effective prediction and management of a CO_2 flooding process for enhancing oil recovery using artificial neural networks. ASME J. Energy Resour. Technol. (2017). https://doi.org/10.1115/1.4038054
25. Van, S.L.; Chon, B.H.: Evaluating the critical performances of a CO_2-enhanced oil recovery process using artificial neural network models. J. Pet. Sci. Eng. **157**(2017b), 207–222 (2017)
26. Pollastro, R.M.; Hill, R.J.; Jarvie, D.M.; Henry, M.E.: Assessing undiscovered resources of the Barnett-Paleozoic total petroleum system, Bend Arch-Fort Worth basin province, Texas. Search and Discovery Article (2003). Accessed 15 Dec 2014
27. Pollastro, R.M.; Jarvie, D.M.; Hill, R.J.; Adams, C.W.: Geologic framework of the Mississippian Barnett Shale, Barnett-Paleozoic total petroleum system, Bend Arch-Fort Worth Basin, Texas. AAPG Bull. **91**(4), 405–436 (2007). https://doi.org/10.1306/2F10300606008
28. Jarvie, D.M.; Hill, R.J.; Ruble, T.E.; Pollastro, R.M.: Unconventional shale-gas systems: the Mississippian Barnett Shale of north-central Texas as one model for thermogenic shale-gas assessment. Am. Assoc. Pet. Geol. Bull. **91**(4), 475–499 (2007)

29. Abouelresh, M.O.; Slatt, R.M.: Shale depositional processes: example from the Paleozoic Barnett Shale, Fort Worth Basin, Texas, USA. Cent. Eur. J. Geosci. **3**(4), 398–409 (2011)
30. Abouelresh, M.O.; Slatt, R.M.: Lithofacies and sequence stratigraphy of the Barnett Shale in east-central Fort Worth Basin, Texas. AAPG Bull. **96**(1), 1–22 (2012)
31. Loucks, R.G.; Ruppel, S.C.: Mississippian Barnett Shale: lithofacies and depositional setting of a deepwater shale-gas succession in the Fort Worth Basin, Texas. AAPG Bull. **91**(4), 579–601 (2007). https://doi.org/10.1306/11020606059
32. Bowker, K.A.: Barnett Shale gas production: Fort Worth Basin-issues and discussion. AAPG Bull. **91**(4), 523–533 (2007). https://doi.org/10.1306/06190606018
33. Bowker, K.A.: Recent development of the Barnett Shale play, Fort Worth Basin. West Texas Geol. Soc. Bull. **42**(6), 4–11 (2003)
34. Fu, Q.; Horvath, S.C.; Potter, E.C.; Roberts, F.; Tinker, S.W.; Ikonnikova, S.; Fisher, W.L.; Yan, J.: Log-derived thickness and porosity of the Barnett Shale, Fort Worth basin, Texas: implications for assessment of gas shale resources. AAPG Bull. **99**(1), 119–141 (2015)
35. Heslop, K.A.: Generalized method for the estimation of TOC from GR and Rt. In: AAPG Search Discov. Article. Article #80117 (2010)
36. Liu, Y.; Chen, Z.; Hu, K.; Liu, C.: Quantifying total organic carbon (TOC) from well logs using support vector regression. GeoConvention 2013 Integr. 6 (2013)
37. Luning, S.; Kolonic, S.: Uranium spectral gamma-ray response as a proxy for organic richness in black shales: applicability and limitations. J. Pet. Geol. **26**, 153–174 (2003). https://doi.org/10.1111/j.1747-5457.2003.tb00023.x
38. Jacobi, D.; Gladkikh, M.; Lecompte, B.; Hursan, G.; Mendez, F.; Longo, J.; Ong, S.; Bratovich, M.; Patton, G.; Hughes, B.; Shoemaker, P.: Integrated petrophysical evaluation of shale gas reservoirs. Paper "SPE-114925" Presented at CIPC/SPE Gas Technology Symposium 2008 Joint Conference held in Calgary, Alberta, Canada, 16–19 June 2006 (2008). https://doi.org/10.2118/114925-MS
39. Gonzalez, J.; Lewis, R.; Hemingway, J.; Grau, J.; Rylander, E.; Pirie, I.: Determination of formation organic carbon content using a new neutron-induced gamma ray spectroscopy service that directly measures carbon. In: Unconventional Resources Technology Conference (2013). https://doi.org/10.1190/urtec2013-112
40. Zhao, T.; Verma, S.; Devegowda, D.: TOC estimation in the Barnett Shale from triple combo logs using support vector machine. SEG New Orleans Annual Meeting (2015). https://doi.org/10.1190/segam2015-5922788.1