Research Article

# Utilization of adaptive neuro-fuzzy interference system and functional network in prediction of total organic carbon content

Osama Siddig[1] · Hany Gamal[1] · Pantelis Soupios[1] · Salaheldin Elkatatny[1]

© The Author(s) 2021   OPEN

## Abstract

This paper presents the application of two artificial intelligence (AI) approaches in the prediction of total organic carbon content (TOC) in Devonian Duvernay shale. To develop and test the models, around 1250 data points from three wells were used. Each point comprises TOC value with corresponding spectral and conventional well logs. The tested AI techniques are adaptive neuro-fuzzy interference system (ANFIS) and functional network (FN) which their predictions are compared to existing empirical correlations. Out of these two methods, ANFIS yielded the best outcomes with 0.98, 0.90, and 0.95 correlation coefficients (R) in training, testing, and validation respectively, and the average errors ranged between 7 and 18%. In contrast, the empirical correlations resulted in R values less than 0.85 and average errors greater than 20%. Out of eight inputs, gamma ray was found to have the most significant impact on TOC prediction. In comparison to the experimental procedures, AI-based models produces continuous TOC profiles with good prediction accuracy. The intelligent models are developed from preexisting data which saves time and costs.

## Article highlights

- In contrast to existing empirical correlation, the AI-based models yielded more accurate TOC predictions.

- Out of the two AI methods used in this article, ANFIS generated the best estimations in all datasets that have been tested.
- The reported outcomes show the reliability of the presented models to determine TOC for Devonian shale.

## Abbreviations

| | | | |
|---|---|---|---|
| TOC | Total organic carbon content | $\rho$ | Bulk density of the formation |
| AI | Artificial intelligence | RO | Ratio between the organic matter and organic carbon |
| ANFIS | Adaptive neuro-fuzzy interference system | | |
| FN | Functional network | $\rho_o$ | Density of the organic matter |
| R | Correlation coefficient | $\rho_{mi}$ | Average bulk density |
| AAPE | Average absolute percentage error | FR | Formation resistivity |
| $\rho_B$ | Matter free rock density | $\Delta t$ | Acoustic transit time |
| | | $\Delta logR$ | Logs separation |

✉ Salaheldin Elkatatny, elkatatny@kfupm.edu.sa; Osama Siddig, g201105870@kfupm.edu.sa; Hany Gamal, g201706870@kfupm.edu.sa; Pantelis Soupios, panteleimon.soupios@kfupm.edu.sa | [1]Department of Petroleum Engineering, King Fahd University of Petroleum and Minerals, Box: 5049, Dhahran 31261, Saudi Arabia.

| LOM | Level of maturity |
| GR | Gamma-ray |
| Tmax | The indicator of maturity |
| RHOB | Bulk density |
| CNP | Neutron porosity |
| SP | Spontaneous potential |
| K | Spectrum logs of potassium |
| Th | Spectrum logs of thorium |
| Ur | Spectrum logs of uranium |
| ANN | Artificial neurons network |
| FL | Fuzzy logic |
| SVM | Support vector machine |
| GPR | Gaussian process regression |
| FB | Forward–backward |
| BF | Backward-forward |
| BE | Backward elimination |
| ES | Exhaustive search |
| FS | Forward selection |
| $X_{given}$ | Measured values |
| $X_{Predicted}$ | Predicted values |

# 1 Introduction

Oil or natural gas reserves are constantly depleting as a result of continued oil and gas exploitation, and the existing reservoir production levels are substantially decreasing [1–5]. Therefore, Source rock and unconventional reservoirs have increasingly piqued interest [6–9]. Comparing to conventional reserves, unconventional reservoirs are tighter, less permeable and more complex, that making their exploration more difficult and costly [10]. Significant unconventional resource discoveries have been reported around the world in recent years, particularly in the Middle East, North and South America, and North Africa, contributing to the global oil reserves [11, 12].

Unlike the conventional reservoirs, unconventional resources are in-situ storing and generating; therefore, it is essential to quantify their potential for hydrocarbon generation. Moreover, unconventional resource characterization, development, and production are complex and expensive processes, all of which indicate the necessity of assessing their potentials precisely and cost-effectively [7, 11].

Total organic carbon (TOC) is being utilized to quantify the potentials of hydrocarbon generation, and therefore it reflects the quality of the unconventional reservoirs [13–16]. Generally, TOC is determined experimentally by the rock pyrolysis test [17, 18], and the number of tests conducted to quantify TOC is limited, because of the high experimental cost. Consequently, obtaining a comprehensive TOC assessment for the formation(s) of concern is quite challenging, which has a significant impact on reservoir evaluation [19].

Several scholars established empirical TOC correlations (summarized in Table 1) based on well logs and

**Table 1** Different correlations of TOC

| References | Model | Remarks |
|---|---|---|
| [25] | $TOC(vol.\%) = \frac{(\rho_B - \rho)}{1.378}$ <br> where the densities are in g/cm$^3$ | Used data from Devonian shale and predicts TOC in volume percentage. The data were taken from seven wells located in Virginia, West Virginia, Kentucky and Ohio. The model needs only the organic matter free rock density ($\rho_B$) and the bulk density of the formation ($\rho$) |
| [26] | $TOC(wt.\%) = \frac{[(100\rho_o) - (\rho - 0.9922\rho_{mi} - 0.039)]}{[(RO*\rho)(\rho - 1.135\rho_{mi} - 0.675)]}$ | Utilized data from 46 wells in Western Appalachian basin in the United States to present a modified model that determines the TOC in weight percentage. RO is ratio between the organic matter and organic carbon. $\rho_o$ is the density of the organic matter and $\rho_{mi}$ is the average bulk density |
| [27] | $\Delta \log R = \log_{10}\left(\frac{FR}{FR_{baseline}}\right) + 0.02 \times (\Delta t - \Delta t_{baseline})$ <br> $TOC = \Delta \log R \times 10^{(2.297 - 0.1688 \times LOM)}$ <br> where transit time is in μs/ft and resistivity is in ohm.m | Widely used and commonly known as ΔlogR model. Using the formation resistivity (FR) and acoustic transit time (Δt), logs separation (ΔlogR) has to be calculated first. Then from ΔlogR and level of maturity (LOM), TOC could be estimated. The case study included data from six wells in different locations (ex. California and Colorado) |
| [28] | $TOC = [\alpha \Delta \log R + \beta(GR - GR_{baseline})] \times 10^{(\delta + \eta T_{max})}$ <br> Where α, β, δ, and η are constants | Wang's Model is a modified ΔlogR model developed using data from Devonian shale. To estimate TOC, It uses the gamma-ray (GR), ΔlogR, and the indicator of maturity (Tmax). The model have been tested in five wells in western Canada |
| [29] | $TOC = a\Delta \log R + b(GR - GR_{baseline}) + c$ <br> where a, b and c are constants | Zhao et al. revised Wang's model and presented a model that doesn't require LOM or Tmax. The authors utilized data from five wells in three different basins in western Canada and china (Sichuan Basin and Ordos Basin) |

their corresponding laboratory-based TOC values of core plugs or drilling cuttings. The obtained models are then suggested to be used in determining TOC for other wells [20–24].

Out of the models presented, ΔlogR model is very common, and many authors presented modifications to enhance its TOC predictability [15, 30, 31]. Charsky and Herron [32] used data from various formations and wells to evaluate the reliability of Schmoker and ΔlogR models and reported significate variation from the actual TOC values.

Artificial intelligence (AI) systems are capable of producing highly accurate models and they have been applied in various sectors such as healthcare [33], mining [34], construction [35] and energy [36]. The accuracy of the empirical correlations presented in Table 1 when applied to different datasets is a major concern, that's why AI techniques were used in numerous studies to predict TOC. Table 2 highlights the various researches that applied different AI tools for TOC estimation from well-logs. These well-logs include bulk density (RHOB), formation resistivity (FR), neutron porosity (CNP), gamma-ray (GR), Spontaneous potential (SP), sonic transit time (Δt), and spectrum logs of potassium (K), thorium (Th) and uranium (Ur). Different AI techniques have been utilized, namely artificial neurons network (ANN), fuzzy logic (FL), adaptive

neuro-fuzzy interference system (ANFIS), functional network (FN), support vector machine (SVM) and Gaussian Process Regression (GPR).

TOC is a significant parameter for evaluating unconventional reservoirs. To estimate TOC, experimental investigation can be employed, but it is time-consuming, costly, and does not provide consistent information of TOC against depth. Alternatively, TOC can be determined using empirical models, however, the accuracy and the generalization associated with these correlations are major concerns. The objective of this paper is to test the effectiveness of two AI approaches in estimating the TOC in Devonian shale formations from logging data. These well-logs consist of bulk density, resistivity, sonic transient time, bulk gamma-ray, spectral GR logs of Th, Ur, and K and neutron log porosity.

The following section presents the methodology used to develop the AI models, including the description of the datasets, data preprocessing, description of the utilized AI methods, tests procedure and the different stages in the model development. This section is followed by the results and discussion, which reports the outcomes of different methods on different datasets and presents a comparison with existing correlations and shows the results of sensitivity analysis of different sets of inputs.

**Table 2** Different researches that utilize AI techniques in TOC prediction

| References | Data source/s | AI tool/s | Inputs | Data points | $R^2$ |
|---|---|---|---|---|---|
| [37] | Kangan-Dalan Formation | FL, ANN | GR, CNP, FR, RHOB, Δt | 124 | 0.85 |
| [38] | Gadvan Formation | ANN, FL | CNP, FR, RHOB, Δt | 2875 | 0.78–0.99 |
| [39] | Kazhdomi and Kangan-Dalan Formations | ANN | GR, FR, Δt, K, Th | 78 | 0.89 |
| [40] | Kazhdomi and Kangan-Dalan Formations | ANN | GR, CNP, FR, Δt, K, Th | 200 | NA |
| [41] | Jiumenchong formation | SVM | GR, CNP, FR, RHOB, Δt, K, Ur, Th | 31 | 0.69 |
| [42] | Sichuan Basin | ANN | GR, CNP, FR, RHOB, Δt, K, Ur, Th | 185 | 0.87–0.91 |
| [13] | Barnett and Devonian shale formations | ANN | GR, Δt, FR, RHOB<br>GR, FR, RHOB, Δt | 442 | 0.89- 0.93 |
| [43] | Ordos basin and Canning basin | GPR | GR, CNP, FR, RHOB, Δt, K, Ur, Th | NA | NA |
| [44] | Barnett shale | ANN | GR, FR, RHOB, Δt | 442 | 0.93 |
| [19] | Kazhdumi formation | FL | GR, CNP, FR, RHOB, Δt, K, Ur, Th | 31 | 0.94 |
| [45] | NA | ANN | FR, Δt | 70 | 0.98 |
| [46] | Tonghua Basin | ANN, SVM | GR, CNP, FR, RHOB, Δt, SP, K, Ur, Th | 215 | 0.9–0.93 |
| [47] | Ras Qattara and Khatatba formations | ANN | GR, CNP, FR, RHOB, Δt | 54 | 0.96 |
| [48] | Beibu Gulf basin | SVM | GR, FR, RHOB, Δt, SP | 18 | 0.75 |
| [14] | Shahejie Formation | ANN | GR, CNP, FR, RHOB, Δt | 125 | 0.83 |
| [23] | Barnett shale | FL | GR, FR, RHOB, Δt | 645 | 0.91 |
| [49] | Barnett shale | ANFIS, FN, SVM | GR, FR, RHOB, Δt | +800 | 0.82–0.87 |
| [50] | Duvernay and Barnett shales | ANN | GR, FR, RHOB, Δt | 460 | 0.98 |
| [22] | Devonian and Barnett shales | FN, SVM | GR, FR, RHOB, Δt | +500 | 0.74–0.77 |

## 2 Methodology

In this study, two AI techniques, the adaptive neuro-fuzzy interference system (ANFIS) and the functional network (FN) were used to estimate the TOC from eight well-logs information.

### 2.1 Data description

Three different wells' TOC experimental data, as well as their corresponding well logs, were obtained. The AI models were trained, tested, and validated using 891, 291, and 82 data points from Well-I, Well-II, and Well-III, in order. All of these wells are in source rock rich in organic liquids known as Devonian Duvernay shale. This basin is located in Canada, Alberta (Fig. 1), and has oil reserves above 60 billion barrels and gas reserves above 400 trillion cubic feet [51, 52]. Table 3 shows the statistical properties of the Well-I dataset.
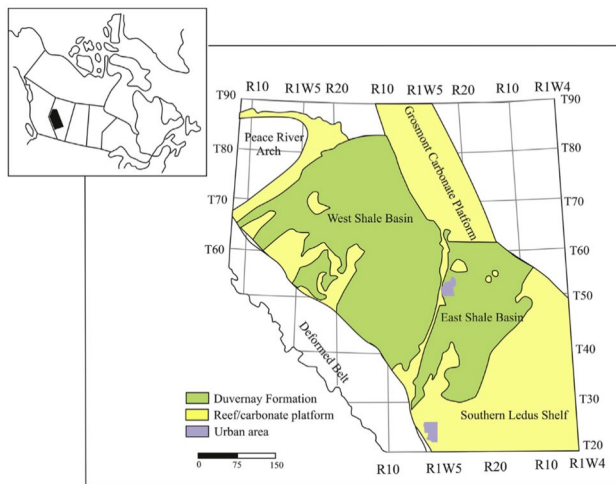


**Fig. 1** Duvernay shale distribution in central Alberta [53]

### 2.2 Well logs

In well's logging, the in-situ properties of rocks around the wellbore are indirectly estimated from electric, acoustic and nuclear indicators. The interpretations of these indicators reflect the existence of hydrocarbon, petrophysical properties, and the lithology of the formation [54, 55]. In this study the following well logs records were used:

*Formation resistivity (FR)* is a measure of electrical resistivity in three increasing depths from the well which is mainly interpreted to know the fluids' saturations and hence the existence of hydrocarbon [56, 57].

*Sonic log* is a measure of the time required for a sound wave to travel for a predetermined distance, which depends on the matrix elasticity and porosity [58], therefore, it is used in the identification of lithology, fractures and porosity.

*Density log* record the bulk density around the well, this density measure covers the matrix and the pores filled with fluid, which can be used to quantify the porosity fraction.

*Neutron log* is a log rely on a neutron source to measure the hydrogen index and consequently the porosity of the formation.

*Gamma-ray log* measures the natural gamma radiations and thus is used to distinguish shales from other sedimentary rocks.

*Spectral gamma-ray log* is a sophisticated measure for gamma-ray that uses the energy of gamma rays and identifies the elements that emitted them.

### 2.3 Samples testing

To determine the TOC of drilling cuttings from several wells, Rock–Eval 6 was used. Experimental procedures -followed what was presented by Chen et al. [59]- are shown in Fig. 2.

### 2.4 Data preprocessing

In the first step, outliers, unrealistic and incomplete data points were removed from the datasets that were used
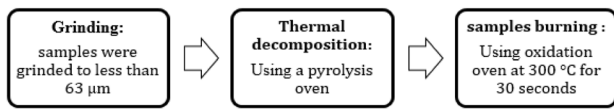
**Table 3** Statistics of Well-I data

| Parameter | FR (Ω.m) | Δt (μsec/ft) | RHOB (g/cm$^3$) | CNP | GR (°API) | Ur (wt%) | Th (ppm) | K (ppm) | TOC (wt%) |
|---|---|---|---|---|---|---|---|---|---|
| CC with TOC | 0.35 | 0.20 | − 0.35 | 0.08 | 0.62 | 0.27 | 0.22 | 0.56 | 1.00 |
| 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 |

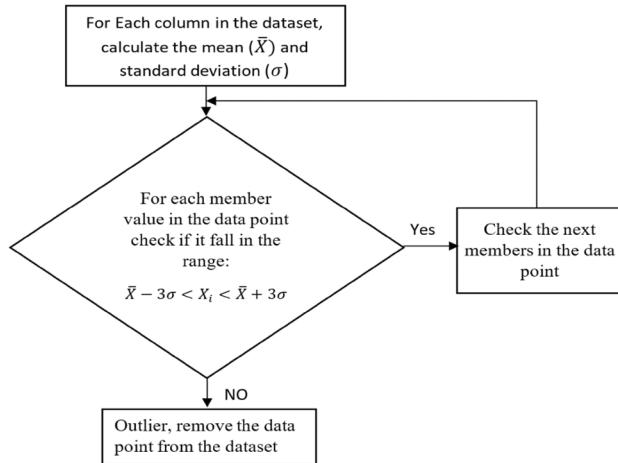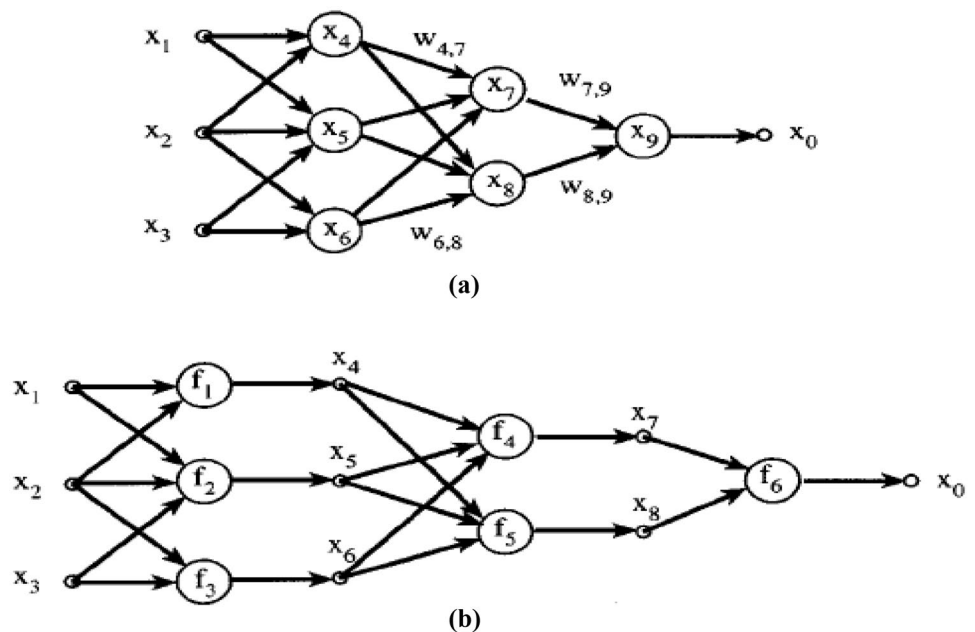**Fig. 2** TOC test procedures



**Fig. 3** Outliers' detection

to build the models. Using a built-in function in Matlab, any data points with a member that has a value that is far from the average with at least triple the standard deviation were designated as outliers. Figure 3 shows the criteria for detecting outliers.

## 2.5 AI methods

In this work, two AI approaches were used, Adaptive neuro-fuzzy inference system (ANFIS) and functional network (FN). Functional Network was introduced in the 1990s as an alternative to ANN [60, 61]. FN uses both domain information and data knowledge, it uses adaptable generalized functional models that change with the learning process [62]. The Functional network contains different elements, such as an input layer, an output layer and a set of intermediate layers, layers of neurons and directed links. Castillo et al. (2000) [63] summarized the difference between ANN and FN as following: in FN, the neuronal functions are multi-argument and can be arbitrary. Figure 4 shows the difference between ANN and FN structures.

Several successful applications of FN related to the oil industry were reported in the literature [64–66]. There are various feature selection techniques associated with the functional network such as [67, 68]:

*Forward selection* starts with minimum variables and adding features.

*Backward elimination* starts with all features and then reduces them.

*Exhaustive search* examine every point which significantly increases the computational time.

Adaptive neuro-fuzzy inference system (ANFIS) was developed in the 1990s and integrate the principles of neural networks and fuzzy logic (FL) [69, 70]. In this method, ANN is used to set the fuzzy rules in FL [71]. This integration of the two methods provides an improved performance [72]. ANFIS has various reported applications

**Fig. 4 a** ANN structure against **b** FN structure [63]



(a)



(b)

in the oil industry [64, 73]. ANFIS structure combines the fuzzy inference system and a neural network as presented in Fig. 5.

## 2.6 Models' development

Well-I's dataset which contains 891 data points with wide ranges of values as displayed in Fig. 6, was used in models training and optimization. The impacts of various parameters inside the algorithms were examined to optimize the models.

To determine the best models, different runs were performed in each approach. This was accomplished by executing the AI runs inside multiple for-loops in MATLAB for each machine learning approach. In FN, five methods were used: forward–backward (FB), backward-forward (BF), backward elimination (BE), exhaustive search (ES), and forward selection (FS). Three types were used with each of these methods, one linear and two non-linear. While in ANFIS, epochs size and cluster radiuses were tested.

The correlation coefficient (R) and the average absolute percentage error (AAPE), were utilized as evaluation criteria for the developed models using Eqs. (1) and (2) respectively:

$$R = \frac{\left[ N \sum_{i=1}^{N} \left( X_{giveni} \times X_{Predictedi} \right) \right] - \left[ \sum_{i=1}^{N} X_{giveni} \times \sum_{i=1}^{N} X_{Predictedi} \right]}{\sqrt{\left[ N \sum_{i=1}^{N} \left( X_{giveni} \right)^2 - \left( \sum_{i=1}^{N} X_{giveni} \right)^2 \right] \left[ N \sum_{i=1}^{N} (X)^2 - \left( \sum_{i=1}^{N} X_{Predictedi} \right)^2 \right]}} \tag{1}$$

$$AAPE = \frac{\sum_{i=1}^{N} \frac{X_{giveni} - X_{Predictedi}}{X_{giveni}} \times 100\%}{N} \tag{2}$$

where N is the size of dataset, $X_{given}$ and $X_{Predicted}$ are the measured and the AI-predicted TOC values respectively.
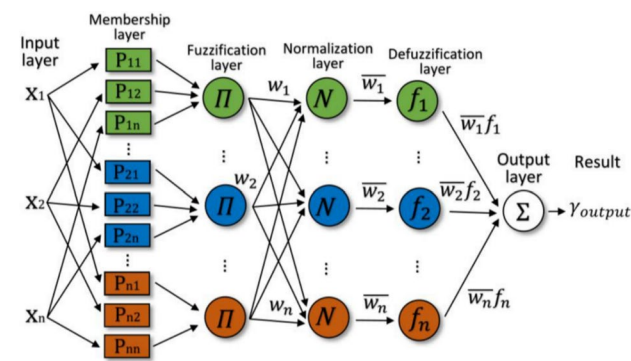


**Fig. 5** ANFIS structure [74]

With the datasets from Well-II (291 data points) and Well-III (82 data points), the generalization of the produced models was internally and externally tested. Similar to Well-I, these two wells are located in the same field. The Schmoker, Zhao et al., and logR models were used to compare the performance of AI-based models.

## 3 Results and discussion

Eight conventional and spectral well logs data were used to train AI models for TOC predictions. The training dataset contained 891 Well-I data points, while the testing dataset contained 291 Well-III data points. The outcomes of each technique are presented in this section.

Different methods of FN have been applied, and the best model was obtained when the Forward–backward method and non-linear type were used. The model yielded R values of 0.902 and 0.879 in training and testing respectively, with AAPE values ranging between 18.9% and 24.4%, as shown in Fig. 7. In Fig. 7, it is obvious that many points were far from the 45° line.

Using ANFIS, various epochs size and cluster radiuses were tested. The estimation yielded from this method was significantly better than FN. The correlation coefficients ranged between 0.899 and 0.983, while the AAPE values were 7.3 and 17.9 in training and testing respectively as shown in Fig. 8. This model has been achieved with 0.25 cluster radios and 100 iterations.

Several runs have been made in each method to achieve the reported results, each run test different parameters/methods inside the algorithms. Multiple for loops in Matlab have been used to test a wide range of possible combinations of algorithms' parameters while reporting the R and AAPE. Figures 9 and 10 present the results of different iterations for the ANFIS and FN respectively. They also indicate the chosen iterations as the best results achieved based on the correlation coefficient and AAPE. The optimized parameters for each method are reported in Table 4.

The two techniques produced a relatively good level of accuracy in predicting Devonian shale's TOC during training and testing, as indicated in the previous results. Dataset from Well-III has been hidden away from the AI tools throughout models' development stages as an additional check to guarantee that the new models
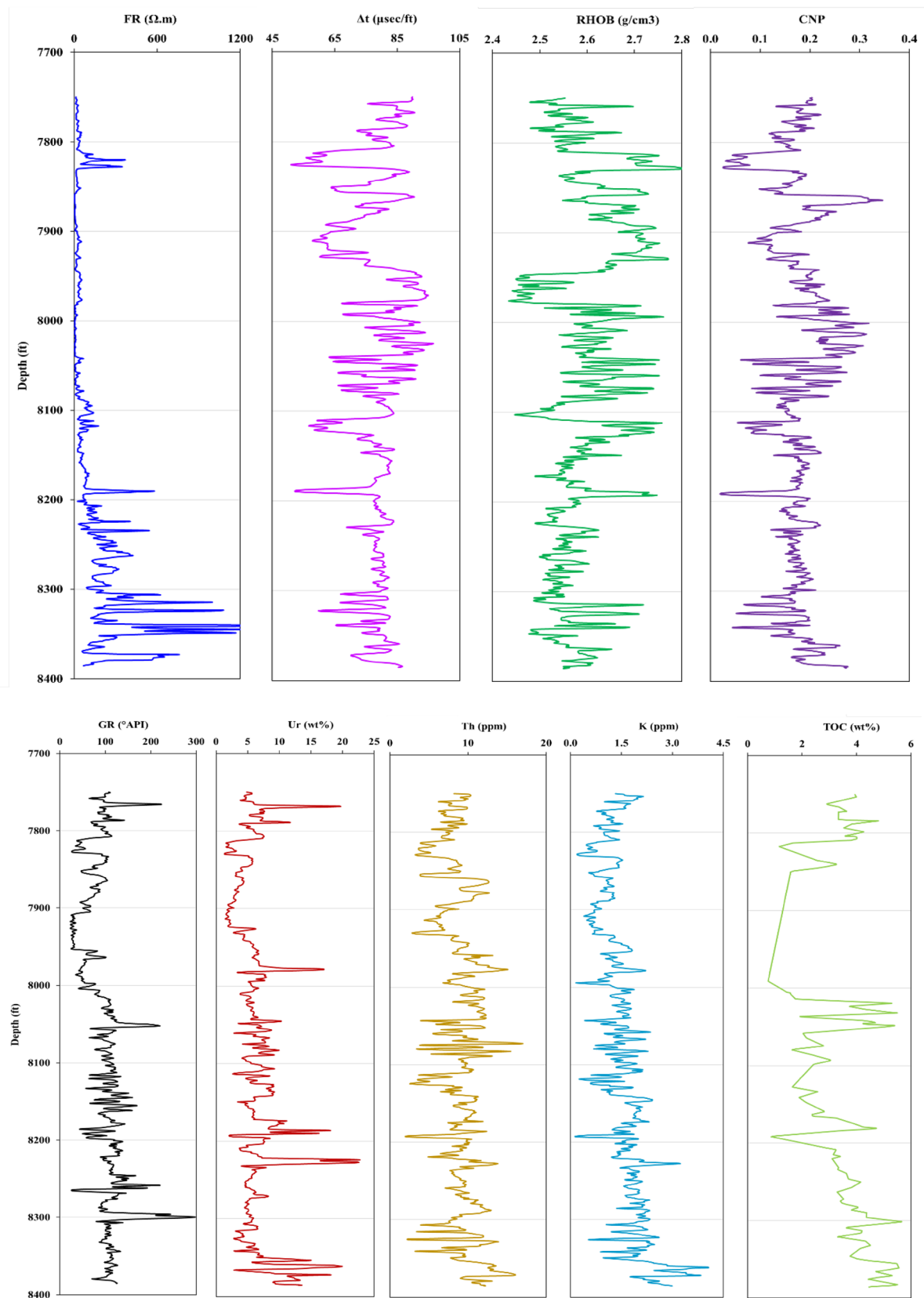
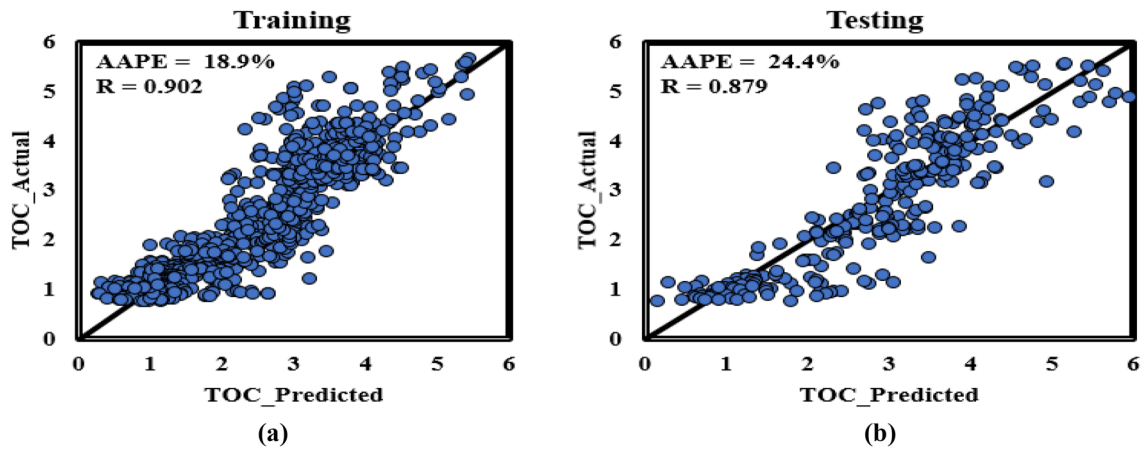**Fig. 6** The well-logs and TOC used to train the models

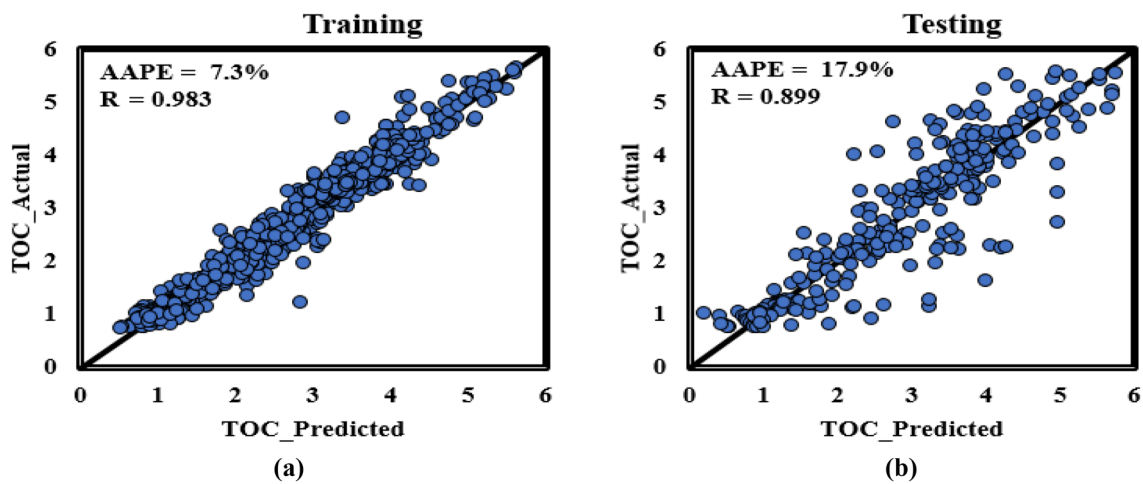**Fig. 7** Cross-plots of the FN-predicted and measured TOC for **a** the training and **b** the testing datasets



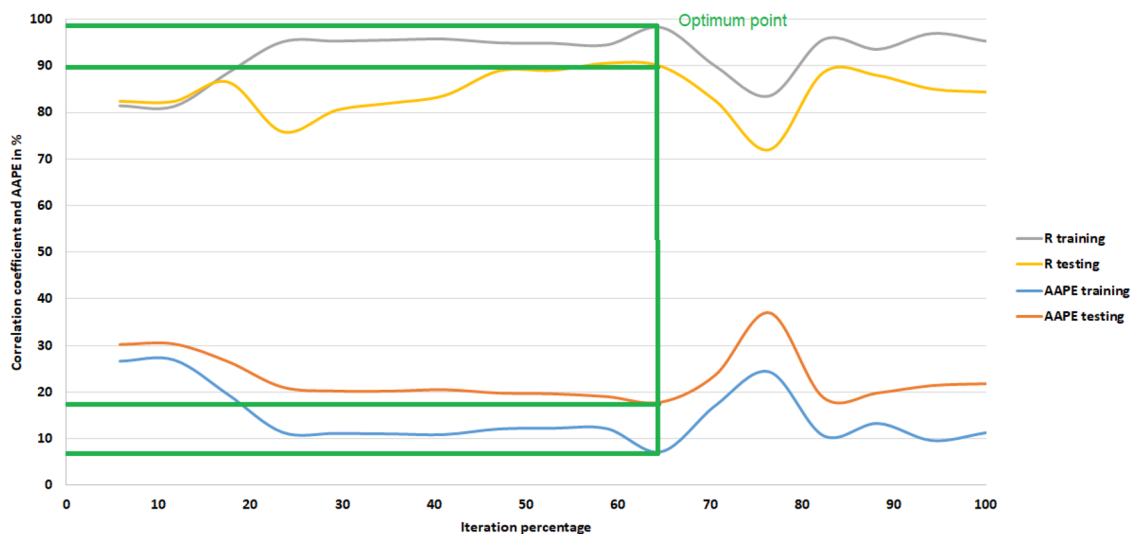**Fig. 8** Cross-plots of the ANFIS-predicted and measured TOC for **a** the training and **b** the testing datasets
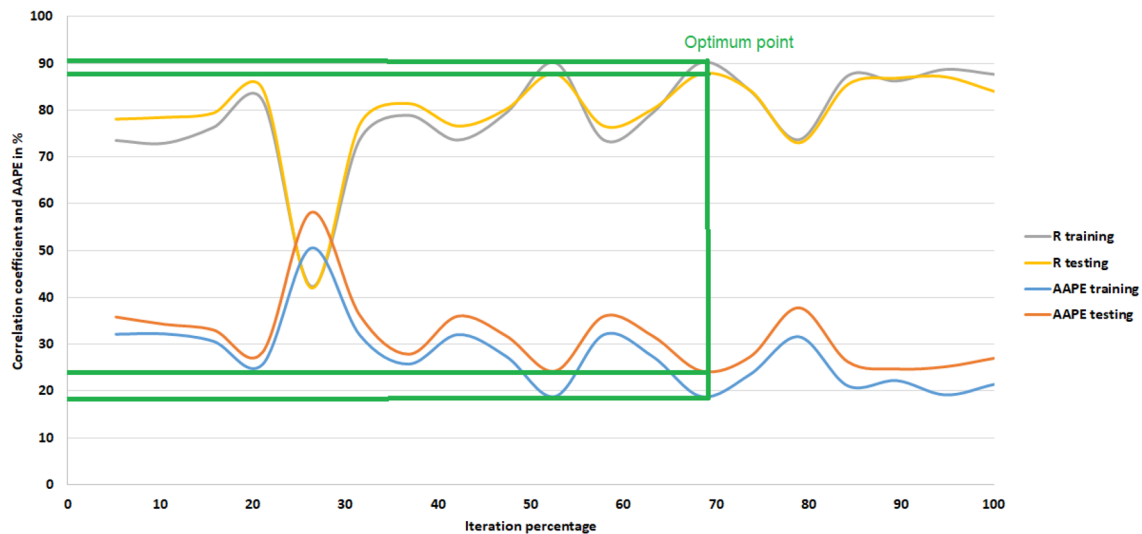


**Fig. 9** Results of different iteration in ANFIS

**Fig. 10** Results of different iteration in FN

**Table 4** Optimized parameters of the AI algorithms

| ANFIS optimized parameters | |
| --- | --- |
| Clustering type | Subtractive Clustering |
| Fuzzy interference system | Sugeno |
| Input membership function | Gaussian: $f = e^{\frac{-(x-c)^2}{2\sigma^2}}$ Where c is the mean and σ is the standard deviation |
| Output membership function | Linear |
| Cluster radius | 0.25 |
| Epochs size | 100 |
| FN optimized parameters | |
| Method | Forward–backward (FB) |
| Type | Nonlinear |

were generalizable. The outcomes of the two AI algorithms were contrasted to the available correlations, namely Schmoker, Zhao et al., and ΔlogR models.

As shown in Fig. 11 that the ANFIS-based predictions were the most accurate with an AAPE of only 10.3% and a high R of 0.95. FN model was less accurate than ANFIS; however, both were better than the other three correlation with R equals 0.87 and AAPE values range between 16 and 20%. Zhao et al. correlation results were close to the FN model, and slightly better than ΔlogR method, with an AAPE value around 20% and R of 0.84. While Schmoker's model yielded the least accurate results. Standing from this comparison, the AI-based models, especially the ANFIS, outperformed the different existing correlations.

## 4 Sensitivity of the inputs

To examine the significance of each input parameter in TOC prediction, various sets of inputs were assessed. Seven sets were tested, the simplest one contains only the conventional well logs except of the GR, and the most comprehensive consist of all available well logs, as shown in Table 5.

The outcomes of different input sets are reported in Table 5. The best performance was achieved in Set 1 where all eight parameters were included, however, it has a significantly higher computational cost. Set 6 and Set 4 performances were followed, those sets excluded the sonic transient time and porosity respectively. This confirms the low correlation coefficient of
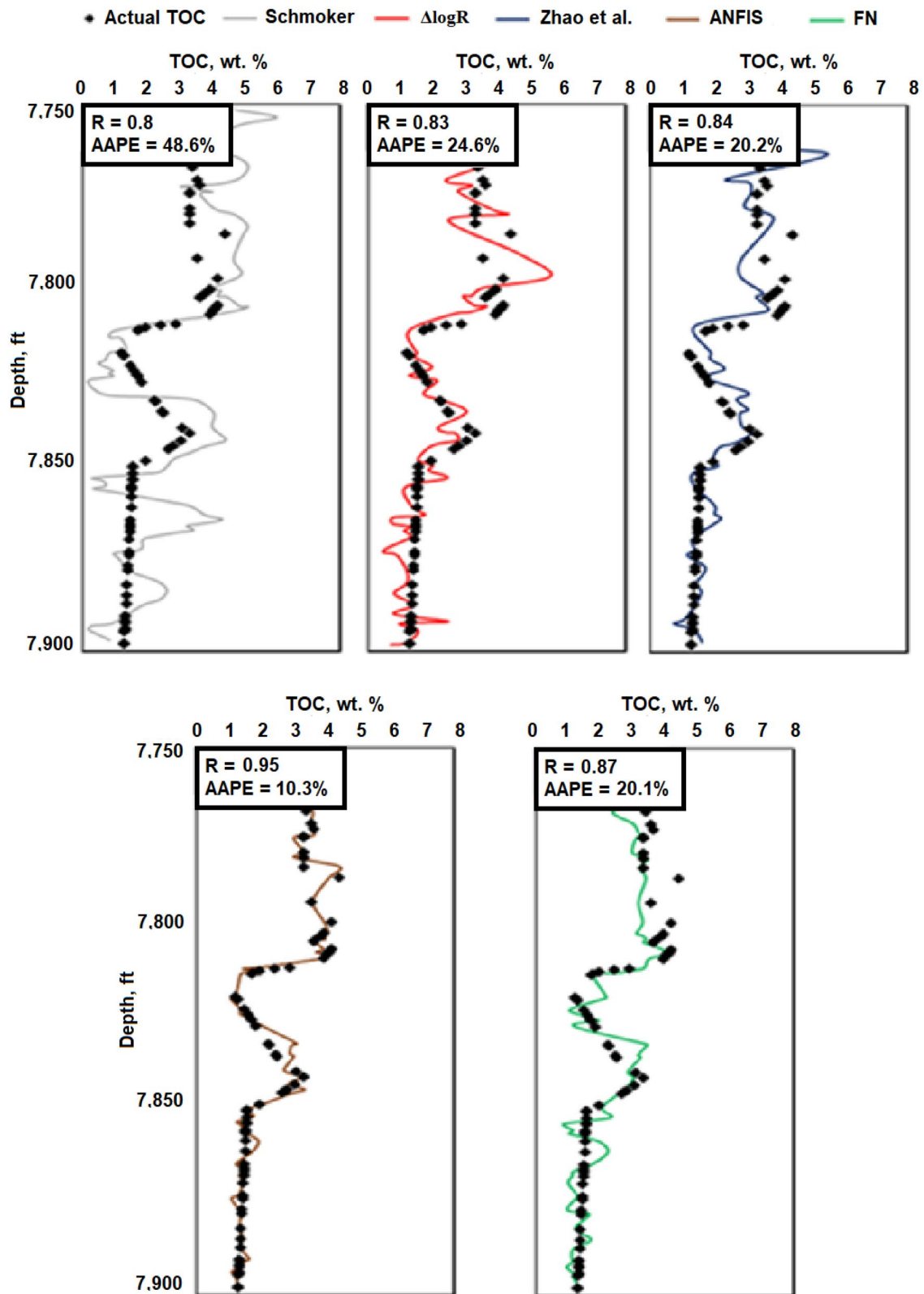
**Fig. 11**  AI-based against empirical correlations TOC prediction of Well-III

**Table 5** Studied input sets

| No | Contain | Inputs |
|----|---------|--------|
| 1 | All logs | GR, CNP, FR, RHOB, Δt, K, Ur, Th |
| 2 | Excluding spectral GR | GR, CNP, FR, RHOB, Δt |
| 3 | Excluding GR and spectral GR | CNP, FR, RHOB, Δt |
| 4 | Excluding porosity | GR, FR, RHOB, Δt, K, Ur, Th |
| 5 | Excluding density | GR, CNP, FR, Δt, K, Ur, Th |
| 6 | Excluding sonic transient time | GR, CNP, FR, RHOB, K, Ur, Th |
| 7 | Excluding resistivity | GR, CNP, RHOB, Δt, K, Ur, Th |

these two parameters with TOC as shown in Table 3. The least favorable case was Set 3 that contains only four variables. According to the results shown in Table 6, The GR and spectral GR have a great impact on the prediction of TOC, while the porosity and sonic transient time showed the least effect. It's noticeable that for all cases, ANFIS outperformed FN, however, the former has slightly higher computational time.

## 5 Models' limitations

The data in this study were acquired from wells in the same area, and the outcomes were restricted to Devonian shale. Therefore, the models' performance cannot be assured if they were utilized in another type of formation or with data ranges that are different from those used in this research.

## 6 Conclusions

Using artificial intelligence techniques and over 1250 data points, two models for TOC prediction from eight well logging information were established in this study. Two artificial intelligence techniques were employed, adaptive neuro-fuzzy interference system (ANFIS) and Functional network (FN). The following is a summary of the outcomes presented in this paper:

- Out of the two AI algorithms utilized in this study, ANFIS yielded the best match in training and testing datasets with correlation coefficients of 0.98 and 0.90 and AAPE values of 7% and 18% for training and testing respectively.
- Data from a different well was hidden entirely from the AI tools and used to verify the built models, the ANFIS model successfully predicted TOC with a 0.95 correlation coefficient and 10% AAPE.
- The presented models were contrasted to three various empirical correlations. The empirical correlations yielded less favorable results with correlation coefficients under 0.85 and AAPE above 20%.

The presented models accurately predicted TOC from well-logs such as CNP, RHOB, GR, t, FR, K, Th, and Ur logs, allowing for continuous TOC profiles with depth without the requirement for core analysis or extra well interventions. To achieve reliable results, we propose using developed models with input parameters within

**Table 6** Outcomes of different sets

| Inputs sets no | R | | | AAPE (%) | | | RMSE | | | CPU time (sec.) |
|---|---|---|---|---|---|---|---|---|---|---|
| | Train. | Test. | Valid | Train. | Test. | Valid. | Train. | Test. | Valid. | |
| | *ANFIS* | | | | | | | | | |
| 1 | 0.98 | 0.90 | 0.95 | 7.3 | 18.0 | 10.3 | 0.24 | 0.61 | 0.33 | 57.4 |
| 2 | 0.90 | 0.82 | 0.87 | 17.2 | 23.8 | 16.9 | 0.57 | 0.80 | 0.55 | 8.8 |
| 3 | 0.84 | 0.72 | 0.88 | 24.4 | 37.1 | 16.9 | 0.72 | 0.98 | 0.52 | 8.3 |
| 4 | 0.96 | 0.89 | 0.91 | 10.7 | 18.9 | 11.5 | 0.38 | 0.65 | 0.47 | 38.7 |
| 5 | 0.94 | 0.88 | 0.92 | 13.4 | 19.9 | 12.9 | 0.46 | 0.67 | 0.42 | 25.9 |
| 6 | 0.97 | 0.85 | 0.91 | 9.7 | 21.5 | 13.5 | 0.32 | 0.75 | 0.44 | 40.7 |
| 7 | 0.95 | 0.84 | 0.96 | 11.4 | 21.9 | 9.7 | 0.39 | 0.77 | 0.33 | 27.5 |
| | *FN* | | | | | | | | | |
| 1 | 0.90 | 0.88 | 0.87 | 18.9 | 24.4 | 20.1 | 0.56 | 0.67 | 0.51 | 50.4 |
| 2 | 0.84 | 0.84 | 0.88 | 23.7 | 27.4 | 17.8 | 0.70 | 0.74 | 0.54 | 5.2 |
| 3 | 0.74 | 0.73 | 0.86 | 31.7 | 37.9 | 19.2 | 0.88 | 0.95 | 0.60 | 5.0 |
| 4 | 0.87 | 0.86 | 0.94 | 21.2 | 26.4 | 13.2 | 0.63 | 0.72 | 0.40 | 39.1 |
| 5 | 0.86 | 0.87 | 0.89 | 22.3 | 24.8 | 15.9 | 0.66 | 0.69 | 0.51 | 17.5 |
| 6 | 0.89 | 0.87 | 0.87 | 19.3 | 25.3 | 19.6 | 0.60 | 0.68 | 0.53 | 16.5 |
| 7 | 0.88 | 0.84 | 0.93 | 21.5 | 27.1 | 15.0 | 0.63 | 0.75 | 0.40 | 16.8 |

the same model's ranges. More data will be collected in future work to further validate the established models, and more machine learning approaches will be used.

## Declarations

## References

1. Tang H, Sun Z, He Y, Chai Z, Hasan AR, Killough J (2019) Investigating the pressure characteristics and production performance of liquid-loaded horizontal wells in unconventional gas reservoirs. J Pet Sci Eng 176:456–465. https://doi.org/10.1016/j.petrol.2019.01.072

2. Zhao P, Ostadhassan M, Shen B, Liu W, Abarghani A, Liu K, Luo M, Cai J (2019) Estimating thermal maturity of organic-rich shale from well logs: Case studies of two shale plays. Fuel 235:1195–1206. https://doi.org/10.1016/j.fuel.2018.08.037

3. Ng CSW, Ghahfarokhi AJ, Amar MN, Torsæter O (2021) Smart proxy modeling of a fractured reservoir model for production optimization: implementation of metaheuristic algorithm and probabilistic application. Nat Resour Res 30:2431–2462. https://doi.org/10.1007/s11053-021-09844-2

4. Amar MN, Ghahfarokhi AJ, Ng CSW, Zeraibi N (2021) Optimization of WAG in real geological field using rigorous soft computing techniques and nature-inspired algorithms. J Pet Sci Eng 206:109038. https://doi.org/10.1016/j.petrol.2021.109038

5. Amar MN, Zeraibi N, Jahanbani Ghahfarokhi A (2020) Applying hybrid support vector regression and genetic algorithm to water alternating CO 2 gas EOR. Greenh Gases Sci Technol 10:613–630. https://doi.org/10.1002/ghg.1982

6. Wu Y, Tahmasebi P, Yu H, Lin C, Wu H, Dong C (2020) Pore-scale 3D dynamic modeling and characterization of shale samples: considering the effects of thermal maturation. J Geophys Res Solid Earth. https://doi.org/10.1029/2019JB018309

7. Zhu L, Zhang C, Zhang C, Zhang Z, Zhou X, Liu W, Zhu B (2020) 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 188:106944. https://doi.org/10.1016/j.petrol.2020.106944

8. Nait Amar M, Larestani A, Lv Q, Zhou T, Hemmati-Sarapardeh A (2021) Modeling of methane adsorption capacity in shale gas formations using white-box supervised machine learning techniques. J Pet Sci Eng. https://doi.org/10.1016/j.petrol.2021.109226

9. Nait Amar M (2020) Modeling solubility of sulfur in pure hydrogen sulfide and sour gas mixtures using rigorous machine learning methods. Int J Hydrogen Energy 45:33274–33287. https://doi.org/10.1016/j.ijhydene.2020.09.145

10. Zou CN, Tao SZ, Bai B, Yang Z (2015) Differences and relations between unconventional and conventional oil and gas. China Pet Explor 20:1–16

11. Kumar S, Das S, Bastia R, Ojha K (2018) Mineralogical and morphological characterization of Older Cambay Shale from North Cambay Basin, India: Implication for shale oil/gas development. Mar Pet Geol 97:339–354. https://doi.org/10.1016/j.marpetgeo.2018.07.020

12. Rani S, Padmanabhan E, Prusty BK (2019) Review of gas adsorption in shales for enhanced methane recovery and CO2 storage. J Pet Sci Eng 175:634–643. https://doi.org/10.1016/j.petrol.2018.12.081

13. Mahmoud AA, Elkatatny S, Mahmoud M, Abouelresh M, Abdulraheem A, Ali A (2017) Determination of the total organic carbon (TOC) based on conventional well logs using artificial neural network. Int J Coal Geol 179:72–80. https://doi.org/10.1016/j.coal.2017.05.012

14. Wang H, Wu W, Chen T, Dong X, Wang G (2019) An improved neural network for TOC, S1 and S2 estimation based on conventional well logs. J Pet Sci Eng 176:664–678. https://doi.org/10.1016/j.petrol.2019.01.096

15. Yang S-C, Wang N, Li M-R, Yu J (2013) The logging evaluation of source rocks of triassic Yanchang formation in Chongxin area, Ordos Basin. Nat Gas Geosci 24:470–476

16. Ma L, Taylor KG, Dowey PJ, Courtois L, Gholinia A, Lee PD (2017) 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. Int J Coal Geol 180:100–112. https://doi.org/10.1016/j.coal.2017.08.002

17. Carvajal-Ortiz H, Gentzis T (2015) Critical considerations when assessing hydrocarbon plays using rock-eval pyrolysis and organic petrology data: data quality revisited. Int J Coal Geol 152:113–122. https://doi.org/10.1016/j.coal.2015.06.001

18. Hazra B, Dutta S, Kumar S (2017) TOC calculation of organic matter rich sediments using Rock-Eval pyrolysis: critical consideration and insights. Int J Coal Geol 169:106–115. https://doi.org/10.1016/j.coal.2016.11.012

19. Bolandi V, Kadkhodaie A, Farzi R (2017) 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. J Pet Sci Eng 151:224–234. https://doi.org/10.1016/j.petrol.2017.01.003

20. Chen Y, Jiang S, Zhang D, Liu C (2017) An adsorbed gas estimation model for shale gas reservoirs via statistical learning. Appl Energy 197:327–341. https://doi.org/10.1016/j.apenergy.2017.04.029

21. Daigle H, Hayman NW, Kelly ED, Milliken KL, Jiang H (2017) Fracture capture of organic pores in shales. Geophys Res Lett 44:2167–2176. https://doi.org/10.1002/2016GL072165

22. Mahmoud AA, Elkatatny S, Ali A, Abdulraheem A, Abouelresh M (2020) Estimation of the total organic carbon using functional neural networks and support vector machine. In: Proceedings of the Day 3 Wed, January 15, 2020, IPTC

23. Mahmoud AA, Elkatatny S, Ali A, Abouelresh M, Abdulraheem A (2019) New robust model to evaluate the total organic carbon using fuzzy logic. In: Proceedings of the Day 4 Wed, October 16, 2019, SPE

24. Mathia EJ, Rexer TFT, Thomas KM, Bowen L, Aplin AC (2019) Influence of clay, calcareous microfossils, and organic matter on the nature and diagenetic evolution of pore systems in mudstones. J Geophys Res Solid Earth 124:149–174. https://doi.org/10.1029/2018JB015941

25. Schmoker JW (1979) Determination of organic content of appalachian devonian shales from formation-density logs: GEOLOGIC NOTES. Am Assoc Pet Geol Bull. https://doi.org/10.1306/2F9185D1-16CE-11D7-8645000102C1865D

26. Schmoker JW (1980) Organic content of Devonian shale in western Appalachian basin. Am Assoc Pet Geol Bull 64:2156–2165

27. Passey QR, Creaney S, Kulla JB, Morett FJ, Stroud JD (1990) A practical model for organic richness from porosity and resistivity logs. Am Assoc Pet Geol Bull 74:1777–1794. https://doi.org/10.1306/0C9B25C9-1710-11D7-8645000102C1865D

28. Wang P, Chen Z, Pang X, Hu K, Sun M, Chen X (2016) Revised models for determining TOC in shale play: example from Devonian Duvernay Shale, Western Canada Sedimentary Basin. Mar Pet Geol 70:304–319. https://doi.org/10.1016/j.marpetgeo.2015.11.023

29. Zhao P, Ma H, Rasouli V, Liu W, Cai J, Huang Z (2017) An improved model for estimating the TOC in shale formations. Mar Pet Geol 83:174–183. https://doi.org/10.1016/j.marpetgeo.2017.03.018

30. Passey QR, Bohacs KM, Esch WL, Klimentidis R, Sinha S (2010) From oil-prone source rock to gas-producing shale reservoir – geologic and petrophysical characterization of unconventional shale-gas reservoirs. In: Proceedings of the All Days, SPE

31. Wang J, Gu D, Guo W, Zhang H, Yang D (2019) Determination of total organic carbon content in shale formations with regression analysis. J Energy Resour Technol. https://doi.org/10.1115/1.4040755

32. Charsky A, Herron S (2013) Accurate, direct total organic carbon (TOC) log from a new advanced geochemical spectroscopy tool: comparison with conventional approaches for TOC estimation. In: Proceeding of the AAPG annual convention and exhibition, Pittsburg, Pennsylvania, 19–22 May

33. Davenport T, Kalakota R (2019) The potential for artificial intelligence in healthcare. Futur Healthc J 6:94–98. https://doi.org/10.7861/futurehosp.6-2-94

34. Jung D, Choi Y (2021) Systematic review of machine learning applications in mining: exploration, exploitation, and reclamation. Minerals 11:148. https://doi.org/10.3390/min11020148

35. Sacks R, Girolami M, Brilakis I (2020) Building information modelling, artificial intelligence and construction tech. Dev Built Environ 4:100011. https://doi.org/10.1016/j.dibe.2020.100011

36. Ahmad T, Zhang D, Huang C, Zhang H, Dai N, Song Y, Chen H (2021) Artificial intelligence in sustainable energy industry: Status Quo, challenges and opportunities. J Clean Prod 289:125834. https://doi.org/10.1016/j.jclepro.2021.125834

37. Kadkhodaie-Ilkhchi A, Rahimpour-Bonab H, Rezaee M (2009) 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, Iran. Comput Geosci 35:459–474. https://doi.org/10.1016/j.cageo.2007.12.007

38. Khoshnoodkia M, Mohseni H, Rahmani O, Mohammadi A (2011) TOC determination of Gadvan Formation in South Pars Gas field, using artificial intelligent systems and geochemical data. J Pet Sci Eng 78:119–130. https://doi.org/10.1016/j.petrol.2011.05.010

39. Alizadeh B, Najjari S, Kadkhodaie-ilkhchi A (2012) 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. Comput Geosci 45:261–269. https://doi.org/10.1016/j.cageo.2011.11.024

40. Sfidari E, Kadkhodaie-Ilkhchi A, Najjari S (2012) Comparison of intelligent and statistical clustering approaches to predicting total organic carbon using intelligent systems. J Pet Sci Eng 86–87:190–205. https://doi.org/10.1016/j.petrol.2012.03.024

41. Tan M, Song X, Yang X, Wu Q (2015) 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 26:792–802. https://doi.org/10.1016/j.jngse.2015.07.008

42. Shi X, Wang J, Liu G, Yang L, Ge X, Jiang S (2016) 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 33:687–702. https://doi.org/10.1016/j.jngse.2016.05.060

43. Yu H, Rezaee R, Wang Z, Han T, Zhang Y, Arif M, Johnson L (2017) A new method for TOC estimation in tight shale gas reservoirs. Int J Coal Geol 179:269–277. https://doi.org/10.1016/j.coal.2017.06.011

44. Mahmoud AA, ElKatatny S, Abdulraheem A, Mahmoud M, Omar Ibrahim M, Ali A (2017) 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, April 26, 2017, SPE

45. Alizadeh B, Maroufi K, Heidarifard MH (2018) Estimating source rock parameters using wireline data: an example from Dezful Embayment, South West of Iran. J Pet Sci Eng 167:857–868. https://doi.org/10.1016/j.petrol.2017.12.021

46. Wang P, Peng S, He T (2018) 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 55:1–15. https://doi.org/10.1016/j.jngse.2018.03.029

47. Shalaby MR, Jumat N, Lai D, Malik O (2019) 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. J Pet Sci Eng 176:369–380. https://doi.org/10.1016/j.petrol.2019.01.055

48. Rui J, Zhang H, Zhang D, Han F, Guo Q (2019) Total organic carbon content prediction based on support-vector-regression machine with particle swarm optimization. J Pet Sci Eng 180:699–706. https://doi.org/10.1016/j.petrol.2019.06.014

49. Mahmoud AA, Elkatatny S, Ali AZ, Abouelresh M, Abdulraheem A (2019) Evaluation of the total organic carbon (TOC) using different artificial intelligence techniques. Sustainability 11:5643. https://doi.org/10.3390/su11205643

50. Elkatatny S (2019) A self-adaptive artificial neural network technique to predict total organic carbon (TOC) Based on well logs. Arab J Sci Eng 44:6127–6137. https://doi.org/10.1007/s13369-018-3672-6

51. Creaney S, Allan J, Cole KS, Fowler MG, Brooks PW, Osadetz K, Macqueen RW, Snowdon L, Riediger CL (1994) Petroleum generation and migration in the western Canada Sedimentary Basin. In: Geological Atlas of the Western Canada Sedimentary Basin, Canadian Society of Petroleum Geologists and Alberta Research Council, 1994. pp 455–468

52. Rokosh CD, Lyster S, Anderson SDA, Beaton AP, Berhane H, Brazzoni T, Chen D, Cheng Y, Mack T, Pana C et al (2012) Summary

of Alberta's Shale-and siltstone-hosted hydrocarbon resource potential. In: ERCB/AGS Open File Report

53. Pengwei W, Zhuoheng C, Zhijun J, Yingchun G, Xiao C, Jiao J, Ying G (2019) Optimizing parameter "total organic carbon content" for shale oil and gas resource assessment: taking west canada sedimentary basin devonian duvernay shale as an example. Earth Sci 44:504–512. https://doi.org/10.3799/dqkx.2018.191

54. Haldar SK (2018) Exploration geophysics. Mineral exploration. Elsevier, Amsterdam, pp 103–122

55. Zhuang H, Han Y, Sun H, Liu X (2020) Introduction. Dynamic well testing in petroleum exploration and development. Elsevier, Amsterdam, pp 1–30

56. Asquith G, Krygowski D (2006) Basic well log analysis. Second Ed. AAPG, ISBN 9780891816676

57. Evenick JC (2019) Introduction to well logs and subsurface maps. 2nd Ed. PennWell Books, ISBN 9781593706487

58. Tixier M, Alger RP, Doh CA (1959) Sonic Logging. Trans AIME 216:106–114. https://doi.org/10.2118/1115-G

59. Chen Z, Jiang C, Lavoie D, Reyes J (2016) Model-assisted Rock-Eval data interpretation for source rock evaluation: examples from producing and potential shale gas resource plays. Int J Coal Geol 165:290–302. https://doi.org/10.1016/j.coal.2016.08.026

60. Castillo E (1998) Functional networks. Neural Process Lett 7:151–159. https://doi.org/10.1023/A:1009656525752

61. Castillo E, Cobo A, Gutiérrez JM, Pruneda RE (1999) Functional networks with applications. Springer, Boston

62. Castillo E, Gutiérrez JM, Hadi AS, Lacruz B (2001) Some applications of functional networks in statistics and engineering. Technometrics 43:10–24. https://doi.org/10.1198/00401700152404282

63. Castillo E, Cobo A, Gutiérrez JM, Pruneda E (2000) Functional networks: a new network-based methodology. Comput Civ Infrastruct Eng 15:90–106. https://doi.org/10.1111/0885-9507.00175

64. Tariq Z, Mahmoud M, Abdulraheem A (2019) Method for estimating permeability in carbonate reservoirs from typical logging parameters using functional network. In: 52nd US Rock Mechanics/Geomechanics Symposium. p 6

65. Tariq Z, Mahmoud MA, Abdulraheem A, Al-Shehri DA (2018) On utilizing functional network to develop mathematical model for poisson's ratio determination. In: 52nd US Rock Mechanics/Geomechanics Symposium. p 6

66. Tariq Z (2018) An intelligent functional network approach to develop mathematical model to predict sonic waves travel time for carbonate rocks. In: SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition. p 16

67. Narisetty NN (2020) Bayesian model selection for high-dimensional data. pp 207–248

68. Hang H-M, Chou Y-M (1995) Motion estimation for image sequence compression**This work was supported in part by the NSC Grant 83–0408-E009012. Handbook of Visual Communications. Elsevier, Amsterdam, pp 147–188

69. Jang J-SR (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23:665–685. https://doi.org/10.1109/21.256541

70. Jang J-SR (1991) Fuzzy modeling using generalized neural networks and kalman filter algorithm. In: Proceedings of the 9th national conference on artificial intelligence, CA, USA pp 762–767

71. Tahmasebi P, Hezarkhani A (2012) A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation. Comput Geosci 42:18–27. https://doi.org/10.1016/j.cageo.2012.02.004

72. Abraham A (2005) Adaptation of fuzzy inference system using neural learning. Fuzzy systems engineering. Springer, Berlin, pp 53–83

73. Shahriar K, Owladeghaffari H (2007) Analysis of Permeability Using BPF, ANFIS and SOM. In: 1st Canada-US Rock Mechanics Symposium, p 5

74. Frizzo Stefenon S, Zanetti Freire R, dos Santos Coelho L, Meyer LH, Bartnik Grebogi R, Gouvêa Buratto W, Nied A (2020) Electrical insulator fault forecasting based on a wavelet neuro-fuzzy system. Energies 13:484. https://doi.org/10.3390/en13020484