

Application of extreme learning machine and neural networks in total organic carbon content prediction in organic shale with wire line logs

Xian Shi ^a, Jian Wang ^{b,*}, Gang Liu ^a, Liu Yang ^c, Xinmin Ge ^d, Shu Jiang ^e

^a College of Petroleum Engineering, China University of Petroleum (Huadong), Qingdao, Shandong 266580, China

^b College of Science, China University of Petroleum (Huadong), Qingdao, Shandong 266580, China

^c State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum (Beijing), Beijing 102249, China

^d CNPC Key Well Logging Laboratory, China University of Petroleum, Qingdao 266580, Shandong, China

^e Energy & Geoscience Institute, University of Utah, Salt Lake City, UT 84108, USA



ARTICLE INFO

Article history:

Received 3 January 2016

Received in revised form

16 April 2016

Accepted 21 May 2016

Available online 24 May 2016

Keywords:

Organic shale

Total organic carbon

Extreme learning machine

Well logs

Artificial intelligence

ABSTRACT

Total organic carbon (TOC) is a critical parameter for source rock characterization in shale gas reservoirs. In this work, the use of extreme learning machines (ELM) for predicting TOC from well logs data have been investigated. We use log data from two wells located in an unconventional shale gas reservoir in the Sichuan Basin, China. Seven wireline logs from this well and a total of 185 TOC observations from core measurements were incorporated. Prediction accuracy of the model has been evaluated and compared with commonly used artificial neural network which is based on Levenberg–Marquardt logarithm (ANN-LM). An Extreme Learning Machine (ELM) network is a single hidden-layer feed-forward network with many advantages over multi-layer networks, such as fast computing speed and better generalization performance. The results demonstrated that TOC prediction by the ELM model and the ANN model, but the ELM method can achieve high accuracy while maintains high running speed. This study shows that ELM technology is a promising tool for TOC prediction, and this work can be incorporated into a software system that can be used in quick ‘sweet spot’ determination and well completion guidance.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Shale gas refers to an unconventional natural gas stored in low permeability shale, thus reservoir characterization is the essential step for shale gas well planning decisions (Pope et al., 2009; King, 2010). Shale gas well productivity depends on reservoir quality and successful hydraulic stimulations (Rickman et al., 2008; Glorioso and Rattia, 2012). Because not all shales are viable targets for economic hydrocarbon production, shale gas companies universally make special shale gas assessment criteria to help them rank appropriate stimulation strategies at wellbore, regional and basin scales (Chong et al., 2010). The definition of productive shale quality is based on many petrophysical properties, such as total organic carbon (TOC) content, thermally mature, permeability, porosity, saturation, rock mineralogy and mechanical properties et al. Most productive shale gas reservoirs can be judged by some critical reservoir properties qualitatively according to previous successful activities in shale plays, a potential commercial shale gas

reservoir typically has at least 2% TOC, and Ro (vitrinite reflectance) ranges within maturity windows (more than 1.4 in gas dry window). Furthermore, it needs to have less than 40% saturation and more than 2% porosity and 100 nanodarcy permeability, which means good gas storage and flow capability. Besides, commercial shales also need to have more than 40% quartz or carbonate in mineralogy, low differential stress and a certain degree of natural fractures, which means good fracability (Sondergeld et al., 2010). Among all the factors, TOC is considered to be a fundamental and important indicator for describing the resource potential of gas in kerogen-rich shale plays, thus a continuous and accurate TOC interpretation profile is highly desirable.

Direct geochemical analysis and well logging are used conventionally for TOC determination in current petroleum industry. However, core TOC data are not available because of the cost and time required for testing and the challenges associated with gathering a representative and intact sample. Although laboratory tests of TOC are difficult, they are still the necessary and preferred methods (Zhu et al., 2003; Jarvie et al., 2015). Moreover, the lab results are often applied as references for further prediction by mathematical approaches. Log-based TOC predictions are more

* Corresponding author.

universally applicable because they can provide continuous TOC profiles that cover the whole interval of interest. Some particular geophysical responses (e.g., gamma ray, neutron, resistivity, density) of the source rock can be detected comparing to surrounding rocks. Empirical mathematical equations are generally used when carrying out log-based TOC prediction. However, the estimation results by logging information rely greatly on equation quality. Meanwhile, the uranium and gamma ray correlation methods are sometimes not suitable for gas marine shale reservoirs that have natural radioactivity in phosphatic fish plates, thus elevated uranium and gamma ray counts cannot reflect TOC. (Passey et al., 1990; Carpentier et al., 1991).

Practically, because the nonlinear and complex relationships between TOC and logs, it is very hard to obtain a universal solution for all wells in one basin. In recent years, the application of artificial intelligence techniques for TOC prediction was found to be more reliable to solve this complicated regression task. In fact, the use of robust artificial intelligence methods approaches has been introduced and successfully employed in many petroleum engineering fields, such as lithology classification, permeability estimation and hydraulic fracturing candidate selection et al. (Shadizadeh et al., 2010; Zoveidavianpoor et al. 2013a,b; Yuan et al. 2014). These methods combine the accuracy of numerical models with the simplicity of analytical approaches, while it is free from constraints of a certain function form. Huang and Williamson (1996) proposed an improved ANN (Artificial neural network) method to map source rock intervals at the Jeanne d'Arc Basin (offshore eastern Canada) by a combination of the 'quickprop' algorithm and 'dynamic node creation' scheme. The Neural-Fuzzy approach proposed by Kamali and Mirshady (2004) is regarded as a good example of TOC prediction through multi-parameter correlation analysis and fuzzy neural networks. Guo et al. (2009) suggested a method integrating cross-plotting, fuzzy ranking and an ANN to predict the TOC content of a mature carbonate resource. Amiri Bakhtiar et al. (2011) adopted a combination of sonic and resistivity logs ($\Delta \log R$ method) and a neural network method to calculate TOC values in the Pabdeh formation in the Ahwaz and Marun oilfields, Iran, but only resistivity and sonic logs were applied for TOC prediction in his paper. Except for neural network approaches, the SVM (support vector machine) approach for regression of TOC in gas-bearing shale was presented, and results show that the SVR technology is more effective and applicable than conventional neural network approaches. (Tan et al. 2015). All of these attempts show a certain degree of success but have some shortcomings. More specifically, some inherent problems are encountered with the algorithms employed in neural network methods. Initially, most feed-forward artificial neural networks use gradient descent algorithms, which need to iteratively update the model weights and biases. Thus, the training process is quite slow. In addition, the solutions may become trapped in local minima in the objective functions, resulting in failure to achieve the global minimum (i.e., the global best fit model). Furthermore, the performance is sensitive to the learning rate, which is difficult to optimally choose ahead of time. Therefore, many new algorithms need to be proposed to overcome these shortcomings with respect to low training speed and generalization ability. (Zoveidavianpoor et al., 2012, 2013a,b) Table 1 summarizes current TOC prediction methods using well logs.

Extreme learning machine (ELM) (Huang et al., 2006a,b) is an algorithm for the single-layer feed-forward network, which is considered to be a distinct artificial intelligence technique from the conventional ANNs. This algorithm is capable of solving problems using gradient descent-based algorithms such as back propagation, which is often applied in ANNs. The ELM is also able to be trained in much less time than an ANN. In fact, it has been shown that by utilizing the ELM, the learning process becomes very fast and leads

to robust performance, which is preferable for small samples and high-dimensional non-linear learning problems. Accordingly, a number of investigations have been successfully carried out with the application of ELM for solving problems in petroleum engineering, particularly in drilling optimization and permeability estimation (Mortaza et al., 2016). Cao et al. (2015) performed reservoir parameter estimation using an extreme learning machine in a heterogeneous sandstone reservoir. Compared to back-propagation (BP) network and SVM approaches, the robust ELM algorithm provides faster and more accurate prediction results. Sunday et al. (2013) investigated the feasibility of ELMs in forecasting permeability from well logs in a Middle Eastern industrial reservoir; the ELM was shown to have better generalization and to be faster in permeability estimation than other methods. Chandra (2013) also applied an ELM algorithm to predict permeability, reaching similar results showing that the ELM was a better predictor than the SVM. However, ELM has not yet been used for prediction of TOC in any hydrocarbon reservoirs. This paper investigates the application of ELM in predicting TOC in a shale gas reservoir. Moreover, in the process of ELM network construction, the influence of well logs on TOC prediction is studied. The results and performance characteristics of the ELM technique are compared to those obtained by the ANN method to evaluate the efficiency of these two networks during the prediction process.

2. Artificial neural networks and extreme learning machines

2.1. Methodology of artificial neural networks

The mathematical models of ANNs are inspired by the functions of the biological nervous system. They are composed of a large number of neurons which are distributed in different layers based on the distinct functions.

Specifically, there are three types of layers: input layer, one or more hidden layers and output layer, each of them consists of a preset number of neurons. It has been rigorously proved that ANN is an universal approximator as it can approximate any continuous function with an arbitrary precision. For feed-forward neural networks, the information moves in the direction: input neurons, activations through the hidden neurons and to the outputs. In supervised learning, the training process is to adjust the weights of the neurons which minimizes the error function between the networks' actual outputs and the desired value over all the example pairs. There are numerous strategies available for training ANNs (Zoveidavianpoor et al. 2013a,b). The most popular of them is the backpropagation (BP) algorithm. It is a type of supervised learning using for feed-forward neural networks. Error function is computed by the difference between the actual and target outputs.

The standard BP algorithm employs the classical optimal method: gradient descent method. Due to its poor performance on real applications, different variants have emerged in the past decades. On the basis of the different optimizing strategies, the common used variants are: Levenberg-Marquardt (LM), Conjugate gradient method with Powell-Beale, Fletcher-Reeves and Polak-Ribiere updates, gradient descent method with momentum term, penalty term and adaptive learning rate, Bayesian regulation and scaled conjugate gradient.

2.2. Methodology of extreme learning machines

ELM was originally proposed as a single hidden-layer feed-forward neural network and later was extended to more generalized single hidden-layer feed-forward neural networks where the hidden layer may not be made of homogeneous neurons (Huang, 2014). ELM can avoid many obstacles faced by back-propagation

Table 1

Summary of log-based TOC prediction models.

Method	Parameters	Explanations and features	References
$\Delta \log R$ method	Porosity and resistivity	(1) LOM range varies in specific areas (2) Baseline and TOC background levels vary regionally (3) Underestimates TOC in over mature shale gas reservoir or formation with an abundance of pyrite	Passe et al. (1990) Guo et al. (2009) Hu et al. (2011) Wang et al. (2016)
Natural gamma method	Gamma intensity	(1) Better for marine source rocks with concentrated uranium (2) A limitation for shale gas formation that is non-radioactive or for those that have phosphatic minerals with elevated uranium content	Beers (1945) Schmoker, J. (1979)
Bulk density method	Density	(1) Need a strong correlation between density and TOC (2) Suitable for reservoirs with similar fluid phases and consistent mineralogy	Schmoker, J. (1979) Schmoker and Hester (1983) Meyer and Nederlof (1984) Decker et al. (1993)
Interval transit time method	Compressional transit-time	(1) Need a strong correlation between compressional transit time and TOC	Autric and Dumesnil, 1985
Spectral gamma ray method	Uranium, potassium, uranium content	(2) Many factors can significantly affect the sonic velocities	Gonfalini M. (1991)
Geochemical Laboratory measurements (Leco TOC and Rock Eval)	(1) S_1 , S_2 , S_3 peaks in rock Eval technique (2) Combust pulverized rock	(1) Not available for all of the wells (2) Expensive	Peters, (1986)
Volume model method	Hydrocarbon and water volume, kerogen volume clay volume, carbonate volume and siliceous volume	(1) Direct measurements from core samples (2) Discrete data points (3) Expensive (1) Composite well log integration (2) The natural gamma or spectral gamma ray logging data are necessary (3) Organic carbon conversion factor (K) varies regionally	Jarvie, D.M. et al. (2015) Mendelson and Toksoz (1985) Huang R.C. et al. (2015)
Multivariate fitting	Composite well logs	(1) Need database establishment (2) Hard to determine the relevant parameters due to nonlinear relationship between different well logs (3) incorporate the random errors of logging	Heidari et al. (2011)
Laser-induced breakdown spectrometry, chemostratigraphy, RockView	The carbon abundance of formation and carbonate mineral in spectroscopy	(1) Core technologies controlled by a few international oil service companies (Schlumberger -Litho Scanner SM , Baker Hughes-RockView SM services et al.) (2) The chemical contents can be determined directly	Pemper et al. (2009) Alyssa and Susan, 2013
Neural networks methods	Composite well logs	(1) Kernel function is significant in SVR methods (2) Data-preprocessing is necessary due to over-fitting input problem (3) Heavy computational work or error if algorithm is not reasonable (4) Suitable for TOC prediction at the early stage of petroleum exploration	Huang et al. (1996) Kamali et al. (2004) Bakhtiar et al. (2011) Tan M.J et al. (2015)

neural networks, such as the selection of stopping criteria, learning rate and training epochs due to its internal mechanisms (Huang et al., 2006a,b; Akusok et al., 2015). Fig. 1 illustrates the structure of the ELM method.

ELM development was motivated by biological learning systems and presented to overcome the challenging problems encountered by BP learning algorithms. Neural network learning is complicated; however, it usually works universally for feature extraction, clustering, regression, and classification and requires little human intervention in many cases. Inspired by biological learning characteristics, a neural network may have random neurons with all of their parameters set independent of the circumstances. This is the case in ELM. Its learning efficiency and effectiveness were justified as early as 2005 (Li et al., 2005). In addition, its universal approximation capability was rigorously proved. (Huang et al., 2006a,b; Huang and Chen, 2007, 2008). The correspondence between

specific biological neural configurations consequently appeared in the literature (Sosulski et al., 2011; Eliasmith et al., 2012; Barak et al., 2013; Rigotti et al., 2013).

Unlike other randomness-based training methods/network models (Igelnik and Pao, 1995), all of the hidden neurons in the ELM model are independent of the training samples and each other. Although hidden nodes are important and crucial for training, they do not need to be tuned, and the hidden weights can be randomly generated before the learning process. Moreover, the architecture of ELM is robust enough when it has enough hidden neurons for the given problems, which is not the case for conventional learning methods that are highly dependent on the data.

For ELM, the connecting weights between the input and the hidden layer do not need to be tuned. The output of the ELM is as follows, using a simple model with one output node as an example:

$$f_L(x) = \sum_{i=1}^L \beta_i h_i(x) = h(x)\beta \quad (1)$$

where $\beta = [\beta_1, \dots, \beta_L]^T$ is the vector of the output weights between the hidden layer and the output node, and $h(\mathbf{x}) = [h_1(\mathbf{x}), \dots, h_L(\mathbf{x})]$ is the output (row) vector of the hidden layer with respect to the input \mathbf{x} . $h(\mathbf{x})$ actually maps the data from the d -dimensional input space to the L -dimensional hidden-layer feature space (ELM feature space), and thus, $h(\mathbf{x})$ is indeed a feature mapping. For binary classification applications, the decision function of ELM is:

$$f_L(x) = \text{sign}[h(x)\beta] \quad (2)$$

In contrast to traditional learning algorithms, ELM tends to reach not only the smallest training error but also the smallest norm of output weights (Huang et al., 2004). According to Bartlett's theory, the smaller weights lead to better generalization performance for a similar training error compared to trained neural networks (Liang et al., 2006). We state that the weights of ELM need not be adjusted for any infinitely differentiable activation function $h(x)$. The input weight can be randomly chosen at the beginning of the training procedure. Then the weights of input layer are fixed in the training. The weights connecting the hidden and output layers can be seemed as to search the least square solution of the following objective function.

The ELM minimizes the training error as well as the norm of the output weights:

$$\text{Minimize : } \|H\beta - T\|^2 \text{ and } \|\beta\| \quad (3)$$

where H is the activation output matrix of hidden-layer.

The essence of ELM is to minimize the training error and the norm of connecting weights β (from the hidden layer to output layer) at the same time. By solving the least squares problem, the hidden layers' output weights are implemented as follows (Huang et al., 2012):

$$\beta = H^\dagger T \quad (4)$$

where H^\dagger is the Moore-Penrose pseudoinverse of matrix H .

The ELM training algorithm can be summarized as follows (Huang et al., 2012):

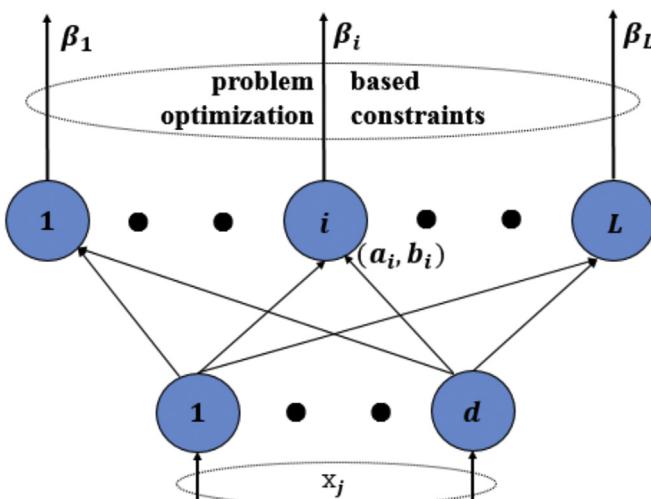


Fig. 1. Structure and schematic diagram of the ELM method used in this paper.

- 1) Randomly assign the hidden node parameters, e.g., the input weights a_i and biases b_i for additive hidden nodes, $i = 1, \dots, L$.
- 2) Calculate the hidden layer output matrix H .
- 3) Obtain the output weight vector

$$\beta = H^\dagger T$$

$$\text{where } T = [t_1, \dots, t_n]^T$$

There are different methods to evaluate the Moore-Penrose pseudoinverse of a matrix such as the orthogonal projection method, the orthogonalization method, the iterative method and singular value decomposition (SVD) (Chorowski et al., 2014). The orthogonal projection method can be employed in two cases: when $H^T H$ is nonsingular and $H^\dagger = (H^T H)^{-1} H^T$, or when HH^T is nonsingular and $H^\dagger = H^T (HH^T)^{-1}$. In light of ridge regression theory, one can add a positive value to the diagonal of $H^T H$ or HH^T . The resultant solution is more stable and tends to have better generalization performance.

In terms of ELM theory, many feature mapping functions $h(x)$ may be employed in ELM design. Thus, ELM has the ability to approximate any continuous target function. Although the actual activation functions of human brain systems are unknown, they are most likely nonlinear piecewise continuous stimuli. In fact, there are two important theorems to support the ELM models that can be categorized into the following two cases: Universal Approximation and Classification Ability.

2.2.1. Universal approximation theorem (Huang et al., 2006a,b; Huang and Chen, 2007, 2008)

Given any bounded, non-constant piecewise continuous function as the activation function in hidden neurons, if by tuning parameters of the hidden neuron activation function SLFNs can approximate any target continuous function, then for any continuous target function $f(x)$ and any randomly generated function sequence, $\{h_i(x)\}_{i=1}^L \lim_{L \rightarrow \infty} \left\| \sum_{i=1}^L \beta_i h_i(x) - f(x) \right\| = 0$ holds with probability one given the appropriate output weights β .

On the basis of the nonlinear piecewise continuous activation functions and their linear combinations, ELMs can not only approximate any continuous mapping but also divide arbitrary disjoint regions of any shapes with enough randomly hidden neurons. In particular, the theorem of ELMs' classification capability has been presented as following:

2.2.2. Classification capability theorem

Given a feature mapping $h(\mathbf{x})$, if $h(\mathbf{x})$ is dense in $C(\mathbb{R}^d)$ or in $C(M)$, where M is a compact set of \mathbb{R}^d , then ELM with random hidden layer mapping $h(\mathbf{x})$ can separate arbitrary disjoint regions of any shapes in \mathbb{R}^d or in M .

This theorem implies that ELMs with bounded non-constant piecewise continuous functions have universal approximation capability. It does not require the special training approaches to update the hidden neurons. That is, the hidden weights in ELMs can be randomly generated without tuning, which is more like the activation of the human brain.

3. Data processing and analysis

3.1. Study area

Well J1 and Well J2 are shale gas exploration wells in the southeast of the Lower Silurian Longmaxi Formation in the Sichuan Basin, China. The strata contain volcaniclastic deposits of the continental island arc on the northern edge of the North China plate to its north. Relative to other districts, Sichuan gas-bearing shale has the advantages of great thickness, wide distribution and good

Table 2

Summary of the recorded logging parameters for well J1.

	DTC (μs/ft)	GR (API)	CNL (%)	DEN (g/cm³)	KTH (%)	TH (ppm)	U (ppm)	TOC (wt%)
Minimum value	200.39	154.12	11.73	2.52	0.81	11.24	7.56	1.11
Maximum value	230.74	191.42	17.22	2.65	1.25	19.63	13.45	3.69
average	219.98	176.15	14.20	2.58	1.03	15.07	10.18	2.55
Standard deviation	6.03	9.42	1.56	0.03	0.12	2.28	1.66	0.57

Table 3

Summary of the recorded logging parameters for well J2.

	DTC (μs/ft)	GR (API)	CNL (%)	DEN (g/cm³)	KTH (%)	TH (ppm)	U (ppm)	TOC (wt%)
Minimum value	195.95	134.82	12.88	2.55	0.95	12.90	3.77	0.19
Maximum value	248.29	214.64	23.85	2.70	1.45	23.81	12.81	3.25
average	224.54	163.31	16.58	2.63	1.12	16.38	7.09	1.78
Standard deviation	8.59	18.89	2.22	0.03	0.11	1.95	2.13	0.73

preservation. In recent years, exploration and production within the Sichuan Basin have been performed by PetroChina, Shell, and Sinopec, among others. The prior reservoir characterization results show that the TOC content of the Longmaxi organic-rich shales ranges from 0.5% to 8%, and the type of kerogen is predominantly type I (rich in algal and amorphous organic matter) and type II (rich in algae and herbaceous matter). The mineral composition of shale samples was measured with X-ray diffraction. The samples have quite similar mineral compositions as the American Ohio and Woodford shales. In the study area, the average contents of clay, quartz, feldspar plus pyrite, and carbonate in the Longmaxi black shales are 35.67%, 45.05%, 6.32% and 17.71%, respectively. The promising shale gas prospects in this area are indicated in appraisal wells.

Database preparation for ELM training is a crucial step, as in any neural network modeling. In this study, TOC was measured with a LECO CS-400 carbon sulfur analyzer (combustion at temperatures over 800 °C). A total number of 185 TOC data points were obtained from laboratory tests, 91 data come from well J1 and 94 data belong to well J2. Available well logs of Well J1 and well J2 include natural gamma ray (GR), compressional wave slowness (DTC), neutron porosity (CNL), lithology density (DEN), and natural spectrum logs including Uranium(U), Thorium (Th), and Potassium (KTH). The summary of the recorded logging parameters and laboratory core TOC data for well J1 and well J2 are given in Tables 2 and 3.

3.2. Data analysis

In general, some abnormal readings are found in well logs (especially gamma ray, porosity and resistivity data) from organic-rich rocks. However, these problems are basin-specific and care needs to be taken because TOC is related simultaneously to density, porosity, resistivity, sonic and gamma measurements, and other characteristics. To evaluate simple relationships of well logs with core TOC data, two analysis techniques were used in this paper to characterize the sensitivity of the TOC values to the log data. This process is also beneficial for the selection of input variables because some irrelevant well logs have to be omitted. Otherwise, the input to the model may be too great, potentially overwhelming the model and modeling the noise instead of the signal. A simple linear regression test (cross plotting) was conducted, and the correction coefficient (R^2) was used as a critical indicator to investigate the influences of different well logs on the laboratory-measured TOC results. The R^2 represents the proportion of the overall variance explained by the model, which can be calculated as:

$$R^2 = 1 - \frac{\sum_{i=1}^n (f(x_i) - y_i)^2}{\sum_{i=1}^n f(x_i)^2 - \sum_{i=1}^n (y_i)^2 / n} \quad (5)$$

Cross plots of well log parameters of well J1 are shown in Fig. 2. It can be observed that the DEN log has a negative relationship with TOC in linear regression, which also has the highest correction coefficient of 0.5061. However, the GR, U and DTC log data appear to have positive relationships with TOC, but these well logs have poor correction coefficients of 0.1796, 0.3306, and 0.1432, respectively. Cross plots of well log parameters of well J2 are shown in Fig. 2. Similar to well J1, the DEN log has a negative relationship with TOC in linear regression, which has the highest correction coefficient of 0.6041. In addition, correction coefficients of GR, U and DTC log data with TOC was found to be, 0.0674, 0.2386, and 0.0641, respectively (see Fig. 3).

In addition to simple linear regression method, we also apply correlation coefficient analysis to analyze the relationships between TOC core data and well logs. As presented in Table 4, the DEN log has the highest correlation coefficient of -0.711 for well J1, while U and DTC log data also correlate well. Both cross plotting and correlation coefficient methods demonstrated that GR, DTC, and U data have more significant relationships with TOC core measurements for well J1. As for well J2 (Table 5), the DEN log also has the highest correlation coefficient of -0.777 and the U log has the second highest correlation coefficient of -0.488. In addition, the other parameters also correlate well with TOC. Both cross plotting and correlation coefficient methods demonstrated that KTH, CNL and U data have more significant relationships with TOC core measurements.

It can be observed that there are no apparent and universal correlations between well logs and TOC laboratory measurements. Thus, the linear or non-linear regression techniques would fail in predicting TOC from well logs, but this analysis is beneficial for determining a general TOC distribution pattern for this area. Considering the purpose of this study, we only analyze four well log data types that have a significant influence on TOC prediction:

(1) Natural gamma ray and spectral gamma ray logs

Organic-rich shale (especially marine shale) produces high gamma ray values because the presence of concentrated uranium or uranium ions in gas-bearing shale strata can result in relatively intense radioactivity. The ability to separate enrichment levels of thorium, potassium and uranium from radiation measurements can be achieved by a spectral gamma ray tool, which can provide stronger evidence for TOC presence than total gamma ray data.

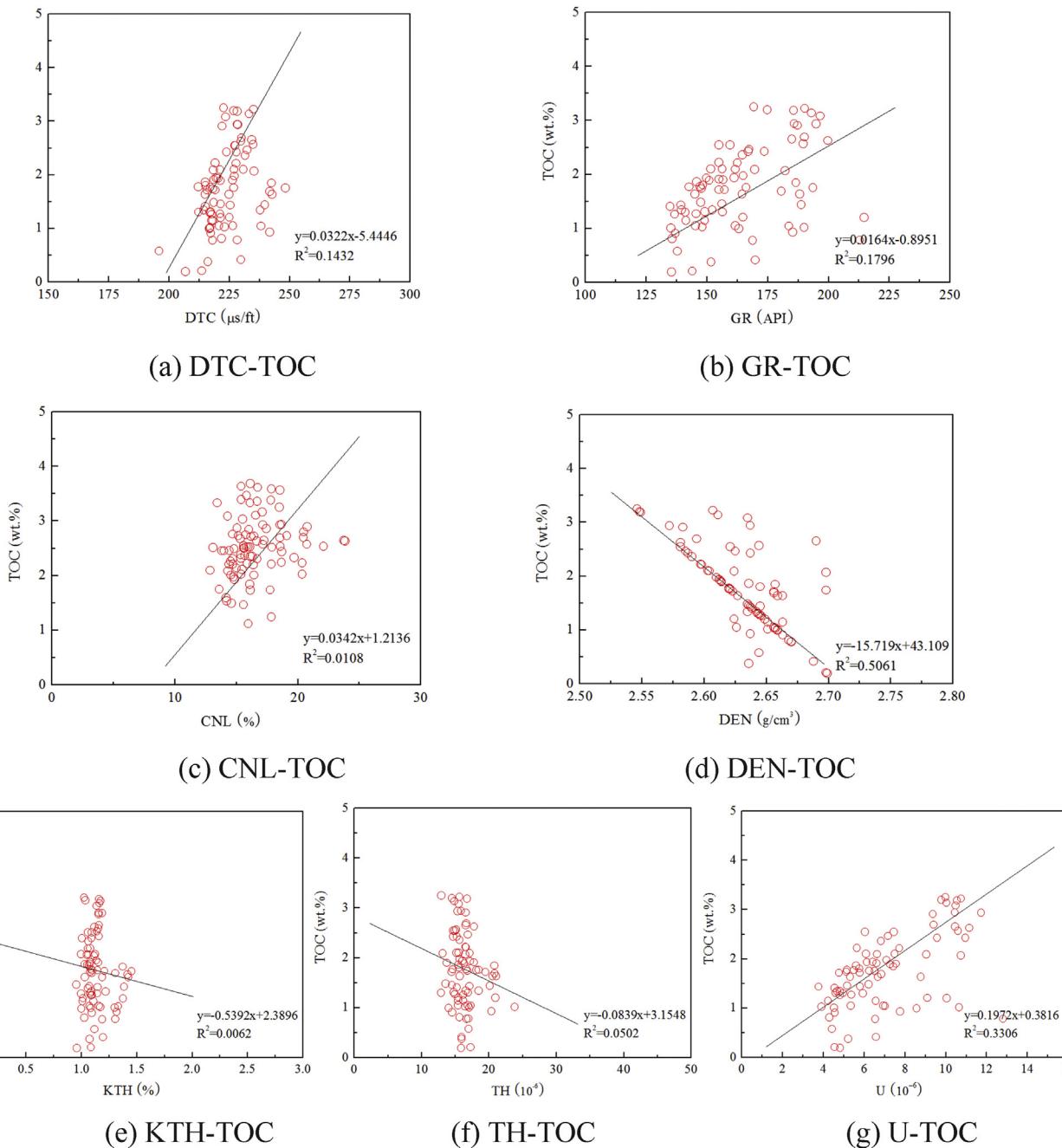


Fig. 2. Cross plots of TOC laboratory-measurements and the logging data for well J1 (a): DTC-TOC (b) GR-TOC; (c) CNL-TOC; (d) DEN-TOC; (e) KTH-TOC; (f) TH -TOC; (g) U-TOC.

Generally speaking, uranium is associated with organic matter and precipitated salts, while thorium and potassium are associated with clay minerals.

(2) Density logs

Strata with high TOC usually have low density because solid organic matter has a similar density to water and is considerably less dense than the surrounding rock. Therefore, when TOC matures, the density decreases. However, a decrease in rock density can be generated from many factors other than organic matter, including increased porosity, clay minerals, etc. Therefore, estimating TOC using density logs is preferable for reservoirs with similar fluid phases and consistent mineralogy.

(3) Sonic logs

The reciprocal of the P-wave velocity can be described by the interval transit time. The presence of organic matter leads to abnormally high observed values of interval transit time. There are correlations between TOC and interval transit time for many cases in previous papers. Most cases show that there is a natural trend of decreasing sonic transit time for source rocks containing more organic matter.

Resistivity and compensated neutron logs are not analyzed in this paper, but some reports describe their clear relationships with TOC in some cases. Typically, the resistivity of rock increases when maturity increases because organic matter is not electrically conductive. The neutron porosity decreases with the decrease in

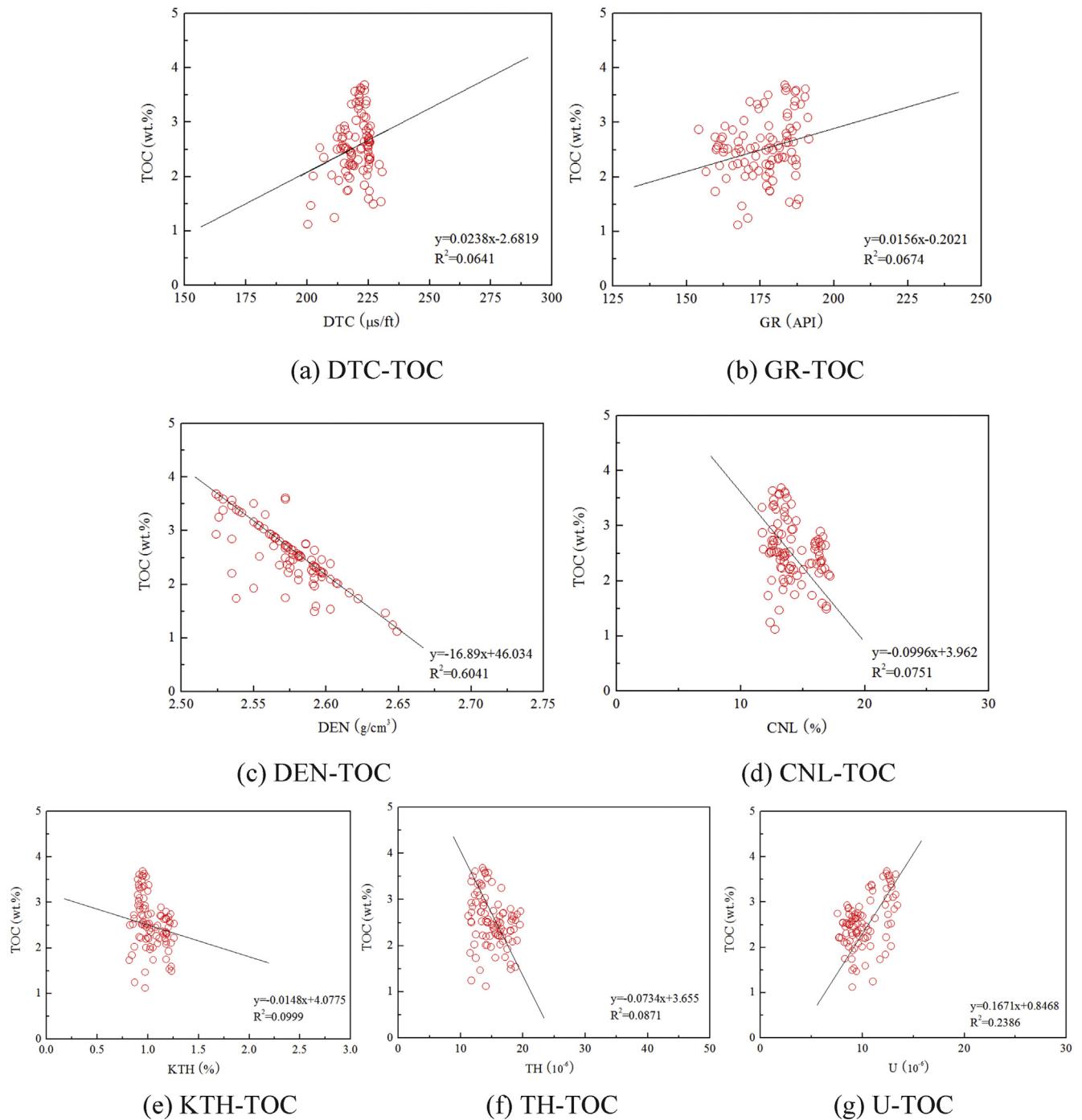


Fig. 3. Cross plots of TOC laboratory-measurements and the logging data for well J2 (a): DTC-TOC (b) GR-TOC; (c) CNL-TOC; (d) DEN-TOC; (e) KTH-TOC; (f) TH -TOC; (g) U-TOC.

Table 4
Correlation matrix of Well J1.

	TOC	GR	KTH	TH	U	CNL	DTC	DEN
TOC	1							
GR	0.424	1						
KTH	-0.078	0.604	1					
TH	-0.224	0.373	0.7861	1				
U	0.575	0.819	0.136	0.021	1			
CNL	0.104	0.513	0.580	0.328	0.234	1		
DTC	0.378	0.609	0.524	0.270	0.4	0.731	1	
DEN	-0.711	-0.410	0.065	0.183	-0.567	-0.020	-0.193	1

Table 5
Correlation matrix of Well J2.

	TOC	GR	KTH	TH	U	CNL	DTC	DEN
TOC	1							
GR	0.260	1						
KTH	-0.316	0.200	1					
TH	-0.295	0.151	0.936	1				
U	0.488	0.454	-0.708	-0.682	1			
CNL	-0.274	0.458	0.824	0.750	-0.373	1		
DTC	0.253	0.556	0.300	0.280	0.216	0.595	1	
DEN	-0.777	-0.272	0.129	-0.353	-0.353	0.110	-0.308	1

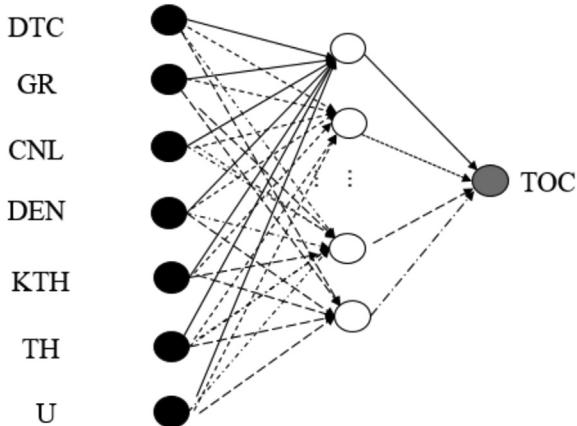


Fig. 4. Schematic of architecture of the neural network models.

hydrogen index, but evaluating the effects of reduced porosity due to the lack of hydrogen in gas and kerogen is quite complex in shale gas formation, given the complex mineral compositions. Regardless, there are no known general relationships between TOC and resistivity/compensated neutron logs.

4. Results and discussion

4.1. Model performance

As mentioned before, 185 data points measured from core analysis are randomly divided into training and testing subsets. In

this study, the training subset was made from 75% of the data, and the testing data consist of the remaining 25%. For comparison, ANN and ELM models were chosen and applied in TOC prediction tests. Schematic of architecture of the ELM and ANN network model can be illustrated as Fig. 4.

At the beginning, all of the data were normalized between 0 and 1 for training and testing dataset in the ANN and ELM model. This pre-processing can guarantee the convergence of neural nets while also increasing the computing speed of network methods.

In this study, a min-max normalization pre-processing step was applied to the data series. This method is used by the following equation:

$$\bar{x}_i = \frac{x_i(k) - \min_k}{\max_k - \min_k} = \{\bar{x}_i(1), \bar{x}_i(2), \dots, \bar{x}_i(n)\}, (i = 0, 1, \dots, m) \quad (6)$$

where x_i is the standardization processing result, $x_i(k)$ denotes the raw data, and x_{\max} and x_{\min} represent the maximum and minimum values from the original dataset, respectively.

The predictive performance of the ANN and ELM models with different well log inputs is evaluated using the coefficient of determination (R^2), as shown in Figs. 5 and 6. Here, we only show the comparison of two cases (in terms of the number of data types used) in ELM and ANN modeling in this paper: one has three logs as inputs, and the other uses seven logs. As a matter of fact, the TOC prediction performances with different well logs were totally compared, we applied ANN and ELM models for TOC prediction

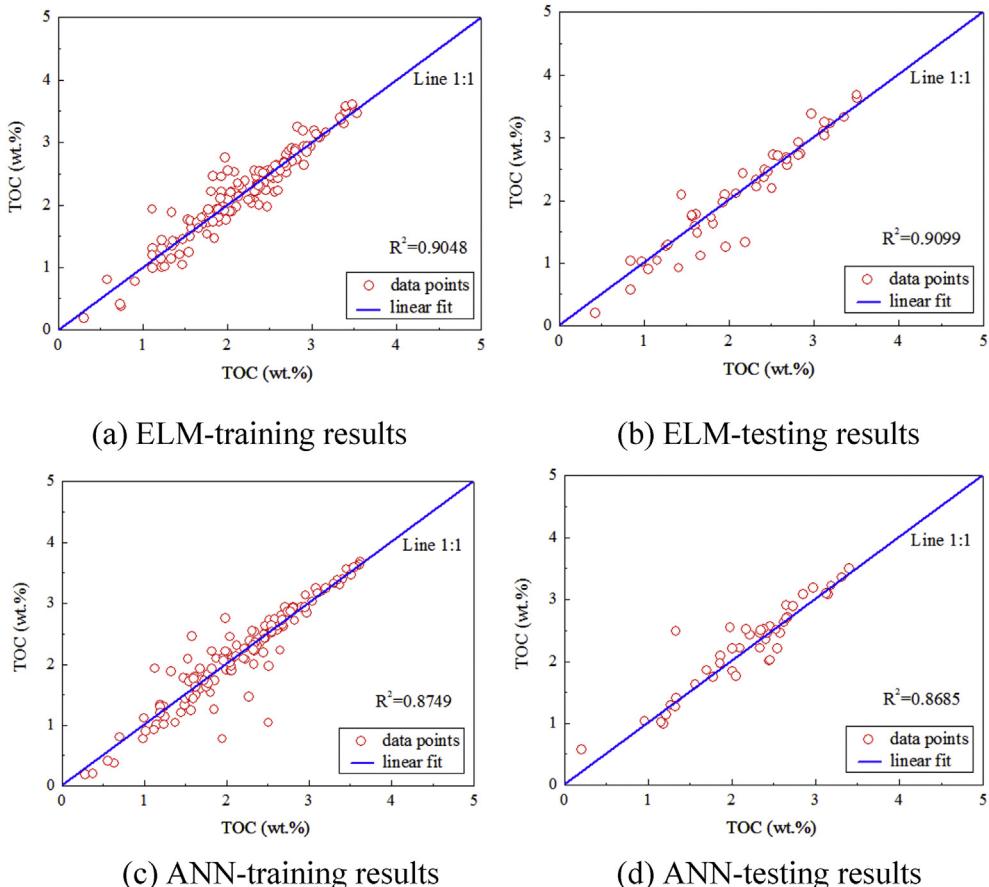


Fig. 5. Comparison of the TOC prediction results with 7 logs as inputs and laboratory measurements of the core samples (a): ELM-training results (b) ELM-testing results; (c) ANN-training results; (d) ANN-testing results.

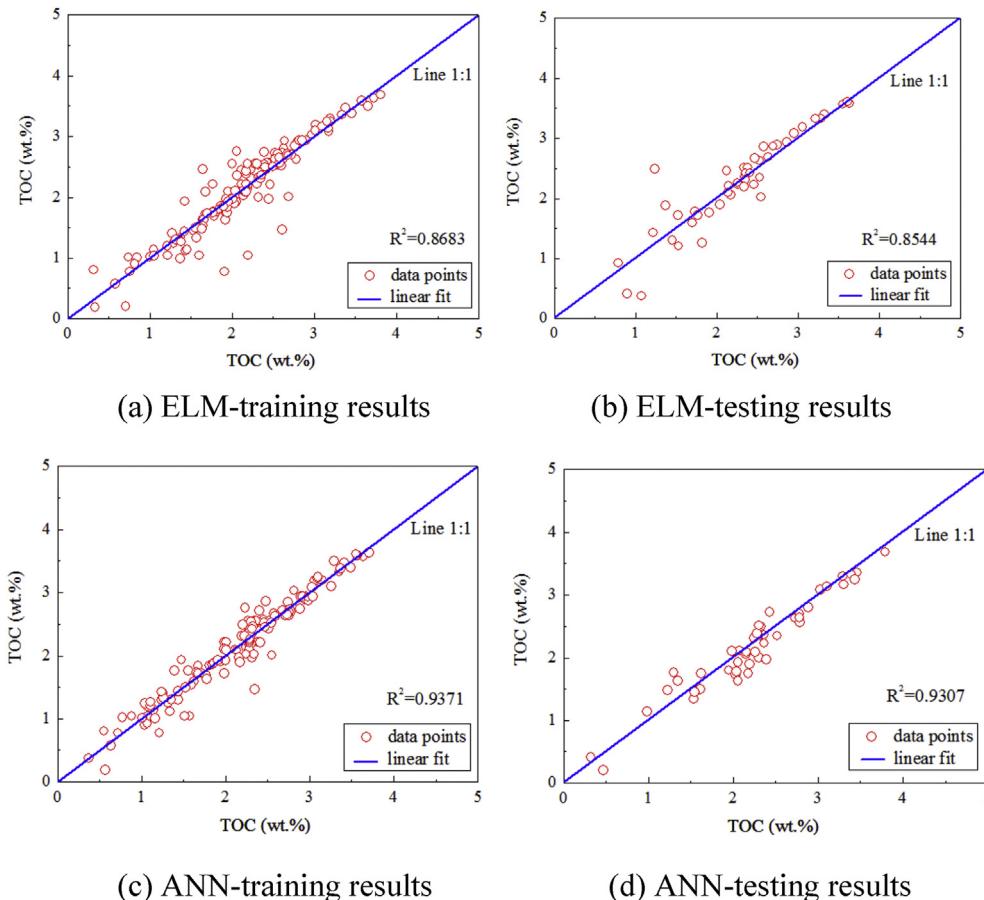


Fig. 6. Comparison of the TOC prediction results with selected 3 logs as inputs and laboratory measurements of the core samples (a): ELM-training results (b) ELM-testing results; (c) ANN-training results; (d) ANN-testing results.

with one well log (DEN) and two well logs (DEN and U), the results of the ANN and ELM model deviate tremendously from the TOC observations. Thus, we do not discuss these two situations. The primary log selection is determined using the correlation matrix ranking. The Levenberg–Marquardt (LM) algorithm has been utilized to update the weights of ANN model. This algorithm effectively accelerates network training speed and minimize the error level in artificial neural networks.

Figs. 5b and 6b show the cross plots of ELM-derived TOC results and the core measurements in which the correlation coefficient between the predicted and observed TOC values was 0.9099 for the seven log model and 0.8544 for the three log model. Figs. 5d and 6d show the cross plots of ANN-derived TOC results and the core measurements in which the correlation coefficient between the predicted and observed TOC values was only 0.8685 for the seven log model and 0.9307 for the three log model. If R^2 is greater than 0.9, model performance can be regarded as premium. In general, both ANN and ELM model yield correlation coefficients that are higher than 0.85, which demonstrates that the predicted TOC was successful. Moreover, comparing the results with different well log inputs, we can see that, the prediction precision is different in both models, but the ELM model accuracy is higher than ANN model with seven well log inputs while the ANN model accuracy is higher with three well log inputs, which demonstrate the advantage of high-dimensions and multilinear regression capability for ELM model. Furthermore, although ANN and ELM are competent for TOC prediction, ANN based models need more attentions on the optimization of many factors, such as initial weights, training ratio,

activation function, number of hidden neurons.

In addition to coefficient of determination (R^2), MAE (mean absolute error), RMSE (root mean square error), and VAF (variance accounted for) criteria were also used to compare prediction performance. These performance indicators can give a sense of how good the performance of a predictive model is relative to the actual value.

MAE can be described as:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n (f(x_i) - y_i) \quad (7)$$

The RMSE is conventionally used as an error function for quality monitoring of the model. Model performance increases as RMSE decreases. RMSE can be calculated by the following equation:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n [f(x_i) - y_i]^2}{n}} \quad (8)$$

VAF is often used to evaluate the correctness of a model, by comparing the measured values with the estimated output of the model. VAF is computed by the following equation:

$$\text{VAF} = \left(1 - \frac{\text{var}[f(x_i) - y_i]}{\text{var}[f(x_i)]} \right) \times 100 \quad (9)$$

where, for Equations (6)–(8), y_i is the measured data and $f(x_i)$ denotes the predicted data, x_1, \dots, x_n are the input wireline logs, and n is the number of samples used for training or testing neural

networks.

The results of MAE, RMSE and VAF are listed in Table 6. The highest VAF and the lowest MAE and RMSE belong to the ANN model with three well logs inputs, indicating that it gives better prediction performance than the ELM model. During the training phase, the decreased predictive accuracy of the ANN model is shown by RMSE values of 0.1995 (3 logs) and 0.2734 (7 logs), compared to the RMSE values of 0.2702 (3 logs) and 0.0692 (7 logs) of ELM model. The VAF values are 86.30 and 89.18 for the ANN and ELM models with seven well log inputs in the testing phase, while the RMSE values of ELM model are lower than that of ANN model. Thus, we can conclude that the ELM model has a higher training accuracy than the ANN model with seven well logs as inputs. In addition, we also compared processing speed and RMSE of the ELM and ANN models with different hidden nodes. The codes in both models were run one thousand times, and the training time and testing RMSE were recorded with 50 hidden nodes. As shown in Fig. 7, ANN requires more training time than ELM does, and training time increases with the increasing number of hidden nodes. However, less than one second was required for the ELM model for all cases because the ELM time requirements are not sensitive to the number of hidden nodes. Furthermore, the ELM can produce relatively more accurate testing results, and the testing RMSE of the ELM model remains relatively steady with increasing hidden nodes. Conversely, the ANN's computational time requirements scale dramatically with the number of hidden neurons; thus, the testing RMSE of the ANN model seems to be random with increasing hidden nodes, which indicates that ANN model is more time-

consuming than ELM in finding the optimum solution.

4.2. Model validation

To estimate the TOC of Well J1 and Well J2, we first need to construct and train the ELM and ANN models with core TOC measurements. To visualize the quality of the prediction, the predicted TOC values from the different models are compared with measurements for the whole dataset and are shown in Figs. 8–11. The left three curves are well logs. The corresponding TOC prediction curves are denoted as ELM or ANN, and the number of logs used is appended as the last character. The corresponding TOC curves with ANN model are denoted as ANN7 with seven logs, ANN6 with six logs, ..., and ANN 3 with three logs as inputs. The corresponding TOC curves with ELM model are denoted as ELM7 with seven logs, ELM6 with six logs, ..., and ELM3 with three logs as inputs. Comparison of the right five curves in Fig. 8 makes us conclude that the ELM-derived results with four, five and six log inputs are consistent with the measured TOC results of core samples. Moreover, the ANN method with four, five and six log inputs also can yield good results (Fig. 9). In the 2670 m–2850 m interval of well J1, it seems that the predicted results from 2760 to 2800 m are much better than the other intervals with the ELM and ANN model. Furthermore, in the 2700 m–2720 m, both ANN and ELM model can only yield acceptable results. In the 2950 m–3130 m interval of well J2, it seems that the predicted results from most intervals are consistent with TOC laboratory data in the ELM and ANN model (Figs. 10 and

Table 6

The calculated RMSE and VAF indicators of ANN model and ELM model.

Data set	Performance indicator	ANN model	ELM model
3-log input training (DEN/U/GR)	VAF	93.95	84.83
	MAE	0.0382	0.0730
	RMSE	0.1995	0.2702
	R ²	0.9371	0.8683
3-log input testing (DEN/U/GR)	VAF	93.04	82.22
	MAE	0.0415	0.0864
	RMSE	0.2038	0.2940
	R ²	0.9307	0.8544
7-log input training (DEN/U/GR/AC/CNL/TH/K)	VAF	85.69	90.41
	MAE	0.0747	0.0692
	RMSE	0.2734	0.2631
	R ²	0.8749	0.9048
7-log input testing (DEN/U/GR/AC/CNL/TH/K)	VAF	86.30	89.18
	MAE	0.0677	0.0827
	RMSE	0.2602	0.1804
	R ²	0.8682	0.9099

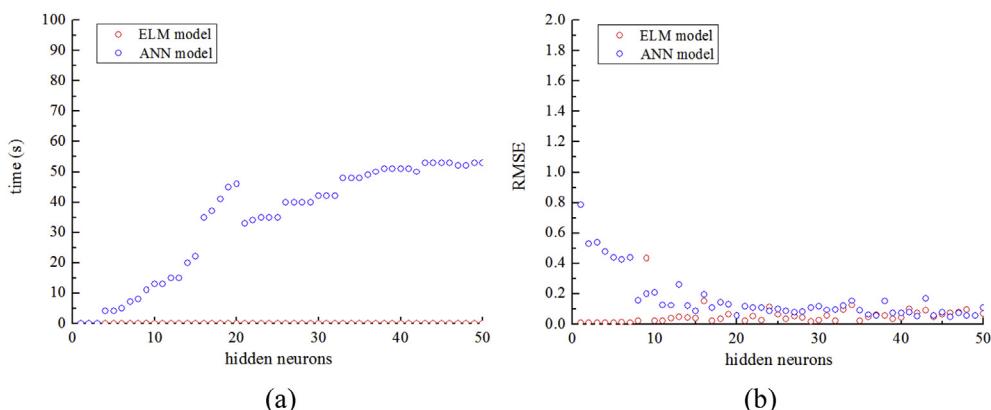


Fig. 7. (a) training time vs. hidden neurons of ANN and ELM models (b) testing RMSE vs. hidden neurons of ANN and ELM models.

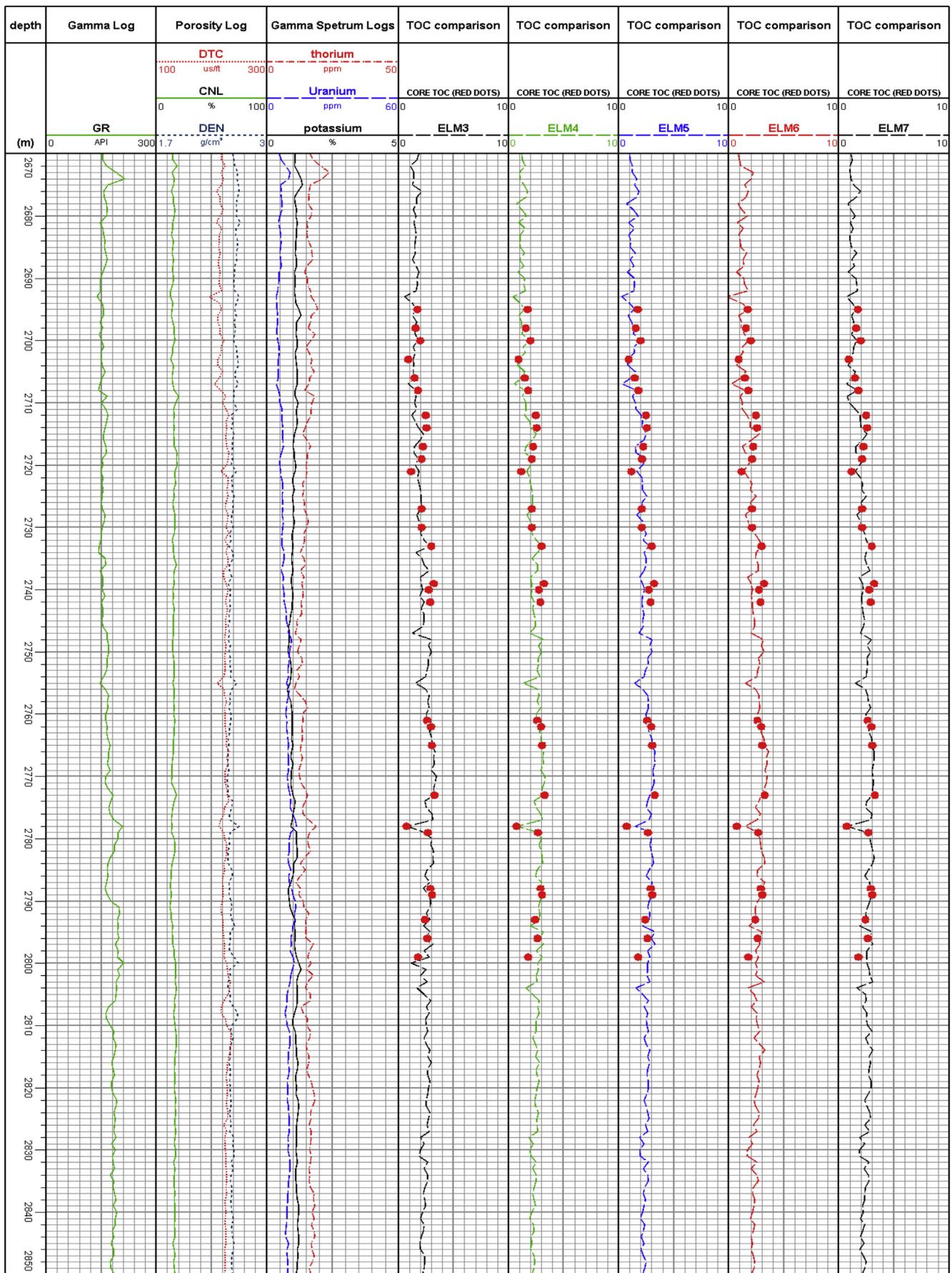


Fig. 8. Log data of Well J1 and comparison of the TOC prediction results with ELM models.

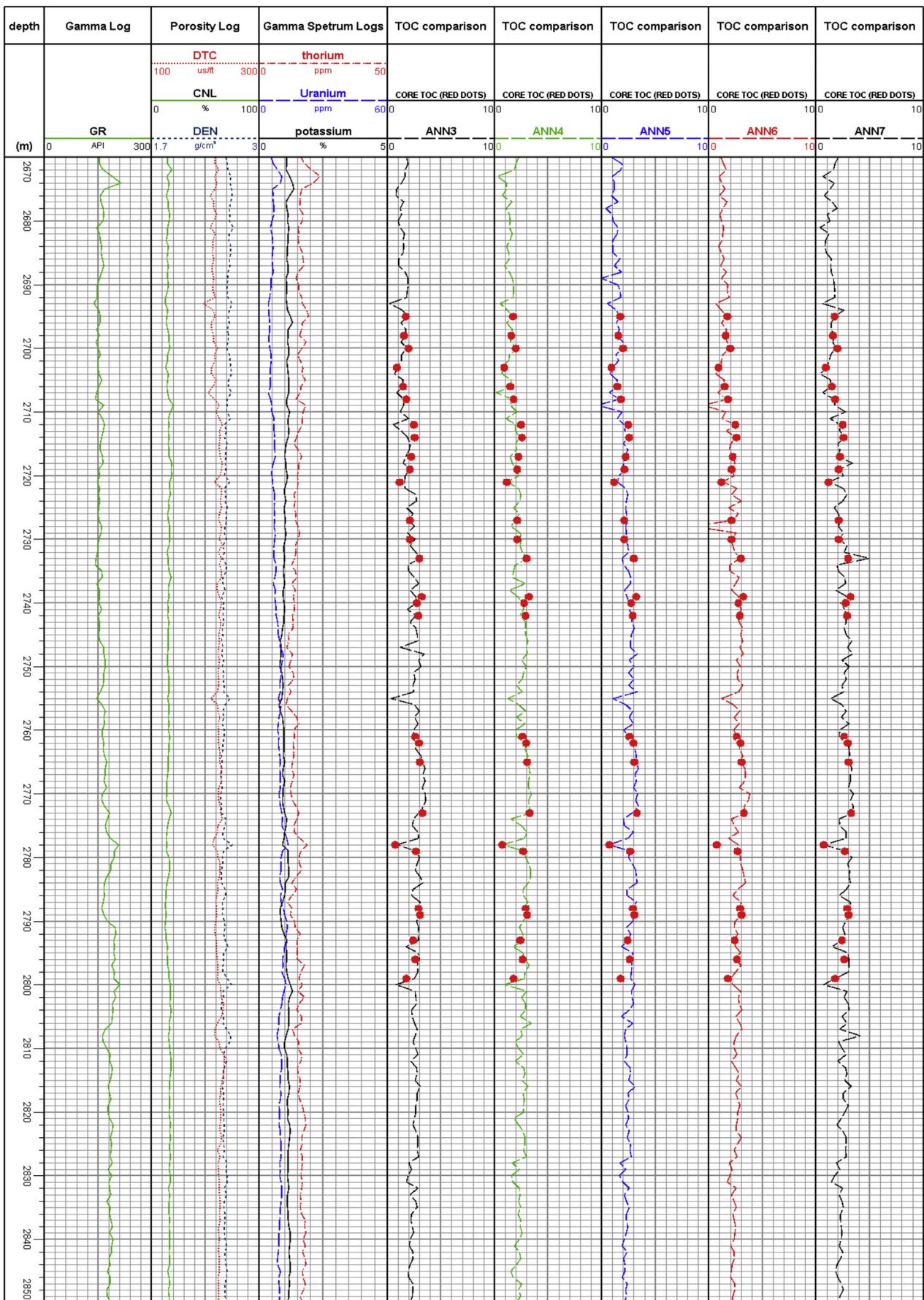


Fig. 9. Log data of Well J1 and comparison of the TOC prediction results with ANN models.

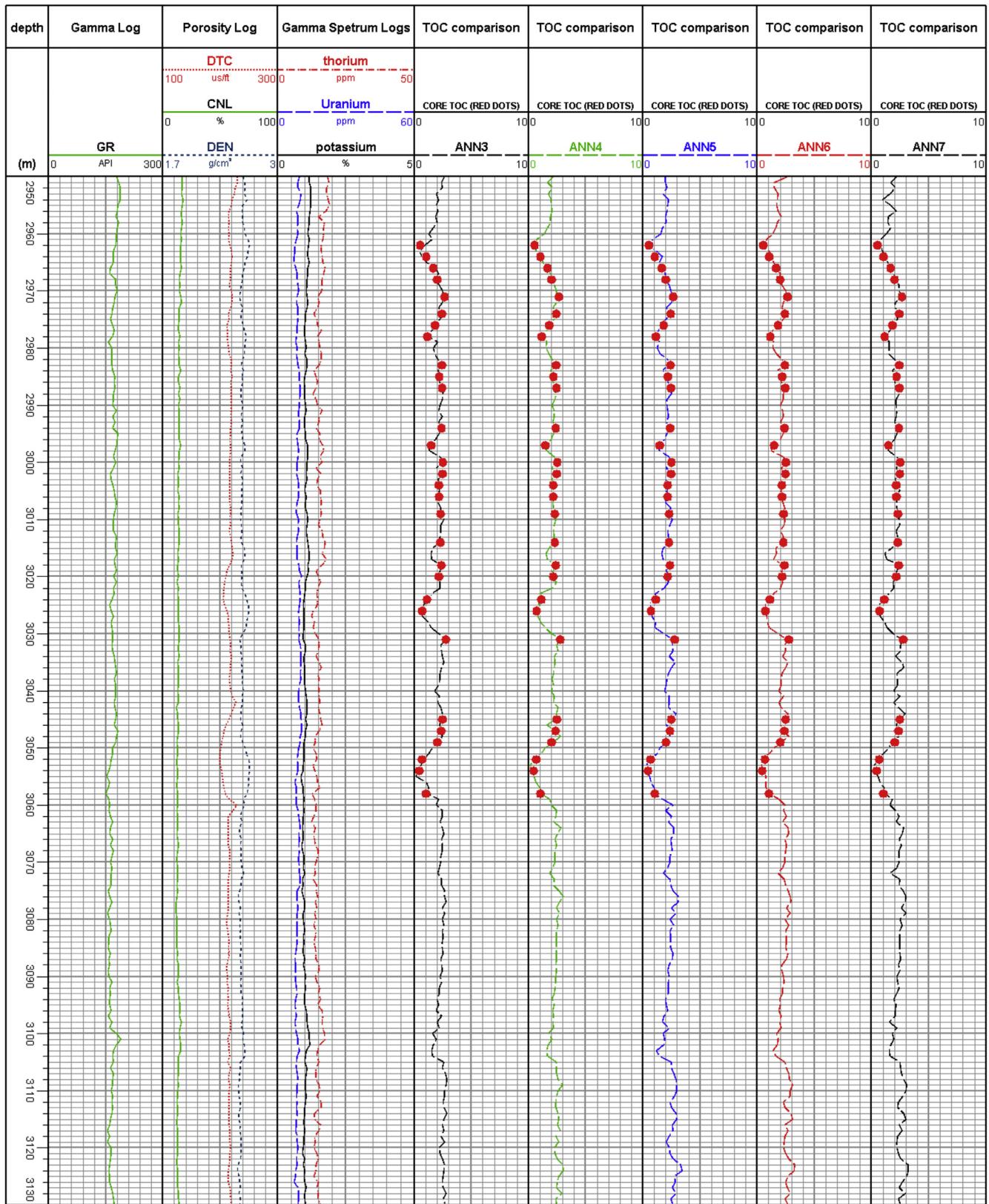


Fig. 10. Log data of Well J2 and comparison of the TOC prediction results with ELM models.

11). Moreover, only in the 2700 m–2720 m interval, both ANN and ELM model have minor errors. Comparison of the right five curves shows that both ANN model and ELM model are competent for TOC

prediction with different well logs. In light of the MAE, RMSE, R^2 and VAF results in Table 7, the ELM model provides slightly better results compared to the ANN model with different well log inputs.

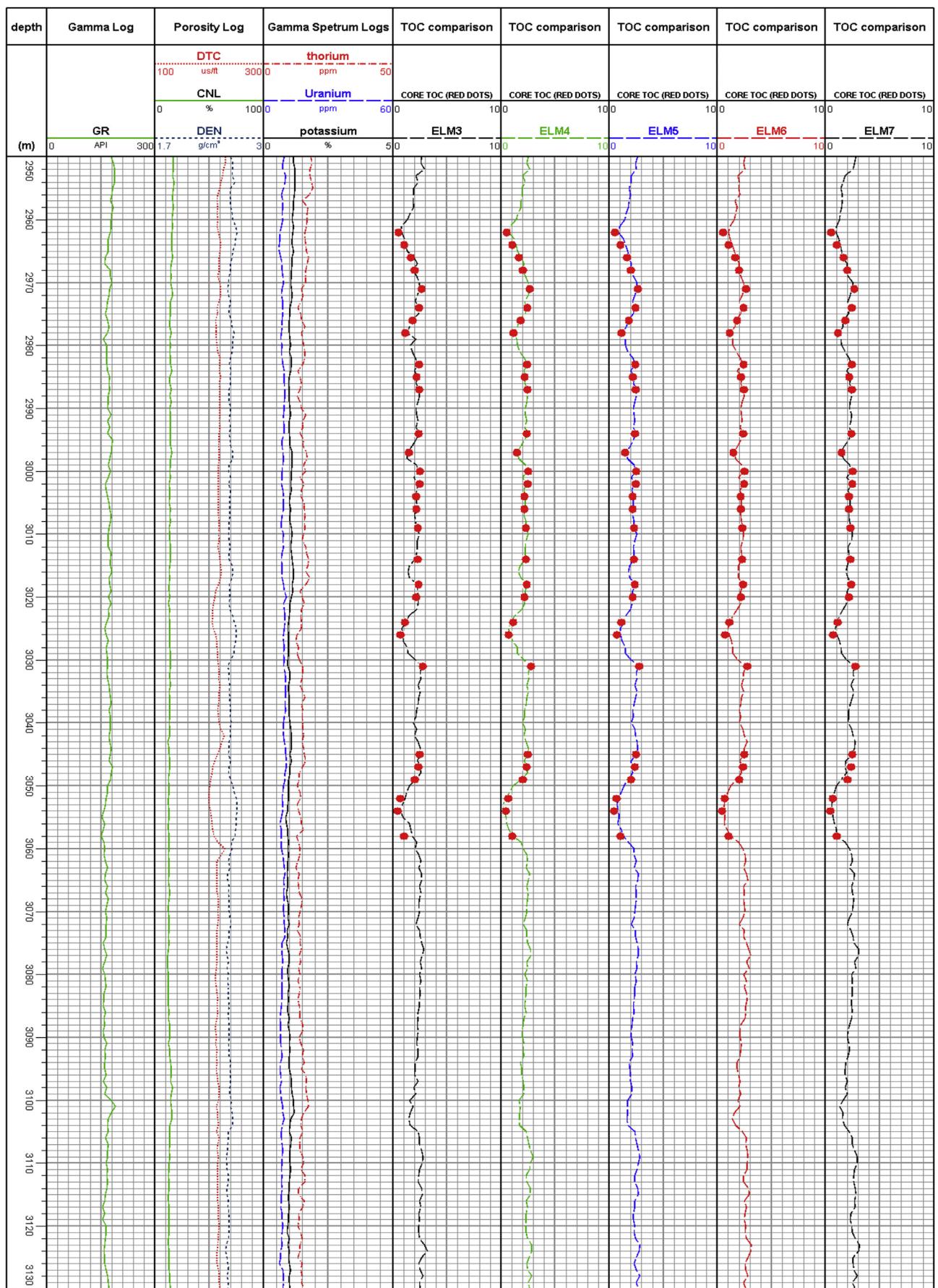


Fig. 11. Log data of Well J2 and comparison of the TOC prediction results with ELM models.

Table 7

The calculated RMSE and VAF indicators of ANN model and ELM model.

Number of logs used as inputs	RMSE	MAE (wt%)	VAF	R ²	Running time (s)
3-log input with ELM (DEN/U/GR)	0.3878	0.1504	82.18	88.34	5
3-log input with ANN (DEN/U/GR)	0.4067	0.1344	75.94	88.36	45
4-log input with ELM (DEN/U/GR/AC)	0.376	0.1381	88.06	91.62	6
4-log input with ANN (DEN/U/GR/AC)	0.3184	0.1014	83.60	93.21	50
5-log input with ELM (DEN/U/GR/AC/CNL)	0.3341	0.0916	81.02	93.12	5
5-log input with ANN (DEN/U/GR/AC/CNL)	0.3118	0.0972	84.21	93.19	52
6-log input with ELM (DEN/U/GR/AC/CNL/TH)	0.3033	0.0920	86.18	93.22	5
6-log input with ANN (DEN/U/GR/AC/CNL/TH)	0.3661	0.1340	80.93	90.60	57
7-log input with ELM (DEN/U/GR/AC/CNL/TH/K)	0.3742	0.1237	77.11	90.50	6
7-log input with ANN (DEN/U/GR/AC/CNL/TH/K)	0.3436	0.1181	83.18	90.60	58

Moreover, the running time of ELM model is well below that of ANN model and most ELM models can finish regression work within 6 s while the ANN models need much more time. Comparing the results of these two artificial networks reveals their strengths and weaknesses, which demonstrate that the ELM is more suitable for real-time TOC prediction task in terms of its fast nonlinear cost function.

5. Conclusions

Current methods for in situ TOC evaluation largely rely on qualitative responses and empirical formulas. Thus, the artificial intelligence techniques for TOC prediction have recently become powerful tools in shale gas exploration. This paper concentrates on a comparison of ELM and ANN artificial intelligence techniques for predicting TOC using conventional well logs. Data necessary for constructing the models were from the TOC core measurements conducted at Sichuan Basin, China. Linear regression and correlation coefficient techniques were performed to evaluate the simple relationships between well logs and TOC core measurements. The correlation analysis indicated that there are no apparent trends and also specified which data were less useful in TOC prediction, therefore allowing us to select the most informative data for model inputs. Error analysis using MAE, R², RMSE and VAF criteria showed that the ELM performed slightly better in both training and validation accuracy. However, the ANN model can also produce good results. According to the training time comparison, the ELM model is much faster than the ANN model, which indicates that ELM should be chosen as the better option if processing speed is important. The above advantages of ELM are quite beneficial for practical use. However, because core test measurements for training neural networks are developed for similar reservoirs with similar mineralogy and maturity, one must be cognizant of the large natural variation of shale gas formation properties and constituents and take care when using the correlations in different areas. In the future, it is necessary to develop an integrated approach to efficiently use the abundant wells with conventional log data to evaluate shale reservoirs at basin or regional scales for shale gas exploration and development.

Acknowledgments

This work was financially supported by the National Key Basic Research Program of China (2015CB251206), Fundamental Research Funds for the Central Universities (No. 15CX05053A, 15CX08011A, No. 15CX02009A) and Qingdao Science and Technology Project (No. 15-9-1-55-jch), National Natural Science Foundation of China (No. 61305075), the Natural Science Foundation of Shandong Province (No. ZR2013FQ004, No. ZR2013DM015), the Specialized Research Fund for the Doctoral Program of Higher Education of China (No. 20130133120014). We are also grateful to

two anonymous reviewers and handling editor for their critical comments and constructive suggestions.

References

- Akusok, A., Bjork, K., Miche, Y., Lendasse, A., 2015. High-performance extreme learning machines: a complete toolbox for big data applications. *IEEE Open Access* 3, 1011–1025.
- Alyssa, C., Susan, H., 2013. Accurate, direct total organic carbon (TOC) log from a new advanced geochemical spectroscopy tool: comparison with conventional approaches for TOC estimation. In: AAPG Annual Convention and Exhibition, Pittsburg, Pennsylvania, May 19–22.
- Amiri Bakhtiari, H., Telmandarrie, A., Shayesteh, M., Heidari Fard, M.H., Talebi, H., Shirband, Z., 2011. Estimating total organic carbon content and source rock evaluation, applying AlogR and neural network methods: Ahwaz and Marun oilfields, SW of Iran. *Pet. Sci. Technol.* 29 (16), 1691–1704.
- Autric, A., Dumesnil, P., 1985. Resistivity radioactivity and sonic transit time logs to evaluate the organic content of low permeability rocks. *Log Anal.* 26 (3), 37–45.
- Barak, O., Rigotti, M., Fusi, S., 2013. The sparseness of mixed selectivity neurons controls the generalization–discrimination trade-off. *J. Neurosci.* 33 (9), 3844–3856.
- Beers, R.F., 1945. Radioactivity and organic content of some Paleozoic Shales. *AAPG Bull.* 29 (1), 1–22.
- Cao, J.H., Yang, J.C., Wang, Y., Wang, D., Shi, Y.C., 2015. Extreme learning machine for reservoir parameter estimation in heterogeneous sandstone reservoir. *Math. Probl. Eng.* 1–10.
- Carpentier, B., Huc, A.Y., Bessereau, G., 1991. Wireline logging and source rocks estimation of organic carbon by the Carbolog method. *Log. Anal.* 32 (3), 279–297.
- Chandra, P.U., 2013. The hybrid of classification tree and extreme learning machine for permeability prediction in oil reservoir. *IJCSI Int. J. Comput. Sci. Issues* 10 (1), 52–60.
- Chong, K.K., Grieser, W.V., Passman, A., Tamayo, H.C., Modeland, N., Burke, B.E., 2010. A completions guide book to shale-play development: a review of successful approaches toward shale-play stimulation in the last two decades. SPE13387. In: Canadian Unconventional Resources and International Petroleum Conference, Calgary, Alberta, Canada, October 19–21.
- Chorowski, J., Wang, J., Zurada, J.M., 2014. Review and performance comparison of SVM-and ELM-based classifiers. *Neurocomputing* 128, 507–516.
- Decker, A.D., Hill, D.G., Wicks, D.E., 1993. Log-based gas content and resource estimates for the Antrim shale, Michigan Basin. SPE25910. In: Low Permeability Reservoir Symposium, Denver, Colorado, USA, April 26–28.
- Eliasmith, C., Stewart, T.C., Choo, X., Bekolay, T., Dewolf, T., Tang, Y., Rasmussen, D., 2012. A large-scale model of the functioning brain. *Science* 338, 1202–1205.
- Glorioso, J.C., Rattia, A.J., 2012. Unconventional reservoirs: basic petrophysical concepts for shale gas. SPE153004. In: SPE/EAGE European Unconventional Resources Conference and Exhibition, Vienna, Austria, March 20–22.
- Guo, L., Chen, J.F., Miao, Z.Y., 2009. Study and application of a new overlay method of the TOC content. *Nat. Gas. Geosci.* 20 (6), 951–956.
- Heidari, Z., Torres-Verdín, C., Preag, W.E., 2011. Quantitative method for estimating total organic carbon and porosity, and for diagnosing mineral constituents from well logs in shale-gas formations. In: SPWLA 52nd Annual Logging Symposium, Colorado Springs, Colorado, USA, May 14–18.
- Hu, H.T., Lu, S.F., Liu, C., Wang, W.M., Wang, M., Li, J.J., Shang, J.H., 2011. Models for calculating organic carbon content from logging information: comparison and analysis. *Acta Sedimentol. Sin.* 29 (6), 1199–1205.
- Huang, G.-B., 2014. An insight into extreme learning machines: random neurons, random features and kernels. *Cogn. Comput.* 6, 376–390.
- Huang, G.-B., Chen, L., 2007. Convex incremental extreme learning machine. *Neurocomputing* 70, 3056–3062.
- Huang, G.-B., Chen, L., 2008. Enhanced random search based incremental extreme learning machine. *Neurocomputing* 71, 3460–3468.
- Huang, Z.H., Williamson, M.A., 1996. Artificial neural network modelling as an aid to source rock characterization. *Mar. Pet. Geol.* 13 (2), 277–290.
- Huang, G.B., Zhu, Q.Y., Siew, C.K., 2004. Extreme learning Machine: a new learning

- scheme of feedforward neural networks. *Int. Jt. Conf. Neural Netw.* 2, 985–990.
- Huang, G.-B., Zhu, Q.-Y., Mao, K.Z., Siew, C.-K., Saratchandran, P., Sundararajan, N., 2006a. Can threshold networks Be trained directly? *IEEE Trans. Circuits Syst. II* 53, 187–191.
- Huang, G.-B., Chen, L., Siew, C.-K., 2006b. Universal approximation using incremental constructive feedforward networks with random hidden nodes. *IEEE Trans. Neural Netw.* 17 (4), 879–892.
- Huang, G.-B., Zhou, H., Ding, X., Zhang, R., 2012. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man, Cybern. Part B Cybern.* 42, 513–529.
- Igelnik, B., Pao, Y.-H., 1995. Stochastic choice of basis functions in adaptive function approximation and the functional-link net. *IEEE Trans. Neural Netw.* 6 (6), 1320–1329.
- Jarvie, D.M., Jarvie, B.M., Weldon, D., Maende, A., 2015. Geochemical assessment of in situ petroleum in unconventional resource systems. SPE178687. In: Unconventional Resources Technology Conference, San Antonio, Texas, July 20–22.
- Kamali, M.R., Mirshady, A.A., 2004. Total organic carbon content determined from wireline logs using DlogR and neuro fuzzy techniques. *Pet. Sci. Eng.* 45, 141–148.
- King, G., 2010. Thirty years of gas shale fracturing: what have we learned? *J. Pet. Technol.* 62 (11), 88–90.
- Li, M.-B., Huang, G.-B., Saratchandran, P., Sundararajan, N., 2005. Fully complex extreme learning machine. *Neurocomputing* 68, 306–314.
- Liang, N.-Y., Huang, G.-B., Saratchandran, P., Sundararajan, N., 2006. A fast and accurate on-line sequential learning algorithm for feedforward networks. *IEEE Trans. Neural Netw.* 17, 1411–1423.
- Mendelzon, J.D., Toksoz, M.N., 1985. Source rock characterization using multivariate analysis of log data. In: SPWLA 26th Annual Logging Symposium. Dallas, Texas, USA, June 17–20.
- Meyer, B.L., Nederlof, M.H., 1984. Identification of source rocks on wireline logs by density/resistivity and sonic transit/resistivity crossplots. *AAPG Bull.* 68, 121–129.
- Mortaza, A., Shahaboddin, S., Meisam, T., Por, L.Y., Yaser, N.L., 2016. The use of ELM-WT to predict energetic performance of a DI diesel engine running on diesel/biodiesel blends containing polymer waste. *Energy* 94, 443–456.
- Passey, Q.R., Creaney, S., Kulla, J.B., Moretti, F.J., Stroud, J.D., 1990. Practical model for organic richness from porosity and resistivity logs. *AAPG Bull.* 74, 1777–1794.
- Pemper, R.R., Han, X., Mendez, F.E., Jacobi, D., LeCompte, B., Bratovich, M., Feuerbacher, G., Bruner, M., Bilven, S., 2009. The direct measurement of carbon in wells containing oil and natural gas using a pulsed neutron mineralogy tool, SPE 124234. In: SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, USA, October 4–7.
- Peters, K.E., 1986. Guidelines for evaluating petroleum source rock using programmed pyrolysis. *AAPG Bull.* 70 (3), 318–329.
- Pope, C., Peters, B., Belton, T., Palish, T., 2009. Haynesville shale -one operator's approach to well completions in this evolving play. SPE 125079. In: Annual Technical Conference and Exhibition, New Orleans, Louisiana, USA, October 4–7.
- Rickman, R., Mullen, M., Petre, E., Grieser, B., Kundert, D., 2008. A practical use of shale petrophysics for stimulation design optimization: all shale plays are not clones of the Barnett shale. SPE115258. In: SPE Annual Technical Conference and Exhibition, Denver, Colorado, USA, September 21–24.
- Rigotti, M., Barak, O., Warden, M.R., Wang, X.-J., Daw, N.D., Miller, E.K., Fusi, S., 2013. The importance of mixed selectivity in complex cognitive tasks. *Nature* 497, 585–590.
- Schmoker, J., Hester, T., 1983. Organic carbon in bakken formation, United States portion of williston basin. *Am. Assoc. Pet. Geol. Bull.* 67, 2165–2174.
- Shadizadeh, S.R., Karimi, F., Zoveidavianpoor, M., 2010. Drilling stuck pipe prediction in Iranian oil fields: an artificial neural network approach. *Iran. J. Chem. Eng.* 7 (4), 29–41.
- Sondergeld, C.H., Newsham, K.E., Comisky, J.H., Rice, M.C., Rai, C.S., 2010. Petrophysical considerations in evaluating and producing shale gas resources. SPE131768. In: SPE Unconventional Gas Conference, Pittsburgh, Pennsylvania, USA, February 23–25.
- Sosulski, D.L., Bloom, M.L., Cutforth, T., Axel, R., Datta, S.R., 2011. Distinct representations of olfactory information in different cortical centres. *Nature* 472, 213–216.
- Sunday, O.O., Ali, S., Abdul, A.R., 2013. Extreme learning machines based model for predicting permeability of carbonate reservoir. *JDCTA Int. J. Digit. Content Technol. Appl.* 7 (1), 450–459.
- Tan, M.J., Song, X.D., Yang, X., Wu, Q.Z., 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.
- Wang, P.W., Chen, Z.H., Pang, X.Q., Hu, K.Z., Sun, M.L., 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.
- Yuan, C., Zhou, C.C., Hu, S., Cheng, X.Z., Dou, Y., 2014. Summary on well logging evaluation method of total organic carbon content in formation. *Prog. Geophys.* 29 (6), 2831–2837.
- Zhu, G.Y., Jin, Q., Zhang, L.Y., 2003. Using log information to analyze the geochemical characteristics of source rocks in jiyang depression. *Well Logging Technol.* 27 (2), 104–109.
- Zoveidavianpoor, M., Samsuri, A., Shadizadeh, S.R., 2012. Fuzzy logic in candidate-well selection for hydraulic fracturing in oil and gas wells: a critical review. *Int. J. Phys. Sci.* 7 (26), 4049–4060.
- Zoveidavianpoor, M., Samsuri, A., Shadizadeh, S.R., 2013a. Prediction of compressional wave velocity by an artificial neural network using some conventional well logs in a carbonate reservoir. *J. Geophys. Eng.* 10, 1–13.
- Zoveidavianpoor, M., Samsuri, A., Shadizadeh, S.R., 2013b. Adaptive neuro fuzzy inference system for compressional wave velocity prediction in a carbonate reservoir. *J. Appl. Geophys.* 89, 96–107.