Selective methods of TOC content estimation for organic-rich interbedded mudstone source rocks

Bang Zeng [a], Meijun Li [a,b,*], Jieqiong Zhu [c], Xin Wang [a], Yang Shi [c], Zhili Zhu [a], Hao Guo [c], Fangzheng Wang [a]

[a] State Key Laboratory of Petroleum Resources and Prospecting, College of Geosciences, China University of Petroleum (Beijing), Beijing, 102249, China
[b] Key Laboratory of Exploration Technologies for Oil and Gas Resources, Ministry of Education, College of Resources and Environment, Yangtze University, Wuhan, 430100, China
[c] Southern Oil Exploration and Development Company, PetroChina, Haikou, 570216, China

ARTICLE INFO

ABSTRACT

The total organic carbon content (TOC) is one of the key parameters for evaluating the hydrocarbon generation potential of source rocks. Petroleum geochemists and geologists usually use conventional logging curves to predict TOC content in order to reduce the errors caused by limited TOC content data. The ΔlogR method is the most widely used TOC prediction method. However, it is not applicable to organic-rich mudstone interlayers, due to the rapidly changing lithology and the difficulty in determining a baseline. Some improved ΔlogR methods and artificial neural network techniques have also been proposed by previous authors. In this paper, a 3D surface fitting technique based on biharmonic interpolation is proposed to optimize an improved ΔlogR method. A total of 76 samples were divided into pure mudstone and mudstone interlayers. The correlation coefficient (R) and the root mean square error (RMSE) between the measured TOC and the predicted TOC obtained using the improved ΔlogR method, the artificial neural network method, and the 3D surface fitting method were calculated. These methods were applied to assess the source rocks of the Eocene Liushagang Formation in the Fushan Depression, South China Sea. The results show that the 3D surface fitting method can effectively distinguish source rocks from non-source rocks and has a higher accuracy, which makes it the most suitable method for assessing the TOC content of organic-rich source rocks with alternating mudstone layers.

1. Introduction

Source rock assessment is a key step in hydrocarbon exploration, and the TOC content is one of the most important parameters for source rock assessment (He et al., 2019; Vega-Ortiz et al., 2020; Lai et al., 2020a). Generally, the most reliable method of obtaining TOC data is to measure core samples in the laboratory. However, the limited number of available core samples and high experimental costs hinder organic matter abundance assessment. Due to their continuous record of the petrophysical properties of rocks and easy accessibility, conventional well logs have been widely applied in the estimation of TOC contents (Schmoker, 1979; Fertl and Chilingar, 1988; Passey et al., 1990; Bodin et al., 2011; Wang et al., 2016; Rui et al., 2019; Li et al., 2020).

Simple linear fitting or empirical formulas were used for TOC content estimation before the 1990s (Schmoker, 1979; Fertl and Chilingar, 1988). For instance, the Schomoker density log based technique

(Schmoker, 1979) has been widely used in previous studies. Equation (1) is the Schomoker model for the Devonian shale.

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

where ρ_B represents the density of the stratum without organic matter (g/cm^3); and ρ represents the stratum bulk density (g/cm^3).

Although this method is convenient, the accuracy of the prediction results is not high. Passey et al. (1990) proposed the ΔlogR method to greatly improve the prediction accuracy, which is still used today. This method superimposes the resistivity and porosity curves (i.e., sonic, density, or neutron) of the non-source rock and calculates the TOC content based on the separation of the two curves, as is described in Eqs. (2) and (3).

$$\Delta logR = \log_{10}\left(\frac{R}{R_{baseline}}\right) + 0.02 \times (\Delta t - \Delta t_{baseline}) \tag{2}$$

$$TOC = \Delta logR \times 10^{(2.2297 - 0.1688 \times LOM)} \quad (3)$$

where $\Delta logR$ is the separation between the resistivity curve and the porosity curve; and R and Δt are the apparent resistivity ($\Omega \cdot m$) and the sonic transit time ($\mu s/ft$) of the target stratum, respectively. $R_{baseline}$ and $\Delta t_{baseline}$ are the resistivity ($\Omega \cdot m$) and sonic transit time ($\mu s/ft$) of the base stratum without organic matter, respectively. LOM is the level of maturation.

However, the accuracy of the Passey model largely depends on the selection of two key parameters, the baseline and LOM, which limits its widespread application. Unlike marine strata, lacustrine sedimentary successions are usually characterized by the extensive occurrence of alternating sandstone and organic-rich mudstone layers (Jiang et al., 2017). Due to these drastic changes in lithology, it is difficult to determine a baseline for the formation, and the logging response is readily influenced by the surrounding rocks. Therefore, the Passey model is not suitable for frequently interbedded sandstone and mudstone strata.

With the development of contemporary computer technology, neural network technology and machine deep learning have been successfully applied in TOC content estimation in previous studies (Tan et al., 2013; Shi et al., 2016; Mahmoud et al., 2017; Yu et al., 2017; Wang et al., 2016; Bai and Tan, 2020). These methods effectively solve the parameter selection problem. However, the research conducted on these
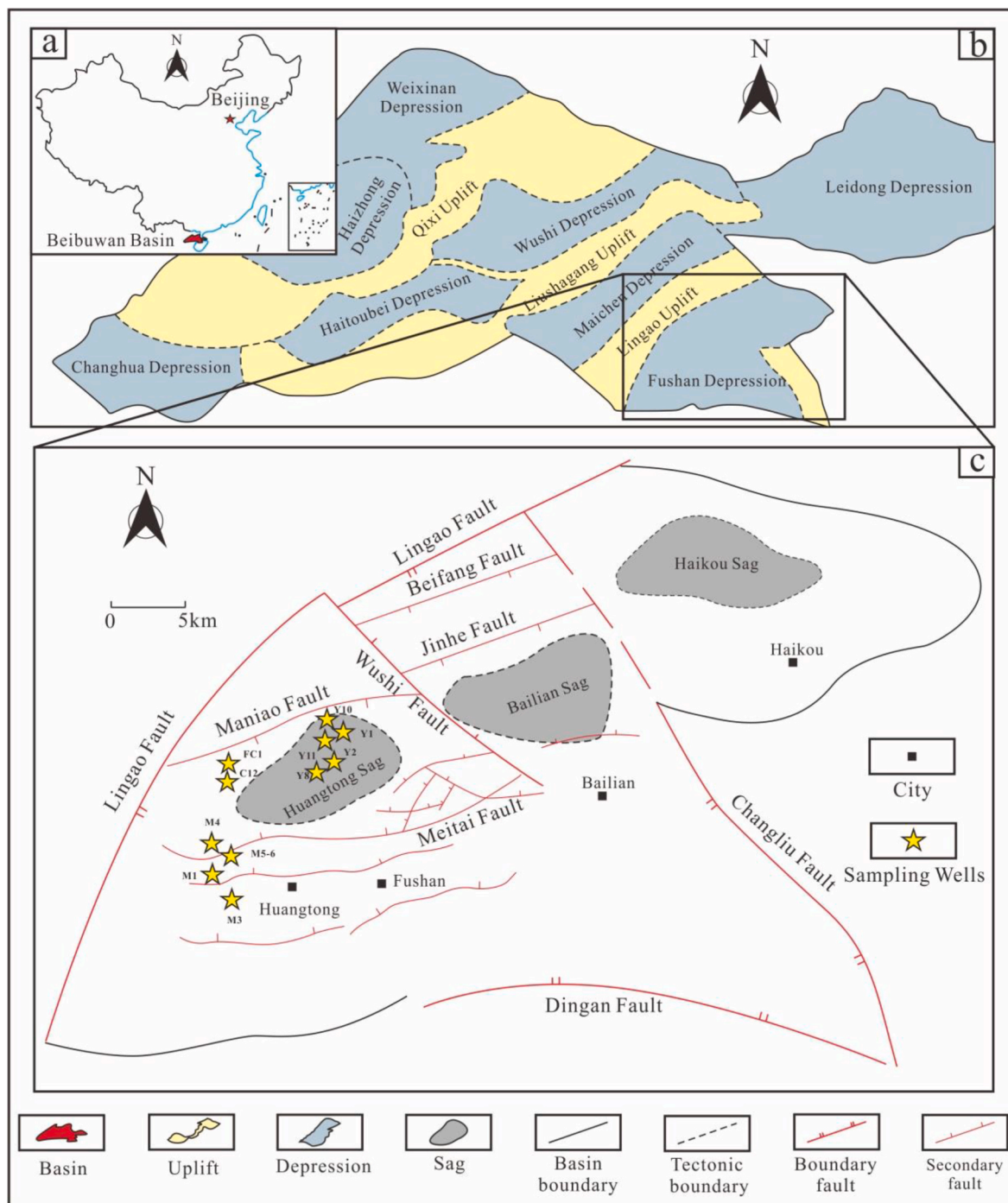

Fig. 1. (a) Location of the Beibuwan Basin. (b) Location of the Fushan Depression (modified after Li et al., 2019). (c) Location of the sampling wells in the Fushan Depression (modified after Liu et al., 2014).

methods focuses more on principles, algorithms, and accuracy and lack significant investigation of practical applications in petroleum geology. These mathematics-based methods are not closely combined with the geological background, and the applicable geological conditions are not definitely constrained, resulting in large prediction errors in many cases.

The Fushan Depression in the Beibuwan Basin is a typical lacustrine basin in southern China. The organic-rich mudstone interlayers occurring in the delta front, turbidite fans, and coastal shallow lakes have a good hydrocarbon generation potential and they form a complete source-reservoir-cap combination with the upper and lower sandstones. However, thus far, TOC prediction and source rock assessment of these mudstone interlayers has not been systematically carried out. Therefore, it is necessary to use conventional logging data to accurately predict the TOC content of this part of the source rock, which is of great significance

to hydrocarbon exploration in this region. In this study the Fushan Depression was used as a case study. The improved ΔlogR, artificial neural network, and three-dimensional (3D) surface fitting methods were used to determine the most effective method for TOC content prediction of source rocks with frequently alternating sandstone and organic-rich mudstone formations.

2. Geologic settings

The Beibuwan Basin is a typical Cenozoic rift basin locating in the South China Sea (Fig. 1a) (Li et al., 2014). The Fushan Depression is a secondary tectonic unit on the southeastern margin of the Beibuwan Basin (Fig. 1b), with a total area of 2920 km². This basin is surrounded by the Linggao, Changliu, and Dingan faults to the northwest, northeast,
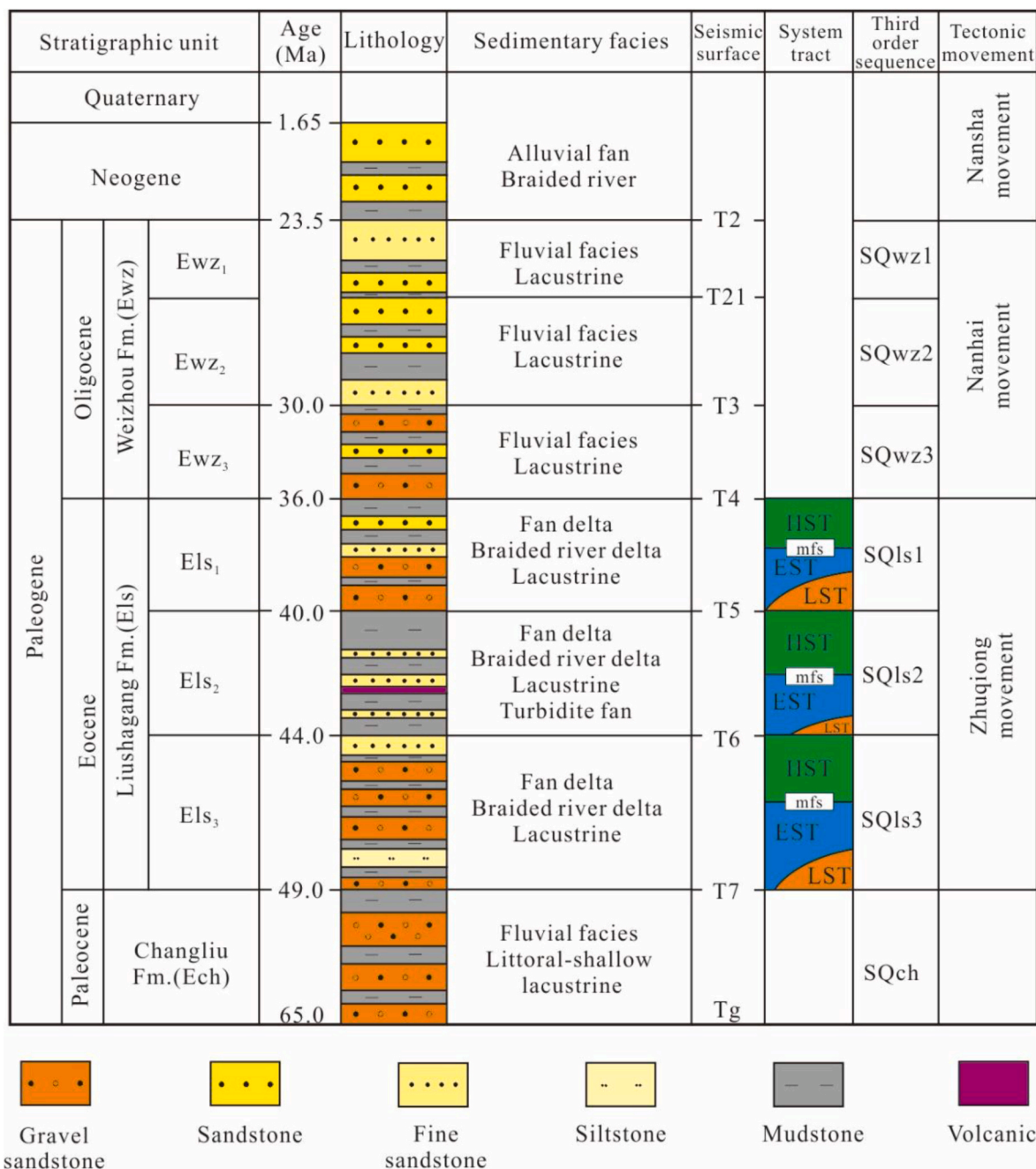

Fig. 2. Stratigraphic histogram of the Fushan Depression (modified after Gan et al., 2020). Els₁, Els₂, Els₃ are Member 1, 2, and 3 of the Eocne Liushagang Formation; Ewz₁, Ewz₂, Ewz₃ are Member 1, 2, and 3 of the Eocene Weizhou Formation.

and south, respectively (Fig. 1a). The sediments in the depression were supplied from three positive tectonic units: the northwestern Lingao Uplift, the northeastern Yunlong Uplift, and the southern Hainan basement (Cao et al., 2016).

Based on the major tectonic movements and unconformities, the Paleogene strata in the Fushan Depression can be divided into three formations: the Changliu Formation (Ech), the Liushagang Formation (Els), and the Weizhou Formation (Ewz) (Fig. 2). The Liushagang Formation contains the most important oil-generating strata in the Fushan Depression. It is composed of braided river delta, fan delta, turbidite fan, and lacustrine facies and deposits of dark gray mudstone frequently intercalated with sandstone and gravel sandstone (Gan et al., 2020; Liu et al., 2014, 2015). Therefore, this study attempts to choose an optimal method for accurately predicting the TOC content of this type of the source rock based on conventional well logs.

3. Materials and methods

The methods used to predict the TOC in this study include the improved ΔlogR, artificial neural network, and 3D surface fitting. TOC data was obtained by a LECO CS-230 carbon analyzer. EXCEL 2019, SPSS 26 and MATLAB 2016a were used for data acquisition and chart drawing. Data processing mainly includes regression and fitting. The correlation coefficient (R) and the root mean square error (RMSE) were used as evaluation criteria.

3.1. Materials

A total of 76 core samples from 11 wells were collected from the Liushagang Formation in the Fushan Depression in order to build a TOC prediction model. 8, 15, 16, and 37 samples were collected from the delta front subfacies, fan delta front subfacies, turbidite fan subfacies, and semi-deep lake subfacies, respectively (Table 1). A total of 136 cutting samples were collected to evaluate the TOC prediction model. The logging response values of the depth-corrected core samples are presented in Table 1. Note that the density of three samples is less than 2.0 g/cm^3, which is mainly due to the sampling location that is close to the bottom of the well, where the logging curve may be slightly inaccurate. In addition, conventional well logs for 11 sampling wells were also collected for this study. The locations of the sampling wells are shown in Fig. 1c.

3.2. Methods

3.2.1. Laboratory experiments

All of the mudstone samples used for the total organic carbon (TOC) analysis were crushed and ground to a diameter of less than 0.2 mm (80 mesh). Then, the carbonates were removed using dilute hydrochloric acid (1.5 mol/L). Finally, all of the samples were washed with deionized water to remove any possible residual contamination. The TOC contents were measured using a LECO CS-230 carbon analyzer.

3.2.2. Improved ΔlogR method

In order to solve the problem of the difficulty to determine baseline, Huo et al. (2011) and Lai et al. (2018) proposed an improved ΔlogR method. The principle of this method is similar to that of the Passey method, and the details of the procedure are as follows. First, Eq. (2) is transformed into Eq. (4).

$$\Delta logR = [\log_{10}(R) + (0.02 \times \Delta t)] - [\log_{10}(R_{baseline}) + (0.02 \times \Delta t_{baseline})]. \quad (4)$$

Second, a new parameter called ΔlogR′ is defined using Eq. (5), and then, Eq. (4) is transformed into Eq. (6).

$$\Delta logR' = \log_{10}(R) + (0.02 \times \Delta t) \quad (5)$$

$$\Delta logR = \Delta logR' - \Delta logR'_{baseline} \quad (6)$$

Third, Eq. (5) is transformed into Eq. (7). The value of ΔlogR′ is the intercept of the straight line with a gradient of −0.02 on the plot of Δt versus $\log_{10}(R)$. ΔlogR'$_{baseline}$ is the minimum intercept of such a straight line.

$$\log_{10}(R) = -0.02 \times \Delta t + \Delta logR' \quad (7)$$

Finally, the numerical relationship between the measured TOC content and ΔlogR is obtained using Eq. (3) (if the data contain non-source rocks) or by mathematical fitting (if the data do not contain non-source rocks).

3.2.3. The principle of the artificial neural network (ANN)

An artificial neural network is a mathematical model that simulates the behavior of animal neural networks and processes information by adjusting the relationships between a large number of internal nodes (Lim, 2003). It consists of different algorithms and structures. In this study, an artificial neural network (ANN) was used to estimate the TOC content. All of the data processing was carried out using SPSS 26 and MATLAB 2016a.

A back propagation artificial neural network (BPANN) is a multi-layer feedforward network trained using an error back propagation algorithm (Kumar et al., 2012). BPANNs have been widely used in data classification and prediction studies (Fath et al., 2020; Urang et al., 2020; Lim, 2003). A BPANN consists of one input layer, one hidden layer, and one output layer (Fig. 3a). Each layer has one or more interconnected neurons. The process of the BPANN is to assign weights and thresholds to the hidden layer neurons, to adjust the weights and thresholds according to the back propagation error calculated by the error function, and to complete the training until the expected output is obtained (Fig. 3a).

Strictly speaking, the multilayer perceptron neural network is a back propagation artificial neural network with one or more hidden layers. In this study, in order to distinguish between artificial neural networks with different structures, a neural network with one hidden layer is called a back propagation artificial neural network (BPANN), and a neural network with two hidden layers is called a multilayer perceptron neural network (MLPNN). The MLPNN includes four layers: the input layer, hidden layer one, hidden layer two, and the output layer (Fig. 3b). The specific training process is the same as that for the BPANN.

3.2.4. 3D surface fitting

3D surface fitting is the use of interpolation to fit discrete points in space onto a surface in order to predict unknown data using known data. Commonly used interpolation methods include the nearest neighbor, linear, and biharmonic methods. Though the former two methods are relatively faster, the surfaces produced by biharmonic interpolation are smoother and are more conducive to obtaining accurate predictions.

The biharmonic interpolation technique was proposed by Sandwell (1987) and has been widely applied in meteorology, marine currents, and manufacturing technology (Adhikari and Marshall, 2013; Sandwell and Smith, 2009; Chen and Wang, 2018). Discrete points in space are interpolated by the biharmonic surface using the Green function (Chen and Wang, 2018). The surface has the characteristic of minimum curvature. In this study, all of the steps of the 3D surface fitting were implemented in MATLAB 2016a.

4. Results

4.1. Evaluation criteria

In order to consider the fitting effect and accuracy, the correlation coefficient (R) and the root mean square error (RMSE) were used as evaluation parameters (Fath et al., 2020). These parameters can be calculated using Eqs. (8) and (9).

Table 1
Types and TOC content of the core samples and their logging response values.

No.	Well	Depth	TOC	AC	GR	RILD	RILM	CNL	DEN	Types
		(m)	(wt%)	(μs/ft)	(API)	(Ω·m)	(Ω·m)	(%)	(g/cm³)	
1	M1	2827.1	1.64	79.73	150.18	8.79	10.72	19.26	2.41	semi-deep lake mudstone
2	M1	3108.4	2.89	101.40	113.12	9.00	6.46	20.17	1.65	delta front mudstone
3	M1	3112.9	1.12	77.37	63.65	28.46	19.75	12.96	2.32	delta front mudstone
4	M1	3169.5	1.51	123.58	116.87	2.22	2.06	60.51	1.30	delta front mudstone
5	M3	2741.5	0.84	75.81	137.99	7.05	6.70	23.24	2.47	delta front mudstone
6	M3	2743.3	0.05	72.59	132.46	7.79	7.85	18.14	2.53	delta front mudstone
7	M4	3218.5	1.15	74.33	166.08	9.31	12.30	20.77	2.18	semi-deep lake mudstone
8	M4	3221.0	1.12	88.38	156.49	5.75	6.02	45.79	2.20	semi-deep lake mudstone
9	M4	3222.5	1.18	91.95	168.65	7.08	7.60	37.41	2.16	semi-deep lake mudstone
10	M4	3223.5	1.57	75.03	101.91	11.62	17.39	13.60	2.38	semi-deep lake mudstone
11	M4	3228.5	1.16	75.47	119.48	9.56	9.96	23.07	2.41	semi-deep lake mudstone
12	M5-6	2347.0	0.57	76.24	125.98	5.27	6.09	39.13	2.477	delta front mudstone
13	M5-6	2519.0	1.68	85.87	132.12	5.49	4.81	45.65	2.20	delta front mudstone
14	M5-6	2624.0	2.82	75.73	146.76	14.10	11.19	47.82	2.51	delta front mudstone
15	M5-6	2671.0	1.47	61.74	156.38	23.15	21.23	43.17	2.43	semi-deep lake mudstone
16	M5-6	2693.0	1.79	73.82	147.31	10.31	10.11	52.89	2.33	semi-deep lake mudstone
17	M5-6	2785.0	1.84	72.05	143.02	13.23	11.30	49.46	2.55	semi-deep lake mudstone
18	M5-6	2918.5	1.92	74.09	159.41	10.71	11.41	40.48	2.49	semi-deep lake mudstone
19	M5-6	3009.5	1.68	74.80	135.34	12.76	10.92	39.23	2.34	semi-deep lake mudstone
20	M5-6	3080.5	1.72	80.35	147.82	6.53	6.20	37.85	2.09	semi-deep lake mudstone
21	M5-6	3147.5	1.77	69.97	131.45	18.97	14.89	55.87	2.52	semi-deep lake mudstone
22	M5-6	3318.5	1.69	65.51	141.60	28.89	17.21	56.01	2.35	semi-deep lake mudstone
23	M5-6	3331.5	1.61	80.72	150.42	7.97	7.62	41.97	2.59	semi-deep lake mudstone
24	M5-6	3395.5	1.69	80.62	141.11	9.43	8.93	36.90	2.46	semi-deep lake mudstone
25	M5-6	3504.0	1.13	81.74	121.39	10.55	7.69	49.71	2.66	semi-deep lake mudstone
26	CH2	2493.5	1.24	80.82	108.49	6.22	6.45	22.29	2.54	fan delta front mudstone
27	CH2	2621.2	0.83	94.16	125.87	5.95	6.31	21.21	2.38	fan delta front mudstone
28	CH2	2622.5	1.26	106.25	120.58	4.69	4.81	23.90	2.34	fan delta front mudstone
29	CH12	2544.3	0.82	72.45	98.59	9.49	9.30	20.20	2.54	fan delta front mudstone
30	CH12	2545.2	1.38	81.51	106.12	7.41	7.25	25.75	2.43	fan delta front mudstone
31	CH12	3473.5	1.86	75.52	109.20	19.18	19.50	18.26	2.62	fan delta front mudstone
32	CH12	3476.0	1.30	75.26	109.37	13.07	14.13	18.97	2.51	fan delta front mudstone
33	CH12	3512.2	1.49	69.10	94.26	20.10	19.52	13.92	2.63	fan delta front mudstone
34	CH12	3513.6	1.23	74.70	127.62	12.30	12.29	24.26	2.60	fan delta front mudstone
35	FC1	2513.0	0.83	93.08	124.04	4.68	3.99	36.77	2.17	fan delta front mudstone
36	FC1	2536.0	0.52	79.63	118.34	6.68	5.89	33.42	2.24	fan delta front mudstone
37	FC1	2694.5	1.14	74.06	135.66	7.14	7.29	23.43	2.29	fan delta front mudstone
38	FC1	3186.5	1.83	56.01	108.46	24.19	16.87	15.73	2.35	fan delta front mudstone
39	FC1	3190.3	1.80	73.59	136.55	8.70	6.03	40.12	2.12	fan delta front mudstone
40	FC1	3373.0	2.03	88.32	118.04	15.96	11.70	36.40	2.38	fan delta front mudstone
41	Y1	2994.0	1.05	77.93	131.55	7.12	7.90	23.67	2.31	semi-deep lake mudstone
42	Y1	2996.0	1.02	94.23	155.89	2.97	3.39	42.17	2.22	semi-deep lake mudstone
43	Y1	3037.0	0.92	96.62	77.92	8.10	13.10	16.99	1.92	turbidite mudstone
44	Y1	3800.1	1.39	88.21	199.56	11.54	10.25	23.32	2.05	turbidite mudstone
45	Y2	2946.8	0.98	75.85	121.81	8.85	8.44	14.96	2.49	turbidite mudstone
46	Y2	2951.8	0.86	74.70	117.07	11.77	11.17	18.61	2.47	turbidite mudstone
47	Y8	2898.3	0.90	78.30	96.71	4.75	5.36	17.74	2.46	semi-deep lake mudstone
48	Y8	2899.4	0.96	89.64	152.43	4.71	4.10	42.95	2.53	semi-deep lake mudstone
49	Y8	2903.0	0.91	73.14	81.83	6.26	6.65	15.27	2.50	semi-deep lake mudstone
50	Y8	2904.0	0.85	80.20	148.53	5.88	5.07	29.07	2.53	semi-deep lake mudstone
51	Y8	2905.0	0.90	85.49	160.53	4.96	4.53	31.09	2.56	semi-deep lake mudstone
52	Y8	3056.3	0.79	76.51	48.05	10.93	10.76	9.00	2.52	turbidite mudstone
53	Y8	3057.3	0.77	78.30	153.32	8.35	6.67	26.39	2.43	turbidite mudstone
54	Y8	3061.8	1.06	73.77	146.80	7.33	7.38	20.40	2.62	turbidite mudstone
55	Y8	3068.8	0.58	63.63	97.16	3.97	4.62	15.15	2.58	turbidite mudstone
56	Y10	3013.8	1.24	89.00	166.44	7.87	5.83	24.91	2.31	turbidite mudstone
57	Y10	3278.2	1.87	73.11	139.88	10.51	9.38	20.86	2.51	turbidite mudstone
58	Y10	3540.5	1.28	82.83	162.38	11.81	8.35	18.84	2.49	turbidite mudstone
59	Y11	3291.5	1.51	68.48	122.56	15.67	10.26	22.41	2.54	semi-deep lake mudstone
60	Y11	3548.5	1.61	70.01	98.06	18.00	14.82	15.38	2.61	semi-deep lake mudstone
61	Y11	3552.0	1.73	71.45	118.19	21.83	16.84	21.67	2.40	semi-deep lake mudstone
62	Y11	3567.4	2.00	74.17	136.29	16.15	13.11	19.05	2.41	semi-deep lake mudstone
63	Y11	3644.3	1.85	71.17	135.86	12.04	11.28	25.17	2.42	turbidite mudstone
64	Y11	3654.9	1.74	84.61	127.33	17.19	15.77	27.48	2.26	turbidite mudstone
65	Y11	3659.5	1.62	63.70	89.11	23.01	15.16	10.31	2.42	turbidite mudstone
66	Y11	3672.6	1.36	80.04	160.75	10.39	10.98	33.42	2.62	semi-deep lake mudstone
67	Y11	3674.3	1.52	82.64	157.78	13.23	12.34	24.30	2.69	semi-deep lake mudstone
68	Y11	3695.0	1.30	81.68	177.23	10.27	9.61	39.56	2.57	semi-deep lake mudstone
69	Y11	3705.5	1.38	65.62	109.71	15.93	14.01	20.91	2.41	semi-deep lake mudstone
70	Y11	3707.8	1.46	75.02	129.76	11.39	11.87	34.11	2.14	semi-deep lake mudstone
71	Y11	3710.2	1.25	76.76	125.78	10.54	9.63	36.64	2.56	semi-deep lake mudstone
72	Y11	3716.0	1.87	81.83	124.27	13.27	8.70	29.96	2.25	semi-deep lake mudstone
73	Y11	3722.0	1.27	82.34	167.01	9.90	8.69	43.03	2.51	turbidite mudstone

Table 1 (*continued*)

No.	Well	Depth	TOC	AC	GR	RILD	RILM	CNL	DEN	Types
		(m)	(wt%)	(μs/ft)	(API)	(Ω·m)	(Ω·m)	(%)	(g/cm³)	
74	Y11	3724.5	1.94	80.93	149.82	9.19	9.67	38.12	2.38	turbidite mudstone
75	Y11	3727.0	1.28	84.98	172.82	9.85	10.20	33.80	2.66	semi-deep lake mudstone
76	Y11	3731.0	1.34	65.02	112.58	17.95	15.90	17.46	2.49	semi-deep lake mudstone
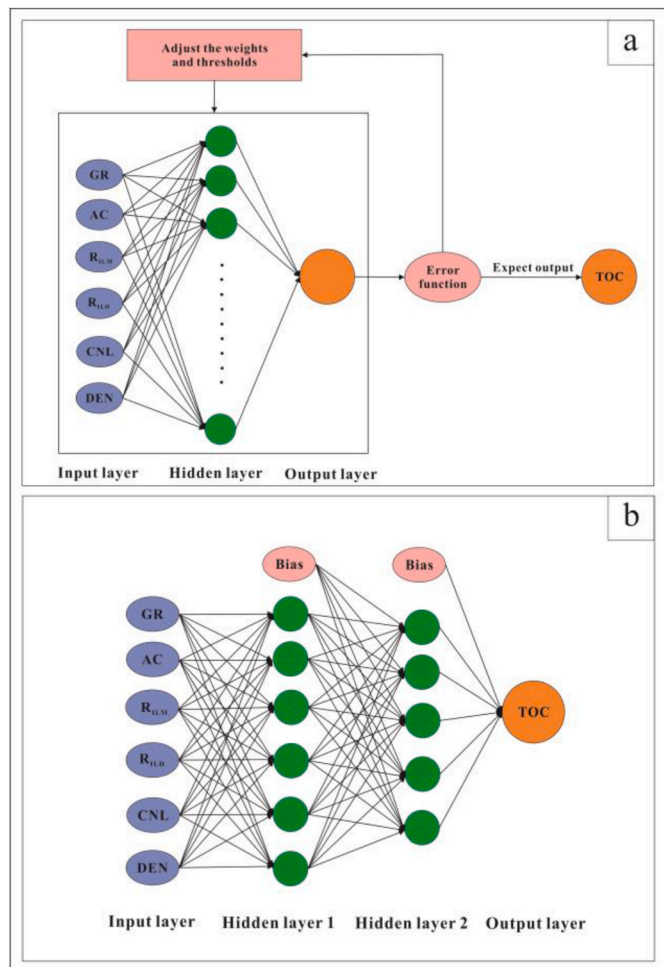
Fig. 3. (a) Diagram of the BPANN algorithm structure. (b) Diagram of the MLPNN algorithm structure.

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n} \left(x_{iexp} - x_{ipre} \right)^2}{\sum_{i=1}^{n} \left(x_{ipre} - \overline{x} \right)^2}} \qquad (8)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(x_{iexp} - x_{ipre} \right)^2} \qquad (9)$$

where x_{iexp} is the target value; x_{ipre} is the predicted value; $\overline{x}$ is the average experimental value; and n is the number of samples.

In addition, all source rock samples were divided into two groups: pure mudstone and mudstone interlayers, respectively. Semi-deep lake mudstone was considered as pure mudstone, where the ratio of the mudstone to the formation thickness is greater than 90%. Delta front mudstone, fan delta front mudstone and turbidite mudstone with a ratio of mudstone to formation thickness of 50–70% were classified as mudstone interlayers.

4.2. Initial data analysis

Previous studies have demonstrated that the organic-sensitive logging parameters mainly include the sonic transit time (AC), natural gamma (GR), deep induction resistivity (R_{ILD}), medium induction resistivity (R_{ILM}), neutron (CNL), and density (DEN), which change as the TOC content changes (Schmoker, 1979; Passey et al., 1990; He et al., 2016; Wang et al., 2016). Generally, minerals, fluids, organic matter, and pressure are the factors that affect the logging response characteristics. In this study, correlation coefficients (R) were used to express the relationship between the TOC and the different log values (Mahmoud et al., 2017).

The results show that for all of the samples, the TOC is positively correlated with GR, CNL, R_{ILM}, and R_{ILD}, is weakly negatively correlated with DEN, and is not correlated with AC. For the pure mudstone, R_{ILD}, R_{ILM}, and AC are more strongly correlated with TOC, while for the mudstone interlayers, all of the well logging correlations are much lower (Fig. 4).

4.3. Performance of the improved ΔlogR method

The Liushagang Formation widely occurs throughout the Fushan Depression, with a stable formation thickness. Although the burial depth varies greatly in the different structural blocks, the sediment supplies and sedimentary backgrounds are similar, and the logging response characteristic value has a small variation range. Therefore, a unified TOC content estimation model can be established for the Liushagang Formation (Lai et al., 2018, 2020a, 2020b).

Overall, the TOC contents of all of the core samples are greater than 0.5%. Mathematical fitting was used to calculate the numerical relationship between the measured TOC content and ΔlogR. The ΔlogR′ obtained from the plot is 1.87 (Fig. 5a), and the quadratic polynomial regression equation for measured TOC content value and ΔlogR is

$$TOC = -1.9908 \times (\Delta logR)^2 + 4.3282 \times \Delta logR - 0.6322 \qquad (10)$$

Eq. (10) is the TOC prediction model obtained using the improved ΔlogR method. The R and RMSE values are 0.5315 and 0.3965,
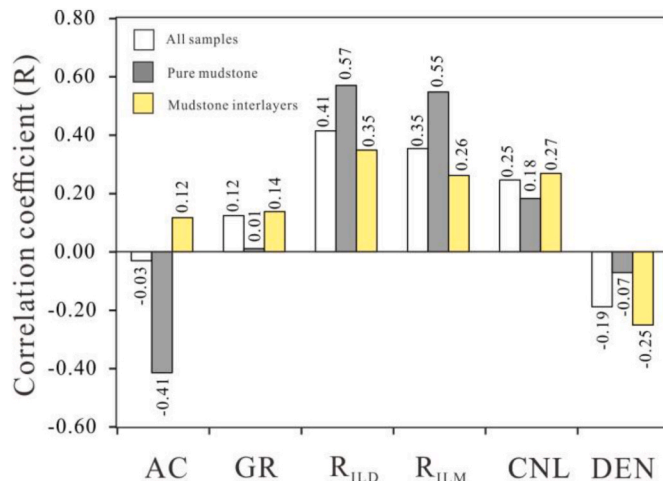

Fig. 4. Correlations between the different well logs and the measured TOC.

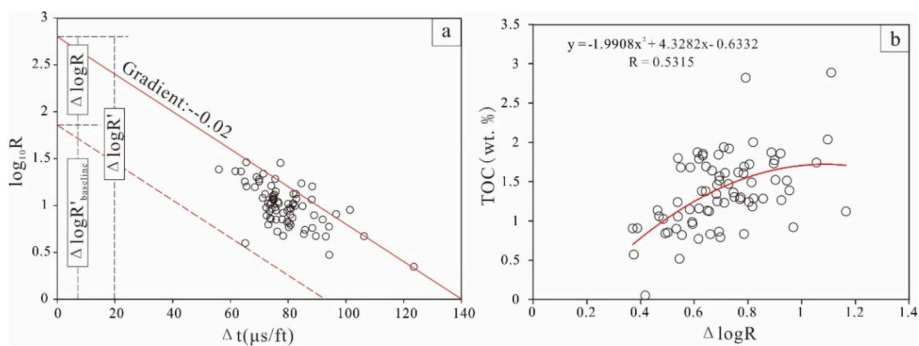**Fig. 5.** (a) Plot of Δt versus $\log_{10}(R)$. (b) The numerical relationship between the measured TOC content and ΔlogR.
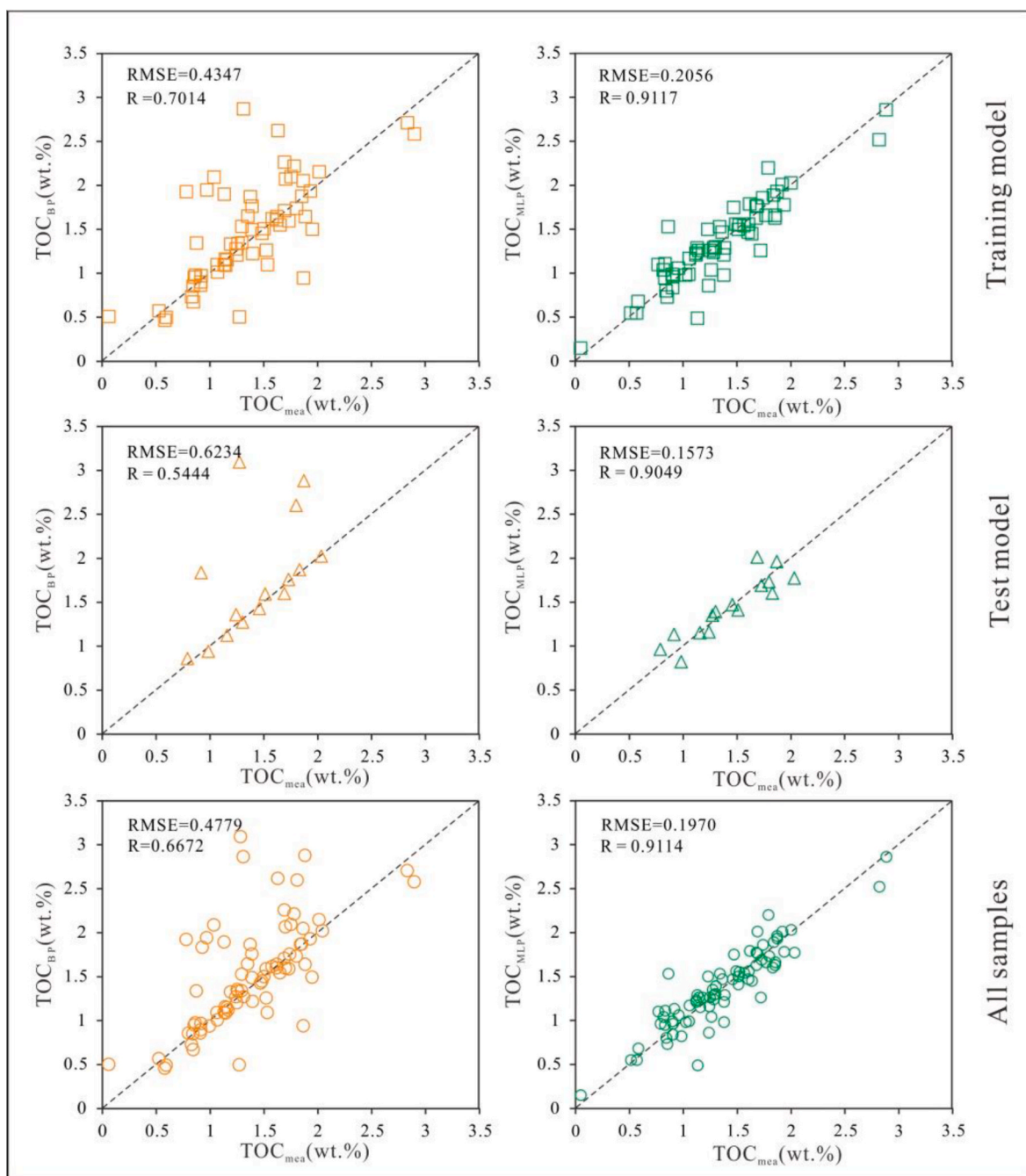

Fig. 6. Comparison of the TOCs calculated using the BPANN and MLPNN. The orange symbols represent the BPANN, and the green symbols represent the MLPNN.

respectively (Fig. 5b).

4.4. Performance of the artificial neural network (ANN)

The selection of the training samples and the input layer is the key issue for artificial neural networks. In this study, 80% of the core TOC data were selected for the training, and 20% of the data were randomly selected for testing using the Bernoulli distribution. Since there was no strong functional relationship with TOC, all of the well logs were considered as the input layer.

The number of hidden layer neurons was determined through extensive experiments. It was found that the prediction accuracy was the highest when 10 neurons were set in the BPANN hidden layer, 6 neurons were set in MLPNN hidden layer one, and 5 neurons were set in MLPNN hidden layer two. For the BPANN training model, the test model, and all of the samples, the R values are 0.7014, 0.5444, and 0.6672 and the RMSE values are 0.4347, 0.6234, and 0.4779, respectively. For the MLPNN training model, the test model, and all of the samples the R values are 0.9117, 0.9049, and 0.9114 and the RMSE values are 0.2056, 0.1573, and 0.1970, respectively (Fig. 6). Since individual samples with abnormal density may affect the accuracy of the model, we removed the samples with a density less than 2.0 g/cm^3 and rebuilt the neural network model. The results show that those samples did not affect the accuracy of the model, which is mainly due to the weak correlation between TOC and DEN. Therefore, we preserved these samples in the neural network model.

4.5. Performance of the 3D surface fitting method

As can be seen from Fig. 5, the prediction accuracy of the improved ΔlogR method is not high. The first reason for this is that the logging response value may be seriously affected by the surrounding rocks due to the frequently changing lithology. Furthermore, the content and type of clay minerals change with increasing maturity, affecting the pores and structure of the mudstone, changing the pore accommodation space, and thus, affecting the adsorption equilibrium state of the organic matter (Lu et al., 2013), causing the poor mathematical relationship between ΔlogR and the measured TOC.

The natural gamma (GR) logging curve, which generally indicates the characteristics of the lithology, clay content, and organic matter, has been used to estimate the TOC content in previous studies (Wang et al., 2016; Heslop, 2010). The apparent clay content was obtained using Eqs.
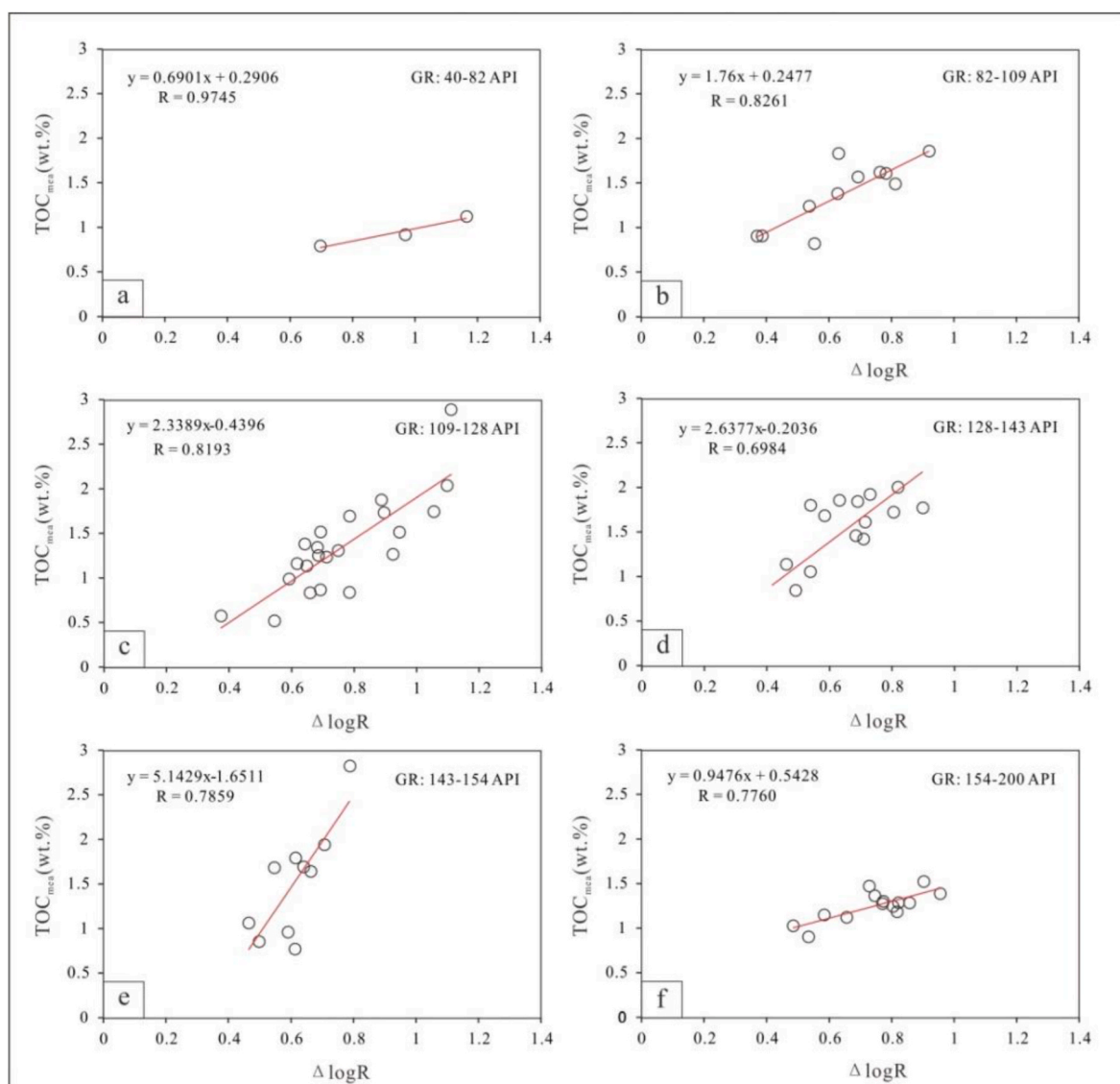

Fig. 7. The numerical relationship between the measured TOC content and ΔlogR of samples with apparent clay contents of (a) 0–10%, (b) 10–20%, (c) 20–30%, (d) 30–40%, (e) 40–50%, and (f) 50–100%.

(11) and (12). The core samples were divided into 6 groups with apparent clay contents of 0–10%, 10–20%, 20–30%, 30–40%, 40–50% and 50–100%. It was found that there is a significant difference in the gradient of the fitted line between $\Delta logR$ and TOC for these 6 groups, which have gradients of 0.6901, 1.76, 2.3389, 2.6377, 5.1429, and 0.9476, respectively (Fig. 7). Therefore, one simple mathematical fitting of $\Delta logR$ and TOC cannot accurately predict the TOC contents of samples with different clay contents.

$$V_{sh} = \frac{2^{GcuR \times V'_{sh}} - 1}{2^{GcuR} - 1}$$ (11)

$$V'_{sh} = \frac{GR - GR_{min}}{GR_{max} - GR_{min}}$$ (12)

where V_{sh} is the apparent clay content. *GcuR* is an empirical coefficient related to the formation, and the value of GcuR is 3.7 for the Paleogene stratum. The GR_{max} of the Liushagang Formation is 200.1 API, and the GR_{min} is 37.9 API.

This study adds GR as a parameter to improve the $\Delta logR$ model of Huo et al. (2011) and Lai et al. (2018). Using $\Delta logR$, GR, and TOC as the coordinates to establish a spatial rectangular coordinate system, a 3D surface was obtained after the biharmonic interpolation (Fig. 8a). The TOC value was determined from GR and $\Delta logR$ using this surface. The ranges of the GR and $\Delta logR$ from low to high TOC content is shown in Fig. 8b. The R and RMSE between measured TOC and the predicted TOC are 0.9656 and 0.1232, respectively (Fig. 9).
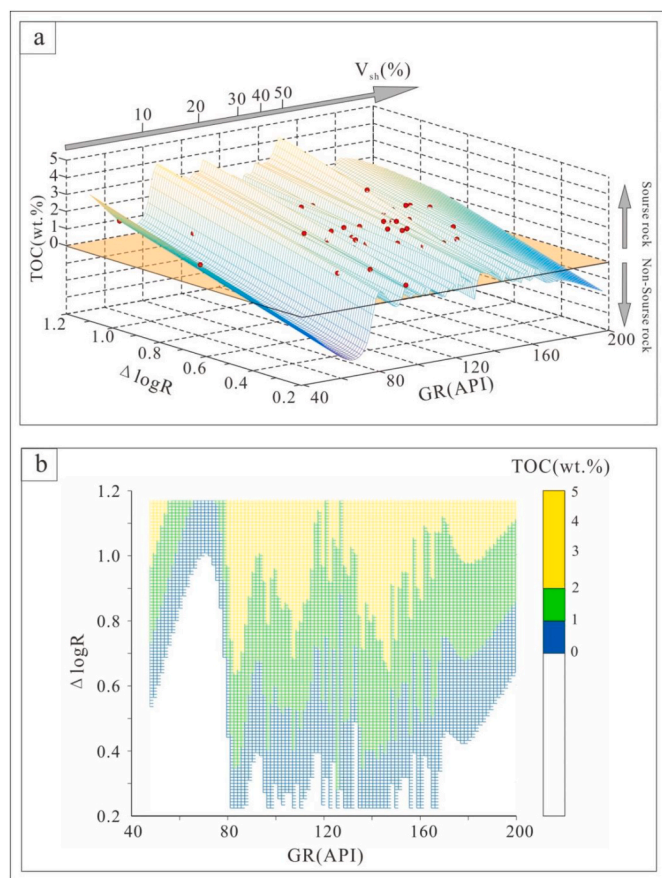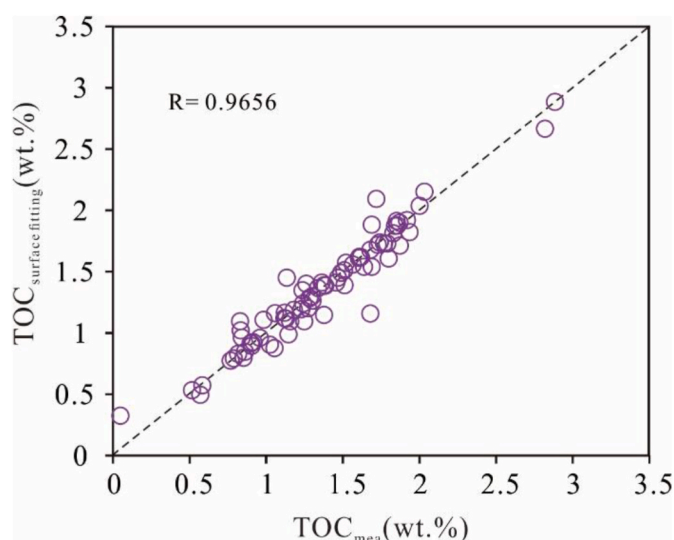

Fig. 9. Prediction results of the 3D surface fitting method.

5. Discussion

5.1. Comparison of the TOC estimation results

Figs. 10 and 11 show the R and RMSE values of the pure mudstone and mudstone interlayers for the different methods. The R values of the pure mudstone and the mudstone interlayers for the 3D surface fitting method (0.9622 and 0.9716, respectively) are much higher than those for the improved $\Delta logR$ method (0.6757 and 0.4884, respectively), the BPANN (0.5622 and 0.7054, respectively), and the MLPNN (0.9028 and 0.9159, respectively). Moreover, the RMSE values of the pure mudstone and mudstone interlayers for the 3D surface fitting method (0.1038 and 0.1392, respectively) are also lower than those for the improved $\Delta logR$ method (0.2443 and 0.5022, respectively), the BPANN (0.4551 and 0.4986, respectively), and the MLPNN (0.1520 and 0.2319, respectively). All of these facts demonstrate that the 3D surface fitting method has the highest prediction accuracy.

The improved $\Delta logR$ method has the lowest accuracy. Note that the R of the pure mudstone is higher than that of the mudstone interlayers, and the RMSE is much lower. This is probably because the well logs are less affected by the surrounding rocks of the pure mudstone, and the TOC has a higher correlation with AC and R_{ILD} (Fig. 4). In contrast, for the mudstone interlayers, the TOC is weakly correlated with AC and R_{ILD}. Therefore, the improved $\Delta logR$ method is more suitable for pure mudstone formations.

The artificial neural network method has a higher accuracy than the improved $\Delta logR$ method. As can be seen from Fig. 10, the BPANN accurately predicts the TOC contents of most of the samples, but a few have large errors; while the MLPNN has a better prediction effect for both the pure mudstone samples and the mudstone interlayer samples. This is mainly due to the different structures of the neural networks, that is, the second hidden layer corrects the prediction results of the first hidden layer, which has larger errors. Moreover, for the MLPNN, the RMSE of the mudstone interlayers is higher than that of the pure mudstone, which demonstrates that the MLPNN method has larger errors in the mudstone interlayers.

Through biharmonic interpolation, the nonlinear relationship between the TOC and $\Delta logR$ of the formations with different natural gamma values was established using the 3D surface fitting method, which greatly broadens the application range of this method and makes it suitable for both pure mudstone and mudstone interlayers.

Fig. 8. (a) The biharmonic interpolation surface based on the GR, $\Delta logR$, and measured TOC. (b) The range of the GR and $\Delta logR$ from low to high TOC content.
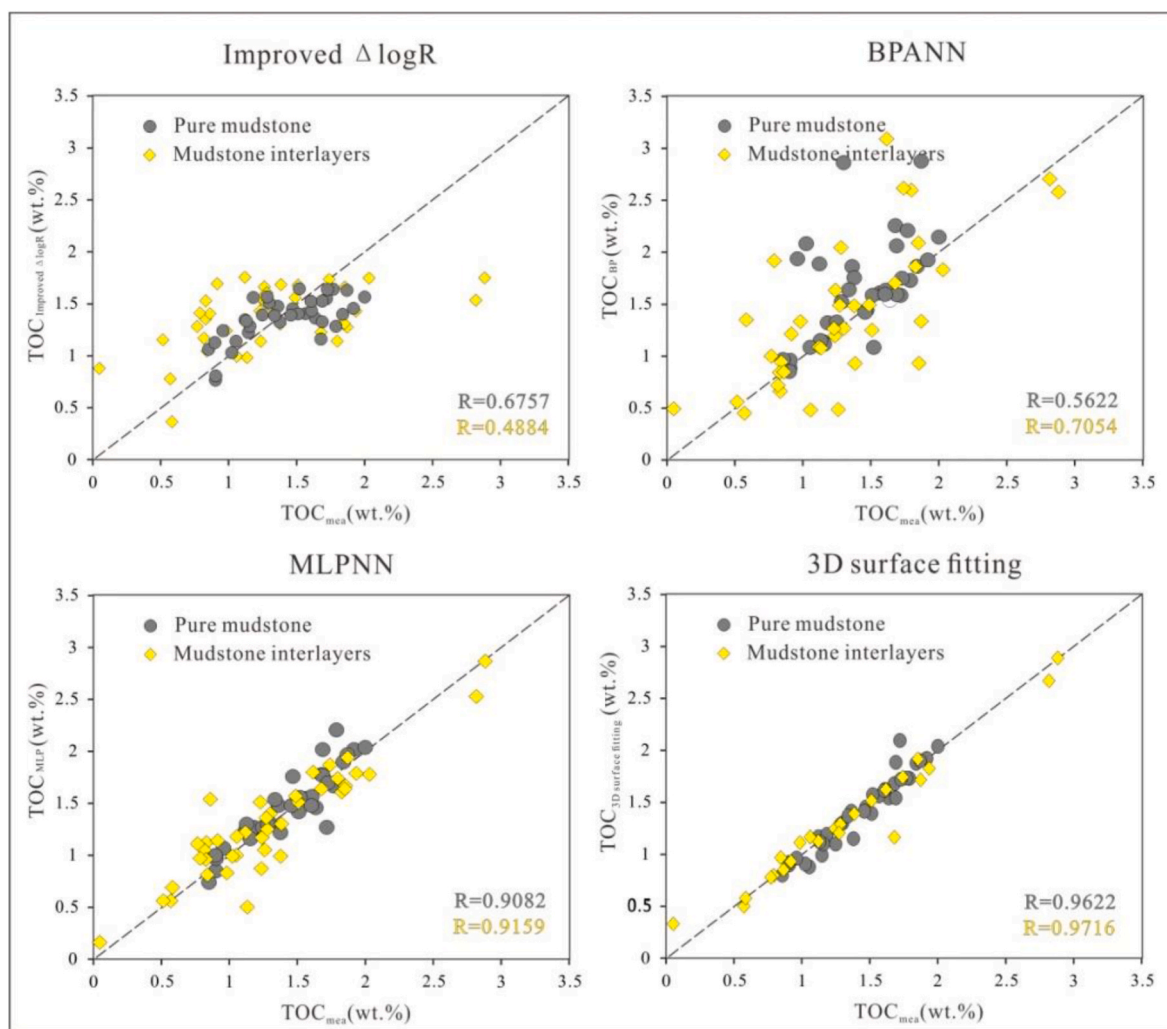
Fig. 10. A comparison of the predicted TOC and measured TOC for the different methods.
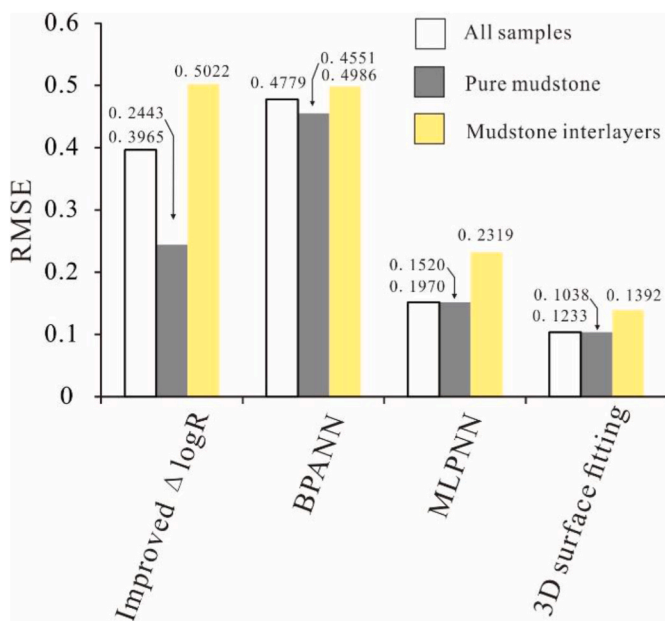

Fig. 11. Root mean square errors (RMSEs) of the different methods.

5.2. Application for TOC content assessment

The Liushagang Formation in well Y11 contains turbidite fan subfacies and semi-deep lake subfacies, and the ratio of the mudstone thickness to the formation thickness is about 80%. The above-mentioned methods were used to process logging data for well Y11 for source rock assessment (Fig. 12). The assessment criteria are described by Peters and Cassa (1994) and Lai et al. (2020a, 2020b). TOC contents of 0%, 0–1%, 1–2%, and >2% represent non-source rocks, poor–fair, good, and very good source rocks, respectively.

The TOC contents of 23 cutting samples were used to evaluate the effects of the various methods (Fig. 12). The results show that the improved ΔlogR and BPANN methods produced large errors and overestimated the TOC content. The accuracies of the MLPNN and 3D surface fitting are relatively high in the depth ranges of 3100–3300 m and 3500–3800 m, which contain pure mudstone. In contrast, in the mudstone interlayers at depths of 3000–3100 m and 3300–3500 m, the prediction precision of the MLPNN method was not very high, while the accuracy of the 3D surface fitting method was not affected by the frequent changes in lithology. Fig. 13 shows the thickness percentages of the non-source rocks to the very good source rocks of the Liushagang Formation in well Y11 obtained using the different methods. According to the MLPNN method, only 3% of the Liushagang Formation is composed of non-source rocks, while the 3D surface fitting result shows that 15% of the formation is composed of non-source rocks, which is similar to the actual sandstone to formation thickness ratio. This is
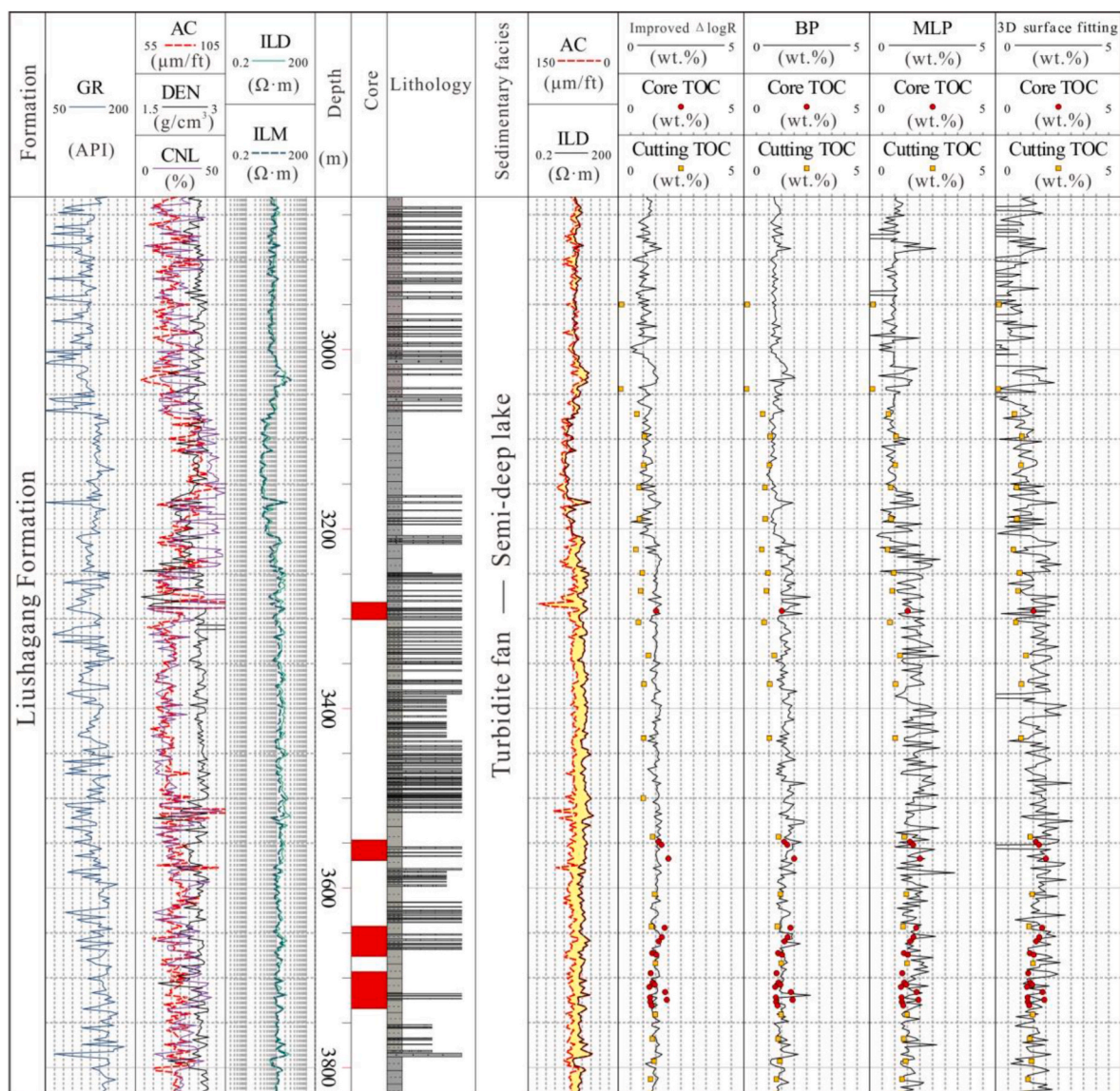
Fig. 12. Source rock assessments of the Liushagang Formation in well Y11 obtained using the different methods.

mainly due to the lack of training for the non-source rock samples in the MLPNN. The TOC estimation results for the other wells are similar to those for well Y11. Therefore, we conclude that the 3D surface fitting method performs better than the other methods in both pure mudstone formations and mudstone interlayer formations.

6. Conclusions

By comparing the above four methods, the following conclusions can be drawn:

(1) The improved ΔlogR method is not suitable for TOC prediction of mudstone interlayers, because mudstone frequently punctuated by sandstone causes the complex mathematical relationship between ΔlogR and the TOC.

(2) Compared to the single hidden layer artificial neural network (BPANN), the double hidden layer neural network (MLPNN) has higher R and lower RMSE and provides more accurate prediction. But this method has poor predictive ability for mudstone interlayers due to the absence of training for the non-source rock samples.

(3) The 3D surface fitting method has R values of higher than 0.96 and RMSE values of lower than 0.14 in both pure mudstone and mudstone interlayers. It can effectively eliminate the errors caused by the frequent lithological changes and provide the best TOC estimation of the mudstone interlayers.

(4) In addition, the 3D surface fitting method can effectively distinguish source rocks from non-source rocks, which has a great significance for source rock assessment and the prioritization of favorable source-reservoir-cap combinations.

Credit author statement

Bang Zeng:provided ideas, write, reviewed and edited the manuscript; Meijun Li:revised the manuscript; Jieqiong Zhu,Yang Shi and Hao Guo:collected samples and materials; Xin Wang, Zhili Zhu, Fangzheng Wang:conducted experiments and analyzed data.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence
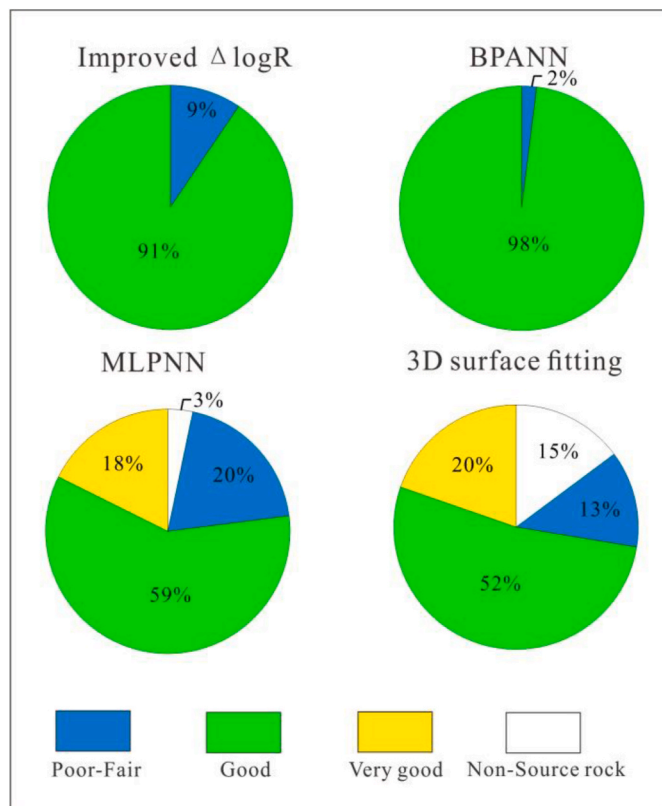
Fig. 13. Thickness percentages of the non-source rocks to the very good source rocks of the Liushagang Formation in well Y11 obtained using the different TOC estimation methods.

the work reported in this paper.

References

Adhikari, S., Marshall, S.J., 2013. Influence of high-order mechanics on simulation of glacier response to climate change: insights from Haig Glacier, Canadian Rocky Mountains. Cryosphere 7 (5), 1527–1541.

Bai, Y., Tan, M., 2020. Dynamic committee machine with fuzzy-c-means clustering for total organic carbon content prediction from wireline logs. Comput. Geosci. 146, 104626.

Bodin, S., Fröhlich, S., Boutib, L., Lahsini, S., Redfern, J., 2011. Early toarcian source-rock potential in the central high atlas basin (central Morocco): regional distribution and depositional model. J. Petrol. Geol. 34, 345–363.

Cao, H., Jin, S., Sun, M., Wang, H., 2016. Astronomical forcing of sedimentary cycles of late Eocene Liushagang Formation in the bailian depression, fushan depression, Beibuwan Basin, south China Sea. J. Cent. South Univ. 23, 1427–1438.

Chen, H., Wang, Q., 2018. Modeling and simulation of the surface topography in ball-end milling based on biharmonic spline interpolation. Int. J. Adv. Manuf. Technol. 99 (9), 2451–2466.

Fath, A.H., Madanifar, F., Abbasi, M., 2020. Implementation of multilayer perceptron (MLP) and radial basis function (RBF) neural networks to predict solution gas-oil ratio of crude oil systems. Petroleum 6 (1), 80–91.

Fertl, W., Chilingar, G., 1988. Total organic carbon content determined from well logs. SPE Form. Evaluation 3 (2), 407–419.

Gan, H., Wang, H., Shi, Y., Ma, Q., Liu, E., Yan, D., Pan, Z., 2020. Geochemical characteristics and genetic origin of crude oil in the Fushansag, Beibuwan Basin, south China Sea. Mar. Petrol. Geol. 112, 104114.

He, J., Ding, W., Zhang, J., Li, A., Zhao, W., Dai, P., 2016. Logging identification and characteristic analysis of marine–continental transitional organic-rich shale in the Carboniferous-Permian strata, Bohai Bay Basin. Mar. Petrol. Geol. 70, 273–293.

He, C., Ji, L., Su, A., Wu, Y., Zhang, M., Zhou, S., Li, J., Hao, L., Ma, Y., 2019. Source-rock evaluation and depositional environment of black shales in the triassic yanchang formation, southern ordos basin, north-central China. J. Petrol. Sci. Eng. 173, 899–911.

Heslop, K.A., 2010. Generalized method for the estimation of TOC from GR and Rt. AAPG Annu. Convention Exhib. Article 80117.

Huo, Q., Zeng, H., Fu, L., Ren, Z., 2011. The advance of ΔlgR method and its application in Songliao Basin. J. Jilin Univ. (Earth Sci. Ed.) 41, 586–591 (in Chinese with English abstract).

Jiang, S., Chen, L., Wu, Y., Jiang, Z., McKenna, E., 2017. Hybrid plays of upper triassic Chang7 lacustrine source rock interval of yanchang formation, ordos basin, China. J. Petrol. Sci. Eng. 159, 182–196.

Kumar, A.R.S., Ojha, C.S.P., Goyal, M.K., Singh, R.D., Swamee, P.K., 2012. Modeling of suspended sediment concentration at Kasol in India using ANN, fuzzy logic, and decision tree algorithms. J. Hydrol. Eng. 17, 394–404.

Lai, H., Li, M., Liu, J., Mao, F., Yang, L., Yang, C., Xiao, H., 2018. Source rock types and logging evaluation in forced regressive systems tract: taking the Termit Basin (Niger) as an example. Acta Sedimentol. Sin. 36 (2), 390–400 (in Chinese with English abstract).

Lai, H., Li, M., Liu, J., Mao, F., Liu, J., Xiao, H., Tang, Y., Shi, S., 2020a. Source rock types, distribution and their hydrocarbon generative potential within the Paleogene Sokor-1 and LV formations in Termit Basin, Niger. Energy Explor. Exploit. 38 (6), 2143–2168.

Lai, H., Li, M., Liu, J., Mao, F., Wang, Z., Liu, W., Yang, L., 2020b. Source rock assessment within a sequence stratigraphic framework of the Yogou Formation in the Termit Basin, Niger. Geol. J. 55 (4), 2473–2494.

Li, Y., Wang, H., Liu, E., Liao, Y., Lin, Z., Ma, Q., 2014. Distribution regularities and control factors for reservoir formation within sequence stratigraphic framework in Fushan Depression, Beibuwan Basin. J. Cent. S. Univ. 45 (5), 1542–1554 (in Chinese with English abstract).

Li, Y., Lin, S., Lin, Z., Ma, Q., 2019. Genesis classification, development mechanism and sedimentary model of deep-lacustrine gravity flow in fushan depression of Beibuwan Basin. J. Jilin Univ. (Earth Sci. Ed.) 49 (2), 323–345 (in Chinese with English abstract).

Li, J., Wang, M., Lu, S., Chen, G., Tian, W., Jiang, C., Li, Z., 2020. A new method for predicting sweet spots of shale oil using conventional well logs. Mar. Petrol. Geol. 113, 104097.

Lim, J., 2003. Reservoir permeability determination using artificial neural network. J. Korean Soc. Geosyst. Eng. 40, 232–238.

Liu, E., Wang, H., Li, Y., Zhou, W., Leonard, N., Lin, Z., Ma, Q., 2014. Sedimentary characteristics and tectonic setting of sublacustrine fans in a half-graben rift depression, Beibuwan Basin, South China Sea. Mar. Petrol. Geol. 52, 9–21.

Liu, E., Wang, H., Li, Y., Leonard, N., Feng, Y., Pan, S., Xia, C., 2015. Relative role of accommodation zones in controlling stratal architectural variability and facies distribution: insights from theFushan Depression, South China Sea. Mar. Petrol. Geol. 68, 219–239.

Lu, L., Cai, J., Liu, W., Teng, G., Wang, J., 2013. Occurrence and thermostability of absorbed organic matter on clay minerals in mudstones and muddy sediments. Oil Gas Geol. 34 (1), 16–26 (in Chinese with English abstract).

Mahmoud, A.A.A., 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.

Passey, Q., Creaney, S., Kulla, J., Moretti, F., Stroud, J., 1990. A practical model for organic richness from porosity and resistivity logs. AAPG Bull. 74, 1777–1794.

Peters, K.E., Cassa, M.R., 1994. Applied source rock geochemistry. In: Magoon, L.B., Dow, W.G. (Eds.), The Petroleum System-From Source to Trap, vol. 60. AAPG Memoir, pp. 93–115.

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. Journal of Petroleum ence and Engineering. J. Pet. Sci. Eng. 180, 699–706.

Sandwell, D.T., 1987. Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data. Geophys. Res. Lett. 14 (2), 139–142.

Sandwell, D.T., Smith, W.H.F., 2009. Global marine gravity from retracked Geosat and ERS-I altimetry: ridge segmentation versus spreading rate. J. Geophys. Res-Sol. Ea. 114 (B1), 1–18.

Schmoker, J., 1979. Determination of organic content of Appalachian Devonian shales from formation-density logs. AAPG Bull. 63 (9), 1504–1509.

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 wireline logs. J. Nat. Gas Sci. Eng. 33, 687–702.

Tan, M., Liu, Q., Zhang, S., 2013. A dynamic adaptive radial basis function approach for total organic carbon content prediction in organic shale. Geophysics 78, D445–D459.

Urang, J.G., Ebong, E.D., Akpan, A.E., Akaerue, E.I., 2020. A new approach for porosity and permeability prediction from well logs using artificial neural network and curve fitting techniques: a case study of Niger Delta, Nigeria. J. Appl. Geophys. 183, 104207.

Vega-Ortiz, C., Beti, D.R., Setoyama, E., McLennan, J.D., Ring, T.A., Levey, R., Martínez-Romero, N., 2020. Source rock evaluation in the central-western flank of the Tampico Misantla Basin, Mexico. J. South Am. Earth Sci. 100, 102552.

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. Petrol. Geol. 70, 304–319.

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.