Oil Shale, 2014, Vol. 31, No. 2, pp. 132–146
doi: 10.3176/oil.2014.2.04
ISSN 0208-189X

A NOVEL WIRELINE LOGS BASED APPROACH FOR ASSESSMENT OF MINERAL CONCENTRATIONS IN ORGANIC SHALE

MAOJIN TAN[(a,b)*], YOULONG ZOU[(b)], XIAOCHANG WANG[(c)], YUE GUO[(b)]

[(a)] Key Laboratory of Geo-detection (China University of Geosciences), Ministry of Education, Beijing, 100083, China
[(b)] China University of Geosciences, Beijing, 100083, China
[(c)] Petroleum Exploration and Production Institute, SINOPEC, Beijing, 100083, China

Abstract. *It is usually difficult to establish a petrophysical model or empirical formula for the assessment of mineral concentrations in oil- or gas-bearing shale. The radial basis function (RBF) network can be used to construct the mapping between well logs and mineral concentration. In this work, the basic principle of RBF is discussed in detail, including network structure, basis function, and network training method. The nearest neighbor algorithm is selected for the network training. Then, one case study for the mineralogy analysis is applied to show how to construct an optimum RBF network. The Gaussian spread in RBF is investigated to improve the mineral composition prediction accuracy through leave-one-out cross validation and optimum wireline logs as inputs are chosen. Finally, the concentrations of minerals such as quartz, feldspars, calcite and pyrite as well as clay minerals are calculated, and they are all in good agreement with X-ray Diffraction (XRD) measurement results. Furthermore, the errors analysis indicates that the RBF method is effective and applicable for the assessment of mineral concentrations in organic shale.*

Keywords: *organic shale, radial basis function (RBF), nearest neighbor (NN) algorithm, mineralogy assessment.*

1. Introduction

Unconventional oil/gas resources become more important because of various difficulties encountered in conventional oil exploration. Shale oil or gas is an important alternative to conventional oil/gas resources. Shale oil/gas forma-

* Corresponding author: e-mail *tanmj@cugb.edu.cn*

tions have complicated pore structure and mineral composition. Gas is partially absorbed and partially free in such formations, which makes formation evaluation more difficult. Unconventional reservoirs evaluation mainly includes mineral components prediction, gas contents and porosity calculation, etc. [1].

In the exploration of shale oil and shale gas, wireline logging is a key borehole geophysical technology, which provides geophysical parameters such as resistivity, formation slowness, formation density, etc. [2, 3]. These well logging data can then be used to predict porosity, permeability, mineral volume and oil/gas saturation by using some petrophysical models in conventional oil/gas formations. In addition, wireline logs are also used to estimate the total organic carbon (TOC) content of source rock [4–10]. Some methods only employ one or a few logs to calculate TOC content, and the TOC prediction data are often incomplete and not accurate for a particular formation [5–8]. Due to the complex composition of unconventional reservoirs such as hydrocarbon-bearing shale and methane-bearing coal, petrophysical models have not been widely used for their mineral composition prediction and the methods employed hitherto are not suitable either. Chen et al. combined a genetic algorithm (GA) and BP network to optimize the neural network, which can predict the mineral components of methane-bearing coal [11]. Certain intelligent algorithms can effectively take advantage of some sensitive logs to predict the mineral components. Yang et al. introduced a stochastic joint inversion method specifically developed to address the quantitative petrophysical interpretation of hydrocarbon-bearing shale, which is based on the rapid and interactive numerical simulation of resistivity and nuclear logs [12]. Compared to traditional deterministic estimation procedures, the new interpretation method explicitly quantifies the uncertainty of properties of hydrocarbon-bearing shale.

In recent years, a radial basis function (RBF) interpolation method has been proposed by several researchers to solve inversion problems for the analysis of lab measurement results [13–16]. The method proposed can approximate smooth and continuous multivariate functions of many variables and derive a nonlinear mapping function to solve well-logging and geophysical inverse problems associated with unknown forward models. Freed was the first to put forward a new approach to determine the characteristics and composition of a fluid sample [13]. Heaton and Freedman used the RBF method to predict the properties of live oil from nuclear magnetic resonance (NMR) relaxation time data [14]. In 2006, Freedman summarized the RBF method and theory, and introduced three case studies, including the viscosity prediction of dead crude oil from NMR data, environment correction of 3D induction tools, and the molecular composition of dead crude oil from near-infrared (NIR) spectra. Overall, all the predicted data were in excellent agreement with lab measurement results [15]. Anand and Freedman adopted the RBF method to determine the molecular properties of hydrocarbon mixtures from NMR data [16]. Huang et al. also employed the

RBF network for well log inversion, and computed the true formation conductivity from the apparent conductivity, which similarly showed that RBF is suitable for parameters prediction from well log data [17]. But can certain intelligent algorithms effectively be applied to predicting the mineral components of organic shale?

Therefore, the current research attempts to calculate the mineral composition for shale oil/gas reservoirs by using the RBF method. In this paper, the principle of the RBF method is first reviewed, including basis function and spread, network structure and nearest neighbor algorithm of the RBF interpolation. Then, a case study for the mineral components prediction of shale gas reservoirs is presented based on the RBF method. The comparison with core measurements proves that the prediction results are accurate, and the RBF method is effective to be used for predicting the mineral components for shale oil/gas reservoirs.

2. The RBF theory

2.1. RBF network model

In 1985, Powell first put forward a multivariable interpolation of the radial basis function (RBF) method [18]. In 1988, Moody and Darken presented a neural network structure, namely RBF network [19]. The neural network is a three-layer feed forward neural network, including input layer, hidden layer and output layer, which are shown in Figure 1. The input layer nodes depicted in the figure are only used to transfer the input signal to the hidden layer. The effect function of hidden nodes, namely basis functions, will produce responses partly to the input signal. When the input signal is close to the central area of basis functions, the hidden nodes will produce a larger output. The output layer node is usually a simple linear function.

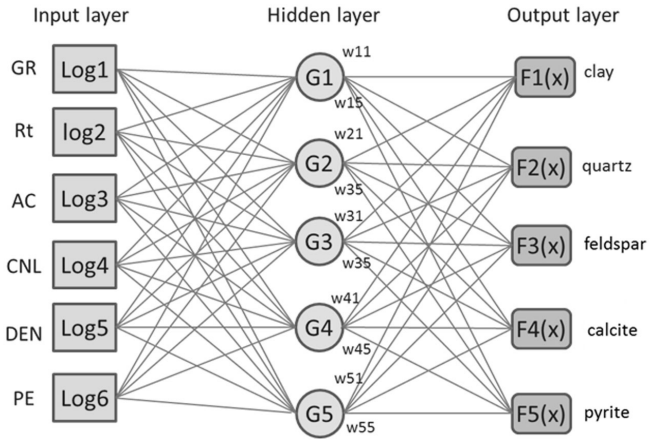

Fig. 1. The diagram of RBF neural network structure for the mineralogy analysis in organic shale.

The pores of organic shale are small and its mineral composition is complex. It is difficult to calculate mineral concentrations. Several wireline logs are sensitive to the logging responses of reservoirs, but the relations between the logs and mineral concentrations are not linear, so it is difficult to determine the volumetric concentration of minerals. If certain sensitive logging attributes are chosen as the input layer, the unknown mineral concentrations are set as the output layer, the RBF network can potentially be used to predict the mineral components in organic shale. Figure 1 shows the neural network model of the radial basis function for predicting mineral content by using logging data. In the figure, the input layer is a sensitive logging curve, the hidden layers are radial basis functions, and the output layer is the unknown mineral content. For organic shale reservoir, the unknown minerals usually include clay minerals, quartz, feldspars, calcite, pyrite, etc., whose RBF network structure is illustrated in Figure 1.

2.2. Basis functions

The basis functions are effect functions of hidden nodes for the RBF neural network, which are generally called radial basis functions. The radial basis functions are commonly expressed as follows [18]:

$$\begin{aligned} f(x) &= \exp\left(-\frac{x}{r}\right)^2 \\ f(x) &= -\frac{1}{(r^2+x^2)^{\alpha}} \\ f(x) &= (\alpha^2+x^2)^{\beta}, \end{aligned} \tag{1}$$

where $\alpha > 0$ and $\alpha < \beta < 1$, respectively.

These functions are radially symmetric. The Gaussian function has advantages such as simple form, radial symmetry, smoothness and good analytical characteristic for the theoretical analysis. So it is commonly chosen as a radial basis function.

In vector space, the Gaussian basis function is chosen as the radial basis function:

$$G_i(x) = \exp\left(\frac{-\|x-c_i\|^2}{2r_i^2}\right), \tag{2}$$

where $i = 1, 2, \ldots, m$; x is the n-dimensional input vector; c_i is the center of the i-th basis function, which has the same dimension with x; r_i is the i-th apperceive variable, which may be chosen freely and determines the width of the basis function around the center c_i; m is the number of apperceive units, namely the number of hidden nodes; $\|x-c_i\|$ is the norm of the $x-c_i$ vector, which usually means the distance between x and c_i; $G_i(x)$ has only a maximum value at c_i, and with increasing $\|x-c_i\|$, $G_i(x)$ will attenuate to

zero rapidly. For a given input vector $x \in R_n$, only a small part near the center of x is activated. This reflects the response characteristics of the brain cortex, which shows the local approximation ability of the network, so the radial basis function network is also called the local apperceive field network.

Form Figure 1 it can be seen that the mapping $x \to G_i(x)$ from the input layer to the hidden layer is nonlinear, and the mapping from the hidden layer to the output layer, $G_i(x) \to y^k$, is linear, namely:

$$y_k = \sum_{i=1}^{m} w_{i,k} G_i(x), \tag{3}$$

where $k = 1, 2, \ldots, m$; y_k is the k-th output node (output variable); $w_{i,k}$ is the output weighting coefficient of the RBF network; m is the number of hidden nodes. As $G_i(x)$ is the Gaussian function and $G_i(x) > 0$ for any x, it has an advantage of the local adjusting weighting constant. In fact, when x is far away from c_i, the basis function $G_i(x)$ is already very small, so it can be treated as 0. When $G_i(x)$ is larger than a certain value, i.e. $G_i(x) = 0.05$, the corresponding weighting coefficient $w_{i,k}$ is not zero, which is meaningful for the RBF network. Considering the treatment above, the RBF network has some advantages of the local approaches network, as well as fast learning speed.

2.3. Nearest neighbor clustering learning algorithm

The learning algorithms of the RBF neural network mainly include random algorithm, self-organizing learning algorithm and the nearest neighbor clustering learning algorithm [17, 18]. These learning algorithms are used to determine the center of RBF. The random and self-organizing learning algorithms are applicable for the off-line learning of a static model, which must get all the possible sample data beforehand, and cannot be used in the on-line learning of a dynamic input mode. Before learning, the number of input data centers, i.e. the number of hidden nodes in the RBF network, must be determined in advance, and this makes it more difficult to solve the problem. The nearest neighbor learning algorithm is a dynamic adaptive RBF network model, which needs no prior determination of the number of hidden units. Furthermore, it is online learning, and the RBF network after clustering may be the optimum.

The process of the nearest neighbor clustering learning algorithm is the following:

(1) An appropriate width r of the Gaussian function is chosen. The vector $A(l)$ is defined to store the sum of output vectors belonging to various samples. The counter $B(l)$ is defined to count the number of various types of samples, where l is the number of categories.

(2) First, start from the first data pair (x^1, y^1), establish a cluster center on x^1, namely, let $c_1 = x^1$, $A(1) = y^1$, $B(1) = y^1$. The established RBF network

has only one hidden layer unit, the center of the hidden layer unit is c_1. The weight vector of the hidden layer units to the output layer is $w_1 = A(1)/B(1)$.

(3) Then, consider the second sample data pair (x^1, y^1), and calculate the distance $\|x^2 - c_1\|$ from x^2 to the cluster center c_1. If $\|x^2 - c_1\| \le r$, c_1 is the nearest neighbor clustering of x^2, and let $A(1) = y^1 + y^2$, $B(1) = 2$, $w_1 = A(1)/B(1)$.

(4) If $\|x^2 - c_1\| > r$, then treat x^2 as a new cluster center, and let $c^2 = x^2$, $A(2) = y^2$, $B(2) = 1$. Add another hidden layer unit to the RBF network above, and the weight vector of the hidden layer units to the output layer is $w_2 = A(2)/B(2)$.

(5) If we consider the *k*-th sample data pair (x^k, y^k), $(k = 3, 4, \cdots, N)$, there are M cluster centers, whose center points are $c_1, c_2, \ldots, c_M$, respectively. So, there are M hidden layer units in the RBF network. And then, calculate the distance $\|x^k - c_j\|$, $(j = 1, 2, \ldots, M)$, from x^k to the *M*-th cluster center. Let $\|x^k - c_k\|$ be the minimum among these distances, c_j is the nearest neighbor clustering of x^k. That is to say, if $\|x^k - c_j\| > r$, treat x^k as a new cluster center. Let $c_{M+1} = x^k$, $M = M + 1$, $A(M) = y^k$, $B(M) = 1$, and keep the value of $A(i), B(i)$, $(i = 1, 2, \ldots, M)$ unchanged. The weight vector in the *M*-th hidden layer units to the output layer is $w_m = A(M)/B(M)$. If $\|x^k - c_j\| \le r$, then $A(j) = A(j) + y^k$, $B(j) = B(j) + 1$. When $i \ne j$, the value of $A(i)$ and $B(i)$ are kept unchanged, the weight vector of the hidden layer units to the output layer is $w_i = A(i)/B(i)$, $(i = 1, 2, \ldots, M)$.

Through such learning algorithm established, the RBF network output is:

$$f(x^k) = \frac{\sum_{i=1}^{m} w_{i,k} \exp\left(-\|x - c_i\|^2 / r^2\right)}{\sum_{i=1}^{m} \exp\left(-\|x - c_i\|^2 / r^2\right)}, \tag{4}$$

where r is the Gaussian width, namely the spread constant. As r is a one-dimensional parameter, the appropriate spread constant could be determined through multi-time cross-validation tests and the errors contrast. The RBF method is much more convenient than the BP neural network, which needs determination of the number of hidden layers and whose results are not always reliable [18–20]. In fact, because each pair of input and output data may create a new cluster, this kind of a dynamic adaptive RBF network is realized through adaptive adjustments of the network structure, including different inputs and parameters at the same time.

3. Case study

A case study well is located in the northeastern wing of the Huangping syncline, Qiannan down warping region of Xinzhuang area, China. The objective of the drilling is to evaluate hydrocarbon of organic shale of the

Jiumenchong group. The formation evaluations include gas bearing layer identification, gas content evaluation, mineral volumetric concentration determination, and lithology and lithofacies identification, with mineral component evaluation being the most important. 27 core samples are subjected to X-ray Diffraction (XRD) analysis, and the mineral volumetric concentrations are measured. The results of XRD analysis show that the shale in the Jiumenchong group is composed of clay, quartz, feldspars, calcite and pyrite. The geophysical logging series of this well include: natural gamma ray (GR), spontaneous potential (SP), dual lateralog (LLD + LLS), compensated acoustic log (AC), compensated density log (DEN), compensated neutron log (CNL) and natural gamma ray spectrum logs (U + TH + K). How to predict the fraction volume of each mineral by using wireline logs is a challenge.

In this research, the RBF network method is adopted to predict mineral volumes. In the RBF method, the construction of an optimum RBF network is a key problem, so, this paper discusses how to find an optimum Gaussian spread through testing dataset validation and cross-validation test, respectively.

The logging data preprocessing is needed. Before carrying out the RBF test, logging data need to be normalized in advance. Furthermore, the resistivity log is transformed to a logarithmic scale.

3.1. Testing dataset validation

For constructing an optimum RBF network, the XRD data of 27 cores are divided into two datasets. One dataset is named a training set and is used to construct the optimum network, and the other, which is named a testing dataset, is used to validate the prediction method.

In each RBF test, the different spread, namely the Gaussian width, is chosen to do some tests to train the network model. The mean absolute error and mean relative error are calculated in the training dataset and testing dataset, respectively:

$$MAE_i = \frac{\sum_{j=1}^{N} \left| V_{i,j} - C_{i,j} \right|}{N}$$
$$MRE_i = \frac{\sum_{j=1}^{N} \frac{\left| V_{i,j} - C_{i,j} \right|}{C_{i,j}}}{N} \times 100\%, \tag{5}$$

where i = 1, 2, 3, …, M; j = 1, 2, 3, … N; $V_{i,j}$ is the predicted volumetric concentration of the i-th mineral corresponding to the j-th core sample; $C_{i,j}$ is the XRD measurement result of the i-th mineral of the j-th core sample; M is the number of minerals; N is the number of core samples; MAE_i and MRE_i are the mean absolute error and mean relative error of the i-th mineral, respectively.

To compare the multimineral evaluation accuracy of the RBF test of each sample or depth, *MAE* and *MRE* are defined as the mean absolute error and mean relative error of *M* minerals of one sample or in one depth, respectively:

$$MAE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left|V_{i,j} - C_{i,j}\right|}{MN}$$
$$MRE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{\left|V_{i,j} - C_{i,j}\right|}{C_{i,j}}}{MN} \times 100\%, \tag{6}$$

where $i = 1, 2, 3, \ldots, M$; $j = 1, 2, 3, \ldots, N$.

Firstly, 9 logs are selected as inputs to carry out RBF tests, which include GR, U, TH, K, AC, CNL, DEN, PE and LLD. Figure 2 illustrates the difference between MAE and MRE with increasing spread. It can be seen that MAE and MRE first gradually decrease and then increase as the spread increases in the predicting dataset. The best value of the Gaussian width is selected at the "inflection point", and the spread is about 2.0. The RBF network with this spread is optimum. In the training dataset, MAE and MRE are about 0.0280 and 11.52%, respectively; and in the testing dataset, MAE is about 0.0110, and MRE is about 3.52%. Table 1 presents the prediction results of RBF with 9 logs as inputs in this case study, in which the expected results refer to the core XRD measurements. The MREs of the majority of samples are lower than 10%, but for some samples they are higher than 20%. So the mineral composition evaluation is accurate enough to meet requirements.

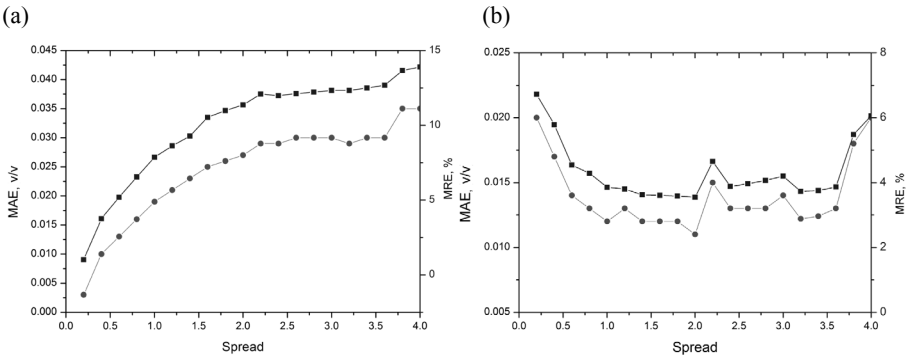

Fig. 2. Comparison of errors of RBF tests with different Gaussian spreads: (a) MAE and MRE of the training dataset; (b) MAE and MRE of the testing dataset.

Table 1. Results of mineral concentration prediction using RBF with 9 logs as inputs in the case study well

No		Expected results					Predicted results					Relative error, %
		Clay, v/v	Quartz, v/v	Feld spar, v/v	Calcite, v/v	Pyrite, v/v	Clay, v/v	Quartz, v/v	Feldspar, v/v	Calcite, v/v	Pyrite, v/v	
1	Training dataset	0.105	0.125	1E-3	0.461	0.075	0.133	0.252	0.037	0.307	0.071	21.17
2		0.358	0.258	0.02	0.285	0.069	0.332	0.212	0.020	0.330	0.072	22.95
3		0.204	0.49	0.133	0.073	0.1	0.24	0.483	0.134	0.069	0.074	4.43
4		0.2	0.555	0.13	0.053	0.062	0.24	0.485	0.134	0.068	0.073	13.23
5		0.172	0.443	0.146	0.175	0.064	0.24	0.479	0.134	0.072	0.076	17.03
6		0.214	0.533	0.131	0.067	0.055	0.239	0.479	0.134	0.072	0.076	10.21
7		0.29	0.397	0.121	0.13	0.062	0.241	0.469	0.136	0.072	0.082	19.32
8		0.238	0.432	0.117	0.144	0.069	0.239	0.476	0.134	0.073	0.078	11.31
9		0.234	0.508	0.148	0.053	0.057	0.241	0.474	0.138	0.065	0.08	6.53
10		0.243	0.511	0.137	0.051	0.058	0.239	0.478	0.134	0.072	0.076	5.87
11		0.218	0.376	0.108	0.129	0.169	0.241	0.469	0.136	0.072	0.082	26.37
12		0.237	0.513	0.126	0.046	0.078	0.242	0.49	0.134	0.064	0.07	4.56
13		0.227	0.482	0.154	0.047	0.09	0.24	0.475	0.134	0.073	0.078	3.7
14		0.251	0.475	0.143	0.059	0.072	0.24	0.483	0.134	0.069	0.074	2.62
15		0.226	0.519	0.136	0.047	0.072	0.241	0.486	0.134	0.067	0.072	6.26
16		0.238	0.502	0.137	0.042	0.081	0.24	0.479	0.134	0.071	0.076	4.08
17		0.241	0.531	0.161	1E-3	0.067	0.241	0.462	0.133	0.075	0.075	11.34
18		0.258	0.505	0.134	0.045	0.058	0.246	0.497	0.135	0.058	0.065	2.44
19		0.243	0.455	0.15	0.075	0.077	0.24	0.478	0.135	0.071	0.077	4.55
20		0.232	0.534	0.14	0.034	0.06	0.24	0.483	0.134	0.068	0.074	8.74
21	Testing dataset	0.229	0.472	0.162	0.059	0.078	0.2340	0.4934	0.1410	0.0589	0.0726	3.87
22		0.25	0.503	0.134	0.047	0.066	0.2327	0.4902	0.1387	0.0645	0.0738	5.77
23		0.253	0.479	0.134	0.072	0.062	0.2328	0.4904	0.1389	0.0641	0.0737	2.34
24		0.243	0.493	0.146	0.05	0.068	0.2341	0.4936	0.1412	0.0585	0.0725	2.63
25		0.24	0.503	0.144	0.057	0.056	0.2337	0.4927	0.1405	0.0603	0.0729	4.44
26		0.261	0.494	0.142	0.046	0.057	0.2336	0.4925	0.1404	0.0606	0.0730	3.75
27		0.244	0.5	0.123	0.063	0.07	0.2331	0.4913	0.1395	0.0626	0.0734	1.81

3.2. Cross-validation tests and optimum parameters

In order to determine the influence of regression algorithms, kernel functions, as well as other parameters on the forecast test results, in this research, the leave-one-out test is used. The test consists in that one from all samples is chosen for mineral volumetric concentration prediction, and the others are regarded as a training dataset to conduct a series of tests with different Gaussian spreads. A Gaussian spread is used to construct an optimum network to get the coefficient matrix. Finally, the mineral volumetric concentration of the left-one sample is calculated. Similarly, the mineral assessment results of other samples are predicted one by one.

When the RBF method is employed to predict the mineral concentrations of the whole logging interval, in order to achieve the best evaluation, all experimental samples are used to construct an optimum neural network. Furthermore, to study the sensitivity to well logs attributes, RBF prediction tests are carried out with a varying number of logs as inputs. In each test, an optimum Gaussian spread is chosen through errors contrast. The final prediction results and errors of RBF tests with a varying number of logs as inputs are listed in Table 2.

Table 2 presents the comparison of errors of different RBF tests. The MAE and MRE of the RBF test with 9 logs as inputs are both minimum, which indicates that in different tests the RBF prediction affords the best results. The MAE and MRE of the 8-log RBF test without the resistivity log are both much larger than those of the 9-log RBF test. It implies that the resistivity log is highly sensitive to the mineral concentration of organic shale, especially, it is probably related to pyrite. The MAE and MRE of the 5-log RBF test with GR/AC/DEN/CNL/PE are both much lower than those of the 4-log RBF test with AC/DEN/CNL/PE. This shows that the gamma ray log is indispensable in the mineral concentration assessment of organic shale. Table 2 shows the results of the 4-log RBF test with GR/U/TU/K, and also demonstrates that MAE and MRE are the largest in 5-log RBF tests. This indicates that only the gamma spectrum log is not enough and other logs are also indispensable for mineral concentration evaluation of organic shale.

Table 3 presents errors of the concentration prediction for five minerals in the 9-log RBF test. The concentrations of clay, feldspar and quartz have been predicted more accurately, but those of calcite and pyrite were both difficult to predict accurately. Figure 3 illustrates the results of the RBF-based prediction with 9 logs as inputs, as well as those from core XRD measurements. From the figure it can be seen that the predicted concentrations of clay, quartz and feldspars are in good agreement with core XRD measurement results, while for pyrite and calcite the predicted results are not in conformity.

Table 2. Results of the RBF-based mineral concentration prediction with a varying number of logs as inputs in the case study well

No	Input	Spread	Errors comparison	
			MAE, v/v	MRE, %
1	GR/U/TH/K/AC/DEN/CNL/PE/logLLD	0.25	0.0151	6.71
2	GR/U/TH/K/AC/DEN/CNL/PE	0.25	0.1018	26.01
3	GR/AC/DEN/CNL/PE	0.25	0.0245	22.38
4	AC/DEN/CNL/PE	0.25	0.0739	25.45
5	GR/U/TH/K	0.25	0.3403	29.39

Table 3. Errors of the RBF-based mineral concentration prediction with 9 logs as inputs in the case study well

Mineral	MAE, v/v	MRE, %
Clay	0.0188	4.05
Quartz	0.0186	1.98
Feldspars	0.0064	2.31
Calcite	0.0178	16.85
Pyrite	0.0139	8.36

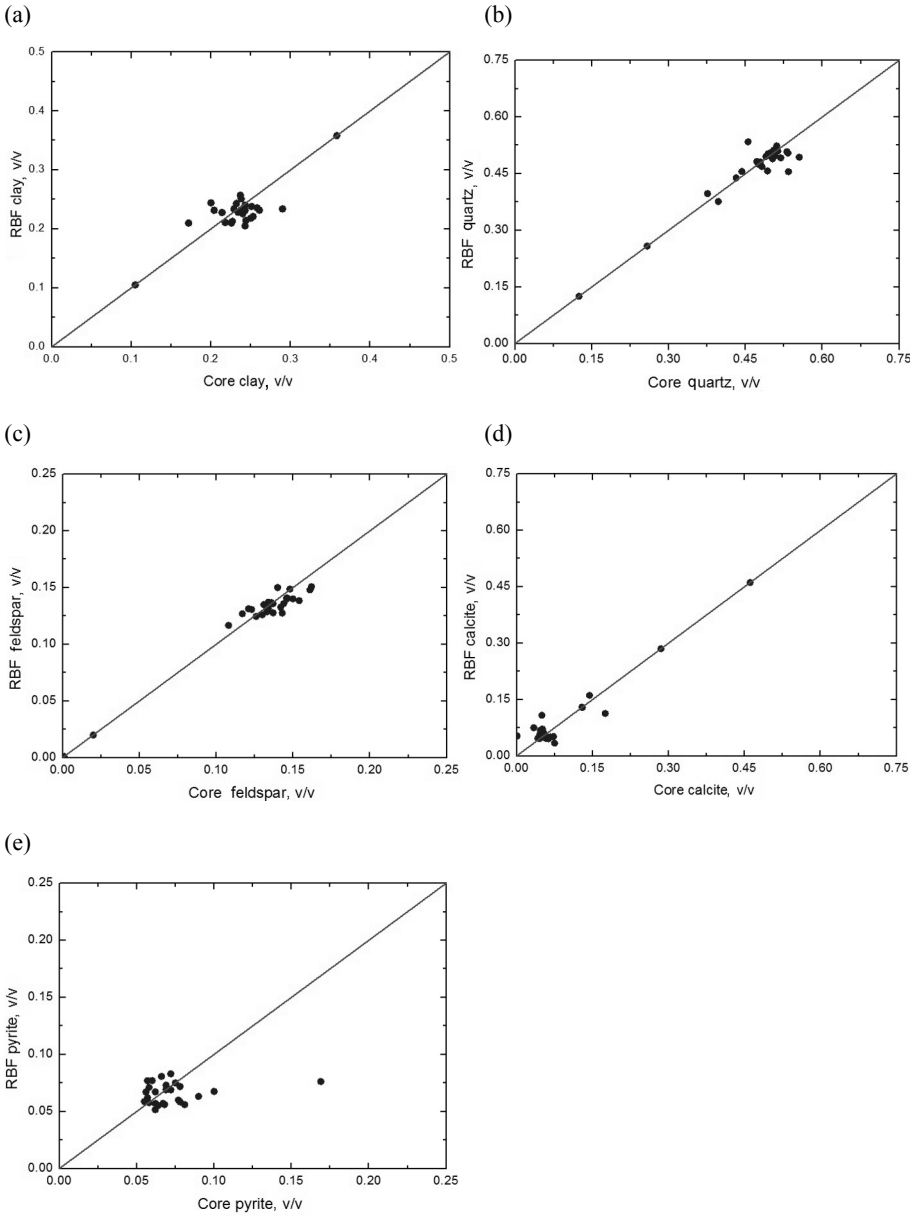

Fig. 3. Comparison of the results of XRD measurements and RBF-based prediction: (a) core clay vs RBF clay, (b) core quartz vs RBF quartz, (c) core feldspars vs RBF feldspars, (d) core calcite vs RBF calcite, (e) core pyrite vs RBF pyrite.

For evaluating the mineral volumetric concentrations of the whole well, all core XRD sample and 9 logging attributes are used to determine the optimum parameters and construct the optimum network. Then the optimum RBF network is applied to predicting the mineral volumetric concentration

of the organic shale interval. Figure 4 shows the RBF predicted concentrations of the above five minerals.

Figure 4 also illustrates the core XRD measurement results for comparison. From the figure it can be seen that the RBF predicted results are in good agreement with core XRD measurements. Based on calculated mineral concentrations, it is indicated that the target formation is relatively homogeneous in

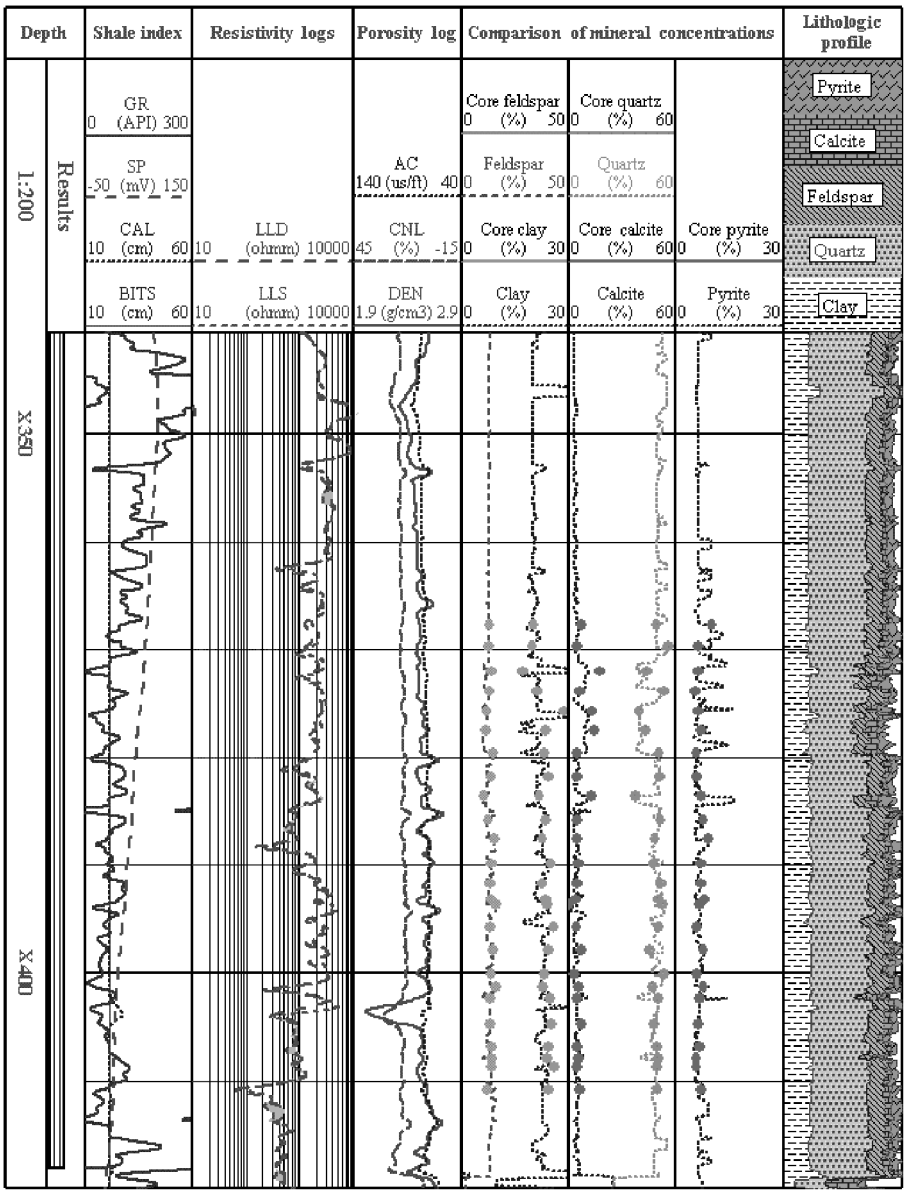

Fig. 4. Results of the RBF-based mineral concentration prediction in the case study well.

minerals and lithology: the clay content is about 0.2, the pyrite volume is approximately 0.6, the feldspars volume is about 0.13, and the quartz volume is approximately 0.5. But in the interval between X370 and X394 m, the quartz volume is about 0.4 and the calcite volume is approximately 0.1. In this interval, because of the high quartz volume, the formation is more brittle, and there may be more induced fractures. In 2011, this interval was successfully stimulated and fractured, and about 400 m^3 gas was obtained per day.

Generally, the authors used the RBF method in one study zone. Using limited core XRD measurements from similar layers of different wells the optimum network can be constructed, and then this optimum network can be applied to predicting the mineral volumetric concentration of other wells in the study area.

4. Conclusions

Through the research and analysis of the RBF method used for organic shale mineral concentration prediction, the following conclusions can be drawn:

(1) The RBF network is a feed forward network with good performance, and has global approximation properties and the best approximation performance. The learning algorithm of the RBF neural network is an online self-adaptive cluster learning algorithm, the training method is rapid and feasible.

(2) The RBF method can be used to easily construct the mapping between various logs and multiple minerals for mineralogy assessment through Gaussian spread optimization. The method can archive the prediction of multimineral concentration at the same time.

(3) For oil- and gas-bearing shale, it is difficult to determine a log interpretation model, so the RBF method can achieve the mineral volume prediction from wireline logs, which is more accurate and more suitable than the empirical formula. The optimum RBF network is optimized through a series of tests with a varying number of logs as inputs. In this case study, 9 wireline logs are all needed, while the gamma ray and resistivity logs are indispensable.

However, the RBF method is dependent on the lab measurement results of core samples, and the training network is suitable for the formation where the training sample data are from. Therefore, the method has still some regional limitations. So, samples for the training RBF network should cover different lithologies, and such an RBF network will have good generalization ability.

Acknowledgments

This paper was sponsored by the National Natural Science Foundation of China (No. 41172130), the Fundamental Research Funds for the Central Universities (No. 265201248), the CNPC Innovation Foundation (No. 2011D-5006-0305), and the National Major Projects "Development of Major Oil and Gas Fields and Coal Bed Methane" (No. 2011ZX05014-001). The support from the China Scholarship Council (No. 201206405002) is acknowledged.

REFERENCES

1. Bodin, S., Frohlich, S., Boutib, L., Lahsini, S., Redfern, J. Early Toarcian source-rock potential in the central high Atlas basin (central Morocco): regional distribution and depositional model. *J. Petrol. Geol.*, 2011, **34**(4), 345–363.
2. Bateman, R. M. *Openhole Log Analysis and Formation Evaluation*. Second Edition. Society of Petroleum Engineers, 2012.
3. Darling, T. *Well Logging and Formation Evaluation.* Gulf Professional Publishing, 2005.
4. Bessereau, G., Carpentier, B., Huc, A. Y. Wireline logging and source rocks – estimation of organic carbon by the Carbolog method. *The Log Analyst*, 1991, **32**(3), 279–297.
5. El Sharawy, M. S., Gaafar, G. R. Application of well log analysis for source rock evaluation in the Duwi Formation, Southern Gulf of Suez, Egypt. *J. Appl. Geophys*., 2012, **80**, 129–143.
6. Lukaszyk, S. A new concept of probability metric and its applications in approximation of scattered data sets. *Comput. Mech.*, 2004, **33**(4), 299–304.
7. Fertl, W. H., Rieke, H. H. Gamma ray spectral evaluation techniques identify fractured shale reservoirs and source-rock characteristics. *J. Petrol. Technol.*, 1980, **32**(11), 2053–2062.
8. Fertl, W. H., Chilingar, G. V. Total organic carbon content determined from well logs. *Soc. Pet. Engineers (SPE) Formation Evaluation*, 1988, 3, 407–419.
9. Passey, Q. R., Creaney, S., Kulla, J. B., Moretti, F. J., Stroud, J. D. A practical model for organic richness from porosity and resistivity logs. *AAPG Bull*., 1990, **74**(12), 1777–1794.
10. Liu, J. M., Peng, P. A., Huang, K. Q., Zhang, L. Y. An improvement in CARBOLOG technique and its preliminary application to evaluating organic carbon content of source rocks. *Geochimica*, 2008, **37**(6), 581–586 (in Chinese).
11. Chen, G. H., Dong, W. W. Application of genetic neural network to coal quality evaluation based on log data. *Well Logging Technology*, 2011, **35**(2), 171–175.
12. Yang, Q. S., Torres-Verdín, C. Joint stochastic interpretation of conventional well logs acquired in hydrocarbon-bearing shale. In: *SPWLA 54th Annual Logging Symposium, 2013,* June 22–26, 15 pp.
13. Freed, D. *Molecular composition from diffusion or relaxation measurements*. U.S. Patent Application Publication, 2004, US 2004/0253743 A1, Published Dec. 16, 2004.

14. Heaton, N., Freedman, R. *Method for determining molecular properties of hydrocarbon mixtures from NMR data*. U.S. Patent 6,859,032, February 22, 2005.
15. Freedman, R. New approach for solving inverse problems encountered in well-logging and geophysical applications. *Petrophysics*, 2006, **47**, 93–111.
16. Anand, V. Freedman, R. New methods for predicting properties of live oils from NMR. In: *SPWLA 50th Annual Logging Symposium Transactions*, June 21–24, 2009.
17. Huang K. Y., Shen, L. C., Weng, L. S. Well log data inversion using radial basis function network. In: *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2011, 4439–4442.
18. Powell, M. J. D. *Radial Basis Functions for Multivariate Interpolation: A Review*. Technical Report DAMPT 1985/NA 12, Dept. of App. Math. and Theor. Physics, Cambridge University, Cambridge, England, 1985.
19. Moody, J., Darken, C. Learning with localized receptive fields. In: *Proceedings of the 1988 Connectionist Models Summer School* (D. Touretzky, G. Hinton, T. Sejnowski, eds.), Morgan Kaufmann, San Mateo, 1988, 133–143.
20. Haykin, S. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, Hamilton, Ontario, Canada, 1999.

Presented by A. Raukas
Received March 6, 2013