

# Research Article

# **Evaluation of Coalbed Methane Content by Using Kernel Extreme Learning Machine and Geophysical Logging Data**

Jianhong Guo<sup>1</sup>, Zhansong Zhang<sup>1</sup>, Guangshan Guo<sup>2</sup>, Hang Xiao<sup>1</sup>, Linqi Zhu<sup>3</sup>, Chaomo Zhang<sup>1</sup>, Xiao Tang<sup>4</sup>, Xueqing Zhou<sup>3</sup>, Yanan Zhang<sup>1</sup>, and Can Wang<sup>5</sup>

<sup>1</sup>Key Laboratory of Exploration Technologies for Oil and Gas Resources, Ministry of Education, Yangtze University, Wuhan 430100, China

<sup>2</sup>CNOOC Research Institute, Beijing 100028, China

<sup>3</sup>Institute of Deep-Sea Science and Engineering, Chinese Academy of Sciences, Sanya 572000, China

<sup>4</sup>Northwest Sichuan Gas Mine of PetroChina Southwest Oil and Gas Field Company, Sichuan, Jiangyou 621709, China

<sup>5</sup>Hubei Institute of Hydrogeology and Engineering Geology, Jingzhou 434020, China

Correspondence should be addressed to Jianhong Guo; 87942024@qq.com and Zhansong Zhang; zhangzhs@yangtzeu.edu.cn

Received 24 June 2022; Revised 22 August 2022; Accepted 15 September 2022; Published 8 October 2022

Academic Editor: Qianyou Wang

Copyright © 2022 Jianhong Guo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Accurate evaluation of coalbed methane (CBM) content plays a momentous role in the identification and efficient development of favorable exploitation blocks of CBM resources, but there are still many technical challenges in the exploration and development of onshore CBM fields. With the development and application of geophysical logging technology, using geophysical logging data to predict the gas content of CBM reservoirs has been proven to be an effective and feasible solution. However, the complex logging response of the CBM reservoirs makes it difficult to characterize the relationship between the gas content and the logging curve response by a simple linear relationship. In this paper, kernel extreme learning machine (KELM), a machine learning method, is combined with the geophysical logging data to predict the vertical variation curve of gas content in CBM wells. In this paper, the laboratory data on coal rock gas content from 12 CBM wells in the Southern Shizhuang block are selected, and a CBM content prediction model based on the KELM method is constructed by selecting the log curves, combining cross-validation and grid-seeking to determine the hyperparameters, and validating the prediction model using the test dataset and a new well in the same block. The application of the model on the test dataset was remarkable, and the vertical variation of CBM content obtained by applying it to the new well was consistent with the laboratory results, which proved the correctness and generalizability of the model. The results of this paper show that the CBM content evaluation model based on the KELM method and geophysical logging data is applicable to the 3<sup>#</sup> coal seam in the target block and can be used to predict the vertical CBM content of CBM wells; compared with the extreme learning machine (ELM) method and the backpropagation neural network (BPNN) method, the KELM method requires fewer hyperparameters to be explored when constructing the CBM content evaluation model, and the model construction is simple and has high prediction accuracy. At the same time, the CBM content model constructed by the KELM method differs for different blocks, coal seams at different depths, and different response ranges of geophysical logging data. The construction of a CBM content prediction model using the KELM method and logging curves is an effective means of characterizing CBM resources, and the model construction process and evaluation criteria studied in this paper can be used to help other blocks evaluate the CBM content, providing guidance for further exploration and development of CBM fields with practical application.

# 1. Introduction

The exploration and development of the coalbed methane (CBM) resources have been widely recognized internation-

ally for their ability to reduce the safety risks of coal mining and reduce the greenhouse effect, and research into CBM exploration and development has been undertaken in several countries [1–4]. China has abundant CBM resources, and the Ordos and Qinshui Basins are already in the commercial development stage [5, 6]. The assessment of CBM resources is an important part of CBM exploration and development and has become a hot topic of research due to the uncertainty of its resource assessment [7]. The CBM resources are influenced by several factors, including the buried depth, thickness, and gas content of CBM reservoirs [8–10]. The accurate evaluation of the CBM content is of paramount significance in determining favorable exploration blocks and in formulating production and development plans.

As an unconventional oil and gas resource, CBM is more complex than other reservoirs due to its storage and seepage mechanism [11]. The gas content of CBM reservoirs is related to geological factors such as the degree of coal rank metamorphism, temperature, pressure, effective depth of burial, thickness, tectonic features, and hydrogeological features [12-15]. Many methods have been proposed internationally for the evaluation of the gas content of coalbeds. Early on, Kim [16] combined the coal quality analysis with reservoir pressure and temperature for the adsorption gas content calculation and proposed an improved form; Ahmed et al. [17] and Hawkins et al. [18] established the isothermal adsorption models and the Langmuir coal order equation calculation methods, respectively, to perform gas content prediction. These techniques, which were created using the experimental data, could be evaluated for the boreholes that had the coring done, but they were impractical to apply for the CBM wells without coring samples and were ineffective at determining the vertical trend of CBM content. Subsequently, the geophysical logging data has been widely used in CBM content prediction studies [19, 20]. The geophysical logging techniques are cost-effective and reliable, and the high vertical continuity of the logging data allows for high-resolution characterization of the subsurface vertical physical properties of the borehole [21]. The geophysical logging data can be used to effectively evaluate the vertical distribution of CBM content in single wells and can be extended to wells without coring experiments, which is of great practical importance.

At present, more scholars have used the geophysical logging data to evaluate the CBM content. Shao et al. [22] used the logging volume model to evaluate the CBM content, and Jin et al. [23] used the background value method to calculate the gas content of CBM reservoirs, both of which achieved certain results, but the choice of parameters in both methods had a large impact on the results and the generalization of such methods was poor, which could only be used for a single well or single layer evaluation. Subsequently, Zhou and Guan [24] used the geophysical logging data and the core industrial fraction to construct an evaluation model and proposed an industrial fraction-based gas content method for the CBM reservoirs, which proved to be feasible, but when evaluating the gas content through intermediate parameters, the relationship between the parameters was not simply linear, and it was difficult to take into account substitution errors. As a result, it is now more common to match the coal-rock laboratory gas content with the geophysical logging data, combined with mathematical methods to construct a prediction model for CBM content. Guo et al. [25]

based on combining a grey multivariate static model with the logging curves, which can continuously and accurately evaluate the gas content curve of the whole coalbed. Liang and Yuan [26] used the geophysical logging data to establish a multivariate regression equation to predict the gas content and used it as the basis for gas content prediction for the corresponding blocks, and the results matched the geological conditions; Huang et al. [27] combined the multivariate linear regression method with the Langmuir equation for the Qinshui Basin to establish a CBM content evaluation model, and the results were highly accurate and valid. When linear relationship is difficult to characterize the relationship between the CBM content and the logging curves, predictions can be made using methods such as machine learning [28, 29], which has a strong nonlinear approximation capability and is dominated by neural network methods, which have been studied by many scholars, where the feature parameters and target parameters are trained through neural networks to form a grid model, and the test dataset is tested for generalizability to evaluate the usefulness of the model and the above methods. For example, the CBM content and the logging curve data were trained by the backpropagation neural network (BPNN) method, and this method was later validated on other wells in the block and found to be highly accurate in predicting the CBM content [30, 31]; more algorithms such as the Support Vector Machine (SVM) method [32, 33] and Random Forest (RF) method [34] were subsequently introduced into the CBM content prediction.

Although there are many methods for evaluating and predicting the CBM content, there is no denying that the complexity of CBM reservoirs is large and the corresponding logging response of CBM reservoirs is also more complex. Although scholars have tried and achieved good results using neural network-like methods with strong nonlinear approximation capability, in actual industrial production, with the increase of experimental samples, the evaluation model should also be updated accordingly. Neural networks and extreme learning machine (ELM) methods are randomly generated initial weights and make the models have a certain degree of randomness, there are many hyperparameters, and the constructed models are vulnerable to human interference. The innovation of this paper is to use the kernel extreme learning machine (KELM) method in combination with the geophysical logging data to construct a model for predicting the CBM content. The KELM method makes the model construction method simple by introducing the Gaussian Radial Basis Function (RBF), with a single hyperparameter and fast construction, and the model is repeatable, which can meet the needs of industrial production to update the model due to the increase of experimental samples, as shown from the available literature and work area data. The KELM method has not yet been applied to the construction of models for predicting the gas content of CBM reservoirs.

Based on the above idea, this paper introduces the KELM method for the evaluation and prediction of the CBM content. The KELM method is used to construct a CBM content prediction model by determining the curves



FIGURE 1: Distribution of coal structure and macrocoal component in the 3<sup>#</sup> coal seam: (a) coal structure; (b) macrocoal component.

to be used in the construction of the model and investigating the effect of different data ratios on the applicability of the model. The KELM method is found to be simpler and more accurate than the BPNN method and the ELM method. It is concluded that the nuclear limit learner model based on geophysical logging data is effective in predicting the gas content of coal seams, which can provide effective guidance for subsequent development and has practical application.

#### 2. Geological Overview and Data Sources

2.1. Geological Overview. The Southern Shizhuang block is one of the more highly explored CBM blocks in China. The southern part of the Qinshui Basin regionally belongs to the southern end of the Qinshui Complex Oblique, which is a monoclinic structure overall. The Qinshui Basin has experienced multiple phases of tectonic movement [35], and the block has an east-west zoned tectonic pattern, with an overall high southeast and low northwest tectonic pattern. The Upper Carboniferous Taiyuan Group (C3t) and the Lower Permian Shanxi Group (P1s) are the main coalbearing seam systems in the area, of which the 3<sup>#</sup> coal seam of the Shanxi Group is the target coal seam for development in the area [36] and is also the target coal seam for this study. The 3<sup>#</sup> coal seam has a stable thickness distribution, mainly in the range of 4.0 m to 8.0 m, with an average of 6.0 m. The coal rock has a high degree of maturity, with the maximum vitrinite reflectance of 2.5% to 3.0%. In the 3<sup>#</sup> coal seam of the target block, the coal structure of the seam can be divided into undeformed coal, cataclastic coal, and granulated coal, among which undeformed coal and cataclastic coal are predominant (Figure 1(a)). From the macrocoal component, semibright coal and semidull coal are dominant, and dull coal has the lowest content (Figure 1(b)).

According to the observation of the collected core coal samples, the cataclastic coal primarily appears as centimeter-level lumps, the undeformed structural coal primarily appears as long columnar and columnar, and some of the undeformed structural coal is broken to a significant degree, along with centimeter-level lumps. The granulated coal primarily appears as powder, along with a small amount of centimeter-level lumps. Combined with the information of macrocoal components, the undeformed coal and the cataclastic coal are dominated by bright coal, followed by mirror coal. The fracture of the coal seam is the key factor to control the permeability and the gas content of CBM reservoirs. As for the fracture development, the exact fracture development cannot be obtained because the coal structure of granulated coal is severely broken. It is known from the fracture description of the undeformed coal and the cataclastic coal samples that the coal samples typically develop one group of fractures or two orthogonal groups of fractures, and the variability of fracture density is large; in one group of fractures, the fracture density may range from 4 to 15 fractures for each 5 cm, and in two groups of fractures, the density of primary fractures may range from 8 to 20 fractures for each 5 cm and 14 to 20 fractures for each 5 cm in the second group. Based on the aforementioned view, the complexity of fracture development and the variety of coal structures also point to the ambiguity of the CBM contents.

2.2. Data Source of CBM Content. For the target block, 12 CBM wells were selected as research boreholes in this paper. Desorption experiments were performed on core coal samples to obtain data on the CBM content, and these samples ranged from 400 to 1000 m from the wellhead. As an example, 12 sets of core desorption samples from the  $3^{\#}$  coal seam in well A6 are shown. The core coal samples were sent to the laboratory from the core confinement tank, and the measurement time, interval time, the reading of the measurement tube, and the volume of gas were recorded, combined with the field temperature and field air pressure for calibration. Figure 2(a) shows the cumulative desorption curves of 12 samples in well A6, which were calibrated, and then, the residual gas content was measured and the lost gas content was determined, and the lost gas content was calculated as Figure 2(b).



FIGURE 2: Continued.



FIGURE 2: Desorption and CBM content calculation of coal core samples: (a) desorption of coal core samples; (b) calculation of the lost gas content.

The collected CBM content data were combined with the volume of coal samples to determine the ratio of adsorbed gas content, lost gas content, and measured residual gas content, as shown in Figure 3. The CBM content results of 12 groups of coal samples from the  $3^{\#}$  coal seam of well A6 are shown in Table 1.

2.3. Preprocessing of Geophysical Logging Data and Experimental Data. In this paper, a total of 12 core boreholes were extracted for the CBM content data, and data preprocessing and log data matching were performed. The geophysical logging data collected from the target block parameter wells were mainly eight geophysical logs, namely, spontaneous potential log (SP), natural gamma log (GR), caliper log (CAL), compensation density log (DEN), acoustic time difference log (AC), compensated neutron log (CNL), and dual lateral resistivity (deep lateral resistivity log (LLD)/shallow lateral resistivity log (LLS)) logging curves,

and some holes contain the flushed-zone resistivity log (RXO) series curves. The preprocessing steps are as follows:

- Depth corrections were made to prevent the effects of drill pipe deformation and stretching on sample depth during the drilling process [37]
- (2) Comparison of geophysical logging response gaps in the dense layer above the 3<sup>#</sup> coal seam from different wells and standardization of logging responses to prevent response differences due to differences between borehole environments and instruments
- (3) Dilation correction of log data. For reservoirs with poor mechanical strength, such as coal seams, varying degrees of dilation during drilling can make the response values of log data abnormal

Along with the data preprocessing work, the samples must be checked and rejected for abnormalities. For

Sample No.A6-3-12	94.04			3.86		10
Sample No.A6-3-11	94.72			3.60		.68
Sample No.A6-3-10	95.04			2.80		16
Sample No.A6-3-9	94.24			3.49		27
Sample No.A6-3-8	94.43		3.98		1	.59
Sample No.A6-3-7	93.64		4.57		1	.79
Sample No.A6-3-6	93.60			4.36		04
Sample No.A6-3-5	96.31			2.09		.60
Sample No.A6-3-4	95.53		2.81		1	.66
Sample No.A6-3-3	93.52		4	4.97		.51
Sample No.A6-3-2	92.64		6.00			1.36
Sample No.A6-3-1	91.46 6.21		2.3	3		
80 100						
Percentage of each type of gas in the coalned methane content (%)						

Percentage of measured residual gas content Percentage of lost gas content

Percentage of adsorbed gas content

FIGURE 3: Proportional distribution of each type of gas content in each sample of well A6.

Sample	Desorption sample mass (g)	Desorption gas content (cm <sup>3</sup> )	Lost gas content (cm <sup>3</sup> )	Residual sample mass (g)	Measured residual gas content (cm <sup>3</sup> )	Vg (cm <sup>3</sup> ·g <sup>-1</sup> )
A6-3-1	1100	12366.44	839.52	400	124.31	12.32
A6-3-2	1100	23380.27	1514.3	400	124.37	22.94
A6-3-3	1250	26864.66	1428.6	400	139.2	22.98
A6-3-4	1500	27967.59	821.99	400	129.26	19.52
A6-3-5	1200	24247.22	526.38	400	134.32	20.98
A6-3-6	1100	20070.15	934.76	400	159.09	19.49
A6-3-7	1000	18825.87	918.46	400	144.2	20.1
A6-3-8	1450	28953.87	1220.8	400	134.25	21.15
A6-3-9	1400	23171.6	858.72	400	159.12	17.56
A6-3- 10	1350	20648.98	607.49	400	139.13	16.09
A6-3- 11	1100	21552.73	818.82	400	139.28	20.69
A6-3- 12	1100	18365.28	754.25	400	149.14	17.75

TABLE 1: CBM content results of 12 sets of core coal samples from the  $3^{\#}$  coal seam in well A6.

example, (1) for samples that do not meet the experiment, such as samples that have been lifted and drilled for longer than the specified time, resulting in samples that have been tested incorrectly such as lost gas content overload, (2) and the noncoal reservoir samples in the gangue section, the gangue is mudstone or carbonaceous mudstone, which usually has an obvious high GR value with a high DEN value (such as sample point 6 in Figure 4(a) as an example), and such points have a prominent response on the rendezvous map, which is also demonstrated in Figure 4(b).



FIGURE 4: Rejection of core samples from the gangue section: (a) logging response trend of the sample at the gangue; (b) analysis of the difference between the logging response of anomalies on the rendezvous plot.

Figure 5 shows the standardized comparison of the porosity series logging curves. The left coordinate of the figure corresponds to the standardized curves (suffixed with \_AS, AS means after standardized), and the AC, DEN, and CNL curves are shown before and after the standardization. The variability of the distribution of the standardized samples with respect to the original samples is not excessive, indicating that there is little human interference and that the apparently high and low values in the

original samples due to the environment and instrumentation are corrected.

After completing the above process, a total of 151 sets of sample data were harvested for the  $3^{\#}$  coal seam CBM content study for subsequent construction of the prediction model, the specific value distribution is shown in Figure 6, and the gas content of the core coal samples from the Southern Shizhuang block is distributed between 4.55 and 26.13 cm<sup>3</sup>/g, with the gas content mainly in the range of 5 to 20 cm<sup>3</sup>/g.



FIGURE 5: Continued.



FIGURE 5: Standardized comparison of logging curves in porosity series: (a) AC curve, (b) DEN curve, and (c) CNL curve.

#### 3. Methods

3.1. Extreme Learning Machine Method. The ELM algorithm was proposed in 2006 [38], which is a new algorithm for single hidden layer feedforward neural networks (Figure 7), mainly to improve the backpropagation algorithm training complexity and low efficiency and to reduce the type of hyperparameters. General backpropagation neural networks such as BPNN input layer to the hidden layer and hidden layer to output layer weights and thresholds need to be constantly changed to achieve the best model results; its gradient descent-based iterative method converges slowly (closely related to the initial weights and thresholds) and is prone to fall into the local optimum trap.

While ELM can obtain the optimal output weights given random weights from the input layer to the output layer and a threshold of neurons in the hidden layer [39], the ELM algorithm is based on the following principle:

For a set of *n* training samples,

$$N = \left\{ \left( x_i, y_i \right) \middle| i = 1, 2, \cdots, n ; x_i \in \mathbb{R}^a, y_i \in \mathbb{R}^b \right\}.$$
(1)

In Equation (1), a is the number of features of the input sample and b is the dimensionality of the output value. An ELM neural network model with n (number of samples) neurons in the input layer, m neurons in the output layer, and l neurons in the hidden layer can be constructed, and this model can be expressed as

$$H\beta = Y.$$
 (2)

In Equation (2),  $Y = [y_1, y_2, \dots, y_n]^T$  is the output of the network and *H* is the output matrix of the hidden layer of

the neural network. In practice, the number of neurons in the hidden layer is usually not set too large, then l < n [38], when  $\beta$  can be solved by the least-squares method, and the solution is

$$\beta = H^+ Y. \tag{3}$$

In Equation (3),  $H^+$  is the Moore-Penrose generalized inverse of the implied layer output matrix H. The orthogonal projection method can be used effectively to calculate. When  $H^TH$  is a nonsingular matrix,  $H^+ = (H^TH)^{-1}H^T$ , or  $HH^T$  is a nonsingular matrix,  $H^+ = H^T(HH^T)^{-1}$ . On this basis, based on the ridge regression theory [40], Huang et al. [41] argue that the solution obtained has better stability and generalization by introducing the regularization factor  $1/\lambda$  into  $H^TH$  or  $HH^T$  when calculating the ELM output weights  $\beta$ . With the introduction of the regularization factor, the output of the ELM is

$$f(x) = h(x)\beta = h(x)H^{T}\left(\frac{I}{\lambda} + HH^{T}\right)^{-1}Y.$$
 (4)

3.2. Kernel Extreme Learning Machine Method. Based on the extreme learning machine, Huang et al. [41] also studied the kernel-based extreme learning machine, which was named KELM (kernel extreme learning machine) to distinguish it from the general extreme learning machine. When the implicit layer activation function is not known, one may choose to use the kernel function as the mapping relation, in which case the KELM kernel matrix is defined as

$$\Omega_{\text{KELM}} = HH^T, \tag{5}$$



FIGURE 6: Experimental distribution trend of coalbed methane content. This figure counts and plots the distribution range of CBM content in the target block.

$$\Omega_{\text{KELM}(i,j)} = h(x_i) \cdot h(x_j) = K(x_i, x_j).$$
(6)

With the introduction of the kernel function, the output of KELM is

$$f(x) = h(x)H^{T}\left(\frac{I}{\lambda} + HH^{T}\right)^{-1}Y = \begin{bmatrix} K(x, x_{1}) \\ K(x, x_{2}) \\ \vdots \\ K(x, x_{n}) \end{bmatrix}^{T}\left(\frac{I}{\lambda} + \Omega_{\text{KELM}}\right)Y.$$
(7)

When using the KELM, there is no need to know the implicit layer activation function; instead, the mapping tool uses the kernel function. The use of kernel functions enables mapping relations in infinite dimensions without the "dimensional catastrophe" of computational complexity. Because the kernel function calculation process does not require knowledge of the specific functional form of the sample data mapped from low to high dimensions, only the specific kernel function form is calculated in the low dimensional space; the kernel function used in this paper is a Gaussian radial basis (RBF) kernel function of the following form:

$$K(x_i, x_j) = \exp\left\{-\frac{|x_i - x|^2}{\sigma^2}\right\}.$$
(8)

In Equation (8),  $\sigma$  is the kernel function parameter, which is mapped as shown in Figure 8.

It should be noted that when the kernel function is introduced into the extreme learning machine, the number of neurons in the hidden layer does not need to be artificially given, nor do the connection weights from the input layer to the hidden layer, and the threshold of neurons in the hidden layer needs to be given, and the KELM only needs to select the



FIGURE 7: Schematic diagram of the parameter calculation method of the extreme learning machine.

appropriate kernel function. This paper will use the KELM for the model construction of CBM content, written and implemented through the MATLAB software platform.

#### 4. Construction of CBM Content Model

In this paper, 151 sets of data from 12 CBM wells are used for the construction of the KELM model, and MATLAB software is used to conduct the training, test, and validation processes in the KELM model. The workflow of the KELM method for evaluating the CBM content is shown in Figure 9 and consists of four sections: data preference normalization, data partitioning, model optimization, and model validation.

4.1. Geophysical Logging Curve Selection. Normally, different geophysical logging responses often represent changes in the physical properties of the rock, and changes in the CBM



FIGURE 8: The kernel function is used to map the low dimensional feature vector to the high dimensional feature space.



FIGURE 9: Workflow of the CBM content prediction based on the KELM method.

content can theoretically be characterized by changes in the logging response. Considering the practicality of constructing the model, the logging curves that were available for all wells were selected and the effective depth of burial was obtained by converting the elevation of the coal rock samples by combining the complement height with the logging depth, and the rendezvous diagrams shown in Figure 10 were drawn for demonstration.

In the logging series characterizing lithology, the natural radioactivity of the coal itself is weak, the natural radioactivity of the coal depends on the clay minerals present in the coal, the clay minerals affect the CBM content by influencing the adsorption properties of the coal, and an increase in clay minerals in the coal reduces the CBM content, while a higher presence of clay minerals in the coalbed leads to an increase in natural radioactivity. The higher the natural gamma value of the coalbed in the logging curve, the less effective pore space in the coalbed, resulting in lower CBM content, a trend that is met by Figure 10(a). In addition, this paper extracted a positive correlation (Figure 11(a)) based on the clay mineral content, which was found to be consistent with the theory by meeting the GR curve. In addition, the SP curve and CAL curve are used more for lithology identification, the SP curve response is related to the nature of the mud filtrate (Figure 10(b)), and the borehole diameter of 21.59 cm can explain the poor mechanical properties of the coalbed susceptible to dilation (Figure 10(c)), which is also corrected for dilation in this paper during preprocessing.



FIGURE 10: Continued.



FIGURE 10: Continued.



FIGURE 10: Continued.



FIGURE 10: Continued.



FIGURE 10: Cross plots of geophysical logging data and Vg: (a) Vg with GR, (b) Vg with SP, (c) Vg with CAL, (d) Vg with DEN, (e) Vg with AC, (f) Vg with CNL, (g) Vg with LLD, (h) Vg with LLS, and (i) Vg with DEP.

In the series of logs characterizing porosity, density values increase with increasing densities in the density log response values, with theoretical increases in densities corresponding to decreases in porosity and CBM content (Figure 10(d)). A significant correlation was found between laboratory apparent density and air-dried basis ash (Figure 11(b)). Conversely, the less dense the coalbed, the looser it is, the higher the CBM content, and the higher the inorganic mineral content such as ash, the more pores and fractures are filled, which is not conducive to methane storage and adsorption, which explains the trend in the response of CBM content to the DEN curve [42]. The higher the content of the higher-order coal specular group, the easier it is to fracture, and the fracture of the coal structure will increase the adsorption area of the CBM, which will increase the CBM content, and the acoustic time difference logging response is sensitive to the CBM content, coupled with the presence of large fractures in the coal rock, which will also increase the AC curve response (Figure 10(e)); the CNL curve is influenced by the coal rock skeleton and gas content multiple parties; in coalbed, the actual porosity of the reservoir is usually low, generally less than 10%, but because the coalbed is composed of carbon, hydrogen, and oxygen and the CBM contains methane, resulting in a high hydrogen content index, compensated neutron logging presents a false high illusion; after compensation, the compensated neutron logging response in this block is negatively correlated with the CBM content (Figure 10(f)).

In the logging series characterizing resistivity, resistivity is influenced by the degree of coal metamorphism, mineral content and distribution, coal structure, and more factors [43] in addition to the CBM content. The trend of change in gas content and resistivity is extremely complex and is more based on the data extracted from the actual workings for statistical purposes [44]. In this paper, statistics for the target block concluded that the resistivity change pattern indicated the relationship between the gas content and the resistivity of the original coalbed was synergetic growth (Figures 10(g) and 10(h)).

The depth of the coalbed determines whether the gas produced by coalification can be preserved. Theoretically, as the depth of the coalbed increases, the degree of coalification and the amount of hydrocarbon production increase, and the corresponding CBM content should increase accordingly [45], but after the depth reaches a certain critical depth, the trend of increasing gas content disappears due to geological and tectonic factors [37]. The effective depth of burial in the target block has a large influence on the CBM content. The depth of burial of the 3<sup>#</sup> coal seam is shallow, and the target boreholes are all located near 1000 m above sea level, and the core depth of coal samples are all located above sea level, corresponding to the effective depth of burial being closer to the sea level, which means the distance from the wellhead is farther, as shown in Figure 10(i), and the gas content is enriched as the depth of burial increases.

In addition to the analysis using the rendezvous diagrams and response principle, the relationship between the logging curves and the CBM content was also quantitatively analyzed, and the Pearson linear correlation index P, Kendall rank correlation coefficient R, and Spearman rank correlation coefficient S were used to quantify the correlation between each logging parameter and the gas content as one of the criteria for curve preference. The three correlation coefficients are calculated as

$$P = \frac{N\sum_{i=1}^{N} x_i y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{\sqrt{N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i^2\right)}} \sqrt{N\sum_{i=1}^{N} y_i^2 - \left(\sum_{i=1}^{N} y_i^2\right)}, \quad (9)$$



FIGURE 11: Cross plot experimental data: (a) relationship between clay mineral content and GR; (b) relationship between ash content and laboratory apparent density of samples under air drying condition.

$$R = \frac{4A}{N(N-1)} - 1,$$
 (10)

$$S = \frac{\sum_{i=1}^{N} (x_i - \bar{x}) (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2 \sum_{i=1}^{N} (y_i - \bar{y})^2}}.$$
 (11)

In Equations (9)–(11), x is for each logging parameter, y is the experimental value of CBM content, i corresponds to the serial number of each sample, and N is the total number of samples for statistics, which is 151 in this paper. A is the logarithm of statistical objects with a consistent size relationship between two attribute values,  $\bar{x}$  is the mean value of the logging curve, and  $\bar{y}$  is the mean value of the CBM content.

In this paper, the correlation coefficients of each logging curve and CBM content were plotted (Figure 12), and the correlation results of the three coefficients were calculated with the same trend. Among the correlations with the CBM content, the five logging curves, including the GR curve, the AC curve, the DEN curve, the AC curve, and the depth curve, have the most significant correlations with the CBM content. The two log curves, the CNL curve, and the LLD curve have slightly weaker correlations with the CBM content. Combined with the actual response mechanism, the correlation between the two log curves, the SP curve, the CAL curve, and the CBM content is the weakest. The above is consistent with the results of the previous analysis.







FIGURE 13: Schematic diagram of cross-validation.

The response of CBM content to geophysical logging data cannot usually be described by a simple linear equation, as illustrated by the analysis of rendezvous plots and quantitative calculations of correlations. Of the resistivity logging series, the LLD logging curve was chosen because it characterizes the resistivity of the original formation, i.e., compared to the LLS series, the LLD is unaffected by the intrusion. In conclusion, six curves, AC, DEN, CNL, GR, depth, and LLD, will be used to construct the CBM content, while logarithmization of the LLD curve is recommended.

4.2. *Kernel Function Search.* The KELM method was used to construct the CBM content model by simply optimizing the kernel function in the model, while the effect of different data scales on the model was investigated to determine the

optimal data scale. Before constructing the KELM method model, due to the large differences in the values of the data collected from different logging series, to avoid a large impact on the prediction accuracy and training speed, the built-in mapminmax function of MATLAB was used to normalize the input logging data to the range of [0, 1], respectively. The normalization equation is

$$Y = \frac{x - x_{\min}}{x_{\max} - x_{\min}}.$$
 (12)

In Equation (12), Y is the standardized logging data, x is the original logging data, and  $x_{min}$  and  $x_{max}$  are the minimum and maximum values in the logging data sequence, respectively.



FIGURE 14: Optimization process of kernel function under different data proportions.

 TABLE 2: Kernel function optimization results and error statistics of

 KELM method under various data ratios.

Datia	Kernel function	Average relative error (%)		
(training:testing)	σ	Training dataset	Testing dataset	Error variation (%)
5:5	6.1	13.72	18.71	4.99
6:4	9.1	14.67	17.95	3.28
7:3	3.9	14.49	15.99	1.5
8:2	5.6	14.54	17.81	3.27
9:1	3.2	12.06	18.63	6.57

<sup>a</sup>Error variation = mean relative error difference between testing dataset and training dataset.

After normalizing the geophysical logging response values, the collected gas-content data were first divided into a training dataset and a test dataset at different scales, and cross-validation was introduced into the model construction to avoid chance in the exploration results [46]. The basic principle is the following: the training dataset is divided into k subsets, each subset in turn as test data, the remaining k-1 subsets as training data, repeated k times to get k models, using the mean square error as the model evaluation method, can get k groups of prediction accuracy, the final k-fold cross-validation results are the average value of kgroups of prediction accuracy, as shown in Figure 13. In this paper, the preprocessed data are set in different ratios, and the data ratios of the training dataset and test dataset are set as 5:5, 6:4, 7:3, 8:2, and 9:1, respectively; the optimal kernel function is then determined in the training using 3-

fold cross-validation combined with grid search, and the average mean square error (MSE) calculated from the cross-validation is used as the evaluation index to complete the kernel function selection, and finally, the parametric model is applied to the training and test datasets. Figure 14 shows the optimization process of the kernel function under each data ratio, the step size of this paper is 0.1, and combined with the actual calculation results, it shows that under different data ratios, with the increase of the value of the kernel function, the cross-validation score decreases rapidly and then increases gradually, and the kernel function corresponding to the lowest value of the average MSE is determined by calculation as the optimal kernel function value. When the proportion of data in the training set exceeds 70%, there may be a drop in the correctness of the randomly sorted test dataset. Table 2 shows the results of the search for merit and the effect of the KELM method on the model construction results by the size of the training data, where the ratio of 8:2 to 9:1 is extracted from the set where the correctness of the test dataset is affected. Analysis of Table 1 shows that the training accuracy is relatively low when the proportion of training data is small, due to the fact that for the whole sample space, the smaller data samples limit the generalized learning ability of KELM, and the error keeps decreasing as the proportion of training data increases. When the proportion of data in the training set exceeds 70%, there is a risk that the relative error gap between the training and test datasets will increase. Two explanations are given for this phenomenon: (i) the training dataset is too large and overfitting occurs, making the test dataset less accurate and the model of no practical application at this time; (ii) it may also be caused by the number of data in the test dataset being compressed



FIGURE 15: Application effect of CBM content model constructed by KELM method: (a) application effect of training dataset; (b) application effect of test dataset.

in the same sample space, which is more likely to be magnified by the quality of the small sample data in the test dataset being constantly compressed, and the data in the test dataset at this time can hardly cover the whole data distribution range, and the correctness of the test dataset cannot effectively characterize the performance of the model due to the chance of random binning of the data. Therefore, in order to effectively demonstrate the effectiveness of the method in this paper, a training data ratio of 7:3 is used to construct the CBM content evaluation model more conducive to the demonstration of the KELM model performance. When using machine learning methods, such as the KELM method used in this paper, there are certain limitations. Similar to the multiple regression approach, the adjustment of the data can have an impact on the overall network structure, just like the slight change of the coefficients in the multiple regression approach. Similarly, the change in data volume will also have an impact on the model. As we know from the previous exploration of the data ratio, the larger the training data share is in the same sample space, the more fully the model learns, but it is not conducive to the testing of the test dataset. When the proportion of training data exceeds 70%, the model is stable



FIGURE 16: Evaluation results of CBM content of well A1.



FIGURE 17: Application of CBM content model constructed by KELM method in a new well.

and the values of the kernel function become stable. The model constructed in this paper is only for the 3<sup>#</sup> coal seam in the Southern Shizhuang block, and it is no longer applicable in different geological regions or when the sources of gasbearing data vary greatly, and the prediction model must be reconstructed with the actual data. Similarly, when predicting the gas content of the development wells, the anomalies in the logging curves affect the accuracy of data normalization and thus the application effect, i.e., the model is not applicable when the response range of the logging is too different from that of the core wells after preprocessing.

There is still room for improving the accuracy of the model in this paper. The hyperparametric search mode can

be improved. This paper uses a stable grid search mode, but it is possible to reduce the search step or increase the optimization search mode, but this will also cost more time. All of the above is the next step in the research of this paper, and there is still a need for breakthroughs in the face of the challenges of onshore CBM exploration.

#### 5. Results and Discussion

5.1. *Model Evaluation*. Based on the comprehensive analysis in Section 3.2, the data ratio between the training and test datasets in the KELM model was determined to be 7:3, and the optimal kernel function was 3.9. The determined



FIGURE 18: Performance of three methods on the same testing dataset.

data proportions were combined with the optimal kernel function for the training dataset back judging and test dataset testing, and the results are shown in Figure 15. Figure 15(a) shows the back-judgment results for the training data, and Figure 15(b) shows the results of the application of the test dataset, with an average relative error of 14.49% and 15.99%, respectively; the training dataset and the test dataset are evenly distributed on both sides of the zero-error line, which means the model application effect is unbiased, and the goodness of fit is 0.83 and 0.82, respectively, indicating the effectiveness of the model constructed in this paper.

Figure 16 shows well A1 in the study block, with the first track being the log depth track; the second track being the log series reflecting lithology, containing CAL, GR, and SP curves; the third track being the resistivity log series, containing LLD, LLS, and RXO curves; the fourth track being the porosity log series, containing DEN, AC, and CNL curves; the fifth track being the prediction result track, containing gas content curves evaluated by the KELM method and core experimental data; and the sixth track being the lithology interpretation track. The predicted curves in well A1 mostly coincide with the core experimental data, and the prediction effect is relatively weak when the core experimental data is low, at which time the response of GR and DEN logging increases, corresponding to the gas content curve in decreasing trend when the mud content and porosity correspondingly decrease, although there is a certain difference in the value but ensures the consistency of the changing trend, as the low-value area is not the main development block, and the gas content curve of the coal seam predicted by the KELM method is indicative, which also shows the applicability of the model in the 3<sup>#</sup> coal seam of the target block.

5.2. Generalizability Evaluation and Error Discussion. The demonstration in Section 5.1 demonstrates the validity and reliability of the model training by evaluating the generalization of the gas content prediction model by evaluating a new well in the same workings, the data from which was not involved in the training and testing of the model. The CBM content curve predicted by the CBM content model based on the KELM method is shown in Figure 17. There are seven core samples for desorption evaluation of CBM content in the  $3^{\#}$  coal seam of this well, with an average relative error of 27.39%. The difference between the predicted and actual results at 646.6 m is too large, with a relative error of 87.79%. Without considering this sample, the average relative error of the remaining six samples is 17.3%, and the error is consistent with the level of the test dataset, indicating that the established KELM model is generalized in the 3<sup>#</sup> coal seam of the target block.

The model application for the new well was analyzed for sample points with large errors, such as the first sample (at 646.6 m) shown from top to bottom in Figure 11, which is the top sampling point of the  $3^{\#}$  coal seam. The corresponding GR and porosity log series response values are high compared to the pure coalbed section, and therefore, the predicted gas content in this section is low. The second sample (at 647.55 m) is located at the spreading point, and when combined with the resistivity series logging response, the intrusion has been observed, and even with the spreading correction, noise has been introduced, making the judgment error larger.

This paper discusses and explains the generation of errors:



FIGURE 19: Performance of the three methods in the same dataset: (a) rendezvous plot of predicted results and laboratory results; (b) cumulative frequency diagram of relative errors of the three methods.

- (1) Field disturbances during the acquisition of logging data during the actual drilling of the well
- (2) It is difficult to avoid human-induced noise in the subsequent preprocessing of logging data, such as standardization of geophysical logging data response

between wells and dilation correction, which is beneficial to the subsequent model construction but difficult to avoid small human-induced errors

(3) The process of the CBM content collection in the laboratory cannot avoid systematic errors, which

TABLE 3: Error results of three methods on the same testing dataset.

Method	KELM	ELM	BPNN
Average relative error (%)	15.99	18.75	21.44
RMSE (cm <sup>3</sup> ·g <sup>-1</sup> )	2.07	2.39	2.87
$R^2$	0.82	0.79	0.77

can usually be accepted. The samples that do not comply with the experimental rules would be rejected from the data preprocessing. In addition, this systematic error can change under different laboratory conditions or when there are large differences in the sources of CBM content data. When the laboratory standards are less different, the obtained experimental data of the CBM content still have some reference value. It should be noted that with the development of industry, especially for the measurement of lost gas content, improved experimental methods may be applied to the measurement of CBM content. In Figure 3, the composition of the coalbed methane content has been shown. The percentage of lost gas content is low and much lower than the percentage of adsorbed gas content. The impact on the experimental improvement depends on the error of the lost gas content in the sample determination. When the difference in values is large, the error in the prediction of the CBM content also increases, and whether it is still valid for the actual project depends on the degree of increase in the error. Similarly, in the actual work block, some core boreholes are developed after the operating block has been in production for a longer period of time. The coal samples from such boreholes usually yield low results when the CBM content is determined in the laboratory, and such data are not informative and cannot be used as supplementary data to participate in modeling or for model validation

(4) In terms of application, as the KELM method requires data normalization in data preprocessing, the geophysical logging data response should be renormalized when it exceeds the response range of the training sample, which also indicates that the geophysical logging data response cannot be standardized in the face of different lateral spreading blocks, different depths of the formation, and different geological backgrounds, and the CBM content prediction model must be reconstructed according to the model construction process

The noise disturbance described for points 1 to 3 is difficult to be removed after separate reconstruction, and the noise in the input data can be disregarded, but the noise will be amplified in the output results. This noise error exists in the CBM content evaluated using the method in this paper, which cannot be attributed entirely to the predictive capability of the KELM method. The analysis of the evaluation of the CBM content in the new well demonstrates the generalizability of the method, while the combination of the error results shows the validity and reliability of the KELM model.

5.3. Comparison of Methods. In Section 3.2, it is shown that the KELM method is based on the ELM method and is similar to the BPNN method. Therefore, this paper compares these three methods, using the same ratio (training dataset: test dataset = 7:3) of the same dataset to train the ELM model and BPNN model, and the models of the three methods are applied to the same test dataset for comparison. It can be analyzed in Figure 18 that some of the samples in the test dataset (e.g., samples 6, 7, and 9) correspond to experimental values with low results, and the three methods discriminate some of the samples and give the correct trend, but there are also errors that are obvious. The predicted results of the three methods were analyzed with the experimental results in the rendezvous plot (Figure 19(a)); for the low-value part of the gas content, some of the predicted results of the three methods fell outside the error line of +15%, and since the experimental values are low, according to the definition of relative error, the relative error of the experimental low values is higher, which is the main source of relative error above 30% in the three methods (Figure 19(b)). For this type of experimental low values, it is not a "sweet spot" section, and a guideline can be given when there is a certain trend difference with the high gas content section. In addition, in the test dataset, sample numbers 37 and 38 had low predicted results, corresponding to such samples in the rendezvous plot falling outside the error line of -15%; such points are mostly affected by the gangue section of the coalbed; although the samples with high ash yield content in the gangue section were rejected in the preliminary data rejection, the gangue on the closer samples also have an impact; the GR curve and the DEN curve values are high, making the prediction results slightly lower than the experimental results; the three methods give a consistent indication that such errors can be artificially discriminated in practical applications, and the predicted results of the intercalated gangue section and its vicinity are not involved in the final statistics. By analyzing the cumulative frequency of the relative errors of the three methods in Figure 19(b), the KELM method has the most samples when the relative errors are within 10%, and the maximum relative errors of the KELM method are lower than the other two methods, which is the reason why the KELM method has the lowest relative errors. Table 3 shows the specific calculations of the mean relative error, root mean square error (RMSE), and the goodness of fit  $(R^2)$  for the combined discrimination. The analysis shows that the KELM method has the lowest average relative error, ELM the second-highest, BPNN the highest average relative error, and the corresponding RMSE trend are the same, while for the goodness of fit, KELM is the highest, ELM the second-lowest, and BPNN the lowest. Among the three methods, KELM is more suitable; on the one hand, its accuracy is outstanding; on the other hand, KELM only needs to find the optimal kernel function when building the model, compared to ELM and BPNN methods, which are relatively complex, and the ELM method and BPNN method have instability due to the initial weights randomly given, so that the KELM method has application advantages in terms of accuracy and ease of operation. It should be noted that the ELM method





(b)

FIGURE 20: (a) Contour map of CBM content of the  $3^{\#}$  coal seam in Southern Shizhuang block (eastern area); (b) contour map of the average effective daily gas production from the  $3^{\#}$  coal seam in the Southern Shizhuang block (eastern area).

Average CBM content (cm<sup>3</sup>·g<sup>-1</sup>)



FIGURE 21: Intersection of CBM content and average effective daily gas production.

and the BPNN method do not guarantee the full performance of the model due to a large number of hyperparameters at the time of construction. Similarly, in practical production applications, the ELM method and the BP neural network method are more cumbersome to update the model as the number of laboratory gas contents of the coal samples in the study block increases, and the advantages of the KELM method in terms of easy model construction and fewer hyperparameters will be more obvious.

5.4. Application and Prospect. By evaluating the CBM content wells in the Southern Shizhuang block, the CBM content contour map is plotted in Figure 20(a), corresponding to the average effective daily gas production contour map of the corresponding block in Figure 20(b). The two plots in Figure 20 show a certain correlation, and the high gas production area often corresponds to the high gas-bearing area, which can indicate that accurate CBM content prediction can provide subsequent construction and development providing guidance. However, the vertical variation of CBM well gas content is complex, and the two-dimensional contour plot has information loss after averaging. Taking the CBM wells in Figure 20 as an example, the stable CBM production data with a long drainage period and output are selected, the average daily production of CBM during the effective production period is calculated, and the rendezvous plot is drawn with the CBM content of the corresponding wells, as shown in Figure 21. It can be analyzed from Figure 21 that the level of CBM content can reflect the average effective daily gas production, especially since the average effective daily gas production of wells with lower CBM content levels is also low, but the  $R^2$  of both is only 0.5 by linear fitting, which also indicates that the gas production of CBM wells is affected by various factors, including fracture modification and human factors such as construction [47]. The exploration and development of onshore CBM need to be explored more deeply, and the contents of this paper provide a new way to predict the CBM content, which provides guidance for the subsequent exploration and development and favorable area identification of the block.

#### 6. Conclusions

Aimed at the difficulty of gas content evaluation of CBM reservoirs, a method of vertical CBM content prediction based on the KELM method combined with geophysical logging data is proposed in this paper. The conclusions of this study are as follows:

- (1) The CBM content evaluation model based on the KELM method and geophysical logging data is established for the 3<sup>#</sup> coal seam in the Southern Shizhuang block of the Qinshui Basin as an example, and the validity and generalization of the model are verified using the test dataset and the new well, demonstrating the applicability of the model in the target block formation. The KELM method, ELM method, and BPNN method were compared by the same set of data. In terms of prediction effect, the KELM method was less disturbed by human factors in the model construction, and the model stability was higher, which was more advantageous in practical use
- (2) The KELM method also has limitations in its use, being affected by data variation and having room for improvement in accuracy, which can be improved by introducing modules and combining hyperparameter search methods in the future

The research content of this paper provides a set of the model construction process and evaluation criteria for the CBM content, and the constructed model cannot be directly applied to other blocks. For different coal seams and different geological blocks, the CBM content evaluation model must be constructed based on the actual core data obtained in conjunction with the process steps in this paper.

#### Nomenclature

CBM:	Coalbed methane
KELM:	Kernel extreme learning machine
ELM:	Extreme learning machine
BPNN:	Backpropagation neural network
RBF:	Gaussian Radial Basis Function
Vg:	Coalbed methane content
AC:	Acoustic time difference log
DEN:	Compensation density log
CNL:	Compensated neutron log
CAL:	Caliper log
SP:	Spontaneous potential log
GR:	Nature gamma log
LLD:	Deep lateral resistivity log
LLS:	Shallow lateral resistivity log
RXO:	Flushed-zone resistivity log
MSE:	Mean square error
RMSE:	Root mean square error
$R^2$ :	Goodness of fit.

# **Data Availability**

The original data used to support the findings of this study are currently under embargo while the research findings are commercialized. Requests for data, 3 months after publication of this article, will be considered by the corresponding author. Data requests can be made by contacting the corresponding author and can be provided upon request as appropriate.

# **Conflicts of Interest**

The authors declare that they have no conflicts of interest.

# Acknowledgments

The authors wish to express their gratitude to CNOOC Research Institute. Also, we sincerely thank Prof. Mianmo Meng and Prof. Qianyou Wang. This work was carried out in the Southern Shizhuang block of the Qinshui Basin and financially sponsored by the Open Fund of Key Laboratory of Exploration Technologies for Oil and Gas Resources, Ministry of Education (Nos. K2021–03 and K2021-08), National Natural Science Foundation of China (No. 42106213), and Natural Science Foundation of Hainan Province of China (No. 421QN281).

# References

 J. Williams, M. Jones, L. Ma, and M. Pourkashanian, "Pollutants from the combustion of solid biomass fuels," *Progress in Energy and Combustion Science*, vol. 38, no. 2, pp. 113–137, 2012.

- [2] T. A. Moore, "Coalbed methane: a review," *International Journal of Coal Geology*, vol. 101, pp. 36–81, 2012.
- [3] T. Hamawand, T. Yusaf, and S. G. Hamawand, "Coal seam gas and associated water: a review paper," *Renewable & Sustainable Energy Reviews*, vol. 22, pp. 550–560, 2013.
- [4] I. Palmer, "Coalbed methane completions: a world view," *International Journal of Coal Geology*, vol. 82, no. 3–4, pp. 184–195, 2010.
- [5] Q. Zhao, "Geological features of the coalbed methane in China and its new exploration domains," *Natural Gas Industry*, vol. 24, no. 5, pp. 4–7, 2004.
- [6] Y. Meng, D. Tang, H. Xu, Y. Li, and L. Gao, "Coalbed methane produced water in China: status and environmental issues," *Environmental Science and Pollution Research*, vol. 21, no. 11, pp. 6964–6974, 2014.
- [7] D. Liu, Y. Yao, D. Tang, S. Tang, Y. Che, and W. Huang, "Coal reservoir characteristics and coalbed methane resource assessment in Huainan and Huaibei coalfields, Southern North China," *International Journal of Coal Geology*, vol. 79, no. 3, pp. 97–112, 2009.
- [8] Y. Gao, M. T. Ma, and G. Ma, "Main controlling factors analysis of enrichment condition of coalbed methane," *Acta Sedimentologica Sinica*, vol. 21, no. 2, pp. 345–349, 2003.
- [9] Y. Wu, Z. Li, and Y. Wang, "The formation mechanisms of abnormal pressure and factor in control of the coal bed gas in Qinshui Basin," *Journal of China Coal Society*, vol. 31, no. 4, pp. 475–479, 2006.
- [10] J. Ye, Q. Wu, and Z. Wang, "Controlled characteristics of hydrogeological conditions on the coalbed methane migration and accumulation," *Journal of China Coal Society*, vol. 26, no. 5, pp. 63–67, 2001.
- [11] J. C. Pashin, M. R. McIntyre-Redden, S. D. Mann, D. C. Kopaska-Merkel, M. Varonka, and W. Orem, "Relationships between water and gas chemistry in mature coalbed methane reservoirs of the Black Warrior Basin," *International Journal* of *Coal Geology*, vol. 126, pp. 92–105, 2014.
- [12] B. Hildenbrand, M. Krooss, A. Busch, and R. Gaschnitz, "Evolution of methane sorption capacity of coal seams as a function of burial history – a case study from the Campine Basin, NE Belgium," *International Journal of Coal Geology*, vol. 66, no. 3, pp. 179–203, 2006.
- [13] S. Hou, X. Wang, X. Wang, Y. Yuan, X. Zhuang, and X. Wang, "Geological controls on gas saturation in the Yanchuannan Coalbed Methane Field, Southeastern Ordos Basin, China," *Marine and Petroleum Geology*, vol. 78, pp. 254–270, 2016.
- [14] J. C. Hower and R. A. Gayer, "Mechanisms of coal metamorphism: case studies from Paleozoic coalfields," *International Journal of Coal Geology*, vol. 50, no. 1–4, pp. 215–245, 2002.
- [15] T. Yan, Y. Yao, and D. Liu, "Critical tectonic events and their geological controls on gas generation, migration, and accumulation in the Weibei coalbed methane field, southeast Ordos basin," *Journal of Natural Gas Science and Engineering*, vol. 27, pp. 1367–1380, 2015.
- [16] G. Kim, Estimating Methane Content of Bituminous Coal Beds from Adsorption Data, vol. 8245, United States Department of the Interior, Report of Investigations-Bureau of Mines, 1977.
- [17] U. Ahmed, D. Johnston, and L. Colson, "An advanced and integrated approach to coal formation evaluation," in SPE Annual Technical Conference and Exhibition, Dallas, Texas, October 1991.

- [18] J. M. Hawkins, R. A. Schraufnagel, and A. J. Olszewsk, "Estimating coal bed gas content and sorption isotherm using well log data," in *SPE Annual Technical Conference and Exhibition*, Washington, D.C., October 1992.
- [19] S. Ghosh, R. Chatterjee, S. Paul, and P. Shanker, "Designing of plug-in for estimation of coal proximate parameters using statistical analysis and coal seam correlation," *Fuel*, vol. 134, pp. 63–73, 2014.
- [20] R. H. Morin, "Hydrologic properties of coal beds in the Powder River Basin, Montana I. Geophysical log analysis," *Journal* of Hydrology, vol. 308, no. 1-4, pp. 227–241, 2005.
- [21] A. Roslin and J. S. Esterle, "Electrofacies analysis using highresolution wireline geophysical data as a proxy for inertiniterich coal distribution in Late Permian Coal Seams, Bowen Basin," *International Journal of Coal Geology*, vol. 152, pp. 10–18, 2015.
- [22] X. Shao, Y. Sun, J. Sun et al., "Log interpretation for coal petrologic parameters: a case study of Hancheng mining area, Central China," *Petroleum Exploration and Development*, vol. 40, no. 5, pp. 599–605, 2013.
- [23] Z. Jin, H. Xue, H. Gao, X. Zhao, and G. Bai, "Technology for evaluation of CBM reservoir logging and its application," *Coal Geology & Exploration*, vol. 41, no. 2, pp. 42–45, 2013.
- [24] F. Zhou and Z. Guan, "Uncertainty in estimation of coalbed methane resources by geological modelling," *Journal of Natural Gas Science and Engineering*, vol. 33, pp. 988–1001, 2016.
- [25] J. Guo, Z. Zhang, C. Zhang et al., "Prediction and application of CBM content based on gray system and logging method," *Geophysical and Geochemical Exploration*, vol. 44, no. 5, pp. 1190–1200, 2020.
- [26] Y. Liang and W. Yuan, "The prediction of the content and distribution of coalbed gas: a case study in the Qinshui coalfield based on logging," *Geophysical and Geochemical Exploration*, vol. 42, no. 6, pp. 1144–1149, 2018.
- [27] Z. Huang, C. Zou, Y. Yang, G. Zhang, and W. Wang, "Coal bed methane reservoir evaluation from wireline logs in TS District, southern Qinshui Basin," *Geoscience*, vol. 26, no. 6, pp. 1275– 1282, 2012.
- [28] B. Zhou and G. O'Brien, "Improving coal quality estimation through multiple geophysical log analysis," *International Journal of Coal Geology*, vol. 167, pp. 75–92, 2016.
- [29] H. Liu, S. Sang, G. G. X. Wang, Y. Li, M. Li, and S. Liu, "Evaluation of the synergetic gas-enrichment and higherpermeability regions for coalbed methane recovery with a fuzzy model," *Energy*, vol. 39, no. 1, pp. 426–439, 2012.
- [30] S. Deng, Y. Hu, D. Chen, Z. Ma, and H. Li, "Integrated petrophysical log evaluation for coalbed methane in the Hancheng area, China," *Journal of Geophysics and Engineering*, vol. 10, no. 3, article 035009, 2013.
- [31] H. Pan and Z. Huang, "Discussion on the interpretation method of coalbed methane content," *Coal Geology&Exploration*, vol. 26, no. 2, pp. 58–60, 1998.
- [32] C. Wu and X. Yue, "Neural network of coal bed gas logging evaluation," *Natural Gas Exploration&Development*, vol. 24, no. 1, pp. 31–34, 2001.
- [33] C. Lian, Y. Zhao, H. Li, F. Qu, F. Cai, and J. Zhang, "Prediction of coal bed gas content based on support vector machine regression," *Journal Center of Xi'an University of Science and Technology*, vol. 28, no. 4, pp. 707–709, 2008.

- [34] J. Yu, L. Zhu, R. Qin, Z. Zhang, L. Li, and T. Huang, "Combining K-means clustering and random forest to evaluate the gas content of coalbed bed methane reservoirs," *Geofluids*, vol. 2021, Article ID 9321565, 8 pages, 2021.
- [35] Y. Yi, Research on the Production Characteristics and Drainage Control of CBM Wells in the Shizhuangnan Block of Qinshui Basin, China University of Geosciences (Beijing), Beijing, 2020.
- [36] G. Guo and L. Xing, "Primary geological controlling factors of coalbed methane enrichment and its exploration potential in Southern Shizhuang block," *Clean Coal Technology*, vol. 21, no. 4, pp. 117–121, 2015.
- [37] X. Fu, Y. Qin, G. G. X. Wang, and V. Rudolph, "Evaluation of gas content of coalbed methane reservoirs with the aid of geophysical logging technology," *Fuel*, vol. 88, no. 11, pp. 2269– 2277, 2009.
- [38] G. Huang, Q. Zhu, and C. K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006.
- [39] S. Ding, H. Zhao, Y. Zhang, X. Xu, and R. Nie, "Extreme learning machine: algorithm, theory and applications," *Artificial Intelligence Review*, vol. 44, no. 1, pp. 103–115, 2015.
- [40] E. Hoerl and R. W. Kennard, "Ridge regression: biased estimation for nonorthogonal problems," *Technometrics*, vol. 42, no. 1, pp. 80–86, 2000.
- [41] G. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems Man and Cybernetics Part B-Cybernetics*, vol. 42, no. 2, pp. 513–529, 2012.
- [42] P. Ren, H. Xu, D. Tang et al., "The identification of coal texture in different rank coal reservoirs by using geophysical logging data in northwest Guizhou, China: investigation by principal component analysis," *Fuel*, vol. 230, pp. 258–265, 2018.
- [43] Z. Liu, W. Wang, J. Yang et al., "Review and prospect of study on conductive properties of coal and CBM reservoirs," *Progress in Geophysics*, vol. 35, no. 4, pp. 1415–1423, 2020.
- [44] L. Zhu, Y. Ma, J. Cai, C. Zhang, S. Wu, and X. Zhou, "Key factors of marine shale conductivity in southern China-part II: the influence of pore system and the development direction of shale gas saturation models," *Journal of Petroleum Science and Engineering*, vol. 209, article 109516, 2022.
- [45] S. Zhang and X. Chen, "Geological controls on content and producibility of coal seam gas," *Natural Gas Industry*, vol. 17, no. 4, pp. 15–19, 1997.
- [46] Z. Mahmood and S. Khan, "On the use of K-fold crossvalidation to choose cutoff values and assess the performance of predictive models in stepwise regression," *International Journal of Biostatistics*, vol. 5, no. 1, 2009.
- [47] Z. Guo, J. Zhao, Z. You, Y. Li, S. Zhang, and Y. Chen, "Prediction of coalbed methane production based on deep learning," *Fuel*, vol. 230, article 120847, 2021.