Prediction of Coalbed Methane Production Based on BP Neural Network

Tianxiang Zhang¹, Yifang Tang²*, Jianjun Wu¹, Zixi Guo¹, Bing Li¹

¹PETROCHINA CBM Institute Of Engineering Technology, Xi’an, Shanxi, 710082, China
²School of Sciences, Southwest Petroleum University, Chengdu, Sichuan, 610500, China
³State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Southwest Petroleum University, Chengdu, Sichuan, 610500, China

Abstract: The low average daily gas production per well and the poor economic benefit of exploration and development have become the main problems restricting the exploration and development of coalbed methane in China. Combining multiple coal seam geological parameters to predict the high-yield area of the block can not only provide guidance for the exploitation of coal-bed methane, but also bring enormous economic benefits. Aiming at the difficulty of coalbed methane dessert discrimination and production prediction, a method of coalbed methane production prediction based on BP neural network is proposed in this paper. Starting from the average daily production of coalbed methane single well, we use the method of grey correlation degree to get the main controlling factors of coalbed methane production. For the main control factors, we use BP neural network with high fitting accuracy and get a good prediction result.

1 Introduction

Coal-bed methane industry has accumulated a lot of geological and engineering data. Field data had the characteristics of large amount of data and many types of information [1]. On the other hand, the data itself had missing values, inconsistent dimensions and irrelevant redundant information. Many factors are weakly correlated with CBM (coalbed methane) production [2], which indicates that the causal relationship between these factors and CBM production is very complex. Comprehensive analysis of geological factors and engineering data to study CBM production prediction is not only conducive to the integration of geological engineering, but also can effectively reduce development costs and bring good economic benefits.

For the main control factors, the main research methods are semi-quantitative research methods [3] such as statistics, comparison and numerical simulation, which are not combined with actual production data. At the same time, the logging series of CBM blocks in China are simple. CBM evaluation only evaluates the parameters such as coal petrology industrial composition and gas content [4], does not carry out in-depth "dessert" evaluation of CBM, and lacks the use of logging information.

Based on BP neural network, this paper presents a single well daily production prediction method for CBM. Aiming at the defect of main control factors diagnosis and production prediction method of CBM production at present, this paper takes the average daily gas production of different single wells as the research target, and calculates the main controlling factors of CBM production. Finally, we use logging data to predict the gas production of CBM wells, and the prediction result is satisfactory.

2 Range normalization

When using raw data, the units and dimensions of each original data are often different [5], and the numerical size and range of variation of each data are different, too. In order to overcome these unreasonable factors, this paper uses range normalization method to plan the values of all original data between 0 and 1. The formula of range normalization is as follows:

$$x = \frac{r_i - \min_{\text{min}} r_i}{\max_{\text{max}} r_i - \min_{\text{min}} r_i} \quad (i = 1, 2, \cdots, n)$$ (1)

2.1. Gray correlation degree analysis

Grey relational degree analysis is a method of grey system analysis and processing of random variables [6], as well as a mapping of data to data. Firstly, we quantify the qualitative data, supplement the incomplete data and normalize the range. Let the data after pretreatment of the selected parameters be $X_i$, and the correlation

*Corresponding Author: Yifang Tang; email: tangyifang_swpu@163.com; phone:15828532038

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
between data \( X_j \) and target data \( X_i \) is \( \gamma_i(k) \). Among them, \( k \) represent the sampling points of the correlation between data \( X_j \) and data \( X_i \). The total number of sampling points is \( n \), then the correlation coefficient can be expressed as follows:

\[
\gamma_i(k) = \frac{\alpha_{\min} + \alpha_{\max} \rho}{\alpha_{ij}(k) + \alpha_{\max} \rho} \quad (k = 1, 2, \ldots, n) \tag{2}
\]

In the formula

\[
\alpha_i(k) = |x_i(k) - x_j(k)|, \quad \alpha_{\min} = \min_{i} \alpha_i(k), \quad \alpha_{\max} = \max_{i} \alpha_i(k)
\]

\( \rho \) is a given constant between 0 and 1, which is usually taken as 0.5 in application.

Thus, the relevance degree \( \gamma_{ij} \) is:

\[
\gamma_{ij} = \frac{1}{n} \sum_{k=1}^{n} \gamma_i(k) \tag{3}
\]

Relevance degree \( \gamma_{ij} \) reflects the degree of correlation between data \( X_j \) and data \( X_i \), and describes the relative changes between data \( X_j \) and data \( X_i \) during the development of the system. By using the ranking order of correlation degree, the influence degree of each parameter on CBM production can be grasped, which lays a foundation for production prediction.

3 Results

3.1. BP neural network

The process of BP neural network can be divided into two stages [7]. The first stage is the forward propagation of the signal, which passes through the hidden layer from the input layer to the output layer. The second stage is the reverse propagation of errors, from the output layer to the hidden layer, and finally to the input layer. In this process, the weights and biases from the hidden layer to the output layer and from the input layer to the hidden layer are adjusted in turn. Next, we begin to construct the whole BP neural network.

(1) Network Initialization

Assuming that the number of nodes in the input layer is \( n \), the number of nodes in the hidden layer is \( l \), the number of nodes in the output layer is \( m \), the weight from input layer to hidden layer is \( \omega_{ji} \), the weight from hidden layer to output layer is \( \omega_{jk} \), the bias from the input layer to the hidden layer is \( \beta_{j} \), the bias from the hidden layer to the output layer is \( \beta_{k} \), the learning rate is \( \eta \) and the excitation function is \( g(x) \).

(2) Output of Hidden Layer

\[
H_j = g \left( \sum_{i=1}^{n} \omega_{ji} x_i + \beta_{j} \right) \tag{4}
\]

(3) Output of Output Layer

\[
t_k = \sum_{j=1}^{l} H_j \omega_{jk} + \beta_{k} \tag{5}
\]

(4) Error calculation

\[
E = \frac{1}{m} \sum_{k=1}^{m} |Y_k - t_k| / Y_k \tag{6}
\]

The above error is the average relative error, where \( Y_k \) is the expected output, \( t_k \) is the actual output.

(5) Updating of Weights

The updating formula of the weights is as follows:

\[
\omega_{ji} = \omega_{ji} + \eta \frac{\partial E}{\partial b_{ji}} \tag{7}
\]

(6) Update of bias

Bias Update from Input Layer to Hidden Layer

\[
\frac{\partial E}{\partial b_{ji}} = (Y_j - t_j) \left( - \frac{\partial H_j}{\partial b_{ji}} \right) = -e_{ji} \tag{8}
\]

Bias Update from Hidden Layer to Output Layer

\[
\frac{\partial E}{\partial \beta_{ji}} = (Y_k - t_k) \left( - \frac{\partial H_k}{\partial \beta_{ji}} \right) = -e_{ji} \tag{9}
\]

Therefore, the updated formula of bias is:

\[
\left\{ \begin{array}{l}
\beta_{j} = \beta_{j} + \eta H_j (1 - H_j) \sum_{k=1}^{m} \omega_{jk} e_{k} \\
\beta_{k} = \beta_{k} + \eta e_{k}
\end{array} \right. \tag{10}
\]
In the above formulas, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, l$; $k = 1, \ldots, m$.

(7) Termination Judgment of Iterative Algorithms

There are many methods to judge whether the algorithm has converged, the common ones are the specified iteration algebra, whether the difference between adjacent two errors is less than the specified value, and so on. The method selected in this paper is to specify the number of iterations. When fitting and predicting with neural network, the number of iterations is 1000 times.

### Table 1. Influencing factors of CBM production

| Parameter Type               | Parameter Name                                                                 |
|------------------------------|-------------------------------------------------------------------------------|
| Geological Factors           | Coal Seam Thickness, Permeability, Porosity, Fracture Pressure, Temperature, Reservoir Pressure, Gas Content |
| Drilling Factors             | Coal Seam Borehole Diameter, Well Construction Period, Drilling Fluid Viscosity, Cement Dosage, Cement Slurry Density |
| Fracturing Factors           | Perforation Thickness, Pre-fluid Volume, Sand Addition, Average Sand Ratio, Total Fluid Volume, Displacement, Oil Pressure |
| Drainage and Production      | Gas Seeking Time, Sheathing Pressure, Drainage Speed, Average Drainage Volume |

First of all, we do data preprocessing to remove abnormal data. Then, the range normalization method is used to plan the values of all the original data between 0 and 1. After that, thirty-three factors affecting CBM production were analyzed by grey correlation degree, and sixteen factors with smaller correlation degree were eliminated. Finally, we get seven main factors that affect the production of CBM. The ranking results of the correlation degree of all factors are as follows:

### Table 2. Relevance ranking results diagram of all factors

| Rank | Parameter Name       | Relevance   | Rank | Parameter Name       | Relevance   |
|------|----------------------|-------------|------|----------------------|-------------|
| 1    | Coal seam thickness  | 0.965764185 | 12   | Sand Addition        | 0.755439345 |
| 2    | Caliper              | 0.968737075 | 13   | Sheathing Pressure   | 0.734497478 |
| 3    | Porosity             | 0.967845103 | 14   | Gas Seeking Time     | 0.712374948 |
| 4    | Gas Content          | 0.966730578 | 15   | Well Construction Period | 0.707058972 |
| 5    | Total Liquid volume  | 0.965928037 | 16   | Cement Dosage        | 0.699700754 |
| 6    | Permeability         | 0.96540494  | 17   | Reservoir Pressure   | 0.689337806 |
| 7    | Average Sand Ratio   | 0.965205188 | 18   | Average Drainage Volume | 0.659584359 |
| 8    | Displacement         | 0.837207129 | 19   | Pre-fluid Volume     | 0.654424481 |
| 9    | Fracture Pressure    | 0.796190106 | 20   | Cement Slurry Density | 0.626949731 |
| 10   | Sand Addition        | 0.762034883 | 21   | Drainage Speed       | 0.612740999 |
| 11   | Oil Pressure         | 0.75702013  | 22   | Temperature          | 0.603235829 |

According to the ranking result of grey relational degree, we finally get seven main controlling factors affecting CBM production in the figure above: coal seam thickness, borehole diameter, porosity, gas content, total liquid content, permeability, average sand ratio. In this way, we reduce the 23-dimensional influencing factors to 7-dimensional through grey correlation, and put these seven main controlling factors into the next stage of calculation and research.

In view of the 88 CBM wells with complete data finally screened from the CBM field data in Linfen area, wells with production less than 400 square meters are regarded as type III dessert, wells with production between 400 and 800 square meters are regarded as type II dessert, wells with production greater than 800 square meters are regarded as type I dessert.

We randomly selected 20 wells including all kinds of dessert as prediction, and built BP neural network to fit the original data of the remaining 68 wells. The number of iterations is set to 1000, the node of hidden layer is selected 12, the activation function of hidden layer is selected 'tansig', and the activation function of output layer is selected 'logsig'. The fitting effect is measured by average relative error. After repeated trial and error, a network with fitting accuracy of 94.06% is obtained. We get the weight of the hidden layer W1, the threshold of the hidden layer b1, the weight of the output layer W2 and the threshold of the output layer b2.

### 4 Discussion

#### 4.1. Prediction of CBM production by BP neural network

According to the field data of CBM in Linfen area, 23 factors affecting production are classified into four categories: geological factors, drilling factors, fracturing factors and drainage and production factors. The various factors involved in different types are listed in the following table (Table 1):
Because it is difficult to predict the value of CBM production within the specified precision range in the actual prediction of CBM production, the general method is classification prediction. That is to say, to divide the dessert types for predicting CBM wells. It can not only improve the prediction accuracy, but also meet the needs of engineering. According to the established network, the weights and thresholds of hidden layers are used for back calculation and prediction, and the output classification of 20 wells to be predicted is realized. The final prediction and classification results are shown in the following table (Table 3):

**Table 3. Output prediction result table of BP neural network**

| Well Number | Real dessert type | Prediction results | Well Number | Real dessert type | Prediction results |
|-------------|------------------|--------------------|-------------|------------------|--------------------|
| Well 1      | I                | I                  | Well 11     | II               | II                 |
| Well 2      | I                | III                | Well 12     | II               | II                 |
| Well 3      | I                | I                  | Well 13     | II               | II                 |
| Well 4      | I                | I                  | Well 14     | III              | III                |
| Well 5      | I                | I                  | Well 15     | III              | III                |
| Well 6      | II               | II                 | Well 16     | III              | III                |
| Well 7      | II               | II                 | Well 17     | III              | III                |
| Well 8      | II               | III                | Well 18     | III              | I                  |
| Well 9      | II               | II                 | Well 19     | III              | III                |
| Well 10     | II               | II                 | Well 20     | III              | III                |

From the above table, it can be seen that the accuracy of classification prediction of coalbed methane production by using BP neural network reaches 85%, and the prediction effect is satisfactory. This prediction method can meet the needs of prediction in practical engineering.

## 5 Conclusions

(1) CBM field data had a large dimension and contains many factors with little influence. In this paper, the dimension reduction of real data is processed, and the method of grey correlation is used to reduce the dimension of multi-dimensional factors on the basis of retaining most of the information of the original data.

(2) BP neural network has a strong fitting effect. By using this method, 20 CBM wells are classified and predicted, and the prediction accuracy reaches 85%.

## Acknowledgements

This research was financially supported by the National Science and technology Major Project of China (Grant No. 2016ZX05042-003, Grant No. 2016ZX05065).

## References

1. Yao, Y., Liu, D., Tang, D., Tang, S., Yao, C., & Huang, W. (2009). Preliminary evaluation of the coalbed methane production potential and its geological controls in the weibei coalfield, southeastern ordos basin, china. International Journal of Coal Geology., 78(1):1-15.

2. Weixu, W., Manjiang, H. E., Xiyou, W., Pei, J., Lisha, P., & Yue, D. U. (2017). Analysis on main controlling factors and comprehensive evaluation of coalbed methane production capacity of junlian block. Coal Science and Technology.

3. Caineng, Z., Zhi, Y., Rukai, Z., Songtao, W., Jinhua, F., & Dewen, L., et al. (2018). Geologic significance and optimization technique of sweet spots in unconventional shale systems. Journal of Asian Earth Sciences, S1367912018302670-.

4. Zhou, & Fengde. (2012). History matching and production prediction of a horizontal coalbed methane well. Journal of Petroleum Science and Engineering., 96-97(10): 22-36.
5. Singh, A., Panda, R. K., Pramanik, N., Cluckie, I. D., Chen, Y., & Babovic, V., et al. (2009). Appropriate data normalization range for daily river flow forecasting using an artificial neural network. Hydroinformatics in Hydrology, Hydrogeology & Water Resources Symposium 4 at the Joint Convention of the International Association of Hydrological Sciences.

6. Deng, A. (2012). The application of grey relevance degree in the structure analysis of grain production. Henan Science.

7. Sadeghi, B. H. M. (2000). A bp-neural network predictor model for plastic injection molding process. Journal of Materials Processing Technology., 103(3):411-416.