Prediction of gas emission based on grey-generalized regression neural network

Yanqiu Chen, Linjiang Zheng, Jing Huang, Zhe Zou and Chunhui Li

College of Computer Science, Chongqing University, Shazheng Street 174#, Shapingba District, Chongqing City, China, 400044

1Email: zlj_cqu@cqu.edu.cn

Abstract. Coal mine gas is one of the main factors endangering mine safety in China, and the amount of gas emission is affected by a variety of related factors, which have the characteristics of non-linearity and uncertainty. In order to predict gas emission more quickly and accurately, a coupling algorithm of grey theory (GM (1,1) and generalized regression neural network (GRNN) is proposed. Thirteen parameters, such as coal seam gas content, depth and GM (1,1) gas emission prediction value, are used as input of the model. The input parameters are normalized and used as training and testing samples of the model. The 10 fold cross validation and minimum root mean square error (RMSE) are used to find the optimal smoothing factor (spread), then a non-linear prediction model of gas emission is established. The gas monitoring data of Qianjiaying Mining Area of Kailuan Mining Group from May 2007 to December 2008 are used in the experiment. The results show that the model has a great improvement in convergence speed and a good accuracy, which can provide theoretical basis for the prevention and control of coal mine gas disasters.

1. Introduction

Coal mine gas is one of the main factors endangering mine safety in China. China's coal mine gas disasters are serious. From 2001 to 2010, there were 2246 coal mine gas disasters in China, causing heavy casualties [1]. Therefore, finding a model or method that can accurately predict the amount of coal mine gas emission is the key to prevent gas disasters [2]. In recent years, with the development of Internet of Things, wireless sensor, neural network, artificial intelligence and other technologies, prediction of gas emission has become a research hotspot of scholars at home and abroad [2,3,4,5,6,7]. Because the gas emission is influenced by many factors such as coal seam thickness and coal seam depth, and considering the highly non-linear characteristics of gas emission data distribution. Therefore, scholars introduce the non-linear prediction method to predict coal mine gas emission. In reference [8], GM (1,1) model in grey system is used to predict gas emission. In reference [3], BP neural network is used to predict gas emission. In reference [4] and reference [5], Kalman filter and principal component regression analysis are used respectively. However, with the deepening of the research, scholars have found that the single non-linear prediction method is affected by the prediction method itself, and the prediction effect is not good. Therefore, some scholars began to improve the prediction accuracy of the algorithm by coupling several prediction models. Reference [6] couples chaotic immune particle swarm optimization algorithm and Elman neural network algorithm, while reference [7] couples genetic simulated annealing algorithm and regression support vector machine (SVR). In reference [9], SFLA-Verhulst combined forecasting model was proposed for gas emission.
prediction with non-linear characteristics; In reference [10], empirical mode decomposition (EMD), modified Drosophila optimization algorithm (MFOA) and limit learning machine (ELM) were coupled to construct a multi-scale time-varying forecasting model of EMD-MFOA-ELM for gas emission prediction. The experimental results show that, compared with a simple algorithm, the coupled algorithm does have the advantage of improving the prediction accuracy, but in general, the convergence speed will be reduced. Fu H et al. [11] proposed a prediction model based on fractional order neural network to improve the convergence speed of BP neural network. The experiment has achieved good results. However, for scenarios requiring real-time prediction or quasi-real-time prediction, the convergence rate still cannot meet the requirements. Therefore, from the perspective of improving the convergence speed without reducing the prediction accuracy, this paper selects the gray prediction model widely used for uncertainty problems and the generalized regression prediction model without training iteration from many nonlinear prediction methods. The prediction model of gas emission based on grey-generalized regression neural network is established, and the applicability of the model in the scene with small sample size and high real-time requirement is proved by experiments with gas measured historical data.

2. Grey system theory

2.1. Grey system theory

Grey system theory was put forward by Professor Deng Julong in 1982, which is called grey theory for short. Grey theory points out that in the process of system research, there are various complex reasons and the influence of human factors, which make researchers unable to fully understand the system objects studied, and data collection is often only a few samples, so the research is uncertain [12]. There are various uncertainties in real life. Gas emission is an uncertain problem affected by many factors. To deal with the uncertain problem of large amount of data, it can be dealt with by statistical methods and probability theory knowledge. However, in view of the uncertainties of gas emission, such as small data, small samples and incomplete known information, such problems often have fuzzy uncertainties. Grey system theory is specially designed to solve the problem of uncertainty with small amount of data [12].

2.2. GM (1,1) model

GM (1,1) model is the most widely used grey prediction model. There is only one variable in the model and it is solved by first-order differential equation. Therefore, in the prediction of coal mine gas emission, according to the existing data information of coal mine gas emission, the irregular initial data sequence is accumulated once to produce a new sequence, thus increasing the regularity of the data, and then the new data sequence is solved by the first-order linear differential equation. The grey modeling steps are as follows:

(1) Set up the original gas data sequence as follows: \( x(t) = \{x(1), x(2), x(3), \cdots, x(n)\} \), accumulate it once, then we can get \( y(t) = \{y(1), y(2), y(3), \cdots, y(n)\} \), that is to say

\[
y(t) = \sum_{i=1}^{n} x(i), k = (1, 2, 3, \cdots, n)
\]

(2) Establishment of first order linear differential equations for

\[
y(t): \frac{dy(t)}{dt} + ay(t) = u
\]

The special solution of the equation is:

\[
Y(t) = [x(1) - u / a]e^{-at} + u / a
\]

\( Y(t) \) is the estimated value of \( y(t) \) sequence, \( a \) is the coefficient of development and \( u \) is the coefficient of coordination in formula (3). The least square method is used to estimate the parameter vectors. By calculating the matrix, it can be concluded that:
\[ a = \{(n-1)[\sum_{i=2}^n x_i z(t)] + [\sum_{i=2}^n z(t)][\sum_{i=2}^n x_i]/D \] 

\[ u = \{[\sum_{i=2}^n z(t)]/[\sum_{i=2}^n x_i z(t)] + [\sum_{i=2}^n x_i][\sum_{i=2}^n z^2(t)]/D \] 

(3) A cumulative reduction of \( Y(t) \) yields the predicted value of \( X(t) \).

\[ X(t) = Y(t) - Y(t-1) \] 

2.3. Grey relational analysis

Grey relational analysis is a method often used to analyze the relativity between two factors with different dimensions and different physical meanings, and the representation of relativity is usually expressed by grey relational degree. The greater the similarity between the two factors (size, direction, order), the greater the correlation between the two factors [13].

Because there are too many factors affecting gas emission, if all the factors are used in network training, the training time will be greatly increased, which is not conducive to improving the convergence speed of the model. Therefore, before training the model, it is necessary to calculate the grey correlation degree of the relevant factors of gas emission, in order to reduce the training time of the network.

The calculation method of grey correlation degree is as follows:

1. Finding the correlation number of each sequence \( \xi_1(k) \):

\[ \xi_1(k) = \frac{\min i \min k |X_i(k) - X_o(k)| + \rho \max i \max k |X_i(k) - X_o(k)|}{|X_o(k) - X_o(k)| + \rho \max i \max k |X_i(k) - X_o(k)|} \]  

(7)

Where \( \rho \) is the resolution coefficient, \( 0 < \rho < 1 \), generally \( \rho = 0.5 \).

2. Finding the relevance degree \( r_j \).

\[ r_j = \frac{1}{N} \sum_1^N \xi_j(k) \]  

(8)

In the formula, \( r_j \) is the correlation degree between sequence \( X_j \) and \( X_o \). The larger \( r_j \) is, the more correlated \( X_j \) and \( X_o \) are.

3. Generalized regression neural network

3.1. Generalized regression neural network

Generalized regression neural network (GRNN) is proposed by Donald F. Specht[14]. It is an improvement on radial basis function. It consists of four layers: input layer, mode layer, summation layer and output layer. The function in the mode layer is Gauss function, which has strong local approximation ability and does not need iterative training, so the convergence speed is faster. At the same time, there is only one smooth parameter in the which needs to be optimized, so it is less affected by human factors. Its structure is detailed in Figure 1:

![Figure 1. The network structure of GRNN.](image)
(1) Input Layer
The number of neurons in the input layer is n, which means the dimension of the input vector X in the training sample is n, and \( x_i \) is the input vector of the \( i \) dimension in the training sample which affects the amount of gas emission.

(2) Model layer
The model layer, also known as the hidden regression layer, has the same number of neurons as the input layer, which is the number of training samples n. The model layer neurons are fully connected with the input layer neurons, and the function is Gauss function. The transfer function of neuron \( i \) is:

\[
p_i = \exp\left[-\frac{(X-X_i)^T(X-X_i)}{2\sigma^2}\right], \quad (i = 1, 2, \ldots, n) \tag{9}
\]

(3) Summation layer
There are two types of neurons in the summation layer. The first one is to sum all the results of formula (9) and its transfer function is as follows:

\[
S_p = \sum_{i=1}^{n} p_i \tag{10}
\]

The connection weight between S and model layer neurons is 1. The second formula is

\[
\sum_{i=1}^{n} Y_i \exp\left[-\frac{(X-X_i)^T(X-X_i)}{2\sigma^2}\right]
\]

and its transfer function is:

\[
S_y = \sum_{i=1}^{n} Y_i p_i \tag{11}
\]

\( S_y \) uses the element \( y_i \) in the output sample \( Y \) of the pattern layer as the connection weight to realize the weighted summation of the output of the model layer.

(4) Output Layer
There is only one neuron in the output layer, which divides the two output values in the previous layer to get the output result.

\[
\hat{Y} = \frac{S_p}{S_y} \tag{12}
\]

3.2. Mathematical model of generalized neural network
GRNN is based on non-linear regression analysis, and the essence of regression analysis of random variable \( y \) relative to \( x \) is to find the \( y \) value of the maximum probability. Let \( f(x, y) \) be a joint density function of \( x \) and \( y \), and the observed value of \( x \) is \( X \) then the regression of \( y \) to \( X \) can be expressed as:

\[
\hat{Y} = E[y | X] = \frac{\int_{-\infty}^{\infty} yf(X, y)dy}{\int_{-\infty}^{\infty} f(X, y)dy} \tag{13}
\]

When \( f(x, y) \) is unknown, it can be estimated by observational values of \( x \) and \( y \), and its non-parametric estimation equation can be expressed as:

\[
f(X, Y) = \frac{1}{(2\pi)^{m+n}/2 \sigma^{m+n}} \sum_{i=1}^{n} \exp\left[-\frac{(X-X_i)^T(X-X_i)}{2\sigma^2}\right] \exp\left[-\frac{(Y-Y_i)^2}{2\sigma^2}\right] \tag{14}
\]

Among them, \( X_i, Y_i \) is the observation value of \( x \) and \( y \), \( \sigma \) is the smoothing parameter; \( n \) is the number of samples; \( m \) is the dimension of \( x \).

By substituting \( \hat{f}(x, y) \) for \( f(x, y) \) and substituting \( \hat{f}(x, y) \) for formula (13), and exchanging the order of integration and summation, we can get the result.
\[
\hat{Y}(X) = \frac{\sum_{i=1}^{n} \exp\left[-\frac{(X-X_i)^T (X-X_i)}{2\sigma^2}\right] \int_{-\infty}^{+\infty} y \exp\left[-\frac{(y-Y_i)^2}{2\sigma^2}\right] dy}{\sum_{i=1}^{n} \exp\left[-\frac{(X-X_i)^T (X-X_i)}{2\sigma^2}\right] \int_{-\infty}^{+\infty} \exp\left[-\frac{(y-Y_i)^2}{2\sigma^2}\right] dy}
\]

For the integral term in formula (15), if we simplify it by using property \( \int_{-\infty}^{+\infty} x \exp(-x^2) dx = 0 \), we can get:

\[
\hat{Y}(X) = \sum_{i=1}^{n} Y_i \exp\left[-\frac{(X-X_i)^T (X-X_i)}{2\sigma^2}\right] \times \left( \sum_{i=1}^{n} \exp\left[-\frac{(X-X_i)^T (X-X_i)}{2\sigma^2}\right] \right)^{-1}
\]

In formula (16), \( \hat{Y}(X) \) is the weighted average estimate of all samples \( Y_i \), and the weight of each sample value \( Y_i \) is the exponent of the square of the Euclid distance between the corresponding \( X_i \) and \( X \).

3.3. Smoothing factor selection

Formula (16) shows that there are no other factors affected by subjective factors in the model except the smoothing parameter \( \sigma \). Therefore, in order to ensure the prediction effect of the model, it is necessary to select \( \sigma \) accurately and adjust the size of \( \sigma \) to obtain the best regression estimation effect. In order to evaluate the merits of smoothing parameters, we usually use network performance indicators to evaluate, or through the establishment of optimization problem model to achieve the optimization of smoothing parameters [15]. In the prediction of gas emission, if \( \sigma \) is optimized by establishing a model, it will greatly increase the difficulty and workload of modeling. And through the network performance evaluation indicators, such as the method of estimating the mean square of the error to determine \( \sigma \), it has the advantages of simplicity and rapidity. The process of determining parameters is also the process of model building. In this paper, the method of estimating mean square error is used to determine the smoothing parameters of the network. The steps are as follows:

1. Setting the smoothing parameter \( \sigma \) to increase the change of the fixed increment in the range of its value;
2. Cross-validate the training samples and get the error sequence between the predicted value and the real value.
3. The RMS of the error sequence is used as the evaluation index to evaluate the network performance. The formula for calculating the RMS of the error sequence is as follows:

\[
E = \sqrt{\frac{1}{n} \sum_{i=1}^{n} [\hat{Y}_i(X_i) - Y_i]^2}
\]

4. Experiments and analysis of gas emission prediction

4.1. Data sources and processing

This paper chooses the gas monitoring data of a mine from May 2007 to December 2008 [5,6,7]. The specific data is shown in Table 1. Grey correlation analysis of the data in the table shows that the correlation degree of all data is greater than 0.5. Therefore, the predicted value of GM (1,1) gas emission and 13-dimensional data of 1-15 groups in Table 1 are directly used for training of neural network, and the remaining 16-18 groups are used to test the training effect, thus forming the training and testing sample set of the network.
Because different vectors have different physical meanings in the sample data and the dimensions of different vectors are different, it is necessary to normalize the data of training sample set before training, so that all the data can be transformed into pure quantities. The mapping is as follows:

\[
F: \quad x \rightarrow y = (y_{\text{max}} - y_{\text{min}}) \times (x - x_{\text{min}}) (x_{\text{max}} - x_{\text{min}})^{-1} + y_{\text{min}}
\]  

In formula (18), \(x_{\text{max}}\) and \(x_{\text{min}}\) are the maximum and minimum values of each dimension vector, \(y_{\text{max}}\) and \(y_{\text{min}}\) are the maximum and minimum values of the result vector, and all the maximum and minimum values are 1 and -1. \(x\) is the result value of the direct prediction of the model, \(y\) is the actual value of the prediction. When the prediction is finished, as shown in Formula (18), the model prediction result \(x\) needs to be normalized, and the final \(y\) is the actual value.

Table 1. Sample data of gas emission.

| Number | Coal Seam Gas Content (m³/t) | Coal seam depth (m) | Coal seam thickness (m) | Coal seam dip (°) | Working face length (m) | Propulsion speed (m/d) | Recovery rate | Gas content in adjacent layers (m³/t) | Thickness of adjacent layer (m) | Layer spacing (m) | Interlayer lithology | Mining Depth (m) | Gas emission (m³/min) |
|--------|-----------------------------|---------------------|------------------------|------------------|------------------------|------------------------|--------------|---------------------------------------|--------------------------|-----------------|----------------------|-----------------|----------------------|
| 1      | 1.92                        | 408                 | 2                      | 10               | 155                    | 4.42                   | 0.96         | 2.02                                 | 1.5                      | 20              | 5.03                 | 1825            | 3.34                 |
| 2      | 2.15                        | 411                 | 2                      | 8                | 140                    | 4.16                   | 0.95         | 2.1                                  | 1.21                     | 22              | 4.87                 | 1527            | 2.97                 |
| 3      | 2.14                        | 420                 | 1.8                    | 11               | 175                    | 4.13                   | 0.95         | 2.56                                 | 1.62                     | 19              | 4.75                 | 1751            | 3.56                 |
| 4      | 2.58                        | 432                 | 2.3                    | 10               | 145                    | 4.67                   | 0.95         | 2.4                                  | 1.48                     | 17              | 4.91                 | 2078            | 3.62                 |
| 5      | 2.4                         | 456                 | 2.2                    | 15               | 160                    | 4.51                   | 0.94         | 2.55                                 | 1.75                     | 20              | 4.63                 | 2104            | 4.17                 |
| 6      | 3.22                        | 516                 | 2.8                    | 13               | 180                    | 3.45                   | 0.93         | 2.21                                 | 1.72                     | 12              | 4.78                 | 2242            | 4.6                  |
| 7      | 2.8                         | 527                 | 2.5                    | 17               | 180                    | 3.28                   | 0.94         | 2.81                                 | 1.81                     | 11              | 4.51                 | 1979            | 4.92                 |
| 8      | 3.35                        | 531                 | 2.9                    | 9                | 165                    | 3.68                   | 0.93         | 1.88                                 | 1.42                     | 13              | 1.82                 | 2288            | 4.78                 |
| 9      | 3.61                        | 550                 | 2.9                    | 12               | 155                    | 4.02                   | 0.92         | 2.12                                 | 1.6                      | 14              | 4.83                 | 3235            | 5.23                 |
| 10     | 3.68                        | 563                 | 3.1                    | 11               | 175                    | 3.53                   | 0.94         | 3.11                                 | 1.46                     | 12              | 4.53                 | 2410            | 5.56                 |
| 11     | 4.21                        | 590                 | 5.9                    | 8                | 170                    | 2.85                   | 0.795        | 3.4                                  | 1.5                      | 18              | 4.77                 | 3139            | 7.24                 |
| 12     | 4.03                        | 604                 | 6.2                    | 9                | 180                    | 2.64                   | 0.812        | 3.15                                 | 1.8                      | 16              | 4.7                  | 3354            | 7.8                  |
| 13     | 4.8                         | 634                 | 6.5                    | 9                | 165                    | 2.77                   | 0.785        | 3.02                                 | 1.74                     | 17              | 4.62                 | 3087            | 7.68                 |
| 14     | 4.8                         | 634                 | 6.5                    | 12               | 175                    | 2.92                   | 0.733        | 2.98                                 | 1.92                     | 15              | 4.55                 | 3620            | 8.51                 |
| 15     | 4.67                        | 640                 | 6.3                    | 11               | 175                    | 2.75                   | 0.802        | 2.56                                 | 1.75                     | 15              | 4.6                  | 3412            | 7.95                 |
| 16     | 2.43                        | 450                 | 2.7                    | 12               | 160                    | 4.32                   | 0.95         | 2                                   | 1.7                      | 16              | 4.84                 | 1996            | 4.06                 |
| 17     | 3.16                        | 544                 | 2.7                    | 11               | 165                    | 3.81                   | 0.93         | 2.3                                  | 1.8                      | 13              | 4.9                  | 2207            | 4.92                 |
| 18     | 4.62                        | 629                 | 6.4                    | 13               | 170                    | 2.8                    | 0.803        | 3.35                                 | 1.61                     | 19              | 4.63                 | 3456            | 8.04                 |

4.2. Grey GM(1,1) prediction

The amount of gas emission from groups 1-15 after data pre-processing was used as the input value for the gray system, and the gas emission amount of groups 16-18 was predicted. Figure 2 shows the comparison of the forecasting results of the grey GM(1,1) forecasting model.

It can be found from the fitting effect chart of the above figure that the non-linear curve of continuous oscillation cannot be well fitted by the simple GM (1,1)-model, but can only show a certain trend.

4.3. Grey-generalized regression neural network prediction

The traditional grey-generalized regression neural network model uses only the gray prediction value of gas emission as the one-dimensional input of the model when predicting the gas emission, ignoring the influence of nonlinear correlation factors. Therefore, the gray prediction value of the gas emission amount and the selected 12 factors affecting the gas emission amount are used as the input of the network to optimize the model. The network output is the amount of gas emission. An optimized grey-generalized regression neural network model with 13-dimensional input and 1D output is constructed by adjusting the smoothing parameters. The value of the network smoothing factor is determined by cross-validation and minimum root mean square error.
First, K-fold cross-validation is performed on the data [13]. K-fold cross-validation is to divide the data into K roughly equal sub-samples, one test data is taken out for verification at a time, and the remaining k-1 sample data are trained. Experiment with the validation data in the resulting model and calculate the current error rate. Cross-validation can be used to learn samples from multiple angles, so that the sample set can be fully trained and learned to avoid falling into local extremum. In this paper, the data of 1-15 rows of Table 1 is used as training data, and k=10 is used for folding cross-validation. This not only can make up for the shortcomings of less training samples, but also avoid over-learning and under-learning. After many attempts, this paper determined that the increment factor of the smoothing factor is 0.1, and the initial value and the ending value are 0.1 and 2, respectively. The RMSE value of the predicted value and the true value is calculated in the cross-validation process. It can be seen from Figure 3 that when the smoothing factor is 0.8, the RMSE value is the smallest, and the smoothing parameter at this time is the optimal value. Therefore, the spread value determined in this paper is 0.8.

The grey-generalized regression neural network combined model is constructed with the determined smoothness factor. The gas emission of 16-18 rows in Table 1 is predicted. The fitting effect chart is shown in Figure 4:

**Figure 2.** Comparison of forecasting effect of grey GM (1,1) forecasting model.

**Figure 3.** Variation of root mean square error with smoothness factor.

**Figure 4.** Comparison of predicted and real gas emission values of different models.
From the comparison results of fitting in Figure 4, it can be seen that the grey-generalized regression neural network with multi-dimensional input has better fitting and prediction effect. In order to compare the fitting results, this paper also constructs a BP neural network model. As can be seen from Figure 4, BP neural network also has a good prediction effect. However, the grey-generalized regression neural network with one-dimensional input can only reflect a certain change trend. Therefore, it is very effective to introduce the prediction method of relevant factors affecting gas emission in the experiment, which provides a new idea for the related research of gas emission prediction in the future.

4.4. Model evaluation

In order to compare the advantages and disadvantages of the model proposed in this paper, we evaluated it from two dimensions: prediction accuracy and training time. Table 2a compares the fitting effects of several models, and Table 2b compares the training time of several models.

Table 2a. Comparison of prediction accuracy of different prediction models.

| Number | True value (m³/min) | One-dimensional Grey GRNN Prediction | BP Neural Network Prediction | Multidimensional Grey GRNN Prediction |
|--------|---------------------|-------------------------------------|----------------------------|-------------------------------------|
|        |                     | Predicted value (m³/min) | Relative error | Predicted value (m³/min) | Relative error | Predicted value (m³/min) | Relative error |
| 16     | 4.06                | 8.058081                  | 0.98474        | 4.067637                  | 0.00188        | 4.597671                  | 0.13243        |
| 17     | 4.92                | 8.058081                  | 0.62651        | 4.624953                  | 0.05997        | 5.133297                  | 0.04335        |
| 18     | 8.04                | 8.058081                  | 0.00849        | 7.2347                    | 0.10016        | 7.869498                  | 0.02120        |
|        | Average relative error |                              | 53.42%         | 5.4%                      | 6.57%          |

Table 2b. Comparison of training time for different prediction models.

| Number of training (times) | One-dimensional grey GRNN training time(ms) | Training time of BP neural network(ms) | Multidimensional Grey GRNN Training Time(ms) |
|----------------------------|--------------------------------------------|--------------------------------------|---------------------------------------------|
| 10                         | 9                                         | 798                                  | 9                                           |
| 50                         | 8.3                                       | 781                                  | 8.3                                         |

References [16-17] indicates that when the average relative error of model fitting and prediction is less than 10% (MRE ≤ 10%), the accuracy and accuracy of prediction are higher, and when MRE ≤ 5%, it is ideal. From Table 2a, it can be seen that the prediction accuracy of grey-generalized regression neural network model with multi-dimensional input is higher, with an average relative error of 6.57%, which is slightly lower than that of BP neural network.

Table 2b shows the training time comparison of different prediction models under MATLAB simulation. The training time of BP neural network is closely related to the number of hidden layers and the number of neurons in hidden layers. In order to compare with the model proposed in this paper, the BP neural network used in this paper contains only one hidden layer, and the number of neurons in the hidden layer is 3. The input layer and output layer of the model are the same as that of the multi-dimensional grey-generalized regression neural network model. In order to achieve the training accuracy of 10⁻⁵, the number of iterations of BP neural network used in this paper is 172. In order to fully illustrate the improvement of convergence speed and eliminate the influence of the number of hidden layer neurons on the experimental results, the training time of BP neural network with the number of hidden layer 3-100 is roughly counted in this paper. Statistics show that the optimal training time is about 600 ms and the worst training time is about 850 ms when the preset accuracy is reached.
In this paper, the average training time of 10 and 50 tests of the three models mentioned in Figure 4 is compared as shown in Table 2b. Because the output of the GM(1,1) model is used as input in the one-dimensional grey GRNN prediction model and the multi-dimensional grey GRNN prediction model, the training time is equivalent to the calculation time of the model. Its threshold distribution is between 8ms and 9ms, while the training time threshold distribution of the BP neural network is between 780ms and 800ms, which is about 100 times of the former. Therefore, the applicability of the multi-dimensional input grey-generalized regression neural network model is stronger in real-time scenarios, which provides a new idea for predicting gas emission.

5. Concluding remarks
In order to improve the accuracy of gas emission prediction model and the convergence speed of the model, this paper proposes an algorithm combining grey system theory with generalized regression neural network by selecting the generalized regression neural network without training and grey GM(1,1) model, and applies the algorithm to the prediction model of gas emission. The actual results of gas emission prediction in a mine show that the application of grey-generalized regression neural network algorithm to predict gas emission can effectively improve the convergence speed of the prediction model when the prediction accuracy is not particularly high. It provides a new way and practical guidance for the effective prevention and control of coal mine gas disasters in theory.

Acknowledgment
The paper is supported by the National Key R&D Program of China (No.2016YFC0801707), the National Key R&D Program of China (No.2017YFC0805208).

References
[1] Li R Q, Shi S L, N Q F and Jiang M 2011 J. Chinese Journal of Safety Sciences 21 143-151
[2] Fu H, Wang X R, Yang B C, Wang Z J, TU N W, Wang Y H and XU Y S 2014 J. Chinese Journal of Sensors and Actuators. 27 1568-1572
[3] Zhu H Q, Chang W J and Zhang B 2007 J. Journal of China Coal Society. 32 504-508
[4] Wang X L, LIU J and Lu J J 2011 J. Journal of China Coal Society. 36 80-85
[5] Lv F, Liang B, Sun W J and Wang Y 2012 J. Journal of China Coal Society. 37 113-116
[6] Fu H, Jiang W and Shan X X 2012 J. Journal of China Coal Society. 37 654-658
[7] Ren Z L, Lin D, Xia B W and Li W 2017 J. Chinese Journal of Sensors and Actuators. 30 247-252
[8] Li M Y, Jiang F C and Chen S F 2013 J. Journal of Safety Science and Technology. 9 188-192
[9] Shi C, Fan Y P, Wang Y S 2018 J. Journal of Safety Science and Technology. 14 72-76
[10] Lu G B, Li X Y, Zu B H and Dong J J 2017 J. Journal of Safety Science and Technology. 13 109-114
[11] Fu H, Wang F J and Chen Z C 2013 J. Transducer and Microsystem Technologies. 32 31-34
[12] Deng Julong 2005 Basic method of grey system: Chinese-English comparison - 2nd Edition.M. Huazhong University of Science and Technology Press.
[13] Liu S F, Cai H, Yang Y J and Cao Y 2013 J. System Eng Theor Pract. 33 2041-2046
[14] F. S D 1991 J. IEEE Transactions on Neural Networks. 2 568-576
[15] Luo Y 2007 Research on passenger volume prediction model based on Grey Theory and generalized regression neural network.D. Southwest Jiaotong University. (in Chinese)
[16] Zhou J, Wang Y and Sun S 2014 J. Applied Mechanics and Materials. 638-640 2036-2041
[17] Song B, Yang Y L, Feng Y X 2017 J. China Digital Medicine. 4 68