Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 11, No. 1, 2024

A new IPCA-CPSO-BP Model for Predicting Gas Emission in Underground Coal Mines

Guangsheng Chen

School of Energy Science and Engineering, Henan Polytechnic University, Jiaozuo, Henan, 454003 China

Abstract: Gas has become the primary factor restricting the safe and efficient production of coal mines. Gas emission prediction model is very important for mine gas emission prediction and gas disaster prevention. The improved principal component analysis (IPCA) was used to reduce the dimension of 13 influencing factors of gas emission, and the CPSO-BP neural network prediction model was built with MATLAB software to predict the absolute gas emission of Zhongtai Mining.The results show that the cumulative contribution rate of principal components is not less than 95%, and the number of principal components after improvement is reduced from the original 6 to 3, which effectively solves the problem of excessive principal components and data redundancy caused by the correlation difference between various influencing factors and gas emission amount. At the same time, the optimized particle swarm optimization algorithm alleviates the trouble that the particle swarm optimization algorithm is prone to fall into the local optimal solution. By coupling the improved particle swarm optimization algorithm with BP neural network, the problem of BP neural network's over-dependence on the initial value of weights and thresholds is solved, and the optimal initial weights and thresholds are provided for BP neural network, improving the accuracy of model prediction results. The average relative error of the original BP neural network is 1.6638%, the normalized mean square error is 0.3155, and the regression correlation index is 0.8157. The IPCA-CPSO-BP prediction model decreased to 0.5749%, 0.1143 and 0.9758, respectively. IPCA improves the data dimensionality reduction ability of principal component analysis, IPCA-CPSO-BP enhances the stability of particle swarm optimization algorithm, and significantly improves the prediction accuracy of BP model. The prediction trend is highly consistent with the actual value, which verifies the reliability of the model and provides strong data support for mine gas prevention and control.

Keywords: Keywords: Prediction of gas emission; Spearman correlation coefficient; Principal component analysis; Particle swarm optimization; BP neural network.

1. Introduction

Mine gas gushing is one of the main causes of gas accidents. According to the requirements of Coal mine Safety regulations, it is necessary to predict the gas emission quantity of coal mining face before mining, so the prediction research of gas emission quantity is very important [1-3].At present, many scholars have carried out a lot of work on the prediction of gas emission and obtained a lot of results. Xie Minglu et al.[4] used the method of multiple linear regression and SPSS software to process the data, carried out multiple regression analysis on the influencing factors of gas emission, and worked out the regression equation for prediction; Lu Guobin et al. [6] used the PCA-BP model to predict the amount of gas emission in the mining face; Zhou Xihua et al.[7] established GA-BP neural network prediction model based on principal factor analysis, and introduced momentum term in the process of weight reverse updating; Xie Xingjun et al.[8] used kernel principal component analysis (KPCA) to reduce the dimensionality of the influencing factors. To solve the problems of slow convergence and easy to fall into local optimal solution in BP neural network (BPNN), they used CMGA to optimize BPNN. The coupled algorithm of CMGA and BPNN (CMGANN) was constructed to predict the gas emission; Lin Haifei et al.[9] based on the monitoring data of absolute gas emission at a coal face in Huangling, Shaanxi Province, applied periodic trend decomposition based on local weighted regression and integrated empirical mode decomposition to construct a SEGS model by optimizing parameters of support vector regression machine through genetic algorithm, and obtained the final prediction result of gas emission. Feng Shengcheng et al.[10] established a prediction model of gas emission in mining face based on PCA and PSO-LS-SVM by taking advantage of the characteristics of fast training speed, global optimal solution and good generalization of vector machine, and successfully applied it in practice.; Xiao Peng et al [11] proposed to combine the wavelet packet decomposition method with the extreme learning machine to build a time-varying prediction model of the gas emission quantity, so as to improve the prediction accuracy and reliability of the gas emission quantity.; Liu Feng [12] used simulated annealing algorithm to optimize particle swarm optimization and established PCA-SAPSO-BP model to predict the amount of gas emission, and the prediction results showed that the model had fast convergence speed and high accuracy.; Fu Hua et al.[13] proposed an algorithm combining kernel principal component analysis and autoregressive integral moving average model, expressed the gas emission amount as a time series, and established a prediction model of the gas emission amount with autoregressive integral moving average, which has high prediction accuracy. Zhang Yucai et al.[14] used WOA, a whale optimization algorithm, to optimize LSTM to predict the amount of gas emission. By comparing the experimental results of different time steps, they found that the model had strong robustness and high accuracy; Zhou Xu et al.[15] proposed a combination of non-negative matrix decomposition NMF and random forest RF, and used the Hunger Games search algorithm HGS to optimize RF parameters, and established a coal seam gas emission prediction model with stable operating results and high precision.

The above gas emission prediction methods have achieved results to a certain extent, but the factors affecting the gas emission amount are complex and diverse, and many influencing factors carry a large amount of redundant information to affect the prediction accuracy and running speed of the model. How to optimize the data pre-processing algorithm and improve the accuracy of the prediction results has become an urgent task for the research of prediction models.

In order to solve the shortcomings of the above prediction models and improve the prediction accuracy, the IPCA-CPSO-BP model of gas emission prediction is proposed in this paper. By introducing Spearman correlation coefficient into the traditional principal component analysis method, the number of principal components decreases and the information carried by the first principal component increases after data dimensionality reduction. Adaptive inertia weight and contraction factor are added to the PSO algorithm to ensure that the PSO algorithm has a good search effect before and after. It is used to optimize the defects of BP neural network, such as excessive dependence on initial weights and thresholds, and slow convergence, so as to improve the prediction accuracy.

2. Principal Component Analysis (PCA) Method Improved

Principal component analysis (PCA) is a common data dimensionality reduction method, which has obvious dimensionality reduction effect in data processing, small error, simple operation, easy implementation and strong practicability, and is widely used in the field of gas emission prediction.

2.1. Influencing factors of gas emission quantity

The approved production capacity of Zhongtai Mining mine is 1.5 million tons/year. It adopts underground mining (longwall caving and falling mining method), multi-level down-hill development of vertical shaft and multi-level down-hill mining of Permian Shanxi Formation 21 coal seam. There are 5 vertical shafts and shafts, and the ventilation method of mine is mixed type and extraction type. The ventilation system of the mine is regionalized, and the gas grade evaluation of the mine from 2017 to 2022 is all outburst mine. Based on the field investigation of gas emission in II_1 coal seam of Zhongtai Mining, the influencing factors of gas emission are selected from geological factors and mining factors [16, 17]： X_1 promotion degree, X_2 recovery rate, X_3 coal thickness, X_4 burial depth, X_5 mining height, X_6 temperature, X_7 underground pressure, X_8 adjacent coal seam spacing, X_9 air volume, X_{10} daily production, X_{11} coal seam inclination, X_{12} adjacent layer gas content and X_{13} mining coal seam gas content, the prediction work of gas emission Y of Zhongtai mining industry is studied. The original data are shown in Table 1:

Table 1. Data of influencing factors of gas emission

Group	X_1/（m/d）	X_2/ %	X_3/ m	X_4/ m	X_5/ m	X_6/ ℃	X_7/ KPa	X_8/ m	X_9/（m^3/min）	X_{10}/（t/d）	X_{11}/ °	X_{12}/（m^3/t）	X_{13}/（m^3/t）
1	4	84.5	5.85	625.17	3	19	87.26	3	1000	3591	8	5.2	4.93
2	4.3	85.7	8.87	640.01	3	20	87.14	4	1008	3857	9	5.63	4.98
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
30	3	84.7	5.04	636.17	3	21	86.95	4.3	967	2856	9	5.41	4.72

2.2. Spearman correlation coefficient improved principal component analysis (PCA)

As principal component analysis (PCA) treats each influencing factor equally in the process of data processing, the impact of the influencing factor that is not strongly correlated with the target variable becomes greater after dimensionality reduction. In order to reduce this impact, Spearman correlation coefficient is used to assign weight to each influencing factor. If the influencing factor is strongly correlated with the target variable, the weight is significant. Small weight, this method has no specific requirements on the distribution of variables, can be applied to small samples (sample size does not exceed 30) and large sample range.

2.2.1. Spearman correlation coefficient

Spearman's correlation coefficient calculation method first arranges n samples according to the value size of the impact factor from small to large, and then arranges the target variables in the same way, which can be evaluated by substituting them into formula (1):

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \tag{1}$$

Where: represents the correlation between the feature variable and the target variable, and the larger the correlation, the stronger the correlation; n represents the number of samples; Represents the ranking difference between the feature variable and the target variable.

2.2.2. Improvement step

The main steps of Spearman correlation coefficient to improve PCA are as follows:

Step 1: Select the original sample matrix $X_{n\times m}$， Where n represents the number of samples and m represents the number of impact factor variables.

Step 2: Spearman correlation coefficient is calculated between each impact factor variable and target variable, and the diagonal matrix formed by the calculation result is denoted as $R_{m\times m}$, as follows:

$$R_{m\times m} = \begin{pmatrix} r_{11} & & & \\ & r_{22} & & \\ & & \ddots & \\ & & & r_{mm} \end{pmatrix} \tag{2}$$

Where: r_{ii} is the Spearman correlation coefficient between the impact factor X_i and the target variable Y.

Step 3: The original sample matrix $X_{n\times m}$ is normalized (de-dimensional) according to formula (3) to obtain the matrix. $Z_{n\times m}$.

$$z_{ij}=\frac{x_{ij}-\frac{1}{n}\sum_{i=1}^{n}x_{ij}}{\sqrt{\frac{\sum_{i=1}^{n}\left(x_{ij}-\frac{1}{n}\sum_{i=1}^{n}x_{ij}\right)^2}{n-1}}} \tag{3}$$

Step 4: Multiply the normalized matrix $Z_{n\times m}$ with the diagonal matrix $R_{m\times m}$ of Spearman correlation coefficient to get the improved PCA initial feature variable matrix $G_{n\times m}$.

Step 5: Calculate the covariance matrix of the eigenvariable matrix $G_{n\times m}$.

Step 6: Calculate the eigenvalues and eigenvectors of the covariance matrix..

Step 7: Calculate the principal component contribution rate of each characteristic variable, and obtain the final principal component as the prediction index according to the cumulative contribution rate greater than or equal to 95%.

2.3. Comparison of data dimensionality reduction before and after improvement

13 principal components were obtained by principal component analysis (PCA) and improved principal component analysis (IPCA) after dimensionality reduction. The results are shown in Table 2 and Figure 1

Table 2. Changes in the contribution rate of principal component analysis before and after improvement

Principal component	PCA		IPCA	
	Contribution contribution (%)	Cumulative contribution rate (%)	Contribution contribution (%)	Cumulative contribution rate (%)
Z_1	**61.2484**	61.2484	**89.6354**	89.6354
Z_2	9.7231	70.9714	3.1107	92.7461
Z_3	8.6822	79.6536	2.4600	**95.2060**
Z_4	8.4483	88.1019	1.6003	96.8063
Z_5	4.5867	92.6886	0.8557	97.6620
Z_6	4.3779	**97.0665**	0.7171	98.3791
Z_7	1.5388	98.6053	0.5624	98.9415
Z_8	0.8199	99.4252	0.4102	99.3517
Z_9	0.2569	99.6821	0.2513	99.6030
Z_{10}	0.1154	99.7975	0.1499	99.7529
Z_{11}	0.0939	99.8914	0.1316	99.8845
Z_{12}	0.0779	99.9694	0.0683	99.9528
Z_{13}	0.0306	100.0000	0.0472	100.0000

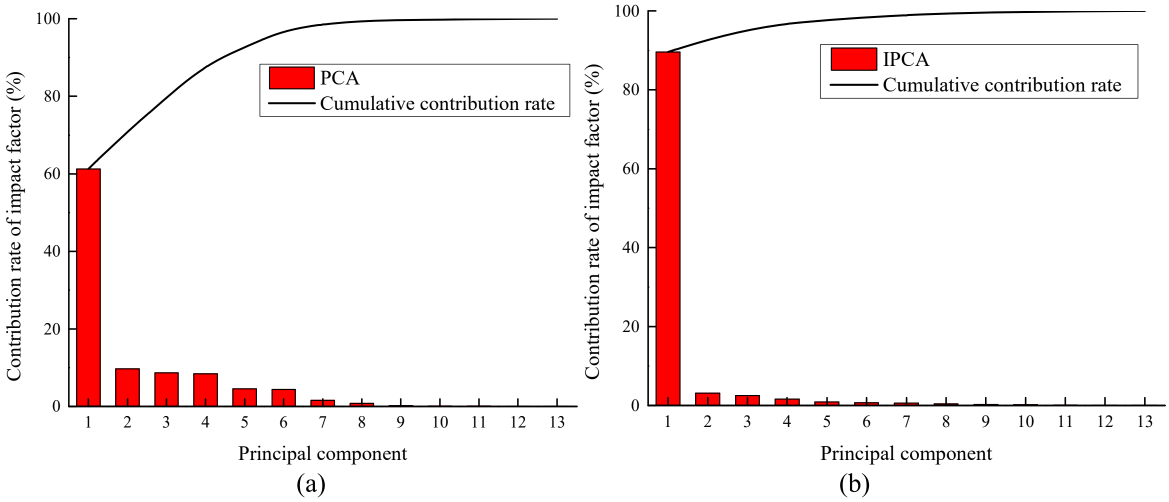

Figure 1. Effect diagram of principal component analysis before and after improvement

Compared with the effect before and after the improvement of principal component analysis, principal components were selected according to the cumulative principal component was greater than or equal to 95%. Principal component analysis extracted 6 principal components with a cumulative contribution rate of 97.0665%, while the improved principal component analysis only needed to extract 3 principal components with a cumulative contribution rate of 95.2060%. Although the cumulative contribution rate of PCA is greater than that of IPCA, the first principal component of IPCA carries more information. While the cumulative contribution rate is greater than or equal to 95%, the number of principal components is reduced, the calculation accuracy of the model is improved, and the calculation amount of the model is effectively reduced and the calculation speed of the model is increased. The improved principal component analysis method has better effect on dimensionality reduction. IPCA method is superior to PCA method, and can be used as the final method to reduce the dimension of gas emission prediction index.

3. BP Neural Network Model Based on CP-SO Optimization

3.1. PSO algorithm improvement

Particle Swarm Optimization (PSO) is a stochastic global optimization algorithm first proposed by Kennedy and Eberhart in 1995[18].

3.1.1. Basic PSO algorithm

The basic PSO algorithm uses formulas (4) (5) to operate on particles:

$$v_i^d = wv_i^d + c_1 r_1 (p_i^d - x_i^d) + c_2 r_2 (p_g^d - x_i^d) \quad (4)$$

$$x_{d+1}^i = x_d^i + v_{d+1}^i \quad (5)$$

Where: w called inertia factor, is non-negative; c_1 and c_2 are called learning factor and is usually 2; r_1 and r_2 are random number between 0 and 1; x_d^i is the position of the i-th particle at step d.

The velocity evolution equation expressed in formula (4) and (5) consists of three parts. The first part is the inertia of the particle itself, and the velocity in a certain direction when the particle carries out the next search; The second part is the learning of the particle itself, which comes from the "experience" of the particle's previous flight, namely the local optimal solution. The third part is the information sharing among particles, that is, the global optimal solution. In addition, when the particle is constantly adjusting its position according to the speed, it is also limited by the maximum speed $v_{\max}$. The termination condition of iteration is generally selected as the maximum number of iterations or the optimal location of the current location of the population meets the predetermined minimum fitness threshold.

3.1.2. Adaptive inertial weight particle swarm optimization

A large inertial weight value is conducive to global search, while a small weight value is more conducive to local search.[19, 20] In order to make PSO more stable, the inertia weight of PSO is changed adaptively. Compared with the original PSO, the inertia weight and the number of iterations are related to the fitness of each particle. For the minimum value problem, the change rule of inertia weight is shown in equation (6). For the maximum value problem, the change rule of inertia weight is shown in equation (7).

$$w_i^d = \begin{cases} w_{\min} + (w_{\max} - w_{\min}) \dfrac{f(x_i^d) - f_j^d}{f_{averge}^d - f_{\min}^d}, & f(x_j^d) \leq f_{averge}^d \\ w_{\max}, & f(x_j^d) > f_{averge}^d \end{cases} \quad (6)$$

$$w_i^d = \begin{cases} w_{\min} + (w_{\max} - w_{\min}) \dfrac{f(x_i^d) - f_j^d}{f_{averge}^d - f_{\min}^d}, & f(x_j^d) \geq f_{averge}^d \\ w_{\max}, & f(x_j^d) < f_{averge}^d \end{cases} \quad (7)$$

Where: $w_{\min}$ and $w_{\max}$ are the minimum and maximum inertia coefficients given in advance, generally 0.4 and 0.9.

The average fitness of all particles at the DTH iteration is shown in equation (8):

$$f_{averge}^d = \sum_{i=1}^{n} (x_i^d)/n \quad (8)$$

Minimum fitness of all particles at the DTH iteration, as shown in equation (9) :

$$f_{\min}^d = \min\{f(x_1^d), f(x_2^d), \cdots, f(x_n^d)\} \quad (9)$$

3.1.3. Compression factor particle swarm optimization

Individual learning factors c_1 and social learning factors c_2 determine the influence of the experiential information of particles and other particles on the trajectory of particles, which reflects the information exchange between particle groups [21-23].Setting a large c_1 value will make the particles search too much in their own range and converge slowly, and a large c_2 value will cause the particles to converge prematurely to the local optimal value. Setting a compression factor k, such as equations (10) and (11) can delay the arrival of the local optimal solution.

$$v_i^d = k[wv_i^d + c_1 r_1 (p_i^d - x_i^d) + c_2 r_2 (p_g^d - x_i^d)] \quad (10)$$

$$k = \frac{2}{\left|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}\right|} \quad (11)$$

Where: $\varphi = c_1 + c_2$, $\varphi > 4$.

3.1.4. Combinatorial particle swarm optimization (CPSO)

n the early stage, the fitness of particle swarm is large, and global search is more necessary. Corresponding to each particle, it requires rapid search within its own scope, and requires large inertia weight w, individual learning factor c_1 and small social learning factor c_2. In the later stage of particle swarm search, fitness decreases gradually, so smaller inertia weights, individual learning factors and larger social learning factors are needed. However, in the later stage, particle swarm will hover around itself too much and easily fall into the local optimal solution. CPSO (Combine adaptive inertia weights and learning factors PSO) algorithm combines "adaptive inertia weights" and "compression factor algorithm". It can balance the local search and global search capabilities of PSO algorithm to avoid the shortcomings of premature convergence of PSO algorithm and falling into local optimal solutions, and the speed update formula is shown in (12).

$$v_i^d = w_i^d v_i^d + k[c_1 r_1 (p_i^d - x_i^d) + c_2 r_2 (p_g^d - x_i^d)] \quad (12)$$

3.2. BP neural network

BP neural network is a kind of multi-layer feedforward neural network which corrects error by error backpropagation algorithm. The signal propagates forward while the error propagates back. In the forward propagation process, the input signal is processed layer by layer through the input layer and hidden layer. When reaching the output layer, if the result or the number of iterations does not meet the requirements, the backpropagation process is entered, the error signal is returned in the original way, and the weight of each layer is modified. The number of hidden layer neurons can be calculated by empirical formula (13) [24]:

$$a=(h+c)^{1/2}+d \tag{13}$$

Where: a is the number of hidden layer nodes; h is the number of nodes in the output layer; c is the number of nodes in the input layer; d is the adjustment constant, generally 1,2,... 10.

3.3. Calculation steps of CPSO-BP model

Step 1: Data normalization, establish BP neural network structure;

Step 2: Set the particle swarm hyperparameters;

Step 3: Conduct training according to mean square error as fitness function of CP-SO objective function (MSE, as shown in equation (14)), iteratively optimize the weights ω and thresholds θ of BP algorithm;

$$MSE=\frac{1}{s}\sum_{i=1}^{s}\sqrt{\left(\hat{y}_i-y_i\right)^2} \tag{14}$$

Where: $\hat{y}_i$ is the predicted value of the training group, y_i is the actual value of the training group.

Step 4: Calculate the fitness value of the particle swarm, compare $pbest$ and $gbest$ update the sum after each calculation, and bring the speed and position of updating the particle swarm into equation (9) and equation (2) until the required accuracy or number of iterations are reached, and the final $gbest$, namely the optimal weight value ω and threshold value θ are obtained;

Step 5: Bring the test group data into the updated BP neural network for calculation until the number of iterations ends or the error condition is reached, and the predicted results are output.

4. IPCA-CPSO-BP Model Application

4.1. IPCA principal component analysis

Through MATLAB software programming, the IPCA algorithm was used to reduce the dimensionality of the original data, and three principal component feature variables were obtained. The eigenvector data of the principal component about the impact factor are shown in Table 3:

Table 3. Improved principal component factor component data

Influence factor	F1	F2	F3
X_1	0.0015	0.0008	-0.0011
X_2	0.0024	0.0017	-0.0030
X_3	0.0088	0.0047	-0.0073
X_4	0.0038	0.0064	-0.0088
X_5	-0.9266	0.3524	0.0779
X_6	0.0343	-0.1115	0.9913
X_7	0.0066	0.0077	-0.0013
X_8	0.0008	0.0009	-0.0001
X_9	0.0046	-0.0068	-0.0079
X_{10}	0.1036	-0.0358	0.0520
X_{11}	-0.3597	-0.9284	-0.0916
X_{12}	0.0035	0.0084	-0.0071
X_{13}	0.0024	0.0015	0.0004

The three principal components obtained after dimensionality reduction by improved principal component analysis become new prediction indexes for gas emission prediction. As can be seen in Table 3.3 of the data obtained by MATLAB analysis, the data information in the first principal component mainly includes the degree of advancement of mining coal seam, coal thickness, recovery rate, mining height, recovery amount and gas content of mining coal seam in the original prediction index. The second principal component includes the underground pressure, the spacing of adjacent coal seams, the inclination of coal seams and the gas content of adjacent layers. The third principal component includes burial depth, temperature and air volume.

4.2. Parameter setting

The five algorithms BP, PSO-BP, PCA-PSO-BP, IPCA-PSO-BP and IPCA-CPSO-BP were programmed and calculated with MATLAB software. BP neural network was constructed with Newff function, with the maximum training times of 1000 and the learning rate of 0.01. Training requires an accuracy of 0.000001. Set groups 1-25 as the training group, 26-30 as the test group, the input, hidden and output layers of the first two groups as 13, 11 and 1, the third group as 8, 9 and 1, and the last two groups as 4, 6 and 1.

4.3. Prediction result analysis

The predicted values and relative errors (hereinafter referred to as errors) of the model are shown in Table 4:

Table 4. Test data results

group	Actual value	BP		PSO-BP		PCA-PSO-BP		IPCA-PSO-BP		IPCA-CPSO-BP	
		forecast	error	forecast	error	forecast	error	forecast	error	forecast	error
26	16.5	17.0189	0.5189	16.7819	0.2819	16.6531	0.1531	**16.4495**	**0.0505**	16.6748	0.1748
27	15.0	14.8042	0.1958	15.4738	0.4738	15.2683	0.2683	**15.0076**	**0.0076**	14.9129	0.0871
28	15.0	15.0634	0.0634	15.0536	0.0536	15.3371	0.3371	14.7403	0.2597	**14.9802**	**0.0198**
29	16.5	**16.6096**	**0.1096**	16.7949	0.2949	16.3260	0.1740	16.1537	0.3463	16.3371	0.1629
30	15.0	14.5828	0.4172	14.9649	0.0351	15.2247	0.2247	15.2073	0.2073	**14.9827**	**0.0173**

Average relative error MRE, normalized root-mean-square error Nmse and regression correlation index R2 were introduced to evaluate the prediction results, as shown in Table 5:

Table 5. Evaluation indeZes of prediction results of each model

Evaluation index	MRE/%	N_{mse}	R^2
BP	1.6638	0.3155	0.8157
PSO-BP	1.4491	0.2811	0.8537
PCA-PSO-BP	1.5034	0.2408	0.8926
IPCA-PSO-BP	1.1138	0.2159	0.9137
IPCA-CPSO-BP	**0.5749**	**0.1143**	**0.9758**

The evaluation indicators of different models were visualized, as shown in FIG. 2 and FIG. 3 are the comparison graphs between the prediction results of different models and the real values.

With the introduction of different algorithms and improved algorithms in a certain order, the average relative error MRE increases first and then decreases, and the normalized root-mean-square error Nmse and the regression correlation index R2 change monotonically. The smaller the mean relative error MRE and the normalized root-mean-square error Nmse, the more accurate the prediction result. The optimization effect of different algorithms on BP neural network is shown in Figure 2. The IPCA-CPSO-BP model is the most accurate in prediction, with MRE value of 0.5749 and Nmse value of 0.1143. Compared with BP prediction model, the average relative error of IPCA-CPSO-BP model is reduced by 1.0889%, and the average relative error of BP prediction model is about 3 times that of IPCA-CPSO-BP model. The normalized root-mean-square error is reduced by 0.2012, and the normalized root-mean-square error of BP prediction model is about 2.8 times that of IPCA-CPSO-BP model, which significantly improves the prediction accuracy of IPCA-CPSO-BP model. The closer the regression correlation index R^2 is to 1, the better the fitting effect of the model is. As shown in Figure 2 and 3, IPCA-CPSO-BP model has the best fitting effect, and the R^2 value is 0.9758. After particle swarm optimization and improved particle swarm optimization are added to the model, the fitting effect is improved and the dispersion of data distribution is low. The lines are more consistent with the true value.

Principal component analysis (PCA) can more effectively remove redundant data information and increase the accuracy of prediction results in the whole model. The improved PCA strengthens this feature. Particle swarm optimization (PSO) mainly controls the stability of the model, and the improved PSO also strengthens this property.

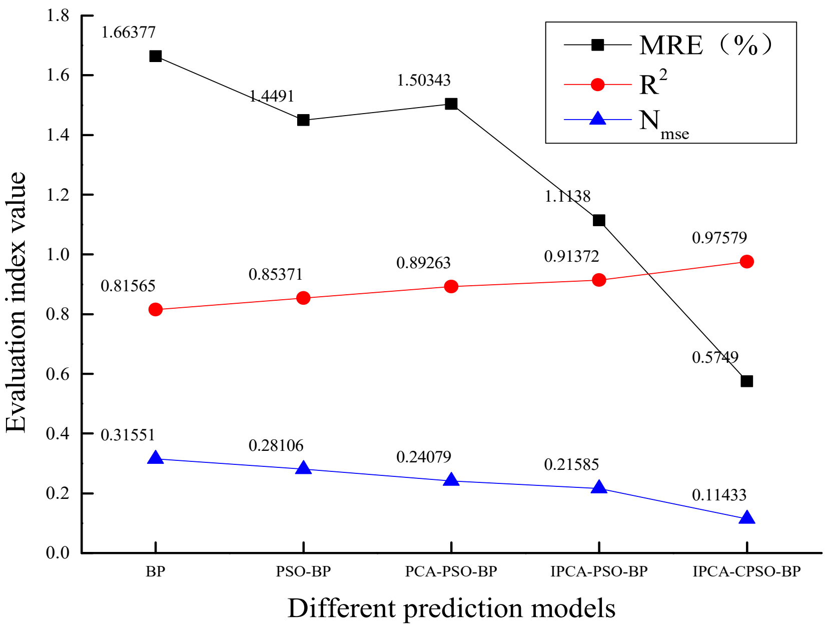

Figure 2. Evaluation indicators of different models

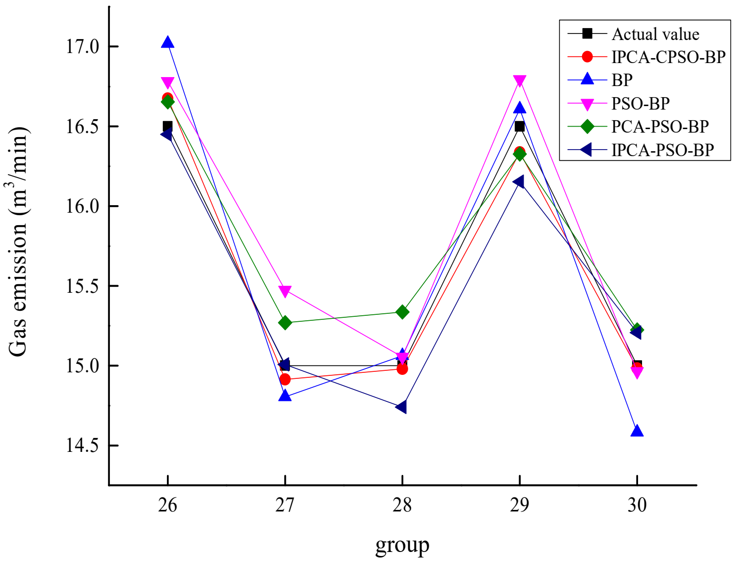

Figure 3. Comparison between the predicted value and the real value of different models

Figure 4 shows the curve of PSO iterations of each model. The model iteration curve is in the shape of a ladder. The model with the fastest convergence speed is IPCA-PSO-BP model, which tends to converge at the 25th step after iteration. The convergence speed of PSO-BP model and IPCA-CPSO-BP model is relatively slow, and the hopping amplitude of fitness value convergence is relatively large. Because PSO-BP model has not undergone data dimensionality reduction, the convergence process is random, and the number of iterations increases correspondingly. IPCA-CPSO-BP model adds adaptive inertia weight and shrinkage factor. The model can have global search particles before and after the search, so it is not easy to fall into the local optimal solution, the convergence rate decreases, and the prediction accuracy increases.

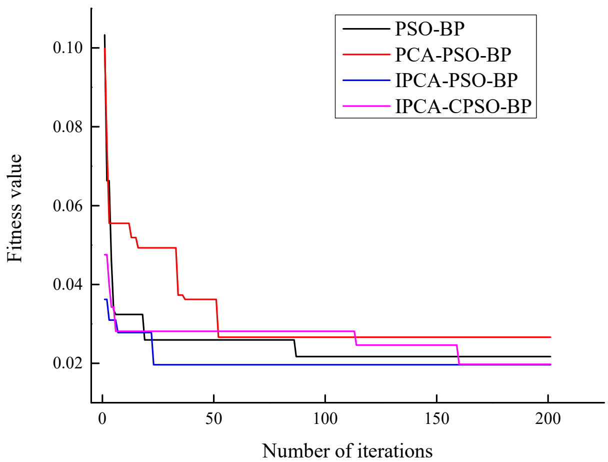

Figure 4. Iteration curves of different models

BP model relies too much on initial parameter values in the process of use, PSO convergence speed in the process of parameter optimization leads to easy to fall into the local optimal solution, PCA is not strong enough to reduce the dimension of data, resulting in the model has randomness in the prediction process. IPCA-CPSO-BP model improves the pre-processing ability of data, slows down the convergence speed and improves the prediction accuracy.

5. Conclusion

(1) The Spearman correlation coefficient (absolute value) was added to the principal component analysis as the weight of each influencing factor, which improved the principal component analysis. After dimensionality reduction, the traditional principal component analysis produced 6 principal components, with a cumulative contribution rate of 97.0665%, while the improved principal component analysis had only 3 principal components, with a cumulative contribution rate of 95.2060%. The improved principal component has better dimensionality reduction under the premise of preserving 95% information.

(2) The traditional particle swarm optimization (PSO) algorithm is improved by adding adaptive inertia weights and factors of compressed search space.The optimized algorithm not only guarantees the search accuracy, but also shows the faster convergence speed, thus achieving double improvement in the optimization efficiency and the accuracy of the predicted results.

(3) The IPCA-CPSO-BP neural network prediction model realized the prediction of gas emission at 3309 working face of Zhongtai Mining, and the average absolute error (MAE) was 0.0924. The mean relative error (MAPE) was 0.5749%. The regression correlation index R2 was 0.9758.

(4) Four models, namely BP model, PSO-BP model, PCA-PSO-BP model and IPCA-PSO-BP model, were established respectively, and the amount of gas emission was predicted, and the predicted results were compared with IPCA-CPSO-BP algorithm. Through the result analysis, it is found that the principal component analysis method controls the stability of the model operation, the particle swarm optimization algorithm controls the prediction accuracy of the model, and the calculation accuracy of IPCA-CPSO-BP model is about 3 times that of BP neural network model. The data confirm the feasibility of the IPCA-CPSO-BP neural network prediction model studied in this paper in the prediction of gas emission quantity, and also has a certain reliability in the prediction effect.

References

[1] QIN Y J, SU W W, JIANG W Z, et al. Research Progress and Development Direction of Mine Gas Emission Forecast Technology in China[J]. Safety in Coal Mines. 2020, 51(10): 52-59.

[2] LI R Q, SHI S L, NIAN Q F, et al. Research on Coalmine Gas Accident Rules in China in Recent Decade[J]. China Safety Science Journal. 2011, 21(09): 143-151.

[3] WU S Y, ZHANG J L, HAN T T, et al. Revision suggestions on “Prediction Methods of Mine Gas Emission Rate” industry standard[J]. China Coal. 2012, 38(09): 117-120.

[4] XIE M J, MA S Q. Prediction of coal mine gas emission based on multiple linear regression theory[J]. Shaanxi Coal. 2021, 40(01): 26-29.

[5] SUN X Y. Research and application of neural network gas emission prediction system[J]. Coal Mine Modernization. 2023, 32(05): 65-68.

[6] LU G B, KANG J K, BAI G, et al. Application of PCA-BP to gas emission prediction of mining working face[J]. Journal of Liaoning Technical University (Natural Science). 2015, 34(12): 1329-1334.

[7] ZHOU X H, SUN J Z. Prediction of Gas Emission based on Principal Factor Analysis and Improved BP Neural Network[J]. Mining Safety & Environmental Protection. 2018, 45(06): 43-47.

[8] XIAO P, XIE X J, SHUANG H Q, et al. Prediction of gas emission quantity based on KPCA-CMGANN algorithm[J]. China Safety Science Journal. 2020, 30(05): 39-47.

[9] LIN H F, LIU S H, ZHOU J, et al. Prediction method and application of gas emission from mining workface based on STL-EEMD-GA-SVR[J]. Coal Geology & Exploration. 2022, 50(12): 131-141.

[10] FENG S C, SHAO L S, LU W J, et al. Application of PCA-PSO-LSSVM model in gas emission prediction[J]. Journal of Liaoning Technical University (Natural Science). 2019, 38(02): 124-129.

[11] XIAO P, XIE X J, SHUANG H Q, et al. Application of wavelet-extreme learning machine in time-varying series prediction of gas emission quantity[J]. JOURNAL OF XI' AN UNIVERSITY OF SCIENCE ANDTECHNOLOGY. 2020, 40(05): 839-845.

[12] LIU F. Prediction of gas emission based on PCA-SAPSO-BP neural network[J]. Safety in Coal Mines. 2023, 54(04): 60-68.

[13] FU H, FU Y, ZHAO J C, et al. Prediction of gas emission based on KPCA-ARIMA algorithm[J]. Journal of Liaoning Technical University (Natural Science). 2022, 41(05): 406-412.

[14] ZHANG Y C, WANG Y, GUO K Y. Research on prediction of gas emission in working face based on WOA-LSTM[J]. Mining Safety & Environmental Protection. 2023: 1-6.

[15] ZHOU X, YANG J P, ZU Y W, et al. Gas emission prediction based on NMF-HGS-RF[J]. Mining Safety & Environmental Protection. 2023, 50(03): 117-123.

[16] LIN Y B, QIN Y, WANG X, et al. Geology and emission of mine gas in Bintang mining area with low rank coal and high mine gas[J]. Journal of China Coal Society. 2019, 44(07): 2151-2158.

[17] GUO C, XIA Y C, SUN X Y, et al. Method and practice of gas geological grading evaluation on coal mining face of high gas m[J]. Journal of China Coal Society. 2019, 44(08): 2409-2418.

[18] EBERHART, KENNEDY. A new optimizer using particle swarm theory[C]. Nagoya, Japan: IEEE, 1995.

[19] KANG Y S, ZANG S L. Improved Particle Swarm Optimization Algorithm Based on Multiple Strategies[J]. Journal of Northeastern University (Natural Science). 2023, 44(08): 1089-1097.

[20] ZHANG H, WANG X L. Adaptive inertia weight particle swarm optimization algorithm[J]. Intelligent Computer and Applications. 2023, 13(09): 5-8.

[21] EBERHART, Y. Shi. Comparing Inertia Weights and Constriction Factors in Particle Swarm Optimization[C]. La Jolla, CA, USA: CEC00, 2000.

[22] VANDENBERGH, ENGELBRECHT. A Cooperative Approach to Particle Swarm Optimization[J]. IEEE Transactions on Evolutionary Computation. 2004, 8(3): 225-239.

[23] LIU Y D, ZHOU M R, XIE J Y, et al. RESEARCH ON SIMULATION MODEL FOR DATA ASSIMILATION IN SOLAR RADIATION PREDICTION BASED ON PARTICLE SWARM OPTIMIZER WITH TIME VARYING CONSTRICT FACTOR[J]. Acta Energeia Solaris Sonica. 2021, 42(04): 181-185.

[24] WANG R B, XU H Y, LI B, et al. Research on Method of Determining Hidden Layer Nodes in BP Neural Network[J]. Computer Technology and Development. 2018, 28(04): 31-35.