

# Calculation Method of Probability Integration Method Parameters Based on MIV-GP-BP Model

Shenshen CHI, Xuexiang YU\*, Lei WANG

**Abstract:** In order to guarantee the precision of the parameters of the probability integral method (PIM), starting from optimizing input and improving algorithm an algorithm integrating the genetic algorithm (GA) and particle swarm optimization (PSO) was put forward to optimize the prediction model of BP neural network and the mean impact value algorithm (MIV) was applied to optimize the input of BP neural network. The mean impact value algorithm (MIV) was applied to optimize the input of BP neural network. The measured data of 50 working faces were chosen as the training and testing sets to build the MIV-GP-BP model. The results showed that among the five parameters, the *RMSE* was between 0.0058 and 1.1575, the *MaxRE* of  $q$ ,  $\tan\beta$ ,  $b$  and  $\theta$  was less than 5.42%, and the *MeaRE* was less than 2.81%. The *RMSE* of  $s/H$  did not exceed 0.0058, the *MaxRE* was less than 9.66% and the *MeaRE* was less than 4.31% (the parameters themselves were small). The optimized neural network model had higher prediction accuracy and stability.

**Keywords:** BP neural network; MIV algorithm; mining subsidence; optimization algorithm; PIM; underground mining

## 1 INTRODUCTION

With China's rapid economic development, there is a growing demand for coal resources. Recently, lots of underground coal resources have been exploited, which has caused a series of environmental problems, such as surface subsidence, cracks, dust, solid waste and etc., thus posing serious threats to the production and life of mining areas.

In order to maximize the production of coal resources and reduce large-scale surface subsidence, scholars have carried out extensive research on the prediction theories of mining subsidence. PIM based on stochastic medium mechanics theory has been widely applied for mining subsidence prediction [1]. Therefore, the acquisition of parameters is very essential, and the commonly-used methods were given. The first method is to use the intelligent optimization algorithm to fit the parameters based on the measured data, such as genetic algorithm, quantum annealing method, etc. [2, 3]. The second method is to apply the machine learning method, such as neural network, support vector machine, etc. [4]. The third method is to use the similarity theory to calculate the lithology of overlying rock [5]. These prediction methods have improved the accuracy of parameter determination to a certain extent.

There are complex non-linear relationships between the basic parameters and geological and mining conditions, so it is hard to describe them by using a specific mathematical model. Both methods 1 and 3 are based on the measured deformation to obtain the parameters. This often requires a large number of surface movement observation stations, which consumes a lot of human and material resources, and the observation period is long, so it is impossible to obtain the parameters of the new mining area. The aforementioned neural network technology has good autonomous learning, self-adaptive and fault-tolerant performance, and can be used to establish complex non-linear mapping relations. It has been extensively applied in pattern classification, clustering analysis, regression fitting analysis and optimization calculation [6, 7].

In order to fully use the existing mining data and establish the relations between the geological and mining conditions and the parameters of PIM, parameter

acquisition method based on neural network learning has been applied extensively. Bian combines PSO algorithm and BP neural network to establish a parameter-seeking model [8]. Zhao established a combined parameter acquisition model based on grey relational analysis and BP neural network [9]. Hu builds prediction method based on improved genetic algorithm [10]. These models have achieved good prediction results. The relevant research has made full use of the measured data, which could provide ideas for the calculation of the PIM parameters in mining areas without measured data.

The prediction accuracy of BP neural network is closely related to the amount of input layers. The more the input layers, the larger the network is needed and the closer to the correct results, which reduces the prediction accuracy to some extent. Therefore, it is useful to decrease the dimensionality of the sample. The average impact value method (MIV method) is used for data dimensionality reduction based on BP neural network, currently widely used. Therefore, in this paper, the MIV method was introduced into the model to analyze the geological and mining conditions, simplify the input layer, reduce the complexity of the neural network and improve the prediction accuracy.

For BP neural network, its weight and threshold selection are random, which makes the standard BP neural network converge slowly and is easy to fall into local optimum. The training efficiency could be improved by using the intelligent optimization algorithm to optimize the parameter optimization process in the BP neural network training process. These algorithms include simulated annealing (SA), GA algorithm, PSO algorithm, and etc. The PSO algorithm is simple and efficient in these optimization algorithms, but it has poor accuracy and easily falls into local optimum. To address the issue, Benvidi proposed a GA-PSO hybrid algorithm to optimize the ANN, and used the algorithm to deal with the relationship between the absorbance data and the concentration of the analyte. The algorithm is simple, the convergence speed is fast, and the convergence precision is high [11]. Mousavi uses fuzzy GA-PSO algorithm to solve the problem of automatic guidance vehicle scheduling optimization. Through simulation experiments, it is proved

that the hybrid algorithm outperforms single GA and PSO algorithms in all aspects [12]. However, there is no relevant research in the field of mining subsidence. In view of the above problems, this paper uses GP algorithm to optimize BP neural network, which is introduced into the mining subsidence field.

This paper collects data from 50 working faces. Firstly, MIV method was applied to screen input variables of the BP neural network, and sort factors affecting the parameters of PIM. It could eliminate lower-order factors, simplify the input of the network and reduce the network complexity. Then the GP algorithm was adopted to optimize the weight and threshold of the BP neural network, thus establishing the MIV-GP-BP prediction model.

## 2 CONSTRUCTION OF MIV-GP-BP PREDICTION MODEL

### 2.1 MIV Algorithm

Proposed by Dombi et al., MIV is applied to show the variation of the weight matrix in the neural network. Later, it is considered to be one of the best indicators for evaluating the most relevant correlation in the neural network [13]. It could be applied to quantitatively evaluate importance of the respective variables for the dependent variable. The algorithm steps are given: (1) After termination of the neural network training, the single independent variable sample will be increased and decreased by a certain ratio based on the original value to form two sets of new samples P1 and P2. In this paper, the regulation rate was 10%, that is, the original sample was increased and decreased by 10% to form two new samples. (2) The new samples P1 and P2 are taken as input layers, and their output layers T1 and T2 are trained. (3) The Impact Value (IV) of the change of the independent variable on the output could be obtained by calculating the difference between the two groups of predicted data T1 and T2. (4) The IV of the obtained sample value is averaged according to the number of observed samples, and the MIV value of the independent variable to the dependent variable is obtained (Fig. 2).

MIV values of each independent variable are sequentially calculated according to the aforementioned steps. The obtained MIV value has positive and negative points. Its symbols represent the direction of influence [14]. In this paper, the training sample data was tested several times to obtain MIV values. To improve the calculation efficiency, the number of operations for each group was 50 times, and the outliers in each group were eliminated. The MIV values of the same independent variable in each group were taken as absolute values and then added and averaged. Finally, the filtering of variables can be achieved by sorting according to the size of the obtained MIV values.

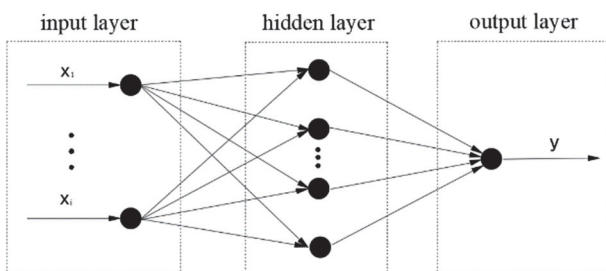


Figure 1 Topology of BP neural network

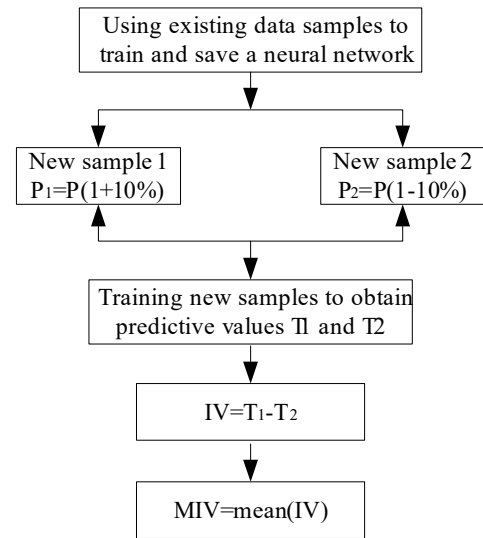


Figure 2 The flow chart of MIV algorithm

### 2.2 GP Algorithm Optimization BP Neural Network

#### 2.2.1 GP-BP Model

GP algorithm optimization BP neural network is to minimize the output error to enhance the weight and threshold parameters.

Structural parameters of the BP neural network need to be encoded into the individual information of the particle swarm. The number of neurons in the input layer, the hidden layer and the output layer is  $x$ ,  $y$  and  $z$ , respectively, then each neuron in the hidden layer includes  $x$  weights and one threshold information. Each neuron in the output layer includes  $y$  weights and one threshold information. The BP neural network needs to optimize  $I = (x + 1)y + (y + 1)z$  parameters. All the parameters to be optimized are expressed as  $D$ -dimensional vectors and they are used as the coding information of the particle swarm individuals.

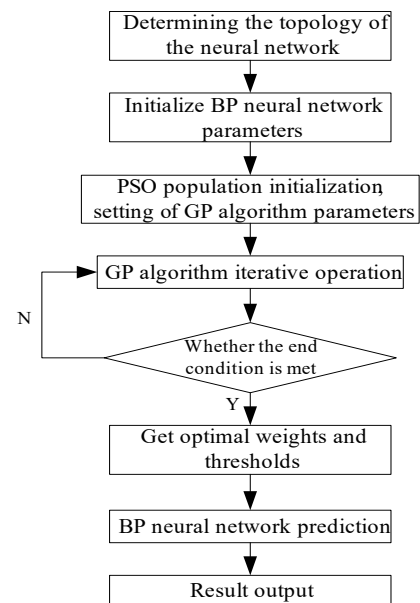


Figure 3 The flow chart of GP-BP algorithm

Fig. 3 shows the GP-BP algorithm flow chart. The main optimization process is described:

(1) Initializing topology of the neural network; that is, setting the number of input layers, hidden layers and output

layer nodes according to the problem;

(2) Initializing the population; encoding the weight and threshold information to be optimized according to a specific coding rule, and initializing the population according to the set population size and initialization method;

(3) Conducting iterative operation of the GP algorithm; the population obtained in Step (2) is iteratively operated according to GP algorithm rule, until the algorithm satisfies the condition. Finally, the global optimal individual is obtained, and the weight and threshold information are obtained by decoding;

(4) Initializing the weights and thresholds; bringing decoded weights and threshold information to obtain the initial weights and thresholds;

(5) GP-BP model operation; the repeated learning training of the network is carried out until the condition is met, and finally the output of the optimized BP neural network is obtained.

### 2.2.2 GP Algorithm

The genetic operation of GA algorithm was introduced into the PSO algorithm in this study. In the iterative process of particle swarm, the position and velocity vector of the particle were respectively cross-existed and mutated. The algorithm firstly evolved a certain algebra by PSO. Then according to the fitness value, particles were sorted from large to small. The population was divided into two parts.  $M$  particles with large fitness values were retained, and the remaining  $N - M$  particles were removed. The remained  $M$  particles were taken as the reference, and  $N - M$  individuals were obtained by copying. Next, the genetic algorithm operator operations such as cross mutation was carried out. Finally, the positional values of the  $M$  particles retained by the PSO and the  $N - M$  particles evolved by the GA algorithm were combined into a new particle swarm for the next generation of operations.

Fig. 4 shows the flow chart of GP algorithm. The algorithm is given, as follows:

(1) Setting of relevant parameters, including the number of particles  $N$ , the number of particles after evolution  $M$ , the probability of variation  $p_m$ , the probability of crossover  $P_c$ , the maximum velocity of particles in the particle swarm algorithm  $V$ , the inertia constant  $\omega$ , the learning factors  $c_1, c_2$  and the number of iterations epoch, the accuracy of the solution  $E$  and etc.

(2) Initialization of the population; setting the size  $N$  of the population and the dimension  $n$  of the particle, using real-number random coding to encode the particle's velocity and position, using fitness function to obtain the individual fitness value of the particle swarm, and improving the individual extremum of the particle and the global extremum. The fitness function used was the Eq. (11).

(3) Iteration of the particle swarm; updating the particle's velocity and position information based on Eq. (1) Eq. (2), and improving individual extremum and the global extremum of the population.

$$v_{i+1} = tv_i + m_1r_1(pb_i - x_i) + m_2r_2(gb_i - x_i) \quad (1)$$

$$x_{i+1} = x_i + v_{i+1} \quad (2)$$

where  $m_1, m_2$  is the accelerating factor,  $t$  is the inertia coefficient,  $r_1, r_2$  is a random number between  $[0, 1]$ ,  $v_i, v_{i+1}$  is the current speed of the particle and the updated speed,  $x_i, x_{i+1}$  is the current and updated position,  $pb_i$  is the current optimal position of the individual extreme value,  $gb_i$  is the global optimal extremum.

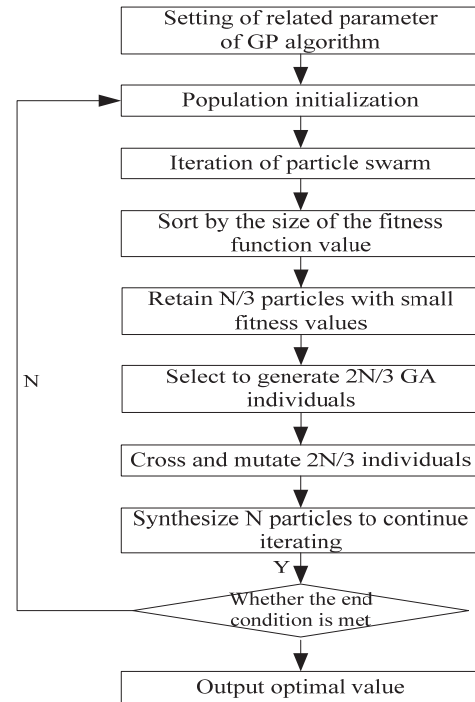


Figure 4 The flow chart of GP algorithm

(4) Sorting  $N$  particles according to the size of the fitness function value, selecting  $N/3$  particles with small fitness value, and selecting  $2N/3$  GA individuals by the position value of the retained  $N/3$  particles.

(5) GA individuals cross and mutate according to Eq. (3) and Eq. (4).

The crossover operation of the  $u$  chromosome  $\theta_k$  and  $v$  chromosome  $\theta_v$  at the  $j$  position is given:

$$\begin{cases} \theta_{uj} = \theta_{uj}(1 - \lambda) + \theta_{vj}b \\ \theta_{vj} = \theta_{vj}(1 - \lambda) + \theta_{uj}b \end{cases} \quad (3)$$

where  $\lambda$  is a random number between  $[0, 1]$ ,  $\theta_{uj}, \theta_{vj}$  is chromosomes obtained using crossover.

The operation method of mutating the  $j^{\text{th}}$  gene of the  $i^{\text{th}}$  individual is:

$$\theta_{ij} = \begin{cases} \theta_{ij} + \theta_{ij} - \theta & \lambda \geq 0.5_{\max} \\ \theta_{ij} + \theta_{ij} & \lambda < 0.5_{\min} \end{cases} \quad (4)$$

where the expression of  $f(d)$  is  $f(d) = \lambda \left( 1 - \frac{d}{H_{\max}^2} \right)$ ,  $d$  is the current number of iterations,  $H_{\max}$  is the largest number of evolutions,  $\lambda$  is a random number between  $[0, 1]$ .

(6) Combining  $2N/3$  individuals generated by GA algorithm with  $N/3$  particles to form  $N$  new particles, calculating the fitness of particle swarm, and updating individual extremum and global extremum.

**3 APPLICATION EXAMPLES AND RESULTS ANALYSIS**

**3.1 Variable Selection by MIV Algorithm**

**3.1.1 Experimental Data**

The literature [15] summarized the geological and mining factors of China's 408 working faces and the corresponding probability integral method parameters. In this paper, 50 working faces were selected as experimental data, in which the first 40 working face data were used as

training set, and the last 10 working face data were taken as testing set to test the model's prediction accuracy.

Commonly-used parameters in the probability integration method selected in this paper include  $q, b, \tan\beta, s$  (can be expressed by  $s/H$ ),  $\theta$ . The geological and mining conditions include  $f, H, M, T, n, W, X$  (first mining uses 0 as the input variable, repeated mining adopts 1 as the input variable). Tab. 1 shows some data.

**Table 1** Geological mining conditions and the probability integral method parameters

Number	Geological mining conditions							Probability integral method parameters				
	$f$	$H/m$	$M/m$	$T/^\circ$	$W/m$	$n$	$X$	$q$	$b$	$\tan\beta$	$\theta$	$s/H$
1	3.2	60	2.1	30	10	1	0	0.64	0.31	1.55	68	0.14
2	3.2	58	2.3	20	9	1	0	0.66	0.2	1.55	72	0.17
3	3.6	120	1.6	38	24	0.98	0	0.65	0.29	1.5	64	0.18
4	5	513	5.47	15	13	0.96	0	0.68	0.25	2.6	79	0.12
...	...	...	...	...	...	...	...	...	...	...	...	...
41	5.2	110	7	9	49	0.63	0	1.31	0.27	1.8	86	0.12
42	5.3	95	2.2	9	48	0.84	1	1.17	0.3	2.2	85.8	0.13
43	4	122	6.6	8	18	0.74	0	0.89	0.2	1.8	87	0.16
44	6	280	2.4	12	17.4	0.46	0	0.655	0.282	1.99	86	0.17
45	3	191	1.88	7	35.1	0.63	0	0.6	0.3	2.3	82	0.17
46	1.72	284	2.9	3.5	112	0.54	1	0.78	0.28	1.65	89	0.1
47	4.56	201	1.4	12	17	0.72	0	0.91	0.16	2	80	0.07
48	4	270	6	5	38	0.85	0	0.86	0.31	2.7	87	0.14
49	6.41	343	2.6	6	76	0.44	0	0.76	0.31	1.98	88	0.16
50	1.5	104	2.2	45	10	0.86	0	0.97	0.24	2.54	71	0.12

**3.1.2 Variable Selection**

Based on MATLAB program, the MIV algorithm program was programmed in this study. The adjustment rate of the MIV algorithm was set as 10%. Each probability integration method parameter was tested by 50 MIV

measurement experiments. After the absolute value of the MIV value obtained in each experiment, the weighted average was performed. Then, the MIV values of the influence factors corresponding to each probability integration method parameter could be obtained, as follows:

**Table 2** Different influencing factors correspond to MIV values

Parameters	$f$	$H$	$M$	$T$	$W$	$n$	$X$
$q$	0.9559	0.9140	0.5050	0.7391	0.7274	0.3744	0.0199
$b$	0.0211	0.0238	0.0246	0.0165	0.0259	0.0226	0.0036
$\tan\beta$	0.2120	0.2706	0.1299	0.1553	0.1396	0.2105	0.0200
$\theta$	0.3654	0.5352	0.3062	0.9606	0.4551	0.2422	0.0371
$s/H$	0.0349	0.0311	0.0185	0.0216	0.0223	0.0532	0.0065

The MIV values of each influencing factor were obtained and sorted from large to small. See Tab. 3.

**Table 3** Sorting results of MIV values of various influencing factors

Parameters	MIV value (big→small)						
	$f$	$H$	$T$	$W$	$M$	$n$	$x$
$q$	$f$	$H$	$T$	$W$	$M$	$n$	$x$
$b$	$W$	$M$	$H$	$n$	$f$	$T$	$x$
$\tan\beta$	$H$	$f$	$n$	$T$	$W$	$M$	$x$
$\theta$	$T$	$H$	$W$	$f$	$M$	$n$	$x$
$s/H$	$n$	$f$	$H$	$W$	$T$	$M$	$x$

According to the sorting results, the top five influence factors with large MIV values were selected as the input layer of the model, thereby achieving the goal of simplifying the network input and improving the prediction accuracy.

**3.2 De-Noise Processing**

Due to the complexity of the observation environment, the acquisition of experimental data also had certain errors, and the data fluctuation range was large. The error of the data itself was also introduced to further expand errors. Therefore, the data should be denoised before prediction to

minimize errors.

The Savitzky-Golay algorithm was applied to denoise the data, which is a filtering method based on local polynomial least squares fitting. It could ensure width and shape of the signal during the denoising process, making the data relatively smooth. The comparison between the processed results and the measured results are shown in Fig. 5a to Fig. 5f.

It can be seen from Fig. 5a to Fig. 5f and Tab. 4 that correlation coefficients after smoothing were all above 0.547, and the correlation coefficient was high. The smoothed experimental data was less up and down, and the curve tended to be smoother. Some points with more fluctuating were smoothed to achieve the effect of reducing the error.

**Table 4** The  $R$  value between the processed data and the measured data

Parameters	$f$	$H$	$M$	$T$	$W$	$n$
$R$	0.6722	0.8018	0.547	0.761	0.7064	0.7487

According to the data processing results, the correlation coefficient  $R$  between the denoised data and the measured data can be obtained, as shown in Tab. 4.

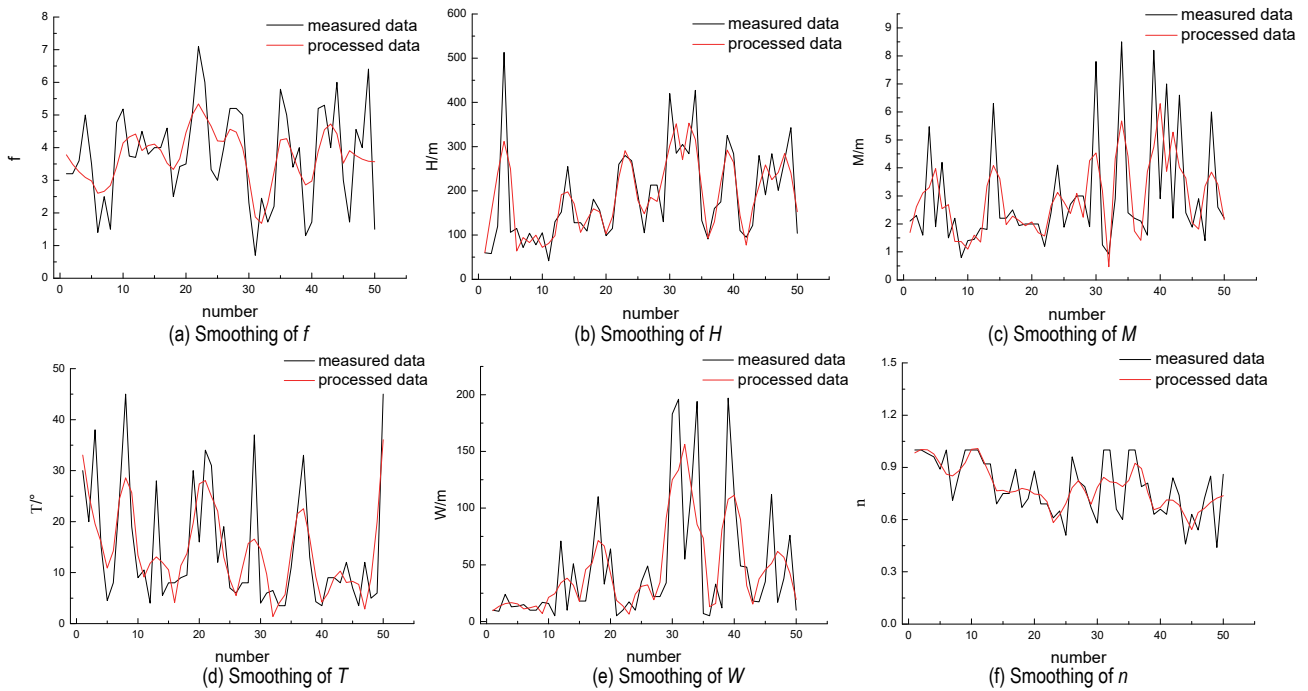


Figure 5 The processed data and the measured data

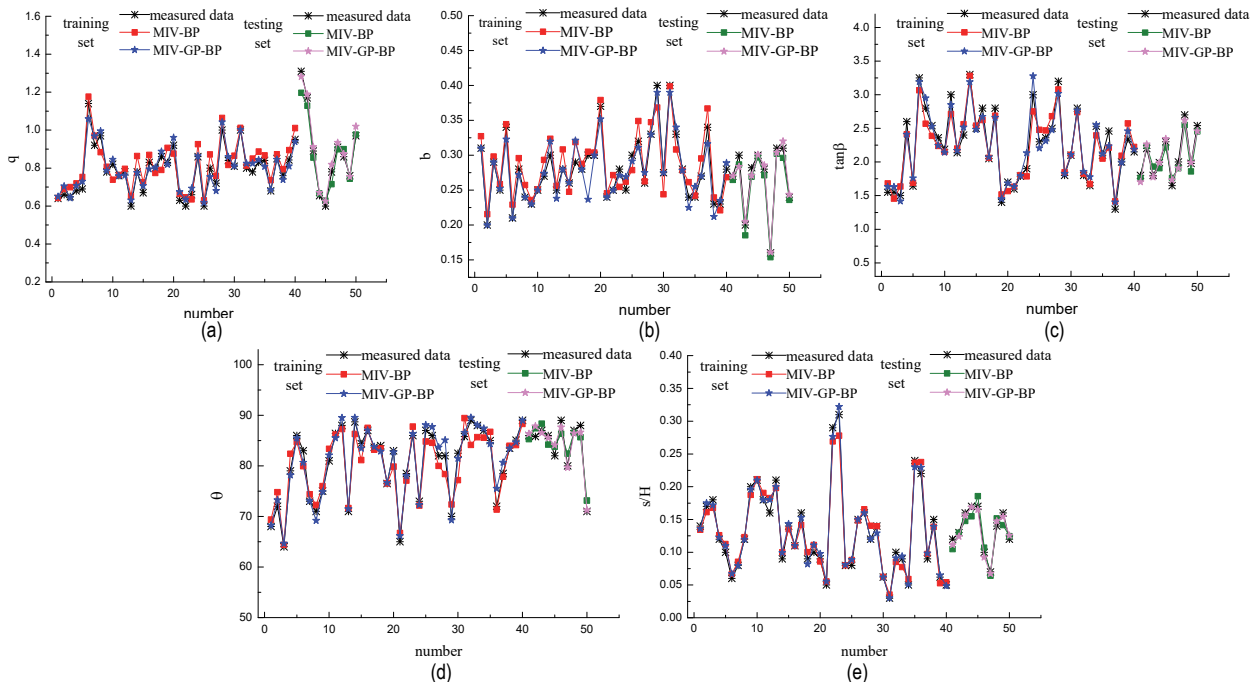


Figure 6 Comparison of predicted values before and after BP neural network optimization

**3.3 Prediction Results and Analysis**  
**3.3.1 The Results of the Model**

To reduce the complex structure of the neural network, a neural network was established for each probability integration method parameter, that is, a many-to-one topology (multiple input layer and single output layer) was adopted.

The parameters were set as follows: the training times were 5000, the learning rate was 0.01 and the training target error was 0.000001. The neural network in Fig.1 used a 3-layer structure. The number of the hidden layer was determined according to the formula  $J = 2m + 1$ , where  $m$  is the number of input layers. The sigmoid function was selected as the transfer function between the input layer

and the hidden layer, and the purelin function was the transfer function between the hidden layer and the output layer. In PSO algorithm, the population of the particles was 20, the maximum number of allowed iterations was 20, the acceleration factors  $c_1, c_2$  were 1.49445, and the maximum speed limit was 5. In the GA algorithm, the population number was 40, the crossover probability was 0.4, the mutation probability was 0.06, and the maximum number of iterations was 50.

To better verify the MIV-GP-BP model, the measured data of the first 40 working faces in Tab. 1 were taken as training set and the 41-50 working faces were taken as the testing set, meanwhile using the MIV-BP model as a predictive comparison. The results are shown in Fig. 6a to Fig. 6e.

It could be seen that after optimization by GP combination algorithm, BP neural network had better generalization ability between geological and mining conditions and parameters of probability integration method. The prediction accuracy was higher than that of non-optimized BP neural network at most points.

To quantitatively analyze the prediction performance differences between the optimized BP neural network prediction model and the non-optimized BP neural network model, the accuracy evaluation indexes of the prediction values of MIV-GP-BP prediction model and MIV-BP prediction model were calculated. In this paper, the maximum relative error (*MaxRE*), the mean relative error (*MeaRE*) and the root mean square error (*RMSE*) of the prediction results were taken as the accuracy evaluation indexes.

$$MaxRE = \max \left\{ \left| \frac{x'_i - x_i}{x_i} \right| \times 100\% \right\}_{i=1}^n \tag{5}$$

$$MeaRE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x'_i - x_i}{x_i} \right| \times 100\% \tag{6}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x'_i - x_i)^2} \tag{7}$$

In Eqs. (5), (6) and (7),  $x'_i$  is the predicted value,  $x_i$  is the actual value,  $n$  is the number of predicted samples.

The calculation results of the evaluation index are shown in Tab. 5.

**Table 5** The evaluation index of prediction model

Parameters	Evaluation indexes	MIV-BP		MIV-GP-BP	
		Training set	Testing set	Training set	Testing set
<i>q</i>	<i>MaxRE</i>	9.96	8.64	7.05	5.14
	<i>MeaRE</i>	6.71	4.93	3.21	2.78
	<i>RMSE</i>	0.0498	0.0478	0.0304	0.027
<i>b</i>	<i>MaxRE</i>	10.11	7.47	9.53	5.42
	<i>MeaRE</i>	6.27	4.54	3.06	2.15
	<i>RMSE</i>	0.0175	0.0101	0.0149	0.0077
$\tan\beta$	<i>MaxRE</i>	12.03	6.78	9.81	5.39
	<i>MeaRE</i>	5.05	4.54	3.79	2.81
	<i>RMSE</i>	0.1357	0.0999	0.1090	0.0666
$\theta$	<i>MaxRE</i>	6.46	3.02	4.89	2.57
	<i>MeaRE</i>	2.17	2.07	1.13	1.09
	<i>RMSE</i>	2.1634	1.8580	1.2076	1.1575
<i>s/H</i>	<i>MaxRE</i>	18.28	12.67	12.57	9.66
	<i>MeaRE</i>	7.89	7.81	4.72	4.31
	<i>RMSE</i>	0.0115	0.0120	0.0075	0.0058

Tab. 5 and Fig. 6a to Fig. 6e showed that the MIV-GP-BP prediction model had much smaller *MaxRE* and *MeaRE* values than those in MIV-BP prediction model in the five probability integration method parameters. The *RMSE* could effectively reflect the degree of deviation between predicted values and true values of the model, so the predicted value of the MIV-GP-BP model with the smallest *RMSE* value was closer to the true value. Considering that some predicted parameter values were small, small changes would cause a large change in relative error. However, the absolute error was small, so the accuracy was still high and could meet the engineering needs.

### 3.3.2 Correctness Validation of MIV Algorithm

In order to prove the effectiveness of MIV algorithm, five factors were randomly selected from the influencing factors as the input layer of the neural network to establish the GP-BP prediction model, and then compared with the MIV-GP-BP prediction model established by the MIV algorithm as the input layer. (At least one factor of the five random factors was not screened by the MIV algorithm). In this paper, three groups of geological and mining conditions were randomly selected. The subsidence factor prediction model and the horizontal movement factor prediction model were selected as the case analysis. The grouping scheme is shown in Tab. 6.

**Table 6** Random extraction of input layer factor scheme

Prediction models		1	2	3	4	5
The subsidence factor prediction model	Group MIV	<i>f</i>	<i>H</i>	<i>T</i>	<i>W</i>	<i>M</i>
	Group 1	<i>f</i>	<i>H</i>	<i>T</i>	<i>n</i>	<i>x</i>
	Group 2	<i>f</i>	<i>W</i>	<i>M</i>	<i>n</i>	<i>x</i>
	Group 3	<i>H</i>	<i>T</i>	<i>W</i>	<i>M</i>	<i>n</i>
The horizontal movement factor prediction model	Group MIV	<i>W</i>	<i>M</i>	<i>H</i>	<i>n</i>	<i>f</i>
	Group 1	<i>W</i>	<i>M</i>	<i>H</i>	<i>T</i>	<i>x</i>
	Group 2	<i>W</i>	<i>n</i>	<i>f</i>	<i>T</i>	<i>x</i>
	Group 3	<i>M</i>	<i>H</i>	<i>n</i>	<i>f</i>	<i>T</i>

For the above 8 sets of different input factor, the GP-BP prediction model was established by the data shown in Table 1. The prediction results are shown in Tab. 7, Tab. 8, Fig. 7 and Fig. 8.

From the prediction values of the four groups in Fig. 7 and Fig. 8 and the prediction evaluation index shown in

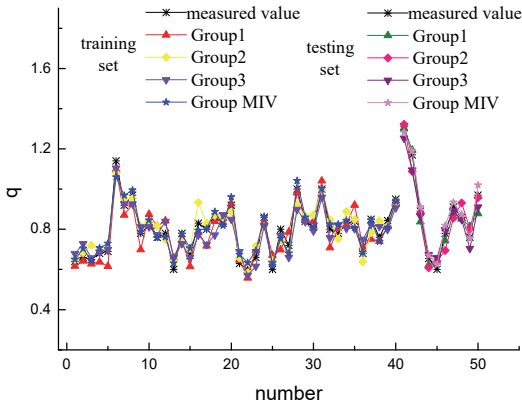
Tabs. 7 and 8, among the five parameters, the *MeaRE*, *MaxRE* and *RMSE* of each parameter have been improved in varying degrees. The MIV algorithm has simplified the input layer, reduced the complexity of the neural network and improved the prediction accuracy, thus proving the feasibility of the MIV algorithm.

**Table 7** Comparison of the evaluation index of the prediction model based on the subsidence factor of different input layers

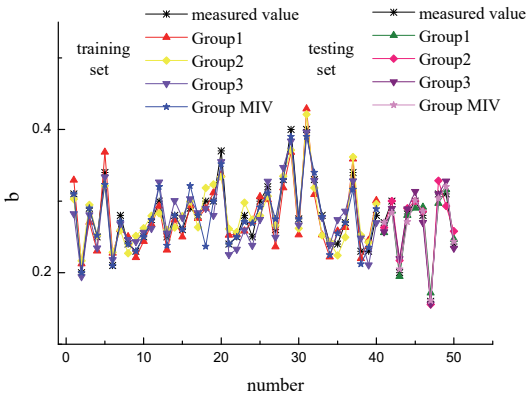
Evaluation index	Group MIV		Group 1		Group 2		Group 3	
	Training set	Testing set	Training set	Testing set	Training set	Testing set	Training set	Testing set
MaxRE	7.05	5.14	11.83	9.50	10.85	10.95	10.39	9.62
MeaRE	3.21	2.78	5.67	4.57	4.99	5.52	5.21	4.49
RMSE	0.0304	0.027	0.0507	0.0445	0.0447	0.0531	0.0469	0.0454

**Table 8** Comparison of the evaluation index of the prediction model based on the horizontal movement factor of different input layers

Evaluation index	Group MIV		Group 1		Group 2		Group 3	
	Training set	Testing set	Training set	Testing set	Training set	Testing set	Training set	Testing set
MaxRE	9.53	5.42	9.71	7.40	9.99	8.50	11.42	9.76
MeaRE	3.06	2.15	5.53	3.68	5.72	3.76	4.72	3.93
RMSE	0.0149	0.0077	0.0176	0.0102	0.0176	0.0118	0.0150	0.0117



**Figure 7** Prediction model predictive value of subsidence factor for different input layers



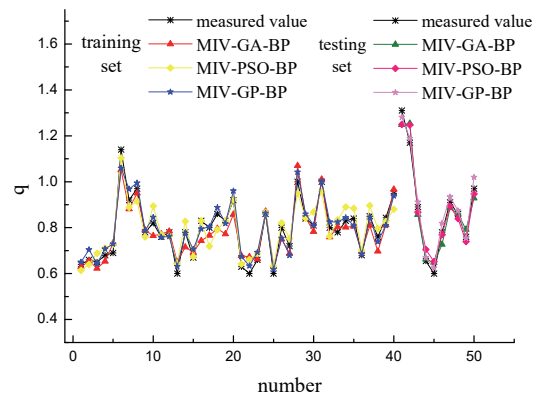
**Figure 8** Prediction model predictive value of horizontal movement factor for different input layers

### 3.3.3 Verify the Superiority of the GP Combination Algorithm

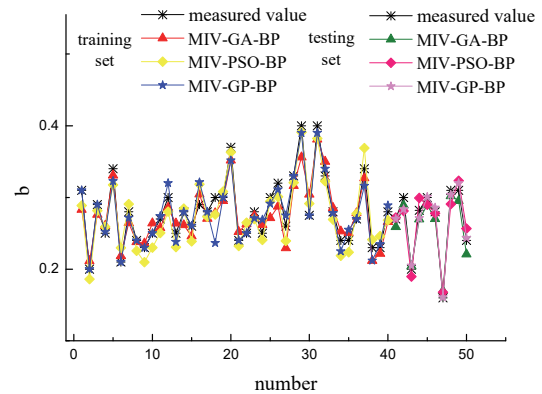
To verify that the established GP algorithm had better computing performance, GA algorithm and PSO algorithm were applied to optimize the BP neural network prediction model (MIV-GA-BP model and MIV-PSO-BP model). The input layer data was obtained from the MIV algorithm (the input layer data was the same). The training data and the testing data were derived from Tab. 1. The parameters of the relevant algorithms were the same as those described above.

This paper only establishes the prediction model of the

subsidence factor and the horizontal movement factor. The prediction results of other models were similar. To analyze the reliability of the results more intuitively, the training set and the testing set were compared and analyzed according to the built models. The results are shown in Tab. 9, Tab. 10, and Fig. 9 and Fig. 10.



**Figure 9** Predicted values of the subsidence factor prediction model based on different algorithms



**Figure 10** Predicted values of the horizontal movement factor prediction model based on different algorithms

From Tab. 9, Tab. 10, Fig. 9 and Fig. 10, we can see that MIV-GP-BP prediction model is superior to MIV-GA-BP prediction model and MIV-PSO-BP prediction model in prediction accuracy. It proves that MIV-GP-BP neural network prediction probability integration method has better accuracy and can meet the actual requirements of engineering technology.

**Table 9** Evaluation index of the subsidence factor prediction model based on different algorithms

Evaluation index	MIV-GP-BP		MIV-GA-BP		MIV-PSO-BP	
	Training set	Testing set	Training set	Testing set	Training set	Testing set
MaxRE	7.05	5.14	10.48	9.18	10.27	8.64
MeaRE	3.21	2.78	4.57	4.27	4.39	4.05
RMSE	0.0304	0.027	0.0427	0.0451	0.0405	0.0412

**Table 10** Evaluation index of the horizontal movement factor prediction model based on different algorithms

Evaluation index	MIV-GP-BP		MIV-GA-BP		MIV-PSO-BP	
	Training set	Testing set	Training set	Testing set	Training set	Testing set
<i>MaxRE</i>	9.53	5.42	10.14	7.78	9.76	6.94
<i>MeaRE</i>	3.06	2.15	4.55	4.27	4.58	4.48
<i>RMSE</i>	0.0149	0.0077	0.0143	0.0117	0.0141	0.0133

#### 4 CONCLUSIONS

The acquisition of parameters of PIM is a hotspot and a difficult point in the data processing of surface deformation monitoring in coal mines. To solve the problem that the parameter accuracy of the traditional BP neural network prediction probability integration method is not high, this paper has proposed a MIV-GP-BP prediction model. This model used a combination of MIV algorithm and GP combinational algorithm to optimize BP neural network. The main conclusions are as follows:

(1) Taking the measured data of 50 typical mining areas as an example, the denoising process and MIV algorithm were applied to optimize the input of BP neural network. The MIV algorithm was used to sort the geological and mining factors affecting the parameters of the PIM. The factors with small influence factors were eliminated and the remaining factors were taken as the input. The smooth denoising method was adopted to eliminate the accidental errors that may be contained in the measured data to replace the original input layers. The prediction accuracy was improved by analysis.

(2) The integration of the algorithm of PSO and GA was applied to optimize the weight and threshold of BP neural network. The MIV algorithm was used to decrease the topology and establish MIV-GP-BP PIM model. The prediction results showed that among the five parameters, the *RMSE* was between 0.0058 and 1.1575, the *MaxRE* of  $q$ ,  $\tan\beta$ ,  $b$  and  $\theta$  was less than 5.42%, and the *MeaRE* was less than 2.81%. The *RMSE* of  $s/H$  did not exceed 0.0058, the *MaxRE* was less than 9.66% and the *MeaRE* was less than 4.31%. MIV-BP model, MIV-GA model and MIV-PSO model has better accuracy than that of BP model.

Due to limited resources, the amount of data collected for modeling in this article is not enough, and more data will be added for verification in the subsequent application process.

#### Acknowledgments

The work was supported by the National Natural Science Foundation of China (Grant numbers 41602357, 41474026), China Coal XinjiEnergy Co. Ltd (Grant number ZMXJ-LZ-JS-2018-25); Huainan Mining Group Co. Ltd (Grant number HNKY-JTJS (2017)-122, HNKY-JTJS (2018)-178).

#### 5 REFERENCES

- [1] Guo, Q. B., Guo, G. L., Lv, X., Zhang, W. X., Lin, Y. K., & Qin, S. Y. (2016). Strata movement and surface subsidence prediction model of dense solid backfilling mining. *Environmental Earth Sciences*, 75(21), 1426. <https://doi.org/10.1007/s12665-016-6237-6>.
- [2] Wang, L., Li, N., Zhang, X. N., Wei, T., Chen, Y. F., & Zha, J. F. (2018). Full parameters inversion model for mining subsidence prediction using simulated annealing based on single line of sight D-InSAR. *Environmental Earth Sciences*, 77(5), 161. <https://doi.org/10.1007/s12665-018-7355-0>
- [3] Zha, J. F., Feng, W. K., Zhu, X. J. (2011). Research on parameters inversion in probability integral method by genetic algorithm. *Journal of Mining and Safety Engineering*, 28(4), 32-36.
- [4] Zhou, D. W., An, S. K., Zhang, J., & Zhang, L. M. (2013). A principal component regression analysis for calculating probability integration method parameters. *Journal of Liaoning Technical University (Natural Science)*, 32(11), 1528-1532.
- [5] Wang, Q. C., Li, J., & Guo, G. L. (2018). Study on analogy analysis on prediction of surface subsidence parameters in mining area of Chongqing City. *Coal Science and Technology*, 46(10), 196-202.
- [6] Yu, R. Y., An, X. M., Jin, B., Shi, J., Move, O. A., & Liu, Y. H. (2018). Particle classification optimization-based BP network for telecommunication customer churn prediction. *Neural Computing and Applications*, 29(3), 707-720. <https://doi.org/10.1007/s00521-016-2477-3>
- [7] Li, D. J., Li, Y. Y., Li, J. X., & Fu, Y. (2018). Gesture recognition based on BP neural network improved by chaotic genetic algorithm. *International Journal of Automation and Computing*, 15(3), 1-10. <https://doi.org/10.1007/s11633-017-1107-6>
- [8] Bian, H. F., Yang, H. C., & Zhang, S. B. (2013). Research on intelligent optimization for predicting parameters of probability-integral method. *Journal of Mining and Safety Engineering*, 30(3), 385-389.
- [9] Zhao, Z. M., Shi, T. W., Dong, W., & Liu, Y. L. (2017). The prediction of probability integral method parameters based on grey relation analysis and BP neural network. *Science of Surveying and Mapping*, 42(7), 36-40.
- [10] Hu, B. N., Zhang, H. X., & Sheng, B. H. (2017). Regulations on mining under buildings, water bodies and railways and coal safety pillars in main roadway, Beijing, China, 99-158.
- [11] Ali, B., Saleheh, A., Sajjad, G., Marzieh, D. T., & Saeed, M. (2017). Spectrophotometric determination of synthetic colorants using PSO-GA-ANN. *Food Chemistry*, 220, 377-384. <https://doi.org/10.1016/j.foodchem.2016.10.010>
- [12] Mousavi, M., Yap, H. J., & Musa, S. N. (2017). A fuzzy hybrid GA-PSO algorithm for multi-objective AGV scheduling in FMS. *International Journal of Simulation Modelling*, 16(1), 58-71. [https://doi.org/10.2507/IJSIMM16\(1\)5.368](https://doi.org/10.2507/IJSIMM16(1)5.368)
- [13] Cui, Z. Q., Fu, X. Y., Zhong, S. S., & Wang, T. C. (2013). Aero-engine arguments selection based on wavelet network mean impact value. *Computer Integrated Manufacturing Systems*, 19(12), 3062-3067.
- [14] Xu, L. B., Wang, W., Zhang, T., Yang, L., Wang, S. Y., & Li, Y. D. (2017). Ultra-short-term Wind Power Prediction Based on Neural Network and Mean Impact Value. *Automation of Electric Power Systems*, 41(21), 40-45.
- [15] State Administration of Work Safety (2017.). Regulations on mining under buildings, water bodies and railways and coal safety pillars in main roadway.



**Contact information:**

**Shenshen CHI**

1) School of Earth and Environment,  
Anhui University of Science and Technology,  
Huainan 232001, China

2) School of Geomatics, Anhui University of Science and Technology,  
Huainan 232001, China

3) Key Laboratory of Aviation-aerospace-ground Cooperative Monitoring and  
Early Warning of Coal Mining-induced Disasters of Anhui Higher Education  
Institutes, Anhui University of Science and Technology, KLAHEI (KLAHEI18015),  
Huainan 232001, China

4) Coal Industry Engineering Research Center of Mining Area Environmental  
and Disaster Cooperative Monitoring,  
Anhui University of Science and Technology,  
Huainan 232001, China  
E-mail: austcss@163.com

**Xuexiang YU**

(Corresponding author)

School of Geomatics, Anhui University of Science and Technology,  
Huainan, 232001, China  
E-mail: xxyu@aust.edu.cn

**Lei WANG**

School of Geomatics, Anhui University of Science and Technology,  
Huainan, 232001, China  
E-mail: austwlei@163.com