

## Research Article

# Advanced Computational Methods for Mitigating Shock and Vibration Hazards in Deep Mines Gas Outburst Prediction Using SVM Optimized by Grey Relational Analysis and APSO Algorithm

Xiang Wu <sup>1</sup>, Zhen Yang <sup>1</sup>, and Dongdong Wu <sup>2</sup>

<sup>1</sup>School of Mines, China University of Mining & Technology, Xuzhou 221116, China

<sup>2</sup>Huayang New Material Technology Group Co., Ltd., Yangquan 045000, China

Correspondence should be addressed to Xiang Wu; lanran@cumt.edu.cn

Received 4 February 2021; Revised 19 March 2021; Accepted 3 April 2021; Published 26 April 2021

Academic Editor: Junfei Zhang

Copyright © 2021 Xiang Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Gas outburst poses a huge threat to the safe production of coal mines. Therefore, the prediction of gas outburst has always been a hot topic for researchers. In recent years, the use of artificial intelligence algorithms for gas outburst prediction has made progress, such as using BP neural network, GA algorithm, and SVM algorithm. Despite these progresses, predicting the gas outburst more accurately and efficiently still remains a great challenge. In this work, an algorithm based on grey relational analysis and SVM using adaptive particle swarm optimization (APSO-SVM) for gas outburst prediction is proposed. Grey relational analysis was used to extract the four most relevant ones from nine gas outburst prediction parameters (geological structure zone distance, coal seam gas content, gas release initial velocity, gas desorption index-K1, drill cuttings volume, coal seam depth, coal seam thickness, coal destruction type, and coal firmness coefficient) to give the needed parameters for SVM model. Higher prediction accuracy was then obtained with the selected parameters composed by coal seam gas content, gas release initial velocity, gas desorption index-K1, and drill cuttings volume. Moreover, adaptive particle swarm optimization (APSO) was used to optimize the penalty factor and kernel parameters of the support vector machine to improve the global search ability and avoid the occurrence of the local optimal solutions. The APSO-SVM model was applied to the prediction of gas outburst in 31004 tunneling face of Xinyuan Coal Mine, Yangquan City, Shanxi Province, China. We further introduced the criteria of accuracy, precision, recall, and  $F_2$ -score to evaluate the prediction results of different models. The results show that, in the gas outburst prediction, the accuracy of the APSO-SVM model is 98.38%, the precision and recall are both 100%, and  $F_2$ -score is 1. Comparative studies confirm that APSO-SVM displayed better performance than SVM and PSO-SVM models for the applied grey relational analysis assisted gas outburst prediction. These obtained results indicate the validity of APSO-SVM model for gas outburst prediction.

## 1. Introduction

In recent years, with the extension of coal mining level and the increase of mining depth, gas outburst has had an increasing impact on safe production of coal mines. Gas outburst threatens the lives of coal miners and also causes huge economic losses to the country [1–3]. For a long time, the study of gas outburst prediction has been a hot topic for researchers. Currently, the commonly used prediction methods for gas outbursts mainly include single index method, comprehensive index method [4], gas geology unit methods [5], and geophysical methods [6, 7]. However, gas

outburst is a nonlinear dynamic process coupled by variety of factors. The prediction accuracy is not satisfactory relying only on a certain indicator or simple combination of few indicators. Therefore, using nonlinear artificial intelligence technology to predict gas outburst has become a research hotspot. At present, some researchers have made good research progress in using artificial intelligence methods to predict gas outburst. Yingjie Li et al. established a coupled forecasting model and achieved good forecasting results by combining Dempster-Shafer theory and traditional forecasting methods [8]. Dong Chunyou and others proposed G-K evaluation and a rough set model to analyze and classify

gas outburst grades [9]; Cheng Xia and Xu Manguan analyzed the factors affecting gas outburst based on the gray system theory and selected the main control factors to simplify the prediction process and increase the prediction speed [10, 11]. Zhou Xihua [12] used the principal component analysis method to assist the reverse BP neural network to reduce the dimensionality of the input factors and predict the gas outburst, which improved the accuracy of prediction results compared with traditional methods. However, the performance of BP neural network depends on the selection of initial value; it is easy to fall into local optimal solution in the process of gradient descent calculation. Liu Yijun, Hao Wenli, and others [13] established a GA-BP hybrid model to predict gas concentration; the GA-BP model used the improved genetic algorithm to optimize the weights and thresholds of BP neural network, which avoided occurrence of the local optimal solutions and achieved a more accurate prediction results compared with conventional BP network. Compared with BP neural network, SVM performs better in small sample data processing [14]. It has been widely used in many fields, such as troubleshooting [15], image recognition [16], and medical diagnosis [17]. Qu Fang used the SVM model to predict gas outburst, and his work showed the effectiveness of SVM in processing small sample data [18]. Qiao Meiyang used LS-SVR to process time series data for gas concentration prediction and achieved good results [19]. The purpose of the SVM algorithm is to find support vectors [20]. The optimization goal is to minimize the structural risk rather than the experience risk, and the performance of the SVM model depends largely on the selection of kernel function parameters and penalty factors. For a specific sample, selecting the most suitable kernel function parameter  $\sigma$  and penalty factor  $c$  can make SVM model have both learning ability and generalization ability. Although the commonly used conventional methods such as cross-validation method and grid search method are relatively simple to implement, they are computationally intensive, and the classification accuracy is not good enough. Therefore, how to obtain the best calculation parameters for a given sample data is a problem to be solved. To solve the problem, this paper proposed an algorithm based on grey relational analysis and SVM using adaptive particle swarm optimization (APSO-SVM). The sample data of the experiment were taken from the 31004 driving face of Xinyuan Coal Mine, Yangquan City, Shanxi Province, China. We first used the grey relational analysis method to analyze the correlation degree of the index parameters used for gas outburst prediction. The 5 parameters with lower correlation degree were eliminated, and the 4 most influential parameters were selected as input parameters to improve the model's performance in stability and accuracy. Then, APSO method was used to optimize the penalty factor and kernel parameters of the support vector machine. The established APSO-SVM model was applied to the gas outburst prediction of the test sample and compared with the SVM model and PSO-SVM model. We further introduce four indicators (accuracy, precision, recall, and  $F_2$ -score) to

evaluate the prediction results. The results show that the accuracy of the APSO-SVM model is 98.38%, the precision and recall are both 100%, and  $F_2$ -score is 1. The APSO-SVM model outperforms the SVM model and the PSO-SVM model.

## 2. Site and Data

The experimental samples in this paper are taken from the #3 coal seam 31004 driving face of Xinyuan Coal Mine, Yangquan City, Shanxi Province, China. Xinyuan Coal Mine is recognized as gassy mine. Its geographic location is shown in Figure 1(a). The #3 coal seam is currently being mined, and the risk of gas outburst is high. The comprehensive histogram of #3 coal seam is shown in Figure 1(b). Using the data collected from this coal seam for experimental analysis is representative.

According to the analysis of the mechanism of gas outburst, the main factors affecting outburst are hydro-geological conditions, gas storage conditions, and coal structural parameters. According to the actual construction conditions of the #3 coal seam, we collected 162 sets of sample data containing 9 parameters (geological structure zone distance, coal seam gas content, gas release initial velocity, gas desorption index-K1, drill cuttings volume, coal seam depth, coal seam thickness, coal destruction type, and coal firmness coefficient) through combination of on-site collection and review of the Xinyuan Coal Mine files. Since it is difficult to obtain samples of gas outbursts that actually occur on site, blowout holes, top drills, and other dynamic phenomena that represent outbursts during drilling are used as the basis for outburst danger. According to the actual situation of the site, if no measures are taken, it is considered that there is no gas outburst risk, and we use the number 1 to indicate this situation. Drilling 12–15 groups of pressure relief holes means that the outburst risk is general, and the number 2 is used to indicate this situation. Drilling 20 groups of pressure relief holes means that the outburst risk is serious, and the number 3 is used to indicate this situation. Detailed data are shown in Table 1.

## 3. Methods

*3.1. Using Grey Relational Analysis to Identify Prediction Parameters.* For SVM models, too many input parameters may cause the model to overfit, and the convergence will also be affected. This paper collected data of 9 indexes as input parameters. There may be correlations between input parameters, and the correlations between input parameters and experimental results are different. Therefore, it is necessary to find and strengthen the input parameters that are highly related to the experimental results and weaken or eliminate input parameters that have low correlation with experimental results to improve the convergence speed and prediction accuracy of the model. Grey relational analysis can effectively analyze the correlation between input parameters

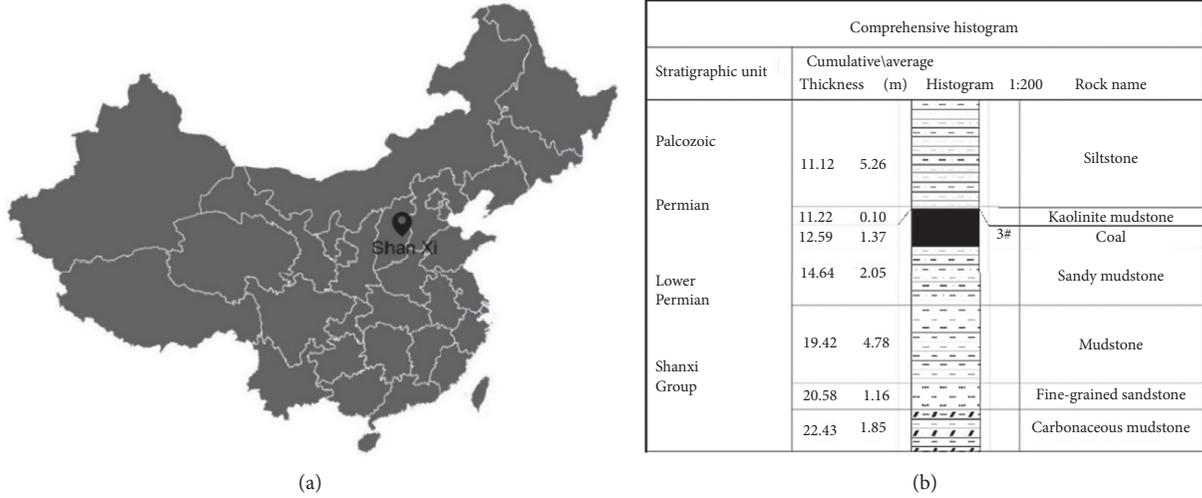


FIGURE 1: The location of the test sample collection and the geological parameters of the coal seam. (a) Location of test sample collection. (b) #3 coal seam geological parameters.

TABLE 1: Part sample data of #3 coal seam 31004 driving face.

Number	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$	$X_9$	Risk
1	3.00	15.50	0.36	7.78	0.30	3.40	36.00	351.00	2.40	1
2	3.00	15.70	0.36	7.76	0.20	3.60	62.00	366.00	2.40	1
3	2.00	15.75	0.36	7.81	0.30	3.20	78.00	366.00	2.40	1
4	2.00	15.56	0.36	7.30	0.30	3.80	47.00	367.00	2.40	1
5	3.00	15.80	0.36	7.05	0.27	3.00	21.00	364.00	2.40	1
6	3.00	18.55	0.36	9.92	0.30	3.4	15.00	363.00	2.60	2
7	3.00	15.69	0.36	7.46	0.28	3.40	35.00	365.00	2.40	1
8	3.00	17.66	0.28	9.32	0.35	0.31	26.00	366.00	2.10	2
9	3.00	17.56	0.36	10.06	0.21	0.32	15.00	366.00	2.10	2
10	3.00	16.69	0.30	13.88	0.38	3.4	9.00	368.00	2.10	3

Note:  $X_1$  means coal destruction type;  $X_2$  means gas release initial velocity;  $X_3$  means coal firmness coefficient;  $X_4$  means coal seam gas content;  $X_5$  means gas desorption index-K1;  $X_6$  means drill cuttings volume;  $X_7$  means geological structure zone distance;  $X_8$  means coal seam depth;  $X_9$  means coal seam thickness; "Risk" means the gas outburst risk.

and experimental results [21]. The specific analysis process is as follows:

- (1) Set the reference sequence, denoted as

$$X_0 = (X_0(1), X_0(2), \dots, X_0(n)). \quad (1)$$

- (2) Set the comparison sequence, denoted as

$$(X_1, X_2, \dots, X_m) = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_m(1) \\ x_1(2) & x_2(2) & \dots & x_m(2) \\ \vdots & \vdots & \vdots & \vdots \\ x_1(n) & x_2(n) & \dots & x_m(n) \end{bmatrix}. \quad (2)$$

In the formula,  $m$  is the number of evaluation objects, and  $n$  is the number of evaluation parameters.  $X_i = (x_i(1), x_i(2), \dots, x_i(n))^T$ ,  $i = 1, 2, \dots, m$ .

- (3) Nondimensionalization data. The physical meaning and dimension of each parameter of gas outburst prediction are different, these parameters are not easy to compare, or it is difficult to get correct conclusions when comparing. Therefore, when performing grey relational analysis, nondimensional data processing is generally required. Commonly used dimensionless processing methods include extreme value method, average value method, and initial value method. Among them, the extreme value method formula is

$$x'_p(q) = \frac{x_p(q) - x_{\min}}{x_{\max} - x_{\min}}. \quad (3)$$

In the formula,  $p$  is the evaluation object,  $q$  is the evaluation parameter,  $i = 1, 2, \dots, m$ ,  $k = 1, 2, \dots, n$ , and  $x_{\max}$  and  $x_{\min}$  are the maximum and minimum values of the sequence.

The calculation formula of the initial value method is

$$x'_p(q) = \frac{x_p(q)}{x_1(q)}. \quad (4)$$

The calculation formula of the averaging method is

$$x'_p(q) = \frac{x_p(q)}{1/m \sum_1^m x_p(q)}. \quad (5)$$

The dimensionless data are as follows:

$$\xi_p(q) = \frac{\min_p \min_q |x'_0(q) - x'_p(q)| + \rho \cdot \max_p \max_q |x'_0(q) - x'_p(q)|}{|x'_0(q) - x'_p(q)| + \rho \cdot \max_p \max_q |x'_0(q) - x'_p(q)|}. \quad (7)$$

Here,  $\xi_p(q)$  is the grey correlation coefficient between the  $p$ -th evaluation object and the  $q$ -th parameter;  $\rho$  ( $0 < \rho < 1$ ) is the resolution coefficient; if  $\rho$  is smaller, the difference between correlation coefficients is larger, and the distinguishing ability is stronger. Generally,  $\rho = 0.5$ .

- (5) Calculate the degree of grey correlation. The correlation coefficient is the value of the correlation between each element in the comparison sequence and the corresponding element of reference sequence; the correlation coefficient of each element of the comparison sequence needs to be averaged as a measure of the degree of correlation between the comparison sequence and the reference sequence, denoted as

$$r_p = \frac{1}{n} \sum_{q=1}^n \xi_p(q). \quad (8)$$

- (6) Sort the grey correlation degree. Sort the calculated gray correlation degree  $r_p$ . If  $r_1 > r_2$ , it means that the correlation between the comparison sequence  $r_1$  and the reference sequence  $r_0$  is greater. The parameter of sequence  $r_1$  plays a greater role in gas outburst than  $r_2$ . If the calculated value of  $r_i$  is the largest, it indicates that the comparison sequence  $r_i$  occupies a dominant position in the entire system.

**3.2. Support Vector Machine Model.** Support vector machine is based on statistical theory [22]. It is currently one of the most popular machine learning methods. Its core lies in the minimum structural risk, and the overall generalization ability is improved by controlling experience risk and confidence range. The method is to map the samples that cannot be classified in the low-dimensional space to the high-dimensional space through a nonlinear transformation to make the samples linearly separable.

For a given sample,  $Z = \{(x_i, y_i)\}_{i=1}^L$ ,  $x \in R^n$ ,  $y \in \{+1, -1\}$ ,  $L$  is the number of samples, and  $n$  is the

$$(X'_0, X'_1, \dots, X'_m) = \begin{bmatrix} x'_0(1) & x'_1(1) & \dots & x'_m(1) \\ x'_0(2) & x'_1(2) & \dots & x'_m(2) \\ \vdots & \vdots & \vdots & \vdots \\ x'_0(n) & x'_1(n) & \dots & x'_m(n) \end{bmatrix}. \quad (6)$$

- (4) Calculate the grey correlation coefficient between corresponding elements of each comparison sequence and reference sequence:

input sample dimension. Construct a classification hyperplane to maximize the interval between all samples. The hyperplane can be expressed as the following formula:

$$w\varphi(x) + b = 0. \quad (9)$$

In order to solve this hyperplane, the problem is transformed into

$$\begin{cases} \min \frac{1}{2} w^2 + c \sum_{i=1}^L \xi_i, \\ \text{s.t.} \begin{cases} y_i(w^T \varphi(x_i) + b) \geq 1 - \xi_i & i = 1, 2, \dots, L, \\ \xi_i \geq 0, \end{cases} \end{cases} \quad (10)$$

where  $c$  is the penalty coefficient and  $\xi_i$  is the slack variable.

Formula (11) is a quadratic optimization problem, introducing Lagrangian multipliers, and the problem can be transformed into a dual problem:

$$\begin{cases} \min_{\alpha} \frac{1}{2} \sum_{i=1}^L \sum_{j=1}^L y_i y_j \alpha_i \alpha_j K(x_i, x_j) - \sum_{i=1}^L \alpha_i, \\ \text{s.t.} \begin{cases} \sum_{i=1}^L \alpha_i y_i = 0, \\ 0 \leq \alpha_i \leq c & i = 1, 2, \dots, L, \end{cases} \end{cases} \quad (11)$$

where  $K(x_i, x_j) = \varphi(x_i) \varphi(x_j)$  is the kernel function.

The solution  $\alpha^*$  of the above problem can be obtained using the SMO algorithm, and then  $w^*$  and  $b^*$  can be obtained. After the above derivation, the optimization function is

$$f(x) = \text{sgn}(w^* \varphi(x) + b^*) = \text{sgn} \left( \sum_{i=1}^L \alpha_i^* y_i K(x_i, x) + b^* \right). \quad (12)$$

The RBF kernel function has a wide range of convergence and can adapt to various samples. Therefore, this paper chooses RBF as the kernel function, and its expression is as follows:

$$k(x, x_i) = \exp\left(-\frac{x - x_i^2}{\sigma^2}\right). \quad (13)$$

**3.3. SVM Optimization Based on Adaptive Weighted Particle Swarm Optimization.** The performance of the SVM model depends on the kernel function parameter  $\sigma$  and the penalty factor  $C$ . The parameter  $\sigma$  is used to control the width of the Gaussian distribution of the sample. If  $\sigma$  is much smaller than the minimum distance between training samples, all samples become support vectors, which results in “overfitting” phenomenon. For new samples, the SVM model has a poor classification ability. If  $\sigma$  is much larger than the maximum distance between training samples, all samples are classified into one category, and the SVM model has no learning ability. The penalty factor  $C$  is the penalty coefficient for the deviation value of the wrong sample, which can adjust the ratio of the empirical risk and the confidence range of the SVM model to improve the generalization ability. When  $C$  becomes larger, it means that all constraint conditions tend to be satisfied, the fit of the data becomes higher, and the classification effect is better, but the generalization ability decreases. When  $C$  becomes smaller, the penalty for empirical error is smaller and the complexity of the model decreases, but the empirical risk increases. Therefore, the SVM model can obtain better learning ability and generalization ability by choosing the appropriate parameters  $\sigma$  and  $C$ . The conventional kernel function parameter optimization method has low efficiency and low accuracy. Compared with the conventional method, the swarm intelligence algorithm has achieved good application effects in the SVM kernel function parameter optimization [23]. At present, swarm intelligence algorithms commonly used in this field include genetic algorithm [24], particle swarm algorithm [25], ant colony algorithm [26], and leapfrog algorithm [27].

The advantage of genetic optimization algorithm’s SVM is that it can avoid the occurrence of the local optimal solutions, but the disadvantage is that it is greatly affected by the initial value. The advantage of ant colony optimization algorithm’s SVM is higher accuracy, but the disadvantage is that the algorithm has slow convergence speed and poor robustness. The advantage of the frog-leap optimization algorithm’s SVM is that the solution accuracy and robustness are high, but the algorithm’s convergence speed is slow and its implementation is complicated. The particle swarm optimization algorithm is currently the most widely used swarm intelligence algorithm in the field of SVM kernel function parameter

optimization. It has fast convergence speed and high global optimization performance, but it is easy to fall into a local optimal solution.

This paper uses an improved particle swarm optimization algorithm to optimize SVM, assuming that the number of particles in the population is  $m$  and the search space of the population is  $D$ -dimensional; the population can be expressed as  $X = (X_1, X_2, \dots, X_m)$ . The space position of the  $i$ -th ( $1 \leq i \leq X_m$ ) particle is  $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ , and the flying speed of the  $i$ -th ( $1 \leq i \leq X_m$ ) particle is  $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ ; the speed and position update formulas are as follows:

$$\begin{aligned} v_{id}(t+1) &= wv_{id}(t) + c_1r_1[p_{id}(t) - x_{id}(t)] \\ &\quad + c_2r_2[p_{gd}(t) - x_{id}(t)], \\ x_{id}(t+1) &= x_{id}(t) + v_{id}(t+1). \end{aligned} \quad (14)$$

In the formula,  $t$  is the number of iterations.  $d$  is the search dimension, and the target parameters optimized in this paper are  $\sigma$  and  $C$ , so  $d = 2$ .  $I = 1, 2, \dots, m$  is the number of particles.

$c_1$  is the individual learning factor, and  $c_2$  is the social learning factor.  $r_1$  and  $r_2$  are random numbers distributed in the interval  $[0, 1]$ .  $p_{id}(t)$  is the historical best position of the particle at the  $t$ -th iteration.  $p_{gd}(t)$  is the best historical position of the  $t$ -th iteration of the population.  $v_{id}(t)$  is the flight speed of the  $i$ -th particle in the  $d$ -dimensional space;  $v_{id} \in [v_{\min}, v_{\max}]$ .  $x_{id}$  is the position of the  $i$ -th particle in the  $d$ -dimensional space.  $w$  is the inertia weight. The flying speed and position of the particles are adjusted by the adaptive inertia weight. The expression is as follows:

$$w = \begin{cases} w_{\min} + \frac{(w_{\max} - w_{\min}) * (f - f_{\min})}{f_{\text{avg}} - f_{\min}}, & f \leq f_{\text{avg}}, \\ w_{\max}, & f > f_{\text{avg}}. \end{cases} \quad (15)$$

When the target value of each particle tends to be the same or tends to the local optimum, the inertia weight will be increased. When the target value of each particle is relatively scattered, the inertia weight will be reduced. At the same time, for the particle whose objective function value is better than the average target value, the corresponding inertia weight is reduced, thereby retaining its particle. On the contrary, for the particle whose objective function value is worse than the average target value, the corresponding inertia weight is increased to make the particle move closer to a better search area. The position  $x_{id}$  updated each time is used as two parameters of SVM for sample training, and the accuracy of training samples is used as the optimization target of particle swarm algorithm.

3.4. *Model Evaluation Index.* This paper introduces four indicators to evaluate the prediction effect of the model:

$$\begin{aligned} \text{accuracy} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}, \\ \text{precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\ \text{recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ F_{\beta} - \text{score} &= (1 + \beta^2) \frac{\text{precision} \cdot \text{recall}}{\beta^2 \cdot \text{precision} + \text{recall}}. \end{aligned} \quad (16)$$

The gas outburst risk is divided into 3 levels: serious, normal, and nonhazardous. We define the serious level as positive sample and the normal and nonhazardous levels as negative sample. TP is the number of correctly classified in the positive sample, FP is the number of incorrectly classified in the negative sample, FN is the number of incorrectly classified in the positive sample, and TN is the number of correctly classified in the negative sample. Accuracy reflects the classification performance of the model. Precision represents the proportion of real positive samples in the data that we predict as positive samples. It reflects a measure of the cost of misjudging a negative sample. Recall is also called “sensitivity,” which represents the proportion of all positive samples that are correctly predicted by us. It reflects the model’s ability to recognize positive samples. In a large-scale data collection, the two indicators “precision” and “recall” are often mutually restricted, so a trade-off needs to be made in the actual evaluation.  $F_{\beta}$  - score is the harmonic value of precision and recall. When  $\beta = 1$ , we think that precision and recall are equally important. If we think that precision is more important, we adjust  $\beta$  to be less than 1, but if we think recall is more important, we adjust  $\beta$  to be larger than 1. For serious level, if it is misjudged as normal or nonhazardous, it will cause major safety hazards. Therefore, we hope that the recall can be as large as possible, while the precision meets the requirements. Recall is more important, and the value of  $\beta$  in this paper is 2.

## 4. Predicting the Gas Outburst and the Results

4.1. *The Results of Identifying Prediction Parameters.* Process the data in Table 1 according to the method described in Section 3.1 of this paper. Set the “Risk” sequence as the reference sequence  $X_0$ . The remaining columns are used as comparison sequences, including coal destruction type  $X_1$ , gas release initial velocity  $X_2$ , coal firmness coefficient  $X_3$ , coal seam gas content  $X_4$ , gas desorption index-K1  $X_5$ , drilling cuttings volume  $X_6$ , geological structure zone distance  $X_7$ , coal seam depth  $X_8$ , and coal seam thickness  $X_9$ . Nondimensional data processing adopts average value method. The grey correlation degree of each comparison sequence and the reference sequence is obtained by calculation, and the grey correlation degree is sorted. The calculation results are shown in Table 2.

TABLE 2: Grey correlation degree between each comparison sequence and reference sequence.

Comparison sequence	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$	$X_9$
Grey correlation	0.51	0.72	0.43	0.81	0.68	0.65	0.61	0.58	0.52
Sort results	8	2	9	1	3	4	5	6	7

According to the calculation results in Tables 2 and 3, sets of experiments are set up to investigate the prediction performance of SVM under different indicator combinations. Take the comparison sequences with an association degree greater than 0.65, comparison sequences greater than 0.6, comparison sequences greater than 0.55, and all comparison sequences, respectively. The accuracy is used to compare the prediction results of the four groups of experiments. The grouping and prediction results are shown in Table 4.

It can be seen from the results in Table 3 that the first group has the highest accuracy. It shows that the number of parameters given in the first group is the smallest, but the predicted result is still the closest to the true value. The reduction of the input parameter dimension can reduce the complexity of the model, thereby reducing the calculation time and improving the efficiency of the entire system. Through the results of the grey correlation analysis experiment, the first set of comparison sequences is finally selected as the main input parameters of the model, which are coal seam gas content  $X_4$ , gas release initial velocity  $X_2$ , gas desorption index-K1  $X_5$ , and drilling cuttings volume  $X_6$ .

4.2. *The Results of the Gas Outburst Prediction.* Using APSO-SVM to predict coal and gas outburst, the prediction process is shown in Figure 2. In order to evaluate the predictive performance of the APSO-SVM model, for the same set of sample data, the SVM model, PSO-SVM model, and APSO-SVM model were used for experiments. There are 162 sets of sample data. The first 100 sets are used as training data, and the last 62 sets are used as test data. The parameters in the SVM, PSO, and APSO algorithms are initialized. For the SVM model, the grid search method is used to optimize the parameters, the minimum value of  $c$  and  $g$  is -8, the maximum value is 8, and the step is 0.8. For PSO-SVM model and APSO-SVM model, set the variation range of the penalty factor  $c$  to (0.1, 100) and the variation range of the kernel function parameter  $g$  to (0.01, 1000),  $c_1$  is 1.5,  $c_2$  is 1.7, the maximum evolutionary generation is 200, and the population size is 20.

Taking the model training results of the first 100 sets to analyze the parameter optimization effects of the three algorithm models, the calculation results are shown in Figure 3. It can be seen from Figure 3(a) that the results of the grid search method are  $c = 16$  and  $g = 0.0625$ , and the accuracy is 89%. Figure 3(b) shows that the PSO algorithm obtains the best parameters ( $c = 62.0514$  and  $g = 0.0265$ ). When the iteration reaches the 10th generation, the accuracy is 91%. Figure 3(c) shows that the APSO algorithm obtains a

TABLE 3: Detailed prediction results of the three models.

Method	Serious	Ordinary	Secure
SVM	4 (true-4, false-0)	28 (true-26, false-2)	30 (true-29, false-1)
PSO-SVM	5 (true-5, false-0)	27 (true-26, false-1)	30 (true-29, false-1)
APSO-SVM	6 (true-6, false-0)	28 (true-27, false-1)	28 (true-28, false-0)
Original	6	27	29

TABLE 4: Test grouping results.

Group	Grey correlation	Parameters combination	Accuracy
1	$r > 0.65$	$X_4, X_2, X_5, X_6$	89
2	$r > 0.60$	$X_4, X_2, X_5, X_6, X_7$	88
3	$r > 0.55$	$X_4, X_2, X_5, X_6, X_7, X_8$	85
4	All	$X_4, X_2, X_5, X_6, X_7, X_8, X_9, X_1, X_3$	79

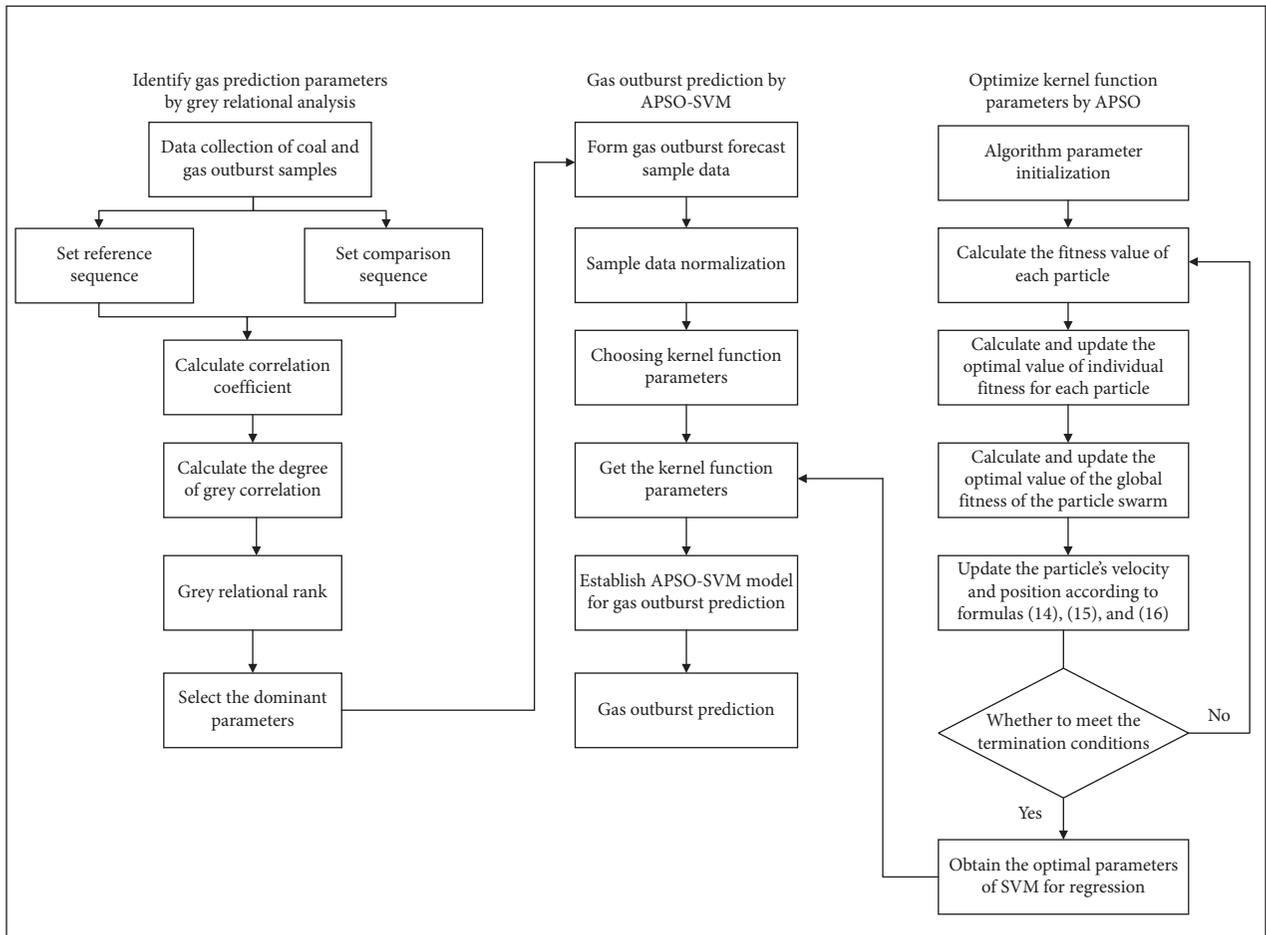


FIGURE 2: The construction of the gas outburst risk prediction based on grey relational analysis and using SVM optimized by APSO algorithm model.

local optimal solution when the iteration reaches the 10th generation, and the accuracy is 93%. When the average fitness gradually approaches 93%, the algorithm adaptively adjusts the update weights so that the particles increase the search range and jump out of local extreme points, and then the search continues. The model obtains a better solution than the first stage when the iteration reaches the 39th

generation, with an accuracy of 95%. Repeat the previous steps; when the average fitness is close to 95%, the algorithm adjusts the update weights again and the search range of the particle is increased. When iterating to the 79th generation, a better solution than that of the previous stage is obtained. The accuracy is 97%. Within the range of 200 iterations, the best accuracy is 97%. Therefore, the solution is regarded as

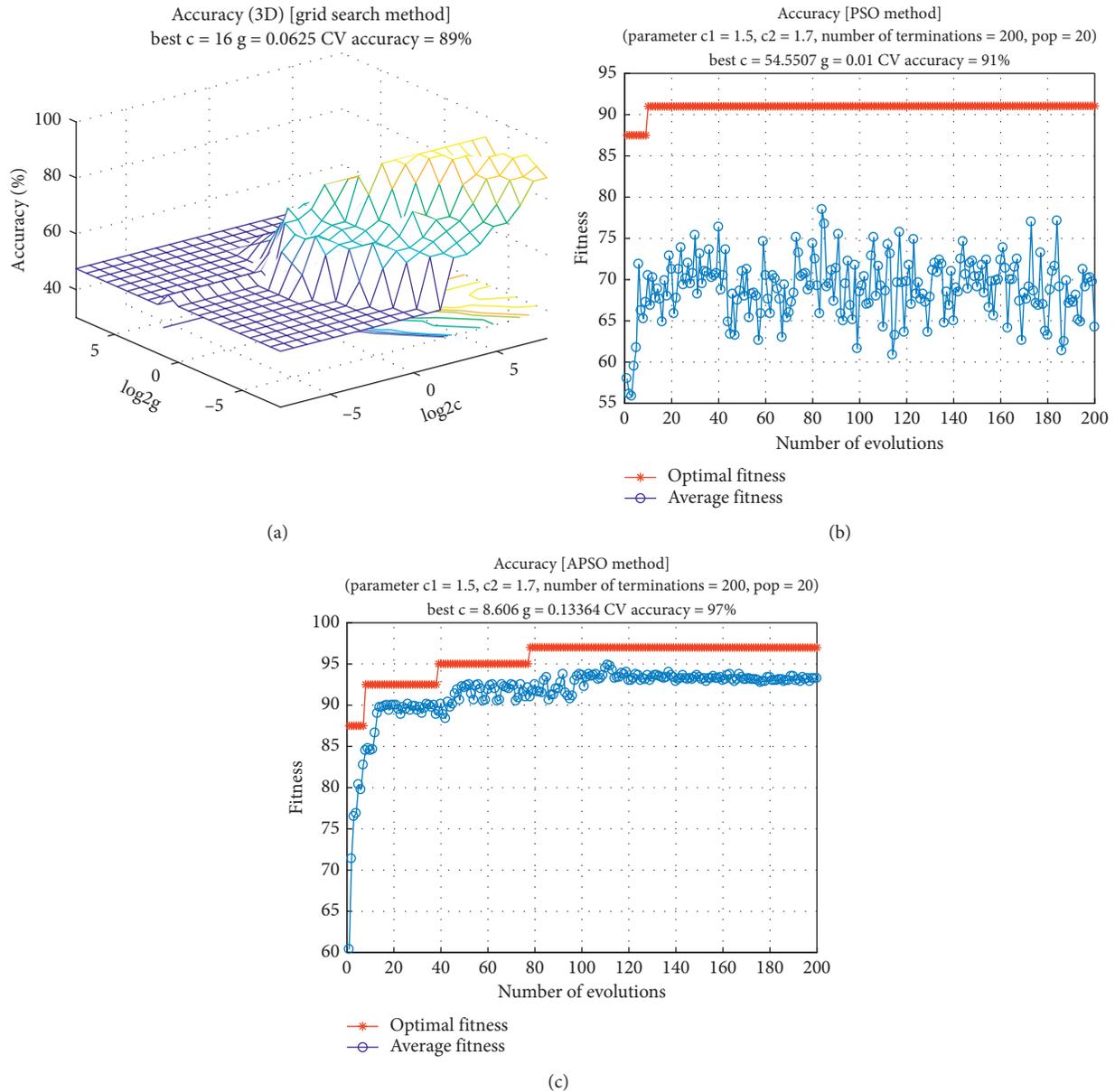


FIGURE 3: Parameter optimization effects of the three algorithm models. (a) Parameter calculation results by grid search method. (b) Parameter calculation results by PSO method. (c) Parameter calculation results by APSO method.

the global optimal solution. The corresponding optimal parameters are  $c = 8.6060$  and  $g = 0.1336$ . According to the above analysis and calculation results, the accuracy of the APSO optimization algorithm is 97%, and the effect is the best. The second is the PSO algorithm, which has an accuracy rate of 91%, and the grid search method is relatively inefficient, with an accuracy rate of 89%.

Figure 4 shows the prediction results of the 62 test sets using 3 algorithm models. Table 3 lists the number of samples predicted by the three algorithm models with serious risk, general risk, no outburst risks, and the misjudgments of each level.

It can be seen from Table 3 that the number of samples with serious risk predicted by the SVM model is 4. Compared with the real sample data, these 4 samples are indeed

serious risk with no misjudgment, but the number of real samples with serious risk is 6, so 2 samples are missed. It can be seen from Figure 4(a) that these two samples with serious risk were misjudged as general risk. The number of samples with general risk predicted by the SVM model is 28. Two of them are misjudgments, so 26 are correct predictions. The number of samples with no outburst risk predicted by the SVM model is 30. There are 29 true nonprotruding risk samples, so 29 of the prediction samples are correct predictions and 1 is wrong prediction.

In actual engineering applications, the serious risk of gas outburst belongs to a small number of samples, so we classify them as positive samples. General risk and no outburst risk of gas outburst are classified as negative samples. Based on the above analysis, it can be calculated that the accuracy of

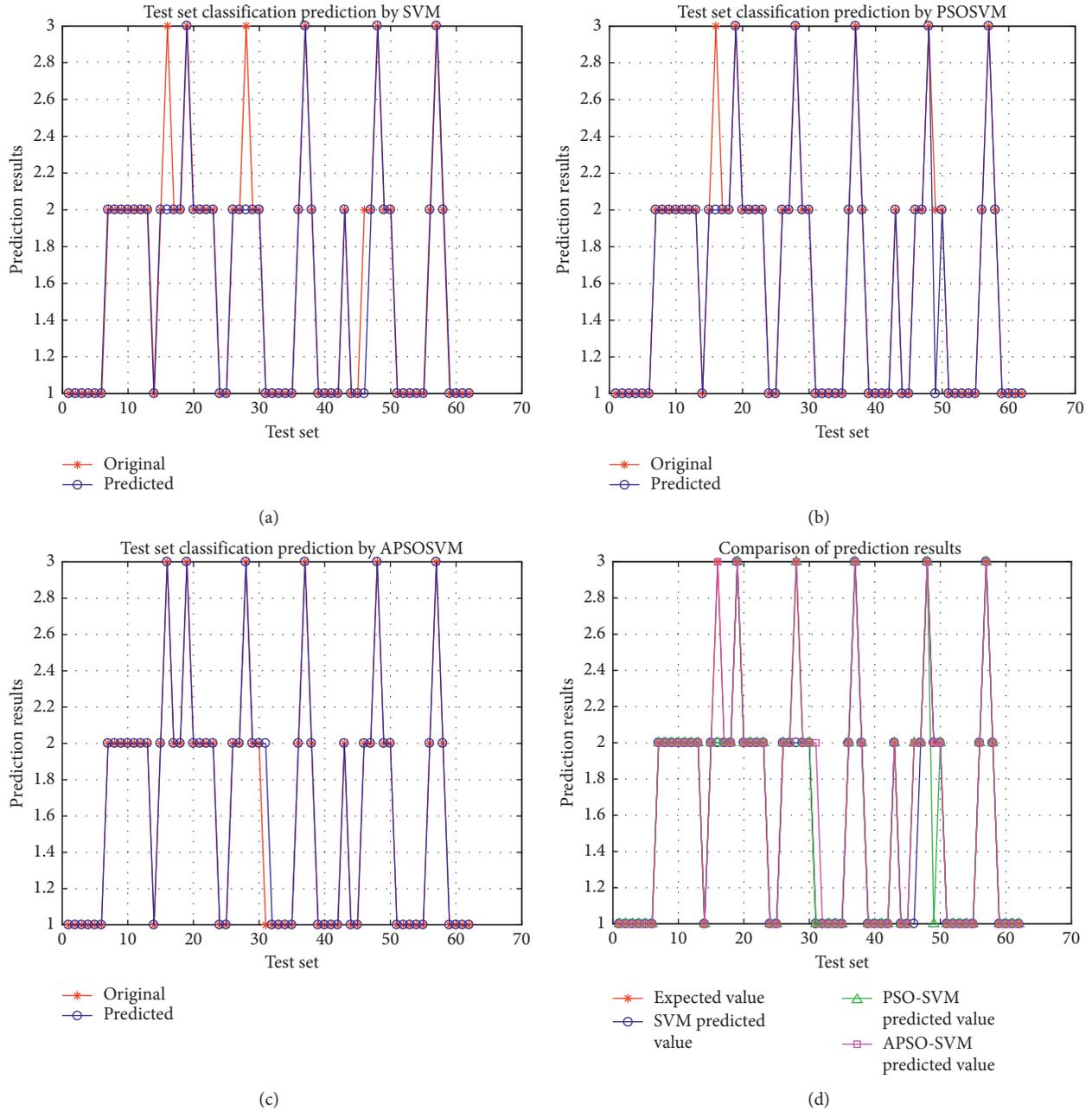


FIGURE 4: Test data set prediction results of three models. (a) Prediction results by SVM model. (b) Prediction results by PSO-SVM model. (c) Prediction results by APSO-SVM model. (d) Comparison of prediction results.

the SVM model is 95.16%, precision is 100%, recall is 66.67%, and  $F_2$ -score is 0.7443. Similarly, according to the above analysis methods, we can calculate the various indicators of the PSO-SVM model and the APSO-SVM model. The specific results are shown in Table 5. From the perspective of accuracy, the APSO-SVM model has the best effect, where its accuracy is 98.38%, followed by the PSO-SVM model with an accuracy of 96.77%, and the SVM model performs relatively poorly. The precision of the three models is 100%, which means that the three models will not misjudge negative samples as positive samples. In actual engineering applications, the three models will not have false

TABLE 5: Index comparison results of the three models.

Method	Accuracy (%)	Precision (%)	Recall (%)	$F_2$ -score
SVM	95.16	100	66.67	0.7443
PSO-SVM	96.77	100	83.33	0.8620
APSO-SVM	98.38	100	100	1.0000

alarms. From the perspective of recall, the APSO-SVM model has the best effect, where its recall is 100%, followed by PSO-SVM with a recall of 83.33%, and the SVM model performs the worst. The results above show that APSO-SVM

is sensitive enough to the serious risk samples, which is very important to ensure safe production. In practical engineering applications, we need to ensure that all the samples with serious risk can be identified, so the importance of recall is emphasized. Here,  $F_2$ -score is calculated. From the calculation results, the APSO-SVM model's score is 1, followed by PSO-SVM with a score of 0.8620, and SVM has the worst realization with a score of 0.7443.

Based on the analysis above, the APSO-SVM model performs the best.

## 5. Conclusions

This paper proposed an algorithm based on grey relational analysis and SVM using adaptive particle swarm optimization (APSO-SVM) for gas outburst prediction. The main conclusions are as follows:

- (1) The parameters of the APSO-SVM model are obtained by the grey correlation analysis method, which avoids the subjectivity of weight determination. 5 parameters are eliminated from the original data by using grey relational analysis, and experiments show that the accuracy of the model is improved after eliminating the parameters with poor correlation. With the continuous increase of sample data in practical engineering applications, too many prediction parameters will bring about great difficulty to model training, so this method is meaningful.
- (2) The APSO algorithm is used to optimize the penalty factor and kernel function parameters of the SVM model, which improves the global search ability of the entire system. Accuracy, precision, recall, and  $F_2$ -score are introduced to analyze the performance of the model. Compared with the SVM model optimized by the grid search method and the PSO method, the APSO-SVM model established in this paper has higher accuracy in the application of gas outburst prediction.
- (3) The APSO-SVM model established in this paper is applied in tunneling heading face 31004 of Xinyuan Coal Mine in Yangquan City, Shanxi Province, China. It should be mentioned that although the proposed prediction model has been proven to be more accurate, gas outburst is a complex nonlinear dynamic process affected by many factors. There are still some important issues to be studied, such as the mechanism of gas outburst and the application of some new development methods. In addition, it is of great significance to apply the APSO-SVM model established in this paper to other disaster predictions in coal mines, such as mine pressure prediction. Therefore, in our future study, we will expand the data range to further test the ability of our model in the accurate and efficient prediction of gas outburst and other disasters.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This research was supported by the Funds for Central Universities (no. 2017CXNL01).

## References

- [1] V. Hudeček, M. Vanek, and I. Černý, "Economic assessment of methods of coal and gas outburst prevention in the extraction of coalfaces," *New Trends in Production Engineering*, vol. 3, pp. 379–393, 2020.
- [2] X. Chen, L. Li, L. Wang, and L. Qi, "The current situation and prevention and control countermeasures for typical dynamic disasters in kilometer-deep mines in China," *Safety Science*, vol. 115, pp. 229–236, 2019.
- [3] K. P. Chen, "A new mechanistic model for prediction of instantaneous coal outbursts - dedicated to the memory of Prof. Daniel D. Joseph," *International Journal of Coal Geology*, vol. 87, no. 2, pp. 72–79, 2011.
- [4] L. I. A. N. G. Yue-qiang, X. I. E. Xue-cai, X. U. De-yu, and L. I. Lei-lei, "Research advance of coal and gas outburst prediction and geological factors," *Coal Technology*, vol. 36, pp. 168–170, 2017.
- [5] S. Li, "Study on gas geological unit division of coal and gas outburst coal seam," *Coal Science & Technology Magazine*, vol. 41, pp. 43–46, 2020.
- [6] L. Zhang, "Comprehensive analysis of acoustic emission evolution characteristics during deformation and failure process of outburst coal," *Journal of China Coal Society*, vol. 43, no. S1, pp. 130–139, 2018.
- [7] Y. Wang, L. Liu, H. Fu, and Y. Xu, "Research on acoustic emission multi-parameter time series based prediction of gas outburst," *China Safety Science Journal*, vol. 28, no. 05, pp. 129–134, 2018.
- [8] Y. Li, Y. Yang, and B. Jiang, "Prediction of coal and gas outbursts by a novel model based on multisource information fusion," *Energy Exploration & Exploitation*, vol. 38, no. 5, pp. 1320–1348, 2020.
- [9] C. Dong, Y. Shang, and X. Liu, "Coal and gas outburst classification analysis based on G-K evaluation and rough set," *Journal of China Coal Society*, vol. 36, no. 07, pp. 1156–1160, 2011.
- [10] C. Xia and X. Chen, "Application of grey correction analysis in simulating test of coal and gas outburst," *Coal Technology*, vol. 37, no. 10, pp. 179–181, 2018.
- [11] M. Xu, S. Gao, and Y. Cao, "Forecast of gas content based on gray theory and multiple regression analysis," *Safety in Coal Mines*, vol. 49, no. 9, pp. 211–214, 2018.
- [12] X. Zhou and J. Sun, "Prediction of gas emission based on principal factor Analysis and improved BP neural network," *Mining Safety & Environmental Protection*, vol. 45, no. 6, pp. 43–47, 2018.
- [13] Y. Liu, Q. Zhao, and H. Wenli, "Study of gas concentration prediction based on genetic algorithm and optimizing BP neural network," *Mining Safety & Environmental Protection*, vol. 42, no. 02, pp. 56–60, 2015.
- [14] X. Yang, Q. Yu, L. He, and T. Guo, "The one-against-all partition based binary tree support vector machine algorithms

- for multi-class classification,” *Neurocomputing*, vol. 113, no. 3, pp. 1–7, 2013.
- [15] A. Widodo, Bo-S. Yang, and T. Han, “Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors,” *Expert Systems With Applications*, vol. 32, no. 2, pp. 299–312, 2005.
- [16] J. Zeng, G.-Z. Cao, Y.-P. Peng, and S.-D. Huang, “A weld joint type identification method for visual sensor based on image features and SVM,” *Sensors*, vol. 20, no. 2, p. 471, 2020.
- [17] M. Wang and H. Chen, “Chaotic multi-swarm whale optimizer boosted support vector machine for medical diagnosis,” *Applied Soft Computing*, vol. 88, 2020.
- [18] F. Qu, W. An, Y. Li, and Z. Li, “Research on prediction model for coal and gas outburst based on GRA-SVM,” *China Coal*, vol. 38, no. 11, pp. 102–106, 2012.
- [19] M. Qiao, X. Ma, J. Lan, and Y. Wang, “Time-series gas prediction model using LS-SVR within a Bayesian framework,” *Mining Science and Technology (China)*, vol. 21, no. 1, pp. 153–157, 2011.
- [20] M. Carl, “The Top Ten Algorithms in Data Mining edited by Xindong Wu Vipin Kumar,” *International Statistical Review*, vol. 78, no. 1, 2010.
- [21] N. Xie and S. Liu, “Research on evaluations of several grey relational models adapt to grey relational axioms,” *Journal of Systems Engineering and Electronics*, vol. 20, no. 2, pp. 304–309, 2009.
- [22] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [23] L. Cui, H. Chen, J. Liu, and Y. Zhang, “Method and application of swarm intelligence optimization algorithm for support vector machine optimization,” *Electronic Measurement Technology*, vol. 42, no. 7, pp. 44–48, 2019.
- [24] Y. Quan and F. Qian, “A hybrid genetic algorithm for twice continuously differentiable NLP problems,” *Computers and Chemical Engineering*, vol. 34, no. 1, pp. 36–41, 2010.
- [25] Y. Dong, Li Zhao, and F. Zong, “Quantum behaved particle swarm optimization algorithm based on artificial fish swarm,” *Mathematical Problems in Engineering*, vol. 2014, Article ID 592682, 10 pages, 2014.
- [26] G. Dong, X. Fu, H. Li, and P. Xie, “Cooperative ant colony-genetic algorithm based on spark,” *Computers and Electrical Engineering*, vol. 60, pp. 66–75, 2017.
- [27] U. Manimegalai-Sridhar, A. Govindarajan, and R. Russell Rhinehart, “Improved initialization of players in leapfrogging optimization,” *Computers & Chemical Engineering*, vol. 60, no. 10, pp. 426–429, 2014.