

# Application of coupled FICA-FCM feature extraction and QPSO-SVM classifier model to predict coal and gas outbursts

**Xuning Liu**

China University of Mining and Technology - Beijing Campus

**Zixian Zhang**

Liaocheng University

**Zhixiang Li**

Shijiazhuang University

**Guoying Zhang** (✉ [zhanglaoshijidian@163.com](mailto:zhanglaoshijidian@163.com))

China University of Mining and Technology - Beijing Campus

---

## Research

**Keywords:** Coal and gas outbursts, Fast independent component analysis, Fuzzy clustering mean, Quantum particle swarm optimization, SVM

**Posted Date:** November 17th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-108416/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Application of coupled FICA-FCM feature extraction and QPSO-SVM classifier model to predict coal and gas outbursts

Xuning Liu<sup>1,2</sup>, Zixian Zhang<sup>1,3</sup>, Zhixiang Li<sup>2</sup>, Guoying Zhang<sup>1,\*</sup>

<sup>1</sup> School of Mechanical Electronic & Information Engineering, China University of Mining and Technology, Beijing, China

<sup>2</sup> Department of Computer Engineering, Shijiazhuang University, Shijiazhuang, China

<sup>3</sup> School of Foreign University, Liaocheng University, Liaocheng, China

Corresponding author: Guoying Zhang (zhanglaoshijidian@163.com).

**Abstract:** The coal and gas outbursts samples data are affected by all kinds of influencing factors, the accuracy of classification on these sample data is not high, the classification of some samples always have errors, which may be inaccurate annotation. In order to reduce the impact of noise data caused by the labeling errors on classification, this paper proposes a combination of classifier and clustering analysis model, which is used to improve the prediction accuracy of coal and gas outbursts. First, the high-order statistical characteristics of fast independent component analysis(FICA) method are used to extract the features of coal and gas outbursts sample data to obtain the independent nonlinear main metadata; Then, the fuzzy clustering mean(FCM) algorithm is used to cluster the coal and gas outbursts samples. Finally, based on improved clustering analysis results, the support vector machine(SVM) classifier based on quantum particle swarm optimization(QPSO) algorithm is used to compare the classification results with the existing labeled results, and the classification results are corrected to improve the classification accuracy. The results show that the performance of coupling model is significantly better than other existing classifiers for coal and gas outbursts prediction.

**Keywords:** Coal and gas outbursts, Fast independent component analysis, Fuzzy clustering mean, Quantum particle swarm optimization, SVM

## 1.Introduction

Coal and gas outbursts is a dynamic process with complex mechanism and many influencing factors, its mechanism and evolution law are not clear so far. Outburst disaster cannot be effectively pre-warning and prevention, which seriously threatens the safety production of coal mines in China. Therefore, rapid and accurate prediction of coal and gas outburst has important practical significance for coal mine safety production[1]. Due to the uncertainty and randomness of coal and gas outbursts, there are many complicated nonlinear relations between outbursts and each index. In the process of coal and gas outbursts prediction, it is very important to extract the complex relationship between coal and gas outbursts indexes, which greatly affects the design and performance of classifier, and is related to the efficiency and accuracy of the whole prediction model [2]. However, it is very difficult to obtain the sample data of coal and gas outburst, and there are few data samples and the labeling results may have problems. The effect of the existing classification model to classify coal and gas outbursts are not very good, the sample data with wrong labels is easy to generate noise data, which affects the classification effect. Therefore, it is necessary to take various methods to find out the wrong labeled samples to correct and improve the prediction accuracy.

In order to solve the above-mentioned key problems, many scholars at home and abroad have carried out research on coal and gas outbursts methods in recent years. The research results mainly focus on the clustering analysis and classification of coal and gas outbursts. At present, the clustering analysis of coal and gas outbursts methods is unsupervised learning. It is a mathematical method for classification of unknown types of data, and its main task is to classify the data according to the similarity or dependence of the characteristics of the patterns to be classified, so that the data of the same class are as similar as possible, and the data of different types are as different as possible. Clustering analysis can intelligently reveal the internal relations and differences between data, and discover the internal structure and law. In practice, the clustering analysis can be used as the preprocessing process of other mining algorithms to analyze the data after clustering, it can be used as a precursor process of other learning tasks such as classification to provide a basis for further data analysis, and it can also be regarded as an independent mining algorithm. By looking for the internal distribution structure of data samples, it reveals the inherent

properties and laws of data samples, so as to study the distribution of data. Although unsupervised learning pattern recognition method cannot give clear classification information as supervised learning method, it can adaptively mine the distribution characteristics of a large number of data according to a certain similarity algorithm, reflecting the similarities and differences between data. The pattern recognition method based on unsupervised learning is mainly used to process unlabeled data, that is, data without clear category information. Therefore, the classification process of unsupervised learning has no supervision and guidance. Its classification principle is mainly to classify with the help of some algorithm and similar principles between data. Therefore, the classification result of data has no specific category information, the classification results can reflect the internal relationship, distribution structure or potential classification rules of these data. The direct distance between samples is used to describe the similarity between samples, and the Euclidean distance between sample vectors is used to judge the classification. The near distance is divided into one class, and the far distance is divided into different classes. A large data set is divided into several subsets or class clusters. Similar target data form a class cluster, and the data objects in different class clusters are different. The typical clustering algorithms in coal and gas outbursts are fuzzy clustering [3], k-means[4], dcaainet[5], projection pursuit clustering [6] and self-organizing mapping neural network[7]. The researchers have applied them to different research fields on this basis combined with other methods. Clustering analysis model can find the rule of sample data characteristics, but the existing clustering analysis model has low computational efficiency, high complexity and low accuracy. The FCM algorithm can mine and analyze the structural information and inherent regularity of the sample data set, and realize the correct classification of samples according to the sample similarity and probability density function estimation method. However, most of the current clustering algorithms generally have the following problems when dealing with practical problems: it is necessary to know the data distribution and the number of clustering categories in advance; the solution results are easy to fall into local minimum; and there is no way to deal with the data sets with unclear class boundaries. The clustering measure does not accurately describe the similarity between the data sample and the clustering prototype, which leads to the low feasibility and effectiveness of the algorithm. Therefore, it is necessary to explore new mining theory and model, and improve some traditional algorithms, or study how to combine the existing algorithms to solve the problem of coal and gas outbursts. The FCM can mine and analyze the structural information and inherent regularity of the sample data set, and realize the correct classification of samples according to the sample similarity and probability density function estimation method. Clustering algorithm can automatically classify the data according to the degree of similarity between data, so that the similarity between classes is as small as possible, while the similarity within classes is as large as possible. FCM is the most common clustering algorithm. At present, most of the clustering algorithms have the following problems when dealing with problems: we need to know the data points in advance. it is necessary to know the data distribution and the number of cluster categories in advance; the solution result is easy to fall into local minimum; there is no ability to do anything for the data set with unclear class boundary, and the clustering measure does not accurately describe the similarity between the data sample and the clustering prototype, which leads to the low feasibility and effectiveness of the algorithm. Therefore, it is necessary to explore new mining theory and model, to improve some traditional algorithms, or to study how to combine the existing algorithms to solve the problem of coal and gas outbursts

Feature extraction has a very important impact on the clustering algorithm design and data processing process, good quality feature can simplify the algorithm complexity, improve the efficiency of the algorithm, and can obtain the ideal clustering effect. Each index feature plays a different role in the clustering process, some of the index features play a leading role in the clustering process, some of the index features do not play a significant role in the clustering process, and some of them cause noise to bring negative impact on the generation of the cluster. It is necessary to select and extract features among clusters, and the task is to find out the most effective features from the numerous features of the data set, and map the data points from the high-dimensional feature space to the low-dimensional feature space, so as to make it easier to distinguish the clusters contained in the dataset. The effective features should be able to highlight the similarities between similar data and the differences between different types of data, and have strong immunity to noise. Reasonable feature extraction can simplify the complexity of clustering algorithm design and improve the efficiency of the algorithm, and can obtain the ideal clustering effect. It can not only save the calculation time, but also can get a more compact and easy learning model. In order to realize the rapid, accurate and dynamic prediction of coal and gas outburst, various influencing factors of coal and gas outbursts are considered, feature extraction methods include principal component analysis, independent component analysis, genetic

algorithm optimization projection pursuit and artificial fish swarm algorithm optimization clustering algorithm [8-10]. These feature extraction methods have their own advantages and disadvantages, improve the accuracy of coal and gas outbursts prediction. However, due to the complexity and diversity of coal and gas outbursts process, there are many uncertain factors in the outbursts characteristics. The existing coal and gas outburst feature extraction methods change the physical distribution of the original characteristics and cannot effectively judge the original, the main influencing factors in the initial features, and it is difficult to mine the complex relationship between the original features, which leads to the low classification accuracy. Due to the particularity of the application field, most of the current clustering algorithms generally have the following problems when dealing with practical problems: it is necessary to know the data distribution and the number of clustering categories in advance; The solution results are easy to fall into local minimum; there is no way to deal with the data sets with unclear class boundaries, the clustering measure does not accurately describe the similarity between the data sample and the clustering prototype, which leads to the low feasibility and effectiveness of the algorithm. Therefore, it is necessary to explore new mining theory and model, and to improve some traditional algorithms, or to study how to combine the existing algorithms to solve the problem of coal and gas outbursts. The measured data and theoretical study show that coal and gas outbursts is affected by many factors, among these factors, some are definite and quantitative, some are fuzzy and qualitative. There is a certain nonlinear correlation between different factors, and the evaluation process is accompanied by strong fuzziness. The measured data and theoretical research show that the data collected in the process of coal and gas outburst often contain a large amount of redundant information, and the variables are usually correlated or collinear. Using these factors directly, a large number of redundant correlation data will affect the prediction accuracy of the model. Therefore, reasonable selection of variables can improve the predictability of the model It's important to be able to. First independent component analysis (FICA) feature extraction can effectively extract non-Gaussian and nonlinear independent components of process data, It reduces the complexity of the detection model; the FCM algorithm can mine and analyze the structural information and internal regularity of the sample data set, realizes the correct classification of the samples according to the sample similarity and probability density function estimation method, and explores the combination of FICA method and FCM method to improve the clustering effect and efficiency of coal and gas outburst. the FCM algorithm can mine and analyze the structural information and intrinsic rules of the sample data set According to the sample similarity and probability density function estimation method, the correct classification of samples is realized. The combination of equidistant feature mapping method and fuzzy mean clustering method is explored to improve the clustering effect and efficiency of coal and gas outburst.

The common classification models of coal and gas outbursts are as follows: SVM and its variant[15], neural network and its variant [12], random forest[13] and decision tree[14]. Random forest is a combination of forecasting methods, is also for non-linear, small sample forecasting methods, but there are low prediction accuracy and generalization ability, and there is no relatively perfect theory to support, for small sample data has a low prediction accuracy. Due to the limitations of BP neural network, such as easy to fall into local optimum, slow convergence rate and large demand for training samples, the prediction results of coal and gas outburst with BP neural network have large error and are not ideal; SVM has certain advantages in solving nonlinear and small samples, which is very suitable for coal and gas outbursts prediction, However, SVM has the problem that the penalty factor is difficult to determine. The experimental results are very sensitive to the data, and improper setting will cause over learning and other problems. The parameters of kernel function of SVM is the key factor affecting the performance of classifier. The proper combination of parameters is very important to the classification performance of classifier. It is very important to obtain the optimal combination of kernel function parameters in this paper. Many scholars have carried out in-depth research on the optimization of kernel parameters of support vector machine. For example, reference [27] uses grid search algorithm to select parameters of support vector machine, but the accuracy of the model obtained by the algorithm is limited. Literature [28] uses genetic algorithm to optimize the parameters of support vector machine, which reduces the time required for training the model, and is easy to fall into local optimization The model parameters of SVM are optimized, but the disadvantages are long time and unstable; The defect of these optimization algorithms is that they are easy to fall into local optimum, the parameters are not optimal, and the efficiency is not high, so we need to choose a better optimization algorithm to optimize the parameters. QPSO algorithm is a global search algorithm, which theoretically guarantees to find a good optimal solution in the search space. Compared with particle swarm optimization, the iterative equation of quantum particle swarm optimization algorithm does not need particle velocity vector, and needs less parameters

to be adjusted, so it can be realized more easily. The experimental results on widely used benchmark functions show that the quantum particle swarm optimization algorithm has better performance than the standard particle swarm optimization algorithm [25-26]. For the parameter optimization problem of the classifier parameters, the quantum particle swarm optimization algorithm [40] will be used to optimize the basic parameters (penalty factor, kernel parameter) and the adjustable parameters of the SVM to obtain the optimal combination of parameters. Through function fitting and optimization of related parameters in coal and gas outbursts prediction, its effectiveness is verified, and the recognition effect of coal and gas outbursts is improved greatly. Aiming at the problems of kernel function selection and parameter intelligent optimization algorithm of SVM, another important research content of this paper is to design effective kernel parameters of SVM based on effective feature extraction. Aiming at the key problems in coal and gas outbursts prediction, a comprehensive model based on feature extraction clustering model and intelligent optimization classification model is proposed in this paper.

In view of the key problems in the classification of coal and gas outbursts, this paper proposes a method which employs a couple of FCM clustering analysis model based on FICA feature extraction and the SVM classification model based on QPSO algorithm for the classification of coal and gas outbursts. The contributions of this paper are as follows: First, FICA can extract the nonlinear components of coal and gas outbursts influencing factors, obtain effective features for FCM clustering analysis, according to the sample similarity and probability density function estimation, FCM performs the clustering on the sample data obtained from FICA, the FICA-FCM model can achieve high accuracy and efficiency; Second, the QPSO-SVM model trained by the results that FICA-FCM obtained is proposed, in QPSO-SVM, the mixed kernel function is presented, and Bayesian optimization algorithm is used to optimize the kernel parameters and weight coefficients, which can improve the accuracy and generality ability; Finally, based on the above FICA-FCM, the classification results are compared on the SVM classifier, and the classification accuracy is used to find the wrong labeled samples, and the wrong labeled samples are correctly labeled in order to reduce the influence of noise data caused by labeling errors on the classification accuracy of SVM classifier, which contributes to the improvement of classification accuracy and efficiency of coal and gas outbursts. The remaining part of the article is organized as follows. Section2 offers the materials and methodologies used in the experiments and introduces our proposed approach. The dataset preparation and configuration environment of the proposed schemes are presented, then the experimental results are analyzed in Section3. Conclusions based on this research and future research targets are highlighted in Section4.

## 2 Proposed methods

### 2.1 FICA-FCM

#### 2.1.1 ICA

ICA[15-16] model assumes that the observed mixed signal variable  $y = (y_1, y_2, \dots)$ . These variables are composed of  $n$  independent unknown source signals  $s = (s_1, s_2, \dots, s_n)$  is a linear mixture, so there is a relationship between the mixed signal  $y$  and the source signal  $s$ , which can be expressed as  $y = ws$  using vectors, and  $w$  is as the unknown mixed coefficient matrix.

$$y = ws + n \quad (1)$$

Where  $n$  is the unknown unpredictable noise, which accords with the Gaussian distribution. ICA recovers the source signal and mixture matrix through  $y$ . The main idea is to remove the correlation between the data variables from the non-Gaussian signal of the original data, so that the components of each variable are statistically independent of each other. ICA has a better separation effect on the cross signal, and can more deeply mine the effective information which is covered due to the cross signal. Superior performance in signal separation, redundancy elimination and noise reduction to realize decomposition and fusion of deformation information. In this paper, formula (5) is analyzed by using FastICA algorithm which is more efficient than traditional ICA.

#### 2.1.2 FICA

Because the FICA[17-19] algorithm takes the maximum negative entropy as a search direction, we first discuss the decision criterion of negative entropy. According to the theory of information theory, the entropy of gauss variable is the largest among all equal variance random variables, so we can use entropy to measure non-gauss, which is often used as the correction form of entropy, namely negative entropy. According to the central limit theorem, if a random variable is composed of the sum of many independent random variables, as long as it has a limited mean and variance, then regardless of its distribution, the random variables are closer to Gaussian distribution. In other words, it is stronger than non-gaussian. Therefore, in the separation process, we can express the mutual independence of the separation results through the non-Gaussian measurement of the separation results. When the non-gaussian measurement reaches the maximum, it indicates that the separation of the independent components has been completed. Definition of negative entropy is as follows:

$$\mathbf{H}(\mathbf{Y}) = - \int \mathbf{p}(\mathbf{y}) \log \mathbf{p}(\mathbf{y}) d\mathbf{y} \quad (2)$$

Where  $\mathbf{H}(\mathbf{Y})$  is a Gaussian random variable with the same variance, and is the differential entropy of the random variable

$$\mathbf{J}(\mathbf{Y}) = \mathbf{H}(\mathbf{Y}_{gauss}) - \mathbf{H}(\mathbf{Y}) \quad (3)$$

According to information theory, among the random variables with the same variance, the random variables with Gaussian distribution have the maximum differential entropy. When there is a Gaussian distribution, the stronger the non-Gaussian property of; is, the smaller the differential entropy is and the larger the value is, so it can be used as a

non-Gaussian measure of random variables. Since it is impractical to calculate the probability density distribution function needed to know for differential entropy according to equation (7), the following approximate formula is adopted:

$$\mathbf{J}(\mathbf{W}) \approx [\mathbf{E}\{\mathbf{G}(\mathbf{W}^T \mathbf{X})\} - \mathbf{E}\{\mathbf{G}(\mathbf{V})\}]^2 \quad (4)$$

Where  $\mathbf{G}(\cdot)$  is some non-quadratic function and  $\mathbf{V}$  is a Gaussian variable of standard normal distribution. In this way, the problem of finding the maximum value of  $\mathbf{J}(\mathbf{W})$  is transformed into the problem of solving the maximum value of  $\mathbf{E}\{\mathbf{G}(\mathbf{W}^T \mathbf{X})\}$ , that is, when solving the value of mixed matrix  $\mathbf{W}$ , each independent component has the strongest non Gaussian property and the best separation effect. The iteration formula (9) is used for calculation

$$\mathbf{w}(k+1) = \mathbf{E}\{\mathbf{Xg}(\mathbf{w}(k)^T \mathbf{X})\} - \mathbf{E}\{\mathbf{Xg}(\mathbf{w}(k)^T \mathbf{X})\} \mathbf{w}(k)^T \quad (5)$$

Where  $\mathbf{g}(\cdot)$  is the derivative of function  $\mathbf{G}(\cdot)$ . At the same time,  $\mathbf{w}(K+1)$  was normalized by (10)

$$\mathbf{w}(k+1) = \mathbf{w}(k+1) / \|\mathbf{w}(k+1)\|_2 \quad (6)$$

Therefore, the specific iterative steps are as follows:

- Step 1: Select the initial value  $\mathbf{w}$ ;
- Step2: Calculate  $\mathbf{w}(k+1)$  according to (9);
- Step3: Normalize  $\mathbf{w}(k+1)$  according to (10);
- Step4: If the adjacent  $\mathbf{w}$  change is less than the given value, the iteration stops; otherwise, it turns to step2

### 2.1.3 FCM clustering analysis

Clustering analysis[20-22] is an unsupervised method that can decomposes data into subgroups or clusters based on the similarity between instances. It can also be divided into two categories: hard clustering and soft clustering. For hard clustering (such as k-means), an instance  $x_i$  belongs to one and only belongs to one cluster which is most similar to  $x_i$ . However, in the fuzzy clustering method, an example on the boundary between multiple clusters does not necessarily belong to one of them. They can belong to multiple clusters at the same time. Each instance has a membership between 0 and 1 that represents part

of its membership. FCM [32-34] is the soft clustering technology, it is a kind of clustering algorithm based on partition.

The idea is to make the similarity between objects divided into the same family the largest, while the similarity between different families is the smallest. Let the sample set be  $X = \{x_1, x_2, x_3, \dots, x_N\}$ , the clustering center is  $\{v_1, v_2, v_3, \dots, v_N\}$ . The membership degree of sample  $x_j$  ( $j = 1, 2, 3, n$ ) to clustering center  $v_i$  ( $j = 1, 2, 3, n$ ) is  $u_{ij}$

$$\min J(\mathbf{U}, \mathbf{V}) = \sum_{i=1}^c \sum_{j=1}^N (u_{ij}^m) \cdot \|x_j - v_i\|^2 \quad (7)$$

$$\text{St } \sum_{i=1}^c u_{ij} = 1 \quad 0 \leq u_{ij} \leq 1$$

Among them,  $M$  is a constant that can control the fuzzy degree of clustering results:  $\|x_j - v_i\|$  represents the Euclidean distance between the  $j$ th sample and the  $i$ th clustering center. Lagrange multiplier algorithm is used to solve the iterative formula of the model.

$$V_i = \frac{\sum_{j=1}^N u_{ij}^r x_j}{\sum_{j=1}^N u_{ij}^r} \quad 1 \leq i \leq c \quad (8)$$

$$u_{ij} = \left( \sum_{i=1}^c \frac{\|x_j - v_i\|^{2/(m-1)}}{\|x_j - v_i\|} \right)^{-1} \quad 1 \leq j \leq N, 1 \leq i \leq c \quad (9)$$

The FCM can complete the required precision for the membership matrix determining the number of iterations, and the precision is calculated using the membership matrix from one iteration to the next iteration, the process is as follows:

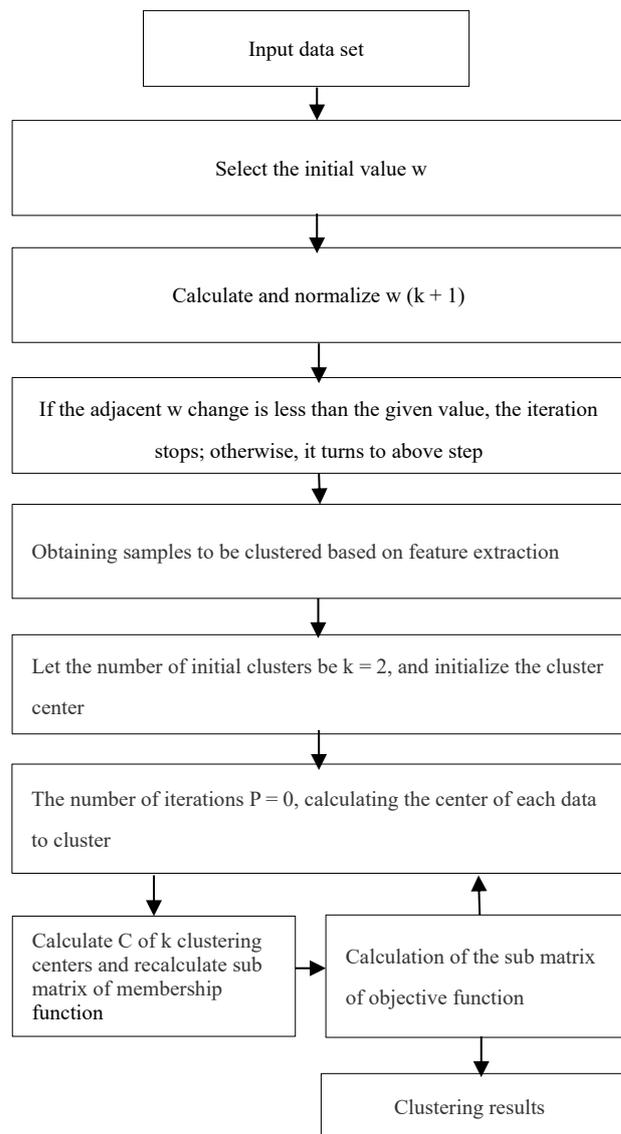
$$\|\delta^{r+1} - \delta^r\| \leq \varepsilon \quad (10)$$

Where,  $\delta^{r+1}$  and  $\delta^r$  are the partition matrix at iteration  $r$  and  $r+1$ , respectively.  $\|\cdot\|$  denotes the matrix norm operator, the iterations will be stopped when the difference between two successive partitions is less than a predefined level of accuracy  $\varepsilon$ .

## 2.1.4 FICA-FCM

FCM clustering is a flexible fuzzy partition scheme. Compared with the hard classification of K-means clustering, FCM belongs to soft classification algorithm. The mathematical theory of FCM clustering algorithm is complete, but the shortcomings of FCM are also prominent. It is very sensitive to noise and outliers, and has poor robustness to noise. The high-order statistics signal processing method adopted by FICA is to extract useful information from the high-order statistics of non-Gaussian signals, which can provide means for many signal processing problems that cannot be solved by the second-order statistics method. Find a nonlinear expression that makes the components statistically independent or as independent as possible from non-Gaussian signals. However, FICA is an unsupervised learning method, which cannot extract various correlation information between class labels and features, which reduces the accuracy and sensitivity of feature extraction [23-25]

Each index feature of coal and gas outbursts influencing factors plays a different role in the clustering process, some of the index features play a leading role in the clustering process and play a positive role in the formation of the cluster, some of the index features do not play a significant role in the clustering process, and some of them cause noise to bring negative impact on the generation of the cluster. It is necessary to select and extract features among clusters. Its task is to find out the most effective features from the numerous features of the data set, and map the data points from the high-dimensional feature space to the low-dimensional feature space, so as to make it easier to distinguish the clusters contained in the dataset. The FCM algorithm can mine and analyze the structural information and inherent regularity of the sample data set, and realize the correct classification of the samples according to the sample similarity and probability density function estimation method, FICA can effectively extract the non-Gaussian components of process data and reduce the complexity of the detection model. The flowchart of proposed model is as follows in Fig1:



**Fig.1.Flowchart of FICA -FCM model**

## 2.2 QPSO-SVM

### 2.2.1 SVM

SVM [26-30] is based on VC theory and structural risk minimization criterion in statistical learning theory, SVM algorithm combines many techniques and methods, such as kernel, sparse solution, loose variables, convex quadratic programming and maximum interval hyperplane, etc. it has advantages in solving problems of small sample, non-linear, high dimension, local minimum, global optimization and generalization performance. It can be divided into linear separable and non-linear SVM. For linear separable, it is to find the optimal classification line with the largest classification gap between the two categories. For non-linear separable problems, the non-linear transformation is realized by kernel function, so that the pattern classification problem in the original space is transformed into a pattern classification problem in the higher dimensional space [41]. According to the characteristics of the data, this paper mainly classifies the nonlinear SVM, and the learning process is the process of finding the optimal separating hyperplane of maximizing the segmentation of the two classes. The principle is as follows:

For a binary classification problem, given the training sample set  $(\mathbf{x}_i, \mathbf{y}_i)_{i=1}^m$ ,  $\mathbf{x}_i \in \mathbf{R}^n$  is the input information,  $\mathbf{y}_i \in \{\pm 1\}$  is the output target information,  $i$  is the number of training samples. To find the hyperplane of classification in new feature space, it is assumed that the separated hyperplane is linear, and the linear estimation function is defined as:

$$\mathbf{f}(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + \mathbf{b} = 0 \quad (11)$$

Where  $\phi(\mathbf{x})$  is a nonlinear mapping function from input space  $\mathbf{x}$  to high-dimensional feature space,  $\mathbf{w}$  is the weight vector,  $\mathbf{b}$  is the bias value, and the classification hyperplane should meet the following constraints:

$$\mathbf{y}_i \mathbf{f}(\mathbf{x}_i) = \mathbf{y}_i (\mathbf{w} * \mathbf{x}_i + \mathbf{b}) \geq 1, i = 1, 2, \dots, m \quad (12)$$

In order to measure the degree of  $\mathbf{w}$  and  $\mathbf{b}$  violating the constraint conditions, we need to introduce a non-negative relaxation variable  $\varepsilon_i$  for each sample  $\mathbf{x}_i$ .  $c \sum_{i=1}^m \varepsilon_i$  is the loss function,  $c$  is the penalty parameter.

$$\begin{aligned} \text{Minimize } & \frac{1}{2} \|\mathbf{w}\|^2 + c \sum_{i=1}^m \varepsilon_i \quad i = 1, 2, 3 \dots m \\ \text{subject to } & \left\{ \begin{array}{l} \mathbf{y}_i (\mathbf{w} * \mathbf{x}_i + \mathbf{b}) \geq 1 - \varepsilon_i \\ \varepsilon_i \geq 0 \end{array} \right\} \end{aligned} \quad (13)$$

By introducing Lagrange multiplication, the above optimization problems can be transformed into quadratic programming optimization problems, namely:

$$\begin{aligned} \text{Maximize } L(\alpha) &= \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{j=1}^m \alpha_j \mathbf{y}_j \mathbf{y}_i (x_i, x_j) \\ \text{Subject to } & \sum_{i=1}^m \alpha_i \mathbf{y}_i = 0, \alpha_i \geq 0, i = 1, 2, \dots, m \end{aligned} \quad (14)$$

In this way, the linear decision function can be defined by the following dual optimization problem:

$$\mathbf{f}(\mathbf{x}) = \text{sign} \left( \sum_{ij=1}^m \alpha_i \mathbf{y}_i (x_i, \mathbf{y}_i) + \mathbf{b} \right) \quad (15)$$

SVM with kernel function can be used in nonlinear classification problem. Nonlinear mapping function can map the original data to linear separable high-dimensional feature space. The nonlinear decision function is as follows:

$$\mathbf{f}(\mathbf{x}) = \text{sign} \left( \sum_{ij=1}^m \alpha_i \mathbf{y}_i k(x_i, \mathbf{y}_i) + \mathbf{b} \right) \quad (16)$$

In the formula,  $k(\mathbf{x}_i, \mathbf{y}_i)$  is the kernel function  $k(\mathbf{x}_i, \mathbf{y}_i) = \phi(\mathbf{x}_i) \phi(\mathbf{y}_i)$ , and the Gaussian kernel function is as follows:

$$k(x_i, y_i) = \exp(-g x_j - x_j^2), \rho > 0 \quad (17)$$

Where  $c, g$  are the parameters of SVM and kernel function, at present, there is no effective method to select kernel function and related parameters in the literatures. The setting of appropriate kernel parameters of SVM has an important impact on the prediction accuracy. Through the optimization algorithm, we can get the most suitable kernel parameters. Relevant research literatures have shown that Gaussian kernel function has significantly advantages in solving nonlinear problems due to its simple, effective and reliable computing ability, and it can realize the nonlinear mapping of input and output data. Therefore, in this paper Gaussian kernel function is generally selected as the kernel function of SVM.

### 2.2.2 QPSO

The main defect of PSO is that the global generalization ability cannot be guaranteed, and there are too many parameters to be set, which is not conducive to find the optimal parameters of the model to be optimized. The change of particle position

is lack of randomness and easy to fall into the trap of local optimum. In order to solve this problem, Sunetal proposed the quantum particle swarm optimization algorithm(QPSO)[30], is inspired by the particle swarm optimization algorithm trace analysis and quantum mechanics. The algorithm cancels the particle's moving direction attribute, and the update of particle's position has nothing to do with the particle's previous motion. In this way, the random particles of particle's position are added, and the particles are moved according to the following iterative formula,

$$mbest_j(t) = \frac{1}{m} pbest_{ij}(t) \quad (18)$$

Where pbest and ij are the best positions of all pbests in particle swarm, and mbest represents the average best position of the particle history of pbest. Where Xij represents the position of the ith particle and M represents the size of the particle swarm.

$$p_{ij}(t) = \phi_{ij} pbest_{ij}(t) + (1 - \phi_{ij}) gbest_j(t) \quad (19)$$

Pij (t) represents the update of the ith particle position. U,  $\phi_{ij}$  are the random value generated by the unified probability distribution in the range of [0,1],  $\alpha$  parameter is the contraction expansion factor, it is the only parameter controlling the particle generalization speed in the quantum particle swarm algorithm, and it is very sensitive to the population number and the maximum number of iterations.

$$X_{ij}(t + 1) = p_{ij}(t) \pm \alpha |mbest_j(t) - x_{ij}(t)| * \ln(1/u) \quad (20)$$

Here we set the value range of  $\alpha$  as (0.5,1) by searching, and the combination of particle swarm optimization and equation(21) is quantum particle swarm optimization.

Algorithm1	QPSO algorithm
	<p>The search space and fitness function f are defined, and the population number m and position <math>x_{ij}</math> are initialized and the dimension of particles</p> <p>For t=1-Maximum Iteration T do</p> <p style="padding-left: 2em;">Compute the mean best position mbest</p> <p style="padding-left: 2em;">B=(1.0-0.5).....(T-t)/T+0.5</p> <p style="padding-left: 2em;">For i=1---population size M do</p> <p style="padding-left: 4em;">If <math>f(X_i) &lt; f(P_i)</math> then</p> <p style="padding-left: 6em;"><math>P_i = X_i</math></p> <p style="padding-left: 4em;">END IF</p> <p style="padding-left: 2em;"><math>P_g = \min(P_{p_i})</math></p> <p style="padding-left: 2em;">For j=1-dimensionD do</p> <p style="padding-left: 4em;"><math>\phi = \text{rand}(0,1); u = \text{rand}(0,1);</math></p> <p style="padding-left: 4em;"><math>P_{ij} = \phi * P_{ij} + (1 - \phi) * P_{gj}</math></p> <p style="padding-left: 4em;">If <math>\text{rand}(0,1) &gt; 0.5</math> then</p> <p style="padding-left: 6em;"><math>X_{ij} = P_{ij} + B</math></p> <p style="padding-left: 4em;">Else</p> <p style="padding-left: 6em;"><math>X_{ij} = P_{ij} - B * (mbest - X_{ij}) * \log(1/u)</math></p> <p style="padding-left: 4em;">End if</p> <p style="padding-left: 2em;">End for</p> <p style="padding-left: 2em;">End for</p> <p>End for</p>

### 2.2.3 QPSO-SVM

At present, the kernel parameter optimization algorithm of SVM is easy to fall into local optimum, and the parameter value may not be the optimal solution, and the operation efficiency is not high. We use the kernel parameter intelligent optimization algorithm based on QPSO and k-fold cross validation to obtain the optimal kernel parameter value of SVM. The specific flowchart of QPSO-SVM is as follows in Fig2:

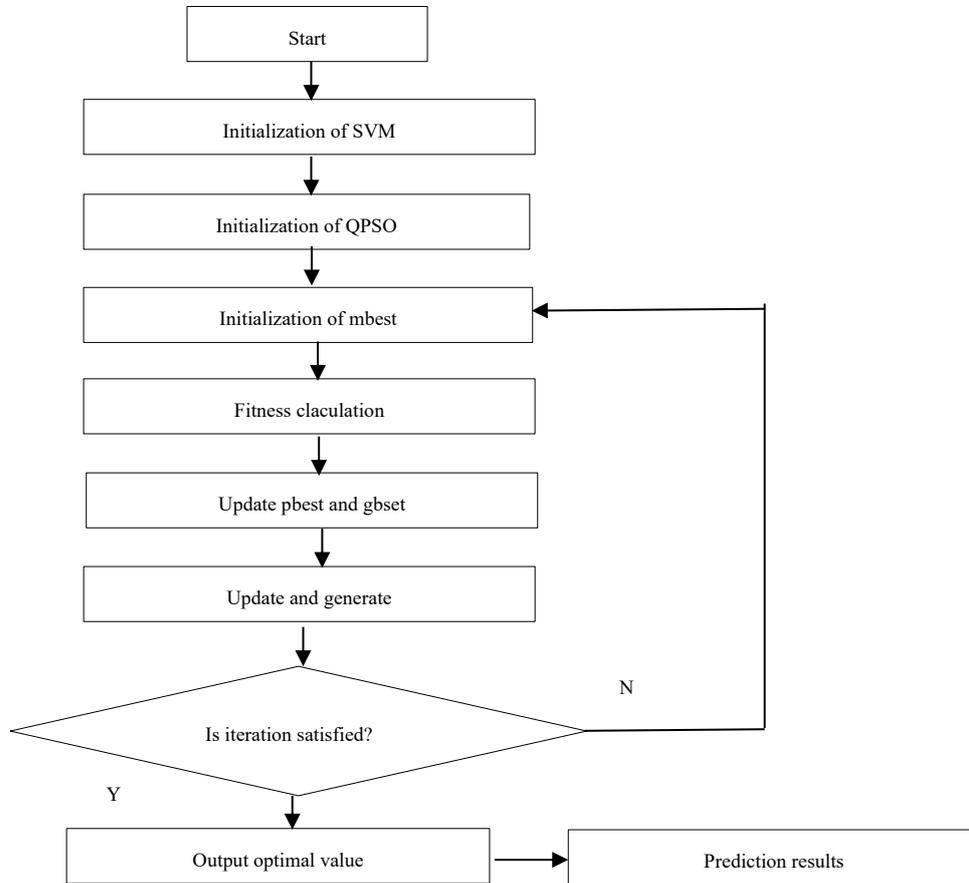


Fig2.Flowchart of QPSO-SVM model

### 2.3 FICA-FCM +QPSO-SVM comprehensive model

Based on different design ideas and learning strategies, clustering analysis can be used as a separate process to reveal the inherent properties and laws of data samples by searching for the internal distribution structure of data samples, and it can also be used as a precursor process of other learning tasks such as classification, which provides the basis for further data analysis because of the complexity and diversity of coal and gas outbursts, the sample data also has a complex relationship. the internal relationship of sample data is very helpful to improve the prediction accuracy. In this paper, we use supervised learning SVM to predict coal and gas outburst samples, the class accuracy is very low, some samples are always wrong in classification, which affects the classification effect of classifiers. In order to mine the essential structural rules of sample data, and analyze the structural information and internal regularity of sample data set, the correct classification of samples is realized according to the estimation method of sample similarity and probability density function. In view of the above problems, this paper combines FICA-FCM unsupervised learning and SVM unsupervised learning classification algorithm. Firstly, the coal and gas outbursts is classified by using SVM, and the preliminary results of classification are obtained. Then, the FCM algorithm based on FICA feature extraction is used to cluster the sample data, and the clustering results are obtained. The classification results are compared and modified, and added to the training set to continuously improve the classification accuracy. A complete unsupervised and supervised combination process mainly includes four stages: FICA feature extraction, FCM clustering analysis, QPSO parameter optimization of SVM, outbursts classification and result analysis. The specific

flowchart of FICA-FCM +QPSO-SVM is as follows in Fig.3:

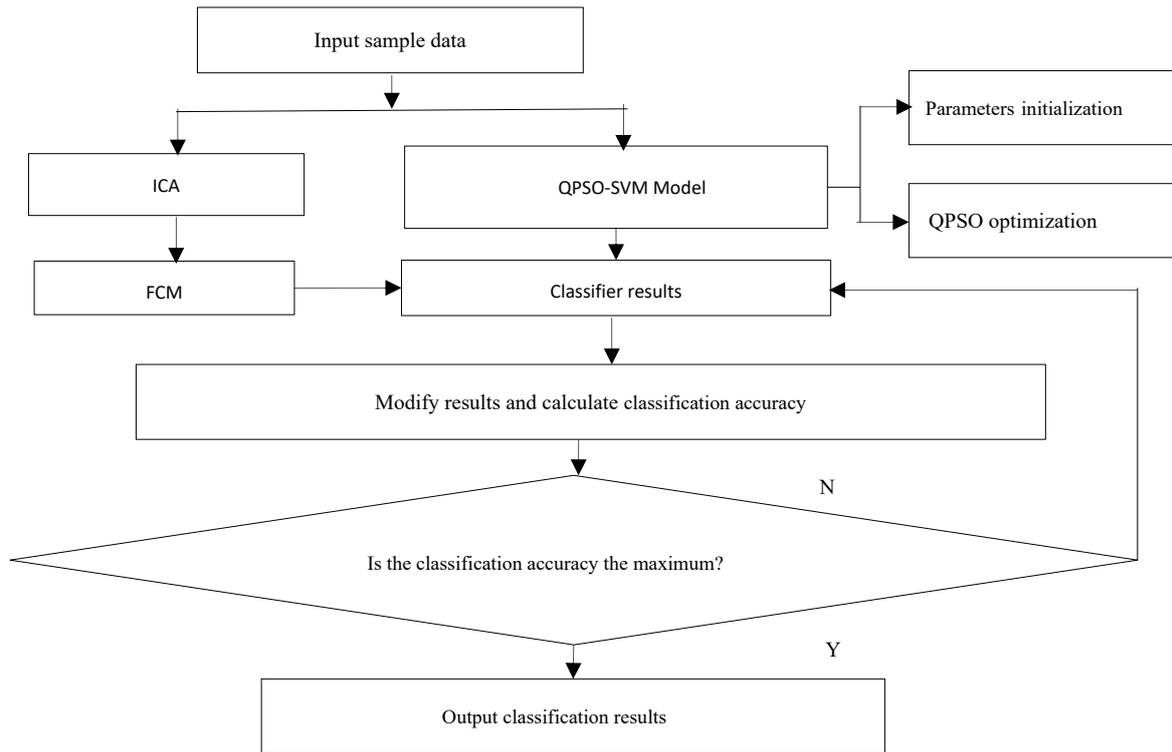


Fig.3 Flowchart of FICA-FCM +QPSO-SVM model

### 3 Experimental result and analysis

#### 3.1 Dataset description and preprocessing

The influencing factors of coal and gas outbursts include geological stress, gas and physical properties of coal seam. The experimental data come from the historical data of Pingdingshan No.8 mine in Henan Province. With reference to previous studies[31], the coal and gas outbursts influencing indexes include the following: gas pressure, initial velocity of gas output, initial velocity of gas emission, coal seam firmness coefficient, structural coal thickness, fault structure complexity. These indexes are operability, extensive and applicable in engineering practice. In order to make the experimental results more objective, the method of ten-fold cross-validation is used to verify the classification effect, that is, the data set is randomly divided into ten parts, one of which is taken as the testing set in turn, the other nine parts are taken as the training set, and then the corresponding classification method is run to classify and learn the data in the training set, and the testing set is used to test, and the corresponding 10 times testing results are obtained. In this paper, coal and gas outbursts prediction is a two-classification (outbursts or normal). According to the data distribution characteristics of coal and gas outbursts, due to the use of different units in the acquisition, the meaning is different, there are great differences in the order of magnitude, which affects the convergence speed and accuracy of the algorithm. Before training the model, this paper uses the normalization method to process these data into new data within the range of [0,1], and removes the dimension to make the indexes more comparable.

#### 3.2 Experimental environment and parameter setting

The validity of the algorithm is further verified by comparing experiments on real datasets. All experiments are completed on a CPU with the main frequency of 2.8GhZ and a PC with 4GB memory. The operating system is windows7, and the algorithm is implemented in Python 3.7 environment. TP indicates the number of samples correctly marked for application, TN indicates the number of samples correctly marked for application, FN indicates the number of samples incorrectly marked for application, and FP indicates the number of samples incorrectly marked for application. According to the above evaluation parameters, to obtain statistically meaningful results, 10 independent experiments are performed for each approach. The

statistical results concerning 1)Accuracy, 2) Precision,3) Sensitivity, 4) Specificity for each approach along with their standard deviations are calculated according to related references, and the runtimes for all approaches are reported as well. In order to eliminate the random factors, we perform ten times with different random seeds on each experiment. The average performance of these ten repetitions is regarded as the final result. The proposed computational method such as (SVM, KNN, NB) are sensitive to the values of their main controlling parameters, the kernel function of SVM is Gaussian kernel function. Two parameters  $\gamma$  and regularization parameter  $C$  determine the best classification. we use the combination of QPSO optimization and the 10-fold cross validation to find out all kinds of parameters structure of the classifier. The parameters of other classifiers are made according to the test. The parameters setting of feature extraction are as follows: The parameter of FICA is set to 3, and the parameters of the KPCA and FA are 4,4 respectively.

### 3.3. Experimental results and analysis

#### 3.3.1 Performance comparison of combination of different feature extraction and clustering analysis

Table1 Comparison of different clustering analysis methods

Method	Accuracy	Precision	Specificity	Sensitivity
NONE	0.73	0.77	0.70	0.75
F-MEANS	0.98	0.97	0.98	0.98
K-MEANS	0.90	0.87	0.85	0.93
SOM	0.91	0.95	0.93	0.85

Table2 Comparison of combination of different feature extraction and FCM

Method	Accuracy	Precision	Specificity	Sensitivity
KPCA-F-MEANS	0.93	0.91	1.0	0.83
FICA-F-MEANS	0.99	1.0	0.93	1.0
FA-F-MEANS	0.98	0.97	0.98	0.98

Table3 Comparison of combination of different clustering analysis and ICA feature extraction

Method	Accuracy	Precision	Specificity	Sensitivity
FICA -FMEANS	0.99	1.0	0.93	1.0
FICA-SOM	0.88	0.87	0.86	0.89
FICA -GS	0.96	0.98	0.98	0.88
FICA -DBSCAN	0.98	1.0	0.93	1.0

In order to verify the effectiveness of the combination of FCM and FICA, we employ different clustering algorithms and different feature extraction methods to predict the coal and gas outbursts. From Table1-3, we can see the performance comparison results using SVM. Table1 gives the performance comparison of different clustering methods, we can see that the accuracy of clustering analysis can achieve higher accuracy than the accuracy without clustering, among which the accuracy of FCM is the highest, reaching 92%, followed by self-organizing neural network(SOM) and K-means, and the values are 90% and 92% respectively. This indicates that the original label of sample data has errors, the error of sample label can be corrected by clustering algorithm, which can improve the prediction accuracy. In order to illustrate the effectiveness of the feature extraction method proposed in this paper, Table2 gives the comparison results of different feature extraction and FCM methods. We can find that the accuracy of FICA -FCM is the highest, reaching 100%, while the accuracy of KPCA and FA are the 0.93,0.98 respectively, which means that these feature extraction methods cannot extract effective feature information from coal and gas outburst sample data, and FICA -FCM is very suitable for coal and gas outbursts prediction. According to FICA feature extraction, the feature points of the original signal can be compressed, the redundant information in the signal can also be removed, and the feature information of the high-order moment of the data set is used to mine the relative independent information in the original signal, so that the feature information covered by the cross signal can be effectively separated from each other, and it can improve the modeling effect of the characteristic difference information. FICA relies on only high-order

statistics information, which can extract non-Gaussian components and global information of process data effectively. The complexity of the detection model is reduced, and extract the sample data effectively, redundancy and noise data can be eliminated, FICA algorithm is more suitable for the feature extraction of coal and gas outbursts, meanwhile, the effective clustering algorithm can improve the clustering effect and efficiency based on the above feature extraction methods. Table3 shows the experimental results of FICA and different clustering methods. We can find that the accuracy of FCM is the highest, reaching 100%, and the others are also lower than the FCM methods. Among different clustering algorithms, FCM algorithm is more suitable for the characteristics of sample data and better clustering data.

### 3.3.2 Comparison performance of different stages of proposed method

Table4 Comparison of different stages of ours

Method	Accuracy	Specificity	Sensitivity	TIME
FICA	0.75	0.75	0.75	3.4
F-MEANS	0.98	0.98	0.98	3.7
FICA- F-MEANS	0.99	0.93	1.0	0.45

In order to verify the importance of each algorithm in the process of mean clustering algorithm based on feature extraction proposed in this paper, we do experiments on the dataset through the methods of FICA feature extraction, FCM and the combination of the two. Table 4 shows the accuracy and efficiency of each stage in the classifier. We can see that the accuracy of FICA-FCM is the highest and the time is the shortest. The indexes are 0.99,0.93,1, the execution efficiency is 0.45 seconds; the accuracy rate of FICA method on classifier is 0.75 and the execution time is 3.4 seconds; the accuracy rate of FCM algorithm without feature extraction is 0.98, and the execution time is 3.7 seconds. This shows that the feature extraction is very necessary because of the redundancy and noise in coal and gas outbursts sample data, the nonlinear low dimensional essential structure can be extracted by FICA feature extraction, then the effect and efficiency of FCM can be improved.

### 3.3.3 Comparison of different classifiers

Table5 Performance comparison of different classifiers

Method	Accuracy	Precision	Specificity	Sensitivity
SVM	0.98	1.0	0.93	1.0
QPSO-SVM	0.99	1.0	0.93	1.0
RF	0.94	0.94	0.89	0.96
DT	0.92	0.87	0.95	0.89
KNN	0.94	0.90	0.95	0.93

Different classifiers have different performance, in order to verify the effectiveness of our proposed model, we choose different classifiers and our improved SVM classifier for comparative experiments. From Table 5 we can find out that our proposed classifier based on QPSO optimization algorithm has the highest classification accuracy of 0.99, which is higher than that of SVM, DT and RF classifier. This is because the combination of the improved SVM model and the FCM model based on feature extraction fully considers the advantages of the two, so it is necessary to select the appropriate feature extraction methods according to the characteristics of the coal and gas outbursts sample data. At the same time, it is necessary to correct the original label error through clustering model. The classification error rate on SVM is the standard to correct the classification label and improve the classification accuracy.

### 3.3.4 Performance comparison with the literatures

Table9 Comparison with different clustering algorithms in literatures

Method	Accuracy	Sensitivity	Time(s)
--------	----------	-------------	---------

Ours	0.99	1	0.45
FCM-SVM[3]	0.92	0.92	2.98
K-MEANS-SVM[4]	0.89	0.92	1.27
AP-RVM[32]	0.90	0.93	2.19
SOM-SVM[9]	0.96	0.98	0.32

To prove the validity of our proposed model, we choose different clustering methods in the literatures, and MKRVM and SVM are employed to carry out comparative tests, the results of accuracy, sensitivity and running time as presented in Table9, we can find that the accuracy of ours, FC-SVM, K-MEANS-SVM, AP-RVM and SOM-SVM are 0.99, 0.92, 0.89, 0.90 and 0.96 respectively, which are lower than the performance indexes of our comprehensive model. The FCM method is sensitive to the initial value of the data and has low accuracy, strongly depends on the quality of initialization data and is easy to fall into local saddle point; K-MEANS is sensitive to the initial selected centroid points and different random seed points have different clustering results, which has a great impact on the clustering results, using iterative method it may only get the local optimal solution, but it cannot get the global optimal solution; The number of hidden neurons in the SOM is difficult to determine, so the hidden neurons are often not fully utilized, and some neurons are far from the learning vector cannot win, thus becoming dead nodes, the learning rate of the clustering network needs to be set artificially, and the learning termination needs to be controlled artificially, which affects the learning progress; At the same time, the FCM based on FICA not only can reveal the potential data structure and the correlation between variables, but also can describe the complex structural features and find the potential dynamic laws. By revealing the inherent and low-dimensional geometric properties, this paper explores the FICA driving process, which can reduce the indicators needed for FCM, accelerate the clustering speed and provide more accurate classification estimation for SVM classifier, the algorithm proposed by us is more suitable for the characteristics of coal and gas outbursts sample data.

At the same time we also compare the performances of different feature extraction methods on different classifiers in terms of computation time. We can see from Table9 that the running time of ours is 0.45 seconds, which takes less total time than other literatures such as reference [3](2.98s), reference [4](1.27s), reference[32](2.19s), only the running time of reference[9] is 0.32s, which is the lowest among them, which is higher than ours. This is due to the fact that FirstICA -FCM method can provide more effective classification information for the classifier, improve the performance and efficiency of classification. In the design of classifier, the proposed SVM employs QPSO algorithm to improve the accuracy with much more time cost. To sum up, our proposed method has higher classification accuracy and much less computation cost compared to other classification methods. Hence, we can conclude that our proposed method is more applicable than other classification methods for classification of coal and gas outbursts in the case of fewer samples, due to its excellent feature extraction and higher accuracy. However, the proposed model has the following defects. The proposed model has only been validated on one available dataset which come from pingdingshan no.8 mine, but the datasets of different mines can be tested in order to achieve better generalization performance. The paper deals with solving a two-class classification problem, however a multi-class outbursts classification problem is highly in demand.

## 4 Conclusion

The exiting classification effect of the model for coal and gas outbursts is not very good, and the classification results of some samples are always wrong, there may be some wrong labeled samples, the wrong sample data are easy to generate noise data, then affect the classification effect of the classifier. The clustering analysis model can find the rule of sample data characteristics, and it can be regarded as the pretreatment of SVM classification, which improves the accuracy of classification, but the existing clustering analysis model has low computational efficiency, high complexity and low accuracy. FICA feature extraction can reveal the potential data structure of data samples and the correlation between variables, describe the complex structural characteristics, and discover the potential dynamic laws; FCM algorithm can mine and analyze the structural information and internal regularity of the sample data set. According to the sample similarity and probability density function estimation method, the correct classification of samples is realized. The combination of FICA and FCM method can be used to improve the clustering effect and efficiency of coal and gas outburst. On the basis of the classification results obtained from

the above clustering analysis and feature extraction, the labeled sample label and the original label are compared, the results of classifier are modified. Considering the sample classification error rate of the SVM classifier, the error labeled samples with low classification accuracy are found, and the error labels are corrected to improve the classification performance. The optimal classification model of SVM is worth studying to reduce the influence of noise data caused by annotation errors on the prediction accuracy. Although the proposed model gives significant results of coal and gas outbursts, there are still some limitations in this study. In the future work, we will find out suitable feature extraction and clustering analysis methods, and modify existing classifiers to improve the comprehensive performance of coal and gas outbursts prediction.

### **Availability of data and material**

The raw/processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing study. Some or all data, models, or code generated or used during the study are available from the corresponding author by request.

### **Declaration of Competing Interest**

The authors declare no competing financial interest.

### **Acknowledgments**

This research was supported by the National Natural Science Foundation of China (U1704242). In addition, the authors also thank the editors, and the anonymous reviewers for their valuable comments that greatly improved the quality of this paper.

### **Reference**

- [1] Chen K P. A new mechanistic model for prediction of instantaneous coal outbursts -Dedicated to the memory of Prof. Daniel D. Joseph. *International Journal of Coal Geology*, 2011, 87(2):72-79.
- [2] Li H Y. Major and minor structural features of a bedding shear zone along a coal seam and related gas outburst, Pingdingshan coalfield, Northern China. *International Journal of Coal Geology*, 2001,47(2):101-113.
- [3] Qin shuyu. Fuzzy gathering classification similar analysis for the coal and gas outburst. *Chinese Journal of Geological Hazard and Control*, 14(4):58-61
- [4] Wang Hongyuan. *Research on coal mine safety monitoring system based on data mining*. Liaoning: Liaoning University of engineering and technology, 2009
- [5] Gui Xiangyou, Yu Zhongming. Analysis of coal and gas outburst degree based on fuzzy cluster correlation analysis, *Journal of Safety and Environment*, 2006,6(4):85-89.
- [6] Cheng Cheng. *Application Research of data mining in coal mine gas monitoring system*. Xi'an: Xi'an University of science and technology, 2013
- [7] Zheng Maojie, *study on coal and gas outburst prediction and emergency response mechanism*, Beijing: China University of mining and technology, Beijing, 2009
- [8] Hou J.(2017).Coal and Gas Outburst Hazard Analysis of Complex Coal Seam Based on Fuzzy Clustering Method. *Safety in Coal Mines*, 48(6):160-165
- [9] Zhang Xiaoming, Zhang Airong, Uuo Yongyi. Application of optimized self-organizing competitive neural network by PSO algorithm for coal and gas outburst prediction, Taiyuan: Taiyuan University of technology, 2011
- [10] YAN Xin, FU Hua, TU Naiwei. Dynamic Prediction of Coal and Gas Outburst Based on Clustering and Case-Based Reasoning, *CHINESE JOURNAL OF SENSORS AND ACTUATORS* 2016,29(4):545-552
- [11] Li Sheng. Study on pattern recognition method for regional prediction of coal and gas outburst, Liaoning: Liaoning University of engineering and technology, 2007,
- [12] Zhang Meimei, Cai Donglei, Cao Qingchun, Coal and gas outburst intensity prediction based on AFSA-PCA-FCM coupling model, *Transducer and Microsystem Technologies*, 2017, 36(12):50-55
- [13] Liang Yueqiang, Guo Deyong, Huang Zhanfeng et al. Prediction model for coal-gas outburst using genetic projection pursuit method. *International Journal of Oil, Gas and Coal Technology*, 2017,16(3): 271-282.
- [14] Lu Guifu, research on feature extraction algorithm and face recognition in the case of small samples, Nanjing: Nanjing University of technology, 2012
- [15] Ye Qiacrlin, Ye ding, Zhang Xun-hua. Extreme Decomposition based mixtures of Kernels and its improvement. *Pattern Recognition and artificial*

Intelligence, 2009, 8(3):366-373

- [16] Wang Zhihui, Fu hua. Application of Improved Support Vector Machine in Coal Mine Safety Evaluation. MINING SAFETY & ENVIRONMENTAL PROTECTION,2007, 34(1), 82-84,87
- [17] Fang Jun, research on swarm intelligence hybrid optimization algorithm and its application, Nanjing: Nanjing University of technology, 2014
- [18] Wen, T., Zhang, B., Shao, L. Prediction of coal and gas outburst based on random forest model. Computer Engineering and Applications, 2014,50 (10): 233-238.
- [19] Liu Jun, Zeng Fanlei, Guo Zhanglin. Study on multiple factors risk evaluation of coal and gas outburst based on RS-SVM model. China Safety Science Journal, 2011, 21(7):21-26.
- [20] Kao, L.-J., et al. Integration of Nonlinear Independent Component Analysis and Support Vector Regression for Stock Price Forecasting. Neurocomputing, 2013,99: 534-542
- [21] BELL B A, SEJNOWSKI T. An information maximization approach to blind source separation and blind deconvolution. Neural Computation, 1995, 7(6): 1129 – 1159.
- [22] HYVARINEN A. Fast and robust fixed-point algorithms for independent component analysis. IEEE Transactions on Neural Networks,1999, 10(3): 626 – 634.
- [23] Subasi A, Gursoy M I. EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Systems with Applications, 2010,37 ( 12) : 8659-8666
- [24] Ding S J, Sun Z Y, Sun Y K, Wang Y G. Combining principal component analysis network with linear discriminant analysis for the classification of retinal optical coherence tomography images. Journal of Image and Graphics, 2019, 24(01):0115-0123.
- [25] DING S F, XU X Z, NIE R. Extreme Learning Machine and Its Applications. Neural Computing and Applications, 2014, 25(3/4): 549—556.
- [26] UZAIR M, SHAFAIT F, GHANEM B, etal. Representation learning with deep extreme learning machines for efficient image set classification//MACINTYRE J. Neural Computing and Applications. London:Springer, 2015:1-13.
- [27] Chen Xin-Quan. Dual clustering method of mixed data set. Computer Engineering and Science, 2013, 35(2):127-132 doi: 10.3969/j.issn.1007-130X.2013.02.022
- [28] Zhou Kaile, Yang Shanlin, Wang Xiaojia, Chen Zhiqiang. Load classification based on improved FCM algorithm with adaptive fuzziness parameter selection, Systems Engineering-Theory & Practice, 2014, 34(5), 1283-1289
- [29] H. Gan, N. Sang, R. Huang, X. Tong, and Z. Dan. Using clustering analysis to improve semi-supervised classification, Neurocomputing, 2013, 101(3): 290–298.
- [30] Feng Chang, Liao Shizhong. Model Selection for Gaussian Kernel Support Vector Machines in Random Fourier Feature Space. Journal of Computer Research and Development, 2016, 53(9): 1971-1978.
- [31] Zhang, Y., Cui, J., Jiao, X., Study of the multi-index coupling forecasting model of coal and gas outburst and its application. Chinese Journal of Engineering, 2018, 40(11):1309-1316.
- [32] An jiyong. Research on Predicting Protein-Protein Interactions Based on Relevance Vector Machine, PhD thesis, China university of mining and technology Beijing campus. 2018

# Figures

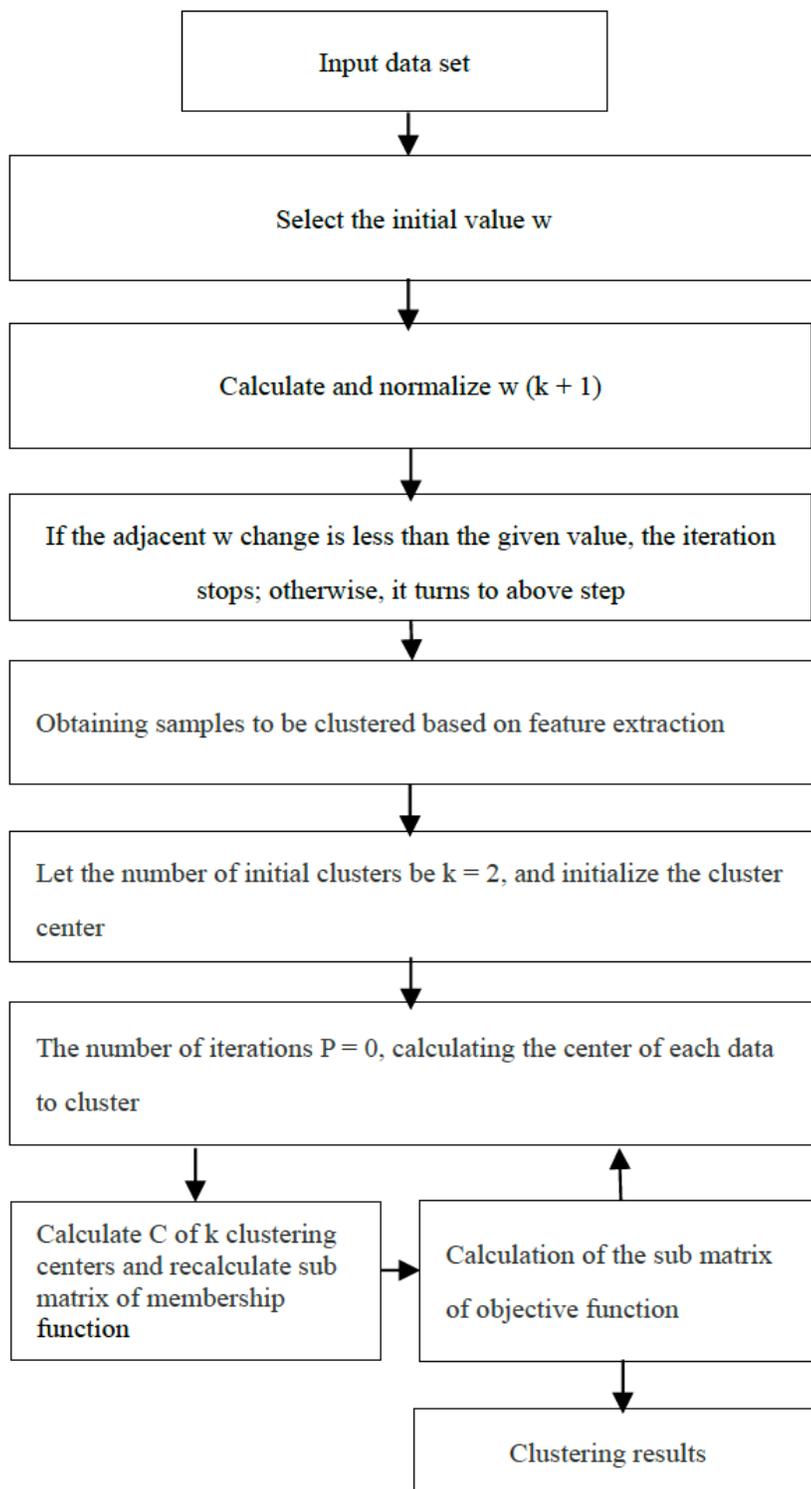
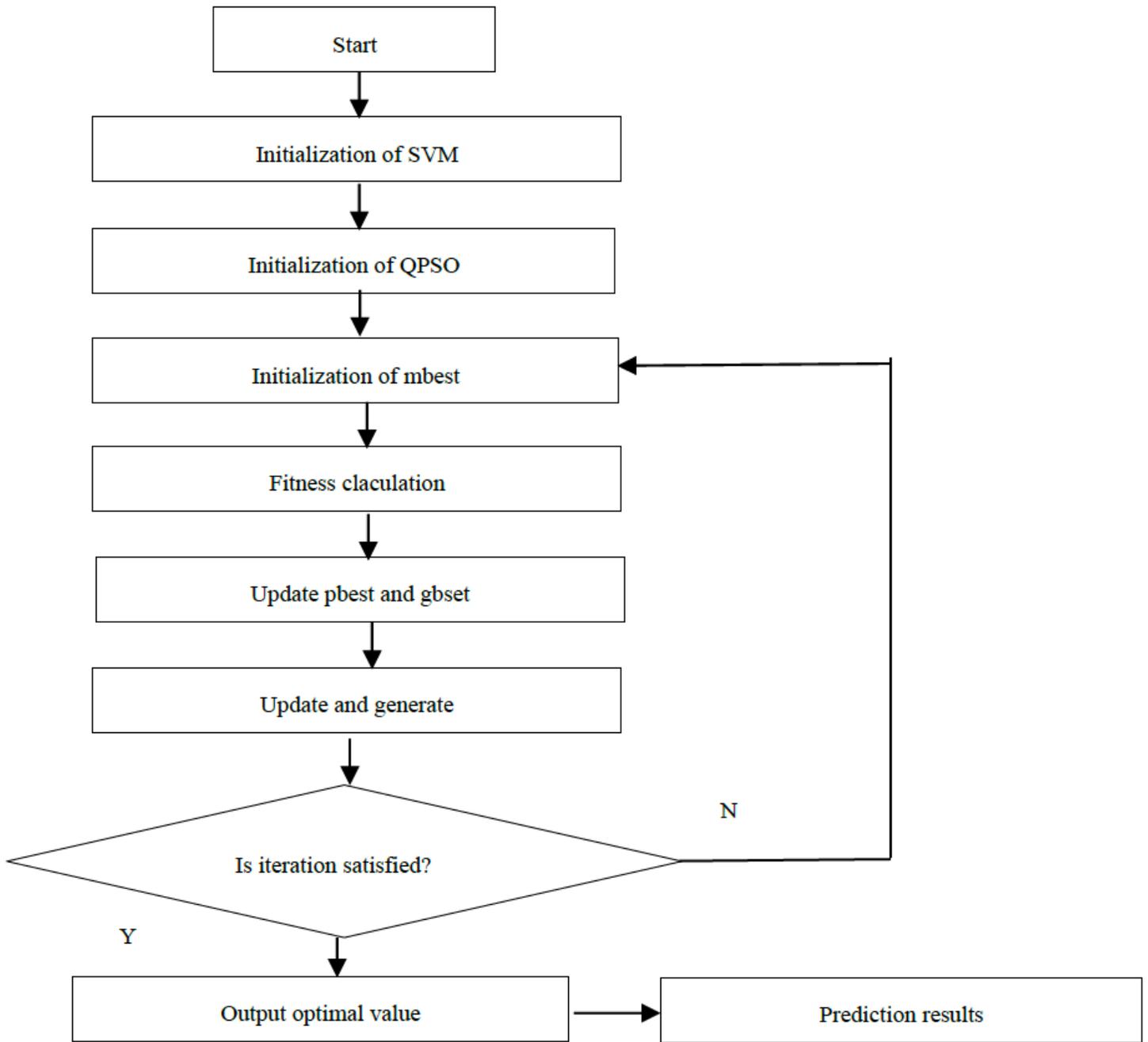


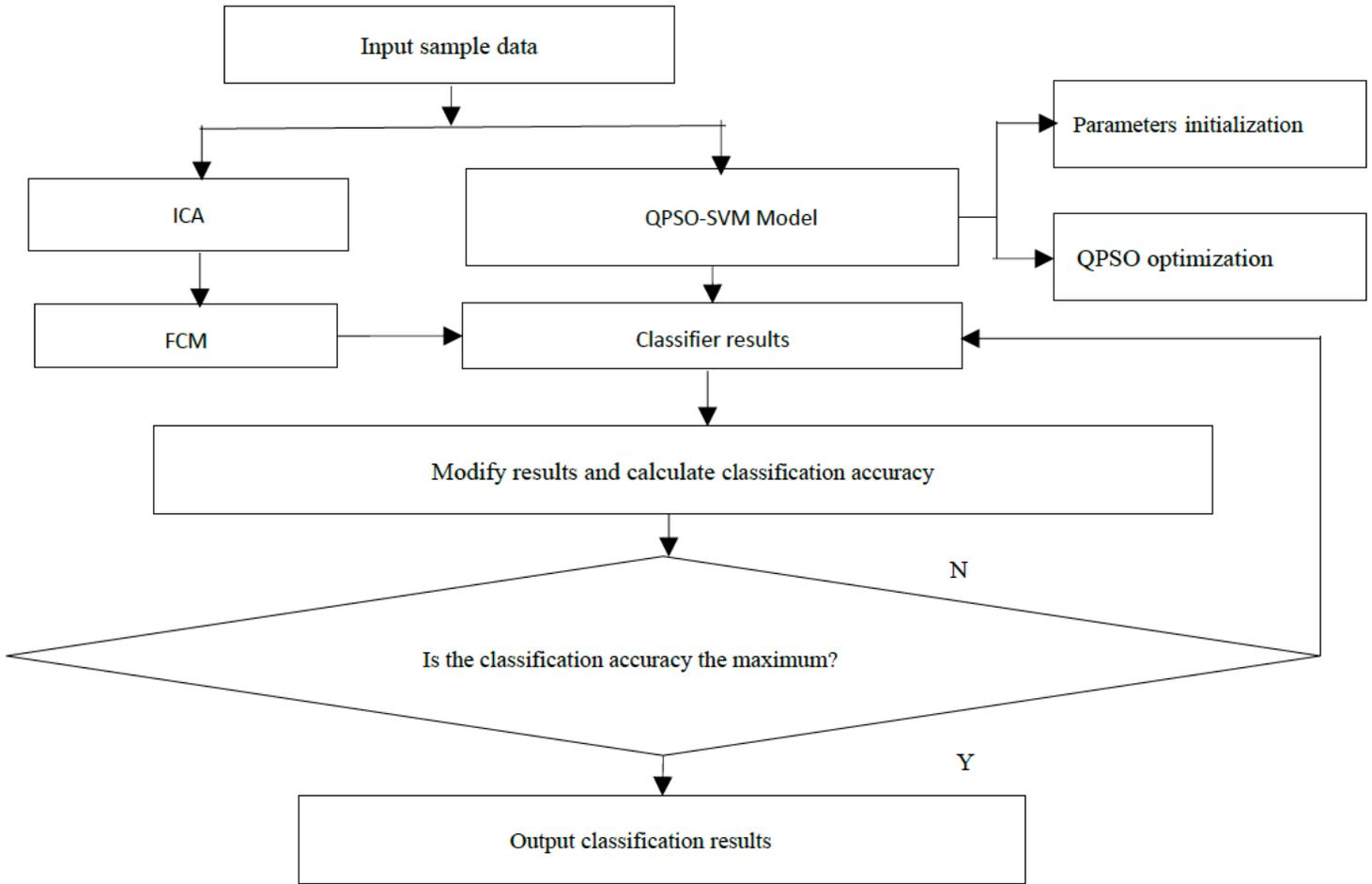
Figure 1

Flowchart of FICA-FCM model



**Figure 2**

Flowchart of QPSO-SVM model



**Figure 3**

Flowchart of FICA-FCM +QPSO-SVM model