A novel tree-based algorithm for real-time prediction of rockburst risk using field microseismic monitoring

Xin Yin (yinxin_engineering@163.com)
Wuhan University
Quansheng Liu
Wuhan University
Yucong Pan
Wuhan University
Xing Huang
Chinese Academy of Sciences

Research Article

Keywords: Rockburst, Intensity prediction, Tree-based algorithm, Microseismic monitoring, Precursory microseismic sequence

Posted Date: July 29th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-188500/v1

Version of Record: A version of this preprint was published at Environmental Earth Sciences on August 5th, 2021. See the published version at https://doi.org/10.1007/s12665-021-09802-4.

A novel tree-based algorithm for real-time prediction of rockburst risk using field microseismic monitoring

Xin Yin [a, b], Quansheng Liu [a, b, *], Yucong Pan [a, b, *], Xing Huang [c]

a. *The Key Laboratory of Geotechnical and Structural Engineering Safety of Hubei Province, School of Civil Engineering, Wuhan University, Wuhan 430072, China*

b. *State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China*

c. *State Key Laboratory of Geomechanics and Geotechnical Engineering, Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, Wuhan 430071, China*

Abstract: Rockburst is a kind of complex and catastrophic dynamic geological disaster in the development and utilization of underground space, which seriously threatens the safety of personnel and environment. Due to the suddenness in time and randomness in space, the prediction of rockburst becomes a great challenge. Microseismic monitoring is capable to continuously capture rock microfracture signals in real time, which offers an effective means for rockburst prediction. With the explosive growth of monitoring data, the conventional manual forecasting methods are laborious and time-consuming. Therefore, artificial intelligence was introduced to improve the prediction efficiency. A novel tree-based algorithm was proposed. Its basic idea was to automatically recognize precursory microseismic sequences for the real-time prediction of rockburst intensity. The database consisting of 1500 microseismic events was analyzed. In order to establish precursory microseismic sequences, dimensionality reduction of the database was first implemented by t-SNE algorithm. Then, *k*-means clustering algorithm was employed for labelling 1500 microseismic events. Before that, canopy algorithm was adopted to

determine the number of clusters. Finally, 300 precursory microseismic sequences were formed by grouping rule. They were further partitioned into two parts through stratified sampling: 70% for training and 30% for validation. The validation results indicated that the precursor tree with pruning achieved higher prediction accuracy of 98.9% than one without pruning on the validation set. And the increase was separately 12.2%, 9.2% and 28.6% on the whole validation set and each classes (low/moderate rockburst). In comparison with low rockburst, moderate rockburst was minority class. The improved accuracy on moderate rockburst suggested that pruning can enhance the recognition ability of precursor tree for minority class. Additionally, two extra rockburst cases were collected from a diversion tunnel in northwestern China, which provided a complete workflow about how to apply the built precursor tree model to achieve field rockburst warning in engineering practice. The tree-based algorithm served as a new and promising way for the real-time rockburst prediction, which successfully integrated field microseismic monitoring and artificial intelligence.

Keywords: Rockburst; Intensity prediction; Tree-based algorithm; Microseismic monitoring; Precursory microseismic sequence

***Corresponding authors.**

E-mail address: liuqs@whu.edu.cn (Quansheng Liu);

pyc1991@whu.edu.cn (Yucong Pan).

1 Introduction

Rockburst is a common geological disaster in the process of deep underground excavation (Cai et al. 2018; Duan et al. 2021; Liang et al. 2020; Ma et al. 2019a; Ma et al. 2021; Wang et al. 2020; Xu et al. 2018). It is extremely destructive, which directly threatens the safety of personnel and equipment, affects the construction progress, and even destroys the entire project and causes earthquake (Afraei et al. 2019). The mechanism of rockburst is complicated (Li and Jimenez 2017; Tang et al. 2010). It is generally assumed that rockburst is induced by both internal and external factors (He et al. 2018). Internal factors include ground stress, physical and mechanical properties of surrounding rock. They determine the energy storage capacity of rock (Xu et al. 2017; Zhang et al. 2018). On the another hand, external factors mainly refer to excavation disturbances, which make the elastic strain energy accumulated in rock mass suddenly released (Sirait et al. 2013). Rockburst has occurred in all mining countries, such as Leipzig coal mine in the United Kingdom (Keneti and Sainsbury 2018), Luer coal mine in German (Baltz and Hucke 2008), Makassar gold mine in Canada (Liang et al. 2019), Hongtoushan copper mine in China (Dong et al. 2013). Except mining industry, there have been more and more rockburst records in the fields of water conservancy and hydropower, traffic tunnel in recent years (Pu et al. 2019a). On November 28, 2009, an extremely strong rockburst happened in the drainage tunnel of Jinping II Hydropower Station (Feng et al. 2017). The support systems were completely damaged, a tunnel boring machine (TBM) was permanently buried and seven people were killed. With the development of underground space, rockburst prediction becomes an urgent challenge

to be solved for safe construction (Ma et al. 2019b; He et al. 2021; Oparin et al. 2017; Qiu et al. 2020).

Numerous empirical criteria were proposed for rockburst prediction from different aspects. Russenes criterion (Russenes 1974), Barton criterion (Barton 2002) and Hoek criterion (Hoek and Brown 1980) were created from the perspective of stress while Goodman criterion (Goodman 1980) and Kidybinski criterion (Kidybiński 1981) were raised on the basis of energy. In addition to stress and energy, rock brittleness and critical depth were used to establish criteria, too (Xue et al. 2020). Since rockburst is jointly controlled by many factors, some comprehensive evaluation methods in applied mathematics were introduced, such as fuzzy comprehensive evaluation (Cai et al. 2016) and expert system (Webber 1996). They can synthesize multiple perspectives to predict rockburst, which surmounts the deficiency of empirical criteria. With the deepening of research, it was gradually realized that there is a highly nonlinear relationship between rockburst and its influencing factors (Zhou et al. 2012). Artificial intelligence has unique advantages for solving nonlinear problems (Zhou et al. 2018). Therefore, it was brought in to rockburst prediction. Feng and Wang (1994) first used artificial neural network to predict rockburst, and later, Sun et al. (2009) and Jia et al. (2013) also carried out the related study. Zhou et al. (2016) adopted ten supervised learning algorithms and compared their performance in rockburst prediction. Pu et al. (2019b) employed support vector machine to build rockburst classifier and successfully applied it to kimberlite pipes at a diamond mine. Except support vector machine, Pu et al. (2018) made use of decision tree to predict rockburst, which tackled the obstacle caused by missing data.

Adoko et al. (2013) developed fuzzy inference systems based on field measurement data for predicting rockburst intensity. Luis et al. (2017) set up naive Bayesian classifier, tree-augmented naive Bayesian classifier and augmented naive Bayesian classifier to access rockburst risk. The above models have achieved favorable prediction results in many projects. Considering the difficulty in obtaining rock mass parameters quickly and accurately during excavation, these models are tough to conduct real-time rockburst prediction, which are more applied to long-term prediction in engineering investigation stage (Zhang et al. 2020; Zhou et al. 2020).

Microseismic monitoring can continuously monitor the evolution process of rock microfractures in real time, whose basic principal is to receive elastic waves generated by rock microfractures and retrieve their source parameters (Liu et al. 2018). It has become a powerful tool for rockburst prediction in underground powerhouse, mines, tunnels and other projects (Ma et al. 2018). Besides, it is also widely applied to slope stability analysis (Xu et al. 2011), water inrush prediction (Cheng et al. 2018) and gas dynamic disaster warning (Zhao and Tan 2010). For different types of rockburst, the microseismic characteristics have been studied deeply (Feng et al. 2019a; Liu et al. 2020). The temporal and spatial evolution laws of source parameters can better reflect the generating process of rockburst. With the explosive growth of field monitoring data, traditional manual forecasting methods have low efficiency. Hence, the combination of microseismic monitoring and artificial intelligence is urgent for rockburst prediction. There are few reports on this. Feng et al. (2019b) built an microseismic monitoring-based intelligent rockburst prediction model and applied it in Jinping II Hydropower

Station.

In this paper, a novel tree-based algorithm was put forward. Its basic idea was to automatically recognize precursory microseismic sequences for real-time prediction of rockburst intensity. A complete procedure consisting of model construction and field warning was introduced in detail. The database including 1500 microseismic events was analyzed and 300 precursory microseismic sequences were established by using a series of data mining algorithms flexibly. First, t-SNE algorithm was implemented to reduce dimensionality of the database. Second, *k*-means algorithm was executed for clustering 1500 microseismic events. Before utilizing *k*-means algorithm, canopy algorithm was employed to determine the number of clusters. Finally, according to the grouping rule, 1500 microseismic events formed 300 precursory microseismic sequences. The precursor tree under two versions (no pruning/pruning) was build. For assessing the prediction performance of precursor tree, 300 precursory microseismic sequences were partitioned into two parts by stratified sampling: training set (70%) and validation set (30%). Through comparing the prediction accuracy on the whole validation set and each class, the precursor tree with pruning was selected as the final precursor tree used for field rockburst warning. Additionally, two extra rockburst cases, collected from a diversion tunnel in northwestern China, provided a complete workflow about how to apply the constructed precursor tree model to achieve warning in engineering practice.

2 Database description

2.1 Data source

All data in the database is collected from a diversion tunnel project in northwestern China. The tunnel has a length of 41.832km and its maximum burial depth is 2268m, which is excavated by the combination of tunnel boring machine (TBM) and drilling-blasting method (DBM) (Deng and Liu 2020; Deng et al. 2020; Liu et al. 2020). Specifically, 32.842km is excavated by TBM and the rest is constructed by DBM. The geological profile along the tunnel is presented in Fig. 1. As observed in Fig. 1, Section 2 is the study area of this paper. The parts with a burial depth greater than 1000m and 2000m account for about 53.3% and 14.4%, respectively. In addition, the geo-stress test results obtained by the hollow inclusion stress-relief method indicate that the maximum principal stress reaches 50MPa. Under the condition of large burial depth and high geo-stress, rockburst poses a serious threat to construction safety.

Fig. 1 Geological profile along the diversion tunnel

ESG (Engineering Seismology Group, Canada) system is utilized to carry out field microseismic monitoring, which capture rock fracture signals of rockburst generating. The hardware mainly includes a portable Paladin data acquisition instrument and six uniaxial accelerometers, which is shown in Fig. 2. The software consists of Hyperion network acquisition system (HNAS) and waveform visualizer (WaveVis). The HNAS can automatically pick the onset time of P-wave and S-wave using STA/LTA algorithm while the WaveVis is a waveform visualization tool. Considering site conditions, three monitoring sections are arranged and each section is equipped with two accelerometers, which are installed symmetrically at the tunnel spandrel. The top view of the layout is shown in Fig. 3. In the ESG microseismic monitoring system, a single-velocity model

is adopted to locate the microseismic events. In order to determine the P-wave velocity, a new method combing the true reflection tomography (TRT) technique is put forward, which has been detailed introduced in Liu et al. (2020).

Fig. 2 Components of ESG microseismic monitoring system

Fig. 3 Top view of the layout of ESG microseismic monitoring system

300 rockburst cases in total are collected from engineering records, occurring from January to September 2018. As shown in Fig. 4, there are 252 low rockburst cases and 48 moderate rockburst cases. The determination of rockburst intensity is in accordance with the national standards of People's Republic of China (The National Standards Compilation Group of People's Republic of China, 2009). The description about rockburst phenomena under different intensity is shown in Table 1. In order to establish microseismic database, five microseismic events before rockburst occurs are extracted from the microseismic monitoring catalog. Therefore, the database is made up of 1500 microseismic events.

Fig. 4 Proportion of rockburst with different intensity

Table 1 Description of rockburst phenomena with different intensity

2.2 Feature determination

The goal of feature determination is to characterize microseismic events, preparing for the quantitative prediction of rockburst intensity. It is required that selected features are able to comprehensively reflect the intrinsic characteristics of microseismic events. Meanwhile, these features should be easy to access.

Microseismic event is caused by rock microfracture. Its source parameters contain

abundant mechanical information of microfracture, which can be interpreted from the waveform. In the research related to rockburst generating mechanism (Ma et al. 2015; Srinivasan et al. 1999), source parameters have been extensively studied. There are five typical source parameters, namely microseismic energy (E_0), seismic moment (M_0), apparent volume (V_A), apparent stress (σ_A) and stress drop ($\Delta\sigma$). In this paper, they are employed to describe microseismic events. The detailed introduction is as follows. The feature information included in the database is shown in Table 2.

Table 2 Original monitoring data

As a measure of microfracture strength, microseismic energy (E_0) is used to assess microseismic activity (Boatwrignt and Fletcher 1984). It is equal to the sum of energy of monitored body waves (P-wave and S-wave), which can be calculated via Formula (1) or Formula (2).

$$E_0 = E_P + E_S \tag{1}$$

$$E_0 = 4\pi\rho v R^2 \frac{J_c}{F_c^2} \tag{2}$$

where, E_0 is microseismic energy; E_P is the energy of monitored P-wave; E_S is the energy of monitored S-wave; ρ is rock mass density; R denotes the distance from microseismic source to the sensor; v represents elastic wave velocity of rock mass; J_c indicates energy flux, which is equal to the integral of particle velocity in the frequency domain; F_c refers to empirical coefficient, which is 0.52 and 0.63 for P-wave and S-wave respectively.

Seismic moment (M_0) is proposed by Aki (1968), which is employed to evaluate seismic magnitude. It can be calculated by Formula (3).

$$M_0 = \frac{4\pi\rho v^3 R\delta_0}{F_c} \tag{3}$$

where, M_0 refers to seismic moment; δ_0 represents low-frequency amplitude of far-field displacement spectrum of body wave (P-wave or S-wave).

Apparent volume (V_A) (Mendecki 1993) and apparent stress (σ_A) (Senatorski 2007) are two crucial parameters that describe deformation distribution and stress state in the source zone. In particular, apparent volume (V_A) refers to rock mass volume with inelastic deformation while apparent stress (σ_A) indicates the degree of stress release. They can be obtained by Formula (4) and (5) respectively.

$$V_A = \frac{M_0}{2\sigma_A} \tag{4}$$

$$\sigma_A = \frac{\mu E_0}{M_0} \tag{5}$$

where, V_A and σ_A represent apparent volume and apparent stress, respectively; μ is rock mass stiffness.

The final source parameter is drop stress ($\Delta\sigma$), which can be calculated through Formula (6) (Mendecki 1997). It reflects the change of stress before and after source rupture.

$$\Delta\sigma = \frac{7M_0}{16r_0^3} \tag{6}$$

where, $\Delta\sigma$ denotes stress drop; r_0 represents source radius.

3 Construction of precursor tree

The procedure of constructing precursor tree is described in this section, which is illustrated in Fig. 5. The establishment of precursory microseismic sequences is the core

of building precursor tree. The primary steps are successively dimensionality reduction (section 3.1), clustering (section 3.2) and grouping (section 3.3). Microseismic events in the database will be assigned a clustering label, which lays the critical foundation for generating precursory microseismic sequences. Further, these precursory microseismic sequences are split into to parts by stratified sampling: 70% used for training precursor tree (section 3.4) and 30% used for engineering validation (section 3.5). Two kinds of precursor tree (pruning/no pruning) are constructed.

Fig. 5 Procedure of constructing precursor tree

3.1 Dimensionality reduction

Dimensionality reduction refers to mapping data in a high-dimensional space into a low-dimensional space (Ayesha et al. 2020). With respect to high-dimensional data, dimension disasters are prone to occur. Normally, dimensionality reduction is divided into two categories: linear and nonlinear dimensionality reduction. Principal component analysis (PCA) and linear discriminant analysis (LDA) are representative algorithms of the former. If there is a nonlinear relationship between features, linear dimensionality reduction cannot maintain the topology between original features. To solve this problem, nonlinear dimensionality reduction is put forward, such as local linear embedding (LLE) and t-distributed stochastic neighbor embedding (t-SNE). t-SNE is particularly suitable for data visualization (Hinton 2008). It can well retain both overall and local topology after reducing dimensionality.

In this study, t-SNE algorithm is implemented to eliminate redundant information contained in the database. When there is too much redundant information, it will greatly

increase calculation burden of subsequent clustering analysis. Information redundancy is positively correlated with correlation strength between features. Pearson correlation coefficient is an effective metric to measure the correlation, which is calculated through Formula (7) (Mu et al. 2018). The relationship between Pearson correlation coefficient and correlation strength is shown in Table 3 (Mohamed Salleh et al. 2015).

$$r_{yz} = \frac{\sum(y_i - \overline{y})\sum(z_i - \overline{z})}{\sqrt{\sum(y_i - \overline{y})^2}\sqrt{\sum(z_i - \overline{z})^2}} \tag{7}$$

where, r_{yz} is the Pearson correlation coefficient between features Y and Z; $\overline{y}$ and $\overline{z}$ denote the mean of Y and Z, respectively.

Table 3 Relationship between Pearson correlation coefficient and strength of correlation

Five features have been chosen in section 2.2. The results of correlation coefficient are shown in Fig. 6. It can be found that microseismic energy (E_0) is strongly correlated with seismic moment (M_0). In addition, there also exists moderate correlation between some features, such as apparent stress (σ_A) and drop stress ($\Delta\sigma$). After transforming the original data into four-dimensional one by t-SNE algorithm, the correlation between features has become weak or very weak. Obviously, information redundancy is greatly decreased compared with the previous. When further converted into three-dimensional one, the correlation between features is entirely very weak. At this time, features can be considered as independent. In other words, the redundant information in the database is eliminated.

Fig. 6 Results of Pearson correlation coefficient under different dimensions

Distance calculation is inevitable in the next clustering. Due to different magnitude of features, the feature with large magnitude occupies a dominant position in distance

calculation. It will weaken the contribution of other features. To avert this phenomenon, the features should be standardized to the same value range before clustering. Formula (8) is used, which normalizes the features to the range [-1, 1]. Table 4 presents the data after dimensionality reduction and its standardization results.

$$x^* = 2 \times \frac{x - x_{\min}}{x_{\max} - x_{\min}} - 1 \tag{8}$$

where, x^* is the normalized value; x is the initial value; $x_{\min}$ and $x_{\max}$ denote the minimum and maximum of feature X, respectively.

Table 4 Data after dimensionality reduction and its standardization results

3.2 Clustering

Clustering is also named unsupervised learning (Jain and Dubes 1988). It aims to partition the dataset into several clusters according to specific criteria, such as distance criterion. In a cluster, there is a certain similarity between data. However, high diversity exists for different clusters.

k-means algorithm is a kind of distance-based clustering method (Hartigan and Wong 1979). Here, distance denotes Minkowski distance, which is a common similarity measurement. Given two sets of data $Q = (q_1, q_2, q_3, \ldots, q_i)$ and $W = (w_1, w_2, w_3, \ldots, w_i)$, Minkowski distance between Q and W can be defined as Formula (9). When p=1 and p=2, it becomes Manhattan distance and Euclidean distance separately.

$$dist = \left(\sum \left(q_i - w_i \right)^{\wedge p} \right)^{\wedge (1/p)} \tag{9}$$

In view of the excellent computation efficiency, *k*-means algorithm is exploited to cluster microseismic events. The procedure of *k*-means algorithm is illustrated in Fig. 7, which is described as follows.

Fig. 7 Flow chart of *k*-means algorithm

(1) Determine the number of clusters k;

(2) Find the minimum and maximum of each feature in the database;

(3) Randomly generate k initial clustering centers via Formula (10);

$$\hat{x}_i = x_{\max} + \lambda_i \cdot \left(x_{\min} - x_{\max} \right) \tag{10}$$

where, $\hat{x}_i$ signifies the value of feature X for the i-th clustering center; $x_{\min}$ and $x_{\max}$ separately denote the minimum and maximum of feature X; λ_i is a random number in the range [0, 1], which needs to be regenerated for different features.

(4) Calculate the distance from the points in the database to clustering centers by Formula (9). Here, p is set to 2, i.e. Euclidean distance is adopted;

(5) Compare the distance and assign these points to the nearest cluster;

(6) Update clustering centers using Formula (11);

$$\hat{x}_i = \frac{1}{n_i} \sum_{G_i} x \tag{11}$$

Where, G_i represents the i-th cluster; x indicates the value of feature X for the points belonging to G_i; n_i refers to the number of points in G_i.

(7) Stop iteration when clustering centers are unchanged before and after updating; otherwise return to step (4) until the termination condition is satisfied.

The number of clusters k is required to be known beforehand when implementing k-means algorithm. In order to avoid human disturbance, some methods are proposed to determine k, such as elbow method (Purnima and Arvind 2014) and canopy algorithm (Mao 2012). When using elbow method, a cost function should be defined first. Then, the scatter plot between the number of clusters and cost function is made. The number

of clusters at the inflection point is the best *k*. However, *k*-means algorithm needs to be executed repeatedly when calculating cost function under different *k*. Canopy algorithm is capable to carry out clustering without specifying *k*. It is usually used together with *k*-means algorithm. Firstly, *k* is obtained using canopy algorithm to conduct clustering. Next, *k* is substituted into *k*-means algorithm for secondary clustering. In comparison with elbow method, canopy algorithm consumes less time. Hence, canopy algorithm is chosen to determine *k*. Its clustering procedure is shown in Fig. 8, which is described in detail as follows.

Fig. 8 Flow chart of canopy algorithm

(1) Randomly select a point *O* from the database as clustering center;

(2) Calculate the distance from other points in the database to *O*;

(3) Merge the points with the distance less than T_1 into *O* and generate a cluster;

(4) Remove *O* and the points with the distance less than T_2 from the database;

(5) Return to step (1) until the database is empty.

T_1 and T_2 ($T_1 > T_2$) are two critical parameters for canopy algorithm. Their geometric diagram is shown in Fig. 9. In the current cluster, the points with the distance less than T_2 cannot be selected as next clustering center. However, the points whose distance is greater than T_2 but less than T_1 could be selected as next clustering center. This indicates that overlap might exist between different clusters. The process to determine T_1 and T_2 is depicted below.

Fig. 9 Geometric diagram of the thresholds T_1 and T_2

(1) For all points in the database (referring to Table 3), calculate the distance with

each other;

(2) Calculate the mean and standard deviation of the distance, which are separately 1.2610 and 0.5316;

(3) Limit the value range for T_2 using three sigma principal in statistics, which is [-0.3338, 2.8558]. Since the negative T_2 is meaningless, the range becomes [0, 2.8558]. 100 values are produced at equal intervals from 0 to 2.8558;

(4) Set T_1 to be twice T_2, i.e. $T_1=2T_2$;

The frequency of cluster number generated by canopy algorithm is counted. The results are shown in Table 5. It can be found that the number of clusters is 5 invariably under 100 sets of T_1 and T_2. Therefore, k is set to 5 for k-means algorithm.

Table 5 Frequency of cluster number generated by canopy algorithm

Fig. 10 illustrates the clustering process of k-means algorithm. After 12 iterations, the clustering centers no longer change, which indicates that clustering is accomplished. Fig. 10(a), Fig. 10(b), Fig. 10(c) and Fig. 10(d) give the clustering results at the 1st, 4th, 8th and 12th iteration, respectively. By means of k-means algorithm, 1500 microseismic events are split into 5 clusters: cluster '1' includes 300 microseismic events; cluster '2' includes 180 microseismic events; cluster '3' includes 337 microseismic events; cluster '4' includes 152 microseismic events and cluster '5' includes 531 microseismic events. Their clustering centers are separately (-0.509, -0.184, 0.801), (-0.073, 0.133, -0.752), (-0.786, -0.422, -0.384), (0.094, 0.846, 0.020) and (0.608, -0.452, 0.185). The clustering results are summarized in Table 6. After that, cluster label l ($l=1,2,3,4,5$) is assigned to each microseismic event, as shown in Table 7.

Fig. 10 Clustering process of *k*-means algorithm

Table 6 Results of *k*-means algorithm

Table 7 Microseismic events after assigning labels

3.3 Grouping

Grouping is the last step to establish precursory microseismic sequences. For each rockburst case, five microseismic events are extracted from microseismic monitoring catalog to build the precursory microseismic sequence. According to time sequence of five microseismic events, their clustering labels are orderly allocated to variables *A*, *B*, *C*, *D* and *E*. Then, *A*, *B*, *C*, *D* and *E* are grouped together, denoted as $S=(A\ B\ C\ D\ E)$. This process is shown in Fig. 11. Corresponding to 300 rockburst cases, 300 precursory microseismic sequences are created.

Fig. 11 Grouping process

Taking the 1st rockburst case in the database as an example, a detailed introduction about the formation of its precursory microseismic sequence is demonstrated. In section 3.2, the clustering labels of five microseismic events have been obtained, which are 4, 5, 3, 5 and 5 respectively. According to the above rule, these labels are grouped together. The grouping result is $(4\ 5\ 3\ 5\ 5)$, which is the precursory microseismic sequence of the 1st rockburst case.

During the training process of precursor tree, precursory microseismic sequence $S=(A\ B\ C\ D\ E)$ is the input variable while rockburst intensity *I* is the output variable. The value of *I* is denoted as 'L' (low rockburst) or 'M' (moderate rockburst). A complete training sample is composed by both input variable and output variable, which is

defined as precursor pattern $P=\{(A\ B\ C\ D\ E)|I\}$. Since the 1st rockburst case belongs to low rockburst, its precursor pattern is $P=\{(4\,5\,3\,5\,5)|\mathrm{L}\}$. A total of 300 precursor patterns are shown in Table 8.

Table 8 Precursor patterns

3.4 Precursor tree modelling

300 precursory microseismic sequences formed in section 3.3 are divided into two parts by stratified sampling: 210 (70%) as the training set and 90 (30%) as the validation set. Since the database is imbalanced, stratified sampling makes the validation set more representative, which can maintain the proportional structure of the database.

Precursor tree is made up by a series of nodes, including root node, internal node and leaf node. Root node is used to store all precursory microseismic sequences, which is the starting point of precursor tree. For each precursor tree, there is only one root node. Internal node represents judgement condition, containing the precursory microseismic sequences that satisfies the judgement condition stored in the previous node. In other words, internal nodes partition a dataset into different subsets according to specific rules. Leaf node consists of the information of rockburst intensity, which is the terminal point of precursor tree. The path from root node to leaf node is a precursor pattern. When deleting leaf node, the path becomes a precursory microseismic sequence.

Except leaf nodes and the last-layer internal nodes, each node has five child nodes, which is equal to the number of clusters in section 3.2. The structure of precursor tree is illustrated in Fig. 12. The child nodes of root node, which are the first-layer internal nodes, are used to judge the attribute A in the precursory microseismic sequence. The

judgement conditions corresponding to these child nodes are A=1, A=2, A=3, A=4 and A=5, respectively. There are five judgement attributes in the precursory microseismic sequence. Thus, five-layers internal nodes exist. They separately indicate the judgement conditions of the attributes A, B, C, D and E. The precursor tree can express all precursor patterns in the database by a hierarchical structure.

Fig. 12 Structure of precursor tree

There are five subtrees in the precursor tree, denoted as PT_1, PT_2, PT_3, PT_4 and PT_5. PT_1 is highlighted in Fig. 12. Due to the limited space, only subtree PT_1 is shown after finishing the training process. Fig. 13 shows the subtree PT_1 of precursor tree trained by 210 precursory microseismic sequences.

Fig. 13 Schematic diagram of subtree PT_1 without pruning

The drawback of precursor tree under the above version is the risk of overfitting. The reason is that some precursory microseismic sequences are too personalized. With respect to them, each is just corresponding to one rockburst case. However, there is also such situation that multiple rockburst cases share a precursory microseismic sequence. Obviously, the latter is conducive to the generalization ability of precursor tree. Pruning is applied so as to prevent overfitting. The specific approach is to cut off the fifth-layer internal node for those precursory microseismic sequences with only one rockburst case. Namely, its attribute E is not considered. The subtree PT_1 of precursor tree after pruning is shown in Fig. 14.

Fig. 14 Schematic diagram of subtree PT_1 with pruning

3.5 Engineering validation

In this section, the prediction performance of the above two versions of precursor tree (no pruning/pruning) is analyzed based on the validation set, and a corresponding comparison is conducted. For ease of expression, the precursor tree without pruning is denoted as the first version and the one with pruning is denoted as the second version.

As for the first version, its prediction accuracy on the validation set is 86.7%. After pruning, the prediction accuracy becomes 98.9%, which is increased by 12.2%. This indicates that pruning can effectively prevent overfitting and enhance the generalization performance of precursor tree. Due to class imbalance of the training set, models trained with such data have a poor recognition ability for minority class. In this study, moderate rockburst is minority class compared with low rockburst. Thereby, it is quite necessary to analyze the prediction accuracy on each class. The prediction accuracy on moderate rockburst is 64.3% and 92.9% for the first version and the second version, respectively. The increase is 28.6%. Obviously, pruning can raise the recognition ability of precursor tree for minority class. Apart from moderate rockburst, the prediction accuracy on low rockburst is also improved slightly after pruning, which is separately 90.8% and 100% for the first version and the second version. The improvement is 9.2%. The relevant results are shown in Fig. 15. In a word, the precursor tree with pruning achieves better prediction performance and has a higher recognition ability for minority class. Hence, it is selected as the final precursor tree for rockburst warning in the engineering practice.

Fig. 15 Validation results of precursor trees

4 Field rockburst warning workflow

In section 3, the precursor tree with a high prediction performance has been built.

At the engineering site, the monitored microseismic data belongs to streaming data and is constantly updated from time to time. The means of forming precursory microseismic sequences is different from that for the non-streaming data described previously. These differences are mainly reflected in the phase of dimensionality reduction and clustering. This section aims to illustrate a complete workflow about how to apply the established precursor tree in engineering practice to achieve field rockburst warning. The flow chart is shown in Fig. 16.

Fig. 16 Field rockburst warning workflow

Two extra rockburst cases are collected from a diversion tunnel in northern China. Its project overview and geological condition have been briefly introduced in section 2.1. One of the cases is low rockburst, which occurred on April 25, 2018. And the another is moderate rockburst, which happened on April 28, 2018. The related field pictures are shown in Fig. 17. A total of 10 microseismic events are extracted from microseismic monitoring catalog, which is shown in Table 9. The former five microseismic events are used to establish the precursory microseismic sequence of the first case and the remain are used to build the precursory microseismic sequence of the second case. After generating precursory microseismic sequences, they are fed into the precursor tree to judge whether rockburst will occur.

Fig. 17 Field pictures related to two cases

Table 9 Monitoring data related to two cases

4.1 Dimensionality reduction using BP neural network

In section 3.1, t-SNE algorithm is implemented to reduce dimensionality for the

database. Since t-SNE algorithm is a nonlinear method, there is no explicit mapping relationship between the features before and after dimensionality reduction. When new data appears, it cannot accomplish the same dimensionality reduction as the previous database. The monitored microseismic data at the engineering site is streaming data and updated in real-time. It is essential about how to achieve the consistent dimensionality reduction for new microseismic data with the database employed for building precursor tree in engineering practice. In this respect, a regressor is trained, in which the original data in the database (referring to Table 2) is used as the input variable and the data after dimensionality reduction and standardization (referring to Table 4) is used as the output variable. After that, new microseismic data can be fed into this regressor to achieve the appropriate dimensionality reduction.

Artificial neural network (ANN) is an intelligent science imitating the structure and function of human brain (Kohonen 1988). It is capable to characterize highly nonlinear relationship. Theoretically, a three-layer neural network can approximate a function with arbitrary precision. The BP (back propagation) neural network is a typical one. It includes an input layer, a hidden layer and an output layer. In this study, it is used as the regressor for realizing dimensionality reduction. The number of neurons in the input layer and output layer is known, which is 5 and 3 respectively. However, the number of neurons in the hidden layer is unknown.

The number of neurons in the hidden layer has an great impact on the prediction performance of BP neural network. In order to determine this parameter, 10-fold cross-validation is utilized to analyze mean square error (MSE) of models with different

parameter values. The model with minimum MSE is selected. The principal of 10-fold cross-validation is illustrated in Fig. 18. The training data is randomly partitioned into 10 equal subsets. In each cross-validation, one subset is used for validation while the remaining nine subsets are used for training. The cross-validation is repeated 10 times to ensure that per subset is traversed as validation set. The final MSE is obtained by averaging.

Fig. 18 Sketch map of 10-fold cross-validation

Before training BP neural network, its initial weights and thresholds are usually randomly assigned. Herein, mind evolutionary algorithm (Liu et al. 2008) is adopted to optimize the initial weights and thresholds to avoid blindness. It is a novel evolutionary algorithm, which has been successfully applied in image processing, automatic control and other fields. Compared with genetic algorithm and particle swarm algorithm, it not only has global search ability but also high precision.

The relationship between MSE and the number of neurons in the hidden layer is shown in Fig. 19. It can be found that MSE is minimum when the number of neurons in the hidden layer is set to 6, which is 0.3930. The topology of BP neural network is shown in Fig. 20. Next, new microseismic data (referring to Table 9) is submitted into the trained network. The prediction results is shown in Table 10.

Fig. 19 Relationship between MSE and the number of neurons in the hidden layer

Fig. 20 Topology of BP neural network

Table 10 Prediction results with BP neural network

4.2 Assigning cluster label

In section 3.2, five clustering centers have been calculated by combined *k*-means algorithm and canopy algorithm. Thereby, only the distance between new microseismic data and clustering centers needs to be computed for assigning the cluster labels to new microseismic events. The distance between new microseismic data (referring to Table 10) and clustering centers (referring to Table 6) is shown in Table 11. Each microseismic event is allocated to the nearest cluster. In correspondence with 10 microseismic events, their cluster labels are successively '1', '5', '5', '5', '5', '5', '5', '5', '5' and '5'.

Table 11 Distance calculation and label assignment

4.3 Grouping

According to the rule described in section 3.3, precursory microseismic sequences can be established, which are separately (1 5 5 5 5) and (5 5 5 5 5) for two cases.

4.4 Rockburst warning

The precursory microseismic sequences are fed into the built precursory tree. The first case is predicted as low rockburst and the second case is predicted as moderate one. The warning results are consistent with the actual situation, which indicates that the precursor tree constructed in this study is successfully applied in engineering practice. Moreover, this warning workflow is understandable and practical.

5 Conclusions

In this research, a novel tree-based algorithm for real-time prediction of rockburst risk using microseismic monitoring technique was proposed. It mainly included model construction and field warning.

Through flexibly using a series of data mining algorithms such as t-SNE algorithm,

k-means algorithm and canopy algorithm, 300 precursory microseismic sequences were formed based on 1500 microseismic events. These sequences were partitioned into two parts by stratified sampling: 70% for training and 30% for validation. Two versions of precursor tree (no pruning/pruning) were trained on the training set. By comparing their prediction accuracy on the whole validation set and each class (low/moderate rockburst), the precursor tree with pruning realized better prediction performance than that without pruning. The increase was separately 12.2%, 9.2% and 28.6%. The prediction accuracy of the former reached 98.9% on the validation set. Since the database was imbalanced, the improvement on moderate rockburst further showed that pruning can improve the recognition ability for minority class.

In view of the above analysis, the precursor tree with pruning was selected to apply in engineering practice. Through two extra rockburst cases, a complete field rockburst warning workflow was illustrated. The related results indicated that this workflow was understandable and practical.

Acknowledgement

This research is supported by National Natural Science Foundation of China under Grant Nos. 41941018 and 41807250, China Postdoctoral Science Foundation Program under Grant Nos. 2019T120686, and National Key Basic Research Program of China under Grant Nos. 2015CB058102. These supports are gratefully acknowledged.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Adoko AC, Gokceoglu C, Wu L, Zuo QJ (2013) Knowledge-based and data-driven fuzzy modeling for rockburst prediction Int J Rock Mech Mining Sci 61:86-95

Afraei S, Shahriar K, Madani SH (2019) Developing intelligent classification models for rock burst prediction after recognizing significant predictor variables, section 1: Literature review and data preprocessing procedure Tunn Undergr Space Technol 83:324-353

Aki K (1968) Seismic displacements near a fault J Geophys Res 73:5359-5376

Ayesha S, Hanif MK, Talib R (2020) Overview and comparative study of dimensionality reduction techniques for high dimensional data Inform Fusion 59:44-58

Baltz R, Hucke A Rockburst prevention in the German coal industry. In: Proceedings of the 27th International Conference on Ground Control in Mining, 2008. pp 46-50

Barton N (2002) Some new Q-value correlations to assist in site characterisation and tunnel design Int J Rock Mech Mining Sci 39:185-216

Boatwrignt J, Fletcher JB (1984) The partition of radiated energy between P and S waves B Seismol Soc Am 74:361-376

Cai W, Dou L, Zhang M, Cao W, Shi J, Feng L (2018) A fuzzy comprehensive evaluation methodology for rock burst forecasting using microseismic monitoring Tunn Undergr Space Technol 80:232-245

Cai W, Dou LM, Si GY, Cao AY, He J, Liu S (2016) A principal component

analysis/fuzzy comprehensive evaluation model for coal burst liability assessment Int J Rock Mech Mining Sci 81:62-69

Cheng S, Li SC, Li LP, Shi SS, Zhou ZQ, Wang J (2018) Study on energy band characteristic of microseismic signals in water inrush channel J Geophys Eng 15:1826-1834

Deng P, Liu Q (2020) Influence of the softening stress path on crack development around underground excavations: Insights from 2D-FDEM modelling Computers and Geotechnics 117:103239

Deng P, Liu Q, Ma H, He F, Liu Q (2020) Time-dependent crack development processes around underground excavations Tunn Undergr Space Technol 103:103518

Dong LJ, Li XB, Peng K (2013) Prediction of rockburst classification using random forest Trans Nonferrous Metals Soc China 23:472-477

Duan Z, Yan X, Sun Q, Tan X, Dong C (2021) Effects of water content and salt content on electrical resistivity of loess Environ Earth Sci 80(14):1-15

Feng GL, Feng XT, Xiao YX, Yao ZB, Hu L, Niu WJ, Li T (2019a) Characteristic microseismicity during the development process of intermittent rockburst in a deep railway tunnel Int J Rock Mech Mining Sci 124:104135-104147

Feng GL, Xia GQ, Chen BR, Xiao YX, Zhou RC (2019b) A method for rockburst prediction in the deep tunnels of hydropower stations based on the monitored microseismicity and an optimized probabilistic neural network model Sustain 11:3212-3228

Feng XT, Liu JP, Chen BR, Xiao YX, Feng GL, Zhang FP (2017) Monitoring, warning,

and control of rockburst in deep metal mines Eng 3:538-545

Feng XT, Wang L (1994) Rockburst prediction based on neural networks Trans Nonferrous Metals Soc China 4:7-14

Goodman RE (1980) An introduction to rock mechanics. John Wiley and Sons, New York

Hartigan JA, Wong MA (1979) A k-means clustering algorithm Appl Stat 28:100-108

He M, Ren F, Liu D (2018) Rockburst mechanism research and its control Int J Mining Sci Technol 28:829-837

He M, Zhang Z, Zhu J, Li N, Li G, Chen Y (2021) Correlation between the rockburst proneness and friction characteristics of rock materials and a new method for rockburst proneness prediction: Field demonstration J Petrol Sci Eng 108997

Hinton GE (2008) Visualizing high-dimensional data using t-SNE J Mach Learn Res 9:2579-2605

Hoek E, Brown ET (1980) Underground excavations in rock. The Institution of Mining and Metallurgy, London

Jain A, Dubes R (1988) Algorithms for clustering data

Jia Y, Lu Q, Shang Y (2013) Rockburst prediction using particle swarm optimization algorithm and general regression neural network Chinese J Rock Mech Eng 32:343-348

Keneti A, Sainsbury B-A (2018) Review of published rockburst events and their contributing factors Eng Geol 246:361-373

Kidybiński A (1981) Bursting liability indices of coal Int J Rock Mech Mining Sci

18:295-304

Kohonen T (1988) An introduction to neural computing Neur Net 1:3-16

Li N, Jimenez R (2017) A logistic regression classifier for long-term probabilistic prediction of rock burst hazard Natural Hazards 90:197-215

Liang W, Sari A, Zhao G, McKinnon SD, Wu H (2020) Short-term rockburst risk prediction using ensemble learning methods Natural Hazards 104:1923-1946

Liang W, Zhao G, Wu H, Dai B (2019) Risk assessment of rockburst via an extended MABAC method under fuzzy environment Tunn Undergr Space Technol 83:533-544

Liu F, Ma TH, Tang CA, Chen F (2018) Prediction of rockburst in tunnels at the Jinping II hydropower station using microseismic monitoring technique Tunn Undergr Space Technol 81:480-493

Liu JX, Wang F, Xie KM (2008) Design of broadband impedance transformer based on improved mind evolutionary algorithm

Liu QS, Wu J, Zhang XP, Tang LX, Bi C, Li WW, Xu JL (2020) Microseismic monitoring to characterize structure-type rockbursts: A case study of a tbm-excavated tunnel Rock Mech Rock Eng

Luis ReS, Tiago M, Rita LeS, Joaquim T (2017) The use of data mining techniques in rockburst risk assessment Eng 3:552-558

Ma D, Duan H, Li X, Li Z, Zhou Z, Li T (2019a) Effects of seepage-induced erosion on nonlinear hydraulic properties of broken red sandstones Tunn Undergr Space Technol 91:102993

Ma D, Duan H, Liu J, Li X, Zhou Z (2019b) The role of gangue on the mitigation of mining-induced hazards and environmental pollution: an experimental investigation Sci Total Environ 664:436-448

Ma D, Zhang J, Duan H, Huang Y, Li M, Sun Q, Zhou N (2021) Reutilization of gangue wastes in underground backfilling mining: overburden aquifer protection Chemosphere 264:128400

Ma TH, Tang CA, Tang LX, Zhang WD, Wang L (2015) Rockburst characteristics and microseismic monitoring of deep-buried tunnels for Jinping II hydropower station Tunn Undergr Space Technol 49:345-368

Ma TH, Tang CA, Tang SB, Kuang L, Yu Q, Kong DQ, Zhu X (2018) Rockburst mechanism and prediction based on microseismic monitoring Int J Rock Mech Mining Sci 110:177-188

Mao DH (2012) Improved canopy-kmeans algorithm based on MapReduce Comput Eng Appl 10:1-8

Mendecki AJ (1993) Real time quantitative seismology in mines. Paper presented at the Proceedings of 3rd International Symposium on Rockbursts and Seismicity in Mines, Kingston, Ontario, Canada,

Mendecki AJ (1997) Seismic monitoring in mines. Chapman & Hall, London

Mohamed Salleh FH, Arif SM, Zainudin S, Firdaus-Raih M (2015) Reconstructing gene regulatory networks from knock-out data using Gaussian Noise Model and Pearson Correlation Coefficient Comput Biol Chem 59:3-14

Mu Y, Liu X, Wang L (2018) A Pearson's correlation coefficient based decision tree

and its parallel implementation Inform Sci 435:40-58

Oparin V, Vostrikov V, Usoltseva O, Mulev S, Rodionova E (2017) Assessment of rockburst hazard based on the data of mine seismology Proc Eng 191:795-801

Pu YY, Apel DB, Lingga B (2018) Rockburst prediction in kimberlite using decision tree with incomplete data J Sustain Mining 17:158-165

Pu YY, Apel DB, Liu V, Mitri H (2019a) Machine learning methods for rockburst prediction-state-of-the-art review Int J Mining Sci Technol 29:565-570

Pu YY, Apel DB, Xu HW (2019b) Rockburst prediction in kimberlite with unsupervised learning method and support vector classifier Tunn Undergr Space Technol 90:12-18

Purnima B, Arvind K (2014) EBK-means: A clustering technique based on elbow method and k-means in WSN Int J Comput Appl 105:17-24

Qiu L, Liu Z, Wang E, Li B (2020) Early-warning of rock burst in coal mine by low-frequency electromagnetic radiation Eng Geol 279:105755

Russenes B (1974) Analyses of rockburst in tunnels in valley sides Norwegian Institute of Technology, Trondheim

Senatorski P (2007) Apparent stress scaling and statistical trends Phys Earth Planet In 160:230-244

Sirait B, Wattimena RK, Widodo NP (2013) Rockburst prediction of a cut and fill mine by using energy balance and induced stress Proc Earth Planet Sci 6:426-434

Srinivasan C, Arora SK, Benady S (1999) Precursory monitoring of impending rockbursts in Kolar gold mines from microseismic emissions at deeper levels Int J

Rock Mech Mining Sci 36:941-948

Sun J, Wang LG, Zhang HL, Shen YF (2009) Application of fuzzy neural network in predicting the risk of rock burst Proc Earth Planet Sci 1:536-543

Tang CA, Wang JM, Zhang JJ (2010) Preliminary engineering application of microseismic monitoring technique to rockburst prediction in tunneling of Jinping II project J Rock Mech Geotec Eng 2:193-208

The National Standards Compilation Group of People's Republic of China (2009) GB50487-2008 Code for Engineering Geological Investigation of Water Resources and Hydropower. China Planning Press, Beijing

Wang L, Guo F, Wang S (2020) Prediction model of the collapse of bank slope under the erosion effect of wind-induced wave in the Three Gorges Reservoir Area, China Environ Earth Sci 79(18):1-17

Webber SJ (1996) Rockburst risk assessment on south african gold mines: An expert system approach. Paper presented at the ISRM International Symposium, Turin, Italy

Xu C, Liu XL, Wang EZ, Zheng YL, Wang SJ (2018) Rockburst prediction and classification based on the ideal-point method of information theory Tunn Undergr Space Technol 81:382-390

Xu J, Jiang J, Xu N, Liu Q, Gao Y (2017) A new energy index for evaluating the tendency of rockburst and its engineering application Eng Geol 230:46-54

Xu NW, Tang CA, Li LC, Zhou Z, Sha C, Liang ZZ, Yang JY (2011) Microseismic monitoring and stability analysis of the left bank slope in Jinping first stage

hydropower station in southwestern China Int J Rock Mech Mining Sci 48:950-963

Xue Y, Bai C, Kong F, Qiu D, Li L, Su M, Zhao Y (2020) A two-step comprehensive evaluation model for rockburst prediction based on multiple empirical criteria Eng Geol 268:105515

Zhang L, Zhang X, Wu J, Zhao D, Fu H (2020) Rockburst prediction model based on comprehensive weight and extension methods and its engineering application B Eng Geol Environ 79:4891-4903

Zhang M, Liu S, Shimada H (2018) Regional hazard prediction of rock bursts using microseismic energy attenuation tomography in deep mining Natural Hazards 93:1359-1378

Zhao ZG, Tan YL (2010) Microseismic monitoring data fusion algorithm and coal and gas outbursts prediction J Meas Sci Instrum 1:315-316

Zhou J, Koopialipoor M, Li E, Armaghani DJ (2020) Prediction of rockburst risk in underground projects developing a neuro-bee intelligent system B Eng Geol Environ

Zhou J, Li X, Mitri HS (2018) Evaluation method of rockburst: State-of-the-art literature review Tunn Undergr Space Technol 81:632-659

Zhou J, Li XB, Mitri HS (2016) Classification of rockburst in underground projects: Comparison of ten supervised learning methods J Comput Civil Eng 30:04016003

Zhou J, Li XB, Shi XZ (2012) Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines Safety Sci

50:629-644

A list of figure titles:

Fig. 1 Geological profile along the diversion tunnel

Fig. 2 Components of ESG microseismic monitoring system

Fig. 3 Top view of the layout of ESG microseismic monitoring system

Fig. 4 Proportion of rockburst with different intensity

Fig. 5 Procedure of constructing precursor tree

Fig. 6 Results of Pearson correlation coefficient under different dimensions

Fig. 7 Flow chart of *k*-means algorithm

Fig. 8 Flow chart of canopy algorithm

Fig. 9 Geometric diagram of the thresholds T_1 and T_2

Fig. 10 Clustering process of *k*-means algorithm

Fig. 11 Grouping process

Fig. 12 Structure of precursor tree

Fig. 13 Schematic diagram of subtree PT_1 without pruning

Note: L/M represent rockburst intensity, which are separately low rockburst and moderate rockburst; the figure in bracket is the number of rockburst cases corresponding to the precursory microseismic sequence.

Fig. 14 Schematic diagram of subtree PT_1 with pruning

Note: L/M represent rockburst intensity, which are separately low rockburst and moderate rockburst; the figure in bracket is the number of rockburst cases

corresponding to the precursory microseismic sequence.

Fig. 15 Validation results of precursor trees

Fig. 16 Field rockburst warning workflow

Fig. 17 Field pictures related to two cases

Fig. 18 Sketch map of 10-fold cross-validation

Fig. 19 Relationship between MSE and the number of neurons in the hidden layer

Fig. 20 Topology of BP neural network

Figures

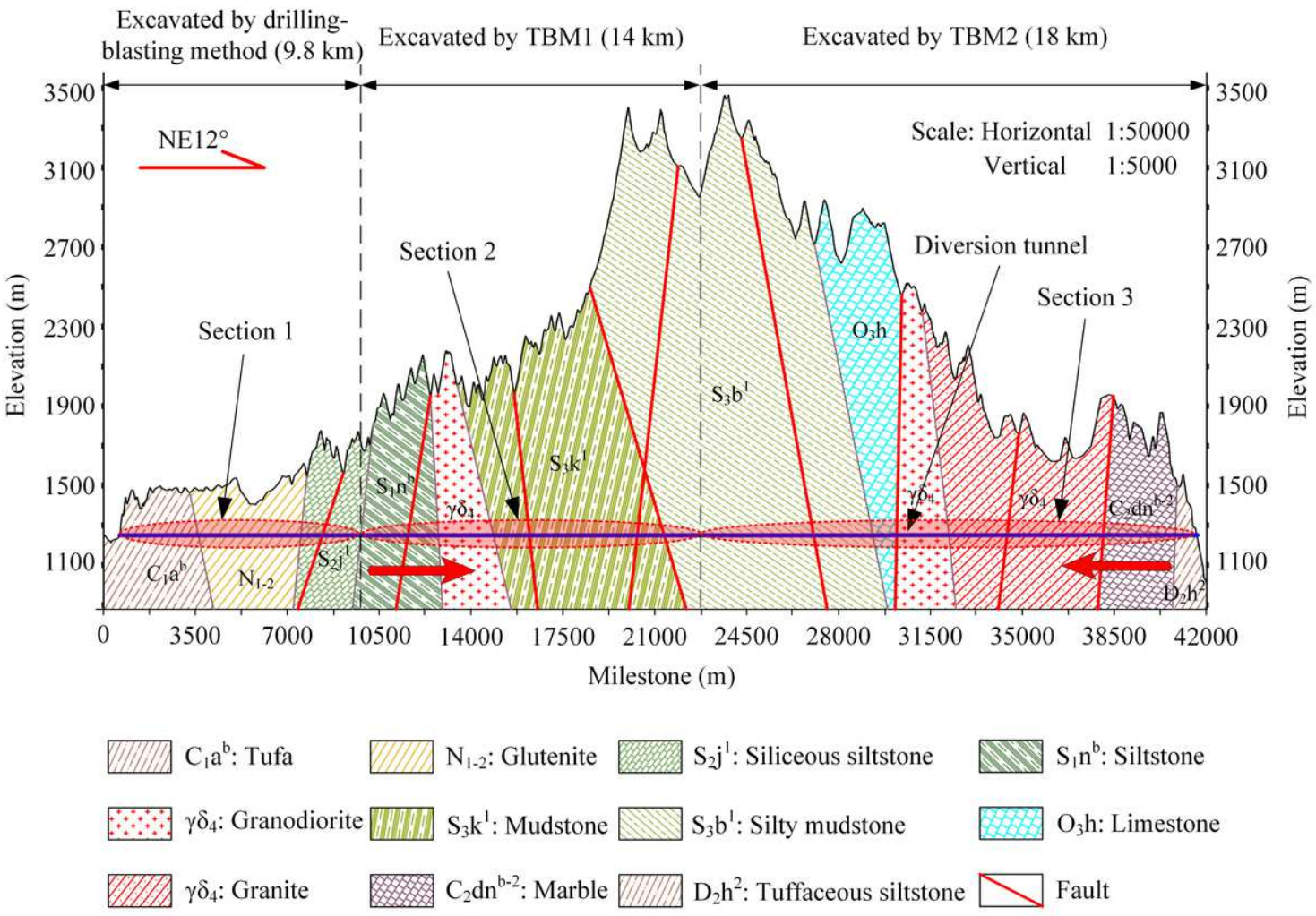

Figure 1

Geological profile along the diversion tunnel

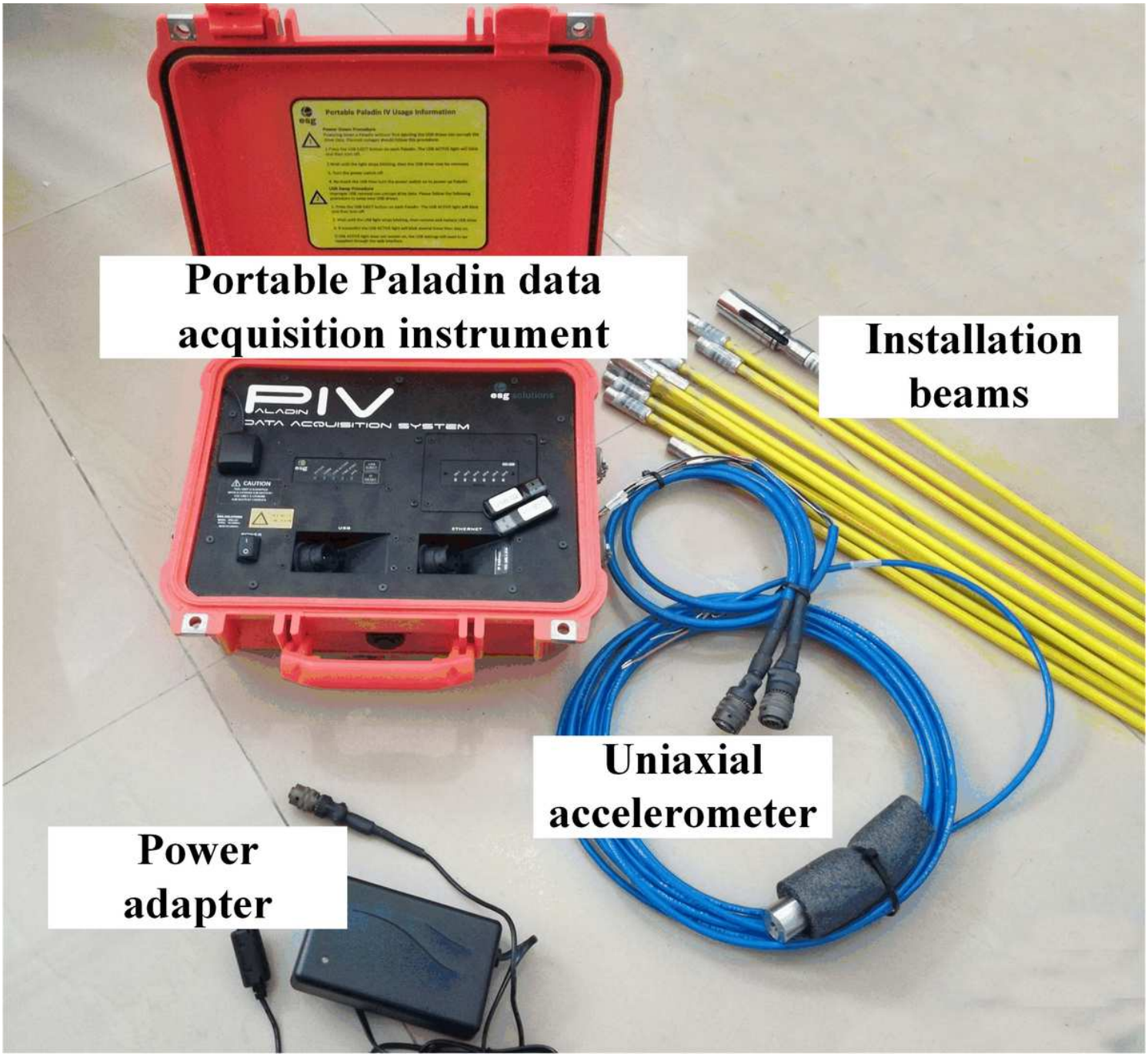

Figure 2

Components of ESG microseismic monitoring system

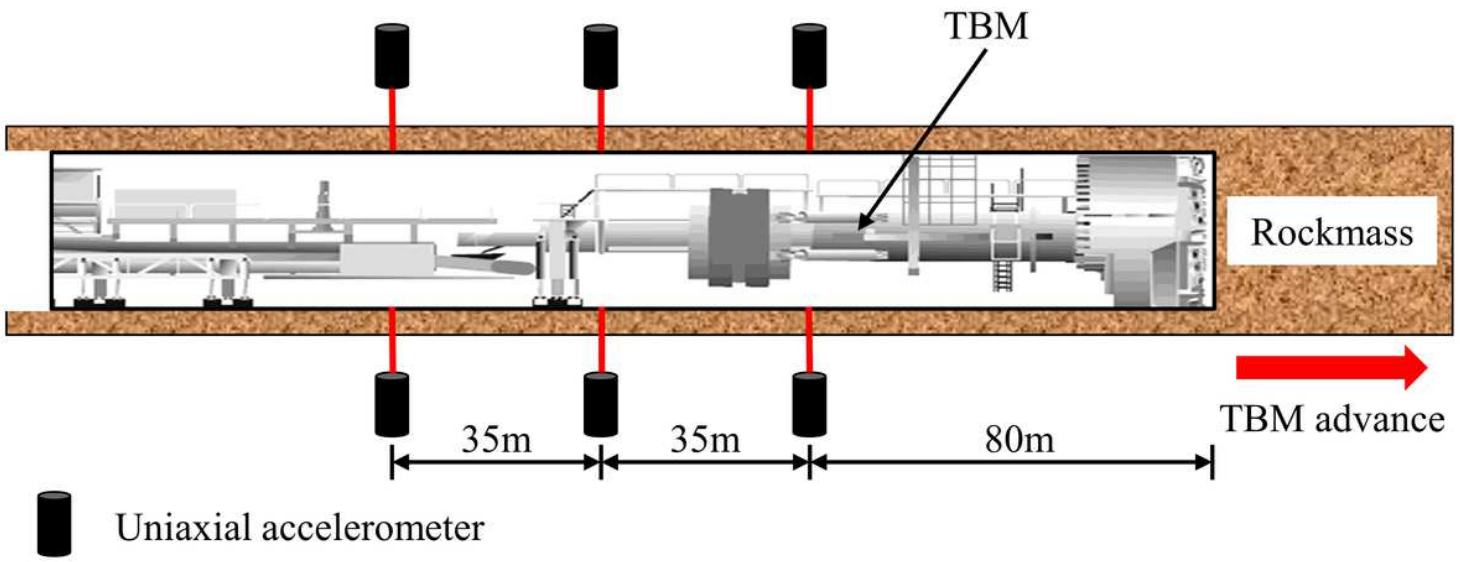

Figure 3

Top view of the layout of ESG microseismic monitoring system

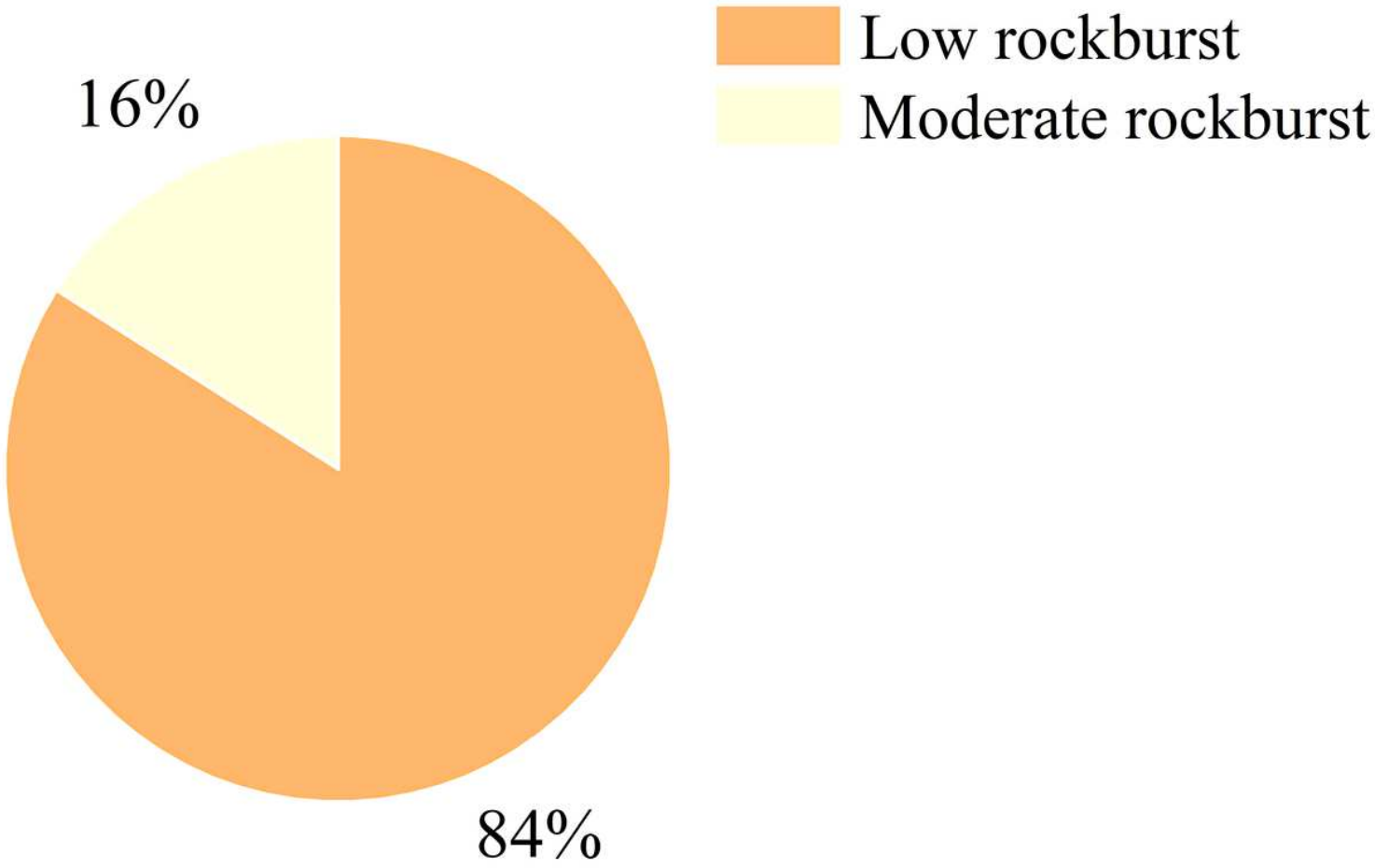

Figure 4

Proportion of rockburst with different intensity

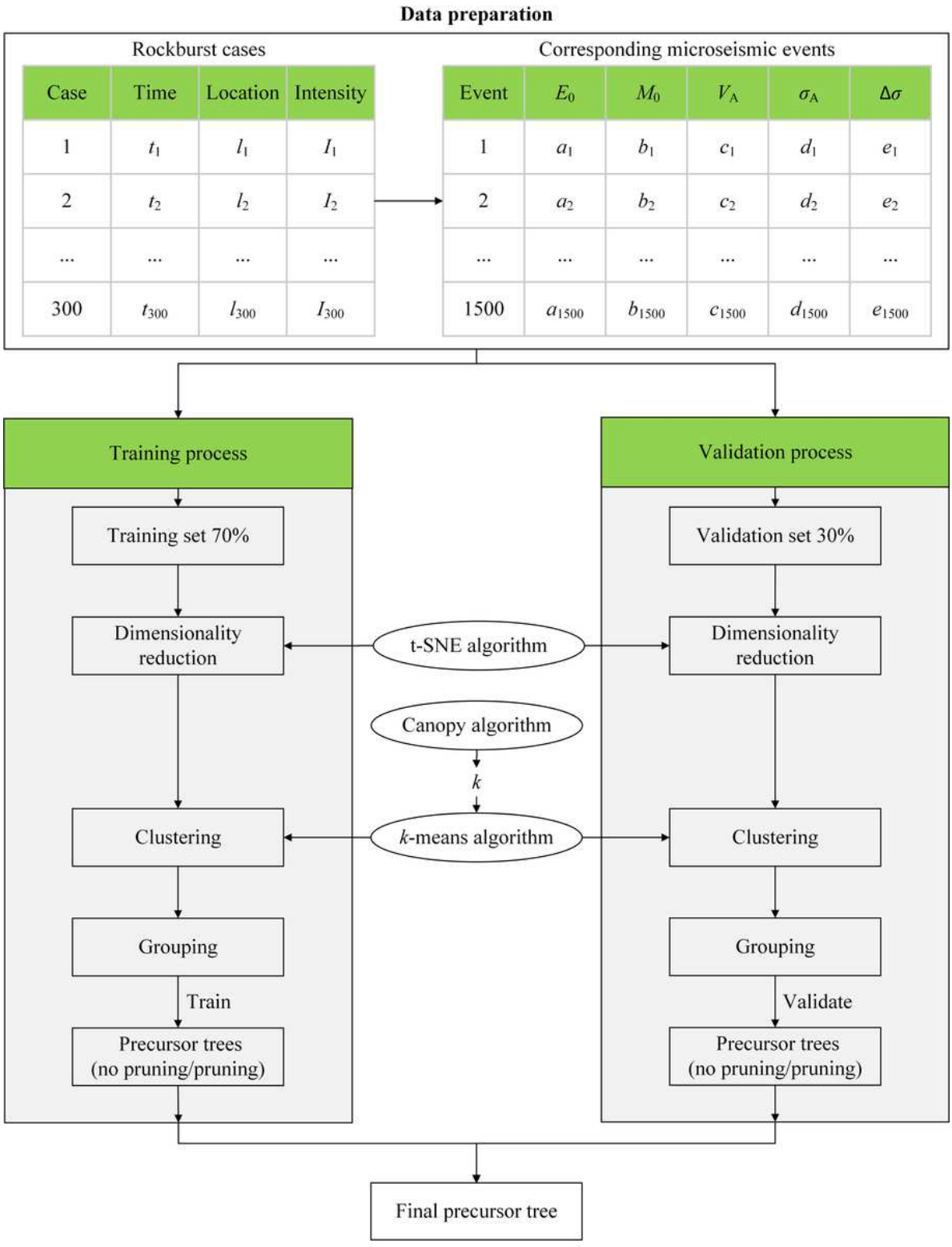

Figure 5

Procedure of constructing precursor tree

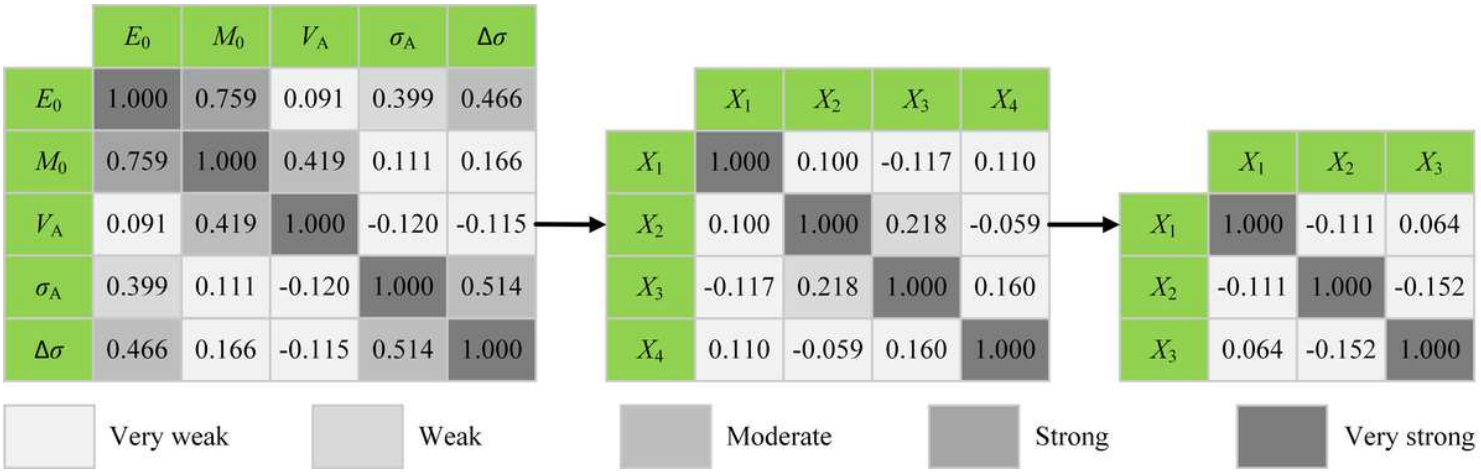

Figure 6

Results of Pearson correlation coefficient under different dimensions

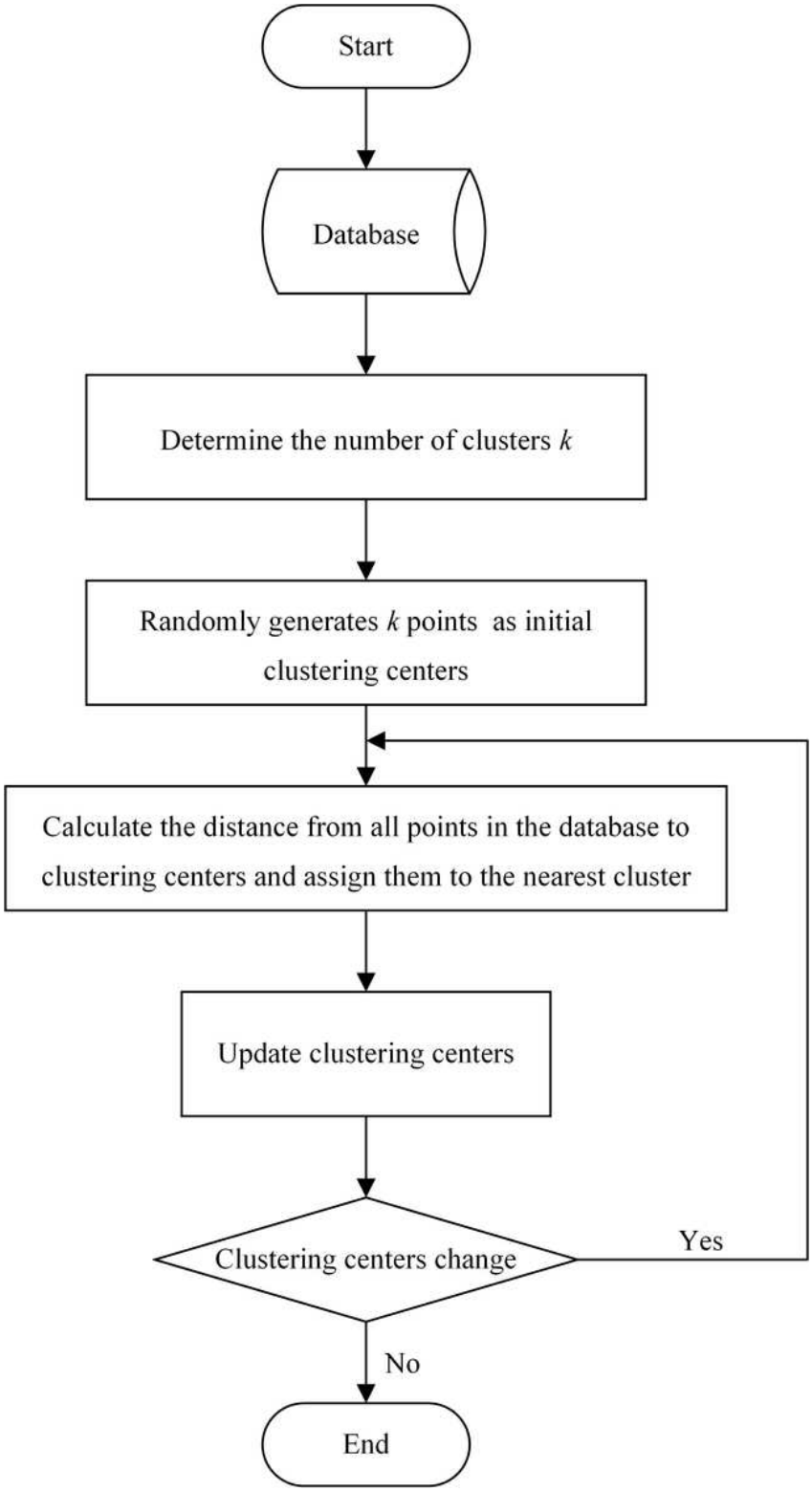

Figure 7

Flow chart of k-means algorithm

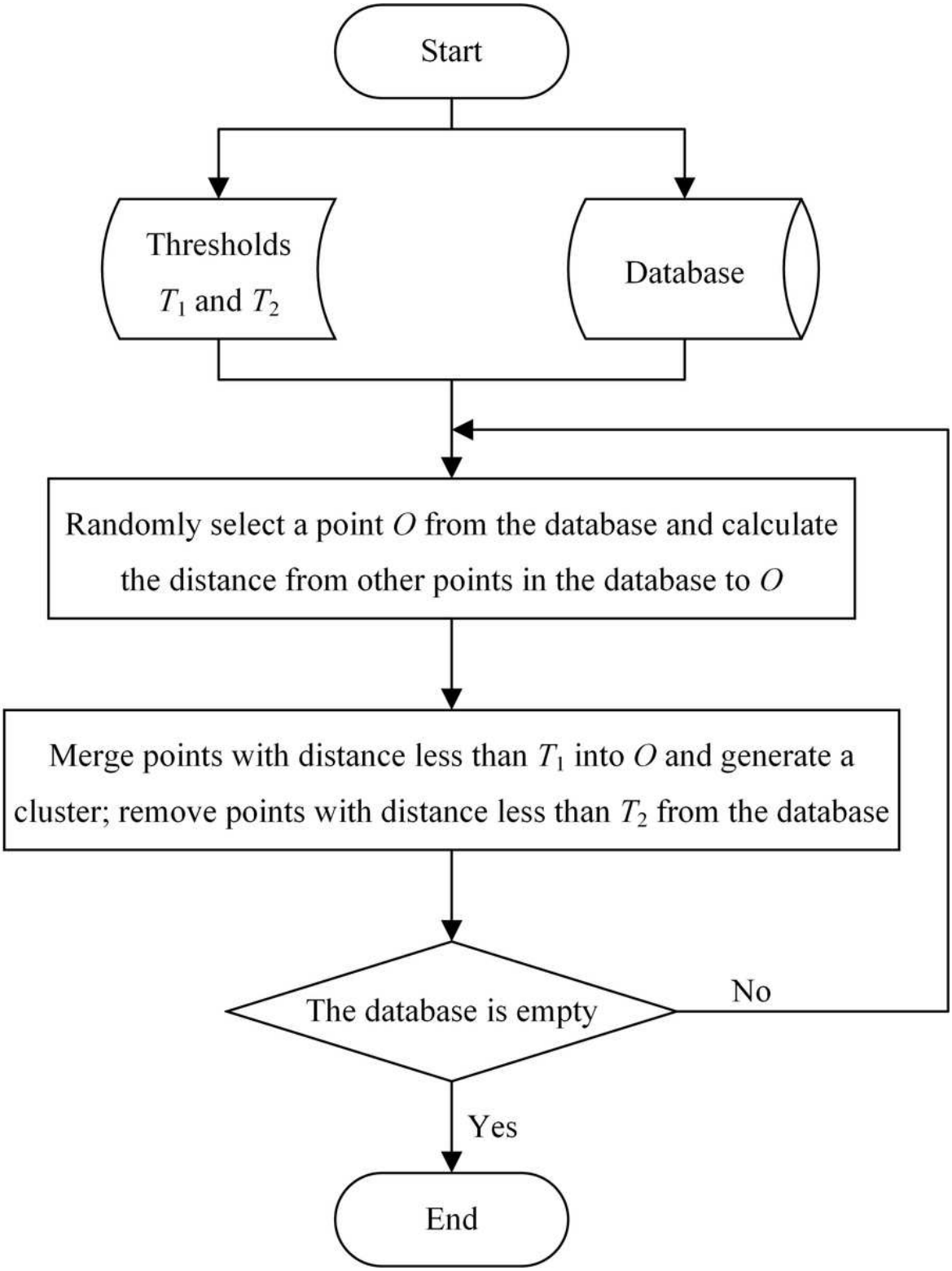

Figure 8

Flow chart of canopy algorithm

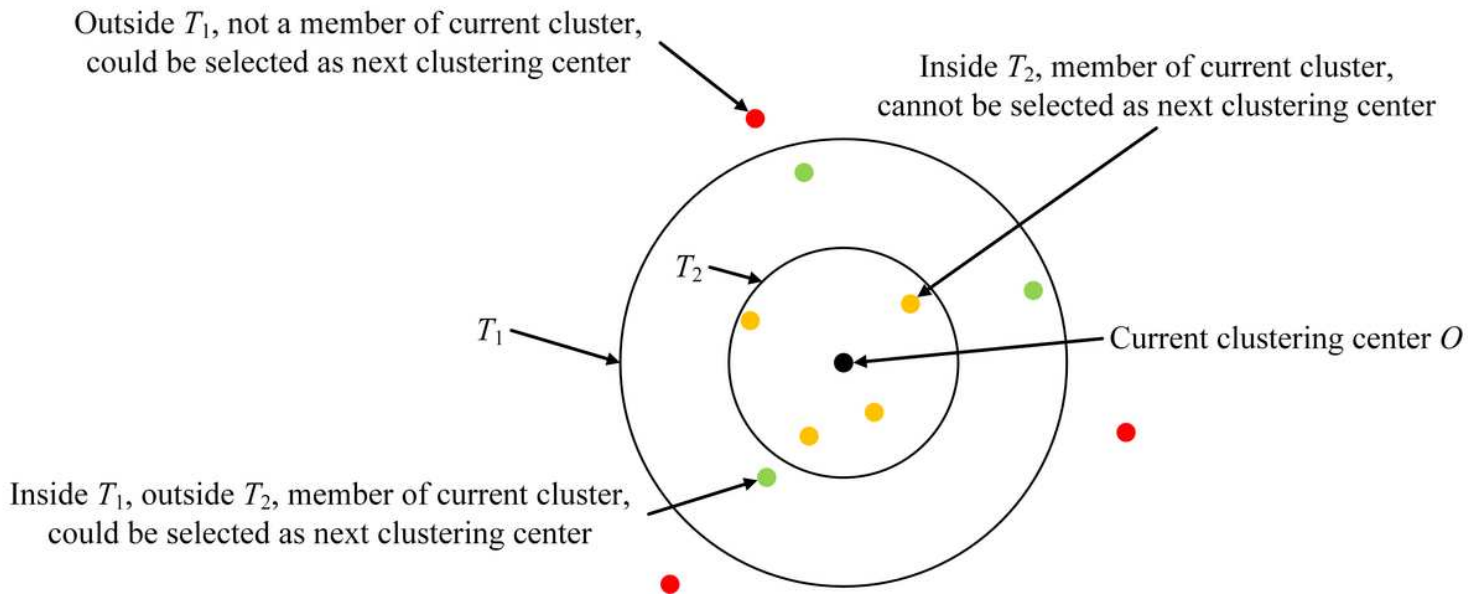

Figure 9

Geometric diagram of the thresholds T1 and T2

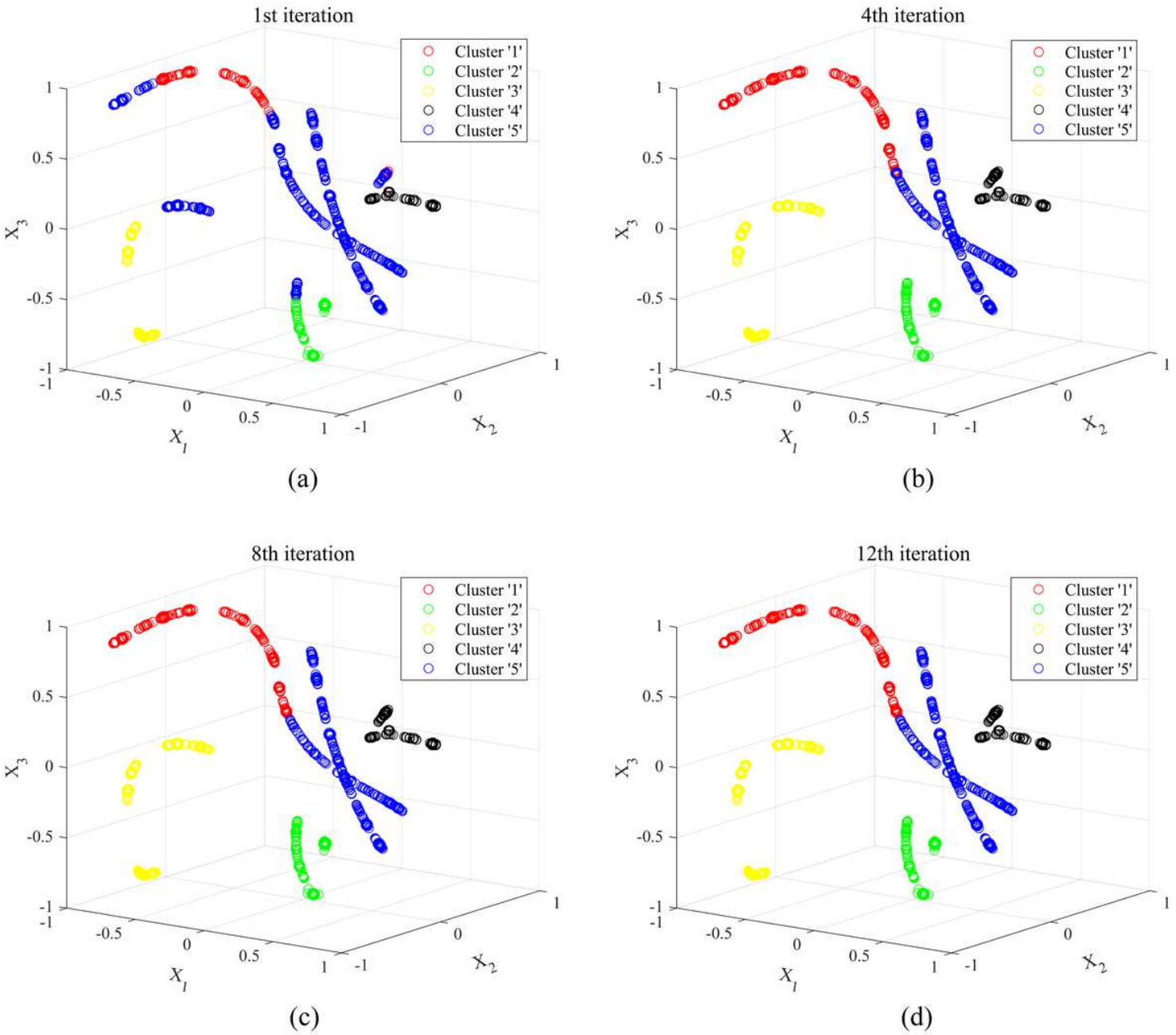

Figure 10

Clustering process of k-means algorithm

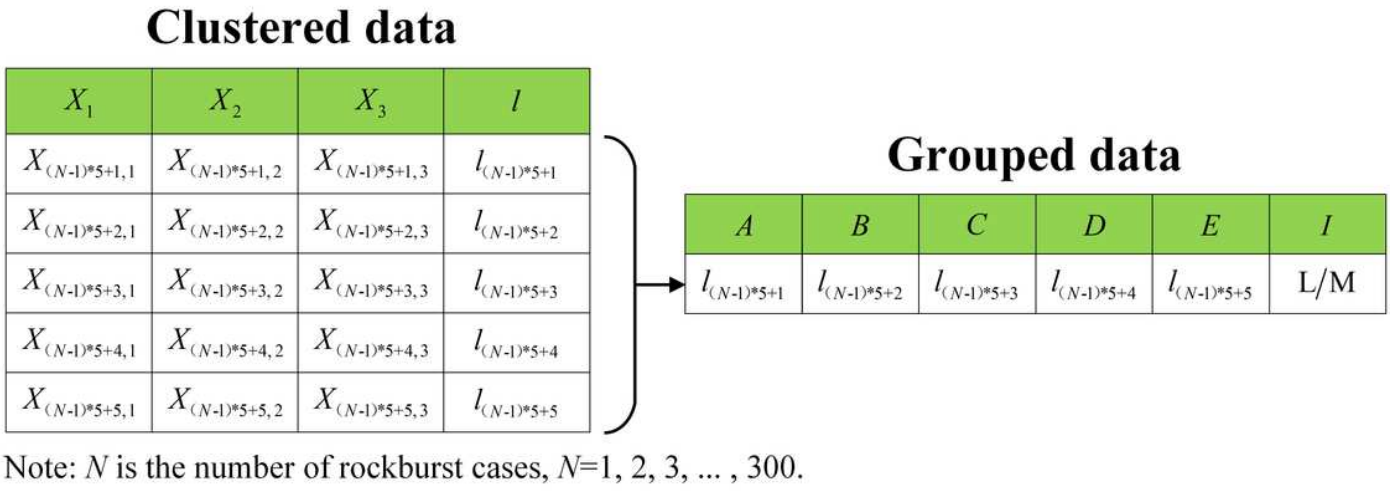

Figure 11

Grouping process

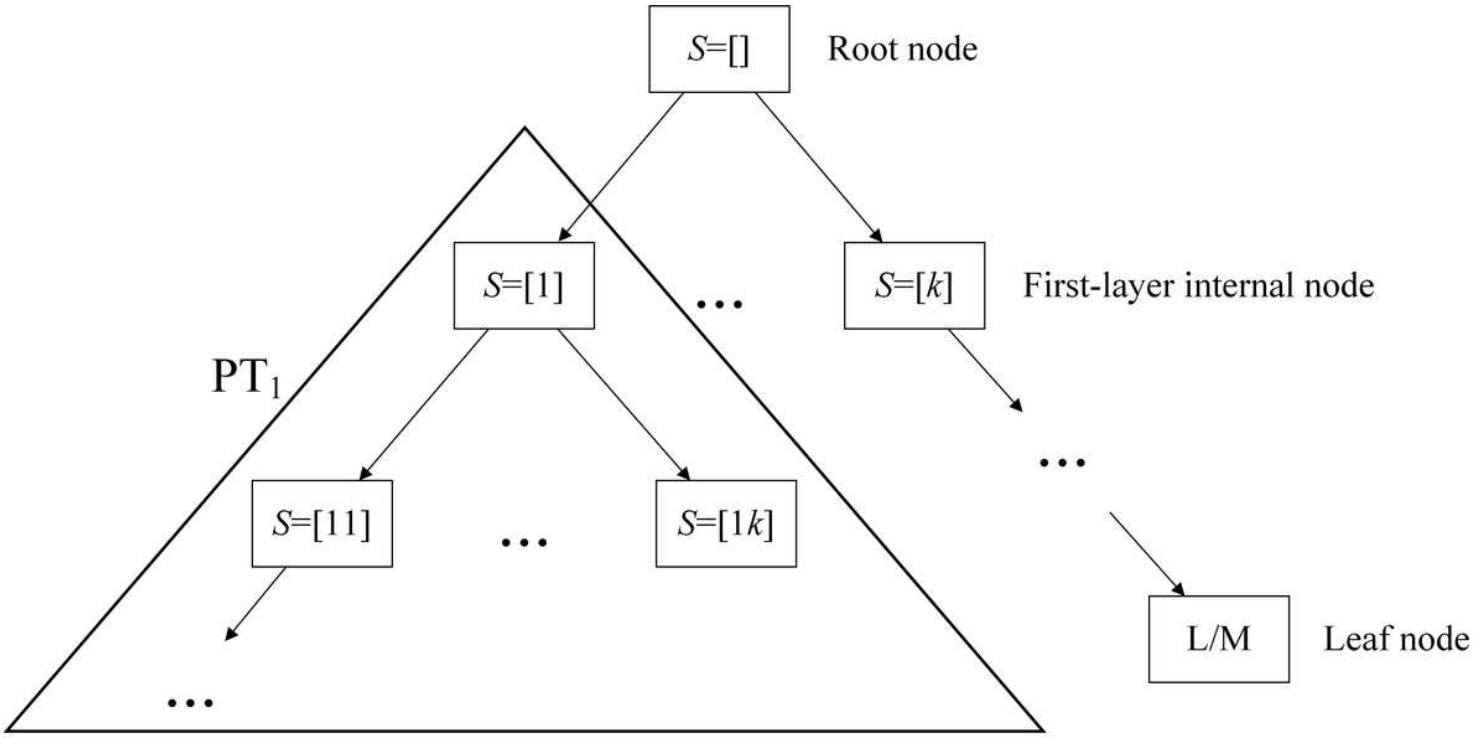

Figure 12

Structure of precursor tree

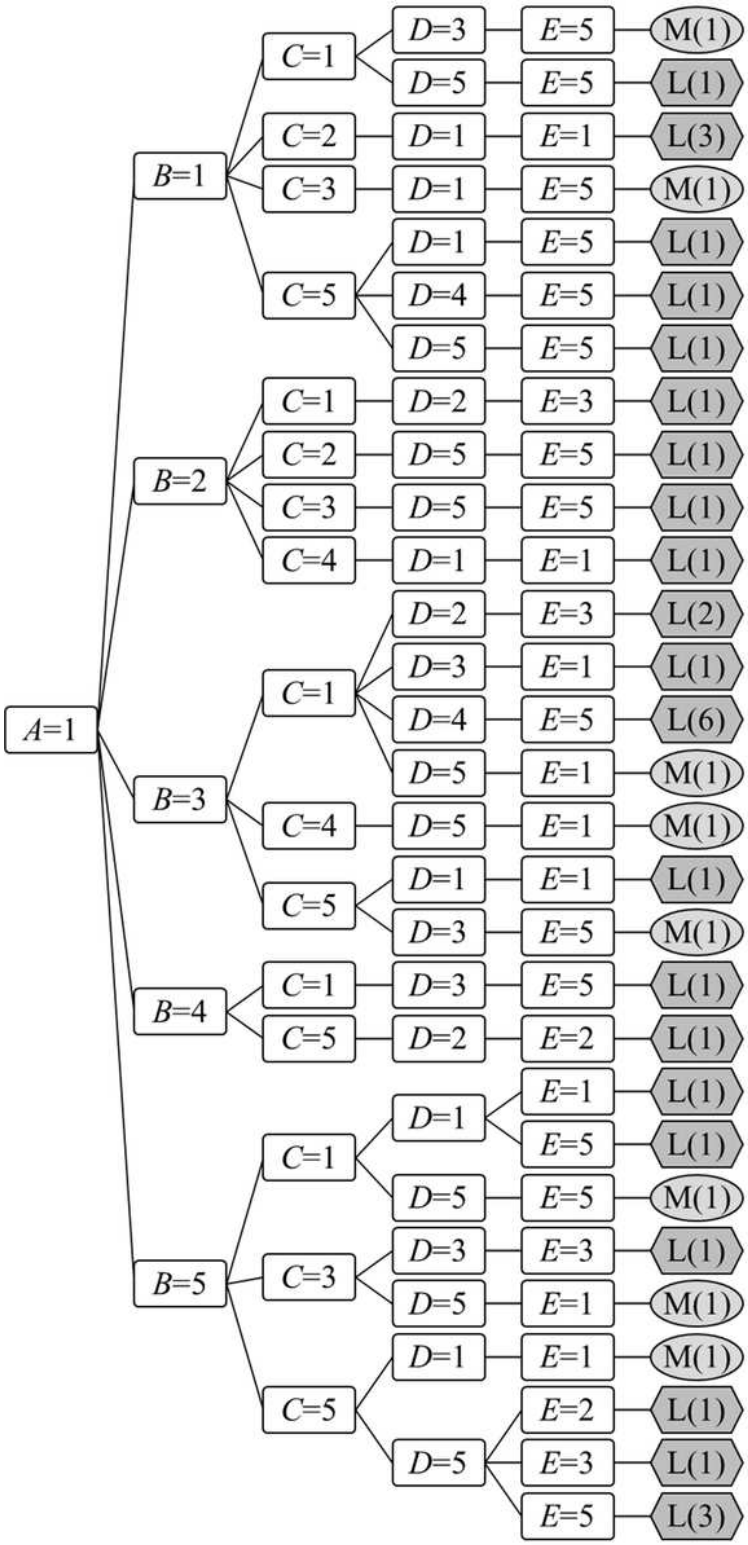

Figure 13

Schematic diagram of subtree PT1 without pruning. Note: L/M represent rockburst intensity, which are separately low rockburst and moderate rockburst; the figure in bracket is the number of rockburst cases corresponding to the precursory microseismic sequence.

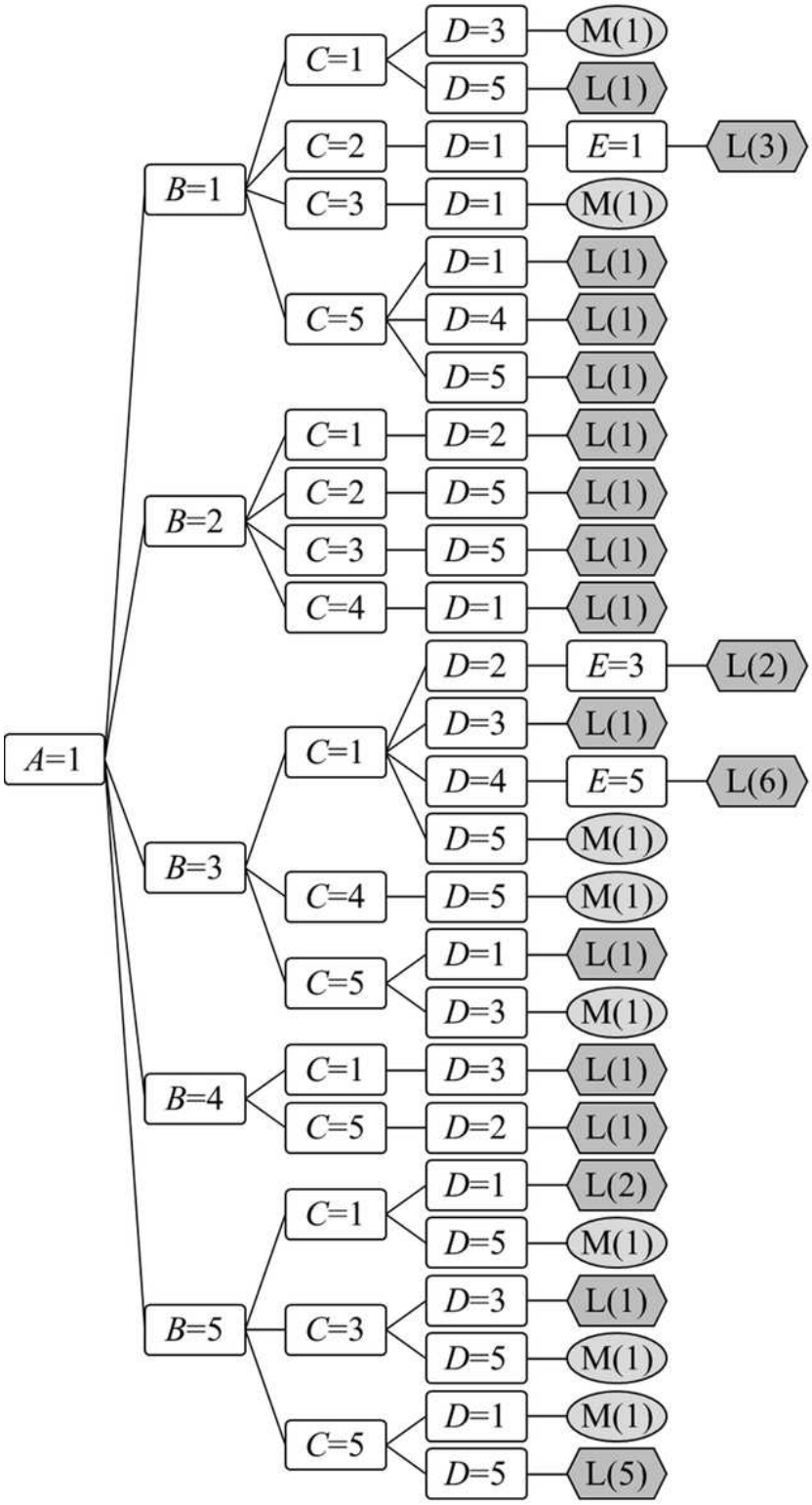

Figure 14

Schematic diagram of subtree PT1 with pruning. Note: L/M represent rockburst intensity, which are separately low rockburst and moderate rockburst; the figure in bracket is the number of rockburst cases corresponding to the precursory microseismic sequence.

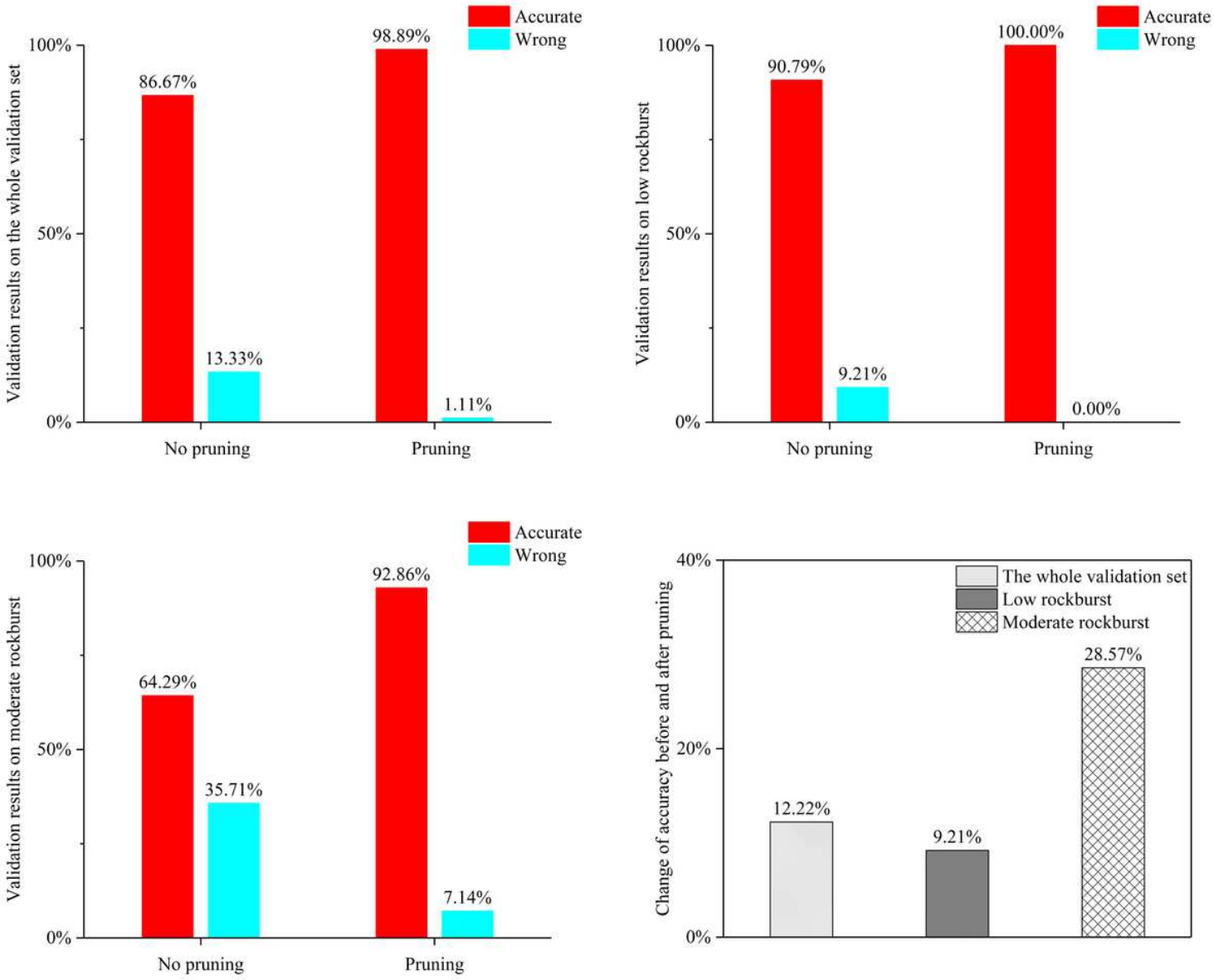

Figure 15

Validation results of precursor trees

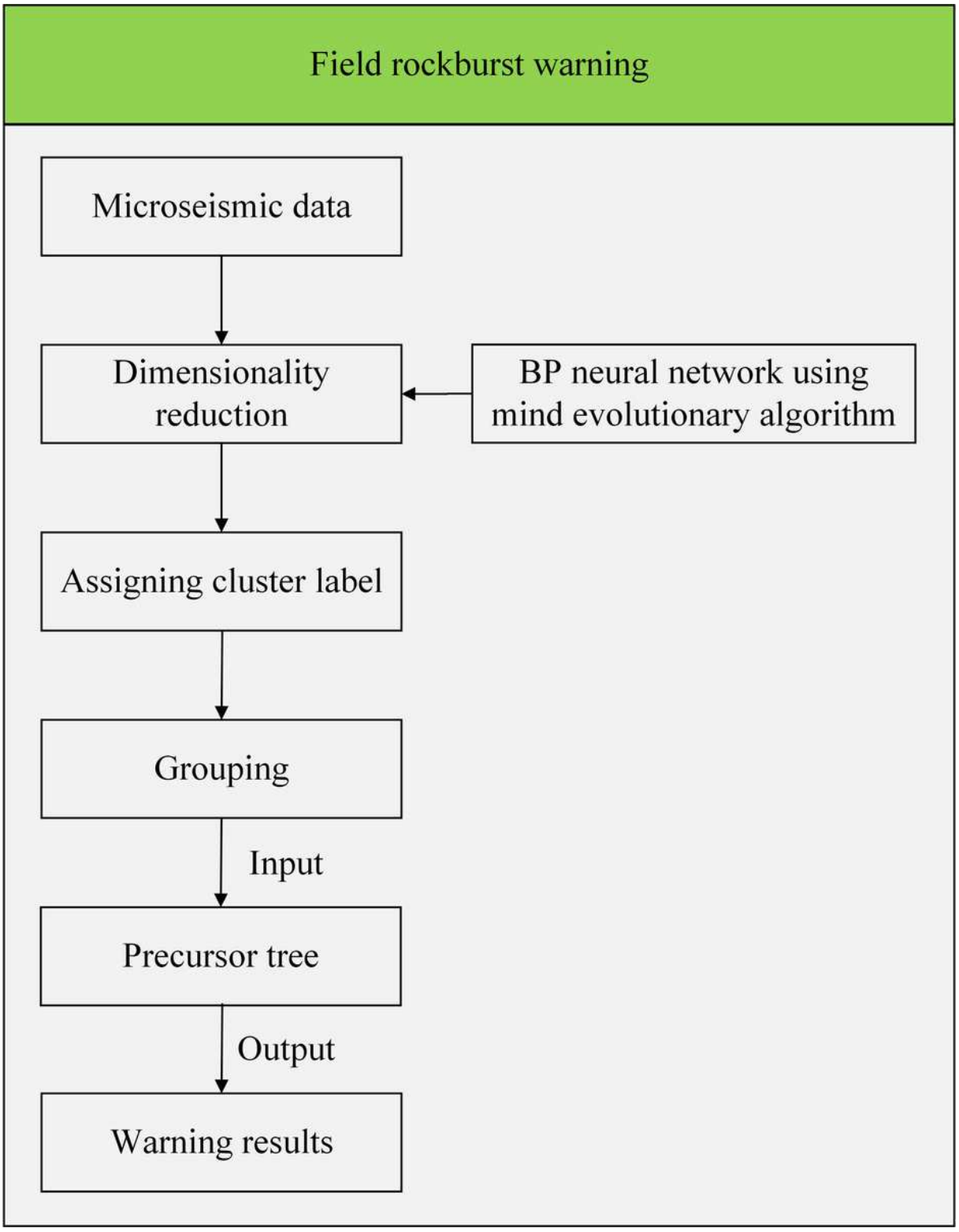

Figure 16

Field rockburst warning workflow

Figure 17

Field pictures related to two cases

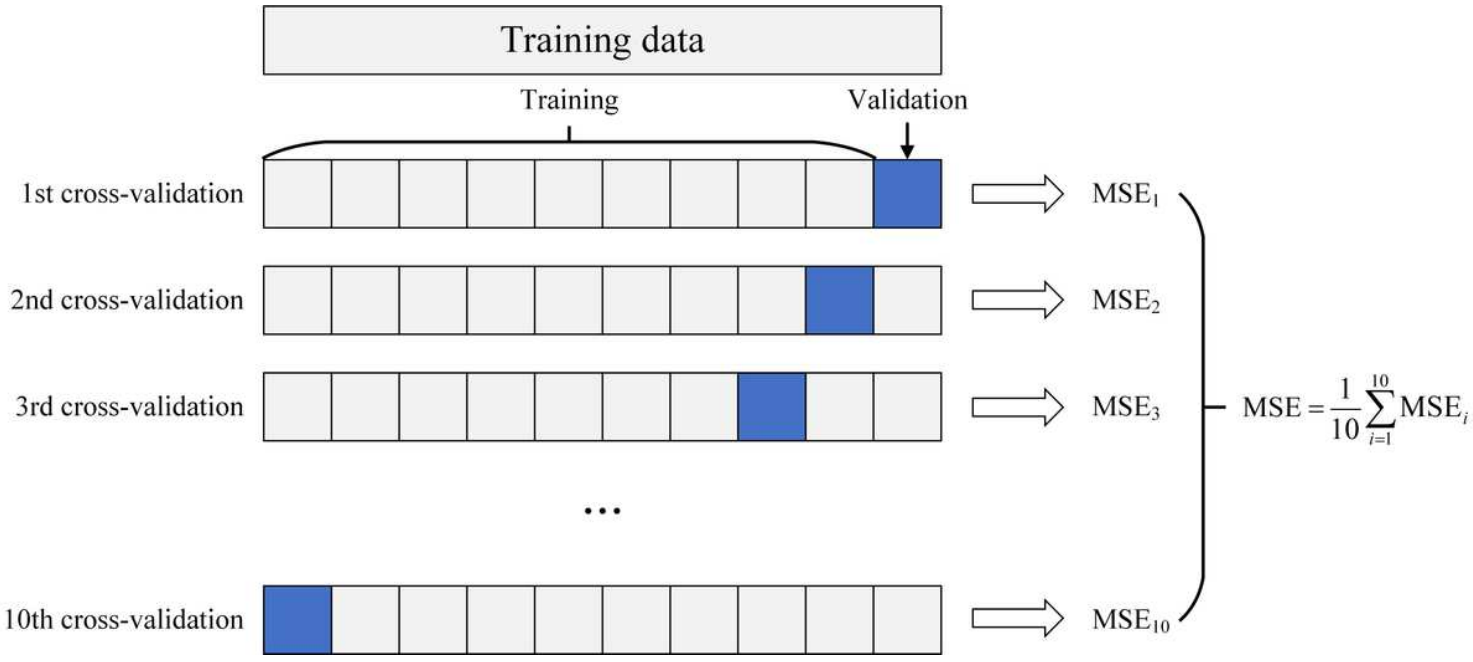

Figure 18

Sketch map of 10-fold cross-validation

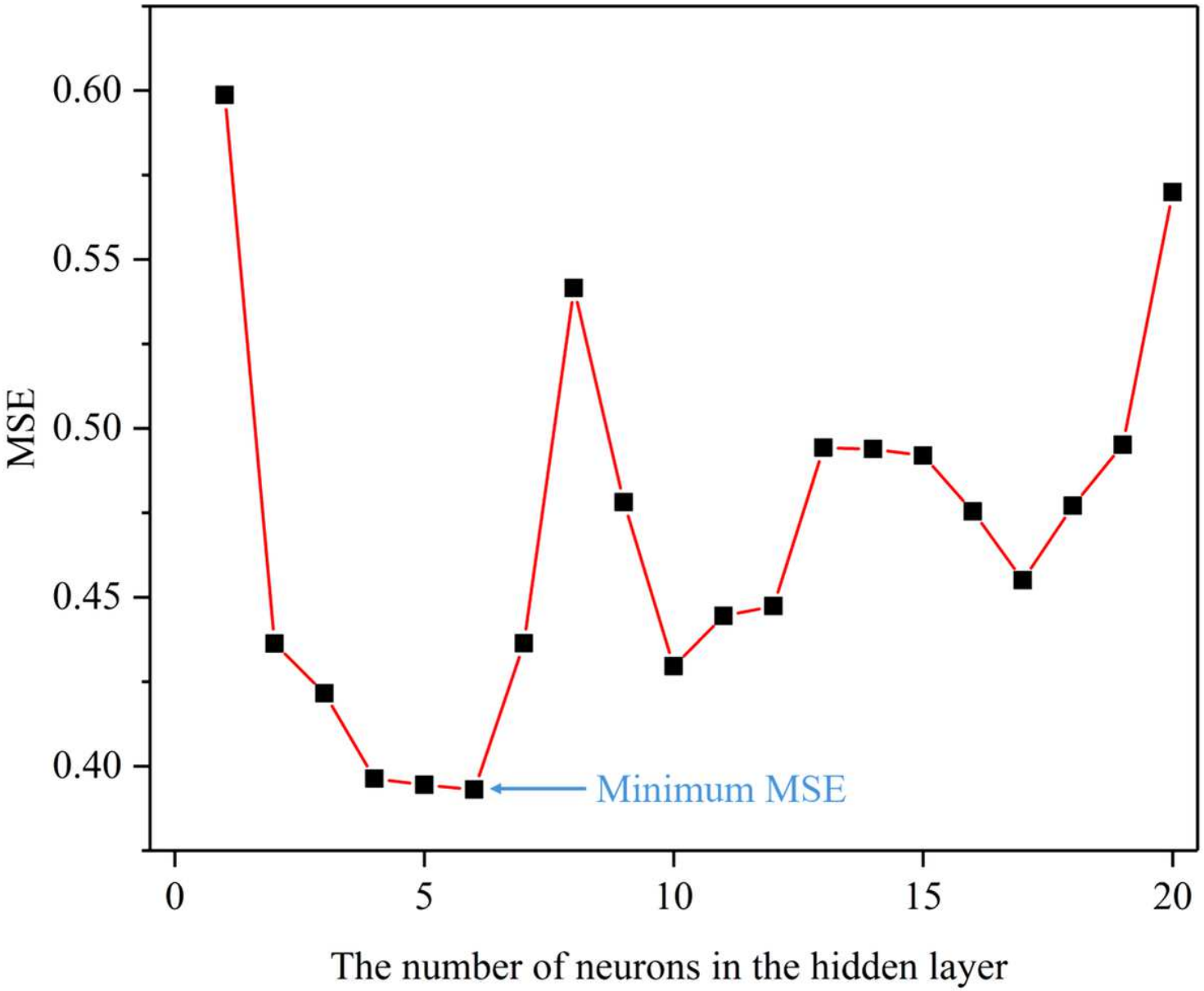

Figure 19

Relationship between MSE and the number of neurons in the hidden layer

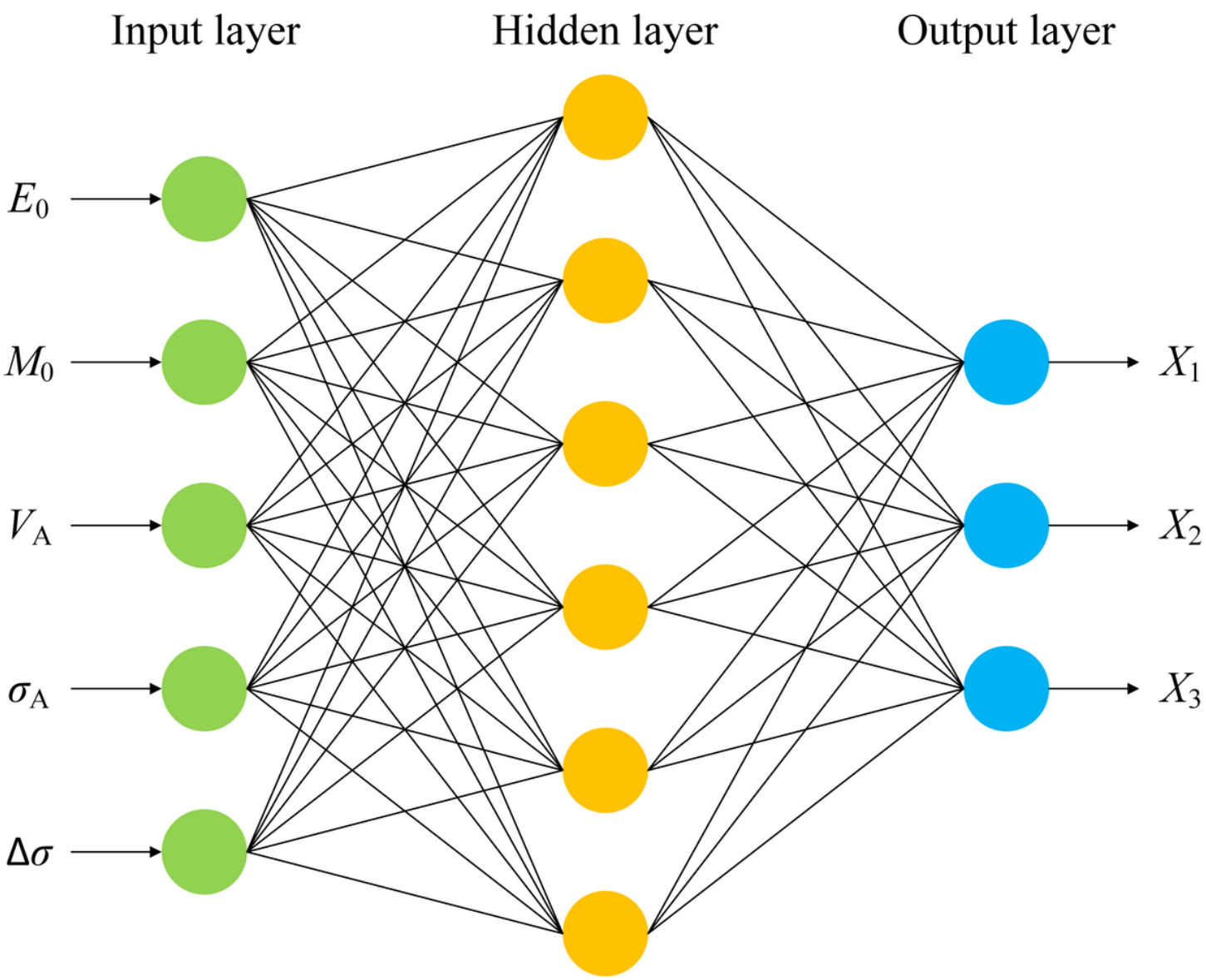

Figure 20

Topology of BP neural network