Prediction of coal and gas outbursts by a novel model based on multisource information fusion

Yingjie Li1,2, Yongguo Yang1,2* and Bo Jiang1,2

Abstract
As one type of geologic disaster, coal and gas outbursts seriously threaten safe production in coal mines, restricting the sustainable development of the mining industry. However, coal and gas outbursts are difficult to forecast due to their uncertainty and the limitation of sample size, which affect the accuracy of the traditional prediction methods to some extent. Therefore, this study developed a novel model based on multisource information fusion to realize the predictive progress of coal and gas outburst disasters. Through the application of Dempster-Shafer theory, a method of multisource information fusion, the proposed model combined the results of different forecasting approaches, including conventional techniques and an emerging method based on artificial intelligence. To enhance the performance of the established model, this study improved Dempster-Shafer theory and verified its effectiveness in dealing with highly conflicting information. We then applied this model to the No. 3 coal seam of the Xinjing coal mine, Shanxi Province, China. The fused prediction accurately reflected the situation of outburst hazards and showed good compensation for false prediction. An analysis of the results concluded that the model based on multisource information fusion increases the credibility of the forecast, which might provide technical support for safe coal mine production.

Keywords
Coal and gas outburst prediction, multisource information fusion, Dempster-Shafer theory, least squares support vector machine, uncertain information

1Key Laboratory of Coalbed Methane Resource & Reservoir Formation Process, Ministry of Education, China University of Mining & Technology, Xuzhou, China
2School of Resources and Geosciences, China University of Mining & Technology, Xuzhou, China

Corresponding author:
Yongguo Yang, Key Laboratory of Coalbed Methane Resource & Reservoir Formation Process, Ministry of Education, China University of Mining & Technology, Xuzhou 221116, China.
Email: ygyang88@hotmail.com
Introduction

Coal and gas outbursts (also referred to as outbursts) are geohazards that involve a sudden and violent disintegration of broken pulverized coal and gas from a coal mine working face to the roadway, permeating most coalbeds (Kissell and Iannacchione, 2014; Lama and Bodziony, 1998; Zhao et al., 2014). When mixed with air and ignition, serious blowouts will produce gas explosions and even fire accidents, therefore triggering fatalities and considerable economic losses. In China, there is a wide distribution of mines vulnerable to outburst disasters because of the unique and complex geological conditions in China, requiring specific and effective prevention and early warning measures to curb outburst disasters (Chen et al., 2019; Jia et al., 2017).

For almost half a century, seam-mining countries have focused on outburst prediction, and many prediction methods have been proposed. Díaz Aguado and González Nicieza (2007) investigated two different mines in the Asturias Central Basin, Spain and combined some predictive indices to establish an assessment of gas outbursts. The stochastic model approach of the Monte Carlo method applied by Wold et al. (2008) suggested that geomechanical variables, such as strength fields, also influence coal and gas outbursts. Islam and Shinjo (2009) found that gas emissions resulted from the lack of mine ventilation and the low atmospheric pressure. These approaches are generally divided into two categories: direct and indirect forecasting methods (Tang et al., 2016). Direct forecasting methods, also called index methods, directly contact the coal mass to measure some parameters that affect the occurrence of outbursts by drilling and exploring for the purpose of prediction of outbursts (Zhao et al., 2019). In terms of applicability in the field, these forecasting approaches are preferred in China. However, there are many complex nonlinear factors influencing coal and gas outbursts, and these indices are sensitive to different occurrences of dynamic phenomena during underground mining, thus possessing different critical values. What is more, in some particular cases, coal and gas outbursts occur even if the measured index value does not exceed the critical value. Therefore, in addition to the direct methods, there is also extensive research on indirect forecasting methods, which provide a new direction for the development of outburst prediction. Mathematical means have been frequently adopted in recent years as an indirect method in newly explored areas. These methods learn and analyse the correlation between the outburst occurrence and outburst factors after being trained by historical data.

Currently, artificial intelligence models are being applied to the mathematical methods of indirect prediction. Zhang et al. (2008) ascertained the relationship between the factors and hazards of coal and gas outbursts by multifactor pattern recognition and then predicted the hazard of each prediction unit. Wang et al. (2010) established a predictive model based on Bayesian discriminant analysis and attained reasonable results that completely coincided with the actual situation. Zhang and Lowndes (2010) applied a coupled artificial neural network and fault tree analysis model to the prediction of coal and gas outbursts, and the results indicated that the combined solution method could explicitly recognize model relationships between the geological conditions and the potential risk of outbursts. Li et al. (2017) used the initial gas flow in a borehole to determine the critical values of coal and gas outbursts based on Fisher’s criterion and predicted the risk of outbursts by a novel linear and continuous method, indicating the proposed method better than the traditional prediction method.
These advanced single artificial intelligence methods have achieved valuable results in the field of outburst prediction; nevertheless, coal and gas outbursts are uncertain geological dynamic phenomena because of their complicated nonlinear systems. Consequently, there are still some inadequacies in practical application when adopting only single methods. For instance, the poor stability of individual models may lead to a high rate of false alarms in field monitoring applications. To address such shortcomings, researchers recommended information fusion methods for predictive systems. For example, Zhang and Yang (2015) established a risk assessment model for coalbed methane development based on multisource information fusion and then evaluated the gas emission risk of eight mining areas of the Luling mine according to the six-layer functional classification model of the fusion system.

Significantly, in multisource information fusion technology, the synthesis algorithm is the theoretical basis, and with the emergence of new methods based on statistical inference, artificial intelligence and information theory, a growing number of efficient methods are being applied to the synthesis algorithm. Among these newly emerging applications of mathematical methods, support vector machine (SVM) and Dempster-Shafer theory (D-S evidence theory) are also appropriate for prediction.

SVM, a promising machine learning technique that analyses data for classification and regression, performs very well in some small sample data processing, which may become more complicated by deep learning models; therefore, it is very suitable for gas data. SVM has been successfully applied in many research fields, such as text categorization (Manochandar and Punniyamoorthy, 2018), image classification (Jac Fredo et al., 2019; Rahmani et al., 2019), health care (Cai et al., 2002; Hua and Sun, 2001), and other engineering fields (Dou et al., 2018; Dou and Yang, 2018; Liu et al., 2019; Zuo & Carranza, 2011). D-S evidence theory is a mathematical framework for analysing uncertain and partial information. D-S evidence theory has been used in a variety of studies and adapted in several fields, such as information fusion (Feizizadeh et al., 2014; Jiang et al., 2017), expert systems (Dymova et al., 2016; Qiu et al., 2018), image identification (Ghasemi et al., 2013; Haouas et al., 2019), and other engineering fields (Jiang et al., 2018).

Owing to the non-linearity and uncertainty of coal and gas outbursts, it is important to establish an accurate model based on particular geological features. Considering the aforementioned advantages and disadvantages of direct and indirect forecasting methods, we attempt to combine multiple methods and conduct the prediction of coal and gas outbursts based on multisource information fusion to improve the accuracy and reliability of the uncertain forecast. Hence, this paper studies the prediction of coal and gas outbursts in the Xinjing coal mine (in Shanxi Province, China) and proposes a new outburst prediction model, with a forecast in the Zhuxianzhuang coal mine as a validation in different geological conditions. The proposed model demonstrates its contributions in the following aspects. (1) The model is capable of flexibly determining the degree of correlation between each factor and outburst occurrences according to the actual conditions of different mining areas by learning historical data, which greatly enhances the model adaptability. (2) The improved D-S evidence theory resolves the failure of the traditional evidence theory when combining highly conflicting evidence, therefore improving the accuracy of convergence. (3) The model succeeds in combining different forecasting methods as well as presenting a novel basic probability assignment from preliminary predictions. (4) By applying the proposed model to research areas, it is indicated that predictions can be made in these areas to provide early warning for safe mine production.
The rest of this paper is organized as follows. In the next section, we briefly introduce the geological background of the Xinjing coal mine, and the prediction methodologies used in this paper (i.e., direct forecasting methods, SVM, and D-S evidence theory) appear in this section. Then, we validate the improved D-S evidence theory and present the application of the proposed fusion-based predictive model as well as the analysis of the converging results. Finally, brief conclusions are made in the last section.

**Geological settings and methodology**

**Geological situation**

The Xinjing coal mine, located in the middle of the North China plate (Figure 1), appears as a southwest-dipping monocline with a few secondary folds. The multiple-periodic, low-order fold structures of the mine provide favourable storage conditions for coalbed methane (Wang et al., 2017). This coal mine has been identified as a mine that is prone to high gas outbursts, and the No. 3 coalbed is an outburst coal seam, characterized by a simple structure, small fractures, and mostly normal faults.

In general, the coal and gas outbursts in the No. 3 coalbed are the result of the comprehensive action of gas pressure, the in situ stress and the physical and mechanical properties of coal. Therefore, research on outburst prediction there typically focuses on the parameters related to these three aspects and our present study adopts such factors to predict coal and gas outbursts: the burial depth of the coalbed (the depth from the coalbed to the bedrock surface), gas pressure, gas content, initial gas emission rate, Protodyakonov coefficient of coal strength, gas emission amount, coal seam thickness, geological structure, pillar collapses, and coal structure. The experiments in this research will rank these factors according to their relationship with the condition of outbursts and then extract them as input datasets for the prediction model.

![Figure 1. Structural outline of Xinjing mine (Li et al., 2019; Wei et al., 2007).](image)
Direct forecasting methods

The indices of the direct forecasting methods often consist of gas pressure \( (P) \), drilling gas inrush initial velocity \( (q) \), drill cutting desorption \( (\Delta h_2) \), drilling cutting weight \( (s) \), Protodyakonov strength index \( (f) \), initial gas emission rate \( (\Delta p) \), and comprehensive index \( (D \text{ and } K) \) (Zhou et al., 2019). According to the provisions of the prevention of coal and gas outbursts (China’s State Administration of Work Safety, 2009), the identification of an outburst-prone coal seam should first be grounded on the field gas dynamic phenomenon. When there are few or no significant characteristics of the dynamic phenomenon, the identification is performed by the forecasting indices, as shown in Table 1.

In addition to using a single index, Wang (1994) proposed a forecasting method using a comprehensive index (including values of \( D \) and \( K \)). The possibility of an outburst of coal seams can be estimated by \( D \) or \( K \)

\[
\begin{align*}
D &= (0.0075H/f - 3)(p - 0.74), \\
K &= \Delta p/f
\end{align*}
\]

where \( D \) and \( K \) are comprehensive indices recognizing whether there will be an outburst, \( H \) is the buried depth, \( f \) is Protodyakonov’s coefficient of coal mass, \( P \) is the gas pressure, and \( \Delta p \) is the initial velocity of gas emission. Generally, the critical values of \( D \) and \( K \) are determined by specific mines corresponding to their observation data, and when there are no reliable measured data, the risk of outburst can be determined by the referential value, as shown in Table 2. In Table 2, if the value calculated at a point is greater than the corresponding critical value, it is considered a dangerous location at which outbursts occur.

Least squares support vector machine: Indirect forecasting method

In practical applications of prediction, limited to sample size, we tend to pursue efficient inductive learning (learning by examples) methods to realize the automatic discovery of regularities in data and therefore apply these methods to new but similar data. Among these methods, SVM, developed from statistical learning theory, is a supervised machine learning algorithm categorizing the input dataset by constructing an optimal separating plane (Vapnik, 2000). This paper uses least squares support vector machine (LSSVM)—

### Table 1. Critical value of single index method.

| Indicators          | Degree of coal fracturing | Initial rate of gas emission (mmHg) | Protodyakonov strength index | Gas pressure (MPa) |
|---------------------|---------------------------|-------------------------------------|-----------------------------|-------------------|
| Critical value      | III, IV, V                | ≥10                                 | ≤0.5                        | ≥0.74             |

### Table 2. Critical value of comprehensive index \( D \) and \( K \).

| Comprehensive index \( D \) | Comprehensive index \( K \) |
|-----------------------------|------------------------------|
| ≥0.25                       | Anthracite: ≥20, Other types of coal: ≥15 |
modified version of SVM proposed by Suykens and Vandewalle (1999)—as one of the three preliminary forecasting methods.

SVM was originally designed for binary classification and then gradually extended to multiclass classification, so the one-against-one method was one of the early and major implementations (Hsu and Lin, 2002), with its principle shown in Figure 2. Given two types of samples (represented by green circles and pink triangles in Figure 2), the task of SVM is to find the optimal classification hyperplane to separate the samples and create rules that will guide the classification of future samples. In Figure 2, hyperplanes $P_1$ and $P_2$ are all capable of segregating the classes well, but $P$ is the optimal classification hyperplane that can maximize the distances—called margin—between the nearest data point of each class.

Assume that there is a training set $x_i \in \mathbb{R}^n$ ($i = 1, \ldots, k$), and $y_i \in \{-1, 1\}$ is the associated label for each vector $x_i$. SVM aims to find the hyperplane $P$ by the following linear equation

$$\omega \cdot x + b = 0$$

(2)

where $\omega \in \mathbb{R}^n$ is the weight vector that determines the direction of $P$, and $b$ is a constant influencing the distance between $P$ and the origin. Suppose that the hyperplane $(\omega, b)$ is an optimal hyperplane, then $\omega \cdot x + b > 0$ if $y_i = +1$ (or $\omega \cdot x + b < 0$ if $y_i = -1$).

As shown in Figure 2, the data points closest to the hyperplane satisfy the following equation

$$y_i (\omega \cdot x_i + b) \geq 1, \quad i = 1, 2, \ldots, k$$

(3)

Figure 2. Linear binary classifier (in the Cartesian plane) (Zhou, 2016).
These points are called “support vectors”, and the sum of the two distances of different support vector classes to the hyperplane is $2/\|\omega\|$, also known as the margin. Hence, the work of finding $\mathcal{P}$ is settled by the solution of $\omega$ and $b$ that meets equation (3) to maximize the margin, i.e., the objective is stated as the following optimization problem

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2$$

$$s.t. \ y_i[\omega \cdot x_i + b] \geq 1, \ i = 1, 2, \ldots, k$$

Then, the solution of equation (4) will deduce a model (equation (5)) corresponding to the optimal classification of hyperplane $\mathcal{P}$

$$g(x) = \omega \cdot x + b$$

To solve the convex quadratic programming of equation (4), the Lagrange function is introduced, and its dual problem (equation (6)) can be obtained

$$L(\omega, b, \alpha) = \frac{1}{2} (\omega \cdot \omega) - \sum_{i=1}^{k} \alpha_i \{y_i[\omega \cdot x_i + b] - 1\}$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ is the Lagrange multiplier. Therefore, an optimal saddle point can be computed by differentiating with respect to $\omega$ and $b$. By substitution and elimination of variables, the dual problem of equation (4) can be found by

$$\max_\alpha \sum_{i=1}^{k} \alpha_i - \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{k} \alpha_i \alpha_j y_i y_j x_i^T x_j,$$

$$s.t. \ \sum_{i=1}^{k} \alpha_i y_i = 0, \ \alpha_i \geq 0, \ i = 1, 2, \ldots, k$$

After solving for the variable $\alpha$ in equation (7), the variables $\omega$ and $b$ can be calculated, and thus far, the final hypothesis is consequently turned into a linear combination of the training points with its discrimination function expressed as

$$g(x) = \text{sign} \left[ \sum_{i=1}^{k} y_i \alpha_i (x_i \cdot x) + b \right]$$

where $k$ is the number of support vectors.
In linearly inseparable cases, a hyperplane with the smallest number of errors can be constructed by the concept of a soft margin. By introducing nonnegative slack variables $e_i$, the constrained optimization problem in equation (4) becomes

$$
\min_{\omega, b, e_i} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{k} e_i,
$$

s.t. $y_i(\omega \cdot x_i + b) + e_i \geq 1, \ i = 1, 2, \ldots, k,$ and $e_i \geq 0$

where $C$ is a cost parameter penalizing the slack variables $e_i$ in the objective function. Therefore, the optimization can be solved through the application of the Lagrange multiplier in the same way as in the separable case (equation (7)) by transforming equation (9) into its dual problem, subject to $0 \leq x_i \leq C, \ i = 1, 2, \ldots, k$.

However, in most classification problems, there may not exist a hyperplane that divides two samples correctly in the primal space. To address the problem, Vapnik initiated kernel functions $\phi$, which can map the data from the original feature space to a higher dimensional feature space $(\phi(x_1), \phi(x_2), \ldots, \phi(x_k))$ where an optimal hyperplane can be found, achieving the solution by a linear algorithm. The kernel function $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ realizes nonlinear transformation (equation (10)) by replacing the inner products $(x_i^T x)$ in the dual Lagrange equation (7)

$$
\max_{\alpha_i} \sum_{i=1}^{k} \alpha_i - \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{k} \alpha_i \alpha_j y_i y_j K(x_i, x_j)
$$

s.t. $\sum_{i=1}^{k} \alpha_i y_i = 0, \ \alpha_i \geq 0, \ i = 1, 2, \ldots, k$

Similar to the method for the separable case, after solving equation (10), the optimization model (equation (11)) can be attained

$$
f(x) = \omega^T \phi(x) + b = \text{sign} \left[ \sum_{i=1}^{k} \alpha_i y_i K(x_i, x) + b \right]
$$

LSSVM, as a modified version of the conventional SVM, transforms quadratic programming into a solution of a set of linear equations and substitutes equality type constraints for inequalities from the classic SVM approach, thus improving the model. LSSVM introduces least squares to SVM by expressing the classification problem

$$
\min_{\omega, b, e} T(\omega, b, e) = \frac{1}{2} \omega^T \omega + \frac{1}{2} \sum_{i=1}^{k} e_i^2
$$

s.t. $y_i(\omega^T \phi(x_i) + b) = 1 - e_i, \ i = 1, 2, \ldots, k$
where $\gamma$ is a regularization factor, like the penalty factor $C$ in SVM, and $e_i$ the slack variable.

Similar to the optimization problem of SVM, Lagrange multipliers $\lambda_i$ are applied and the objective function (equation (13)) can be obtained

$$
\mathcal{L}(\omega, b, e; x) = T - \sum_{i=1}^{k} \lambda_i \{ y_i [\omega^T \phi(x_i) + b] - 1 + e_i \}
$$

Then the conditions for optimality are acquired by partially differentiating by $\omega$, $b$, $e$, and $x$. After eliminating the intermediate variables $\omega$ and $e$, $\lambda_i$ and $b$ can be obtained by solving the linear equation (Suykens et al., 2002). The final LSSVM becomes

$$
g(x) = \text{sign} \left[ \sum_{i=1}^{k} \lambda_i K(x_i, x) + b \right]
$$

$K(x_i, x)$ is also a kernel function in equation (14). General kernel functions include the sigmoid kernel, the radial basis function kernel (Gaussian RBF kernel) and the polynomial kernel. When prior knowledge is absent, the radial basis function kernel performs better and is appropriate for samples in any distribution with sound generalization; therefore, this paper adopts it as kernel functions.

**Dempster-Shafer theory and its improvement**

Both direct and indirect methods have their own advantages. However, the range of the identification results of the three forecasting methods, i.e., the direct forecasting methods and LSSVM, may vary widely; thus, a comprehensive consideration based on the uncertainty reasoning method may be required to increase the reliability of coal and gas outburst prediction.

As one of the uncertainty reasoning methods, D-S evidence theory is becoming increasingly attractive. Compared with the theories of statistical inference, such as probabilistic reasoning, D-S evidence theory is more flexible in measuring indeterminacy and more concise in reasoning mechanisms and is closer to the general convention of human thinking, especially when handling unknown information.

**Basic concept.** Let $\Theta$ be a finite and exhaustive set of $N$ mutually exclusive elements, indicated by $\Theta = \{\theta_1, \theta_2, \ldots, \theta_i, \ldots, \theta_N\}$, where $\Theta$ is called the frame of discernment, and $\theta_i$ is one of the events in frame $\Theta$ ($i = 1, 2, \ldots, N. N$ is the number of events) (Shafer, 1976). All the possible subsets of $\Theta$ constitute the power set of $\Theta$, represented by $2^\Theta$.

**Definition 2.1.** $\forall A \in \Theta$, a basic probability assignment (BPA) is a mapping ($m$) from $2^\Theta$ to $[0, 1]$, formally defined by

$$
m : 2^\Theta \rightarrow [0, 1]
$$

that also satisfies the following condition

$$
\begin{cases}
m(\emptyset) = 0, \\
\sum_{A \in \Theta} m(A) = 1
\end{cases}
$$
In Shafer’s original definition, the condition $m(\emptyset) = 0$ is required, which corresponds to the “closed-world assumption”. However, in some recent papers, this condition is often omitted, corresponding to the “open-world assumption” (Jousselme et al., 2001). The rest of the paper considers the closed-world assumption.

We can regard a BPA as our belief at an observation time: $m(\emptyset) = 0$ represents no belief in a hypothesis, $m(\emptyset) = 1$ represents total belief, and $m(\emptyset) \in (0, 1)$ represents partial belief. A BPA with $m(\emptyset) = 0$ is also known as a mass function (Liu, 2006), and an element of $2^\Theta$ having a non-zero mass is called a focal element.

**Definition 2.2.** Assuming that $m_1$ and $m_2$ are two BPAs defined on the frame of discernment $\Theta$, the combined BPA by Dempster’s combination rule is thus defined by

$$m(A) = m_1 \oplus m_2 = \frac{\sum_{A_i \cap B_j = A} m_1(A_i)m_2(B_j)}{1 - \mu(\emptyset)}$$

where $\mu(\emptyset) = \sum_{A_i \cap B_j = \emptyset} m_1(A_i)m_2(B_j)$. The case of $\mu(\emptyset) = 0$ implies no conflict between $m_1$ and $m_2$, while $\mu(\emptyset) = 1$ represents that $m_1$ completely contradicts $m_2$.

When there are $n$ ($n > 2$) pieces of evidence, i.e., $m_1, m_2, \ldots, m_n$ are BPAs defined on $\Theta$, the combined BPA $m$ is formulated as follows

$$m = m_1 \oplus m_2 \oplus \cdots \oplus m_n$$

The final orthogonal sum $m$ represents our belief in the most likely object.

**Improvement of the D-S evidence theory.** In the traditional D-S rule, a wrong convergence contradicting reality may be yielded when $\mu(\emptyset) \rightarrow 1$, i.e., there is a presence of relatively higher contradiction between some pieces of evidence. Thus, some sophisticated alternatives to the conventional D-S evidence theory have been proposed and are categorized into combination modification (Lefevre et al., 2002) and source modification (Murphy, 2000). Combination modification argues that the reason for the conflict is the improper disposal of evidence when empty sets are present. However, source modification states that a disturbed evidence source triggers conflicts; therefore, there is a need to pre-process the evidence. This paper considers the two ideas, presenting a novel conflict measurement and adopting an optimized combination to solve the paradox that may occur when conflicting evidence exists.

**Step 1. Measurement of conflict.**

The degree of conflict between two pieces of evidence can be described by the distance between two BPAs. In the direct application of the D-S evidence theory, $\mu(\emptyset)$, explained in the previous section, is generally regarded as the measure of conflict. The $\mu(\emptyset)$ changes the value of BPAs in all the focal elements in both pieces of evidence to $(1 - \mu(\emptyset))$ times the original amount. However, the conflict is not the result of all the focal elements. Therefore, many scholars have studied the correlation between evidence and proposed the inter-evidence distance to measure both conflict and similarity, which mainly includes the following three distances.
(1) Jousselme’s distance.

Assume that \( m_1 \) and \( m_2 \) are two BPAs defined on the frame of discernment \( \Theta \). Then, the distance between \( m_1 \) and \( m_2 \) is denoted as (Jousselme et al., 2001)

\[
d(m_1, m_2) = \sqrt{\frac{1}{2} (m_1 - m_2)^T D (m_1 - m_2)}
\]

(19)

where \( D \) is a \( 2^N \times 2^N \) positively defined matrix, with its elements

\[
D(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad A, B \in R(\Theta)
\]

(20)

\(|A|\) denotes the cardinality of \( A \), and \( D \) conveys the information that the distance
\( d(m_1, m_2) \) is a function concerning \(|A|\), \(|B|\), \(|A \cap B|\), and \(|A \cup B|\). The conflict between \( A \) and \( B \) can be measured by \(|A \cap B|\): when \(|A \cap B| = 0\), there is no common element in \( A \) and \( B \), then it is asserted that \( A \) strongly conflicts with \( B \).

(2) Pignistic probability.

Pignistic probability transformation (PPT) (Ristic and Smets, 2006) is a probability transforming the BPA to probability. Let \( m(A) \) be a BPA on the D-S frame \( \Theta = \{x_1, x_2, \ldots, x_n\} \), and then its PPT on \( \Theta \) is defined by

\[
BetP_m(A) = \sum_{A \subseteq \Theta} \frac{|A \cap B|}{|B|} \frac{m(B)}{1 - m(\Theta)}
\]

(21)

where \( B \subseteq \Theta \). In addition, the PPT for the singleton is

\[
BetP_m(\{x\}) = \sum_{A \subseteq \Theta, x \in A} \frac{m(A)}{|A|}
\]

(22)

\( BetP_m \) can also be generalized to power set \( 2^\Theta \)

\[
BetP_m(A) = \sum_{x \in A} BetP_m(x)
\]

(23)

(3) Liu’s distance.

Liu (2006) presented an alternative quantitative measure to reveal the relationship among beliefs using the difference of the PPT among evidence to describe the degree of conflict. Let \( m_1 \) and \( m_2 \) be two BPAs on frame \( \Theta \), and \( BetP_{m_1}(A) \) and \( BetP_{m_2}(A) \) be their respective PPTs. Then, the distance between betting commitments of the two BPAs can be given by

\[
difBetP^m_{m_2} = \max_{A \subseteq \Theta} |BetP_{m_1}(A) - BetP_{m_2}(A)|
\]

(24)

where \( difBetP^m_{m_1} = difBetP^m_{m_2} \), and \( 0 \leq difBetP^m_{m_i} \leq 1 \).
When there are \( n \) sources on \( \Theta \) with their corresponding BPA \((m_1, m_2, \ldots, m_i, \ldots, m_n)\), the PPT among these \( n \) sources is

\[
difBetP = \max_{i < j} difBetP^m_{m_i}
\]  

(25)

The above three methods are commonly used as the measurement of the distance between evidence, which consider the distance not only from the perspective of dissimilarity but also from consistency. Inspired by Jousselme, who viewed a BPA as a special case of vectors, many researchers adopt this geometrical interpretation of BPA when studying the distance (Wen et al., 2008); hence, we can measure the conflict among evidence using the cosine of vectors. Assume that there are \( n \) hypotheses on frame \( \Theta \), and every BPA of their evidence is a vector:

\[
m_i = \frac{1}{C_{P(\Theta)}} \sum_{k=1}^{P(\Theta)} m_i(A_k)
\]

where \( Y_{P(\Theta)} \) is the vector space generated by elements of \( P(\Theta) \). Then, we can use

\[
\cos(m_i, m_j) = \frac{d(m_i, m_j)^2 + difBetP^m_{m_i}^2}{2}
\]

(26)

is called the parameter of conflicting degree between \( m_1 \) and \( m_2 \).

In equation (26), \( d(m_i, m_j) \) is the Jousselme distance, \( difBetP^m_{m_i} \) is the Pignistic distance, and \( \cos(m_i, m_j) \) is the cosine distance. Additionally, equation (26) satisfies the requirements for conflicting measurement (nonnegativity, symmetry, and boundedness).

Step 2. Calculation of the degree of support.

Deng et al. (2004) proposed a method for expressing the weight of evidence according to the definition of the degree of support among the evidence; when the evidence is largely supported by other evidence, this evidence is more similar to other evidence and thus possesses a greater weight. In this part, we redefine the supporting degree of evidence with Deng’s weight for reference:

Definition 2.4. Let \( E_i \) be the evidence on the frame of discernment \( \Theta \), given that the conflict between two BPAs is \( conf(m_i, m_j) \), \( \forall i, j \in [1, n] \). Then, the similarity between evidence \( i \) and evidence \( j \) is

\[
sim(m_i, m_j) = \begin{cases} 
1 - conf(m_i, m_j), & i \neq j \\
1, & i = j
\end{cases}
\]

(27)
Considering the positive correlation between $\text{sim}(m_i, m_j)$ and the similarity of evidence, the support degree of evidence can, therefore, be denoted by

$$\text{sup}(m_i) = \sum_{j=1, j\neq i}^{n} \text{sim}(m_i, m_j), \quad i, j = 1, 2, \ldots, n$$

(28)

On a frame $\Theta$, when a piece of evidence is notably against other evidence, it can be convinced that the information it provides is inadequate; consequently, the support degree of evidence $\text{sup}(m_i)$ should be allocated relatively smaller with the purpose of eliminating or reducing the impact on convergence.

The relative support degree of evidence is then obtained by the normalization of the support degree

$$\text{norm}(m_i) = \frac{\text{sup}(m_i)}{\max \text{sup}(m_i)}, \quad i = 1, 2, \ldots, n$$

(29)

where $0 \leq \text{norm}(m_i) \leq 1$, obviously, and $\text{norm}(m_i)$ represents the reliability of a piece of evidence supported by the whole evidence source.

Step 3. Modification of the evidence source.

Considering that highly conflicting evidence yields counterintuitive results, in this step, we rectify the evidence by their assistant measuring index. The modified BPAs are defined as follows

$$\begin{align*}
  m'_i(A_r) &= \text{norm}(m_i) \cdot m_i(A_r), \\
  m'_i(\Theta) &= \left(1 - \text{norm}(m_i)\right) + \text{norm}(m_i) \cdot m_i(\Theta)
\end{align*}$$

(30)

where $m'_i(A_r)$ is the new BPA allocated to $A_r$.

Step 4. Cluster.

In order to improve the efficiency of fusion, we implement a clustering analysis to the new modified BPAs, combining evidence with different degrees of conflict in different combination rules. Agglomerative hierarchical clustering is chosen here due to its character of merging from singletons to a root cluster. At the first stage of this process, every BPA ($m'_i$) is a maximal cluster ($P_i$), and their $d(m'_i, m'_j)$ can be calculated to measure the conflict. In addition, a threshold ($\delta$), determined by practical application, is adopted as the standard to assemble similar BPAs into a cluster: If $d(m'_i, m'_j) > \delta$, there is an extreme divergence among the evidence, therefore requiring clustering to different classes. When all the $d(m'_i, m'_j)$ between newly formed clusters ($P$) are greater than $\delta$, the process ends, producing $x$ clusters ($P_1, P_2, \ldots, P_x$). Similar to the definition of the relative support degree of evidence, the relative support degree of the cluster is deduced as follows.

Definition 2.5. Let ($P_1, P_2, \ldots, P_x$) be clusters after the merging of $n$ evidence, with each $P_i$ involving $q$ piece of evidence ($q \geq 1$). The relative support degree of the cluster is then

$$\text{norm}(P_i) = \text{norm}(m_1) + \text{norm}(m_2) + \cdots + \text{norm}(m_q)$$

(31)
Step 5. Combination.

Evidence within a certain cluster \( P_i \) will still be combined by Dempster’s rule; therefore, \( t \) clusters of evidence \((t \leq n)\) as well as their relative support degree can be generated. The conflicting clusters should be converged by a new combination rule that requires a rectification factor (Yang et al., 2011):

**Definition 2.6.** A rectification factor of two clusters is

\[
E_{ij} = \sum_r |P_i(A_r) - P_j(A_r)|, \quad r = 1, 2, \ldots, 2^N
\]

where \( E_{ij} \) measures the absolute difference between \( P_i \) and \( P_j \) and is positively correlated to the difference. The relative difference measurement is thus defined as

\[
E_{ij}(A_r) = \frac{|P_i(A_r) - P_j(A_r)|}{\sum_r |P_i(A_r) - P_j(A_r)|}, \quad r = 1, 2, \ldots, 2^N
\]

where \( E_{ij}(A_r) \) is the relative difference on focal element \( A_r \).

Then, we can define a novel combination rule based on clustered evidence:

**Definition 2.7.** The novel combination rule of conflicting clusters is

\[
m(A_r) = \begin{cases} 
E_{ij}(A_r) \times \text{conf}(P_i, P_j) + \sum_{X \cap Y = A_r, \forall X, Y \subseteq \Theta} P_i(X) \times P_j(Y), & X \cap Y \neq \emptyset \\
0, & X \cap Y = \emptyset
\end{cases}
\]

where \( \text{conf}(P_i, P_j) = \frac{P_i(A_r) \times P_j(A_r)}{P_i(A_r) + P_j(A_r) - P_i(A_r) \times P_j(A_r)} \).

\[
m(\emptyset) = 1 - \sum_r \left( E_{ij}(A_r) \times \text{conf}(P_i, P_j) + \sum_{X \cap Y = A_r, \forall X, Y \subseteq \Theta} P_i(X) \times P_j(Y) \right)
\]

Verifiably, the new combination satisfies requirements for BPAs, i.e., \( \sum_{A_r \subseteq \Theta} m(A_r) = 1 \).

**Results and discussion**

**Validation and evaluation of the improved D-S evidence theory**

In order to validate the improvement of the modified D-S evidence theory, this paper fuses five pieces of evidence from different sources and compares the convergence results with the combination rule in other studies. Assume that the focal element \( A \) is the correct object, and five different pieces of evidence with their BPAs on the frame of discernment \( \Theta \) are

\[
m_1 : m_1(A) = 0.5, \quad m_1(B) = 0.2, \quad m_1(C) = 0.3,
m_2 : m_2(A) = 0, \quad m_2(B) = 0.9, \quad m_2(C) = 0.1,
m_3 : m_3(A) = 0.55, \quad m_3(B) = 0.1, \quad m_3(C) = 0.35,
m_4 : m_4(A) = 0.55, \quad m_4(B) = 0.1, \quad m_4(C) = 0.35,
m_5 : m_5(A) = 0.55, \quad m_5(B) = 0.1, \quad m_5(C) = 0.35
\]
The convergence results of the different combination rules are shown in Table 3. From Table 3, it can be seen that D-S combination rule (Dempster, 2008) consistently regards that focal element $A$ should be assigned no belief only because of $m_2(A) = 0$; and with increasing evidence, beliefs will increasingly gather to focal element $C$, which reflects that the merging result is influenced considerably by $m_2$. The Yager rule (Yager, 1987) attributes the conflicting information to the universal set, so this operation reduces the belief on element $B$; nevertheless, with increasing combination times, almost all the BPA convergence concentrates on the frame $\Theta$, which is practically helpless for assistant decisions. Sun’s method (Sun et al., 2000) introduces reliability to the combination, whose accuracy is better than that of the Yager rule, but $m(\Theta)$ is still greater than 0.5, also not beneficial to the decision. The Murphy rule (Murphy, 2000) combines the weighted average evidence four times, reasonably locates belief to hypothesis $A$, and gives less belief to $B$ and $C$. The proposed rule in this paper correctly provides the possibility of the occurrence of $A$ in a way that is quicker than Murphy’s combination. Furthermore, when completing the convergence with only three pieces of evidence, $A$ emerges as a high probability event. Thus, the proposed method is appropriate for conflicting materials in a more reasonable and rapid way.

### Empirical data preparation

Because there are many complicated factors that cause coal and gas outbursts, it is a necessity to rank those elements and thereby determine the main controlling factors. The present research uses grey relational analysis (GRA) to analyse and investigate the key factors of outburst events. It is characterized by its convenient calculation, availability to any data volume, and fewer errors compared with that of qualitative analysis (Yang et al., 2012).
It regards the similarity or dissimilarity of the factor trend as the measurement of the relational degree among the elements (Xiong et al., 2018).

Considering the outburst probability and the degree of coal damage in the research area, we reclassify the coal body structure in Table 1 into three types from a macro perspective—primary constructional coal and ruptured coal (coal I), mortar coal (coal II), and mylonitic coal (coal III)—with numbers 1, 2, and 3 representing coal I, II, and III in the following paragraphs, respectively (Zhang et al., 2016). Obviously, coal III is the coal prone to outburst. Based on the historical data in the study area, the ranking of the degrees of correlation is acquired, as shown in Table 4. The order of the contribution of each factor is as follows: coal structure > gas pressure > coalbed burial depth > gas content > initial gas emission rate > Protodyakovskov coefficient of coal strength > coal seam thickness > gas emission amount > geological structure > pillar collapses. The first seven elements whose degrees of correlation are greater than 0.55 are selected as the prediction indices of coal and gas outbursts.

### Performance of the predictive model

The proposed predictive model is executed on 57 datasets of field geological data in the No. 3 coal seam of the Xinjing coal mine. Twenty samples are selected randomly for coal and gas outburst prediction, and the remaining data are used as the training sets of LSSVM. After ranking the relational degree of the factors, the initial gas emission rate is chosen for the single index method, the Protodyakovskov coefficient of coal strength accompanied by the initial gas emission rate are chosen for the composite index method \((K)\), and the remaining five factors are chosen for LSSVM. The field-derived geological data for the prediction of coal and gas outbursts at the Xinjing coal mine are shown in Table 5, where \(R_1 = [1]\) represents the outburst event and \(R_2 = [-1]\) signifies the non-outburst event. The results from three prediction methods (single index, composite index, and LSSVM) will be taken as BPAs of the D-S evidence theory for the final prediction. The process of the predictive model based on multisource information fusion is shown in Figure 3.

### Construction of the predictive model

Let the D-S frame of discernment be \(\Theta = \{R_1, R_2\}\), where \(R_1\) is the presence of an outburst hazard, and \(R_2\) is its absence. Each of the testing samples has three pieces of evidence (evidence A, B, and C) from the different forecasting methods.

| Factors                          | Correlation degrees | Rank |
|----------------------------------|---------------------|------|
| Coal structure                   | 0.64                | 1    |
| Gas pressure                     | 0.60                | 2    |
| Coalbed burial depth             | 0.59                | 3    |
| Gas content                      | 0.58                | 4    |
| Initial gas emission rate        | 0.57                | 5    |
| Protodyakovskov coefficient      | 0.56                | 6    |
| Coal seam thickness              | 0.54                | 7    |
| Gas emission amount              | 0.53                | 8    |
| Geological structure             | 0.43                | 9    |
| Pillar collapses                 | 0.41                | 10   |
Evidence A (The first evidence from the LSSVM prediction model).

First, we use 37 field datasets collected from previously mined zones in the Xinjing coal mine for leave-one-out cross validation, which is suitable for the proof of small samples, to verify the performance of the LSSVM model. The overall datasets are divided into two sets (a training set with 36 groups of data and a testing set with one group of data) for cross-validation testing in succession.

The experiment is performed in a MATLAB® R2016a environment, and the LSSVM model is built on the dataset with the help of LSSVMlab Toolbox (version 1.8) (Pelckmans et al., 2002). Since the kernel function adopted in this paper is the Gaussian RBF kernel, two parameters needed to be tuned: the regularization parameter ($\gamma$) and the squared bandwidth ($\sigma^2$). The LSSVMlab Toolbox has a built-in tune function applying the state-of-the-art global optimization technique, coupled simulated annealing (CSA), to find the suitable parameters and then fine-tune these parameters by grid search as a second optimization procedure. During the parameter tuning, the performance is estimated by 10-fold cross validation.

Before being input to the LSSVM model, the empirical datasets should be normalized within the range of [0, 1]: (given that the number of samples is $n$ and the number of variables is $m$)

$$x_{i}(k)_{\text{norm}} = \frac{x_{i}(k) - \min_{k} x_{i}(k)}{\max_{k} x_{i}(k) - \min_{k} x_{i}(k)}, \quad i = 1, 2, \ldots, m; k = 1, 2, \ldots, n$$

Table 5. Field derived geological data at the Xinjing coal mine.

| Data set no. | Coal seam thickness | Coal structure | Gas content | Coalbed burial depth | Gas pressure | Initial gas emission rate | Protodyakonov coefficient of coal strength | K | Actual situation of outbursts |
|--------------|---------------------|----------------|-------------|----------------------|--------------|--------------------------|-------------------------------------------|---|-----------------------------|
| 1            | 2.255               | 3              | 12.157      | 578.564              | 2.056        | 14.654                   | 0.400                                     | 36.638 | $R_{1} = [1]$ |
| 2            | 2.466               | 2              | 17.200      | 566.687              | 1.921        | 7.676                    | 0.555                                     | 13.819 | $R_{2} = [-1]$ |
| 3            | 2.076               | 2              | 9.517       | 663.447              | 1.519        | 20.856                   | 0.497                                     | 41.991 | $R_{1} = [1]$ |
| 4            | 2.310               | 2              | 15.887      | 593.607              | 2.057        | 17.536                   | 0.507                                     | 34.603 | $R_{1} = [1]$ |
| 5            | 2.011               | 2              | 9.561       | 600.304              | 2.207        | 19.458                   | 0.301                                     | 64.601 | $R_{1} = [1]$ |
| 6            | 2.492               | 2              | 14.804      | 564.902              | 1.271        | 7.269                    | 0.555                                     | 13.101 | $R_{2} = [-1]$ |
| 7            | 2.570               | 2              | 13.810      | 485.590              | 2.167        | 23.617                   | 0.533                                     | 44.310 | $R_{1} = [1]$ |
| 8            | 1.712               | 2              | 5.344       | 603.868              | 2.020        | 19.613                   | 0.445                                     | 44.073 | $R_{1} = [1]$ |
| 9            | 2.177               | 2              | 11.777      | 540.384              | 2.259        | 18.998                   | 0.405                                     | 46.913 | $R_{1} = [1]$ |
| 10           | 2.498               | 2              | 16.556      | 564.542              | 1.479        | 7.227                    | 0.552                                     | 13.102 | $R_{2} = [-1]$ |
| 11           | 2.223               | 2              | 9.705       | 605.107              | 2.078        | 21.004                   | 0.328                                     | 64.028 | $R_{1} = [1]$ |
| 12           | 2.300               | 3              | 7.548       | 495.515              | 2.286        | 24.090                   | 0.204                                     | 118.091| $R_{1} = [1]$ |
| 13           | 1.899               | 2              | 8.868       | 663.411              | 1.482        | 18.885                   | 0.538                                     | 35.086 | $R_{1} = [1]$ |
| 14           | 2.531               | 2              | 15.061      | 562.041              | 1.401        | 6.996                    | 0.560                                     | 12.504 | $R_{2} = [-1]$ |
| 15           | 2.070               | 2              | 7.110       | 539.026              | 2.259        | 19.948                   | 0.379                                     | 52.661 | $R_{1} = [1]$ |
| 16           | 2.207               | 3              | 11.637      | 560.520              | 2.050        | 9.574                    | 0.390                                     | 24.521 | $R_{1} = [1]$ |
| 17           | 2.263               | 2              | 10.172      | 592.702              | 1.544        | 17.388                   | 0.415                                     | 41.863 | $R_{1} = [1]$ |
| 18           | 1.606               | 3              | 6.381       | 634.082              | 2.173        | 19.417                   | 0.405                                     | 47.960 | $R_{1} = [1]$ |
| 19           | 1.394               | 3              | 5.894       | 650.234              | 2.076        | 22.081                   | 0.285                                     | 77.358 | $R_{1} = [1]$ |
| 20           | 2.500               | 2              | 17.031      | 561.477              | 1.620        | 7.248                    | 0.555                                     | 13.064 | $R_{2} = [-1]$ |
where $x_i(k)$ is the value of sample $k$ under factor $i$, and $x_i(k)_{\text{norm}}$ is the normalized value of $x_i(k)$. The training dataset is denoted as $T = \{(x_i, y_i), i = 1, 2, \ldots, 37\}, x \in R^m, y_i \in \{-1, 1\}$, where $y_i$ are levels of actual outburst occurrences (non-outburst danger (-1), occurrence of an outburst (1)).

After the pre-processing of the data and the search of the suitable tuning parameters, we train the LSSVM model and then simulate and evaluate it at test points. The classified results are described in Table 6 (the misclassified values are in bold) showing that LSSVM misjudges the situation at No.18 and No.33.

We measure the efficacy of this trained LSSVM model by popular performance metrics (including accuracy, precision, recall, and F1 score) for this classification task, where the occurrence of an outburst is treated as positive. Note that “positive” here is the definition of the presence of events rather than their absence, not the actual effect on coal mines. Accuracy is the proportion of correct predictions to the total sample size, while precision is specific for results and represents the proportion of actual positive samples (correctly predicted occurrence of outbursts) in the total outburst predictions. Recall is specific for the original sample, referring to the correct predictions of outbursts among the total outburst events, and the F1 score is based on precision and recall: $F1 = 2 \times (\text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$. From the results, the accuracy of predicted outbursts is calculated as 0.95, the precision is calculated as 1, and the recall is calculated as 0.89. Therefore, the F1 score is 0.94, showing that the performance of this trained LSSVM model fits well with the field data.
Since the validity is demonstrated of LSSVM in coal and gas outburst prediction, it can be applied as a preliminary analysis for the prediction model proposed in this paper. In this section, we adopt the Bayesian posterior class probability of each predicted sample as the constructed BPA (equation (37)) in each outburst prediction

\[
\begin{align*}
    m_1(R_1) &= p(R_1), \\
    m_1(R_2) &= 1 - p(R_1)
\end{align*}
\]

where \( p(R_1) \) is the posterior class probability of the occurrence of outbursts. The predictive results and BPAs of each point from the first evidence source are then obtained, as shown in Table 7, where \( m_1(R_1) \) represents the probability of outburst occurrence, and \( m_1(R_2) \) represents the probability of non-outburst forecasted by evidence A.

(2) Evidence B (The second evidence from the single index method).

The initial gas emission rate (\( \Delta p \)) method is employed as the second evidence. Nonetheless, it is not very precise and reasonable to determine whether coal and gas outbursts occur simply depending on a fixed critical value (Miao, 2009). Actually, this boundary range is a fuzzy critical region, and the predictive results of the discernment framework are fuzzy sets. In addition, determining the BPAs of D-S evidence theory is still an open issue. Therefore, we take \( R_1 \) and \( R_2 \) as two fuzzy sets and construct the membership functions of \( R_1 \) and \( R_2 \) according to the critical values of the single index method in this coal and gas outburst prediction (with the graph of fuzzy membership shown in Figure 4)

\[
\omega_{R_1}(\Delta p) = \begin{cases} 
1, & \Delta p \geq 10 \\
\exp\left[-\left(\frac{\Delta p - 10}{\sigma_2^2}\right)^2\right], & \Delta p < 10
\end{cases}
\]

\[
\omega_{R_2}(\Delta p) = 1 - \omega_{R_1}(\Delta p)
\]

where \( \sigma_2^2 = 2.5 \) is an approximate standard deviation of the initial gas emission rate of evidence B deduced from recorded data. The closer the value of degree of membership \( \omega_{R_1}(\Delta p) \) is to 1, the higher the probability that \( \Delta p \) belongs to the occurrence of an outburst;

![Figure 4. The graph of fuzzy membership of two indices.](energy-exploitation-38-5-1338-fig4.png)
therefore, the fuzzy membership of $R_1$ and $R_2$ for the measured data can be determined by equations (38) and (39) and regarded as their corresponding BPAs

\[
\begin{align*}
  m_2(R_1) &= \omega_{R_1}(\Delta p) \\
  m_2(R_2) &= \omega_{R_2}(\Delta p)
\end{align*}
\]

(40)

Table 6. Comparison of LSSVM prediction results.

| Data set no. | Coal seam thickness | Coal structure | Gas content | Burial depth of coalbed | Gas pressure | Actual situation of outbursts | Prediction |
|--------------|---------------------|----------------|-------------|------------------------|--------------|------------------------------|------------|
| 1            | 2.528               | 2              | 13.545      | 561.769                | 1.226        | -1                           | -1         |
| 2            | 2.257               | 2              | 13.077      | 591.097                | 2.053        | 1                            | 1          |
| 3            | 2.198               | 2              | 8.052       | 533.959                | 2.259        | 1                            | 1          |
| 4            | 1.481               | 3              | 6.783       | 649.181                | 1.998        | 1                            | 1          |
| 5            | 2.478               | 2              | 9.532       | 471.438                | 2.286        | 1                            | 1          |
| 6            | 2.482               | 2              | 16.790      | 564.035                | 1.799        | -1                           | -1         |
| 7            | 2.424               | 2              | 13.568      | 573.967                | 1.275        | -1                           | -1         |
| 8            | 2.394               | 2              | 8.297       | 479.708                | 2.286        | 1                            | 1          |
| 9            | 2.509               | 2              | 10.154      | 438.124                | 1.883        | 1                            | 1          |
| 10           | 1.988               | 3              | 8.449       | 570.936                | 2.050        | 1                            | 1          |
| 11           | 2.448               | 2              | 9.567       | 460.840                | 2.236        | 1                            | 1          |
| 12           | 2.529               | 2              | 14.319      | 556.677                | 1.381        | -1                           | -1         |
| 13           | 2.359               | 2              | 8.103       | 484.044                | 2.286        | 1                            | 1          |
| 14           | 2.510               | 2              | 15.188      | 563.630                | 1.381        | -1                           | -1         |
| 15           | 2.481               | 2              | 15.802      | 566.745                | 1.385        | -1                           | -1         |
| 16           | 2.446               | 2              | 16.078      | 570.203                | 1.660        | -1                           | -1         |
| 17           | 2.026               | 3              | 7.680       | 560.334                | 2.052        | 1                            | 1          |
| 18           | 2.334               | 2              | 12.831      | 594.511                | 1.346        | 1                            | -1         |
| 19           | 2.547               | 2              | 13.839      | 558.984                | 1.305        | -1                           | -1         |
| 20           | 2.449               | 2              | 15.013      | 571.063                | 1.289        | -1                           | -1         |
| 21           | 2.452               | 2              | 15.037      | 569.974                | 1.433        | -1                           | -1         |
| 22           | 2.528               | 2              | 13.545      | 561.769                | 1.226        | 1                            | 1          |
| 23           | 2.476               | 2              | 16.684      | 565.970                | 1.660        | -1                           | -1         |
| 24           | 1.879               | 2              | 6.203       | 584.933                | 2.066        | 1                            | -1         |
| 25           | 2.296               | 2              | 14.286      | 605.465                | 1.756        | 1                            | 1          |
| 26           | 1.466               | 3              | 5.616       | 629.505                | 2.076        | 1                            | 1          |
| 27           | 2.468               | 2              | 15.806      | 567.600                | 1.496        | -1                           | -1         |
| 28           | 1.831               | 2              | 6.511       | 588.407                | 2.070        | 1                            | 1          |
| 29           | 2.522               | 2              | 15.218      | 557.347                | 1.415        | -1                           | -1         |
| 30           | 1.361               | 3              | 5.899       | 656.446                | 2.051        | 1                            | -1         |
| 31           | 2.490               | 2              | 13.606      | 565.107                | 1.179        | -1                           | -1         |
| 32           | 1.715               | 2              | 5.159       | 601.279                | 2.049        | 1                            | 1          |
| 33           | 2.350               | 2              | 17.110      | 587.714                | 1.984        | 1                            | -1         |
| 34           | 2.374               | 3              | 18.659      | 585.470                | 2.437        | 1                            | 1          |
| 35           | 2.523               | 2              | 15.274      | 554.695                | 1.474        | -1                           | -1         |
| 36           | 2.431               | 2              | 13.690      | 572.837                | 1.269        | -1                           | -1         |
| 37           | 2.537               | 2              | 14.274      | 552.234                | 1.318        | 1                            | -1         |

Note: Non-outburst (-1), occurrence of outburst (1).
Hence, the second evidence with its BPAs for D-S evidence is derived (shown in Table 8), where $m_2(R_1)$ represents the probability of outburst occurrence, and $m_2(R_2)$ represents the probability of non-outburst forecasted by evidence B.

(3) Evidence C (The third evidence from the comprehensive index method).

This study employs a comprehensive index $K$ to execute the prediction of coal and gas outbursts. Similar to the acquisition method of the second evidence, it is necessary to construct the membership function

$$\omega_{R_1}(K) = \begin{cases} 1, & K \geq 15 \\ \exp \left[ - \left( \frac{K-15}{\sigma_2^2} \right)^2 \right], & K < 15 \end{cases}$$  \hspace{1cm} (41)

$$\omega_{R_2}(K) = 1 - \omega_{R_1}(K)$$  \hspace{1cm} (42)

where $\sigma_2^2 = 2.25$ is an approximate standard deviation of $K$ of evidence C deduced from recorded data. Thus, the fuzzy membership and their corresponding BPAs are determined by (with the graph of fuzzy membership shown in Figure 4)

$$\begin{cases} m_3(R_1) = \omega_{R_1}(K) \\ m_3(R_2) = \omega_{R_2}(K) \end{cases}$$  \hspace{1cm} (43)

| Data set no. | Actual situation of outbursts | Prediction | $m_1(R_1)$ | $m_1(R_2)$ |
|--------------|-------------------------------|------------|------------|------------|
| 1            | 1                             | 1          | 0.99950981 | 0.00049019 |
| 2            | -1                            | -1         | 0.011722725| 0.988277275|
| 3            | 1                             | 1          | 0.999813211| 0.000186789|
| 4            | 1                             | 1          | 0.980144009| 0.019855991|
| 5            | 1                             | 1          | 0.999993328| 6.67201E-06|
| 6            | -1                            | -1         | 1.44E-07   | 0.99999856 |
| 7            | 1                             | 1          | 0.968130528| 0.031869472|
| 8            | 1                             | 1          | 0.999999165| 8.34742E-07|
| 9            | 1                             | 1          | 0.999982708| 1.72921E-05|
| 10           | -1                            | -1         | 7.97E-07   | 0.99999203 |
| 11           | 1                             | 1          | 0.999987324| 1.26761E-05|
| 12           | 1                             | 1          | 0.999982409| 1.75906E-05|
| 13           | 1                             | 1          | 0.999954713| 4.52866E-05|
| 14           | -1                            | -1         | 4.13E-07   | 0.99999587 |
| 15           | 1                             | 1          | 0.999999455| 5.45168E-07|
| 16           | 1                             | 1          | 0.999571299| 0.000428701|
| 17           | 1                             | 1          | 0.51336827 | 0.48663173 |
| 18           | 1                             | 1          | 0.99990048 | 9.9522E-06 |
| 19           | 1                             | 1          | 0.999967892| 3.2108E-05 |
| 20           | -1                            | -1         | 6.85E-06   | 0.999993154|

Table 7. Prediction and BPAs of each point from evidence A.
The third evidence with its BPAs for D-S evidence is shown in Table 8, where $m_3(R_1)$ represents the probability of outburst occurrence, and $m_3(R_2)$ represents the probability of non-outburst forecasted by evidence C.

**Table 8. BPAs of each point from evidence B and evidence C.**

| Data set no. | Initial gas emission rate | $m_2(R_1)$ | $m_2(R_2)$ | $K$ | $m_3(R_1)$ | $m_3(R_2)$ |
|--------------|--------------------------|------------|------------|-----|------------|------------|
| 1            | 14.654                   | 1          | 0          | 36.63829745 | 1          | 0          |
| 2            | 7.67591                  | 0.42138    | 0.57862    | 13.81928854 | 0.75929    | 0.24071    |
| 3            | 20.856                   | 1          | 0          | 41.99124176 | 1          | 0          |
| 4            | 17.5362                  | 1          | 0          | 34.60290775 | 1          | 0          |
| 5            | 19.4583                  | 1          | 0          | 64.60130276 | 1          | 0          |
| 6            | 7.26853                  | 0.30308    | 0.69692    | 13.10067571 | 0.49038    | 0.50962    |
| 7            | 23.6171                  | 1          | 0          | 44.30967296 | 1          | 0          |
| 8            | 19.6134                  | 1          | 0          | 44.07297633 | 1          | 0          |
| 9            | 18.9978                  | 1          | 0          | 46.91278151 | 1          | 0          |
| 10           | 7.22659                  | 0.29209    | 0.70791    | 13.10244858 | 0.49103    | 0.50897    |
| 11           | 21.0039                  | 1          | 0          | 64.02808177 | 1          | 0          |
| 12           | 24.0899                  | 1          | 0          | 118.0912184 | 1          | 0          |
| 13           | 18.8848                  | 1          | 0          | 35.08614171 | 1          | 0          |
| 14           | 6.9959                   | 0.236      | 0.764      | 12.5035075  | 0.29197    | 0.70803    |
| 15           | 19.9482                  | 1          | 0          | 52.66058969 | 1          | 0          |
| 16           | 9.57446                  | 0.97144    | 0.02856    | 24.52103806 | 1          | 0          |
| 17           | 17.3877                  | 1          | 0          | 41.86276634 | 1          | 0          |
| 18           | 19.4172                  | 1          | 0          | 47.95957181 | 1          | 0          |
| 19           | 22.0814                  | 1          | 0          | 77.35808074 | 1          | 0          |
| 20           | 7.24794                  | 0.29766    | 0.70234    | 13.06434169 | 0.47706    | 0.52294    |

The third evidence with its BPAs for D-S evidence is shown in Table 8, where $m_3(R_1)$ represents the probability of outburst occurrence, and $m_3(R_2)$ represents the probability of non-outburst forecasted by evidence C.

**Prediction by multisource information fusion.** Finally, we use the improved combination rule proposed in the section Dempster-Shafer theory and its improvement to fuse the three predicted results and then analyse the probability of outburst risk. The predicted results are listed in Table 9, where $m(R_1)$ represents the probability of outburst occurrence, and $m(R_2)$ represents the probability of non-outbursts converged by improved D-S evidence, showing that the prophetic model is capable of forecasting the coal and gas outbursts of the No. 3 coal seam in the study area. According to the field experience of coal and gas outbursts, the points with the probability of $m(R_1)$ greater than 0.6 are classified as the occurrence of outbursts, and the points with the probability of $m(R_2)$ greater than 0.6 as non-outbursts. The fused prediction demonstrates that there will be 15 sites predicted to be safe and 5 sites predicted to have outbursts, and all of these predictions are consistent with the actual situation in the mine.

From the comparative table of 20 samples, we can learn that the advanced D-S evidence theory assesses four positions of non-outburst where all three preliminary forecasting methods also predict that no outburst occurs, including points No. 6, No. 10, No. 14, and No. 20. In particular, three means indicate that there is no outburst at point 2, but the approximation of the comprehensive index at the critical value will mislead the judgement. Therefore, after the convergence, the possibility of non-outbursts here is 0.6957, which provides convincing evidence for decisions. The points where coal and gas outbursts will appear can also
be divided into two cases. The predictive credibility of points that are difficult to identify, i.e., at points No. 16 and No. 17, increases after the application of D-S evidence theory. In the remaining 13 sites, where the symptoms of outbursts are obvious, the probability of outburst existence is forecasted to be 1 as a good reflection of the strong agreement among the three initial methods.

It is worth mentioning that the prediction for points No. 16 and No. 17 in particular presents the advantage of multisource information fusion by the improved D-S evidence theory. By congregating the multisource evidence, the estimation is consistent with the real situation with great certainty so that the prediction can be significantly improved by “compensating” the incorrect solution. If only the solitary index method was adopted, the initial gas emission rate at point No. 16 is smaller than the critical value in the provisions, while its BPA is greater than 0.5 after the assignment of the membership function. Likewise, the LSSVM forecasts that there will be coal and gas outbursts at point No. 17, but its Bayesian posterior class probability is only 0.51337, causing slight ambiguity for judgement. After the application of D-S evidence theory, the prediction result supports the occurrence with a probability of 0.62858, more consistent with the actual geological situation.

Therefore, compared with the existing predictive methods, which ordinarily focus on individual models and are limited in their flexibility to variable critical values, the model presented in this paper shows higher accuracy and better stability, effectively providing a sound way to solve the uncertain information of coal and gas outbursts.

Furthermore, to test the performance of the proposed approach, the present work compares it with two other single machine learning algorithms, i.e., LSSVM and artificial neural

### Table 9. Final prediction by D-S evidence theory.

| Data set no. | m₁(R₁) | m₂(R₁) | m₃(R₁) | m₁(R₂) | m₂(R₂) | m₁(R₁), R₂) | Actual situation of outbursts |
|--------------|--------|--------|--------|--------|--------|-------------|-----------------------------|
| 1            | 0.99951 | 1      | 1      | 1      | 0      | 0           | Outburst                   |
| 2            | 0.01172 | 0.42138| 0.75929| 0.3043 | 0.6957 | 0           | Non-outburst               |
| 3            | 0.99981 | 1      | 1      | 1      | 0      | 0           | Outburst                   |
| 4            | 0.98014 | 1      | 1      | 1      | 0      | 0           | Outburst                   |
| 5            | 0.99999 | 1      | 1      | 1      | 0      | 0           | Outburst                   |
| 6            | 1.4E-07 | 0.30308| 0.49038| 0.06784| 0.93216| 0           | Non-outburst               |
| 7            | 0.96813 | 1      | 1      | 1      | 0      | 0           | Outburst                   |
| 8            | 1      | 1      | 1      | 1      | 0      | 0           | Outburst                   |
| 9            | 0.99998 | 1      | 1      | 1      | 0      | 0           | Outburst                   |
| 10           | 8E-07   | 0.29209| 0.49103| 0.06266| 0.93734| 0           | Non-outburst               |
| 11           | 0.99999 | 1      | 1      | 1      | 0      | 0           | Outburst                   |
| 12           | 0.99998 | 1      | 1      | 1      | 0      | 0           | Outburst                   |
| 13           | 0.99995 | 1      | 1      | 1      | 0      | 0           | Outburst                   |
| 14           | 4.1E-07 | 0.236  | 0.29197| 0.01478| 0.98522| 0           | Non-outburst               |
| 15           | 1      | 1      | 1      | 1      | 0      | 0           | Outburst                   |
| 16           | 0.99957 | 0.97144| 1      | 1      | 3.11E-09| 0           | Outburst                   |
| 17           | 0.51337 | 1      | 1      | 0.62858| 0.14416| 0.22726     | Outburst                   |
| 18           | 0.99999 | 1      | 1      | 1      | 0      | 0           | Outburst                   |
| 19           | 0.99997 | 1      | 1      | 1      | 0      | 0           | Outburst                   |
| 20           | 6.8E-06 | 0.29766| 0.47706| 0.06166| 0.93834| 0           | Non-outburst               |
network (ANN), with the same training set and testing set. The predictions by LSSVM regarding whether coal and gas outbursts occur are presented in Table 7.

We use the classic ANN model, that is, the feed-forward neural network, for coal and gas outburst prediction. The hidden (intermediate) layer is established to one layer owing to its reasonable precision for the approximation of any continuous multivariate function. The hidden nodes of the model are adjusted by 4, 3 and 2 to test the predictive performance. The results show that the numbers of the total accurate predictions by different ANN types are 20, 5, 19 with respective accuracies of 100%, 25% and 75%, extremely unstable.

The reason why the predictions by the feed-forward neural network vary sharply is that in small-sample prediction, the optimization function tends to fall into the local optimal solution (i.e., overfitting, which has a good fitting effect on the training samples but a poor effect on the test set), and this “trap” increasingly deviates from the true global optimal. Therefore, it is necessary to manually adjust the node. LSSVM performs better with small sample datasets. More importantly, after the offset by D-S evidence, the situation of outbursts can be forecasted more convincingly.

Performance evaluation in another research area. To test the general applicability of this predictive method in Chinese coal mines with different geological conditions, we apply it to the outburst forecast in the Zhuxianzhuang coal mine, Anhui Province. This coal mine, located in northern China, is a large-scale high-gas-outburst mine, and its No. 8 coal seam has a high outburst frequency (Liu and Cheng, 2015). We conduct this predictive experiment with 45 datasets of field geological data, selecting 15 samples at random for validation and using the remaining data as the training sets of the LSSVM model for preliminary forecasting.

After the ranking of the key factors of outburst events in the Zhuxianzhuang coal mine by GRA, the indices of coal and gas outburst forecasting determined by the predictive model are as follows: the Protodyakonov coefficient of coal strength, gas content, coalbed burial depth, initial gas emission rate, coal structure, gas pressure, and gas emission amount.

As in the predictive procedure in the Xinjing coal mine, the initial gas emission rate and the Protodyakonov coefficient of coal strength are chosen for index methods, and the remaining factors are chosen for the LSSVM model. In addition, the outburst event is represented by $R_1 = [1]$, and the non-outburst event is represented by $R_2 = [-1]$, constituting the D-S frame of discernment. Each of the results of 15 test datasets is derived from the three preliminary methods, which are then taken as BPAs of the improved D-S evidence theory for the final prediction.

The predicted results are listed in Table 10, where $m(R_1)$ represents the probability of outburst occurrence, and $m(R_2)$ represents the probability of non-outburst occurrence forecasted in the Zhuxianzhuang coal mine, demonstrating that this prediction model can also be applied to the No. 8 coal seam in the Zhuxianzhuang coal mine. From the results, it can be observed that the proposed method corrects the mistake of the single index method at points 2, 9, 10 and 12 and the misjudgement of LSSVM at point 14. Thus, the final prediction determines that there will be six sites at a safe level and nine sites likely to experience outbursts, all of which are in line with the actual situation in the mine. The case study in this research area shows that our proposed forecasting method can flexibly adjust the relevant factors and fuse the different predictive results, thereby enhancing the reliability of coal and gas outburst prediction.
Conclusions and future work

Based on the work in previous studies on coal and gas outburst prediction and on machine learning theory and application, this study establishes a predictive model by integrating the multisource information of gas geology and predictive results from different methods, achieving the following conclusions.

1. By executing a ranking of the degrees of correlation between different factors and outburst events, the model first determined some dominant parameters as input variables. The flexible index-selection method by GRA can achieve dimension reduction and individual adaptation to specific geological conditions.

2. Taking advantage of LSSVM in the data processing of high-dimension and small samples, we trained the model and subsequently predicted coal and gas outbursts through the LSSVMLab MATLAB toolbox.

3. Three forecasting models (single index method, comprehensive index method and LSSVM) were applied to preliminarily predict coal and gas outbursts. However, there were still some inconsistencies for some forecasted points, which needed further prediction.

4. The numerical experiments demonstrated that the improved D-S evidence theory performed better than the traditional D-S combination rule in defining a new concept of evidence distance ($\text{conf}(m_i, m_j)$) and modifying the evidence source.

5. New definitions of BPAs from different preliminary forecasting methods were proposed, and the improved D–S evidence theory was finally employed in outburst prediction. The fusing forecasts were consistent with the record of the actual outburst conditions, showing the good performance of the model in improving prediction reliability. The analysis of the results demonstrated that the probability obtained by the novel predictive model was more responsible and reasonable than that obtained by a solitary prediction method, avoiding incorrect judgement as a result of an accidental error. The experiments

| Data set no. | $\Delta p$ (mmHg) | $K$ | $m_1(R_1)$ | $m_2(R_1)$ | $m_3(R_1)$ | $m(R_1)$ | $m(R_2)$ | $m(R_1, R_2)$ | Actual situation of outbursts |
|-------------|-----------------|----|-------------|-------------|-------------|-----------|-----------|--------------|-----------------------------|
| 1           | 4.651           | 8.395 | 0.00788     | 0.01027     | 0.00018     | 3.01E-07  | 1         | 0            | Outburst                  |
| 2           | 8.235           | 27.92  | 0.95778     | 0.60751     | 1           | 0.99928   | 0.00072   | 0            | Non-outburst              |
| 3           | 4.65            | 8.185  | 3.4E-05     | 0.01026     | 0.0001      | 9.46E-11  | 1         | 0            | Outburst                  |
| 4           | 4.468           | 8.452  | 0.04527     | 0.00748     | 0.00021     | 1.54E-06  | 1         | 0            | Outburst                  |
| 5           | 10.67           | 27.94  | 0.92494     | 1           | 1           | 0         | 0         | 0            | Outburst                  |
| 6           | 4.723           | 8.403  | 0.00037     | 0.01161     | 0.00018     | 1.60E-09  | 1         | 0            | Non-outburst              |
| 7           | 4.436           | 8.292  | 0.01172     | 0.00705     | 0.00014     | 2.91E-07  | 1         | 0            | Outburst                  |
| 8           | 4.588           | 8.868  | 0.09952     | 0.00921     | 0.00059     | 5.89E-06  | 0.99999   | 0            | Outburst                  |
| 9           | 8.011           | 70.93  | 0.97158     | 0.53108     | 1           | 0.99956   | 0.00044   | 0            | Outburst                  |
| 10          | 9.504           | 22.04  | 0.9102      | 0.96147     | 1           | 0.99992   | 7.87E-05  | 0            | Non-outburst              |
| 11          | 12              | 85.59  | 0.81348     | 1           | 1           | 0         | 0         | 0            | Outburst                  |
| 12          | 9.717           | 22.09  | 0.91586     | 0.98726     | 1           | 0.99999   | 8.48E-06  | 0            | Outburst                  |
| 13          | 11.98           | 85.06  | 0.80657     | 1           | 1           | 0         | 0         | 0            | Outburst                  |
| 14          | 11.94           | 82.5   | 0.48719     | 1           | 1           | 0.61391   | 0.14892   | 0.23717      | Non-outburst              |
| 15          | 10.42           | 22.81  | 0.91848     | 1           | 1           | 0         | 0         | 0            | Outburst                  |
conducted in both the Xinjing coal mine and the Zhuxianzhuang coal mine also revealed that the presented model outperforms the previously used models due to its flexibility in the selection of the main controlling factors and fusion by the improved D-S evidence theory.

Although we successfully demonstrate the potential of multisource information fusion technology in coal and gas prediction, there are still some limitations to be addressed in the future. Since we do not have enough data to carry out experiments due to the difficulty in obtaining mine data, it is necessary to analyse and apply this method to additional geological conditions. This is the major obstacle to the study. In addition, this paper simply uses the recorded historical data to analyse and forecast outbursts. Many researchers now apply time series for coal and gas outburst prediction. Therefore, we will consider studying field data in chronological order in future work.

Acknowledgements
The MATLAB experiment framework used LSSVMLab MATLAB toolbox download from http://www.esat.kuleuven.be/sista/lssvmlab/, a special thanks here. The authors also gratefully acknowledge anonymous reviewers for their constructive comments.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Projects of the National Science Foundation of China (Grant Nos. 41872248, 41430317, and 41672324), the Priority Academic Program Development of Jiangsu Higher Education Institutions, the Natural Science Foundation of the Jiangsu Higher Education Institutions (Grant No. 19KJB520029) and the Planning Subject for the 13th Five Year Plan of Jiangsu Education Sciences (C-a/2018/03/02).

ORCID iD
Yongguo Yang https://orcid.org/0000-0003-4290-7383

References
Cai YD, Liu XJ, Xu XB, et al. (2002) Prediction of protein structural classes by support vector machines. Computers & Chemistry 26(3): 293–296.
Chen XJ, Li LY, Wang L, et al. (2019) The current situation and prevention and control countermeasures for typical dynamic disasters in kilometer-deep mines in China. Safety Science 115: 229–236.
China’s State Administration of Work Safety (2009) Provisions on Prevention of Coal and Gas Outburst. Beijing: Coal Industry Press (in Chinese).
Dempster AP (2008) Upper and lower probabilities induced by a multivalued mapping. In: Yager RR and Liu L (eds) Classic Works of the Dempster-Shafer Theory of Belief Functions. Berlin: Springer Berlin Heidelberg, pp. 57–72.
Deng Y, Shi WK, Zhu ZF, et al. (2004) Combining belief functions based on distance of evidence. *Decision Support Systems* 38(3): 489–493.

Díaz Aguado B and González Nicieza C (2007) Control and prevention of gas outbursts in coal mines, Riosa–Olloniego coalfield, Spain. *International Journal of Coal Geology* 69(4): 253–266.

Dou XM and Yang YG (2018) Evapotranspiration estimation using four different machine learning approaches in different terrestrial ecosystems. *Computers and Electronics in Agriculture* 148: 95–106.

Dou XM, Yang YG and Luo JH (2018) Estimating forest carbon fluxes using machine learning techniques based on eddy covariance measurements. *Sustainability* 10: 203–229.

Dymova L, Sevastjanov P and Kaczmarek K (2016) A Forex trading expert system based on a new approach to the rule-base evidential reasoning. *Expert Systems with Applications* 51: 1–13.

Feizizadeh B, Jankowski P and Blaschke T (2014) A GIS based spatially-explicit sensitivity and uncertainty analysis approach for multi-criteria decision analysis. *Computers & Geosciences* 64: 81–95.

Ghasemi J, Ghaderi R, Karami Mollaei MR, et al. (2013) A novel fuzzy Dempster–Shafer inference system for brain MRI segmentation. *Information Sciences* 223: 205–220.

Haouas F, Solaiman B, Ben Dhiaf Z, et al. (2019) Multi-temporal image change mining based on evidential conflict reasoning. *ISPRS Journal of Photogrammetry and Remote Sensing* 151: 59–75.

Hsu CW and Lin CJ (2002) A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks* 13(2): 415–425.

Hua SJ and Sun ZR (2001) A novel method of protein secondary structure prediction with high segment overlap measure: Support vector machine approach. *Journal of Molecular Biology* 308(2): 397–407.

Islam MR and Shinjo R (2009) Numerical simulation of stress distributions and displacements around an entry roadway with igneous intrusion and potential sources of seam gas emission of the Barapukuria coal mine, NW Bangladesh. *International Journal of Coal Geology* 78(4): 249–262.

Jac Fredo AR, Abilash RS, Femi R, et al. (2019) Classification of damages in composite images using Zernike moments and support vector machines. *Composites Part B: Engineering* 168: 77–86.

Jia TR, Feng ZD, Wei GY, et al. (2017) Shear deformation of fold structures in coal measure strata and coal-gas outbursts: Constraint and mechanism. *Energy Exploration & Exploitation* 36(2): 185–203.

Jiang S, Lian MJ, Lu CW, et al. (2018) SVM–DS fusion based soft fault detection and diagnosis in solar water heaters. *Energy Exploration & Exploitation* 37(3): 1125–1146.

Jiang W, Xie CH, Zhuang M, et al. (2017) Failure mode and effects analysis based on a novel fuzzy evidential method. *Applied Soft Computing* 57: 672–683.

Jousselme AL, Grenier D and Bossé É (2001) A new distance between two bodies of evidence. *Information Fusion* 2(2): 91–101.

Kissell FN and Iannacchione AT (2014) Chapter 9 – Gas outbursts in coal seams. In: Thakur P, Schatzel S and Aminian K (eds) *Coal Bed Methane* Oxford: Elsevier, pp. 177–184.

Lama RD and Bodziony J (1998) Management of outburst in underground coal mines. *International Journal of Coal Geology* 35(1): 83–115.

LeFevre E, Colot O and Vannoorenberghe P (2002) Belief function combination and conflict management. *Information Fusion* 3(2): 149–162.

Li FL, Jiang B, Cheng GX, et al. (2019) Structural and evolutionary characteristics of pores-microfractures and their influence on coalbed methane exploitation in high-rank brittle tectonically deformed coals of the Yangquan mining area, northeastern Qinshui basin, China. *Journal of Petroleum Science and Engineering* 174: 1290–1302.

Li XW, Jiang CL, Tang J, et al. (2017) A Fisher’s criterion-based linear discriminant analysis for predicting the critical values of coal and gas outbursts using the initial gas flow in a borehole. *Mathematical Problems in Engineering* 2017: 1–11.
Liu HB and Cheng YP (2015) The elimination of coal and gas outburst disasters by long distance lower protective seam mining combined with stress-relief gas extraction in the Huaiabei coal mine area. *Journal of Natural Gas Science and Engineering* 27: 346–353.

Liu WR (2006) Analyzing the degree of conflict among belief functions. *Artificial Intelligence* 170(11): 909–924.

Liu ZJ, Wu D, Liu YW, et al. (2019) Accuracy analyses and model comparison of machine learning adopted in building energy consumption prediction. *Energy Exploration & Exploitation* 37(4): 1426–1451.

Manochandar S and Punniyamoorthy M (2018) Scaling feature selection method for enhancing the classification performance of Support Vector Machines in text mining. *Computers & Industrial Engineering* 124: 139–156.

Miao YZ (2009) *Study on multi-sensor information fusion and its application on the gas outburst prediction system*. PhD Thesis, China University of Mining and Technology, China.

Murphy CK (2000) Combining belief functions when evidence conflicts. *Decision Support Systems* 29(1): 1–9.

Pelekkmans K, Suykens JAK, Gestel TV, et al. (2002) LS-SVMLab: A Matlab/c toolbox for least squares support vector machines. Available at: www.esat.kuleuven.be/sista/lssvmlab/ (accessed 4 March 2020).

Qiu SQ, Sallak M, Schön W, et al. (2018) A valuation-based system approach for risk assessment of belief rule-based expert systems. *Information Sciences* 466: 323–336.

Rahmani H, Scanlan C, Nadeem U, et al. (2019) Automated segmentation of gravel particles from depth images of gravel-soil mixtures. *Computers & Geosciences* 128: 1–10.

Ristic B and Smets P (2006) The TBM global distance measure for the association of uncertain combat ID declarations. *Information Fusion* 7(3): 276–284.

Shafer G (1976) A theory of statistical evidence. In: Harper WL and Hooker CA (eds) *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*. Dordrecht: Springer Netherlands, pp. 365–436.

Sun Q, Xq Y and Gu WK (2000) A new combination rule of evidence theory. *Acta Electronica Sinica* 28(8): 117–119.

Suykens JAK and Vandewalle J (1999) Least squares support vector machine classifiers. *Neural Processing Letters* 9(3): 293–300.

Suykens JAK, De Brabanter J, Lukas L, et al. (2002) Weighted least squares support vector machines: Robustness and sparse approximation. *Neurocomputing* 48(1): 85–105.

Tang J, Jiang CL, Chen YJ, et al. (2016) Line prediction technology for forecasting coal and gas outbursts during coal roadway tunneling. *Journal of Natural Gas Science and Engineering* 34: 412–418.

Vapnik VN (2000) *The Nature of Statistical Learning Theory*. New York: Springer-Verlag.

Wang C, Wang EY, Xu JK, et al. (2010) Bayesian discriminant analysis for prediction of coal and gas outbursts and application. *Mining Science and Technology (China)* 20(4): 520–523.

Wang X, Li Y, Chen TJ, et al. (2017) Quantitative thickness prediction of tectonically deformed coal using extreme learning machine and principal component analysis: A case study. *Computers & Geosciences* 101: 38–47.

Wang YA (1994) *Technical Manuals of Coal Mine Safety*. Beijing: China Coal Industry Press. [In Chinese].

Wei CH, Qin Y, Wang GGX, et al. (2007) Simulation study on evolution of coalbed methane reservoir in Qinshui basin, China. *International Journal of Coal Geology* 72(1): 53–69.

Wen CL, Wang YC, Xu XB (2008) Fuzzy information fusion algorithm of fault diagnosis based on similarity measure of evidence. In: Sun FC, Zhang JW, Tan Y, et al. (eds) *5th International Symposium on Neural Networks*, Beijing, China, 24–28 September 2008, pp. 506–515. Berlin: Springer Berlin Heidelberg.
Wold MB, Connell LD and Choi SK (2008) The role of spatial variability in coal seam parameters on gas outburst behaviour during coal mining. *International Journal of Coal Geology* 75(1): 1–14.

Xiong F, Wang DF, Zhang S, et al. (2018) Lightweight optimization of the side structure of automobile body using combined grey relational and principal component analysis. *Structural and Multidisciplinary Optimization* 57(1): 441–461.

Yager RR (1987) On the Dempster-Shafer framework and new combination rules. *Information Sciences* 41(2): 93–137.

Yang JP, Huang HZ, Miao Q, et al. (2011) A novel information fusion method based on Dempster-Shafer evidence theory for conflict resolution. *Intelligent Data Analysis* 15: 399–411.

Yang YJ, Forrest J and Liu SF (2012) A brief introduction to grey systems theory. *Grey Systems: Theory and Application* 2(2): 89–104.

Zhang HW, Song WH, Yang H, et al. (2008) Pattern recognition prediction of coal and gas outburst hazard in the sixth mine of Hebi. *Journal of Coal Science and Engineering (China)* 14(2): 248–251.

Zhang KP, Jiang B, Li M, et al. (2016) Identification and distribution of structure of seam No. 3 in Xinjing Mine on the basis of well logs. *Coal Geology & Exploration* (1): 123–127.

Zhang RL and Lowndes IS (2010) The application of a coupled artificial neural network and fault tree analysis model to predict coal and gas outbursts. *International Journal of Coal Geology* 84(2): 141–152.

Zhang YC and Yang YG (2015) Research of risk evaluation model for CBM development based on multi-information fusion. *The Open Fuels & Energy Science Journal* 8(1): 337–340.

Zhao ZG, Yan JP and Liu X (2014) Research on gas occurrence regularity in Fengcheng Mining Area of Jiangxi Province. *Energy Exploration & Exploitation* 32(3): 591–599.

Zhao W, Cheng YP, Pan ZJ, et al. (2019) Gas diffusion in coal particles: A review of mathematical models and their applications. *Fuel* 252: 77–100.

Zhou ZH (2016) *Machine Learning*. Beijing: Tsinghua University Press.

Zhou BZ, Yang SQ, Wang CJ, et al. (2019) Experimental study on the influence of coal oxidation on coal and gas outburst during invasion of magmatic rocks into coal seams. *Process Safety and Environmental Protection* 124: 213–222.

Zuo RG and Carranza EJM (2011) Support vector machine: A tool for mapping mineral prospectivity. *Computers & Geosciences* 37(12): 1967–1975.