To investigate the effect of different proximate index on minimum ignition temperature (MIT) of coal dust cloud, 30 types of coal specimens with different characteristics were chosen. A two-furnace automatic coal proximate analyzer was employed to determine the indexes for moisture content, ash content, volatile matter, fixed carbon and MIT of different types of coal specimens. As the calculated results showed that these indexes exhibited high correlation, a principal component analysis (PCA) was adopted to extract principal components for multiple factors affecting MIT of coal dust, and then, the effect of the indexes for each type of coal on MIT of coal dust was analyzed. Based on experimental data, support vector machine (SVM) regression model was constructed to predicate the MIT of coal dust, having a predicing error below 10%. This method can be applied in the predication of the MIT for coal dust, which is beneficial to the assessment of the risk induced by coal dust explosion (CDE).

Keywords: Coal dust explosion, minimum ignition temperature, principal component analysis, SVM predication
wysokości minimalnej temperatury zapłonu, co jest ważnym aspektem w ocenie ryzyka wybuchu pyłu węglowego.

Słowa kluczowe: wybuch pyłu węglowego, minimalna temperatura zapłonu, analiza podstawowych składników, prognozowanie z wykorzystaniem regresji metodą wektorów nośnych

1. Introduction

CDE disaster is one of major calamity accidents happened in coal mines. The coal dust cloud involving air and coal dust are often encountered in the mining process of coal mine, and rapidly oxidized under the effect of heat sources including high temperature or a certain ignition energy, eventually, explosion is produced. The MIT of coal dust is an important characteristic parameter of CDE and an important index used in evaluating the sensitivity of CDE, also provides an important basis for coal dust and explosion prevention (Babrauskas, 2003; Cashdollar, 2000; Deng, 2014; Eckhoff, 2003).

Illustrative case studies and past accident analyses reflect the high frequency, geographic spread, and damage potential of dust explosions across the world. The sources and triggers of dust explosions, and the measures with which different factors associated with dust explosions can be quantified are reviewed alongside dust explosion mechanism by Abbasi (2007). One hundred and six coal dust explosion accidents that occurred in China between the years of 1949 and 2007 were investigated through statistical methods so as to review the overall situation and provide quantitative information on coal dust explosions. Statistical characteristics about accident-related factors such as space, time, volatile ratio of coal dust, ignition sources, and accident categories were analyzed by Zheng et al. (2009). Zhang (1993) extended the energy equations of Krishna and Berlad to include coal volatile matter (VM) release and the homogeneous ignition of VM in the gas phase of the cloud. The effects of particle size on the ignition temperature and ambient oxygen concentration on ignition mechanisms, predicted by this extension, are then examined. The influence of humidity on dust explosions of metallic and organic materials was studied by TRaoré (2009). Nifuku (2006, 2007) investigated the ignitability of aluminium and magnesium dusts, the relations between particle size and the minimum explosive concentration, the minimum ignition energy, the ignition temperature of the dust cloud, etc. were studied experimentally. An investigation of the ignition behaviour of iron sulphide dusts has been undertaken by Amyotte (2003), Commercial samples of FeS and FeS 2 and mine samples of pyrrhotite and pyrite were tested for minimum ignition temperature (MIT) using a device known as the BAM oven. The MIT of three coal dusts was studied by Wu et al. (2016, 2014), a modified steady-state mathematical model based on heterogeneous reaction was presented to interpret the observed experimental phenomena and estimate the ignition mechanism of coal dust cloud under MIT conditions. An experimental study into the hot surface ignition of coal dust layers was conducted by Prabhakar et al. (1998), the effect of admixed the dust layer ignition temperature was analyzed using a steady-state thermal explosion model. Danzi et al. (2015) conducted the experiment of MITs of the mixing of two different combustible dusts by using a Godbert Greenwald furnace and a hot plate, the experiment data showed that MITs increased as the inert content was increased. Li et al. (2011) studied the coal dust explosion characteristics of anthracite, bituminous coal and lignite, the characteristic parameters of coal dust explosion under different experimental conditions were obtained, and the quantitative evaluation was conducted. The minimum ignition temperatures
(MIT) of hybrid mixtures had been investigated by Addai et al. (2016, 2016) performing several series of tests in a modified Godbert-Greenwald furnace. Further, seven mathematical models for prediction of the minimum ignition temperatures (MIT) of dust/air mixtures were presented of which three were selected for further study and verified by the experimental results based on the availability of the input quantities needed and their applicability. Mittal and Guha (1997) presented a model developed for determining the minimum ignition temperature for an organic dust cloud, polyethylene, simulating the conditions in the furnace. The model correlates the particle size, as well as the dust concentration with the minimum ignition temperatures, the results of the computations were compared with the experimental values in his study. Many groups of Mg-Al alloy dust were used to test MIT by Wang (2016) using Godbert-Greenwald furnace, based on the nonlinear fitting in regression analysis and evaluation of six indexes, the model equation of the influence of concentration and particle size on MIT was fitted by R language and 1st Opt.

Most scholars only carry out qualitative analysis around the minimum ignition temperature of coal dust cloud and the factors affecting the minimum ignition temperature of coal dust cloud, and the quantitative research such as principal component analysis method was ignored.

If the minimum ignition temperature of coal dust cloud can be quantitatively controlled by changing the composition of coal dust, it will be helpful to the study of preventing coal dust explosion and coal dust explosion suppressor and so on. In this paper, the MIT measurement device of coal dust cloud and two-furnace automatic coal proximate analyzer were adopted to investigate the MIT of coal dust cloud and the indexes of proximate analysis, fitting and analyzing data by PCA and SVM regression analysis.

2. Experiments and methods

2.1. The collection and preparation of coal specimens

To satisfy the universality demands of experimental results, 30 types of coal specimens with different characteristics were chosen and numbered successively. The specimens were obtained from the advancing working face and mining face in the coal mines of different regions.

a. The surface of the coal seam was cleaned before sampling, and then two parallel lines were drawn along the vertical direction of coal seam. The vertical distance between two parallel lines is 150 mm when the thickness of the coal seam is smaller than 1 m, while it is 100 mm when the thickness of the coal specimens exceeded 1 m.

b. The sites for storing the coal specimens were cleaned so as to be free from the pollution.

c. The obtained samples after removing gangue were packaged for use.

Subjected to relevant national standards aforementioned, a jaw crusher and multi-functional hermetically-sealed sample preparation machine were used to conduct ultrafine grinding on coal specimens, as demonstrated in Fig. 1. Afterwards, an automatic sieving machine was utilized to screen the coal dust with a grain size of smaller than 0.074 mm, and then, three-slot receptacle was adopted to split the samples into three parts, with an aim for determination of the indexes for the MIT and the proximate analysis of coal samples, as well as the storage of coal samples.

Before the experiment, the coal sample was placed in 105°C constant temperature drying box for two hours to remove the external moisture, and the dried sample was placed in the silica gel drying dish.
At the same time, the laboratory was equipped with the experimental screens with a aperture of 58 μm ~ (200) mesh, 48 μm ~ (300) mesh, 38 μm ~ (400) mesh, and 25 μm ~ (500) mesh. Mesh number is the approximate number of sieve openings per linear inch. The aperture of screen mesh decreases with the increase of mesh number. Because of the difference of the opening rate and the thickness of the mesh, the national standards and regulations are different.

Fig. 1. The coal specimens and corresponding crushing device

2.2. The test of the MIT for coal dust

According to the standard of the determination of the minimum ignition temperature of dust cloud for GB/T16429-1996 Coal samples, FCY-II dust cloud MIT measurement device is used, as shown in Fig. 2. This instrument adopts photoelectric detection technology and intelligent dust spray control system. The measurement of ignition temperature of coal dust cloud is carried out automatically. The schematic diagram of the device is shown in figure 3. The device consists of a heating furnace, pressure dusting system, temperature control system and temperature recording system. The furnace is installed in a hood that is not affected by air flow and can pump out coal dust and toxic and harmful gases. The heated quartz in the heating furnace is installed vertically, and the external wall has a total resistance of 13 Ω on heating electrical wire, of which the middle and lower parts are respectively provided with thermocouples connected to the temperature controller and the temperature recorder. the lower end of the heated quartz is connected with the atmosphere, and the upper end is connected to the dust storage device through an adapter. The furnace is mounted on a support seat with a mirror mounted on the bottom of the heated quartz.

Fig. 2. Minimum ignition temperature meter of dust cloud
tube to observe the internal condition of the heated quartz tube. The pressure dusting system consists of a small air compressor, a gas storage tank (500 mL), U-tube, solenoid valve and dust storage device. The air compressor is connected to the solenoid valve to control and record the temperature of heating furnace by temperature controller and temperature recorder, and to control the dust injection pressure through U-tube.

![Diagram of dusting system](image)

**Fig. 3. Principle diagram of minimum ignition temperature measurement device for dust clouds.**

1. heating furnace; 2. connecting head; 3. dust accumulator; 4. solenoid valve; 5. gas storage tank; 6. gate valve; 7. U-tube; 8. stable power supply; 9. temperature control instrument; 10. temperature recorder

### 2.2.1. Test method for minimum ignition temperature of coal dust cloud

In the experiment, 0.1 g coal dust was loaded into the dust storage device with electronic balance weighing, and the temperature of the heating furnace was set to 500°C, and the dust injection pressure was set to 10 kPa. When the solenoid valve was opened, the dust was injected into the furnace, and the burning situation of coal dust cloud was observed by reflecting mirror. If there was no ignition, raising the furnace temperature gradually at the step of 50°C, and redoing the experiment with the coal dust of the same quality until the furnace temperature reaches the upper limit of the instrument.

When a tongue of flame was observed, changing the coal dust quality and dust injection pressure until the most intense ignition occurred. When coal dust cloud was on fire at its most intense, keeping the coal dust mass and jet pressure constanted, and reducing the temperature of the heating furnace gradually at the step of 20°C, until there were 10 experiments without flames. According to GB / T 16429-1996, in the experiment, the quanlity of coal dust was selected from 0.01 g, 0.02 g, 0.03 g, 0.05 g, 0.1 g, 0.2 g, 0.3 g, 0.5 g, 1 g, 2 g and 3 g, and the dust pressure was selected from 2 kPa, 3 kPa, 5 kPa, 10 kPa, 20 kPa, 30 kPa, 50 kPa, the allowable deviation is ±5%. If the temperature drops at 300°C with no flame, then the temperature setting of the furnace was reduced gradually by 10°C each time to increase the measuring accuracy.
It should be noted that the maximum temperature of the device is 990°C, the upper limit of ignition temperature is set to 900°C in this experiment. To ensure the normal use of the device and the safety of the experiment process, if the dust is not exploded at 900°C, the experiment will not be continued.

2.2.2. determination standard of ignition and determination of minimum ignition temperature of coal dust cloud

A high resolution camera was installed in the outside of the observation chamber. The camera which pointing at the reflector fixed in the bottom of the observation chamber can realize the real-time monitoring and recording of videos for coal dust explosion. Through observing the image captured in the slow playback of video at 1/8 magnification, flame length of dust cloud was recorded as illustrated in Fig. 4. When flame erupts or delays to erupt according to the observation result of the reflector, the flame length greater than 3mm was judged as ignition.

The coal dust cloud measured by the above method is on fire, if the minimum temperature of the furnace is greater than 300°C, the final result minus 20°C; if the minimum temperature of the furnace is less than or equal to 300°C, the final result minus 10°C, that is the minimum ignition temperature of coal dust cloud. Because the temperature recorder works to record the thermocouple’s temperature in real time, there is a certain deviation from the internal temperature of the furnace, so the correction of the result can make it more close to the real situation. In order to eliminate the accidental error of the experiment, each experiment of each coal sample must be repeated five times. If the results of this five experiments are ignition, the minimum ignition of the coal dust cloud can be further confirmed.

Fig. 4. The image for the flame length of coal specimens

2.3. The proximate analysis of coal specimens

The proximate indexes included the moisture content, ash content, volatile matter and fixed carbon of coal were used, among which, the fixed carbon content was calculated using a subtraction method. Based on ASTM standard D7582-12, a two-furnace automatic coal proximate analyzer integrating the functions of electronic temperature control and automatic weighing was used to measure the moisture content, ash content, and volatile matter, as illustrated in Fig. 5.
2.3.1. The determination of moisture content

The moisture content of the samples was conducted according to ASTM standard D3173 using moisture meter of the two-furnace automatic coal proximate analyzer. Heat the empty capsules under the conditions at which the sample is to be dried, place the cover on the capsule, cool over a desiccant for 15 to 30 min, and weigh. Dip out with a spatula from the sample bottle approximately 1 g of the sample, afterwards, put this quickly into the capsule, close, and weigh at once to the nearest ±0.1 mg.

After removing the covers, quickly place the capsules in a preheated oven (at 107°C) through which passes a current of dry air. Close the oven at once and heat for 1 h. Open the oven, cover the capsules quickly, cool in a desiccator over desiccant, and weigh as soon as the capsules have reached room temperature.

Calculate the percent moisture in the analysis sample as follows:

$$\text{Moisture in analysis sample, } \% = \left( \frac{A - B}{A} \right) \times 100$$  \hspace{1cm} (1)

where:

- $A$ = grams of sample used and
- $B$ = grams of sample after heating.

2.3.2. The determination of ash content

According to ASTM standard D3174, Ash is the residue remaining after burning the coal sample. Using the dried coal from the moisture determination in part 2.3.1. After removing the cover, place the capsule containing the sample in a cold furnace and heat gradually at such a rate that the temperature reaches 450 to 500°C in 1 h. Then heat coal samples so that a final temperature of 700 to 750°C is reached by the end of the second hour. Continue to heat at the final temperature for additional 2 h, if the sample reaches a constant weight at 700 to 750°C in less than 4 h, the 4-h time limit can be reduced.

Finally, remove the capsule from the muffle, place the cover on the capsule, cool under conditions to minimize moisture pickup, and weigh. As for nonreactive coals may require ad-
ditional time. If unburned carbon particles are observed, or if duplicate results are suspect, the samples should be returned to the furnace for sufficient time to reach a constant weight (±0.001 g).

Calculate the ash percent in the analysis sample as follows:

\[
\text{Ash in analysis sample, } \% = \left[ \frac{(A - B)}{C} \right] \times 100
\]

(2)

where:
- \(A\) = weight of capsule, cover and ash residue, g,
- \(B\) = weight of empty capsule and cover, g,
- \(C\) = weight of analysis sample used, g.

### 2.3.3. The measurement of volatile matter

According to ASTM standard D3175, volatile matter is determined by establishing the mass loss resulting from heating a coal under rigidly controlled conditions. The measured mass loss, corrected for moisture as determined in part 2.3.1 establishes the volatile matter content. The experiments use the test device of volatile matter in the two-furnace automatic coal proximate analyzer.

Record the mass of the crucible and cover to the nearest ± 0.0001 g. Dip out with a spatula from the sample bottle approximately 1 g of the sample. Put this quickly into the capsule, close, and weigh to the nearest 0.1 mg. Then close with the cover which fits closely enough so that the carbon deposit from coals does not burn away from the underside. Record the total mass of the crucible, sample and cover to the nearest ± 0.0001 g again. Afterwards, place the crucible on nickel-chromium wire supports and insert directly into the furnace chamber, which is maintained at a temperature of 950 ± 20°C, and lower immediately to the 950°C zone. After heating for a total of exactly 7 min, remove the crucible from the furnace and without disturbing the cover, allow it to cool.

To ensure uniformity of results, keep the cooling period constant and do not prolong beyond 15 min. The percentage loss in weight minus the percent moisture in accordance with Test Method D3173, is the volatile matter. Calculate the percent volatile matter in the analysis samples as follows:

\[
\text{Volatile matter in analysis sample, } \% = \left[ \frac{(B - C)}{(B - A)} \right] \times 100 - D
\]

(3)

where:
- \(A\) = mass of crucible and cover, g,
- \(B\) = mass of crucible and cover and contents before heating, g,
- \(C\) = mass of crucible and cover and contents after heating, g,
- \(D\) = moisture in analysis sample, %, as determined by part 2.3.1.

### 3. The PCA process

#### 3.1. Multi-index correlation analysis

The aim of this study was to reveal the influences of the indexes for the proximate analysis on MIT of coal. Based on the experiment of measuring the indexes in Part 2.2, six indexes
measured are air-dried moisture Mad (\(X_1\))%, air-dried ash Aad (\(X_2\))%, dry-base ash Ad (\(X_3\))%, air-dried volatile matter Vad (\(X_4\))%, dry-base volatile matter Vd (\(X_5\))%, and air-dried fixed carbon FCad (\(X_6\))%.

To avoid the correlation among different indexes which influences the analysis results, these indexes were used as variables to conduct correlation analysis, and the data were performed dimension reduction so as to employ fewer new variables to reflect all data information. Due to the work limits, the first ten types of coal samples among the 30 types of coal samples were carried out PCA, with observation matrix \(X\) for the six indexes as follows:

\[
X = \begin{bmatrix}
    x_{1,1} & x_{1,1} & \cdots & x_{1,1} \\
    x_{2,1} & x_{2,1} & \cdots & x_{2,1} \\
    \vdots & \vdots & \ddots & \vdots \\
    x_{6,1} & x_{6,1} & \cdots & x_{6,1}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
    1.80 & 15.20 & 15.80 & 6.10 & 6.40 & 76.90 \\
    0.22 & 14.80 & 15.20 & 10.10 & 10.30 & 74.88 \\
    0.62 & 18.10 & 18.30 & 13.30 & 13.50 & 67.98 \\
    0.55 & 19.60 & 19.40 & 17.10 & 17.20 & 62.75 \\
    1.30 & 12.70 & 13.30 & 22.30 & 23.30 & 63.70 \\
    1.90 & 22.00 & 23.20 & 23.60 & 25.70 & 52.50 \\
    1.54 & 7.30 & 7.50 & 30.40 & 32.10 & 60.76 \\
    2.30 & 14.70 & 15.70 & 36.50 & 39.10 & 46.50 \\
    2.80 & 7.00 & 7.70 & 38.20 & 41.60 & 52.00 \\
    1.70 & 4.60 & 4.70 & 42.70 & 45.10 & 51.00
\end{bmatrix}
\]

\(X\) is a 10×6 matrix; other six columns respectively represent 6 indexes, while 10 rows reflect ten types of different coal samples. Based on the test of the MIT for coal dust illustrated in Section 2.2, the MITs \(X_i\) of coal dust are displayed in Table I.

By performing correlation analysis of \(X_1, X_2, X_3, X_4, X_5, X_6\) with \(X_i\), the relevant coefficients were obtained as shown in Table II.

### Table I
The MIT of 10 types of coal specimens

<table>
<thead>
<tr>
<th>Coal specimens No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIT °C</td>
<td>851</td>
<td>820</td>
<td>605</td>
<td>581</td>
<td>546</td>
<td>543</td>
<td>525</td>
<td>485</td>
<td>437</td>
<td>412</td>
</tr>
</tbody>
</table>

### Table II
Correlation coefficients

<table>
<thead>
<tr>
<th></th>
<th>(X_1)</th>
<th>(X_2)</th>
<th>(X_3)</th>
<th>(X_4)</th>
<th>(X_5)</th>
<th>(X_6)</th>
<th>(X_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(X_1)</td>
<td>1.0000</td>
<td>-0.4020</td>
<td>-0.3498</td>
<td>0.6612</td>
<td>0.6846</td>
<td>-0.6588</td>
<td>-0.5246</td>
</tr>
<tr>
<td>(X_2)</td>
<td>-0.4020</td>
<td>1.0000</td>
<td>0.9975</td>
<td>-0.6531</td>
<td>-0.6476</td>
<td>0.2707</td>
<td>0.3564</td>
</tr>
<tr>
<td>(X_3)</td>
<td>-0.3498</td>
<td>0.9975</td>
<td>1.0000</td>
<td>-0.6303</td>
<td>-0.6226</td>
<td>0.2403</td>
<td>0.3420</td>
</tr>
<tr>
<td>(X_4)</td>
<td>0.6612</td>
<td>-0.6531</td>
<td>-0.6303</td>
<td>1.0000</td>
<td>0.9993</td>
<td>-0.9044</td>
<td>-0.8662</td>
</tr>
<tr>
<td>(X_5)</td>
<td>0.6846</td>
<td>-0.6476</td>
<td>-0.6226</td>
<td>0.9993</td>
<td>1.0000</td>
<td>-0.9084</td>
<td>-0.8635</td>
</tr>
<tr>
<td>(X_6)</td>
<td>-0.6588</td>
<td>0.2707</td>
<td>0.2403</td>
<td>-0.9044</td>
<td>-0.9084</td>
<td>1.0000</td>
<td>0.8958</td>
</tr>
<tr>
<td>(X_i)</td>
<td>-0.5264</td>
<td>0.3564</td>
<td>0.3420</td>
<td>-0.8662</td>
<td>-0.8635</td>
<td>0.8958</td>
<td>1.0000</td>
</tr>
</tbody>
</table>
According to the correlation coefficients shown in Table II, autocorrelation coefficient of various variables lives up to 0.8958, indicating that there are remarkable autocorrelation among variables. To make the data in processing free from overlapping to a certain degree, PCA was used to eliminate autocorrelation so as to realize the data dimension reduction (Li, 2015). The most of correlation coefficients of the MIT $X_t$ with $X_1$, $X_2$, $X_3$, $X_4$, $X_5$ and $X_6$ are higher than 0.3, the highest correlation reaches 0.8958, suggesting there is significant correlation the MIT and the indexes used in proximate analysis of coals.

### 3.2. PCA of multi-indexes

PCA is a statistic method for dimension reduction. It uses orthogonal transformation to let first largest variance of data projected on the new coordinate system locating at the first coordinate (the first principal component), while the second largest variance is placed on the second coordinate (Second principal component), the same to other variances. PCA cannot also reduce the dimension number of data sets, but also make data sets contribute to most prominent characteristics of variance.

The covariance matrix $S$ of the samples based on the matrix $X$ is written as:

$$
S = \begin{bmatrix}
0.6633 & -1.8809 & -1.6687 & 6.7819 & 7.6296 & -5.5643 \\
-1.8809 & 33.0089 & 33.5644 & -47.2544 & -50.9133 & 16.1264 \\
-1.6687 & 33.5644 & -46.4893 & -46.4893 & -49.8993 & 14.5936 \\
6.7819 & -47.2544 & 158.5890 & 158.5890 & 172.2112 & -118.1165 \\
7.6296 & -50.9133 & 172.2112 & 172.2112 & 187.2734 & -128.9275 \\
-5.5643 & 16.1264 & -118.1165 & -118.1165 & -128.9275 & 107.5543
\end{bmatrix}
$$

The results show that the diagonal element difference was largest in the matrix $S$: The principal components extracted based on the matrix $S$ tend to focus on the variables with large variances, resulting in false results. Hence, the principal components were extracted based on the matrix of correlation coefficients for samples, and the correlation coefficient matrix $R$ is obtained as follows:

$$
R = \begin{bmatrix}
1.0000 & -0.4020 & -0.3598 & 0.6612 & 0.6846 & -0.6588 \\
-0.4020 & 1.0000 & 0.9975 & -0.6531 & -0.6476 & 0.2707 \\
-0.3598 & 0.9975 & 1.0000 & -0.6306 & -0.6226 & 0.2403 \\
0.6612 & -0.6531 & -0.6306 & 1.0000 & 0.9993 & -0.9044 \\
0.6846 & -0.6476 & -0.6226 & 0.9993 & 1.0000 & -0.9084 \\
-0.6588 & 0.2707 & 0.2403 & -0.9044 & -0.9084 & 1.0000
\end{bmatrix}
$$

Principal components were carried out on the six indexes influencing the MIT $X_t$ of coal dust based on $R$, as shown in Table III.

According to Table 3, first principal component – eigenvalue is 4.2646, which is able to explain 71.08% of all data information; while the second eigenvalue is 1.2878, accounting for 21.4% of all data information; the third principal component – eigenvalue is 0.4462, demonstrating 7.44% of all data information. As seen in the Table, the cumulative contribution rate of first
three principal components lives up to 99.97%, indicating they can nearly reflect all information of the six indexes, therefore, the first three principal components were chosen in the analysis.

TABLE III

<table>
<thead>
<tr>
<th>Principal components</th>
<th>Eigenvalue</th>
<th>Contribution rate of variance /%</th>
<th>Cumulative variance contribution rate /%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_1$</td>
<td>4.2646</td>
<td>71.08</td>
<td>71.08</td>
</tr>
<tr>
<td>$F_2$</td>
<td>1.2878</td>
<td>21.46</td>
<td>92.54</td>
</tr>
<tr>
<td>$F_3$</td>
<td>0.4462</td>
<td>7.44</td>
<td>99.97</td>
</tr>
</tbody>
</table>

The results of PCA

The coefficients of each variable in the first standardized principal components (corresponding Unit orthogonal eigenvectors are $e_1^*$, $e_2^*$ and $e_3^*$), as displayed in Where $x_i^*$ denotes the standardized variable corresponding to $x_i (i = 1, 2, ..., 6)$, the first three principal components were derived as follows:

$$F_1 = -0.3587x_1^* + 0.3755x_2^* + 0.3630x_3^* - 0.4710x_4^* - 0.4722x_5^* + 0.3921x_6^*$$

$$F_2 = 0.2762x_1^* + 0.5549x_2^* + 0.5828x_3^* + 0.1310x_4^* + 0.1434x_5^* - 0.4882x_6^*$$

$$F_3 = 0.8896x_1^* - 0.0631x_2^* - 0.0046x_3^* - 0.2667x_4^* - 0.2236x_5^* + 0.2889x_6^*$$

3.3. The multivariate PAC

The first principal component $F_1$ reflects 71.08% of all data information, the corresponding eigenvectors, represent the contribution rates of each dimensional data to the first principal component. The symbols and absolute values respectively indicate the natures and values of the contribution of each dimensional data to the first principal component. In $F_1$ function, the coefficients in the front of $x_4^*$ and $x_5^*$ are larger and negative, which means volatile matter content
exerts greatest influence on the MIT of coal dust, namely, the higher the volatile matter content, the lower the MIT, the more possibility inducing CDE, \( F_1 \) is therefore considered volatile matter factor; for the second principal component, \( x_2^* \) and \( x_3^* \) in the function \( F_2 \) are larger and positive, suggesting the higher the ash content, the higher the MIT, \( F_2 \) is therefore defined ash content factor; while the coefficients in \( F_3 \) excluding \( x_1^* \) and \( x_6^* \) in the third principal component function \( F_3 \) are all negative and have absolute values, proving the higher moisture content and fixed carbon, the higher MIT of coal dust, hence, \( F_3 \) is defined as moisture content and fixed carbon factors.

4. The SVM predication of the MIT for coal dust

4.1. SVM regression algorithm

SVM is seen as a machine learning method constructed based on statistical learning theory. Unlike the neural network approaches which are based on empirical risk minimization principle of conventional, SVM follows structural risk minimization principle and it can deal with the high-dimensional nonlinear issues of small sample size, therefore it can be well used in regression predication.

Assumed that the training sample sets is \( \{x_i, y_i\} \), where input variable is \( x_i \in R^n \), while output variable is \( y_i \in R \), \( i = 1, 2, ..., l \), \( l \) refers to the number of training samples, \( R \) is real number, while \( n \) represents vector dimensions. The basic idea of SVM regression is to use preselected nonlinear value \( \phi \) for mapping input variables into a high-dimensional eigenspace \( F \) in which a linear regression function is adopted to carry out linear regression. It is an approach used for transforming low-dimensional nonlinear issues into high-dimensional linear issue. The process of the approach can be divided into two steps:

a. A nonlinear mapping datum is used to transformed input variables to a high-dimensional eigenspace \( F \);

b. Based on linear characteristics of mapping variables in eigenspace, the linear regression is conducted and its function is written as

\[
f(x) = \omega^T \phi(x) + b
\]

where \( \omega \) is weight and \( b \) is deviation.

The objective is to seek suitable \( \omega \) and \( b \) based on risk function minimization. As \( \phi(x) \) is preselected nonlinear mapping function, it is a constant, hence optimal regression function based on SVM is minimization objective function on the basis of structural risk minimization principle and expressed as

\[
\min R = \min \left[ \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{n} (\xi_i^* + \xi_i) \right]
\]

\[
\begin{cases}
    y_i - \omega^T \phi(x_i) - b \leq \xi_i + \varepsilon \\
    \omega^T \phi(x_i) + b - y_i \leq \xi_i^* + \varepsilon \\
    \xi_i, \xi_i^* \geq 0
\end{cases}
\]
Where, $C$ is penalty factor, to compromise belief risk and empirical risk, the larger the $C$ value, the higher fitting capability the $C$ data. $\xi_i, \xi_i^*$ are relaxation factors, $i = 1,2,\ldots,l$, $\varepsilon$ is the minimized value.

Lagrangian function is introduced as

$$L = \frac{1}{2} ||\phi||^2 + C \sum_{i=1}^{n} (\xi_i + \xi_i^*) - \sum_{i=1}^{n} \alpha_i \left[ \xi_i + \varepsilon - y_i + f(x_i) \right] - \sum_{i=1}^{n} \alpha_i^* \left[ \xi_i^* + \varepsilon + y_i - f(x_i) \right] - \sum_{i=1}^{n} (\gamma_i \xi_i + \gamma_i^* \xi_i^*)$$  (7)

Where $\alpha_i$ and $\gamma$ are Lagrange multiplier.

Partial differentials are solved using $\omega, \xi, \xi^*$ and $b$ in Eq. (7) as follows:

$$\begin{align*}
\omega &= \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) x_i \\
\sum_{i=1}^{l} (\alpha_i - \alpha_i^*) x_i &= 0 \\
C - \alpha_i - \gamma_i &= 0, i = 1,\ldots,l \\
C - \alpha_i^* - \gamma_i^* &= 0, i = 1,\ldots,l
\end{align*}$$  (8)

By eliminating $\omega$ and $\gamma$, $\alpha$ is calculated and $f(x)$ equation is expressed as:

$$f(x) = \sum_{i,j=1}^{l} (\alpha_i - \alpha_i^*) \left[ \phi(x_i) \phi(x_j) \right] + b$$  (9)

To effectively solve the Eq. (9) aforementioned, kernel function $k(x_i,x_j) = \phi(x_i)\phi(x_j)$ is introduced to reflect the similarity of support vectors and unknown samples, while support vector regression issue is converted into a quadratic programming issue to further acquire SVM regression function as

$$f(x) = \sum_{i,j=1}^{l} (\alpha_i - \alpha_i^*) k(x_i,x_j) + b$$  (10)

4.2. The prediction of the MIT for coal dust

Based on the six measured indexes of proximate analysis (air-dried moisture Mad ($X_1$)%, air-dried ash Aad ($X_2$)%, dry-base ash Ad ($X_3$)%, air-dried volatile matter Vad ($X_4$)%, dry-base volatile matter Vd ($X_5$)%, and air-dried fixed carbon FCad ($X_6$)%), as well as MIT $X_t$, SVM algorithm was used to establish a regression model.

The process of SVM training and prediction of the MIT for coal dust was described as follows:

(1) The preparation of data sets

The data formats supported by SVM were constructed by taking the seven indexes aforementioned of the coal samples No. 11-30 as training samples, and the seven indexes for the coal samples No. 1-10 as prediction samples.
(2) Input and output setting

\[ X \text{ denotes input, while } Y \text{ represents output.} \]

\[
X = (X_1, X_2, X_3, X_4, X_5, X_6) \quad Y = X_i
\]

(3) The compiling of the program

Two function – svm\_train and svm\_pridect were used: among which, svm\_train function was adopted to train and predicate samples; while svm\_pridect function was used for test (Wang, 2013).

(4) The predicated results

There are 20 samples being created and trained in total to conduct model training. The predication accuracy based on training samples is 100%.

The created model was used to predicate the coal samples No. 1-10, with the results as illustrated in Fig. 6. The errors of predicated data and raw data are displayed in Table V.

![Fig. 6. The predicking performance](image)

### TABLE V

<table>
<thead>
<tr>
<th>Coal specimens No.</th>
<th>Sample data</th>
<th>Predicated values</th>
<th>Errors</th>
<th>Percentage error %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>851</td>
<td>823</td>
<td>28</td>
<td>3.40</td>
</tr>
<tr>
<td>2</td>
<td>820</td>
<td>798</td>
<td>22</td>
<td>2.76</td>
</tr>
<tr>
<td>3</td>
<td>605</td>
<td>630</td>
<td>−25</td>
<td>3.97</td>
</tr>
<tr>
<td>4</td>
<td>581</td>
<td>552</td>
<td>29</td>
<td>5.25</td>
</tr>
<tr>
<td>5</td>
<td>546</td>
<td>560</td>
<td>−14</td>
<td>2.50</td>
</tr>
<tr>
<td>6</td>
<td>543</td>
<td>526</td>
<td>17</td>
<td>3.23</td>
</tr>
<tr>
<td>7</td>
<td>525</td>
<td>499</td>
<td>26</td>
<td>5.21</td>
</tr>
<tr>
<td>8</td>
<td>485</td>
<td>520</td>
<td>−35</td>
<td>6.73</td>
</tr>
<tr>
<td>9</td>
<td>437</td>
<td>450</td>
<td>−13</td>
<td>2.89</td>
</tr>
<tr>
<td>10</td>
<td>489</td>
<td>456</td>
<td>33</td>
<td>7.24</td>
</tr>
</tbody>
</table>
According to the error comparison in the table, the error of the predicated results using SVM was mainly in the range of 2% ~ 4%, the highest error was 7.24%.

5. Conclusions

The following conclusions are obtained in the research:
• The PCA was conducted on the indexes of multi-component coals by extracting three principal components (volatile matter, ash content, as well as moisture content and fixed carbon. Among them, volatile matter accounted for 71.08% of all the data, which meant volatile matter content played primary role in decreasing MIT.
• The SVM regression model built in the research for forecasting proximate indexes of coal dust and the MIT had a small forecasting error, which confirmed that the model can be applied in the predication of the MIT for coal dust.

References


