scientific reports

OPEN

An air door opening and closing time identification and stage division method based on the wind speed data of a single sensor

Ying Song[1✉], Shan Li[1] & Wentian Shang[2]

In mines, tunnel ventilation is monitored using wind speed sensors to measure the stability of the mine ventilation system. However, opening and closing the air door will cause violent fluctuations in the monitoring data of the wind speed sensors. When false alarms are triggered, the staff can diagnose only the mine ventilation system based on their experience. A numerical simulation method is adopted to explore the changes in the flow field during the opening and closing of the air door to address this issue. In addition, a method that is based on the wind speed data of a single sensor is proposed to identify the time and divide the stages of air door opening and closing. The experimental results showed that the proposed method can successfully identify the air door opening and closing time and apply stage division when needed.

In mine ventilation systems, air doors are commonly built in tunnels to control the airflow direction and speed to maintain the airflow distribution in conformance with the designed route[1,2]. However, during the process of opening and closing the air door in the tunnel, the flow field at the air door's location and the associated tunnel will change drastically, resulting in large fluctuations in the wind speed sensor monitoring data for the tunnel[3]. The fluctuations in the wind speed during this process are greater than those during other activities, such as car operation[4]. Large data fluctuations can cause wind speed sensor alarms, which interfere with normal mine production activities, making the application of mine ventilation management systems extremely difficult[5–8]. Therefore, it is crucial to identify the time duration and classify the air door opening and closing stages by monitoring the wind speed sensor data.

The essence of wind speed monitoring data is time series data. Therefore, fluctuations in the wind speed sensor monitoring data, which are caused by opening and closing the dampers, can be considered time series anomalies. In recent years, time series anomaly identification methods have been applied in many fields[9,10]. In blast furnace ironmaking, Zhou et al.[11] proposed a method that integrates principal component analysis (PCA) and independent component analysis (ICA) to monitor and diagnose furnace anomalies during blast furnace ironmaking by using operational data from the BF ironmaking process. In semiconductor fabrication, Hsu et al.[12] developed a multiple time series convolutional neural network (MTS-CNN) model to distinguish between normal and abnormal wafers by collecting monitoring data from multiple sensors. In oil and gas extraction, Soriano-Vargas et al.[13] studied a visualization and analysis method that is based on an interactive visualization of the time series data for anomaly detection through monitoring the data of oil and gas reservoirs. In aviation automation, He et al.[14] used an anomaly detection and mitigation algorithm that is based on online subspace tracking to achieve high accuracy in anomaly detection and low error in data recovery through online unmanned aerial vehicle (UAV) flight data. In power quality signals, Rodriguez et al.[15] adopted a recurrent long short-term memory (LSTM) recurrent neural network (RNN) for power quality disturbance classification through a deep convolutional autoencoder and stacking to identify and classify disturbances through power quality signals. In speech signal recognition, Cao et al.[16] employed an urban noise identification method using convolutional neural networks (CNNs) to classify urban noise by utilizing the monitored acoustic signals within a city. In pipeline transportation, Zang et al.[17] utilized a small leak detection method based on virtual sample generation (VSG) and unified feature extraction (UFE) techniques, which improved accuracy in detecting small leaks in pipelines through deep mining of basic information and machine learning model training.

[1]School of Management Science and Engineering, Shandong Technology and Business University, Yantai 264005, China. [2]College of Safety Science and Engineering, Liaoning Technical University, Huludao 125105, China. ✉email: 201613602@sdtbu.edu.cn

In mine ventilation, several methods exist for the effective detection of fluctuations in wind speed sensor data. Huang et al.[18] proposed a hybrid coded adaptive evolutionary strategy (ES) algorithm, which can identify the location of faults in a ventilation network and predict the range of changes in wind speed data using data from multiple wind speed sensors. Zhao et al.[19] introduced a method for building a tunnel fault wind speed range library using a 0–1 sensitivity matrix and applied an improved support vector machine (SVM) method for fault diagnosis and localization. However, the above methods have three major limitations: (i) They cannot be used to obtain additional information about a fault after acquiring the reason or location of its occurrence. (ii) They are used only for permanent or prolonged failures such as tunnel collapse, broken air doors, or fan malfunction. Normal operation with minor variations in time, such as the normal opening and closing of dampers, cannot be recognized. (iii) Their performance is significantly influenced by sensor location.

The identification method for air door opening and closing proposed in this paper aims not only to determine the air door opening and closing state but also to explore and clarify the relationship between the air door opening and closing and the abnormal fluctuation data of the wind-velocity sensor data to conduct intelligent identification. To better identify fluctuations in wind speed data caused by the opening and closing of air doors, this paper uses image recognition to identify time series data, which is different from the above methods.

This article is organized into five sections, including the current section. In Section "Numerical simulation", a numerical simulation is used to explore the air door opening and closing process and identify the specific flow field changes in the associated tunnel after closing. In Section "The proposed time identification and stage division method", a method for identifying air door opening and closing times and stage division via data enhancement via multiscale sliding windows, the extraction of deep features from the data via DWT, the classification of the data samples via a classification model, and the correction of the identification results via a regression model are proposed. The advantages of the SVM and least absolute shrinkage and selection operator (LASSO) models are validated in this paper in terms of their accuracy in air door opening and closing time recognition and stage division compared to other classification and correction models through a significant number of experiments, as shown in Section "Experimental studies". Conclusions and future work are presented in the last section.

Numerical simulation

In this section, several numerical simulations were conducted to explore the specific changes in the tunnel's flow field during the opening and closing of the air door. The model contains four parts: (1) a brief description of the physical model is given in the first part, (2) the mesh generation method and the quality of the final generated mesh are discussed in the second part, (3) the selections of the solution model and the parameter settings for the numerical simulations are given in the third part, and (4) the numerical simulation results are presented, and the flow field laws are summarized in the fourth part.

Physical model

The numerical simulation model in this study refers to the experimental tunnel of the Laboratory of Mine Thermodynamic Disasters and Control of Ministry of Education. The tunnel section is rectangular, with a length of 2.5 m and a width of 3 m. The air door is located in the middle of the left and right sides of the tunnel. The detailed dimensions of the tunnel are shown in Fig. 1a.

Mesh generation

Given the complex geometric structure of the physical model, partition and local encryption methods were used to divide the meshes. Tetrahedral meshes were used in areas VI and VII, with air doors within 5 m. Hexahedral meshes were used in areas I, II, III, IV, and V. Dense meshes were used within 5 m of the air door, and sparse meshes were used in the area.

After the grid dependency test, the final model in this study consisted of 434,700 meshes in the area within 5 m of the air door, 79,200 meshes in areas VI and VII, and 112,500 meshes in areas I, II, III, IV, and V, for a
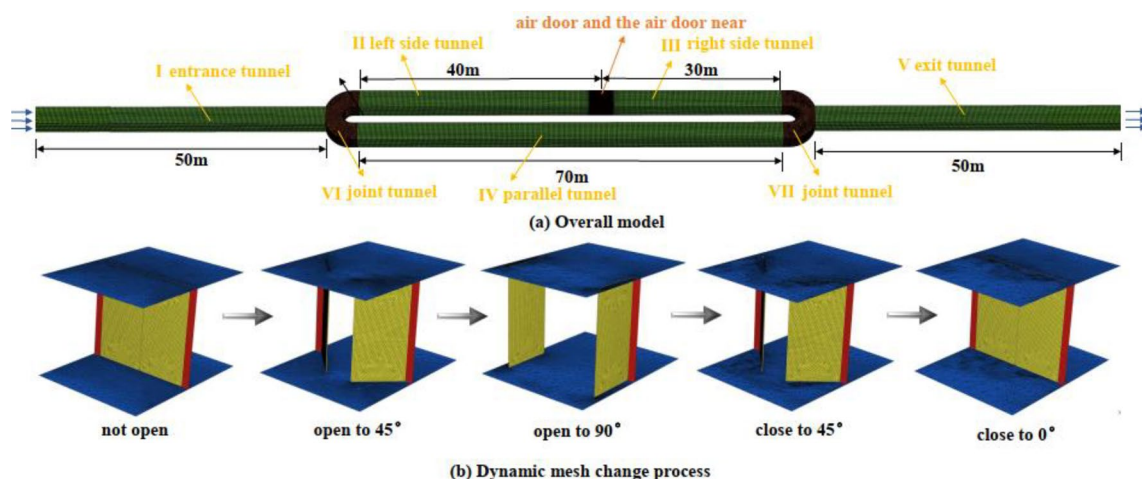

Figure 1. Physical model after meshing.

total of 626,400 meshes. Skewness is one of the quality inspection criteria of the mesh basic unit, where a value of 0 is the best quality and 1 is the worst quality. After the mesh division, the average skew of the entire model was approximately 0.283, and the maximum deflection of the mesh in the entrance area was 0.794. All of these values fall within the requirements of the mesh's maximum deflection being less than 0.97 to ensure good mesh quality. The mesh generation results are shown in Fig. 1, where Fig. 1a shows the overall mesh of the model and Fig. 1b shows the mesh change process in the area within 5 m of the air door.

Model selection and parameter setting

According to the actual mine ventilation system situation, the inlet and outlet boundaries are set as the velocity inlet and pressure outlet, respectively, and the physical quantities of the fluid within the physical model are set with reference to air. According to the literature[20–22], the temperature is 20 °C, the humidity is 50%, the pressure is atmospheric pressure, the realizable k-ε turbulence model is selected as the solution model, and the coupled method with second-order upwind accuracy is selected for the solution. According to the air door motion law, the dynamic mesh is updated using the spring analogy model and the local redrawing model. After the time step sensitivity analysis, the selected time step is 0.0125 s.

Numerical simulation results

In this section, numerical simulations were conducted based on several condition parameters: (1) the inlet wind speed before opening the air door was 3 m/s; (2) the opening and closing angle was 90°; (3) the opening and closing speed was 15°/s; and (4) the time duration to open the air door to a fixed angle was 20 s. Figure 2 shows the numerical simulation results and the field velocity.

Figure 2 shows that a sudden increase in the wind speed develops during the opening and closing of the air door, where its influence is limited to the tunnel area where the air door is located. This is due to the narrower cross-section between the two doors compared to the tunnel section. Ultimately, the air door opens and closes the overall process, which dramatically changes the flow field. In the tunnel where the air door is located, the wind speed is significantly higher, and the flow field changes are more disturbing. In the tandem tunnel, the wind speed is slightly higher, and the flow field changes are more stable. In parallel tunnels, the wind speed is significantly lower, and the flow field changes are more stable. The flow field can still change even after the air door is close to 0°, after which it becomes stable.

The proposed time identification and stage division method

This section describes the specific process of identifying the air door opening and closing time and the stage division method. It includes five parts. First, the overall architecture of the proposed time identification and stage division method is introduced. Second, the preprocessing step, including discrete normalization and multiscale sliding window discretization, is described in detail. Third, the classification steps, including feature vector extraction and the classification process, are described. Fourth, we introduce the merging and selection steps. Finally, the correction steps of the four regression models and their feature vectors are introduced.

The architecture of the method

Based on the research discussed in the previous two sections, the air door opening and closing time identification and stage division method is proposed based on the numerical simulation results. Figure 3 shows the architecture of the proposed method.

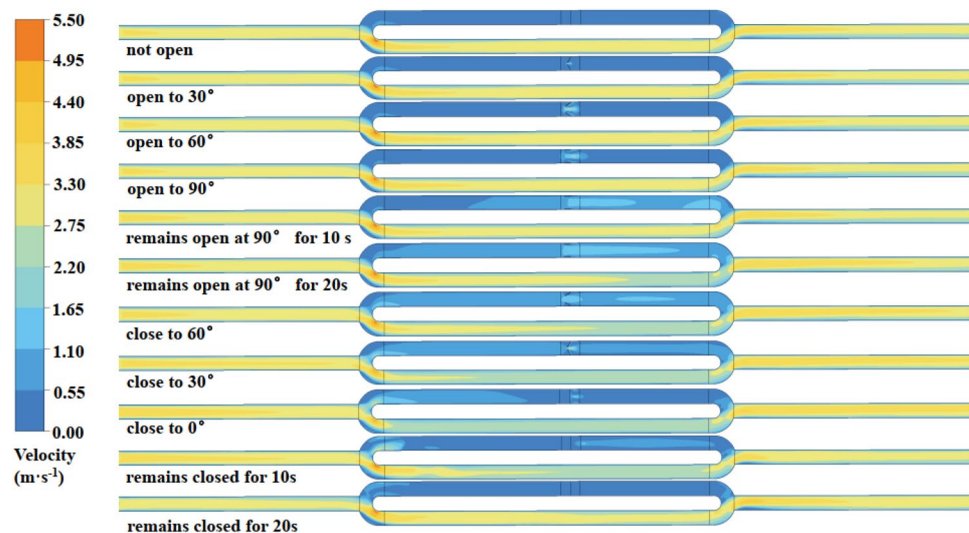The proposed method includes four main steps, as follows:

Figure 2. The flow field changes due to opening and closing the air door.
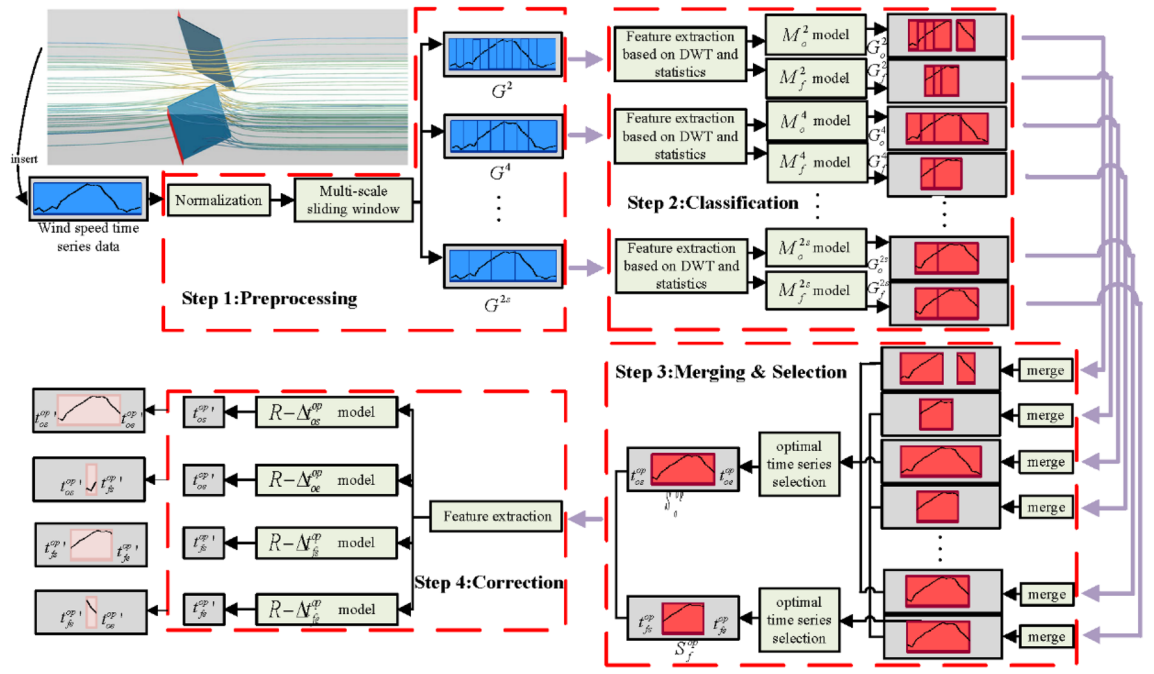
Figure 3. The process of identifying the air door opening and closing time and stage division.

Step 1: Preprocessing. The wind speed sensor data segment is processed into time series data with a varied range of [0, 1] using outlier normalization. Then, multiple original samples $\{G^2, G^4, \ldots, G^{2s}\}$ composed of subtime series data are generated using the multiscale sliding window.

Step 2: Classification. The traditional statistics and discrete wavelet transform are used to extract the feature vectors from all the subtime series data of each original sample, and multimodels that are dedicated to this sample are used to complete the classification process. For example, in the original sample (G^k) classification process, one of the two classification models (M_o^k) is for classifying the air door opening and closing time, and the other (M_f^k) is for classifying the air door fixed angle opening time. In summary, two types of samples can be generated based on the original sample classification results. The first type is for the time identification samples $\{G_o^2, G_o^4, \ldots, G_o^{2s}\}$ that are composed of different sets containing subtime series data that belong to the air door opening and closing category and its classification confidence. The second type is for stage division samples $\{G_f^2, G_f^4, \ldots, G_f^{2s}\}$, which are composed of multiple sets containing subtime series data that belong to the air door's fixed angle opening category and its classification confidence.

Step 3: Merging and Selection. All the subtime series data in each time identification sample or stage division sample are merged using the merge set approach. After the merging is completed for all the samples of both types, the optimal air door opening and closing time series data (S_o^{op}) and the optimal air door fixed angle opening time series data (S_f^{op}) are selected using the intersection over union (IoU) metric with classification confidence.

Step 4: Correction. Twelve basic features for the two optimal time series are extracted as the input features for using the four regression models $\left\{R - \Delta t_{os}^{op}, R - \Delta t_{oe}^{op}, R - \Delta t_{fs}^{op}, R - \Delta t_{fe}^{op}\right\}$. The two optimal time series starting and ending times $\left\{t_{os}^{op}, t_{oe}^{op}, t_{fs}^{op}, t_{fe}^{op}\right\}$ are corrected. Using the four correction times $\left\{t_{os}^{op\prime}, t_{oe}^{op\prime}, t_{fs}^{op\prime}, t_{fe}^{op\prime}\right\}$, the air door opening and closing times $\left[t_{os}^{op\prime}, t_{oe}^{op\prime}\right]$ can be identified and divided into the air door opening stage $\left[t_{os}^{op\prime}, t_{fs}^{op\prime}\right]$, air door fixed angle opening stage $\left[t_{fs}^{op\prime}, t_{fe}^{op\prime}\right]$, and air door closing stage $\left[t_{fe}^{op\prime}, t_{oe}^{op\prime}\right]$.

Preprocessing

The wind speed sensor data are continuous time series data. Before processing the data using machine learning, data discretization and normalization are performed to obtain good classification performance[23–25].

Dispersion standardization

Data normalization is an important preprocessing step that converts all the data of several ranges to fit in the range [0, 1], making different samples comparable[26].

Normalization can be performed via various methods. In this study, deviation standardization is used to process the wind speed sensor data. The expression is shown in Eq. (1):

$$x' = \frac{(x - x_{\min})}{(x_{\max} - x_{\min})} \tag{1}$$

where x' represents the normalized data, x represents the wind speed sensor data, x_{min} represents the minimum value of the wind speed sensor data, and x_{max} represents the maximum value of the wind speed sensor data.

The wind speed monitoring data $\{x_1, x_2, \ldots, x_L\}$ change to $\{x'_1, x'_2, \ldots, x'_L\}$ after dispersion standardization is applied.

Multiscale sliding window discretization

Discretization, which converts continuous data into discrete data with a finite number of intervals, is one of the most basic data partitioning techniques. Using discretized data allows for the construction of more efficient machine learning models[27–29]. Similar to most other time sequences, wind speed sensor data have multiscale properties. Figure 4 shows that the data at different scales can show different patterns. Therefore, the adopted discrete method in this study is the multiscale sliding window.

Based on the Coal Mine Safety Regulations in China and the data characteristics of the wind speed sensor data, the parameter constraints are obtained[4]. The parameters that are selected for the multiscale sliding window obey the following two constraints:

$$P_{min} \leq \{W|w_i\} \leq P_{max} \tag{2}$$

$$\begin{cases} t_{i,k} = \dfrac{w_i}{2} & l_{i,k} \geq \dfrac{w_i}{2} \\ t_{i,k} = l_{i,k} & l_{i,k} < \dfrac{w_i}{2} \end{cases} \tag{3}$$

Equation (2) is the constraint of the sliding window scale, where W is the set of the sliding window scale, w_i is the sliding window scale, and P_{min} and P_{max} represent the minimum and maximum values of a reasonable sliding window scale, respectively. When P_{min} is greater than or equal to 2, P_{max} is less than the quickest air door opening and closing time. Equation (3) is the constraint of the sliding distance, where $t_{i,k}$ is the sliding distance of the w_i scale sliding window and $l_{i,k}$ is the remaining length of the time series data after the w_i scale sliding window has been slid k times.

According to the two constraints of the multiscale sliding window, a segment of the normalized wind speed sensor data $\{x'_1, x'_2, \ldots, x'_L\}$ is discretized by a multiscale sliding window to generate multiple original samples $\{G^2, G^4, \ldots, G^{2s}\}$. The expression of each sample is shown in Eq. (4):

$$\{G^k|S_i^k\}, i = 1, 2, \ldots, n \tag{4}$$

where G^k represents the sample discrete by the k-scale sliding window and S_i^k represents the subtime series data within the ith k-scale sliding window. $S_i^k = \{x'_{i\cdot k+1}, x'_{i\cdot k+2}, \ldots x'_{i\cdot k+k}\}$ when $i \neq n$; $S_i^k = \{x'_{L-k}, x'_{L-1}, \ldots, x'_L\}$ when $i = n$.
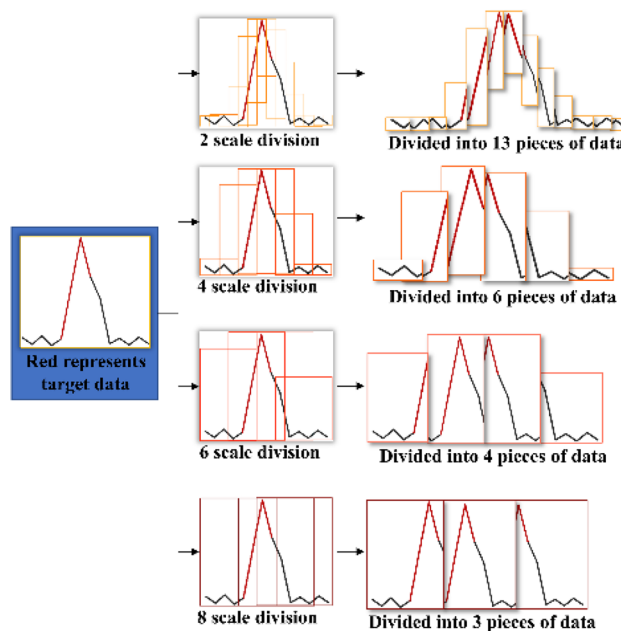

Figure 4. An example of continuous data discretized by a multiscale sliding window.

Classification

This section introduces two major points: the composition of the feature vectors extracted from each subtime series and the classification process for the specific flow.

Feature extraction method based on the DWT and statistics for the subtime series data
Statistical features can fully express the global information of subtime series data. Each subtime series statistical feature includes the mean value $\overline{x'}$, the minimum value $x'_{\min}$, the maximum value $x'_{\max}$, and the standard deviation σ. The expressions of the features used in this study are listed in Table 1.

However, the limited statistical features cannot fully express the sample's fluctuation information. A method for extracting fluctuation characteristics based on the DWT is proposed to mine the hidden information from subtime series data.

The DWT is a signal represented by a finite length or a fast-decaying oscillatory waveform that is scaled and panned to match the input data. It overcomes short-time Fourier transform (STFT) deficiencies with a good time–frequency local analysis capability and multiresolution analysis characteristics, which makes it widely used in signal and image recognition and detection. The process of a discrete wavelet transformation of the subtime series data can be expressed using Eq. (5):

$$WT_f(q, w) = \int_R f(t) \cdot \overline{\psi_{q,w}(t)} \, dt \tag{5}$$

where $f(t)$ is subtime series data, $WT_f(q, w)$ is subtime series data resulting from discrete wavelet transform data, t is the sequence of data, q is a scale parameter, w is the translation parameter along the time axis, and $\psi_{q,w}(t)$ is a wavelet base function. In this study, the db1 wavelet is used.

After processing by the DWT, the subtime series data are decomposed into several layers, each consisting of several high- and low-frequency coefficients. The fluctuation features are obtained by performing entropy sum calculations on the high- or low-frequency coefficients obtained from the multilayer decomposition. The entropy sum formula for any layer of the low- and high-frequency coefficients is calculated using Eq. (6):

$$\begin{cases} e_{js} = \sum_{i=1}^{n_s} -cs_{ji} \times \log_2^{cs_{ji}} \\ e_{jd} = \sum_{i=1}^{n_d} -cd_{ji} \times \log_2^{cd_{ji}} \end{cases} \tag{6}$$

where e_{js} represents the entropy sum of the low-frequency coefficients in layer j, e_{jd} represents the entropy sum of the high-frequency coefficients in layer j, cs_{ji} represents the i th low-frequency coefficient in layer j, cd_{ji} represents the ith high-frequency coefficient in layer j, n_s represents the total number of low-frequency coefficients in layer j, and n_d represents the total number of high-frequency coefficients in layer j.

Since this article uses a db1 wavelet with a filter length of 2, the number of decomposition layers is calculated using Eq. (7):

$$l = \left\lceil \log_2^k \right\rceil \tag{7}$$

where l represents the number of decomposable layers.

According to Eq. (6) and Eq. (7), the fluctuation features of each subtime series data are composed as shown in Eq. (8):

$$\{C_f | e_{is}, e_{id}\} i = 1, 2, \ldots, l \tag{8}$$

where C_f represents the set of fluctuation features, e_{is} represents the entropy sum of the low-frequency coefficients in the i-layer, and e_{id} represents the entropy sum of the high-frequency coefficients in the i-layer.

In summary, each subtime series feature vector consists of several statistical and fluctuating features. The subtime series data feature vector C_i^k is shown in Eq. (9):

$$C_i^k = \left(\overline{x'}, x'_{\min}, x'_{\max}, \sigma^2, e_{1s}, e_{1d}, \ldots, e_{Ls}, e_{Ld} \right)^T \tag{9}$$

Statistical features	Expression
Mean value	$\overline{x'} = \frac{\sum_i^k x'_i}{k}$
Minimum value	$x'_{min} = \min(S_i^k)$
Maximum value	$x'_{max} = \max(S_i^k)$
Standard deviation	$\sigma = \frac{\sum_i^k (x'_i - \overline{x'})}{k}$

Table 1. Four statistical features.

To illustrate the feature vector clearly, Table 2 shows the composition of the feature vectors that are extracted by 2, 4, and 8 data lengths with the subtime series data covering the air door closing time and other times.

Classification process for the specific flow
Multiple models ($\{M_o^2, M_o^4, \ldots, M_o^{2s}\}$ and $\{M_f^2, M_f^4, \ldots, M_f^{2s}\}$) are used in the classification process. Their training and testing data can be found in Section "Experimental studies" of this article. Through the classification process, multiple time identification samples $\{G_o^2, G_o^4, \ldots, G_o^{2s}\}$ and multiple stage division samples $\{G_f^2, G_f^4, \ldots, G_f^{2s}\}$ can be obtained. This is an important prerequisite for air door opening and closing time identification and stage division. Algorithm 1 shows the procedure for the classification process.

Inputs: multiple origin samples $\{G^2, G^4, ..., G^{2s}\}$, multiple classification models $\{M_o^2, M_o^4, \ldots, M_o^{2s}\}$ for classifying the air door opening and closing time from other times, multiple classification models $\{M_o^2, M_o^4, \ldots, M_o^{2s}\}$ for classifying the air door fixed angle opening time from other times;

Outputs: multiple time identification samples $\{G_o^2, G_o^4, ..., G_o^{2s}\}$, multiple stage division samples $\{G_f^2, G_f^4, ..., G_f^{2s}\}$.

1. **establish** multiple empty samples $\{G_o^2, G_o^4, ..., G_o^{2s}\}$ and $\{G_f^2, G_f^4, ..., G_f^{2s}\}$
2. **for** k in 2, 4, … ,2 s **do**
3.   j, h = 0, 0
4.   **for** I in 1, 2, …, n **do**
5.   Realize $S_i^k \rightarrow C_i^k$ by Eq. (6), Eq. (7) and Table 1. Expression
6.   classify C_i^k by M_o^k
7.   **if** C_i^k Category = air door opening and closing time **then**
8. G_o^k append [S_i^k ,Category confidence o_i^k] $_j$
9.   j = j + 1
10.   **end if**
11.   classify C_i^k by M_f^k
12.   **if** C_i^k Category = air door fixed angle opening time **then**
13. G_o^f append [S_i^k ,Category confidence f_i^k] $_h$
14.   h = h + 1
15.   **end if**
16.   **end for**
17. **end for**

Algorithm 1. Classification process.

Merging and selection
This section introduces two main processes. The first introduces the merging process of the subtime series. The second introduces the process of selecting the optimal time series based on the IoU and the confidence.

Subtime series merging
Each time identification sample or stage division sample has many overlapping or nonoverlapping subtime series data, which are all part of the target time series that is selected by the classification model, as shown in Fig. 5. Therefore, we need to overlap the judgements and merge the two subtime series.

The IoU is used to determine whether two subtime series of data overlap. The IoU between two subtime series of data can be calculated using Eq. (10):

$$IoU_{ab} = \frac{S_a^k \cap S_b^k}{S_a^k \cup S_b^k} \tag{10}$$

where S_a^k and S_b^k represent the two subtime series data.

Data curve	Data length	Data time	$\bar{x}$	x_{min}	x_{max}	σ^2	e_{1s}	e_{1d}	e_{2s}	e_{2d}	e_{3s}	e_{3d}
	2	Other	0.500	0.442	0.558	0.082	0.048	0.500	–	–	–	–
	2	Air door closing	0.851	0.798	0.903	0.075	0.042	− 0.774	–	–	–	–
	4	Other	0.476	0.356	0.558	0.047	0.031	0.959	0.128	0.129	–	–
	4	Air door closing	0.834	0.798	0.903	0.050	0.050	− 1.332	0.011	− 4.110	–	–
	8	Other	0.404	0.308	0.558	0.101	0.041	1.919	0.136	0.650	0.191	− 0.501
	8	Air door closing	0.764	0.548	0.903	0.110	0.102	− 1.369	0.169	− 5.942	0.182	− 10.401

Table 2. Comparison of 2-, 4-, and 8-length target and non-target subtime series data features.
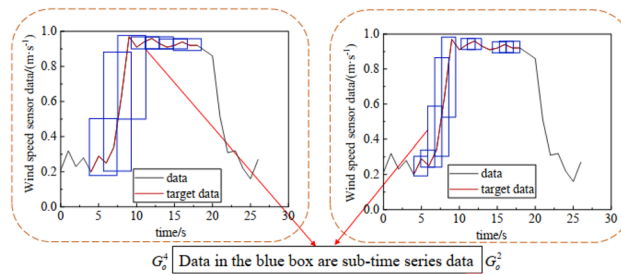
Figure 5. Example of different sample subtime series data before the merge.

When the IoU is greater than 0, the two time series data overlap. The union method is used to merge the two subtime series of data. The merging of the two overlapping subtime series data and the confidence calculation of the merged subtime series data are shown in Eq. (11):

$$\begin{cases} S_m^k = S_a^k \cap S_b^k \\ c_m^k = \dfrac{c_a^k + c_b^k}{2} \end{cases} \tag{11}$$

where S_m^k represents the merged subtime series data, c_m^k represents the confidence of the merged subtime series data, and c_a^k and c_b^k represent the confidence of S_a^k and S_b^k, respectively.

Based on the overlapping judgement and merging of the two time series datasets, a merging method for each sample time series is proposed. Algorithm 2 shows the merging process of the multiple time identification samples $\left\{ G_o^2, G_o^4, \ldots, G_o^{2s} \right\}$ using this method. The merging process of the stage division samples $\left\{ G_f^2, G_f^4, \ldots, G_f^{2s} \right\}$ using this method is the same.

Inputs: multiple time identification sample $\{G_o^2, G_o^4, ..., G_o^{2s}\}$

Outputs: multiple merged time identification sample $\{G_o^{2\,\prime}, G_o^{4\,\prime}, ..., G_o^{2s\,\prime}\}$

  1. **establish** multiple empty samples $\{G_o^{2\,\prime}, G_o^{4\,\prime}, ..., G_o^{2s\,\prime}\}$

  2. **for** k in 2, 4, … ,2 s **do**

  3.   s,d,f= 0

  4.   **for** j in 1, 2, …,n-1 **do**

  5.   calculate $IoU_{j,i+1}$ by Eq. (11)

  6.   **if** $IoU_{j,i+1} \geq 0$ and f=d **then**

  7.   **if** f=d **then**

  8.   merge S_j^k and $S_{j+1\ \ j+1}^k \rightarrow S_s^k$ by Eq. (12)

  9.   d= d+1

10.   **end if**

11.   **if** f<d **then**

12.   merge S_s^k and $S_{j+1\ \ j+1}^k \rightarrow S_s^k$ by Eq. (12)

13.   **end if**

14.   **end if**

15.   **if** $IoU_{j,i+1} \leq 0$ **then**

16. $G_o^k\,\prime$ append [S_s^k ,Category confidence o_s^k] $_s$

17.   f = f+1

18.   **end if**

19.   **end for**

20.**end for**

Algorithm 2. Time identification samples time series data merging process.

Optimal time series selection based on the IoU and the confidence
For multiple merged identification samples $\left\{ G_o^{2\prime}, G_o^{4\prime}, \ldots, G_o^{2s\prime} \right\}$ or multiple merged stage division samples $\left\{ G_f^{2\prime}, G_f^{4\prime}, \ldots, G_f^{2s\prime} \right\}$, their subtime series data are considered preliminary identification results. These results include an error result and an accuracy gap between the correct results, as shown in Fig. 6. Therefore, an optimal time series selection method based on the IoU and the confidence interval is proposed.

In this method, the IoU threshold is used to eliminate incorrect results. Initially, an IoU threshold is set. If the average IoU value of a result is smaller than the threshold, the IoU value is eliminated. In this study, the IoU threshold is set to 0.1. After eliminating the error, the result with the highest confidence is selected as the optimal recognition result. The average IoU of each recognition result is calculated using Eq. (12):
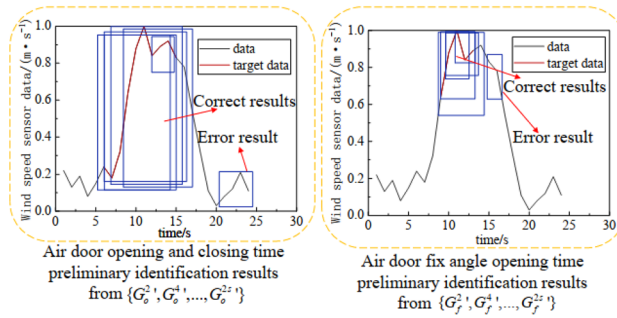
Figure 6. An example of the preliminary identification results.

$$\overline{IoU_m^v} = \frac{\sum\limits_{k=1}^{n_k} \left(\sum\limits_{s=1}^{n_s} \frac{S_m^v \cap S_s^k}{S_m^v \cup S_s^k} \right)}{n_s \cdot n_k} \tag{12}$$

where $\overline{IoU_m^v}$ represents the average IoU of the S_m^v subtime series and the other subtime series data, n_s represents the number of time series data in G_o^k or G_f^k, and n_k represents the number of time identification samples or stage division samples.

Using the optimal time series selection based on the IoU and the confidence method, the optimal air door opening and closing time series data (S_o^{op}) and its confidence (c_o^{op}) can be selected from the time identification samples $\left\{ G_o^{2\prime}, G_o^{4\prime}, \dots, G_o^{2s\prime} \right\}$. The optimal air door fixed angle opening time series data (S_f^{op}) and its confidence (c_f^{op}) can be selected from the stage division samples $\left\{ G_f^{2\prime}, G_f^{4\prime}, \dots, G_f^{2s\prime} \right\}$.

Correction

Four regression models are established in this method. $R - \triangle t_{os}^{op}$ and $R - \triangle t_{oe}^{op}$ models with the start time correction $\triangle t_{os}^{op}$ and end time correction $\triangle t_{oe}^{op}$ of the optimal air door opening and closing time series data S_o^{op}. $R - \triangle t_{os}^{op}$ and $R - \triangle t_{oe}^{op}$ models with the start time correction $\triangle t_{fs}^{op}$ and end time correction $\triangle t_{fe}^{op}$ of the optimal air door opening and closing time series data S_f^{op}. In the training or testing of these regression models, the feature vector can be expressed using Eq. (13):

$$C_l = \left(t_{os}^{op}, t_{oe}^{op}, T_o^{op}, \overline{x}_o^{op}, \sigma_o^{op}, c_o^{op}, t_{fs}^{op}, t_{fe}^{op}, T_f^{op}, \overline{x}_f^{op}, \sigma_f^{op}, c_f^{op} \right)^T \tag{13}$$

where C_l represents the feature vector, T_o^{op} is the duration of the optimal air door opening and closing time series, $\overline{x}_o^{op}$ is the mean of the optimal air door opening and closing time series, σ_o^{op} is the mean of the optimal air door opening and closing time series, T_f^{op} is the start time of the optimal air door fix angle opening time series having a fixed angle, $\overline{x}_f^{op}$ is the mean of the optimal air door fix angle opening time series, and σ_f^{op} is the mean of the optimal air door fix angle opening time series.

By applying four corrections for the time $\{\Delta t_{os}^{op}, \Delta t_{oe}^{op}, \Delta t_{fs}^{op}, \Delta t_{fe}^{op}\}$, the four times $\left\{ t_{os}^{op}, t_{oe}^{op} t_{fs}^{op}, t_{fe}^{op} \right\}$ are corrected as $\left\{ t_{os}^{op\prime}, t_{oe}^{op\prime}, t_{fs}^{op\prime}, t_{fe}^{op\prime} \right\}$. The air door opening and closing time series data $S_o^{op\prime}$ can be located by $\left[t_{os}^{op\prime}, t_{oe}^{op\prime} \right]$. The air door opening stage $S_{oo}^{op\prime}$, air door fixed angle opening stage $S_f^{op\prime}$, and air door closing stage $S_c^{op\prime}$ can be characterized by $[t_{os}^{op\prime}, t_{fs}^{op\prime}]$, $[t_{fs}^{op\prime}, t_{fe}^{op\prime}]$, and $[t_{fe}^{op\prime}, t_{oe}^{op\prime}]$, respectively.

Experimental studies

In this section, some experiments are conducted to verify the effectiveness of the proposed method. This section includes three parts: (1) the experimental system and data description, (2) the evaluation indices of the proposed method, and (3) experimental studies on the identification and division of air door opening and closing times.

Experimental system and data description

This section is composed of two parts. The first part explains the principles of the experimental system design and the equipment configuration. The second part describes the experimental data.

Experimental system

The experimental model is designed according to the numerical simulation model and the flow similarity principle. To satisfy the geometric similarity between the experimental model and the original model, the overall similarity scale was taken as 1:16, and the rate of change in the length direction was taken as 2[30].

According to the literature[30–32], within two geometrically similar models, the flow field enters the second self-simulation zone when the Euler number (EU) is independent of the Reynolds number (RE), satisfying the flow similarity principle. We explored the similarity between the EU and RE by changing the wind speed to obtain the relationship between them in the experimental model and the numerical simulation model. Figure 7 shows the results of the EU with the RE within the experimental and numerical simulation models.
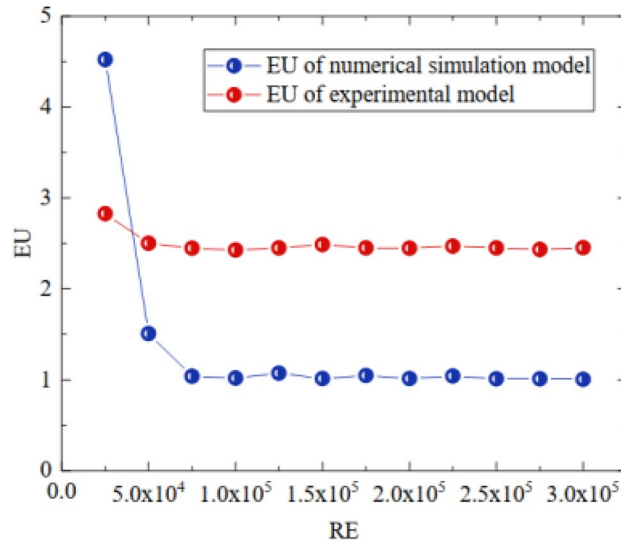
Figure 7. Curves of the variation in EU with respect to RE within the numerical simulation model and the experimental model.

Figure 7 shows that the Euler numbers of both the experimental and numerical simulation models do not change when the Reynolds number is greater than 0.75×10^5. Therefore, when the inlet wind speed is greater than 7.9 m/s in the numerical simulation model and greater than 0.49 m/s in the experimental model, the dynamics of the two flow fields can be considered similar.

According to the above, an experimental system with variable air door opening and closing parameters was designed, as shown in Fig. 8. Figure 8a shows the size and principles of the experimental system, whereas Fig. 8b is an entity diagram of the experimental system.

The experimental system included a parallel connection ventilation pipeline, an air supply system, a wind speed monitoring system, and an electric air door. The wind speed monitoring system was arranged according to the position regulations of the sensors in the mine. The parallel connection ventilation pipeline was made using
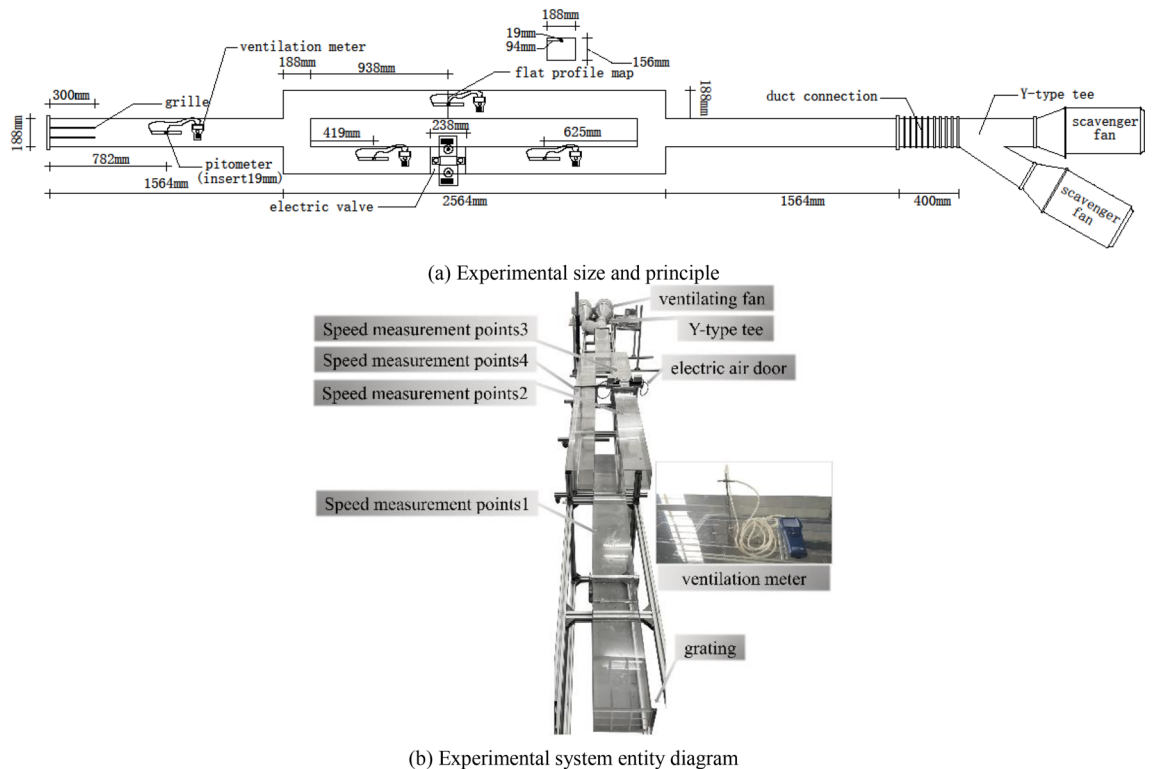

(a) Experimental size and principle

(b) Experimental system entity diagram

Figure 8. Experimental system.

a 4.8 mm thick acrylic board. The air supply system consisted of two ventilation fans connected in parallel, with a maximum ventilation capacity and rotating speed of 2206 m³/h and 2350 r/min, respectively, for every fan. The wind speed monitoring system is composed of four ventilation meter (model TSI-9565). The electric air door's maximum opening angle was 90°, and the fastest opening and closing speed was 15°/s.

Experimental data
There were four variable parameters for the experimental system: inlet wind speed, air door opening and closing speed, air door opening and closing angle, and air door fixed angle opening time. Under the premise that the air door starts opening at 10 s, experiments were carried out for 240 working conditions by arranging and combining the different values of the four parameters, as shown in Table 3.

Some of the working conditions are shown in Table 4, and their data are displayed in Fig. 9.

According to Fig. 9, the abnormal fluctuation time and amplitude of the wind speed data at each measurement point under different working conditions are different, but the trend is consistent with that of the flow field in the numerical simulation.

Evaluation indices of time identification and stage division
To verify the identification effect of the proposed method on the air door opening and closing time, three indicators, namely, the accuracy (*AC*), precision (*PR*), and recall (*RE*), were selected. The average value of these indicators in every air door opening and closing stage was taken to evaluate the effect of the opening and closing stage division. These were calculated using the numbers of true positives (*TP*), true negatives (*TN*), false negatives (*FN*), and false positives (*FP*). True positive (TP) refers to instances correctly identified as the event of interest. True negative (TN) refers to instances correctly identified as not being the event of interest. False negative (FN) refers to instances where the event of interest is not identified when it should have been. False positive (FP) refers to instances incorrectly identified as the event of interest when they are not. Equations (14)–(16) are the expressions used in these calculations.

$$AC = \frac{TP + TN}{TP + FP + TN + FN} \tag{14}$$

$$PR = \frac{TP}{FP + TP} \tag{15}$$

$$SE = \frac{TP}{TP + FN} \tag{16}$$

Experimental studies
In this section, 960 experimental data points representing 240 working conditions and 4 speed measurement points were used as the dataset. Fifty per cent of the dataset was used as the training set for the classification model, 30% was used as the training set for the regression model, and 20% was used as the test set for the overall

Parameter	Values
1	3°/s, 5°/s, 10°/s, 15°/s
2	5 s, 10 s, 20 s, 30 s, 40 s
3	45°, 60°, 75°, 90°
4	8.5 m/s, 9.5 m/s, 10.5 m/s

Table 3. Specific parameters of each component of the air door opening and closing factors.

Work condition	Parameter 1	Parameter 2 (s)	Parameter 3	Parameter 4 (m/s)
1	3°/s	5	60°	9.5
2	0°/s	0	0°	10.5
3	3°/s	5	60°	10.5
4	10°/s	5	60°	10.5
5	3°/s	30	60°	10.5
6	3°/s	5	90°	10.5
7	10°/s	30	90°	10.5

Table 4. Setting of each parameter condition.

(a) Wind speed data of the tandem tunnel where the air door is located

(b) Wind speed data of the Parallel tunnel where the air door is located

(c) Wind speed data on the ventilation side of the tunnel where the air door is located

(d) Wind speed data on the windward side of the tunnel where the air door is located
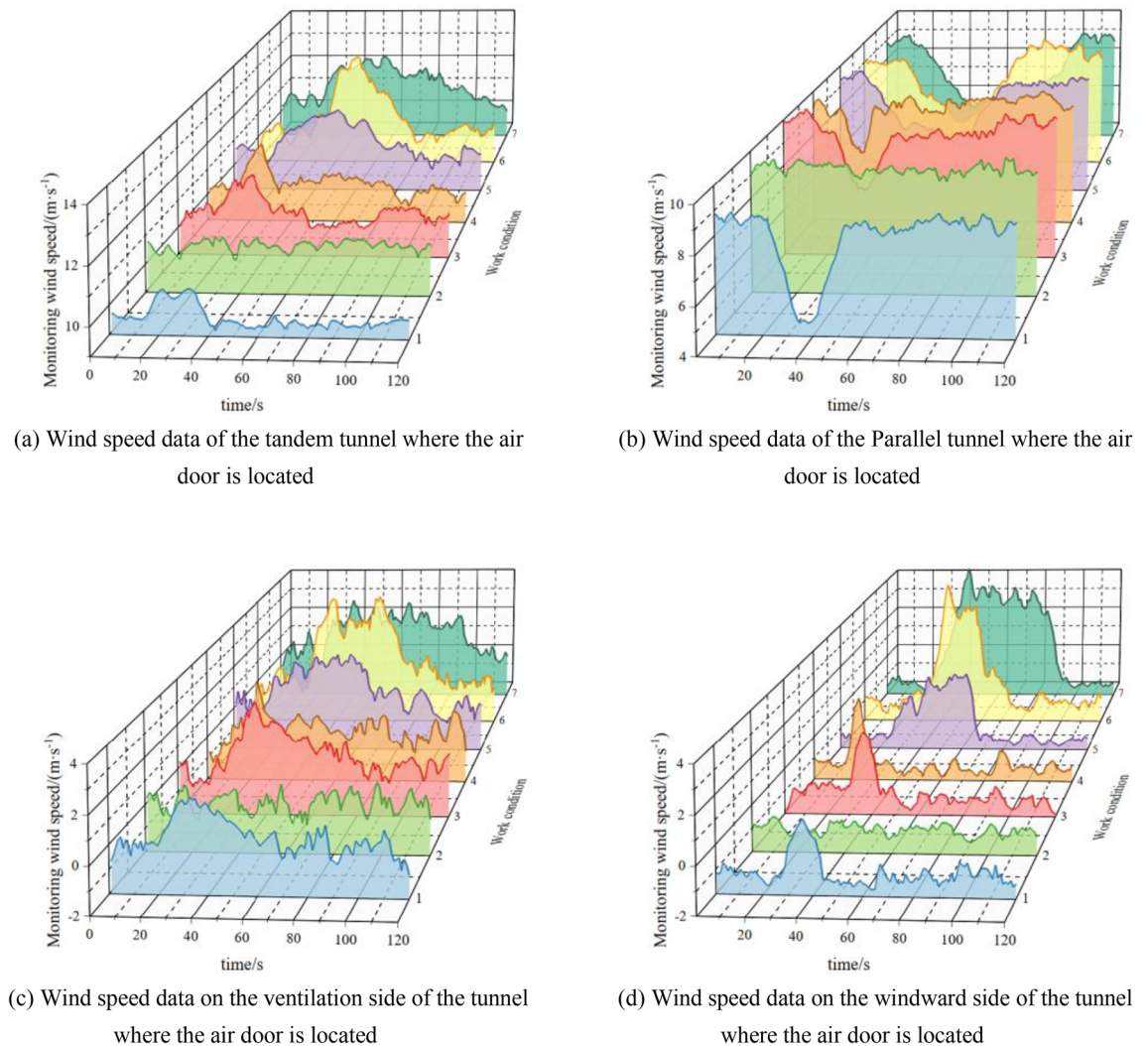
Figure 9. Variation in the wind speed data at each position with different parameters.

method. To avoid any experimental bias, 10 cross-validations of the method effect were performed after each parameter change.

A comparison of the effects of time identification and stage classification of the ten model combinations was conducted to select the optimal combination of the classification and regression models. Among the ten model combinations, the classification models that were used were SVM[33], random forest (RF)[34], gradient boosting decision tree (GBDT)[35], Bayesian network (BN)[36], and backpropagation neural network (BPNN)[37]. The LASSO[38] and elastic net regression (ENR)[39] regression models were used.

A comparison of each indicator is shown in Fig. 10. The accuracy, precision, and recall of this method for air door opening and closing time identification and stage division are optimal when using the SVM model for classification and the LASSO model for regression.

To determine the optimal number of sliding windows, the effects of the method when the sliding window scale varied from 1 to 9 were compared. A comparison of all the indicators is shown in Fig. 11. The results show that when the number of sliding windows reaches 8, the accuracy, precision, and recall reach stability. Therefore, the optimal number of sliding windows is 8.

Through the above work, the optimal parameters of the method were obtained for air door opening and closing time identification and stage division. The accuracy, precision, and recall rate of the method using the optimal parameters for air door opening and closing time identification and stage division are above 90% and
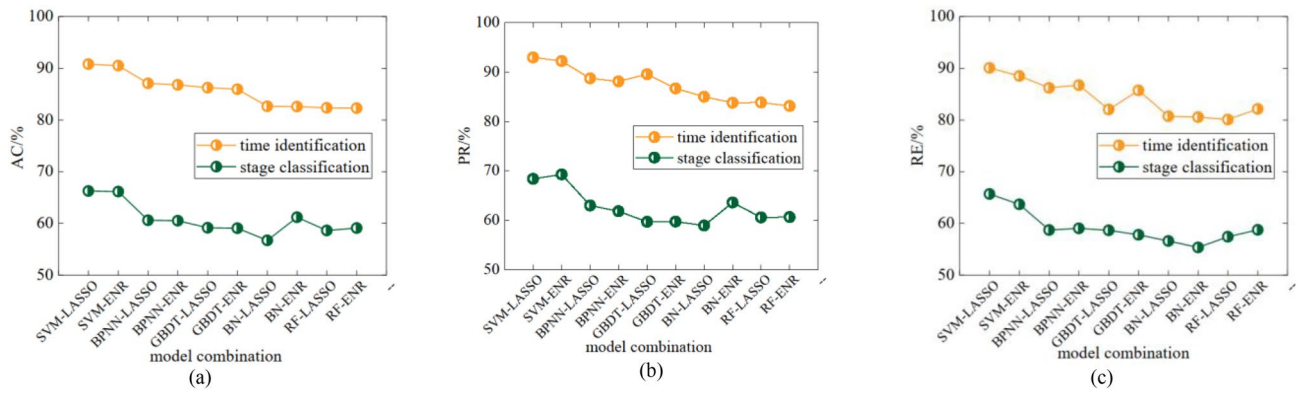
Figure 10. Individual indicator comparative analysis of different method combinations when the number of sliding window scales is 8.
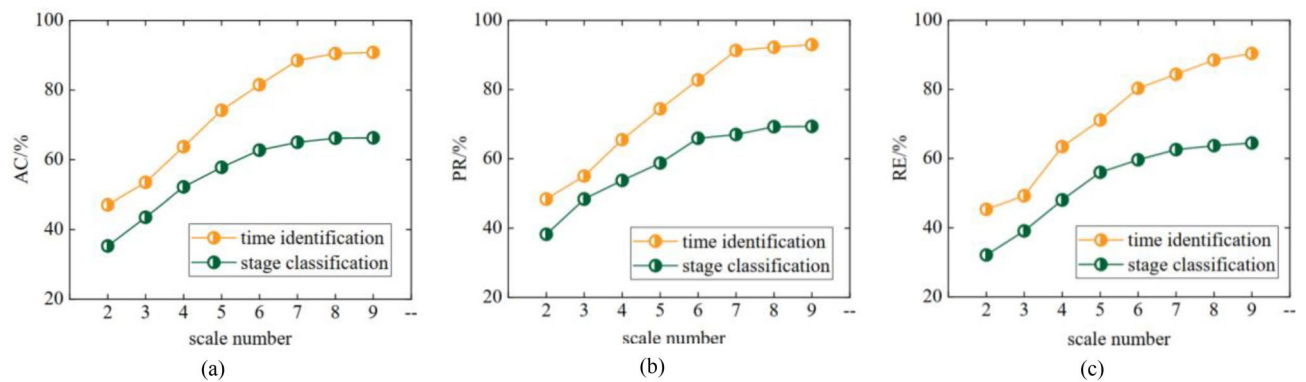

Figure 11. Individual indicator comparative analysis of different sliding window scale numbers when the method combination is SVM-LASSO.

62%, respectively. Figure 12 shows the effect of the method on the time identification and phase classification for some data.

Conclusion

In this paper, a numerical simulation of the air door opening and closing process was conducted. The simulation results suggest the applicability of the proposed method in identifying the time of air door opening and closing, which can further be used for dividing the stages. Furthermore, this method was verified using an experimental approach. This can assist in realizing a basic theory for intelligent mine ventilation.

The proposed method in this study has four important implications. First, the method is based on single wind speed sensor data for air door opening and closing time identification and stage division, using less information to obtain more data while reducing the dependence on both the number and location of sensors. Second, the proposed feature extraction method, which is based on the DWT and statistical methods, can mine local fluctuation information and global information for subtime series data, and the extracted features are rich and interpretable. Third, the sequence of steps that need to be identified and then corrected can accurately solve the inconsistency between the air door opening and closing times and the flow field change time. Finally, with some additional improvements in the framework used in this method, it can be applied to the identification of other production activities that can cause abnormal fluctuations in wind speed monitoring data, such as mine car operation and cage hoisting.

The current method is suitable only for opening and closing air doors. Other production activities, time identification and stage division, and multiproduction activity classification are topics that should be considered in future investigations.
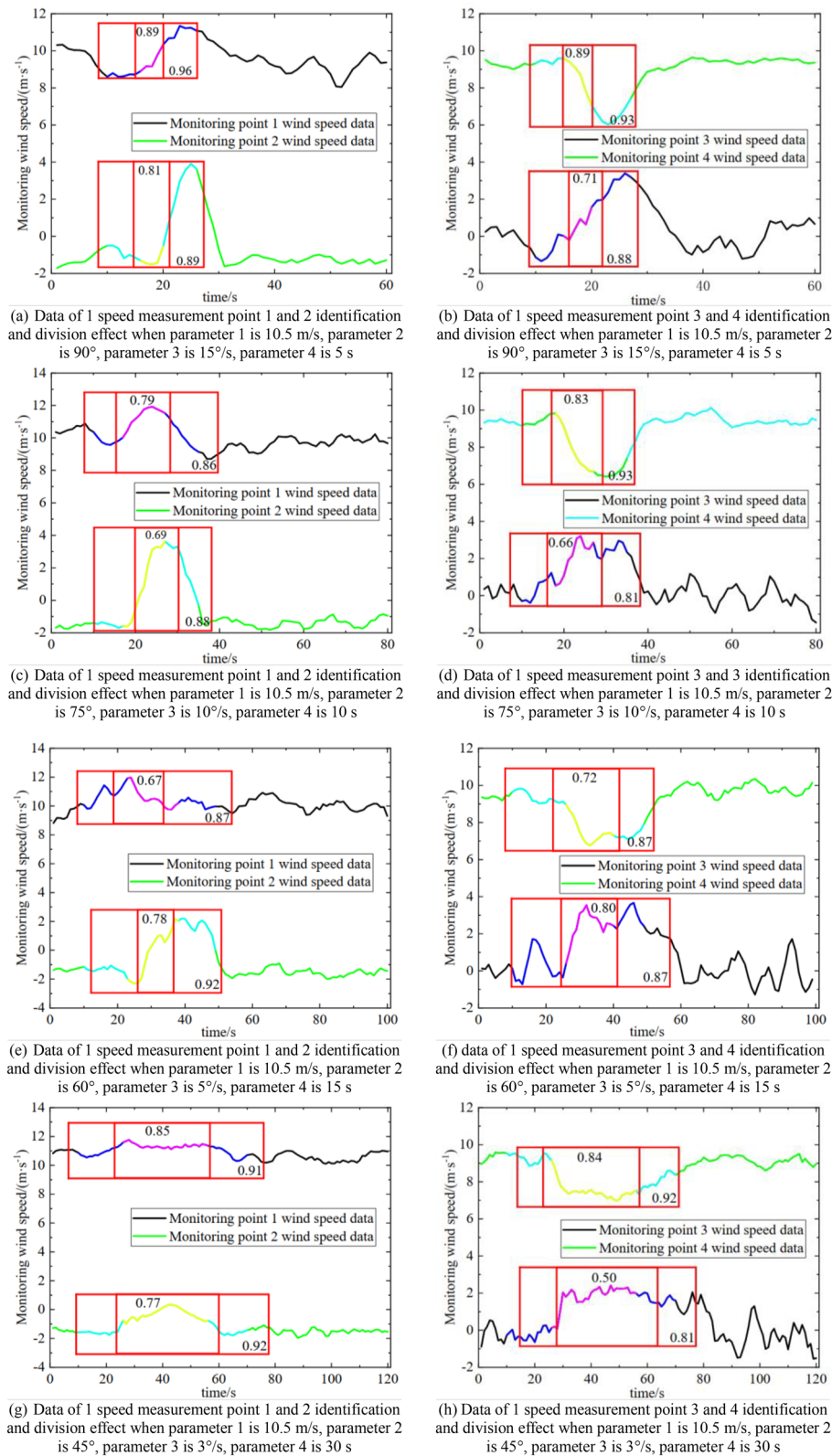
(a) Data of 1 speed measurement point 1 and 2 identification and division effect when parameter 1 is 10.5 m/s, parameter 2 is 90°, parameter 3 is 15°/s, parameter 4 is 5 s

(b) Data of 1 speed measurement point 3 and 4 identification and division effect when parameter 1 is 10.5 m/s, parameter 2 is 90°, parameter 3 is 15°/s, parameter 4 is 5 s

(c) Data of 1 speed measurement point 1 and 2 identification and division effect when parameter 1 is 10.5 m/s, parameter 2 is 75°, parameter 3 is 10°/s, parameter 4 is 10 s

(d) Data of 1 speed measurement point 3 and 3 identification and division effect when parameter 1 is 10.5 m/s, parameter 2 is 75°, parameter 3 is 10°/s, parameter 4 is 10 s

(e) Data of 1 speed measurement point 1 and 2 identification and division effect when parameter 1 is 10.5 m/s, parameter 2 is 60°, parameter 3 is 5°/s, parameter 4 is 15 s

(f) data of 1 speed measurement point 3 and 4 identification and division effect when parameter 1 is 10.5 m/s, parameter 2 is 60°, parameter 3 is 5°/s, parameter 4 is 15 s

(g) Data of 1 speed measurement point 1 and 2 identification and division effect when parameter 1 is 10.5 m/s, parameter 2 is 45°, parameter 3 is 3°/s, parameter 4 is 30 s

(h) Data of 1 speed measurement point 3 and 4 identification and division effect when parameter 1 is 10.5 m/s, parameter 2 is 45°, parameter 3 is 3°/s, parameter 4 is 30 s

Figure 12. Part of the working conditions wind speed data time for air door opening and closing recognition and stage division effect diagram.

Data availability

A summary of the data used in this study is included in the paper, and a detailed data sample will be available upon request by contacting the corresponding author (201613602@sdtbu.edu.cn).

References

1. Sui, X., Wang, L., Miao, D. The design of mine air door based on the finite element analysis. In *Second International Conference on Mechanic Automation and Control Engineering* 804–806. https://doi.org/10.1109/MACE.2011.5987049 (2011).
2. Acuña, E. I. & Lowndes, I. S. A review of primary mine ventilation system optimization. *Interfaces* **44**(2), 163–175. https://doi.org/10.1287/inte.2014.0736 (2014).
3. Chen, W. & Wang, X. Coal mine safety intelligent monitoring based on wireless sensor network. *IEEE Sens. J.* **21**(22), 25465–25471. https://doi.org/10.1109/JSEN.2020.3046287 (2020).
4. Shang, W., Deng, L., Liu, J. & Zhou, Y. Multi-disturbance identification from mine wind-velocity data based on MSSW and WPT-GBDT. *PLoS One* **18**(4), 0284316. https://doi.org/10.1371/journal.pone.0284316 (2023).
5. Liu, J., Li, X. & Song, Y. Experiment study on uncertainty mechanism of mine air velocity and pressure with non-external disturbance. *J China Coal Soc* **41**(06), 1447–1453. https://doi.org/10.13225/j.cnki.jccs.2015.1347 (2016).
6. Wallace, K., Prosser, B. & Stinnette, J. D. The practice of mine ventilation engineering. *Int. J. Min. Sci. Technol.* **25**(2), 165–169. https://doi.org/10.1016/j.ijmst.2015.02.001 (2015).
7. De Souza, E. Application of ventilation management programs for improved mine safety. *Int. J. Min. Sci. Technol.* **27**(4), 647–650. https://doi.org/10.1016/j.ijmst.2017.05.018 (2017).
8. Liu, J. *et al.* Overview on key scientific and technical issues of mine intelligent Ventilatio. *Saf. Coal Mines* **51**(1), 108–117. https://doi.org/10.13347/j.cnki.mkaq.2020.10.016 (2020).
9. Esling, P. & Agon, C. Time-series data mining. *ACM Comput. Surv.* **45**(1), 1–34. https://doi.org/10.1145/2379776.2379788 (2012).
10. Fu, T. *et al.* A review on time series data mining. *Eng. Appl. Artif. Intell.* **24**(1), 164–181. https://doi.org/10.1016/j.engappai.2010.09.007 (2011).
11. Zhou, P. *et al.* Data-driven monitoring and diagnosing of abnormal furnace conditions in blast furnace ironmaking: An integrated PCA-ICA method. *IEEE Trans. Ind. Electron.* **68**(1), 622–631. https://doi.org/10.1109/TIE.2020.2967708 (2020).
12. Hsu, C. Y. & Liu, W. C. Multiple time-series convolutional neural network for fault detection and diagnosis and empirical study in semiconductor manufacturing. *J. Intell. Manuf.* **32**(3), 823–836. https://doi.org/10.1007/s10845-020-01591-0 (2021).
13. Soriano-Vargas, A. A visual analytics approach to anomaly detection in hydrocarbon reservoir time series data. *J. Petrol. Sci. Eng.* **206**, 108988. https://doi.org/10.1016/j.petrol.2021.108988 (2021).
14. He, Y., Peng, Y., Wang, S. & Liu, D. ADMOST: UAV flight data anomaly detection and mitigation via online subspace tracking. *IEEE Trans. Instrum. Meas.* **68**(4), 1035–1044. https://doi.org/10.1109/TIM.2018.2863499 (2018).
15. Rodriguez, M. A., Sotomonte, J. F., Cifuentes, J., Bueno-López, M. Power quality disturbance classification via deep convolutional auto-encoders and stacked LSTM recurrent neural networks. In *International Conference on Smart Energy Systems and Technologies* 1–6. https://doi.org/10.1109/SEST48500.2020.9203082 (2020).
16. Cao, J. *et al.* Urban noise recognition with convolutional neural network. *Multimed. Tools Appl.* **78**(20), 29021–29041. https://doi.org/10.1007/s11042-018-6295-8 (2019).
17. Zang, D., Liu, J. & Qu, F. Pipeline small leak detection based on virtual sample generation and unified feature extraction. *Measurement* **184**, 109960. https://doi.org/10.1016/j.measurement.2021.109960 (2021).
18. Huang, D., Liu, J. & Deng, L. A hybrid-encoding adaptive evolutionary strategy algorithm for windage alteration fault diagnosis. *Process Saf. Environ. Protect.* **136**, 242–252. https://doi.org/10.1016/j.psep.2020.01.037 (2020).
19. Zhao, D. & Shen, Z. Study on roadway fault diagnosis of the mine ventilation system based on improved SVM. *Min. Metall. Explor.* **39**(3), 983–992. https://doi.org/10.1007/s42461-022-00595-z (2022).
20. Liu, X., Liu, J., Wang, D. & Zhao, L. Experimental and numerical simulation investigations of an axial flow fan performance in high-altitude environments. *Energy* **234**, 121281. https://doi.org/10.1016/j.energy.2021.121281 (2021).
21. Carneiro, R., Gaspar, P. D., Silva, P. D. & Domingues, L. C. Numerical modelling of cold room's hinged door opening and closing processes. *AIP Conf. Proc.* **1738**(1), 370004. https://doi.org/10.1063/1.4952149 (2016).
22. Carneiro, R., Gaspar, P. D. & Silva, P. D. 3D and transient numerical modelling of door opening and closing processes and its influence on thermal performance of cold rooms. *Appl. Therm. Eng.* **113**, 585–600. https://doi.org/10.1016/j.applthermaleng.2016.11.046 (2017).
23. Li, C., Xie, Z., Qin, Y., Jia, L. & Chen, Q. A multi-scale image and dynamic candidate region-based automatic detection of foreign targets intruding the railway perimeter. *Measurement* **185**, 109853. https://doi.org/10.1016/j.measurement.2021.109853 (2021).
24. Niemczewska-Wojcik, M. & Wojcik, A. The multi-scale analysis of ceramic surface topography created in abrasive machining process. *Measurement* **166**, 108217. https://doi.org/10.1016/j.measurement.2020.108217 (2020).
25. Zheng, L., Wei, H. & Yong, M. Raw wind data preprocessing: A data-mining approach. *IEEE Trans. Sustain. Energy* **6**(1), 11–19. https://doi.org/10.1109/TSTE.2014.2355837 (2014).
26. Singh, D. & Singh, B. Investigating the impact of data normalization on classification performance. *Appl. Soft Comput.* **97**, 105524. https://doi.org/10.1016/j.asoc.2019.105524 (2020).
27. Tsai, C. F. & Chen, Y. C. The optimal combination of feature selection and data discretization: An empirical study. *Inf. Sci.* **505**, 282–293. https://doi.org/10.1016/j.ins.2019.07.091 (2019).
28. Maldonado, S., Gonzalez, A. & Crone, S. Automatic time series analysis for electric load forecasting via support vector regression. *Appl. Soft Comput.* **83**, 105616. https://doi.org/10.1016/j.asoc.2019.105616 (2019).
29. Shawon, S. M. A. Z. *et al.* Surface modified hybrid $ZnSnO_3$ nanocubes for enhanced piezoelectric power generation and wireless sensory application. *Nano Energy* **92**, 106653. https://doi.org/10.1016/j.nanoen.2021.106653(2022) (2022).
30. Wang, Y., Xie, Y., Liu, H. & Ren, R. Physical model experiment on semi transverse ventilation air inlet and outlet of subsea tunnel. *China J. Highw. Transp.* **23**(03), 76–82. https://doi.org/10.19721/j.cnki.1001-7372.2010.03.012 (2010).
31. Li, F., Luo, J., Wang, L., Guo, D. & Gao, L. Analysis of aerodynamic effects and load spectrum characteristics in high-speed railway tunnels. *J. Wind Eng. Ind. Aerodyn.* **216**, 104729. https://doi.org/10.1016/j.jweia.2021.104729 (2021).
32. Xin, B., Dang, W., Yan, X., Yu, J. & Bai, Y. Dispersion characteristics and hazard area prediction of mixed natural gas based on wind tunnel experiments and risk theory. *Process Saf. Environ. Protect.* **152**, 278–290. https://doi.org/10.1016/j.psep.2021.06.012 (2021).
33. Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297. https://doi.org/10.1007/BF00994018 (1995).
34. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32. https://doi.org/10.1023/A:1010933404324 (2001).
35. Friedman, J. H. Greedy function approximation: a gradient boosting machine. *Ann. Stat.* **29**, 1189–1232 (2001).
36. Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (Morgan Kaufmann, Burlington, 1988).

37. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**(6088), 533–536. https://doi.org/10.1038/323533a0 (1986).
38. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **58**(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x (1996).
39. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **67**(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x (2005).

Author contributions

Y.S. conceptualized this project and revised the manuscript; S.L. and W.S. carried out the simulation study and drafted the manuscript.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.