Health status identification of catenary based on VMD and FA-ELM

Lingzhi Yi, You Guo, Nian Liu, Jian Zhao, Wang Li and Junyong Sun

Abstract
Catenary works as a key part in the electric railway traction power supply system, which is exposed outdoors for a long time and the failure rate is very high. Once a failure occurs, it will directly affect the driving safety. Based on the above, a model of identifying the health status for the catenary based on firefly algorithm optimized extreme learning machine combined with variational mode decomposition is proposed in this paper. Variational mode decomposition is used to decompose the original detection curve of catenary into a series of intrinsic mode function components, and the intrinsic mode function components filtered by the correlation coefficient method after decomposing each detection curve are input into the firefly algorithm optimized extreme learning machine model to realize health status identification. Compared with some other models, the results show that the proposed model has better health status identification effect.

Keywords
Catenary, firefly algorithm, variational mode decomposition, extreme learning machine, health status identification

Received 3 September 2020; revised 6 March 2021; accepted 26 May 2021

Introduction
Catenary is one of the “three components” in electrified railway, and its working status is directly related to the safety of train operation. Catenary erects along the railway track and arranges outdoors. It is vulnerable to the wind, sun and impact of locomotive pantograph of high-speed train, which becomes the main point of failure in the traction power supply system. At present, most of railways use catenary inspection vehicles to detect catenary faults. The fault judgments are made by humans based on experience, and their efficiency and accuracy cannot meet the requirements of high-speed and large transportation volume. Therefore, it is particularly important to find an accurate and efficient method for identifying the health status of the catenary.

Fault identification has always been a research hotspot of scholars. Zhang and Chen presented a rolling bearing fault identification method based on empirical mode decomposition (EMD) and spectral kurtosis. First, the fault signal was decomposed by EMD to obtain multiple mode functions, and then suitable intrinsic mode function (IMF) components were selected for reconstruction to achieve a certain denoising effect, but EMD is an empirical decomposition method, which is prone to end effects and mode mixing problem, so the identification results will be affected. Liu et al. applied ensemble empirical mode decomposition (EEMD) to extract the fault characteristics of train bogie bearings. Although EEMD effectively suppressed the mode mixing problem of EMD, it required multiple EMD processing, which increased the running time. Wu et al. presented a fault feature extraction method that combining variational mode decomposition (VMD) and dispersion entropy, the waveform method was used to determine the mode.
number \( K \), and the dispersive entropy of decomposition signal was input as eigenvalue into the support vector machine (SVM) to realize fault identification. This method resolved the problem of mode mixing. However, multiple classifiers need to be constructed when SVM implements multi-classification, and the running time is also too long. Geng et al.\(^8\) proposed an energy-saving prediction model of analytic hierarchy process (AHP) combined with extreme learning machine (ELM), which achieved good results. Because the parameters of ELM are obtained randomly without repeated iterations, the running time is greatly shortened. However, due to the random parameters, the accuracy is unstable. Yi et al.\(^9\) proposed a method of contact network fault identification using particle swarm optimization (PSO) to optimize the extreme learning machine, which improved the identification accuracy. However, as a typical optimization algorithm, particle swarm optimization may fall into local optimum in the process of optimizing the parameters of the extreme learning machine. Pei et al.\(^10\) identified failures by installing failure indicators on the catenary, but they were vulnerable to environmental, climatic, and other natural factors.

To solve these problems, a catenary health status identification method based on VMD and ELM optimized by firefly algorithm (FA) is proposed in this paper. Firstly, the original detection curve of contact wire height, contact pressure and contact wire gradient are decomposed by VMD, and then appropriate IMF components are selected by correlation coefficient method, then the selected components are normalized. Finally, normalized data are classified by FA-optimized ELM to achieve health status identification. This method is an offline status identification method. The method can identify the health status of catenary according to the data collected in advance, and provide effective reference for maintenance personnel.

**Firefly optimization algorithm**

The Firefly algorithm is a swarm intelligence heuristic algorithm proposed by Professor Xin-She Yang of Cambridge University.\(^11,12\) The algorithm simulates the biological characteristics of fireflies, and then makes its companions aware of it. The brightness of fireflies is used as the criterion for judging whether its position is optimal. These fireflies search for their companions nearby and move to the brighter companions, thus achieving iteration and optimization of the location of fireflies. Therefore, the position of firefly is equal to the size of the objective function in solving the problem. In the process of iteration and optimization, the target is optimized by moving to the companions with higher brightness. The steps are as follows:

**Step 1:** Initial parameter settings: population size, spatial dimension, maximum number of iterations, and so on.

**Step 2:** Randomly initializes the position of fireflies. Calculate the relative brightness \( I \) and attraction \( \varepsilon \) of each firefly in the population. The position updating direction of fireflies is determined according to the relative brightness \( I \). The calculation formulas of \( I \) and \( \varepsilon \) are as follows:

\[
I = I_S \times e^{-\gamma t} \quad \varepsilon = \varepsilon_S \times e^{-\gamma t^2}
\]

\[
r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}
\]

where \( I_S \) is the biggest brightness of firefly at the light source, \( \varepsilon_S \) is the biggest attractiveness in the light source, \( \gamma \) is the light intensity absorbed coefficient that is different in different media, \( r_{ij} \) is the distance between firefly \( i \) and firefly \( j \).

**Step 3:** During the iterations, firefly \( i \) moves towards the brighter firefly \( j \). The formula for updating the position is shown in equation (3)

\[
x_{i+1} = \delta(t) \times x_i + \varepsilon \times (x_j - x_i) + \alpha \times (\text{rand} - 0.5)
\]

where \( \delta(t) \) is the inertia weight function; the second term describes the effect of attraction, \( x_i \) and \( x_j \) are the positions of two fireflies respectively; the third term is the random disturbance term, \( \alpha \) is the random parameter belonging to \((0, 1)\), and “rand” is the random number between \([0,1]\).

In order to obtain the best global search ability, inertia weight function \( \delta(t) \) is introduced:

\[
\delta(t) = \delta_1 + (\delta_2 - \delta_1) \frac{K}{K}
\]

where \( \delta_1 \) and \( \delta_2 \) are the minimum and maximum values of \( \delta \), \( K \) is the current iteration number, and \( K \) is the maximum iteration number.

**Step 4:** Re-calculate the brightness of each firefly after the position is updated. When the stop condition is met or the maximum number of iterations is achieved, output the final position of each firefly, otherwise, go to step 3 and continue to iterate and optimize.
Finally, fireflies’ final position $x$ are the optimal parameters. In this paper, FA is used to select the optimal input weight matrix $\omega$ and hidden layer bias matrix $b$ of ELM. It can avoid the divergence and low identification accuracy caused by random generation of $\omega$ and $b$ in traditional ELM.

In order to verify the convergence of FA, the following test functions are used to verify the convergence of FA:

$$f_1(x, y) = \exp\left(-\left(x - 4\right)^2 - \left(y - 4\right)^2\right) + \exp\left(-\left(x + 4\right)^2 - \left(y - 4\right)^2\right) + 2\exp\left(-x^2 - y^2\right), \quad |x| \leq 5, |y| \leq 5$$

First, 50 fireflies are arranged randomly. The initial population is shown in Figure 1, and the solid dots represent fireflies.

After testing, through the attraction of the brighter fireflies, all fireflies will gather near the brighter fireflies. The results are shown in Figure 2, which shows that FA has good convergence.

| Algorithms | Parameter setting |
|------------|------------------|
| FA         | Step factor is set as 0.2; biggest attractiveness is set as 1; light intensity absorbed coefficient is set as 1 |
| DEA        | Variation coefficient is set as 0.3; crossover operator is set as 0.8 |
| GA         | Encoding length is set as 20; crossover probability is set as 0.8; variation probability is set as 0.05 |

Table 1. Parameter setting of 3 algorithms.

| Parameter setting |
|-------------------|
| $x$    | $\tau$ | tol |
| 2000   | 0      | 1.0E-7 |

Table 2. VMD parameter configuration.

In order to further verify the optimization performance of FA, FA was compared with differential evolution algorithm (DEA) and genetic algorithm (GA). Test functions $f_1(x, y)$ and $f_2(x, y)$ was selected to show optimizations of 3 algorithms. The
maximum value of $f_1(x, y)$ is 2, and the maximum value of $f_2(x, y)$ is 1.

$$f_2(x, y) = 0.5 - \frac{\left(\sin \sqrt{x^2 + y^2}\right)^2}{1 + 0.001 \times (x^2 + y^2)}$$  \hspace{1cm} (6)

The parameter setting of 3 algorithms is shown in Table 1. Where the population size is set as 50 and the iterations is set as 40.

The optimization results of 3 algorithms for test functions $f_1(x, y)$ and $f_2(x, y)$ are shown in Figures 3 and 4 respectively.

It can be seen from Figures 3 and 4 that FA is superior to DEA and GA in convergence accuracy and convergence speed, and has better optimization ability. Therefore, FA is used to optimize the parameters of ELM in this paper.

VMD-FA-ELM health state identification model

Variational mode decomposition

VMD is a new complex signal decomposition method proposed by K. Dragomiretsk and D. Zosso on the basis of EMD.\textsuperscript{13,14} It decomposes the signal into limited bandwidth with different center frequencies according to the preset number of modes, updates each mode and its center frequency by using alternating direction multiplier, adjusts each mode to the corresponding base frequency band by band, extracts each mode and its corresponding center frequency, and obtains the components with different center frequencies.

In VMD, the desired decomposition mode number $K$ need to be given in advance. VMD is a kind of signal processing method with adjustable scale. By selecting the appropriate $K$ value, the modal aliasing and end effect in EMD can be effectively improved. In the process of signal processing, VMD method is adaptive to nonlinear and nonstationary signals, and can restrain the noise in the signal. Therefore, VMD is used to preprocess the detection curve of catenary in this paper.

The parameters involved in VMD are mode number $K$, penalty factor $\alpha$, noise tolerance $\tau$, and convergence stop condition $tol$. The selection of the decomposition mode number $K$ directly determines the decomposition effect. If the mode number $K$ is too large, the mode mixing or additional noise will be produced, and if it is too small, the mode will be under-decomposed.\textsuperscript{15} In this paper, the mode number $K$ is selected by observing the center frequency. Other parameters are configured by default as shown in Table 2.

The center frequencies of the IMF components decomposed by VMD are distributed from low to high. If the mode number $K$ is taken from small to large, the problem of insufficient decomposition will not occur when the center frequency of the last IMF component reaches its maximum for the first time. If the value of the maximum center frequency remains

| $K$ | IMF1    | IMF2   | IMF3   | IMF4   | IMF5   | IMF6 | IMF7 |
|-----|---------|--------|--------|--------|--------|------|------|
| 2   | 3.72E-07| 0.2516 |        |        |        |      |      |
| 3   | 3.72E-07| 0.184  | 0.304  |        |        |      |      |
| 4   | 3.72E-07| 0.103  | 0.254  | 0.441  |        |      |      |
| 5   | 3.72E-07| 0.101  | 0.188  | 0.305  | 0.442  |      |      |
| 6   | 3.72E-07| 0.1    | 0.177  | 0.249  | 0.311  | 0.443|      |
| 7   | 3.72E-07| 0.054  | 0.11   | 0.218  | 0.304  | 0.384| 0.446|

| $K$ | IMF1    | IMF2   | IMF3   | IMF4   | IMF5   | IMF6 | IMF7 |
|-----|---------|--------|--------|--------|--------|------|------|
| 2   | 4.0E-06 | 0.27   |        |        |        |      |      |
| 3   | 4.0E-06 | 0.178  | 0.36   |        |        |      |      |
| 4   | 3.92E-06| 0.108  | 0.268  | 0.368  |        |      |      |
| 5   | 3.83E-06| 0.084  | 0.179  | 0.274  | 0.418  |      |      |
| 6   | 3.82E-06| 0.083  | 0.178  | 0.271  | 0.362  | 0.420|      |
| 7   | 3.54E-06| 0.048  | 0.151  | 0.212  | 0.276  | 0.363| 0.420|

| $K$ | IMF1    | IMF2   | IMF3   | IMF4   | IMF5   | IMF6 | IMF7 |
|-----|---------|--------|--------|--------|--------|------|------|
| 2   | 6.4E-05 | 0.19   |        |        |        |      |      |
| 3   | 5.82E-05| 0.052  | 0.3    |        |        |      |      |
| 4   | 5.78E-05| 0.051  | 0.193  | 0.31   |        |      |      |
| 5   | 5.77E-05| 0.051  | 0.187  | 0.3    | 0.41   |      |      |
| 6   | 5.68E-05| 0.047  | 0.123  | 0.2    | 0.302  | 0.41 |      |
| 7   | 5.67E-05| 0.047  | 0.123  | 0.198  | 0.298  | 0.378| 0.415|

The center frequencies of the IMF components decomposed by VMD are distributed from low to high. If the mode number $K$ is taken from small to large, the problem of insufficient decomposition will not occur when the center frequency of the last IMF component reaches its maximum for the first time.
relatively stable when $K$ value increasing, the value of $K$ is considered to be optimal value at this time. When $K$ takes different values, the VMD is carried out by taking the three detection curves of contact wire height, contact pressure and contact wire gradient, and the center frequency of each IMF component varying with distance is obtained as shown in Table 3.

From Table 3, when $K \geq 4$, the maximum value of center frequency of the contact wire height detection curve tends to be stable; when $K \geq 5$, the maximum value of center frequency of contact pressure and contact wire gradient detection curve tends to be stable. Therefore, the $K$ values of contact wire height, contact pressure and contact wire gradient are taken as 4, 5 and 5 respectively, and the decomposed modal components and their frequency spectrum are shown in Figure 5.

The correlation coefficient method is employed to determine whether each IMF component is a noise component. The correlation coefficient formula is shown in equation (7). The correlation coefficient

![Figure 5. VMD decomposition results. (a) Modal components and their spectrum of the detection curve of contact wire height by VMD. (b) Modal components and their spectrum of the detection curve of contact pressure by VMD. (c) Modal components and their spectrum of the detection curve of contact wire gradient by VMD.](image)
between each mode component and the original signal calculated by equation (7) is shown in Table 4.

\[ R(k) = \frac{\sum_{t=1}^{N} (u(t) - \bar{u})(u_k(t) - \bar{u}_k)}{\sqrt{\sum_{t=1}^{N} (u(t) - \bar{u})^2 \times \sum_{t=1}^{N} (u_k(t) - \bar{u}_k)^2}} \]  

(7)

where \( u \) is the original signal, \( u_k \) is the modal component, \( \bar{u} \) and \( \bar{u}_k \) are the mean values of the original signal and modal component \( u_k \) respectively.

As shown in Table 4, for all three kinds of signals, the correlation coefficients of IMF1 components are far greater than those of other mode components. Therefore, the decomposed IMF1 components of the three signals can be used for health status identification.

The comparison curves of the contact wire height, contact pressure and contact wire gradient before and after decomposition are shown in Figure 6. The instability and pulsation of the decomposed IMF1 component are significantly reduced, as shown in Figure 6. The similarity of IMF1 component with the original signal is high, and the denoising effect is well.

**FA optimized ELM**

ELM is an algorithm model proposed by Professor Guangbin Huang for solving the single hidden layer feed-forward neural networks.\textsuperscript{18,19} Compared with the neural network algorithm using gradient descent method to update weights, the output weights of ELM are calculated by least square method, so ELM has better generalization performance and faster learning speed. Compared with deep learning algorithm, ELM takes less computing resources. Therefore, ELM is used to identify the health status of catenary in this paper.

Although ELM has the advantages of fast convergence speed and less parameters to be set, the input weight matrix \( \mathbf{a} \) and hidden layer transmission bias matrix \( \mathbf{b} \) are randomly generated, which may lead classified accuracy to be worse. Therefore, this paper introduces FA to select the best \( \mathbf{a} \) and \( \mathbf{b} \) of ELM, and its optimization process is shown in Figure 7.

**Combination model**

Based on the FA, VMD and ELM model mentioned above, the VMD-FA-ELM health status identification model is built in this paper, as shown in Figure 8, and the steps are as follows:

**Step 1:** Decompose the original detection curve of contact wire height, contact pressure and contact wire gradient by VMD. Determine the decomposition mode number \( K \) by the method of center frequency, and then select the effective mode components by the correlation coefficient method.

**Step 2:** Normalize the IMF1 components after decomposition of the three detection curves to \([-1,1]\) by “premnmx” function. The Function expression is shown in equation (8).

\[ y = \frac{2(x - x_{\text{min}})}{x_{\text{max}} - x_{\text{min}}} - 1 \]  

(8)

where \( x \) is the input data to be normalized; \( x_{\text{max}} \) and \( x_{\text{min}} \) are the maximum and minimum values of input
data respectively; $y$ is the output data after normalization.

**Step 3:** Use FA to select the optimal input weight matrix $\omega$ and hidden layer bias matrix $b$ in ELM. Train and test the normalized data respectively, and then obtain the final identification results.

**Step 4:** Calculate the identification accuracy, mean square error (MSE), decision coefficient ($R^2$) and running time ($t$). Compare the abovementioned results with VMD-GA-ELM, FA-ELM, GA-ELM, VMD-ELM and ELM model.

## Experimental verification

### Identification results

The catenary detection data of a power section in 2019 is selected as the data source. Among them, there are 2400 sampling points for training samples and 2400 sampling points for testing samples. Select three types of catenary health status: contact wire height, contact pressure and contact wire gradient. Then design labels for all data based on 8 indicators as shown in Table 5.

The experimental parameters are set as follows:

Parameters of ELM: The number of input layer nodes is 3, the number of hidden layer nodes is 50, the

### Table 5. Labels of health status.

| Labels | Contact wire height | Contact pressure | Contact wire gradient |
|--------|---------------------|------------------|-----------------------|
| F1     | Normal              | Normal           | Normal                |
| F2     | Fault               | Normal           | Normal                |
| F3     | Normal              | Fault            | Normal                |
| F4     | Normal              | Normal           | Fault                 |
| F5     | Normal              | Fault            | Fault                 |
| F6     | Fault               | Normal           | Fault                 |
| F7     | Fault               | Fault            | Normal                |
| F8     | Fault               | Fault            | Fault                 |
number of output layer nodes is 1, and activate function is “sig” function.

Parameters of FA: The size of firefly is set as 30, the spatial dimension is set as 250, the biggest attractiveness \(s_{\infty}\) is set as 1, the light intensity absorbed coefficient \(c\) is set as 1, the step factor \(a\) is set as 0.5, the maximum number of iterations is set as 100, and

---

**Figure 8.** Structure of VMD-FA-ELM combined model.

---

**Figure 9.** Identification result of VMD-FA-ELM.

---

**Figure 10.** Confusion matrix of VMD-FA-ELM.
the minimum and maximum values of inertia weight $\delta$ are set as 0.3 and 0.9 respectively.

The parameters of VMD are shown in Table 2.

The identification results based on VMD-FA-ELM are shown in Figure 9, and the confusion matrix of identification results is shown in Figure 10. Where the diagonal box represents the number of samples with the same prediction category as the actual category. The remaining squares from F1 to F8 represent the number of samples with wrong identification. The last row of the matrix represents the Accuracy (to be used to measure the accuracy of system identification, which is equal to the number of samples correctly identified/the number of samples identified). The last column of the matrix represents the Recall (to be used to measure the recall rate of the system identification, which is equal to the number of samples correctly identified/the number of real samples). The box in
the bottom right corner represents the total identification accuracy.20

It can be seen from Figure 10 that the identification accuracy of VMD-FA-ELM is 98.75%, and the accuracy and recall rate are higher than 99% except for F5 and F6, which indicate that the new method has good effect on health status identification.

Comparison and analysis

On the basis of section “Identification results”, five models, including VMD-GA-ELM, FA-ELM, GA-ELM, VMD-ELM and ELM, are used to compare the identification effect. The identification effect is reflected by the identification accuracy, MSE, \( R^2 \) and running time.

In order to ensure the fairness and rationality of comparative analysis, the same test samples and training samples as in section “Identification results” are used to test the above five models. The confusion matrices are drawn as shown in Figure 11.

It can be seen from Figures 10 and 11 that the VMD-FA-ELM model has the highest identification accuracy, and both the accuracy and recall rates are above 93%. VMD-FA-ELM has a 28.6% improvement in identification accuracy compared to ELM. The identification accuracy of the VMD-GA-ELM model is second (increased by 22.1%), but there are some cases where the accuracy or recall is low (less than 70%).

Compared with ELM model, the identification accuracy of FA-ELM model and GA-ELM model is improved by 16.1% and 8.8% respectively. Compared with the VMD-ELM model, the identification accuracy of VMD-FA-ELM model and VMD-GA-ELM model is improved by 18.8% and 11.8% respectively. This indicates that FA is better than GA in improving the model. Compared with ELM model and FA-ELM model, the identification accuracy of VMD-ELM model and VMD-FA-ELM model are improved by 9.3% and 11.8% respectively. This shows that denoising by VMD has a certain optimization effect on this model.

Since the effect of each run is random, in order to reconfirm the high identification effect of the VMD-FA-ELM model, this paper tests the above six models ten times respectively. Table 6 shows the running time of six models under ten tests.
Combined with Table 6, Figures 10 and 11, it can be seen that the identification accuracy of VMD-GA-ELM model is slightly lower than that of VMD-FA-ELM model, but its running time is more than twice that of VMD-FA-ELM. This will reduce the identification efficiency and does not meet the requirements of rapid health status identification of catenary. The traditional ELM does not need to decompose the original data, the parameters are obtained randomly, so the running time is the lowest. However, the identification accuracy of traditional ELM is too low to meet the requirements of catenary health status identification.

Figure 12 shows the box-plot of the identification effect of different models after ten tests. The meaning of each box in the figure is the identification accuracy, MSE and $R^2$ of each model under ten tests. The abscissa is the model type.

According to the results of ten runs in Figure 12, VMD-FA-ELM not only has higher identification accuracy than other models, but also the convergence of identification accuracy is better than other models. The MSE of VMD-FA-ELM is significantly lower than other models, and its convergence is also better. And the $R^2$ is better than other models.

Conclusions

In this paper, the health status identification of electrified railway catenary is studied, and a health status identification method based on VMD-FA-ELM is proposed. The conclusions are as follows:

1. Decomposing the original detection curve of the catenary by VMD can stabilize the non-stationary and complex detection curve of the catenary. Experiments have shown that using the decomposed IMF1 component with the highest correlation coefficient for health status identification has a higher identification accuracy than the accuracy before decomposition.

2. FA has good parameters optimization ability. The ability can effectively solve the problem of low identification accuracy caused by random generation of parameters in ELM learning, and makes full use of the excellent identification ability of ELM. Compared with the typical GA algorithm, the optimized effect on the model is more prominent. The comparison with VMD-ELM indicates the effectiveness of optimizing ELM by FA.

3. Experiments indicate that the identification model based on VMD-FA-ELM can accurately and quickly identify the health status of catenary, which provides a new method for identifying the health status of the catenary. It has certain application value in the era of comprehensive popularization of electrified railway.

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 is supported by the National Natural Science Foundation of China (61572416), Hunan province Natural science Zhuzhou United foundation (2020J6009) and Key Laboratory Open Project Fund of Disaster Prevention and Mitigation for Power Grid Transmission and Transformation Equipment.

ORCID iD

You Guo https://orcid.org/0000-0001-9008-7856

References

1. Huang YM, Yuan TC and Yang J. Research on identification method of catenary fault based on SVM. Comput Simul 2018; 35: 145–152.
2. Yang J, Song RG and Fang Y. Stability of the pantograph-catenary system of urban railway vehicle. J China Railw Soc 2011; 33: 31–35.
3. Liu ZG, Liu WQ and Han ZW. A high-precision detection approach for catenary geometry parameters of electrical railway. IEEE Trans Instrum Meas 2017; 66: 1798–1711.
4. Liu ZG, Song Y, Han Y, et al. Advances of research on high-speed railway catenary. J Southwest Jiaotong Univ 2016; 51: 495–518.
5. Zhang C and Chen JJ. A fault diagnosis method of roller bearing based on EMD de-noising and spectral kurtosis. Mech Sci Technol Aerosp Eng 2015; 10: 252–256.
6. Liu JQ, Zhao ZB, Wu CZ, et al. An intelligent fault diagnosis for bogie bearings of trian based wavelet packet decomposition and EEMD. J China Railw Soc 2015; 37: 40–45.
7. Wu SJ, Feng FZ, Wu CZ, et al. Research on fault diagnosis method of tank planetary gearbox based on VMD-DE. J Vib Shock 2020; 39: 170–179.
8. Geng ZQ, Qin L, Han YM, et al. Energy saving and prediction modeling of petrochemical industries. A novel ELM based on FAHP. Energy 2017; 122: 350–362.
9. Yi LZ, Zhao J, Yu WX, et al. Catenary fault identification based on PSO-ELM. JPHCS 2019; 1302: 032017.
10. Pei XQ, Zhang L, Liu Y, et al. Brief analysis of using fault indicator to identify catenary faults. Electr Railw 2013; 24: 37–39.
11. Su HJ, Cai Y and Du Q. Firefly-algorithm-inspired framework with band selection and extreme learning machine for hyperspectral image classification. IEEE J Sel Top Appl Earth Obs Remote Sens 2017; 10: 309–320.
12. Yi LZ, Sui YB and Yu WX. The faults diagnostic analysis for analog circuit faults based on firefly algorithm and extreme learning machine. *J Autom Control Eng* 2016; 4: 413–417.

13. Dragomiretskiy K and Zosso D. Variational mode decomposition. *IEEE Trans Signal Process* 2014; 62: 531–544.

14. Wang CG, Li HK, Huang GJ, et al. Early fault diagnosis for planetary gearbox based on adaptive parameter optimized VMD and singular kurtosis difference spectrum. *IEEE Access* 2019; 7: 31501–31516.

15. Wang J, Li X, Zhou XD, et al. Ultra-short-term wind speed prediction based on VMD-LSTM. *Power Syst Prot Control* 2020; 48: 45–52.

16. Yang DW, Feng FZ, Zhao YD, et al. A VMD sample entropy feature extraction method and its application in planetary gearbox fault diagnosis. *J Vib Shock* 2018; 37: 198–205.

17. Huang WX and Liu DW. Mine microseismic signal denoising based on variational mode decomposition and independent component. *J Vib Shock* 2019; 38: 56–63.

18. Tang JX, Deng CW and Huang GB. Extreme learning machine for multilayer perceptron. *IEEE Trans Neural Netw Learn Syst* 2016; 27: 809–821.

19. Huang GB, Zhou HM, Ding XJ, et al. Extreme learning machine for regression and multiclass classification. *IEEE Trans Syst Man Cybern B Cybern* 2012; 42: 513–529.

20. Liu JW, Li Q, Chen WR, et al. Fast fault diagnosis method of evaporatively cooled fuel cell system based on online sequential extreme learning machine and principal component analysis. *Trans China Electrotech Soc* 2019; 34: 3949–3960.