

Online Dynamic Working-state Recognition Through Uncertain Data Classification

Xiaozhen Yan^{1,2*}, Qinghua Luo^{1,2,3*}, Jianyu Sun³, Zhenhua Luo⁴, Yunsai Chen⁵

¹ School of Information Science and Engineering, Harbin Institute of Technology at Weihai, No. 2 Wenhua West Road, Weihai, Shandong Province, China, zip code: 264209; E-Mail: yxz_heu@163.com, luoqinghua081519@163.com

² The Auto Test and Control Institute, Shandong Institute of Shipping, No. 2 Wenhua West Road, Weihai, Shandong Province, China, zip code: 264209, E-Mail: luoqinghua081519@163.com

³ Shandong New Beiyang Information Technology Co., Ltd., Weihai, Shandong Province, China, zip code: 264203, E-Mail: sunjianyu@newbeiyang.com

⁴ School of Water, Energy and Environment, Cranfield University, Cranfield, MK43 0AL, UK, E-Mail: Z.Luo@cranfield.ac.uk

⁵ China National Deep-Sea Center, Qingdao, Shandong Province, China, zip code: 264209, cys@ndsc.org.cn

* Xiaozhen Yan (corresponding author), Qinghua Luo (corresponding author), School of Information and Electrical Engineering, Harbin Institute of Technology at Weihai, No. 2 Wenhua West Road, Weihai, Shandong Province, China, zip code: 264209; E-Mail: luoqinghua081519@163.com; Tel.: +86-631-5678-234; Fax: +86-631-5678-220.

Abstract: The satellite must continue working properly under different working environments and working loads. The power system is an essential component. Due to different working tasks, loads, and attitudes, a power system has many diverse working states. Therefore, it is necessary to accurately recognize the working state online for fault diagnostics and health management. However, under different working loads, measurement errors, environmental noise, environmental interference, and other uncertain factors, the output voltage value of a satellite power system has different levels of uncertainties. If these uncertainties and various working states are not considered, the recognition results can be of low quality. To address this problem and the uncertainty factors, we present an online dynamic working-state recognition system for satellite power systems based on uncertain data classification. In the system, we first explore the uncertain-data clustering center to model the working state. Then, with a slide-window processing strategy, we compute the distances between the uncertain cluster centers and the uncertain voltage data

for the satellite power system online. Thus, we can obtain more accurate dynamic working-state recognition results. The evaluation results of real data demonstrate that the presented system is valid for working-state recognition and can be applied to a satellite power system.

Keywords: Satellite power; Working-state recognition; Uncertainty data; Classification

List of abbreviations

FCM	Fuzzy C-means
RST	Rough set theory
IPCA	Improved principal component analysis (IPCA)
EKF	Extended Kalman filter

1. Introduction

From application requirements and cost perspective, it is crucial for the satellite to work steadily under various conditions [1-5]. During this process, the satellite power system is a critical part and serves a crucial function in the satellite system [4-6]. Currently, the voltage of the satellite power system is measured to recognize its working state and health conditions [4-7]. However, as there are several factors, e.g., attitude adaption, rotation, changes of working load, the satellite power system experiences many different working states [8]. On the one hand, the satellite must adapt its attitude in space, or due to the rotation, its solar panel may face toward the sun or away from the sun. Thus, the output voltage value of the power system may change over time. On the other hand, the satellite has various working loads. These working loads also generate different working states for the satellite power system.

In addition to the various working conditions, there are many other uncertain factors [9-14], e.g., environmental noise and environmental interference, measurement errors, or minimum attitude changes for the satellite. These negative factors lead to different levels of uncertainties in the output voltage values in the satellite power system. If not considering these working states with uncertainties, we may obtain incorrect recognition results.

Working-state recognition has attracted the attention of many scholars. Specific recognition methods of the working state have been studied in research works, including fuzzy-based technology [7, 12, 13], K-medoid clustering-based recognition [14], filter-based methods [15-16], amplitude-based recognition [17-18], statistical and processing-based technology [19-20], statistical mean-based recognition [21], and statistical standard deviation-based recognition [22-24]. SVM-based methods have been proposed to recognize the working state [21, 25-26]. However, data uncertainty and a variety of working states have not been considered in these methods, leading to poor recognition accuracy. In this paper, we consider the working states with different levels of uncertainties of the satellite power system and utilize an uncertain data classification to recognize the working-state online

to improve the recognition accuracy. This method can also provide an important reference value for other applications.

In this paper, we make the following contributions:

(1) We consider the uncertainty of output voltage values in the satellite power system and resort to interval data to express distribution information of the uncertain voltage value. We also explore the distance calculation between uncertain **cluster-centers** and uncertain voltage values.

(2) To **recognize** the working state online, we propose and adopt an online dynamic working-state recognition method based on uncertain data classification to achieve higher recognition accuracy.

(3) During online recognition, we consider the unknown working states, which did not appear during the modeling stage. We take advantage of a two-level cluster window and utilize a candidate cluster window to save it. If an unknown working state is confirmed, **it will update** to the current cluster center window.

We organize the structure of this paper as follows: Part II reviews the related works on working-state recognition. In part III, we illustrate the working-state recognition method and its implementation. Part IV evaluates the performance of our proposed method and compares it with other related **works**. Finally, we conclude this paper.

2. Related works

Many researchers have focused on the problems with modeling and diagnosis of various application systems in recent years. They have proposed many working-state modeling methods, fault diagnosis, and working-state recognition methods for the satellite power systems. For working-state recognition and fault diagnosis, there are various methods, including the fuzzy C-means (FCM)-based method [7, 12, 13], the amplitude-based method [17-18], the statistical mean-based method [21-22], the statistical standard deviation-based method [23-24], and the support vector machine (SVM)-based method [25-26].

2.1. Working-state Modeling Methods

For a peak power tracker architecture for small satellites, a dynamic model was proposed to verify the solar array and the battery size under different operating conditions [3]. [4] proposed a modeling method and diagnostic technique for the power system of a satellite. A power model was built in a component-based modeling manner. The fault evolution, coupling, unreliable tests, and other problems were considered [5], and a framework of a model-based diagnosis system was presented to minimize uncertainty during the diagnostics of the satellite, which was called LYDIA-NG. The framework utilized active-testing algorithms to improve recognition accuracy. FCM and improved principal component analysis (IPCA) techniques were utilized to model the fault mode of the power transformer [7]. As duplicate or conflicting objects are the main negative factors for accurate models, [13] utilized rough set theory (RST) and FCM to model rotating machinery faults. The faults of both sensors and

actuators were considered [15], and two extended Kalman filters (EKFs) were designed to estimate the attitude information. Control chart pattern recognition (CCPR) is also an important tool; in [20], it was used to perform multivariate processes and recognize abnormal patterns.

2.2. Working-state Recognition Methods

Fuzzy logic is a powerful tool used to process uncertainty. In [7], FCM and IPCA technology were exploited to perform recognition modeling and recognition of a power transformer. The recognition accuracy was improved at the cost of high computational complexity. The uncertainty was considered, and an uncertain data clustering-based method was proposed to improve the recognition accuracy of a satellite power system [9]. To implement model matching and pattern recognition, RST, and FCM technologies were used to conduct fault diagnosis of rotating machinery [13]. To assess a fault diagnosis problem in a satellite flywheel bearing, reference [14] proposed a method based on 3D correlation dimension clustering technology, and the recognition accuracy was very high. Using the attitude estimates provided by the EKFs, [15] designed a fault detection and isolation scheme to diagnose faults in the system, but the evaluation was performed through simulations. A predictive filter and empirical mode decomposition (EMD) were exploited to estimate and diagnose the fault of the satellite [16]. To assess a nonlinear effects problem during diagnostics of a medium structure, [17] proposed a method using finite amplitude. To assess a diagnostic problem with high-temperature fuel cells, [18] presented a new two-stage diagnostic method based on amplitude transient technology. Because the vibration spectrum may deliver information on the type of defects, [19] calculated the statistical information of recorded vibration signatures. Vibration spectra were then utilized to diagnose the fault of a ball bearing. In [20], CCPR was used to conduct statistical calculations to diagnose a fault. The bottom-up and top-down strategies were compared during the evaluation. The diagnostics of incipient faults for analog circuits is very important and challenging. [21] utilized a statistical vector and particle swarm optimization (PSO) based on the Mahalanobis distance (MD) to diagnose incipient faults in an analog circuit. To improve the fault diagnosing accuracy, [22] utilized dual triangle, alternative triangle, and standard deviation information to identify the fault. The statistical means and standard deviation are very useful information, and [23-24] exploited statistical information and uncertain clustering technology to perform model recognition and matching for higher accuracy. The computational complexity was very low. However, they did consider the online processing problem of uncertain data streams. SVM is a powerful tool that is used for pattern recognition. It is utilized to recognize the working state [25-26], but it is weak at recognizing the dynamic working state.

3. The Working-state Recognition Method based on Uncertainty Data Classification

In this section, we first illustrate the framework of our proposed method. We then demonstrate the details of each submodule, including the theoretical basis and implementation details.

3.1. The Framework of the Working-state Recognition Method

The framework of the working-state recognition method is illustrated in Fig. 1. It is composed of three main components: the **uncertainty modeling of output voltage values**, **working-state modeling**, and **online working-state recognition**. We first perform an uncertainty modeling of output voltage values for the satellite power system and express it in terms of interval data. We then utilize an offline or historical uncertain voltage value to perform the working-state modeling of the satellite power system. Finally, we conduct online working-state recognition with the uncertain data stream and determine the working state result, which can provide an important reference value for fault diagnostics and health management.

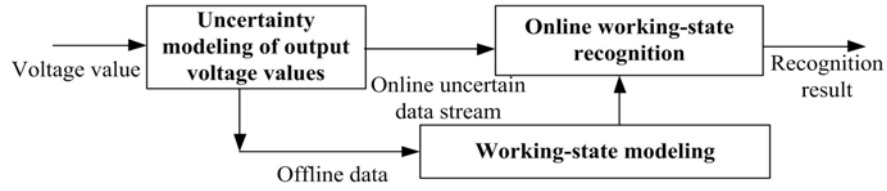


Figure 1. The framework of the working-state recognition method.

Uncertainty modeling of output voltage values: As the output voltage values of the satellite power system have different levels of uncertainties, we model the voltage values in terms of interval data and consider distance calculation between two interval data.

Working-state modeling: Considering the uncertainty of the output voltage value of the satellite power system, we utilize the offline uncertain data and interval data clustering center to model the working state with a sliding window strategy.

Online working-state recognition: With the working-state model, we perform online working-state recognition on the uncertain voltage data stream and determine the working state.

3.2. The Uncertain Modeling of the Output Voltage Values

Interval data are a powerful tool for expressing uncertain data and related calculations. In this paper, we use interval data to represent uncertain output voltage values and their models. **We first give some related definitions of interval data and distance calculations between interval data.** Next, we represent uncertain output voltage values in terms of interval data through statistical computation. **During this procedure, we should consider the outliers.** Therefore, we conduct outlier detection and eliminate them.

3.2.1. Definitions of Interval Data

Interval data: Suppose there are two real data $x_L, x_R \in \mathbf{R}$, and $x_L \leq x_R$; the collection $x = [x_L, x_R]$ can be noted as interval data. The interval data are not traditional certain data; they represent a value range, not a specific value. As shown in Fig. 2, x_L is the lower bound of interval data x . x_R is the upper bound of interval data x .



Figure 2. Interval data x .

Midpoint and radius of interval data: Suppose there are interval data $x=[x_L, x_R]$, $m_x=(x_L+x_R)/2$ is denoted as the midpoint of interval data x , and $r_x=(x_R-x_L)/2$ is treated as the radius of interval data x . The midpoint and radius of interval data x can be calculated as follows.

$$x_L = m_x - r_x \quad (1)$$

$$x_R = m_x + r_x \quad (2)$$

Therefore, data x is also represented as $x=[m_x-r_x, m_x+r_x]$ in terms of midpoint m_x and radius r_x .

Distance computation between two interval data: Suppose there are two interval data $x=[m_x-r_x, m_x+r_x]$ and $y=[m_y-r_y, m_y+r_y]$, $m_x, m_y \in \mathbf{R}$, $r_x, r_y \in \mathbf{R}$; the geometrical relationship between x and y can be illustrated with Fig. 3. We perform the following analysis.

Separated: As illustrated in Fig. 3(a), when x and y are separated, the minimum distance d between the two interval data is $|m_x-m_y|-r_x-r_y$, and the maximum distance d between them is $|m_x-m_y|+r_x+r_y$.

Adjoined or Overlapped: As described in Fig. 3(b) or Fig. 3(c), when x and y are adjoined or overlapped, the minimum distance d between them is 0, and the maximum distance d between them is $|m_x-m_y|+r_x+r_y$.

Contained: As shown in Fig. 3(d), when the interval data x contain the other interval data y , the minimum distance d between them is 0, and the maximum distance d between them is $|m_x-m_y|+r_x+r_y$.

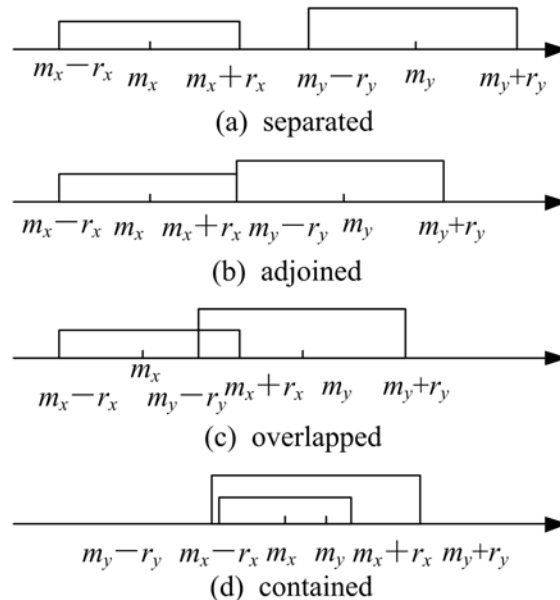


Figure 3. The geometrical relationship between interval data x and y .

From the analysis above, we can see that the distance d is defined by equations (3-5), which is between the two interval data x and y .

$$d = [d_{\min}, d_{\max}] \quad (3)$$

$$D_{\min} = \begin{cases} |m_x - m_y| - \alpha_x - \alpha_y, & |m_x - m_y| - \alpha_x - \alpha_y \geq 0 \\ 0, & |m_x - m_y| - \alpha_x - \alpha_y < 0 \end{cases} \quad (4)$$

$$D_{\max} = |m_x - m_y| + \alpha_x + \alpha_y \quad (5)$$

3.2.2. Statistical Computation and Uncertain Expression in Terms of Interval Data

The power system of the satellite works all the time. Therefore, the output voltage value of the satellite power system is generated from time to time; i.e., it is generated in the form of a data stream. We should pursue the sliding window strategy when performing statistical computation and other processing methods on the system. More specifically, we perform a statistical calculation on the output voltage values of the satellite power system within a sliding window with a specific time range, which slides over time. We obtain the statistical mean m_y and the statistical standard deviation r_y , and the uncertain voltage value can be expressed in terms of interval data $y_t = [m_y - r_y, m_y + r_y]$ with the statistical characteristic. Here, t denotes the time stamp. The statistical mean m_y and statistical standard deviation r_y can be calculated as follows.

$$m_y = \sum_{i=t-w}^t v_i / w \quad (6)$$

$$r_y = \sqrt{\sum_{i=t-w}^t (v_i - m_y)^2 / w} \quad (7)$$

where w denotes the width of the sliding window, which determines the performance of statistical computation. When w is too large, statistical information cannot track the dynamic change in the output voltage values. Alternatively, when it is too small, the statistical information cannot be accurately computed. The optimized value of the sliding window width should be analyzed and determined.

3.3. Working-state Modeling

The output voltage values of the same working state can form a specific cluster. They share the same statistical characteristics in the same cluster. We resort to the cluster center to represent the statistical characteristics of each cluster. Therefore, we first model the working state with uncertain offline sample data through statistical computation technology and represent the statistical features of the working state with interval data.

The output voltage values have different levels of uncertainties. The interval data are utilized to model the working state. In the modeling method, statistical computation is applied to the sample data

with the same working state. According to the error theory, if voltage values obey the normal distribution, suppose μ is the statistical mean of voltage values and σ is the statistical standard deviation of voltage values. The probability of the value in the interval $[\mu - \sigma, \mu + \sigma]$ is 68.3%. The probabilities of the voltage value in the intervals $[\mu - 2\sigma, \mu + 2\sigma]$ and $[\mu - 3\sigma, \mu + 3\sigma]$ are 95.4% and 99.7%, respectively. Therefore, the statistical interval can be generally expressed as $[\mu - k\sigma, \mu + k\sigma]$, where k is the coverage factor, and $\{k \in \mathbf{R} | 0 \leq k \leq 3\}$.

Supposing the satellite power systems have N kinds of working states, $N \in \mathbf{N}$. For the i th working state, the statistical mean and statistical standard deviation of the voltage values are μ_i and σ_i , respectively. The voltage values of the same working state can compose a cluster, and the corresponding cluster center c_i can be expressed as $c_i = [\mu_i - k\sigma_i, \mu_i + k\sigma_i]$. The valid range of clusters is considered, which indicates that if the value is located out of $3\sigma_i$ of the cluster center c_i , it is treated as an outlier. In the same manner, we perform the statistical calculation and clustering center representation of the other working states. The overall cluster centers can be expressed as $\mathbf{C} = \{[\mu_1 - k^*\sigma_1, \mu_1 + k^*\sigma_1], [\mu_2 - k^*\sigma_2, \mu_2 + k^*\sigma_2], \dots, [\mu_i - k^*\sigma_i, \mu_i + k^*\sigma_i], \dots, [\mu_N - k^*\sigma_N, \mu_N + k^*\sigma_N]\}$; then, we can obtain the radius $R = \{r_1, r_2, \dots, r_i, \dots, r_N\}$ and the corresponding working state $M = \{m_1, m_2, \dots, m_i, \dots, m_N\}$.

3.4. Online Working-state Recognition

This subsection is composed of four parts. We first describe the uncertain classification-based working-state recognition principle. **We then calculate the square of the distance between interval data for embedding in the online recognition algorithm.** Next, we consider the unknown working-state processing; i.e., we adopt the learning and updating strategy to enhance the recognition method. Finally, we present the online working-state recognition algorithm.

3.4.1. Working-state Recognition Principle

Based on the working-state modeling, considering the variety of working states and the uncertainty of the output voltage values in the satellite power system, an online uncertain data classification strategy is utilized to perform working-state recognition, i.e., the distance $\{d_i\}$ between interval data y and the clustering centers $\{c_i\}$; here, the data y are represented in terms of statistical information calculated through statistical computation and are expressed in terms of interval data. Next, the minimum distance d_j within the specific radius $\{r_i\}$ is confirmed, and **we treat the corresponding working-state as the working-state recognition result y .**

3.4.2. Calculation of Square Distance

To embed the distance d_i in the online recognition algorithm, a correlation factor λ is introduced to combine the minimum distance square value and the maximum distance square value, as shown in (8), and $\{\lambda \in \mathbf{R} | 0 \leq \lambda \leq 1\}$. The distance square $d^2(c_i, y)$ can cover every possible value between $d_{\min}^2(c_i, y)$ and $d_{\max}^2(c_i, y)$.

$$d^2(c_i, y) = \lambda d_{\min}^2(c_i, y) + (1 - \lambda) d_{\max}^2(c_i, y) \quad (8)$$

3.4.3. Outlier Detection and Unknown Working-state Processing

The output voltage values may contain outliers. Therefore, during statistical computation, we consider outlier detection and elimination. According to the 3σ criterion, if the residual between the voltage value and its statistical mean is higher than three times its statistical standard deviation, the current value may be an outlier. However, as the working states vary, the output voltage value may change rapidly.

During the working-state modeling stage, the sample data may not be comprehensive; i.e., the corresponding voltage values of some working state may not be collected during working-state modeling. Therefore, at the recognition stage, some unknown working states may appear that did not **emerge** during the working-state modeling stage. A two-level classification window-based online recognition algorithm is proposed. The structure of the algorithm is composed of a current window and a candidate window, as shown in Fig. 4. As illustrated in Fig. 4, the current window stores the cluster centers gained during the modeling stage. The candidate window is used to store the outlier data, which are located outside the valid range of every cluster center in the current window. After being determined to be an outlier in the current window, the outlier is determined to be located in the valid range of temporary cluster centers in the candidate window. If it belongs to any cluster center, the information of the cluster is updated. Otherwise, the outlier is stored as a candidate cluster center in the candidate window to further judge whether it belongs to one kind of unknown working state or not. If the number of members of the cluster center in the candidate window is greater than four, **the cluster center will update** to the current window; i.e., the cluster is one kind of unknown working state of the satellite power system. In this way, the unknown working state can be learned, determined, and updated into the current window. Otherwise, the outlier is eliminated from the voltage values.

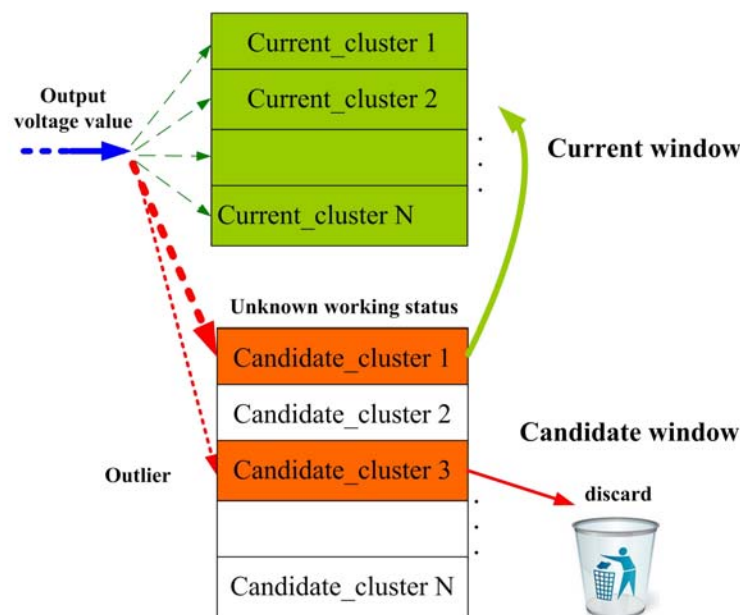


Figure 4. Diagram of unknown working-state processing.

3.4.4. Dynamic Working-state Recognition Algorithm

At the recognition stage of working states, for a measured output voltage stream of the satellite system, a statistical computation is conducted online with a sliding window strategy. The statistical features are expressed in terms of interval data y_t . We describe the pseudocode of the proposed dynamic working-state recognition through interval data classification (DWSRIDC).

Algorithm: DWSRIDC ()

1. **Input:** $y_t, k, \lambda, \{m_i\}, \{c_i\}, \{r_i\}, (0 \leq i \leq N)$
2. **Output:** m_y // working-state recognition output
3. $\{current_c_i\} = \mathbf{C}$ // initial current window
4. $\{candidate_c_i\} = \{\emptyset\}$ // initial candidate window
5. do while (y_t)
6. for $i=1$ to N
7. compute $d^2(current_c_i, y_t)$
8. end
9. find $\min(d^2(current_c_i, y_t)) = d^2(current_c_j, y_t)$
10. if ($d^2(current_c_j, y_t) < r_j^2$)
11. $m_y = m_j$
12. else
13. if ($candidate_c_i = \emptyset$)
14. $candidate_c_i = y$
15. else
16. if ($y_t \in \{candidate_c_i\}$)
17. if $\text{size}(candidate_c_i) > 4$
18. update $candidate_c_i$ to the current window
19. $m_y = u_j$
20. $candidate_c_i = \emptyset$
21. else
22. $\text{size}(candidate_c_i)++$
23. $m_y = \emptyset$
24. end

```

25.         else
26.              $m_y = \emptyset$ 
27.             if ( $\{ candidate\_c_i \}$  is full)
28.                  $candidate\_c_{old} = y_t$ 
29.             else
30.                  $candidate\_c_{i+1} = y_t$ 
31.             end
32.         end
33.     end

```

From the pseudocode, the computational time complexity of DWSDIDC is $O(Nt)$. As the value of N is very small, the complexity is approximately $O(N)$, and the complexity is lightweight.

4. Experimental Evaluation

We perform experiments to validate the presented working-state recognition method. We also evaluate the performance of the method and compare it with that of other related recognition methods. In this section, **we illustrate the experimental setup**. We then evaluate the feasibility of the method. Finally, we assess the performance of the proposed method.

4.1. Experimental Setup

The experimental data are from our research partner. They are the real output voltage values of a satellite power system. As illustrated in Fig. 5, there are many working states. Some negative factors can lead to different levels of uncertainties. The working state also changes over time. We also cooperate with experts from our research partner to confirm the validation and feasibility of our method. We show the actual working state in Fig. 6, which will be adopted as a reference working-state value.

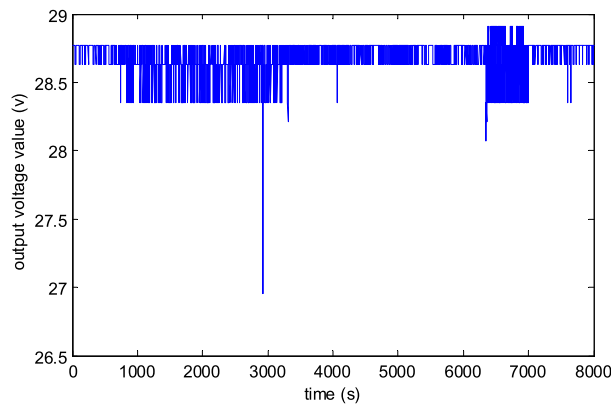


Figure 5. The satellite power system output voltage value.

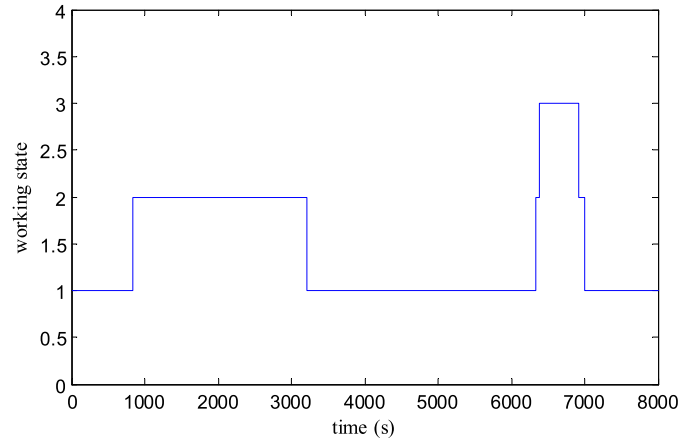


Figure 6. The working state of the satellite power system.

The performance of the working-state recognition method is evaluated in terms of accuracy and efficiency. The accuracy is computed in terms of the similarity between the recognition results and the reference state value; A higher accuracy indicates high accuracy. The efficiency is evaluated in terms of the processing time of the working-state recognition method. The shorter the processing time, the higher the efficiency is.

The configuration of the working-state recognition evaluation platform is described as follows. CPU, Intel G2020@2.9 GHz; RAM, 4 Gb; operating system, Windows 7 32-bit; evaluation environment, MATLAB R2009b.

The reference working-state recognition methods include the FCM-based recognition method [7, 12], the amplitude-based recognition method [17, 18], the SVM-based recognition method [26-27], the statistical mean-based recognition method [21], and the statistical standard deviation-based recognition method [22-24]. In [7, 12], the uncertainty was considered, and FCM clustering was adopted to perform working-state recognition. The amplitude was used to **recognize** the working state [17, 18]. The SVM-based strategy was utilized to recognize the working state in [25-26]. The statistical mean and the statistical standard deviation are calculated from the uncertain output voltage value. Therefore, this approach can be utilized to improve recognition accuracy [22-24]. To be fair, we do not implement every aspect of these methods; we adopt the concept to perform working-state recognition.

4.2. Feasibility Evaluation

In our proposed working-state recognition method, there are three key parameters; i.e., the coverage factor k , correlation factor λ , and width w of the sliding window. We first conduct a correlation analysis of the three factors. We then evaluate their effects on the recognition accuracy of the working-state recognition method.

4.2.1. Correlation Analysis of Factors

1) Correlation Analysis Between Correlation Factor k and Coverage Factor λ

According to the parameter setting, we adopt different values of correlation factor k and coverage factor λ and evaluate the working-state recognition accuracy. We illustrate the result in Fig. 7. From Fig. 7, we can see that there is no correlation between the two parameters. More specifically, the recognition accuracy almost changes with increasing coverage factor. Therefore, we can determine the optimized value of the two parameters through impact analysis on the recognition accuracy.

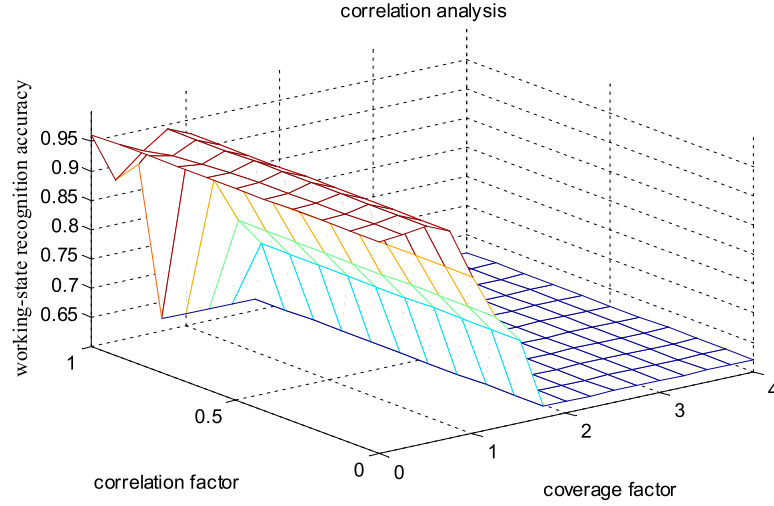


Figure 7. The correlation analysis between correlation factor k and coverage factor λ .

2) Correlation Analysis Between Window Width w and Coverage Factor λ

We adopt different values of window width w and coverage factor λ and evaluate the working-state recognition accuracy. The result is shown in Fig. 8. We can see that there is no correlation tendency between the two parameters. Therefore, we can determine the optimized value of the two parameters through impact analysis on the recognition accuracy.

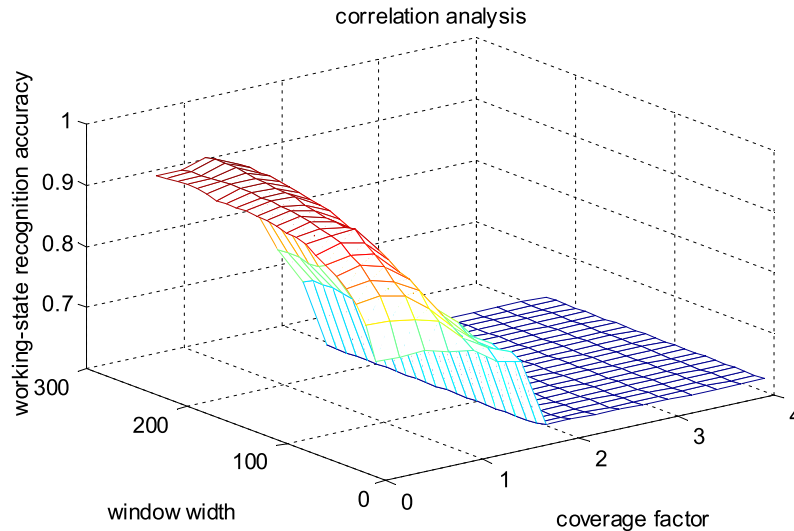


Figure 8. The correlation analysis between window width w and coverage factor λ .

3) Correlation Analysis Between Window Width w and Correlation Factor k

With different values of window width w and correlation factor k , we perform the working-state recognition and show the recognition accuracy in Fig. 9. We can see that there is no correlation tendency between the two parameters. Therefore, we can determine the optimized value of the two parameters through an impact analysis on the recognition accuracy.

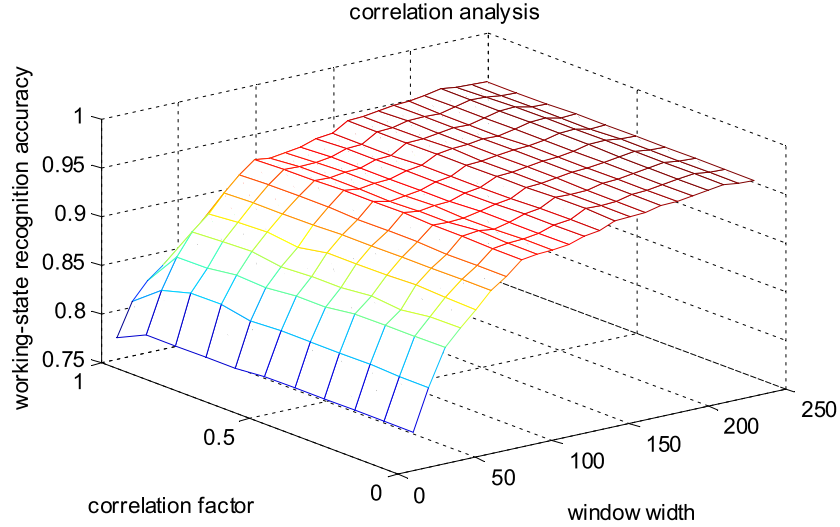


Figure 9. The correlation analysis between window width w and coverage factor λ .

4.2.1. Effect Evaluation of Coverage Factor k

The coverage factor k determines the valid range of uncertainty data. We adopt different values of the coverage factor k and evaluate the recognition accuracy of our proposed method. The width w of the time sliding window is set to 200, and the correlation factor λ is set to 0.5. We illustrate the recognition accuracy in Fig. 10.

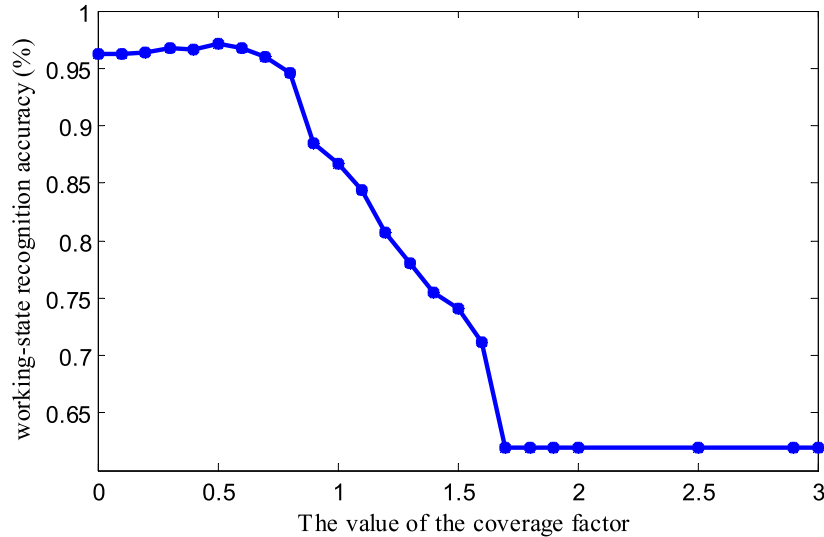


Figure 10. The effect of the coverage factor on the recognition.

Fig. 10 illustrates that the coverage factor k has an obvious impact on recognition accuracy. When its value is between 0 and 1, the recognition accuracy is very high—almost 90%. When it takes a value of 0.5, the recognition accuracy is 97.18%. However, with increasing coverage factor, the recognition accuracy decreases. Therefore, we select 0.5 as the optimized value of coverage factor k .

4.2.2. Effect Evaluation of the Correlation Factor λ

The correlation factor λ determines the distance calculation between two interval data. We take different values of the correlation factor λ and conduct working-state recognition to evaluate its impact on recognition accuracy. The width w of the time sliding window is set to 200, and the coverage factor k is set to 0.5. We illustrate the recognition accuracy in Fig. 11.

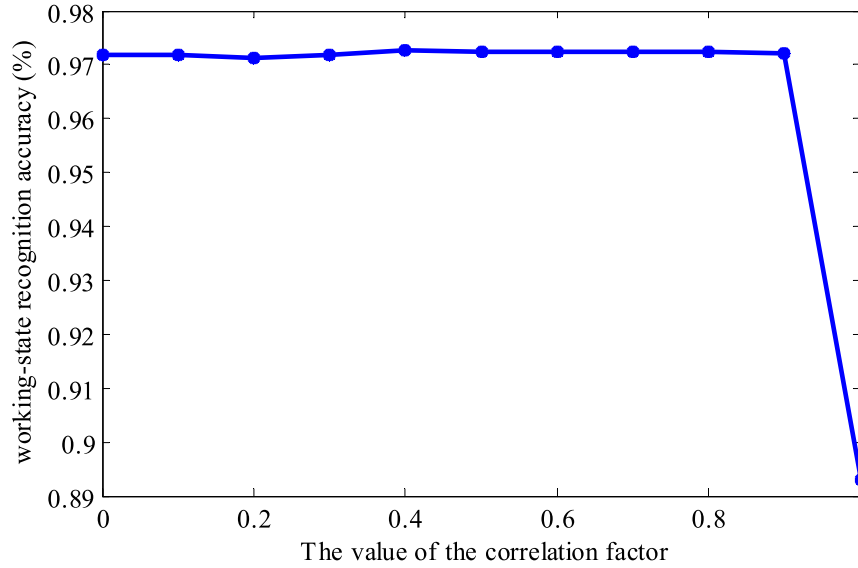


Figure 11. The effect of the correlation factor on the recognition accuracy.

Fig. 11 indicates that the correlation factor λ does not have an obvious impact on the recognition accuracy, except when it takes the value of 1. When its value is between 0 and 0.9, the recognition accuracy is approximately 97.20%. When it takes a value of 0.4, the highest recognition accuracy is 97.28%. Therefore, we take 0.4 as the optimized value of the correlation factor.

4.2.3. Effect Evaluation of Sliding Window Width w

The sliding window width w determines the number of voltage values when we perform statistical computation. As discussed before, when its value is too large, the statistical information cannot track the dynamic change in the output voltage value. When its value is too small, the statistical information cannot be accurate. Therefore, we adopt different values of sliding window width w to conduct working-state recognition and evaluate its impact on the recognition accuracy. We changed its value from 10 to 600. The coverage factor k is set to 0.5. The correlation factor λ takes a value of 0.4. We illustrate the recognition accuracy in Fig. 12.

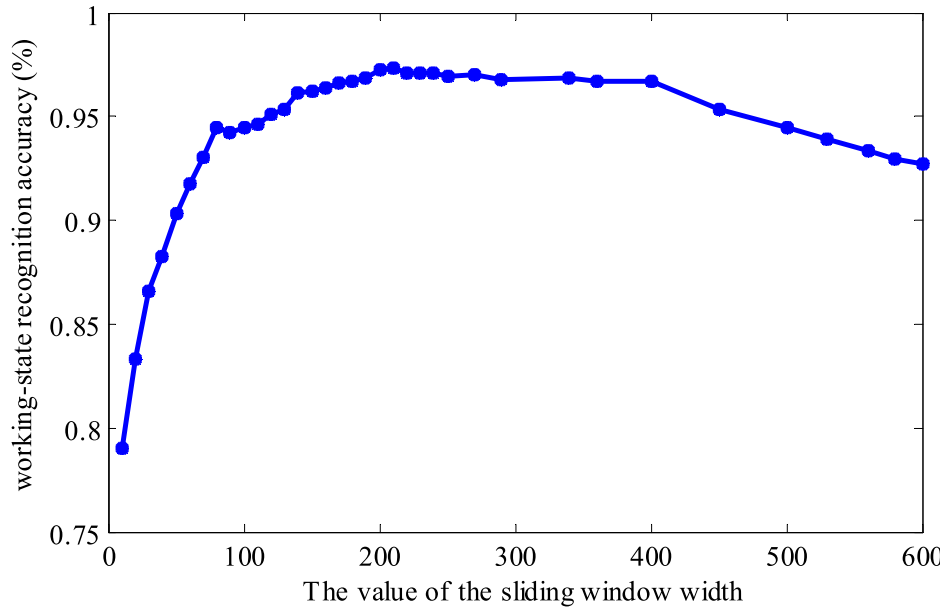


Figure 12. The impact of the sliding window width on the recognition accuracy.

Fig. 12 shows that the sliding window width w has a significant impact on recognition accuracy. More specifically, with increasing sliding window width w from 0 to 210, the recognition accuracy increases from 79% to 97.34%. When the sliding window width w takes the value of 210, the recognition accuracy is the highest, up to 97.34%. With increasing sliding window width w from 220 to 600, the recognition accuracy decreases from 97.10% to 92.74%. Therefore, we select 210 as the optimized value of the sliding window width w .

4.3. Recognition Performance Evaluation

We evaluate the performance of the working-state recognition method in terms of recognition accuracy and recognition efficiency.

4.3.1. Recognition Accuracy Evaluation of Dynamic Recognition Methods

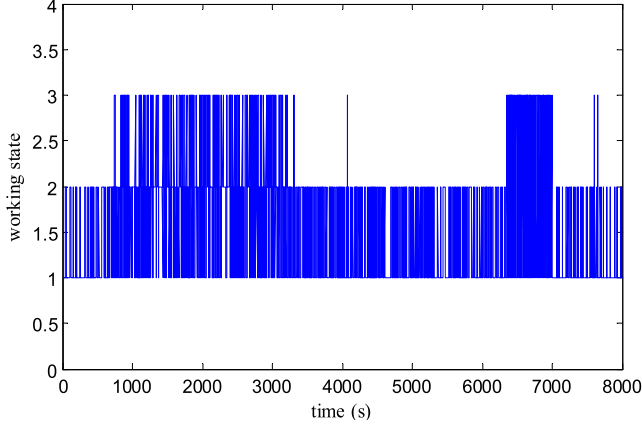
After evaluating the feasibility and optimizing the value determination of parameters, we determine the working-state recognition accuracy of our presented method. The parameter settings are listed in Table I.

TABLE I The Parameter Setting.

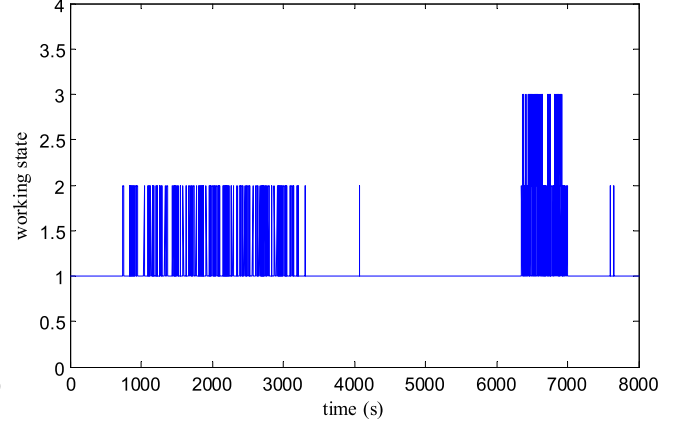
Parameter	Coverage factor	Correlation factor	Sliding window width
Value	0.5	0.4	210

We also evaluate the recognition accuracy of the reference methods, including the FCM-based method, the amplitude-based method, the SVM-based method, the statistical mean-based method, and the statistical standard deviation-based method. We compare the recognition accuracy of these methods

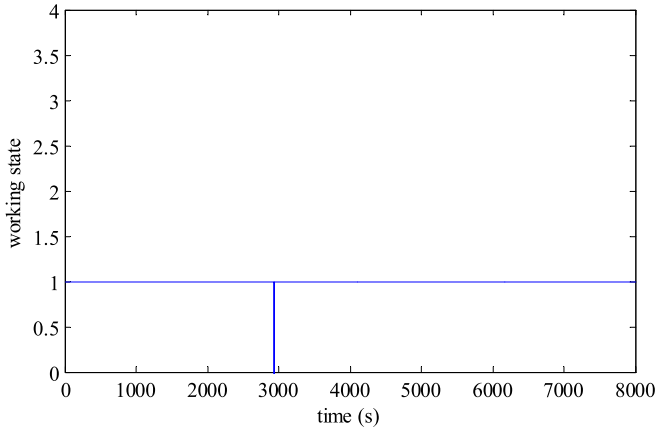
and illustrate them in Fig. 13. We also describe the recognition accuracy of these methods in Fig. 14 and Table II.



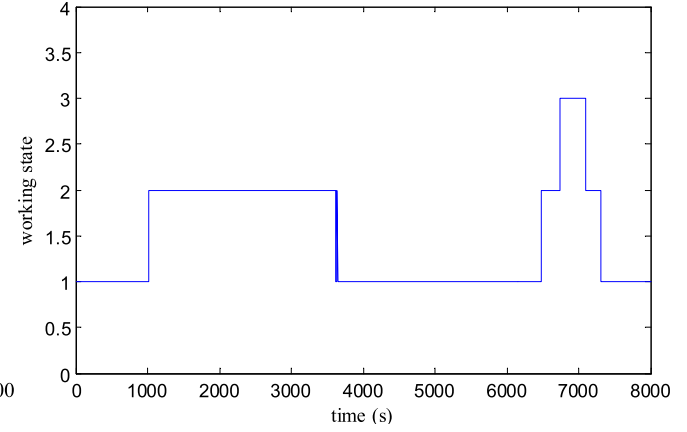
(a) Recognition result of the FCM-based method.



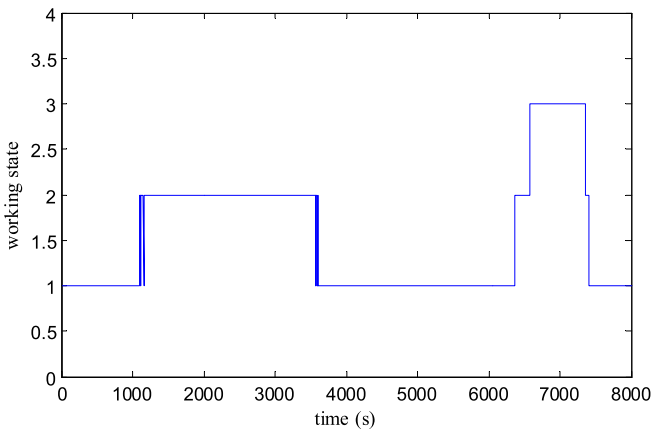
(b) Recognition result of the amplitude-based method.



(c) Recognition result of the SVM-based method.



(d) Recognition result of the statistical mean-based method.



(e) Recognition result of the standard deviation-based method. (f) Recognition result of the uncertain data classification-based method.

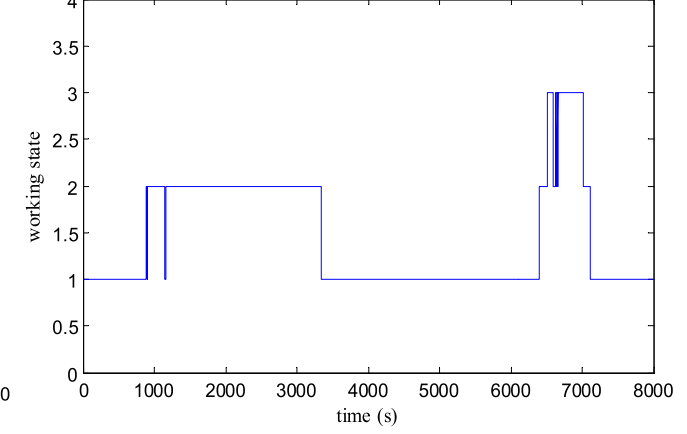


Figure 13. The working-state recognition results of methods.

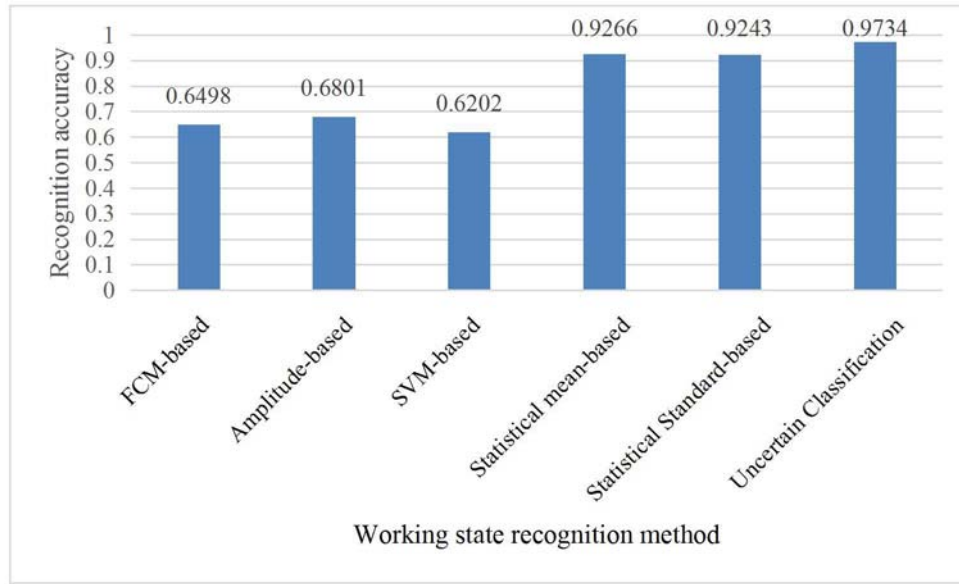


Figure 14. The working-state recognition accuracy of the methods.

TABLE II THE RECOGNITION ACCURACY OF THE METHODS.

Method	FCM-based	Amplitude-based	SVM-based	Statistical Mean-based	Statistical Standard-based	Uncertain Classification
Accuracy	0.6498	0.6801	0.6202	0.9266	0.9243	0.9734
Improvement	49.80%	43.13%	56.95	5.05%	5.31%	-

Fig. 13, Fig. 14, and Table II illustrate that the recognition accuracy of the presented method is much higher than that of the reference recognition methods. Compared with the FCM-based, amplitude-based, SVM-based, statistical mean-based, and statistical standard deviation-based methods, the presented method enhances the recognition accuracy by 49.80%, 43.13%, 56.95%, 5.05%, and 5.31%, respectively. Relative to the FCM-based, amplitude-based, and SVM-based methods, the improvement is obvious. **That** is mainly because the presented method considers the variety and uncertainty of the working state and explores the sliding window-based interval data classification strategy. Relative to the statistical mean-based and the statistical standard deviation-based methods, the proposed method considers both the statistical mean and standard deviation information and explores the interval data classification strategy to improve the recognition accuracy.

4.3.2. Recognition Efficiency Evaluation of the Dynamic Recognition Methods

The recognition efficiency of the working state is evaluated in terms of recognition processing time. For each method, we run the working-state recognition 10,000 times and calculate the average processing time. The recognition processing time of each method is given in Table III.

TABLE III THE RECOGNITION EFFICIENCY OF THE METHOD.

Method	FCM-based	Amplitude-based	SVM-based	Statistical Mean-based	Statistical Standard-based	Uncertain Classification
processing time (s)	0.0581	1.5444e-4	6.13	0.1194	0.3022	0.3882

Table III describes the average recognition processing time of these methods. The results **indicate** that the average recognition processing time of the uncertain classification-based method is longer than that of other methods. This is mainly because the presented method contains distance computations and interval data classification operations. The other methods contain lower computational amounts than the presented method.

The results also indicate that the recognition processing time of the statistical-based methods, including the statistical mean-based, the statistical standard deviation-based, and the presented method are of the same level.

However, considering comprehensively the recognition accuracy and efficiency of these methods, the presented method is more feasible for application to the working-state recognition of satellite power systems.

4.4. Discussion of the Proposed Method Application to Multidimensional Data

To improve the recognition accuracy of the working state of the satellite power system, we propose a one-dimensional recognition method. In real applications, the data may be multidimensional. In our future work, we will explore multidimensional data processing for satellite systems. During this procedure, multidimensional interval data processing will be one of the key points, especially the distance metric calculation. We will also research the generality of our proposed method in other multidimensional applications.

4.5. Discussion of the Recognition Accuracy

The methods evaluated in this paper are classified into two main groups. One is the statistical-based recognition methods: the statistical mean-based, statistical standard deviation-based, and presented method are among statistical-based recognition methods. The other method is ordinary-based methods, including FCM-based and amplitude-based methods. From the evaluation result, we can see that the statistical-based recognition methods can gain higher accuracy than ordinary-based methods. The reason is that the statistical calculation can obtain the key features of the working state.

5. Conclusion

For the working-state recognition problem of the satellite power system in a complex environment, we consider the uncertainties and variety of the output voltage value and present a recognition method. In this method, we first model the uncertainty of the output voltage and express it in terms of interval

data and statistical information. We then explore the interval data classification and sliding window process strategy to perform working-state recognition. The experimental results of actual data show that the presented recognition method has better performance than the other related works and can be applied to satellite power systems. Multidimensional interval data processing-based recognition with high accuracy will be a focus of future research directions to improve the recognition accuracy required from satellite power systems.

Acknowledgments

The research presented in this paper is supported by the National Natural Science Foundation of China (61671174, 61601142, and 51909039), the China Postdoctoral Science Foundation, the Shandong Postdoctoral Innovation Project, the Weihai Research Program of Science and Technology, the Key Lab of Weihai, the Engineering Center of Shandong Province, the Guangxi Key Laboratory of Automatic Detecting Technology and Instruments (YQ18206, YQ15203), the Key Research and Development Program of Shandong Province (2019JMRH03**), the Major Science and Technology Innovation Project of Shandong Province (2019JZZY020702), the Natural Science Foundation of Guangdong Province (Grant No 2020A1515010370), the Characteristic Innovation Projects of Universities in Guangdong Province (Grant No 2019KTSCX094), and the Technology Innovation Strategy Fund Project of Guangdong Province (Grant No. 2018A03017).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References and Notes

- [1]. Y. Li, J. Chen, X. Huang, R. Hong, "Life state recognition of slewing bearing with heavy load and low speed based on point density function with fuzzy C means". Transactions of the Institute of Measurement and Control, vol 41, no.5, pp.1256-1265, March 1, 2019.
- [2]. F. Jawad, M. Ehsan, "Demagnetization Modeling and Fault Diagnosing Techniques in Permanent Magnet Machines under Stationary and Nonstationary Conditions: An Overview ". IEEE Transaction on Industry Applications, vol 53, no. 3, pp. 2772-2785, May, 2017.
- [3]. X. Bi, J. Lin, F. Bi, et al, "Engine Working State Recognition Based on Optimized Variational Mode Decomposition and Expectation Maximization Algorithm". IEEE Access, 2020, 8, pp. 33545-33559.
- [4]. M. Elisa, I. Luigi, S. Salvatore, et al, "Modeling and power management of a LEO small satellite electrical power system", In proceedings of 2018 European Control Conference, ECC 2018, pp. 2738-2743, November 27, 2018.
- [5]. S. Zhang, L. Wang, Y. Liu, et al., "Real Time Fault Diagnosis with Tests of Uncertain Quality for Multimode Systems and its Application in a Satellite Power System", *Journal of Electronic Testing: Theory and Applications (JETTA)*, vol. 34, no. 5, pp. 529-545, 2018.
- [6]. F. Feldman, A., D. Castro, H. Vicente, et al. "Model-Based Diagnostic decision-support system for satellites", In proceedings of *IEEE Aerospace Conference*, March 9, 2013.
- [7]. K. Tusongjiang, W. Gao. "Power transformer fault diagnosis using FCM and improved PCA", *Journal of Engineering*, vol 2017, no. 14, pp. 2605-2608, 2017.

- [8]. F. Wang , T. Feng , X. Chen. "Design and optimisation of a wireless power transfer system for satellite application". *JET Power Electronics*, v 12, n 10, pp. 2586-2598.
- [9]. X. Yan, Q. Luo, Y. Yang, Z. Yang and J. Wang. "a working pattern recognition method for satellite power system based on uncertain data clustering strategy", *The 10th IEEE Prognostics and System Health Management Conference (PHM 2019)* October 26-27, 2019.
- [10]. Q. Luo, Y. Peng, J. Li, X. Peng, "MWPCA-ICURD: density-based clustering method discovering specific shape original features", *Neural Computing and Applications*, vol. 28, no. 9, pp. 2545-2556, 2017.
- [11]. X. Yan, Q. Luo, Y. Yang, S. Liu, H. Li, and C. Hu, "ITL-MEPOSA: An Improved Trilateration Localization with Minimum Uncertainty Propagation and Optimized Selection of Anchor nodes for wireless sensor networks", *IEEE Access*, 2019, vol 7, no. 1, pp. 53136-53146.
- [12]. M. Suo, M. Zhang, D. Zhou, et al, " Fault diagnosis of satellite power system using variable precision fuzzy neighborhood rough set ", In proceedings of Chinese Control Conference (CCC), 26-28, July, 2019.
- [13]. R. Li, Chen, J. Chen, X. Wu, "Fault diagnosis of rotating machinery based on SVD, FCM and RST", *Journal of Vibration and Shock*, vol 24, no 4, pp. 46-49, 2005.
- [14]. C. Chen, T. He, D. Wu, Q. Pan, et al, " A Fault Diagnosis Method for Satellite Flywheel Bearings Based on 3D Correlation Dimension Clustering Technology ", *IEEE Access*, vol. 6, pp. 78483 - 78492, December 2018.
- [15]. S. Gao, Z. Zhang, W. Zhang, et al, " Fault Diagnosis for Satellite Attitude Control System with Using Extended Kalman Filter ", In proceedings of Chinese Control Conference (CCC), 27-30, July 2019.
- [16]. Y. Shen, Y. Zhang, Z. Wang, " Satellite fault diagnosis method based on predictive filter and empirical mode decomposition", *Journal of Systems Engineering and Electronics*, vol. 22, no. 1, pp. 83-87, February 2011.
- [17]. V.O. Vakhnenko, V.A. Danylenko, A.V. Michtchenko, "Diagnostics of the medium structure by long wave of finite amplitude", *International Journal of Non-Linear Mechanics*, vol 35, no 6, pp. 1105-1113, 2000.
- [18]. B. De, B. Chris, S. Paul, et al., "Online diagnostics of HTPFM fuel cells using small amplitude transient analysis for CO poisoning", *IEEE Transactions on Industrial Electronics*, vol 62, no 8, pp. 5175-5186, 2015.
- [19]. T. Karacay, N. Akturk, "Experimental diagnostics of ball bearings using statistical and spectral methods", *Tribology International*, vol 42, no 6, pp. 836-843, 2009.
- [20]. P. Jahani, S. M. Alexander, "Control charts pattern recognition and fault diagnostics in Multivariate Processes", *IIE Annual Conference and Expo 2015*, pp. 1601-1610, 2015.
- [21]. B. Long, M. Wei, M. Li, H. Wang, "Improved diagnostics for the incipient faults in analog circuits using LSSVM based on PSO algorithm with Mahalanobis distance", *Neuro Computing*, vol 133, pp. 237-248, 2014.
- [22]. G.K. Irungu, A.O. Akumu, J.L. Munda, "Fault diagnostics in oil filled electrical equipment: Review of duval triangle and possibility of alternatives", In proceedings of *34th Electrical Insulation Conference, EIC 2016*, p 174-177, August 22, 2016.
- [23]. Q. Luo, Y. Peng, X. Peng, et al., "Uncertain data clustering-based distance estimation in Wireless Sensor Networks", *Sensor*, vol. 14, no. 4, pp. 6584-6605, 2014.
- [24]. Q. Luo, X. Yan, et al., " DDEUDSC: Dynamic distance estimation using uncertain data stream clustering in mobile wireless sensor networks", *Measurement*, vol 55, pp. 423-433, 2014.
- [25]. J. Huang, D. Wang, X. Wang, et al, "Intelligent recognition method for working-cycle state of hydraulic excavator", *Journal of Zhejiang University (Engineering Science)*, vol. 53, no. 9, pp. 1663-1673, September 1, 2019.
- [26]. L. Mao, Y. Xue, H. Lin, A. Yang, "A recognition method of working state of excavator based on video image", *System Engineering Theory and Practice*, vol. 39, no. 3, pp. 797-804, March 1, 2019.



DSpace

<https://dspace.org/>

Online dynamic working-state recognition through uncertain data classification

Yan, Xiaozhen; Luo, Qinghua; Sun, Jianyu; Luo, Zhenhua; Chen, Yunsai

2020-11-28

Elsevier

<https://doi.org/10.1016/j.ins.2020.11.022>

Yan X, Luo Q, Sun J, et al., (2021) Online dynamic working-state recognition through uncertain data classification. Information Sciences, Volume 555, May 2021, pp. 1-16