Heliyon 10 (2024) e23420

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

New energy electric vehicle battery health state prediction based on vibration signal characterization and clustering

Liping Lu, Huiying Zhai, Yun Gao *

Henan Polytechnic, Zhengzhou 450046, China

ARTICLE INFO

Keywords:
Vibration signal
K-mean clustering algorithm
EEMD
Prediction
Lithium battery

ABSTRACT

The health status of the battery of new energy electric vehicles is related to the quality of vehicle use, so it is of high practical application value to predict the health status of the battery of electric vehicles. In order to predict the health status of lithium battery, this study proposes to optimize the empirical modal decomposition method and obtain the ensemble empirical modal decomposition algorithm, and use this algorithm to collect the vibration signal of the battery, then use wavelet transform to pre-process the collected signal, and finally combine K-mean clustering and particle swarm algorithm to cluster the signal types to complete the prediction of battery State of Health. The experimental results show that the ensemble empirical modal decomposition algorithm proposed in this study can effectively perform signal acquisition for different state types of batteries, and the K-mean clustering-particle swarm algorithm predicts a 63 % decrease in the health state of the battery at 600 cycles, with a prediction error of 2.6 %. Therefore, the algorithm proposed in this study is feasible in predicting the battery health state.

1. Introduction

With the intensification of global energy crisis and environmental problems, electric vehicles (EVs) as a kind of green transportation have gained wide attention and application. As one of the key components of electric vehicles, the State of Health (SOH) of the battery directly affects the performance and reliability of the electric vehicle [1]. In the battery SOH prediction, vibration signal is a common monitoring signal. The vibration signal can reflect the mechanical and electrical faults inside the battery and can provide effective data support for predicting the battery health status [2]. However, there is relatively little research on techniques such as vibration signal feature extraction and clustering analysis, and further exploration and research are needed. The current common method for health prediction of Li-ion batteries is the discharge test, which determines the battery's endurance time by connecting a load and discharging it continuously. However, the accuracy of this method is low and easily interfered by external environmental factors. In addition to this, there are many other scholars who still use different methods to predict the healthy life of the battery, Liu et al. constructed an improved lithium-ion battery decay model using a data-driven framework with particle filters. The battery prediction rate of this model was less than 8 % [3]. Liu et al. constructed a battery cycle prediction model using a combination of EMD and long and short-term memory algorithms, which can achieve more than 95 % prediction accuracy. The performance of the prediction model constructed in this study has better performance [4]. Therefore, this study innovatively proposes to combine vibration signal characteristics and clustering algorithms to predict battery SOH. Based on the Empirical Mode Decomposition Method (EMD), the research

* Corresponding author.
E-mail address: gyauto@126.com (Y. Gao).

https://doi.org/10.1016/j.heliyon.2023.e23420
Received 18 July 2023; Received in revised form 23 October 2023; Accepted 4 December 2023
Available online 8 December 2023

first optimizes it to obtain the Ensemble Empirical Mode Decomposition (EEMD) algorithm. The optimized EEMD algorithm has better adaptability, can average different Gaussian noise sequences, suppress modal aliasing, and improve the reliability and operational accuracy of the decomposition results. In addition, the algorithm also has a relatively simple operational process [5]. And then uses the improved EEMD algorithm to collect the battery Then, we use the improved EEMD algorithm to collect the vibration signals of the battery, and then optimize the K-mean clustering algorithm by using the Particle Swarm Optimization (PSO) algorithm, and use the optimized K-mean clustering-PSO algorithm to classify the collected battery vibration signals to complete the battery SOH prediction. In order to complete the above-mentioned research, the article will be divided into five parts for discussion. The first part is an overview of the research direction and content of the full paper. The second part is a study of the battery SOH prediction technique combining vibration signal features and clustering algorithm, which is divided into three subsections. The first subsection is a discussion of the SOH features of lithium batteries, the second subsection is an optimization study of the vibration acquisition method, and the third subsection is a study of the optimization of the clustering algorithm and battery prediction. The fourth section is the performance analysis of the algorithm proposed in the study. The fifth section is a summary of the research content and shortcomings of the whole text.

2. Related works

EMD is a signal feature extraction method, which obtains several intrinsic mode functions (IMFs) by decomposing the signal, and then obtains the feature vector of the signal by feature extraction of each IMF. EMD has been widely used in various fields because of its efficient signal feature processing capability. In the field of energy regeneration, Ghimire et al. combined EMD and bidirectional long- and short-term memory networks to construct a prediction model for renewable energy, aiming to improve the control efficiency of the grid system. Experimental results show that the model predicts MAE values up to 9.769 W/m in 5 min^2, and MAPE values up to 5.657 % [6]. Zhang et al. used EMD improved filters for wind energy generation renewable energy prediction and corrected the prediction results based on real-time weather information by depth gradient method. The proposed method provides a feasible technical solution for the development of wind energy generation systems [7]. In the technical field, Qiu et al. proposed to introduce a formal error construction factor in the conventional EMD to satisfy the extraction of time series signal values. The experimental results show that compared with the traditional EMD, the error of the improved EMD is reduced by 14.52 % and the component share is reduced by 9.08 %, which is more suitable for the extraction of time series signal values [8]. Wang et al. applied the EMD to rolling bearing fault early warning by using an acceleration sensor to collect its vibration signal and process the signal value based on the EMD, and then obtain the fault wear. The method was applied to rolling bearing fault warning. The experimental results show that the early warning accuracy of the method is above 94 % and the alarm response time is 0.27s [9].

K-mean clustering algorithm is a common clustering algorithm that discovers the clustering structure and patterns in a dataset by randomly dividing the dataset into K subsets. It is widely used in different fields because of its high clustering performance. In the field of data processing, Zhang et al. designed a data security privacy protection scheme using K-mean clustering algorithm and proposed to update data centers without disclosing data privacy. Experimental results show that the proposed K-mean clustering protection method can complete data protection in a faster response time [10]. In the field of earthquake prediction, Jeong et al. designed an unsupervised machine learning geophone for earthquake signal wave monitoring, and after obtaining the signal wave data, the data were analyzed using a K-mean clustering algorithm to achieve the capability of earthquake prediction. The experimental results show that the prediction method is feasible in the field of seismic [11]. In the field of localization, Pinto et al. proposed an indoor localization system, which uses a K-mean clustering algorithm to characterize the indoor environment and then combines it with a Bayesian model for predicting the environmental location [12]. The experimental results show that the algorithm has good localization accuracy and improves the prediction accuracy by 12 % compared with the traditional scheme, and the system can be effectively applied in the field of indoor localization. In industrial applications, Zheng designed a static friction detection scheme for industrial loops by combining K-mean clustering algorithm and moving windows. The scheme is used for viscous loop friction detection in viscous control valves and aims to improve the control of control valves. Experimental results show that the scheme not only has high detection accuracy, but also can be used to detect the unexpected closing of valves with high feasibility [13].

In summary, EMD algorithm and K-mean clustering algorithm have good application effects in different fields, but the application of both algorithms in the health state prediction of new energy electric vehicle batteries is still relatively rare. In order to investigate the application effect of the two algorithms on battery SOH prediction, this study proposes to use the improved EMD algorithm to extract and process the vibration signal of the battery, and then use the optimized K-mean clustering algorithm to cluster and identify the signal values to achieve the health state prediction of lithium batteries.

3. Research on prediction of lithium battery health state based on vibration signal and clustering

As the main on-board power system of new energy electric vehicles, it is of high research value to predict the health status of their operation of lithium batteries. Based on this, the study intends to explore the SOH characteristics of the battery, use the improved EEMD to collect the battery vibration signal, and combine the K-mean clustering-PSO algorithm for signal feature clustering, aiming to complete the battery SOH prediction.

3.1. Study of SOH characteristics of new energy electric vehicle lithium batteries

As the main power source of new energy electric vehicles, the health status of batteries greatly affects the quality of vehicles.

Among them, Li-ion battery has become the main battery category of new energy electric vehicles nowadays because of its advantages of high energy density and long life cycle.

Fig. 1 shows the charging and discharging principle diagram of lithium ion. It can be seen that in the charging state, the positive electrode of the battery will provide lithium ions in the free state, and the lithium ions move continuously through the electrolyte and the electric diaphragm and finally reach the negative electrode of the battery. Due to the charge conservation theorem, an equal amount of electrons will arrive from the negative terminal to the positive terminal through the external circuit. The discharge process is the opposite of the charging process, where the lithium ions move from the negative terminal to the positive terminal and the electrons move from the positive terminal to the negative terminal. During the continuous charging and discharging operation, the battery will gradually age and its SOH performance will change and deteriorate accordingly [14]. The deterioration of battery SOH is mainly reflected in two aspects. One is the decrease of capacitance content. The second is the rise in resistance, resulting in a decrease in power.

Fig. 2 shows the factors influencing the deterioration of battery SOH. It can be seen that the factors of battery SOH deterioration are divided into internal factors and external factors. There are three internal factors, the first one is electrode corrosion leading to the decrease of capacitance content. The cyclic transfer of lithium ions can lead to corrosion between the positive and negative electrodes, causing a decrease in capacitance capacity. The second one is the loss of electrolyte solvent leading to the decrease of capacitance content. Since the moving environment of lithium ion is located in the electrolyte, the long time moving will bring the loss of electrolyte solvent, which will affect the capacitance capacity. The third is the increase in resistance caused by the aging of the electrical diaphragm. When lithium ions move in positive and negative cycles, they will continuously pass through the electric diaphragm, and over a long period of time the electric diaphragm will suffer from loss, which will cause the resistance to rise. External factors are divided into four, the first is the charge multiplier, multiplier will be too large will lead to battery polarization, multiplier is too small will prolong the charging time. The second is the cut-off voltage, if the cut-off voltage is too low, it will cause the battery energy loss, resulting in a sharp rise in resistance and accelerate the rate of battery performance decline. The third is the use of temperature, too high and too low temperature will affect the reaction rate of the battery, thus affecting the use of performance. The fourth is the battery overcharge and discharge. When overcharging, lithium ions will be deposited in the negative electrode, and when over discharging, lithium ions will be shed. Both phenomena lead to impaired capacitance content [15].

The performance of battery SOH is mainly reflected in the capacitance and resistance of the battery, and when the capacity of the battery decreases and the resistance increases, it means that the performance of the battery is decreasing. Thus, the performance of the battery SOH can be indirectly characterized by measuring the capacity and resistance of the battery. The formula for calculating the loss of battery capacitance is shown in equation (1).

$$S_C = \frac{C_0 - C_{now}}{C_0} \times 100\% \tag{1}$$

In equation (1), C_{now} is the current maximum capacitance content of the battery. C_0 is the maximum capacitance content of the initial state of the battery. S_C is the degree of loss of battery capacitance. When the capacitance of the battery reaches the minimum use standard of the battery life, the battery needs to be disposed of at end-of-life. The formula for calculating the change in battery resistance is shown in equation (2).

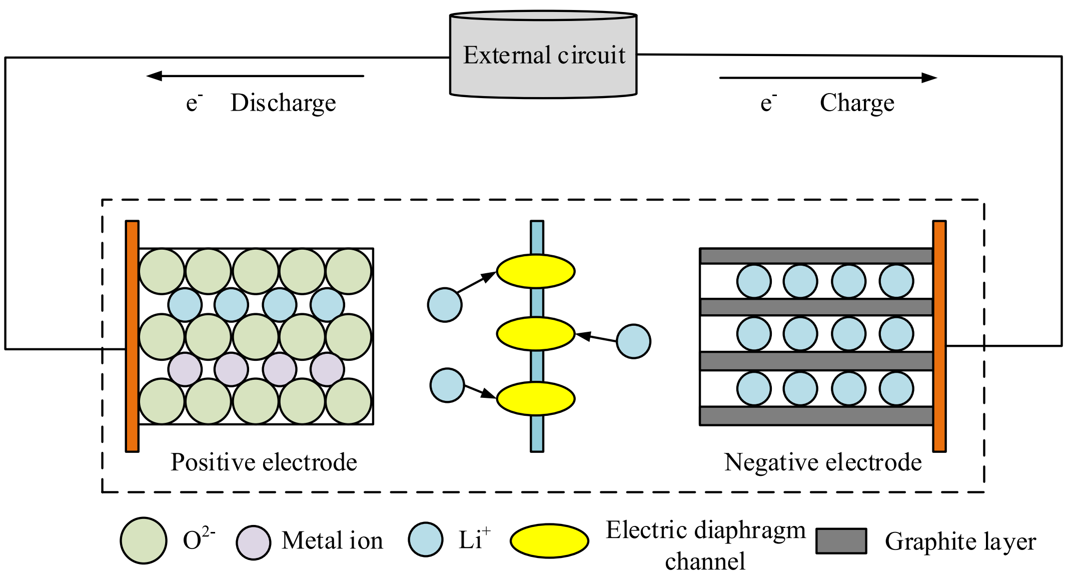

Fig. 1. Charging and discharging principle diagram of lithium battery.

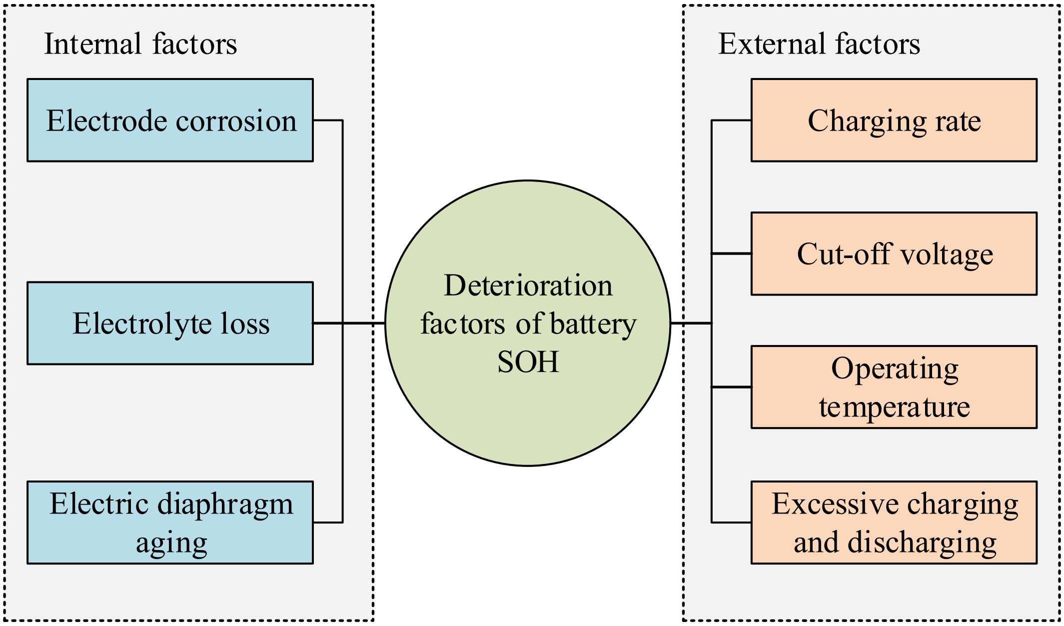

Fig. 2. Factors influencing the deterioration of cell SOH.

$$S_R = \frac{R_e - R_{now}}{R_e - R_0} \times 100\% \tag{2}$$

In equation (2), R_e is the resistance of the battery at the minimum usage standard. R_{now} is the resistance of the battery at the present time. R_0 is the initial resistance of the battery. S_R is the degree of change in resistance. When the magnitude of the resistance reaches twice the initial resistance, the battery needs to be disposed of at end-of-life. Therefore, in order to judge the change of battery performance, the signal characteristics need to be collected for battery performance analysis.

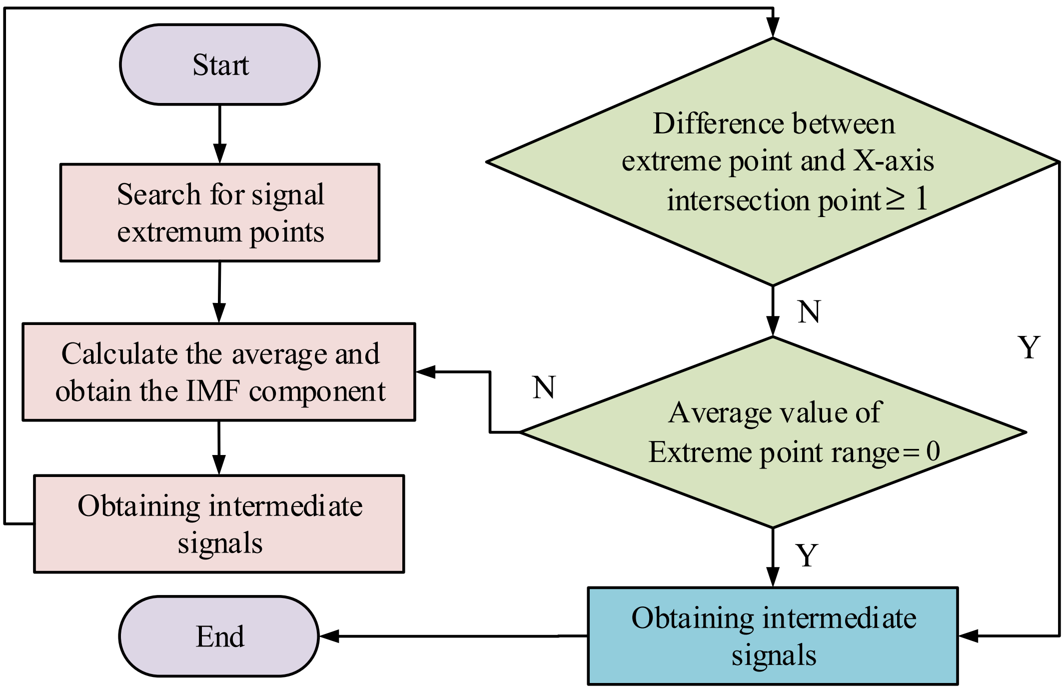

Fig. 3. EMD operation flow.

3.2. Study of battery vibration signal feature extraction based on empirical modal decomposition

For the battery, when its performance changes, the value of the vibration signal generated by it will also fluctuate. A reasonable acquisition frequency helps the subsequent analysis of the mnemonic eigenvectors [16]. According to Shannon's theorem, the basis of acquisition frequency judgment can be obtained as shown in equation (3).

$$\varpi > 2\varpi_c \tag{3}$$

In Equation (3), ϖ is the actual acquisition frequency. ϖ_c is the cutoff frequency. The value of the cutoff frequency is generally determined based on the actual acquisition. EMD is widely used in signal processing because of its wide processing range [17].

Fig. 3 shows the operation flow of EMD. The first step of the operation is to search for the extreme points of the signal, and the search process is to cover the original fluctuation range by enveloping and connecting the extreme points of the local original signal with an envelope. The average value of the signal points within the coverage range is then calculated as the Intrinsic Mode Functions (IMF). Then the original signal value and the average IMF component are subtracted to obtain the intermediate signal value. If the difference between the extreme value point of the intermediate signal value and the number of intersection points of the curve and X-axis is within 1, the intermediate signal value can be determined as the IMF component. If the above conditions are not satisfied, but when the mean value within the extreme value of the envelope is 0, it can also be determined as the IMF component. If both conditions are not satisfied, a new intermediate signal value is retrieved until the IMF component condition is satisfied. The conventional EMD is subject to modal mixing in the process of signal extraction. In order to solve the above problems, this study improves it and proposes an EEMD for lithium battery vibration signal feature extraction [18]. Firstly, different levels of white noise are randomly added to the original signal values to obtain the ensemble signal, whose expression is shown in equation (4).

$$s(t) = x(t) + n(t) \tag{4}$$

In Equation (4), $s(t)$ is the ensemble signal. $x(t)$ is the original signal. $n(t)$ is the white noise signal. Then the ensemble signal $s(t)$ is decomposed to obtain different levels of IMF components $c_i(t)$ and its residual term $r_m(t)$, whose expressions are shown in Equation (5).

$$s(t) = \sum_{i=1}^{m} c_i(t) + r_m(t) \tag{5}$$

Then a cyclic value N is selected, and a cyclic operation is performed on Eqs. (4) and (5), and a different value of white noise signal is added in each operation to homogenize it, and the operation expression of this process is shown in Equation (6).

$$c_i = \frac{1}{N}\sum_{j=1}^{N} c_{j,i} \tag{6}$$

The mean values of different IMF components can be obtained by equation (6), which is the output result of EEMD signal

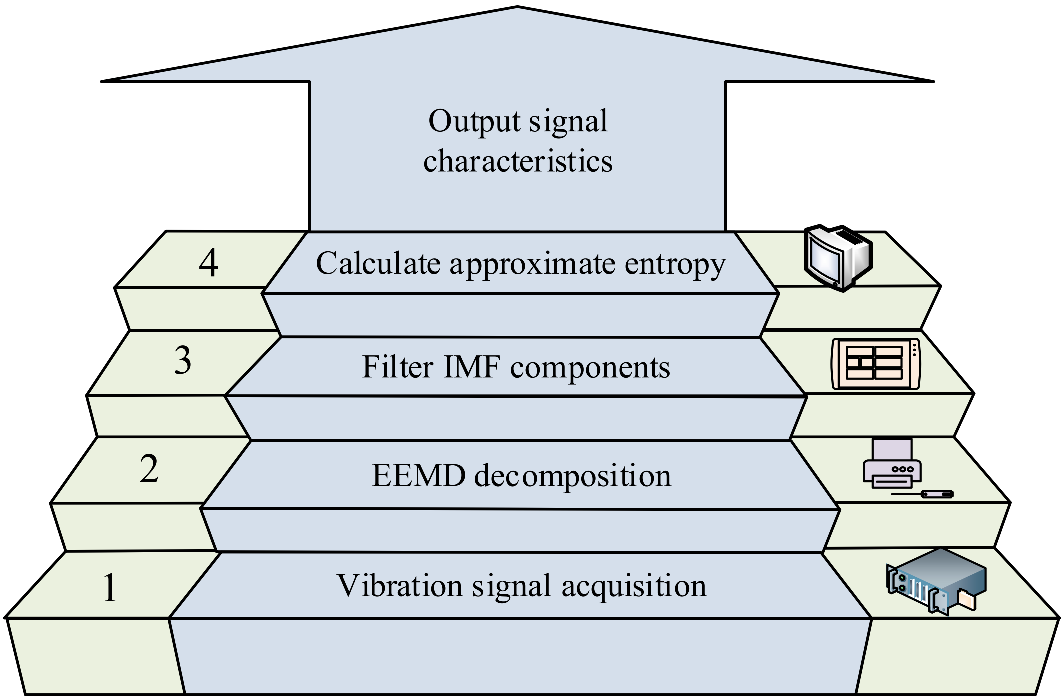

Fig. 4. EEMD-ApEn cell vibration signal feature extraction process.

processing. Since the signal characteristics of lithium battery are more complex and have more interference, an approximate entropy (ApEn) is introduced in this study to process the signal characteristics of its EEMD output. ApEn can quantize the noise-containing feature vectors to reduce the interference effect. The quantization formula of ApEn is shown in equation (7).

$$ApEn(m,r)=\lim_{M\to\infty}\left[5^{m}(r)-5^{m+1}(r)\right] \tag{7}$$

In Equation (7), m is the predetermined number of dimensions. r is the approximate capacity. M is the number of points. 5 is the entropy value.

Fig. 4 shows the process of EEMD-ApEn battery vibration signal feature extraction. It can be seen that the extraction of battery vibration signal features firstly determines the signal acquisition frequency, then performs the signal value decomposition of EEMD to obtain different IMF components, then quantifies the IMF components by ApEn operation, and finally uses the quantified signal feature values as the output.

3.3. Research on battery vibration signal feature recognition based on improved K-mean clustering algorithm

In order to predict the health status of a battery based on the characteristic signals collected from the battery, it is necessary to identify the signal type. This study uses the K-mean clustering algorithm to identify the signal features. The K-mean clustering algorithm is based on distance as an indicator for identification, which has the advantage of fast cluster identification. Since the collected battery vibration signal is not directly used for algorithm recognition, it needs to be pre-processed with noise reduction [16]. The wavelet transform noise reduction method has a good noise reduction effect because it can be adaptively adjusted according to the difference between the noise signal and the sample signal. The essence of the wavelet transform is the wavelet basis function $\psi(t)$, which is the product of the signal value and the acquisition time. $\psi(t)$ The expression of is equation (8).

$$\psi_{a,b}(t)=\frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) \tag{8}$$

In equation (8), a is the scale factor. b is the translation, and t is the time of signal acquisition. Assuming that there exists a signal characterized by $f(t)$, the expression of its wavelet transform is shown in Equation (9).

$$\langle f(t)|\psi_{a,b}(t)\rangle=\frac{1}{\sqrt{a}}\int f(t)\psi\left(\frac{t-b}{a}\right)dt \tag{9}$$

The collected battery signal values can then be pre-processed by equation (9). The processed signal values are then set to the data set X, on which the K-mean clustering operation is performed, with the expression shown in equation (10).

$$X=\{x_i|x_i\in R^m, i=1,2,\ldots,n\} \tag{10}$$

In equation (10), x_i is the i th data in the data set X. R^m is the sample set. m is the number of dimensions of the sample set. n is the

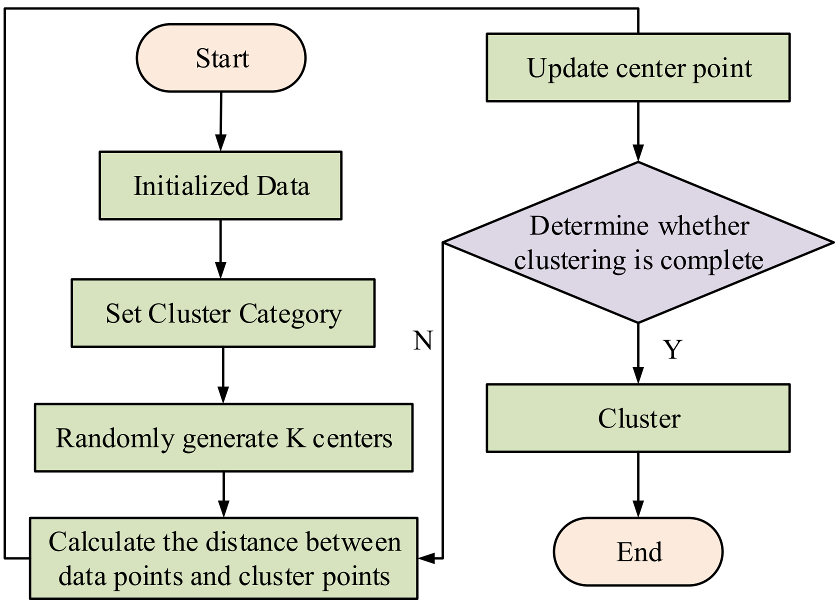

Fig. 5. K-mean clustering algorithm flow chart.

number of samples contained in the sample set. The expression of the category set μ of the sample set R^m is shown in equation (11).

$$\mu = \{\mu_j | \mu_j \in R^m, j = 1, 2, ..., p\} \tag{11}$$

In equation (11), μ_j is the category j of the category set μ. p is the number of categories. Meanwhile, the distance between two samples in the sample set is calculated using the Euclidean distance algorithm, whose calculation formula is shown in Equation (12).

$$D(x_i, x_j) = \sqrt{(x_i - x_j)^T (x_i - x_j)} \tag{12}$$

In equation (12), $D(x_i, x_j)$ is the Euclidean distance between samples x_i and x_j. Then, k samples are selected as cluster centers, and the distances between samples and different cluster centers are calculated according to equation (12), and the nearest distance is used for cluster identification of samples.

Fig. 5 shows the flow chart of K-mean clustering algorithm. Its operation process is to first process the sample data in the sample set, then set up random k clustering centers, and then calculate the distance between samples and clustering centers based on equation (12), and complete the sample clustering with the minimum distance as the index. When all samples are finished classification, then update the center point and repeat the above clustering process until the distance between all sampling points reaches the shortest. Although the traditional K-mean clustering algorithm has the advantage of strong recognition ability, its recognition process is more complicated and the globalization is poor [19]. In order to optimize the above problem, this study improves it by using PSO algorithm, which is an algorithm suitable for global optimal solution search with simple operation, and combines it with the traditional K-mean clustering algorithm to optimize its operation performance.

Fig. 6 shows the operational flow chart of the K-means clustering-PSO algorithm. Firstly, the samples should be initialized, and the process is to set the initial position and velocity of the sample particles, and then update them by calculating the best position of the sample individuals in the group and the velocity fitness value, and optimize them with the K-mean clustering algorithm. After that, the samples are classified according to the Euclidean distance between the clustering center and the sample particles, and if the maximum number of iterations of the algorithm is reached, the sample classification result is output, otherwise, the sample particle iteration calculation is continued. The different clustered sample feature data in the output result represent the real-time status of the battery, and the health status of the battery can be analyzed in real time according to the output signal features, and its health status is predicted according to equations (1) and (2), and an alarm is issued when the SOH of the battery reaches the minimum usage standard [20].

4. Performance analysis of battery SOH prediction based on EEMD and K-mean Clustering-PSO algorithm

In this study, in order that the battery SOH prediction accuracy can be improved, an improved EEMD algorithm is proposed to collect the battery vibration signals to improve the accuracy of signal acquisition. Then the type of vibration signal is identified by

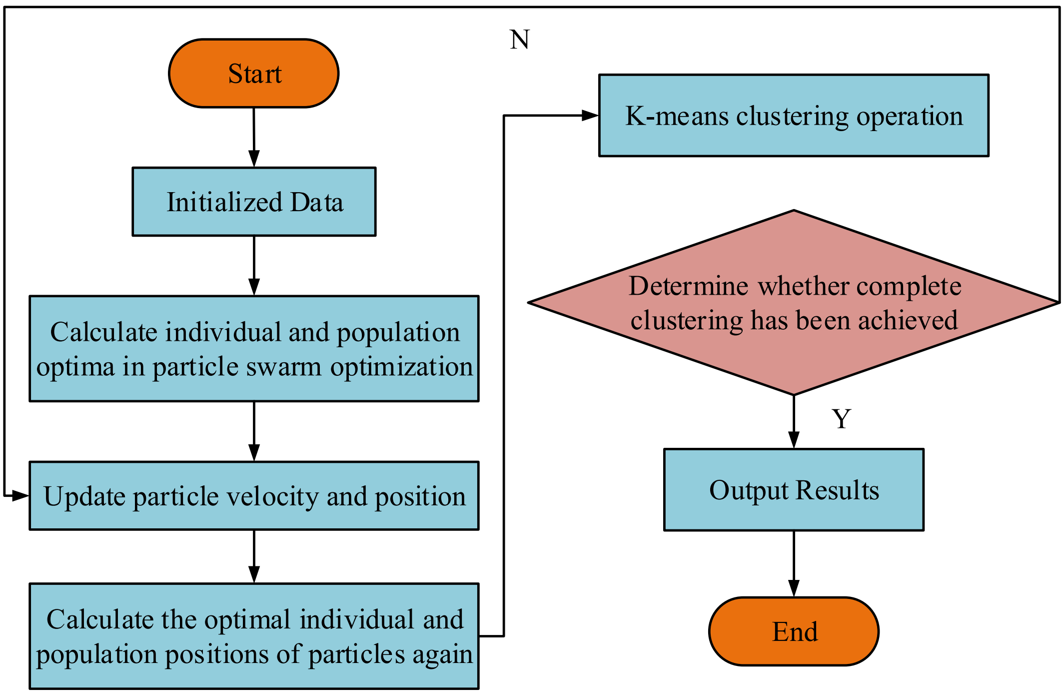

Fig. 6. Operational flow chart of K-means clustering-PSO algorithm.

combining K-mean clustering algorithm and PSO algorithm to achieve the battery SOH prediction. In order to verify the performance of the algorithm proposed in the study for practical applications, several sets of experiments were set up in this study for analysis.

4.1. Performance analysis of EEMD-based battery vibration signal algorithm

In this study, in order to verify the performance of the optimized EEMD algorithm for battery vibration signal extraction, traditional EMD and traditional EEMD were used as comparison algorithms for experimental analysis. The vibration data of the battery in three states of normal state, overcharge and overdischarge are selected for the experiments, and the signal values are extracted using the above two comparison algorithms and the vibration data of the EEMD-ApEn algorithm proposed in the study, and the extracted signal values are compared with the real signal values, and the performance analysis of the algorithm is carried out according to the comparison results.

Table 1 shows the experimental environment table. The signal acquisition frequency and cell specifications for this experiment are shown in Table 1. NCM111 battery is a lithium-ion battery composed of cobalt, nickel, manganese and other elements. The rated capacity of the battery is 3000 mA h. Generally, constant current and constant voltage charging is used. The charging current should be controlled at around 0.5C of the rated capacity of the battery. The charging voltage should be controlled at around 4.2 V. When the battery voltage reaches 4.2 V, it should be switched to constant voltage charging until the charging current drops to around 0.05C. The discharge conditions of NCM111 batteries are generally between 0 °C and 45 °C Within the temperature range of. The discharge current should be determined based on the rated capacity of the battery, and it is generally recommended to control it between 0.2C and 1C of the rated capacity of the battery. The experiments were conducted using the conventional EMD algorithm, the conventional EEMD algorithm and the modified EEMD-ApEn algorithm to acquire signals at an acquisition frequency of 20 kHz, and then the algorithms were analyzed based on the acquired vibration signal values.

Fig. 7 shows the IMF signal components of different algorithms in the normal state of the battery. Among them, Fig. 7(a), (b), (c) and (d) show the IMF signal components under the real value, EMD algorithm, EEMD algorithm and EEMD-ApEn algorithm, respectively. It can be seen that among the three algorithms, the IMF signal components of EEMD-ApEn are more consistent with the real values, followed by the EEMD algorithm and finally the EMD algorithm. The EEMD-ApEn algorithm can achieve higher accuracy in the signal detection process because it has the ability to effectively avoid the problems of modal aliasing and noise interference. Therefore, the EEMD-ApEn algorithm can effectively detect the normal state vibration signal value of the battery.

Fig. 8 shows the IMF signal components for different algorithms in the battery overcharge state. Among them, Fig. 8(a), (b), (c) and (d) show the IMF signal components under the real value, EMD algorithm, EEMD algorithm and EEMD-ApEn algorithm, respectively. It can be seen that the detected signal values of the EEMD-ApEn algorithm are closer to the real vibration signal in the state of battery overcharge. Between 4 and 7s, the vibration signal values detected by the EMD algorithm are somewhat different from the real values. Therefore, the EEMD-ApEn algorithm can effectively detect the vibration signal in the battery overcharge state.

Fig. 9 shows the IMF signal components of different algorithms in the over-discharged state of the battery. Among them, Fig. 9(a), (b), (c) and (d) show the IMF signal components under the real value, EMD algorithm, EEMD algorithm and EEMD-ApEn algorithm, respectively. It can be seen that the fluctuation range of the vibration signal detected by the three algorithms is approximately the same as the real value in the over-discharged state of the battery, and the fluctuation trend of the EEMD-ApEn algorithm is more consistent with the real value. Therefore, the EEMD-ApEn algorithm is feasible in detecting the vibration signal in the over-discharged state of the battery.

4.2. Performance analysis of battery SOH prediction based on K-mean clustering-PSO algorithm

In order to verify the prediction performance of the battery SOH by the K-mean clustering-PSO algorithm proposed in the study, this experiment uses the traditional K-mean clustering and PSO algorithm as the comparison algorithm, and collects the vibration signal data of 10 groups of batteries in three states: normal state, overcharge and overdischarge, and then uses the above three algorithms to classify and identify them respectively, and according to the rules of SOH determination, the battery The SOH of the battery is predicted and analyzed, and the SOH of the battery is determined and the prediction error of the above three algorithms is analyzed according to the classification results of the three algorithms.

Fig. 10 shows the comparison of the battery SOH classification performance with different algorithms. Among them, the algorithms used in Fig. 10(a), (b) and (c) are K-mean clustering algorithm, PSO algorithm and K-mean clustering-PSO algorithm, respectively. It can be seen that among the three algorithms, the K-mean clustering algorithm-PSO algorithm identifies the classification of batteries

Table 1
Table of experimental environment.

Parameter Type	Parameter Specifications
Battery Model	NCM111
Sampling frequency	20 kHz
Initial capacitance	100UF
Initial resistor	1 Ω
Cutoff capacitance	1000UF
Cutoff resistor	2 Ω

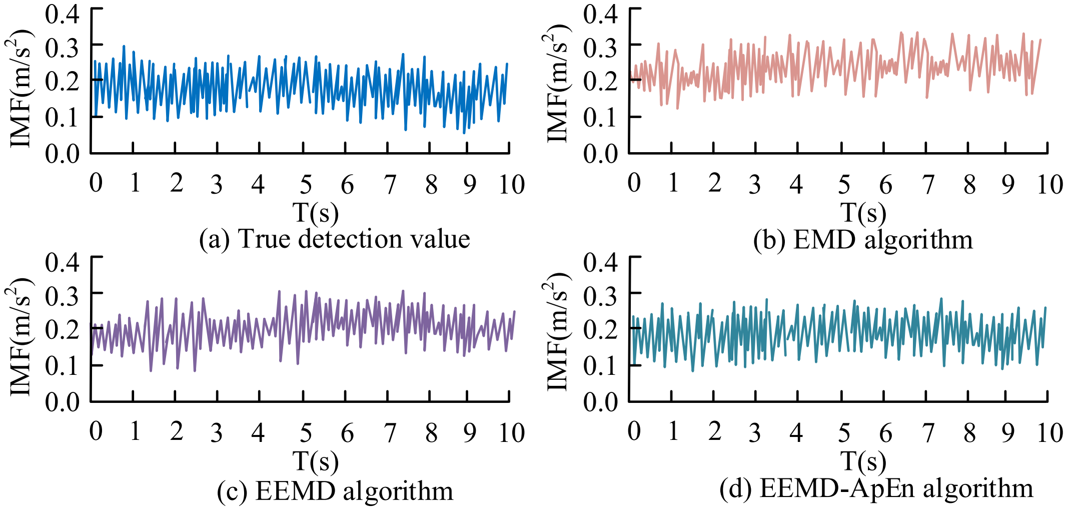

Fig. 7. IMF signal components of different algorithms in the normal state of the battery.

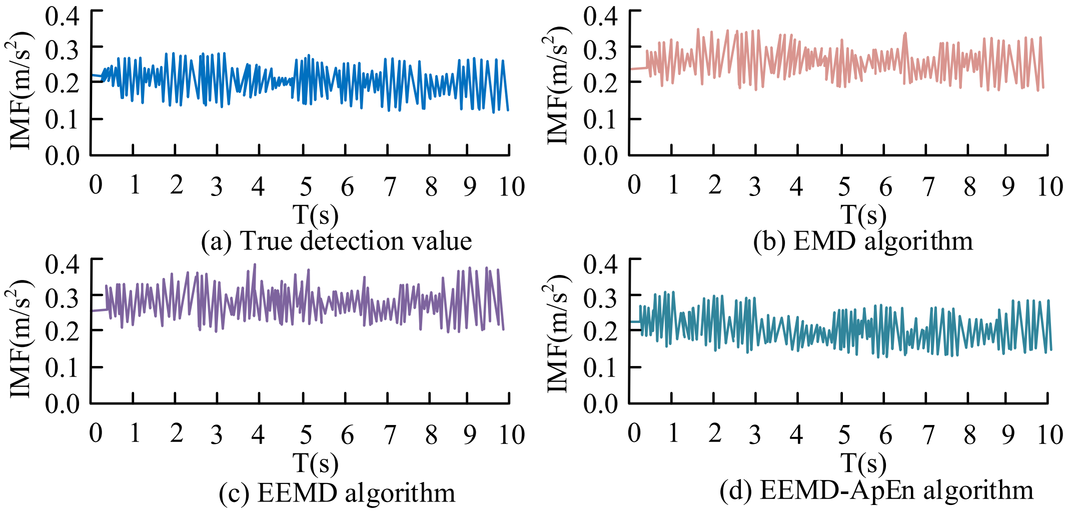

Fig. 8. IMF signal components of different algorithms in the battery overcharge state.

under different SOH more clearly, and the traditional K-mean clustering algorithm and PSO algorithm both show a slight overlap, which can affect the judgment of battery SOH, so the K-mean clustering-PSO algorithm has a better performance in the classification of battery SOH.

Fig. 11 shows the prediction curves of SOH of the battery by different algorithms. The predicted value of battery SOH by the K-mean clustering-PSO algorithm is most consistent with the real state of the battery, and when the battery is cycled up to 600 times, the predicted health state of the battery by this algorithm is 37 % of the initial state, a decrease of 63 %, while the real value is 38 %. In summary, the K-mean clustering-PSO algorithm can effectively predict the battery SOH.

Fig. 12 shows the prediction errors of different algorithms for the battery SOH. From the figure, it can be seen that the overall prediction error of the battery error by different algorithms increases with the increase of the number of cycles. This is because as the number of battery cycles increases, there will be more and more factors affecting the decline of battery SOH, and the overall state of the battery is more complex. Among the three algorithms, the K-mean clustering-PSO algorithm has a relatively lower prediction error of battery SOH, and the prediction error of the algorithm is only 2.6 % when the number of cycles reaches 600. Therefore, the K-mean clustering-PSO algorithm has higher feasibility for the prediction of the overall battery state.

5. Conclusion

As a clean and efficient transportation, new energy electric vehicles are receiving more and more attention. As an important component of electric vehicles, the health status of the battery has an important impact on the performance of the whole vehicle. In

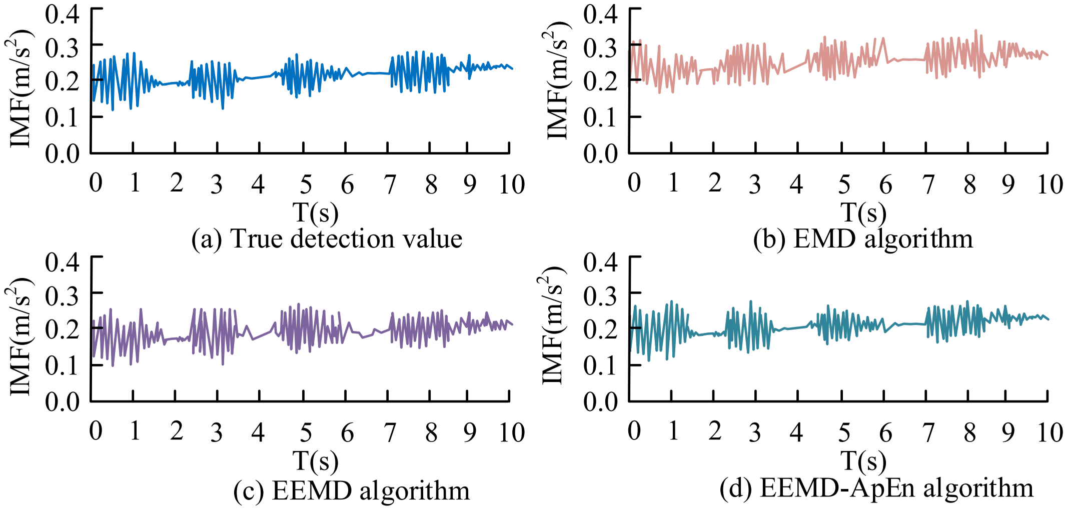

Fig. 9. IMF signal components of different algorithms in the over-discharged state of the battery.

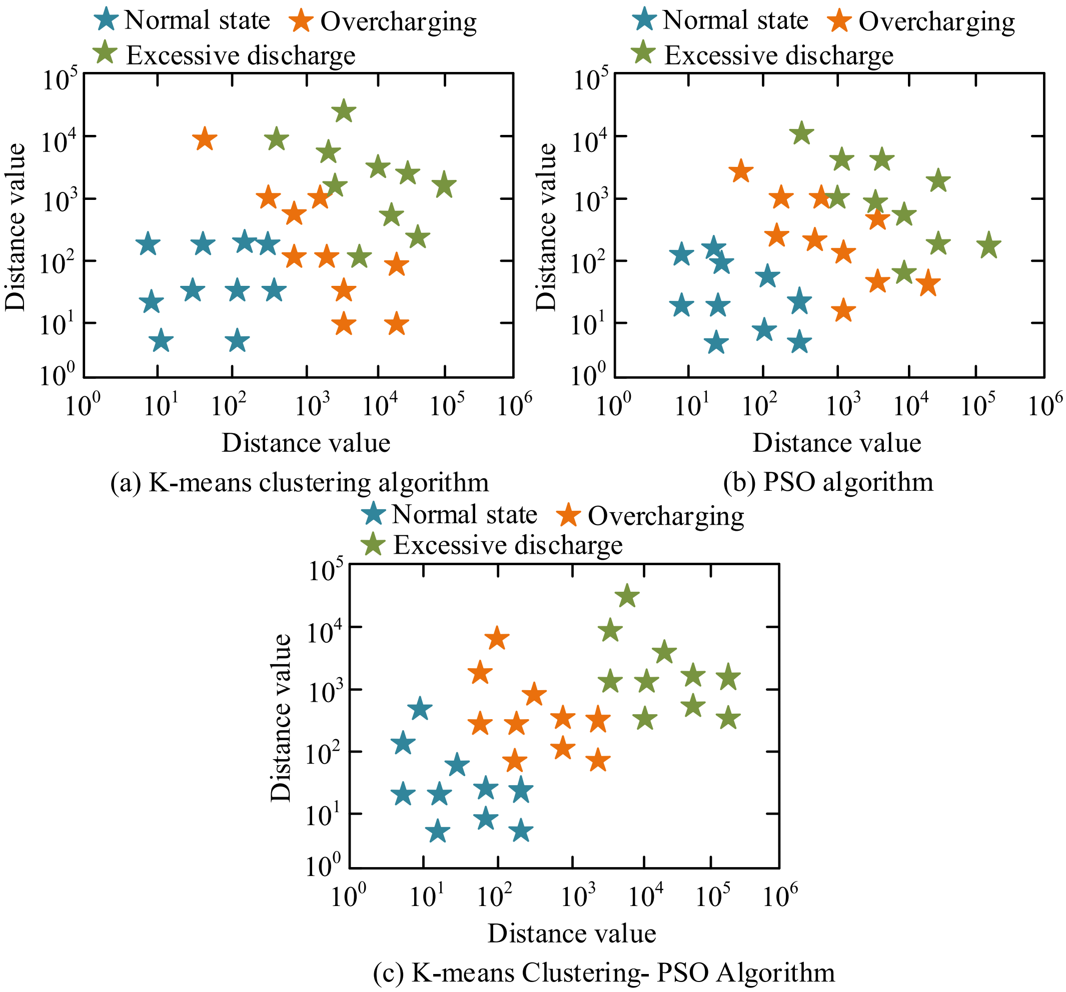

Fig. 10. Comparison of battery SOH classification performance under different algorithms.

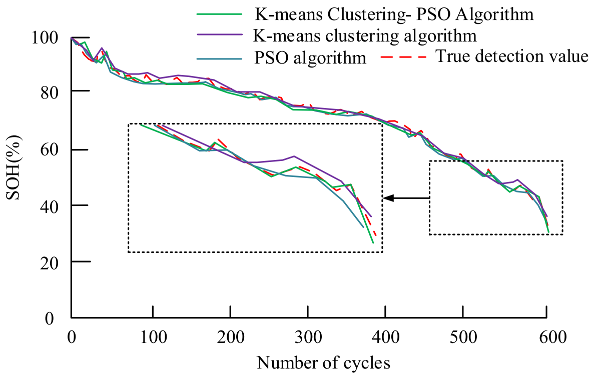

Fig. 11. Prediction curves of cell SOH by different algorithms.

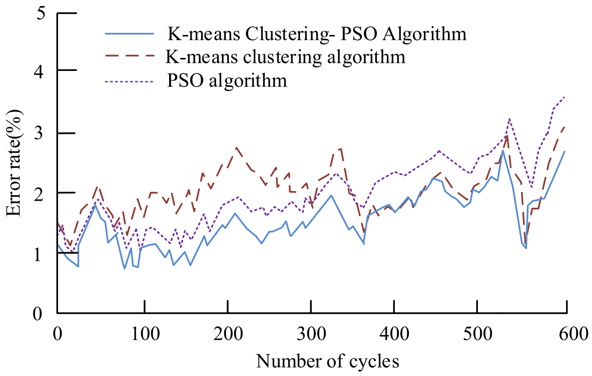

Fig. 12. Prediction errors of different algorithms for cell SOH.

order to predict the health status of the battery, this study obtained the EEMD algorithm by optimizing on the basis of EMD. The optimized EEMD algorithm was used to collect and process the vibration signal of the battery, and then the optimized K-mean clustering algorithm-particle algorithm was used to cluster the collected signal values to achieve the prediction effect. The experimental results show that the battery signal components in different states obtained by using the optimized EEMD algorithm proposed in the paper are highly consistent with the real values, and their signal fluctuations range between 0.1 and 0.4 m/s^2. The K-mean clustering-PSO algorithm can classify the battery signals in different state types more explicitly, and predict that the battery performance decreases by 63 % when it is cycled up to 600 times, and the algorithm has a prediction error of only 2.6 %. So the algorithm proposed in this study has some feasibility in SOH prediction of new energy electric vehicle batteries. The proposed algorithm can be used to predict the SOH of different types of batteries in subsequent experiments to verify the application scope of the algorithm.

Funding statements

No funding was received.

Data availability statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

CRediT authorship contribution statement

Liping Lu: Writing – original draft, Resources, Data curation. **Huiying Zhai:** Writing – review & editing, Formal analysis, Data curation. **Yun Gao:** Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R.R. Ardeshiri, C. Ma, Multivariate gated recurrent unit for battery remaining useful life prediction: a deep learning approach, Journal of Energy Research 45 (11) (2021) 16633–16648.
[2] L. Ren, J. Dong, X. Wang, Z. Meng, L. Zhao, D. Jamal, A data-driven auto-CNN-LSTM prediction model for lithium-ion battery remaining useful life, IEEE Trans. Ind. Inf. 17 (5) (2021) 3478–3487.
[3] K. Liu, Y. Shang, Q. Ouyang, W.D. Widanage, A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery, IEEE Trans. Ind. Electron. 68 (4) (2020) 3170–3180.
[4] C. Liu, Y. Wang, Z. Chen, Degradation model and cycle life prediction for lithium-ion battery used in hybrid energy storage system, Energy 166 (1) (2019) 796–806.
[5] Z. Wang, S. Zeng, J. Guo, Understanding the influence of state of health on the range anxiety of battery electric vehicle drivers, IET Intell. Transp. Syst. 15 (2) (2021) 286–296.
[6] S. Ghimire, R.C. Deo, C. David, S. Sancho, Improved complete ensemble empirical mode decomposition with adaptive noise deep residual model for short-term multi-step solar radiation prediction, Renew. Energy 190 (5) (2022) 408–424.
[7] H. Zhang, D. Yue, C. Dou, K. Li, G. Hancke, Two-step wind power prediction approach with improved complementary ensemble empirical mode decomposition and reinforcement learning, IEEE Syst. J. 16 (2) (2021) 2545–2555.
[8] X. Qiu, F. Wang, Y. Zhou, S. Zhou, Weighted empirical mode decomposition for processing GNSS position time series with the consideration of formal errors, Acta Geodyn. Geomater. 18 (3) (2021) 399–408.
[9] P. Wang, D. Li, N. Zhang, Research on early warning of rolling bearing wear failure based on empirical mode decomposition, Int. J. Mater. Prod. Technol. 63 (1/2) (2021) 72–85.
[10] E. Zhang, H. Li, Y. Huang, S. Hong, L. Zhao, C. Ji, Practical multi-party private collaborative k-means clustering, Neurocomputing 467 (7) (2022) 256–265.
[11] W. Jeong, M.S. Almubarak, C. Tsingas, Quality control for the geophone reorientation of ocean bottom seismic data using k-means clustering, Geophys. Prospect. 69 (7) (2021) 1487–1502.
[12] B. Pinto, R. Barreto, E. Souto, H. Oliveira, Robust RSSI-based indoor positioning system using k-means clustering and bayesian estimation, IEEE Sensor. J. 21 (21) (2021) 24462–24470.
[13] D. Zheng, X. Sun, S.K. Damarla, A. Shah, J. Amalraj, B. Huang, Valve stiction detection and quantification using a k-means clustering based moving window approach, Ind. Eng. Chem. Res. 60 (6) (2021) 2563–2577.
[14] M. Li, Y. Li, S.S. Choi, Dispatch planning of a wide-area wind power-energy storage scheme based on ensemble empirical mode decomposition technique, IEEE Trans. Sustain. Energy 12 (2) (2021) 1275–1288.
[15] U. Riaz, S. Aziz, M.U. Khan, S. Zaidi, M. Ukasha, A. Rashid, A novel embedded system design for the detection and classification of cardiac disorders, Comput. Intell. 37 (4) (2021) 1844–1864.
[16] Z. Yang, W. Huang, X. Chen, Mitigation of rain effect on wave height measurement using x-band radar sensor, IEEE Sensor. J. 22 (6) (2022) 5929–5938.
[17] F. Xia, K. Wang, J. Chen, State of health and remaining useful life prediction of lithium-ion batteries based on a disturbance-free incremental capacity and differential voltage analysis method, J. Energy Storage 64 (2023), 107161.
[18] Y. Fan, Y. Liu, H. Qi, F. Liu, X. Ji, Anti-interference technology of surface acoustic wave sensor based on k-means clustering algorithm, IEEE Sensor. J. 21 (7) (2021) 8998–9007.
[19] F. Xia, K. Wang, J. Chen, State-of-Health prediction for lithium-ion batteries based on complete ensemble empirical mode decomposition with adaptive noise-gate recurrent unit fusion model, Energy Technol. 10 (4) (2022), 2100767.
[20] X. Zhao, F. Nie, R. Wang, X. Li, Improving projected fuzzy K-means clustering via robust learning, Neurocomputing 491 (28) (2022) 34–43.