ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

# Fault Diagnosis Method for Lithium-Ion Battery Packs in Real-World Electric Vehicles Based on K-Means and the Fréchet Algorithm

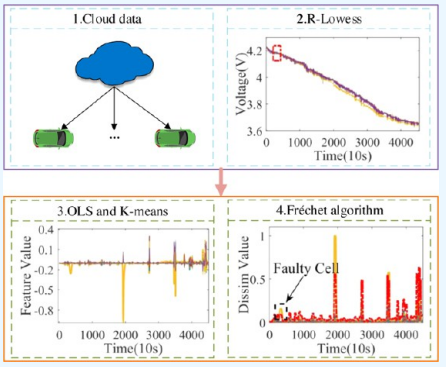
Minghu Wu, Wanyin Du, Fan Zhang,* Nan Zhao, Juan Wang, Lujun Wang, and Wei Huang

Cite This: *ACS Omega* 2022, 7, 40145−40162

Read Online

ACCESS | Metrics & More | Article Recommendations

**ABSTRACT:** Battery failure has traditionally been a major concern for electric vehicle (EV) safety, and early fault diagnosis will reduce many EV safety accidents. However, the short-circuit signal is generally very weak, so it is still a challenge to achieve a timely warning of battery failure. In this paper, an initial microfault diagnosis method is proposed for the data of electric vehicles in actual operation. First, a robust locally weighted regression data smoothing method is proposed that can effectively remove noisy data and retain fault characteristics. Second, an ordinary-least-squares-based voltage potential feature extraction method is proposed, which can effectively capture the small fault features of battery cells and achieve early warning. Third, a reference cell selection method based on K-means clustering is proposed, which can effectively reduce the false alarms caused by the inconsistency of each cell. Fourth, the Fréchet algorithm is introduced into the field of battery pack fault diagnosis and combined with thresholds for battery pack fault diagnosis and localization to accomplish the diagnosis and early warning of minor faults. Finally, the fault diagnosis method is validated by three actual running electric vehicles to verify the effectiveness, reliability, and robustness of the method.




## 1. INTRODUCTION

The rapid development of electric vehicles in the world has made lithium-ion batteries a popular development as clean energy in the coming years.[1−3] Compared with traditional fuel vehicles, electric vehicles use rechargeable and dischargeable batteries as the power system, which can reduce the environmental pollution caused by fuel consumption. However, while electric vehicles are commonly used, often some car safety accidents are caused by battery failure, which greatly affects driving safety.[4] Battery failure is generally caused by mechanical abuse, electrical abuse, and thermal abuse, which in serious cases can trigger thermal runaway and lead to spontaneous combustion.[5] Therefore, realizing early warning of battery failure can effectively decrease safety accidents and bring a safer driving environment for drivers.

The current research on battery for electric vehicles has been mentioned in many types of literature, such as battery fault diagnosis, estimation of remaining useful life for batteries, state of the health estimation, etc.[6−12] And the research approaches in the literature about fault diagnosis can be broadly classified into three categories: knowledge-based, model-based, and data-driven fault diagnosis approaches.[13−18] Among them, the knowledge-based fault diagnosis method uses some historical and empirical knowledge of the battery to design some diagnostic rules for fault diagnosis.[19] The model-based approach is to establish a physical model of the battery, which is generally capable of accurately calculating the values of the parameters of the battery. This value is subsequently compared with the collected battery parameters. If their difference exceeds a threshold value, a fault is considered to have occurred.[20] Data-driven approaches based on data generally start with the collection of data, extraction of the features of the data, and performing fault diagnosis using algorithms such as outlier detection.[4] For example, in ref 21, authors connect a constant voltage source in parallel with the battery to be detected and then observe the current direction between the battery and the constant voltage source. Based on the current flow direction, it can be determined whether the battery is malfunctioning or not. However, the constant voltage source is sometimes difficult to add and is very inconvenient for real-time monitoring of electric vehicles. In ref 22, a partial-differential-equation (PDE) method was proposed to detect and estimate the severity of battery thermal faults in real time. However, it is sometimes difficult to apply the method in practice because of the difficulty in collecting the battery temperature of electric vehicles in actual operation. The authors utilized an observer based on an electrochemical model and a fuzzy logic algorithm that can be implemented in real time. A battery internal fault diagnosis

Received: August 5, 2022
Accepted: October 6, 2022
Published: October 25, 2022





https://doi.org/10.1021/acsomega.2c04991
*ACS Omega* 2022, 7, 40145−40162

method was developed using the relationship of residuals, which can reliably detect various faults inside lithium-ion batteries.[23] However, the method requires a large amount of historical fault data for rule building and fewer fault data in actual operation. A battery fault diagnosis method was developed in ref 24 using LSTM networks in combination with a battery equivalence model. The method was studied based on actual operational data and can effectively identify faulty monoliths. However, the method is sometimes difficult to train the model and requires a large amount of historical data. The authors in ref 25 performed fault diagnosis based on a battery with a ring topology and were able to locate the faulty battery using a recursive least-squares-based approach. However, the portability of the method is poor. The authors in ref 26 use the Kernel Principal Component Analysis (KPCA) approach to train a nonlinear data model for internal short-circuit detection of lithium-ion batteries. However, the method requires a large amount of historical data for offline training. The authors in ref 27 calculate the correlation coefficients of adjacent cells within a cell group and are able to quickly locate cells with abnormal fluctuations in voltage data. However, this method has low robustness, is easily affected by noisy data, and is difficult to define the threshold value. The authors in ref 28 use a correlation coefficient for faulty monomer detection of battery packs. The method first estimates the SOH parameters of the battery using the extended Kalman filter. Then, the SOH correlation between the cells is calculated and the faulty monomer number is determined using the threshold method. This method has difficulty in SOH estimation and is difficult to implement in real vehicles. A feature that most of the internal short-circuit faults have a larger SOC variance is proposed in ref 29. A mutual information method for fault cell identification is proposed for this feature. However, this method requires high model accuracy, and the SOC accuracy in actual operation is sometimes difficult to meet. The authors in ref 30 estimated the battery and system parameters for fault diagnosis and proposed a fuzzy clustering method for fault diagnosis. However, the model of this method is sometimes difficult to calculate, and the actual operating electric vehicle data cannot satisfy the algorithm.

In summary, many of the current methods are fault diagnostics studied in a laboratory environment and are difficult to apply on actual vehicles. In addition, the fault diagnosis algorithm developed based on the model has high requirements for hardware and additional equipment when applied on actual operating vehicles due to the high requirements for model accuracy. The fault diagnosis algorithm developed based on knowledge needs a large amount of fault data support, and the lack of fault data in practical applications is therefore difficult to establish model rules. Therefore, it is difficult to be applied to the battery management system (BMS). The data-driven approach, on the other hand, can directly start from the more easily collected voltage data and perform fault diagnosis by the change characteristics of the voltage data. Therefore, this paper develops a data-driven early warning algorithm for lithium-ion batteries based on data driven for minor faults.

Based on the voltage data, this paper develops a fault warning algorithm for electric vehicle lithium-ion battery packs based on K-means and the Fréchet algorithm. And the actual collected EV driving data are used to verify. First, due to the noise of the EV data collected in actual operation, it will affect the accuracy of the diagnosis algorithm. Robust locally weighted regression algorithm (R-Lowess) is proposed for data cleaning. Second, it is difficult to be detected in the BMS because the fault characteristics of the faulty battery are very slight. Therefore, ordinary least squares (OLS) is proposed for extracting trend components and potential features of voltage data and normalizing them. Third, there may also be inconsistencies between batteries of the same model due to errors in the manufacturer's production process, which can affect the diagnostic accuracy of the algorithm. To solve this problem, a K-means-based clustering method is proposed to select a reference cell to represent the current process operation. Finally, to achieve automatic fault detection and localization, the discrete Fréchet algorithm is proposed, and the extracted new features are used as the input to the algorithm. And adaptive thresholds are set for the detection and localization of faulty cells. To the best of our knowledge, the discrete Fréchet algorithm is presented for the first time in the field of faulty detection of battery packs.

The remainder of this paper is organized as follows. In Section 2, the fault diagnosis algorithm is introduced in detail, including the data cleaning process, feature extraction process, reference cell selection process, fault detection process, etc. In Section 3, three actual operational vehicles are selected to validate our fault diagnosis algorithm. Section 4 is the summary and outlook.

## 2. PRINCIPLE AND STRUCTURE FOR FAULT DIAGNOSIS ALGORITHM

**2.1. Real-World Vehicle Data Description.** The real vehicle data used in this paper comes from the big data management platform for new energy vehicles, which currently serves more than 200,000 vehicles and can collect more than 10 million pieces of data a day. In this paper, as shown in Table 1,

**Table 1. Parameters on the Three Vehicles**

| no. | vehicle type | faulty cell | cell number | sample interval (s) |
|---|---|---|---|---|
| #C1 | normal | | 95 | 10 |
| #C2 | potential fault | #47 | 95 | 10 |
| #C3 | potential fault | #31 | 96 | 10 |

we select one vehicle #C1 that did not show any fault during driving and two vehicles #C2 and #C3 that showed short-circuit fault for algorithm experimental verification. The vehicles are described as follows.

Vehicle #C1 is a normal vehicle with no faults as of the latest data collection moment. Figure 1 shows the unprocessed discharge process data of normal vehicle 1. The experimental data were selected as the last voltage data of the discharge process transmitted to the big data platform. Vehicle #C1 consists of 95 battery cells connected in series, so each cell has a different voltage value, while the current value is the same. Vehicle #C1 had a sudden voltage drop at the 10th sampling point for #Cell 1−23 at the starting moment due to a sensor anomaly. This phenomenon is not brought about by the failure of the battery cells but is data noise, which is not part of the fault range and requires data cleaning.

In Figure 2, Vehicle #C2 was a failed vehicle with a power supply system consisting of 95 battery cells connected in series to form a power battery pack. #Cell 47 in the battery pack showed a sudden voltage drop at the 425th sampling moment, which was confirmed to be caused by a weak internal short circuit in the battery cell. However, due to the equalization mechanism of the battery pack and other factors, the voltage drop in the cell did not produce a large fluctuation after the occurrence in the microshort circuit but showed a similar trend


https://doi.org/10.1021/acsomega.2c04991
ACS Omega 2022, 7, 40145−40162

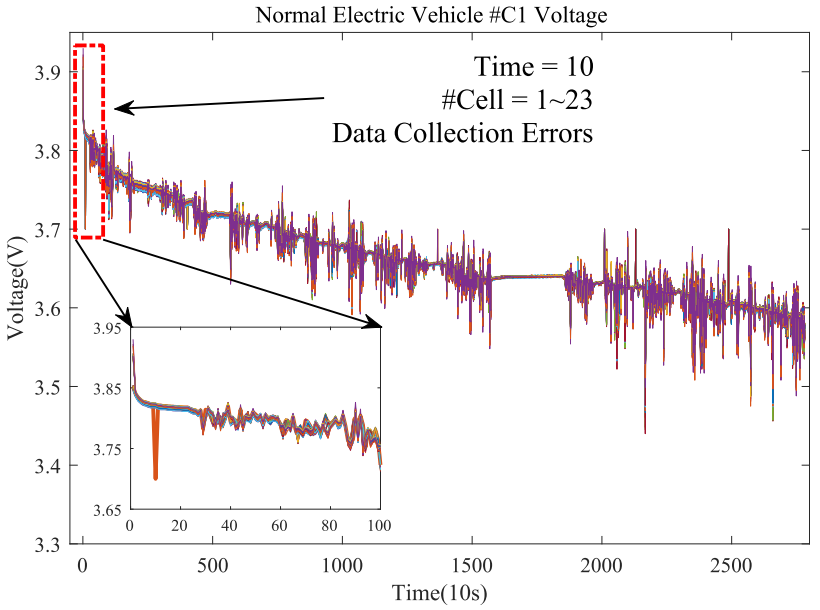



**Figure 1.** Normal vehicle #C1 discharge process raw data.

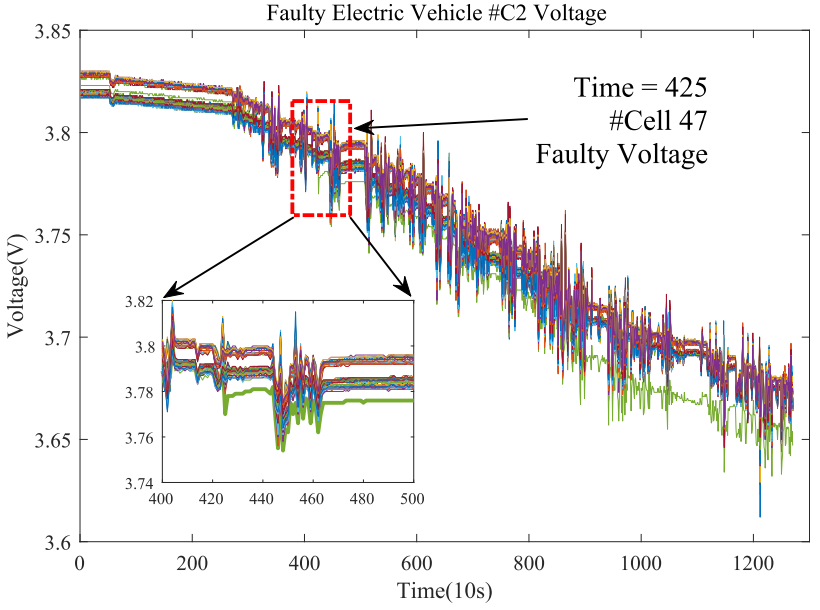



**Figure 2.** Faulty vehicle #C2 discharge process raw data.

in other normal cells. Until the 900th, 1100th, and other sampling moments, the evolution of the internal short circuit of the monomer intensifies, the voltage drops suddenly again, and a more noticeable voltage difference appears. Eventually, the battery management system (BMS) alarmed.

Similar to vehicle #C2, Figure 3 shows that vehicle #C3 is an electric vehicle that has experienced a failure with a battery pack of 96 battery cells. At the 350th sampling moment, battery #Cell 31 experienced a weak internal short circuit and a sudden voltage drop. Subsequently, #Cell 31 had a worsening internal short circuit at the 1960th sampling moment and the voltage dropped again. Eventually, the BMS alarm was raised. And at the 3600th sampling point, the cell showed a sudden voltage increase due to a sensor acquisition error, which was not part of the cell failure and was a data acquisition error.

**2.2. Data Cleaning Based on R-Lowess.** Different from laboratory-collected data, there is a large amount of noise in the real vehicle data. And often this type of noise can greatly affect the detection results of the algorithm. In addition, the weak internal short-circuit features of the battery are small and usually hidden in this type of noise and cannot be detected. Therefore, there is a need to retain the characteristics of battery failure while data noise reduction is an issue that must be solved. Also, data cleaning can improve data quality and reduce algorithmic false positives.

In this paper, we introduce a data smoothing method based on R-Lowess, which is very effective for outlier removal and preserves fault features. R-Lowess is proposed by Cleveland, which is based on the improvement of locally weighted regression and has strong robustness.[31,32]



https://doi.org/10.1021/acsomega.2c04991
ACS Omega 2022, 7, 40145−40162

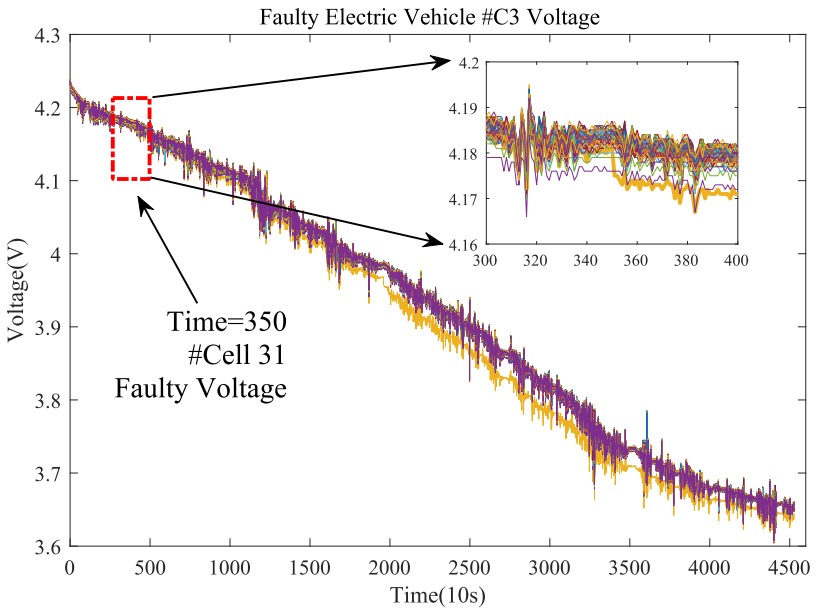



**Figure 3.** Faulty vehicle #C3 discharge process raw data.

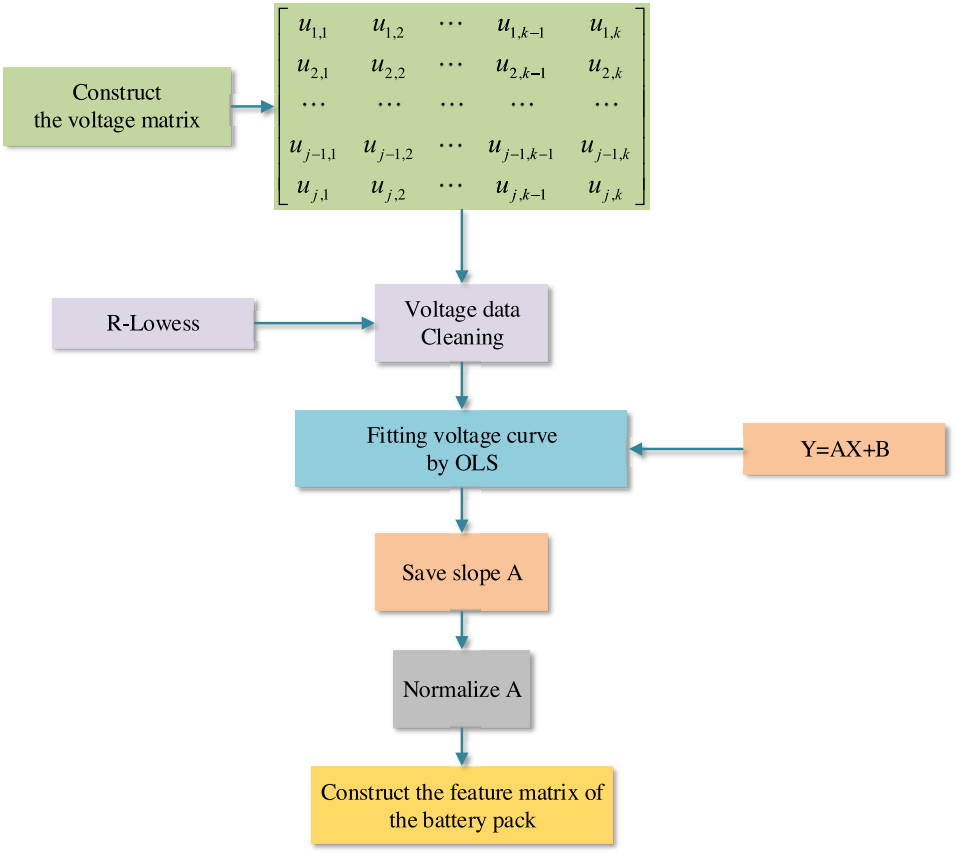



**Figure 4.** Discharge process feature extraction flow chart.

Strongly locally weighted regression is for each point $x_i$ on all data points within a window of $k$. The weights $\omega_k(x_i)$ are obtained from the weight function, and a polynomial fit $x_i$ is performed using the weighted least-squares method with the right values to obtain $y_i$. Then, the residual $\delta$ between the estimated and actual values is determined, and $\delta\omega_k(x_i)$ is used to estimate again to get the new fitted value. The procedure is repeated several times to obtain the final $\hat{y}_i$, which is the final smoothed value.

The weight function is defined as follows.

When we set the weight function to $W(c)$, the weight function generally is of $(o, p)$ type, and there are also other weight functions, which are not listed here



https://doi.org/10.1021/acsomega.2c04991
ACS Omega 2022, 7, 40145−40162

$$W(c) = \begin{cases} (1 - |c|^o)^p, & |c| \le 1 \\ 0, & |c| > 1 \end{cases} \quad (1)$$

where $c = e_i/6S$, $e_i$ denotes the residual of the $i$th voltage value fit and $S$ denotes the median value of the residual.

Using R-Lowess, voltage data with noise can be cleaned and the data quality is improved. R-Lowess has a weighting mechanism that effectively corrects the anomaly to a value similar to the surrounding points, thereby eliminating the noise voltage. When data collected due to external factors such as sensor errors is noisy data, the anomaly is transient and will return to normal values at subsequent sampling points, while when an internal battery failure occurs, the anomaly is persistent in character and will not return to normal in a short period of time. Therefore, when R-Lowess processes noisy data, a smaller weight is given to the noisy point to remove it since it is surrounded by normal. And if the subsequent data has a continuous voltage drop, the weight of that point will be enhanced and therefore retained. That is, it is insensitive to that fault point, which will reduce algorithm underreporting and improve the diagnostic accuracy of the algorithm.

**2.3. Fault Feature Extraction Based on OLS.** The fault characteristics of the voltage data collected by the big data platform are generally small and easily hidden among the normal fluctuations of the voltage, which is not conducive to the detection of fault diagnosis algorithms. Generally, when a fault occurs, such as a short circuit in a cell, the cell will show a decreasing trend in the voltage value compared to other normal cells. Therefore, using the OLS method to fit the cell voltage curve and extract this change trend feature can effectively highlight the fault characteristics in this cell. The least-squares method was proposed by French scientist Adrien-Marie Legendre in 1806 and is widely used in parameter estimation, optimization problems, and other fields.[33−35]

The feature extraction process is shown in Figure 4.

We assume that there are $n$ collected data $(x_1,y_1)$, $(x_2,y_2)$,⋯, $(x_{n-1}, y_{n-1})$,$(x_n,y_n)$, where $x$ denotes the independent variable of the function and $y$ denotes the dependent variable.

We substitute the above observed data points into the function to be fitted

$$y = \alpha_0 + \alpha_1 r_1(x) + \cdots + \alpha_m r_m(x) \quad (2)$$

The calculation gives

$$\hat{y}_i \approx \alpha_0 + \alpha_1 r_1(x_i) + \cdots + \alpha_m r_m(x_i), \quad i = 1, 2, 3, \cdots, n \quad (3)$$

The objective of OLS curve fitting is to determine the value of $\alpha_0$, $\alpha_1$, $\alpha_2$,⋯,$\alpha_n$ It is an optimization-seeking problem whose objective function is

$$\min_{\alpha_0,\alpha_1,\alpha_2,\cdots,\alpha_n} \sum_{i=1}^{k} (y_i - \hat{y}_i)^2 \quad (4)$$

This process of finding the solution to eq 4 is known as the OLS curve fitting of the function $y$.

The least-squares-based feature extraction process is as follows.

The voltage data in this process is first used to construct the voltage matrix $U'$

$$U' = \begin{bmatrix} u'_{1,1} & u'_{1,2} & \cdots & u'_{1,k-1} & u'_{1,k} \\ u'_{2,1} & u'_{2,2} & \cdots & u'_{2,k-1} & u'_{2,k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ u'_{j-1,1} & u'_{j-1,2} & \cdots & u'_{j-1,k-1} & u'_{j-1,k} \\ u'_{j,1} & u'_{j,2} & \cdots & u'_{j,k-1} & u'_{j,k} \end{bmatrix}$$

Data cleaning by the R-Lowess method in Section 2.2 is performed, $U'$, to form the new voltage matrix $U$

$$U = [v_1, v_2, \cdots, v_{k-1}, v_k] = \begin{bmatrix} u_{1,1} & u_{1,2} & \cdots & u_{1,k-1} & u_{1,k} \\ u_{2,1} & u_{2,2} & \cdots & u_{2,k-1} & u_{2,k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ u_{j-1,1} & u_{j-1,2} & \cdots & u_{j-1,k-1} & u_{j-1,k} \\ u_{j,1} & u_{j,2} & \cdots & u_{j,k-1} & u_{j,k} \end{bmatrix}$$

Curve fitting is performed in the voltage matrix $U$ using least squares for each discrete sequence of cell voltages $v$ with a sliding window $win$. To facilitate feature extraction and computational speed, this method utilizes linear function fitting

$$Y = AX + B \quad (5)$$

The slope value of each cell under each window is saved and the slope characteristic matrix $A$ is constructed

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,k-1} & a_{1,k} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,k-1} & a_{2,k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{j-win-1,1} & a_{j-win-1,2} & \cdots & a_{j-win-1,k-1} & a_{j-win-1,k} \\ a_{j-win,1} & a_{j-win,2} & \cdots & a_{j-win,k-1} & a_{j-win,k} \end{bmatrix}$$

Finally, the slope characteristic matrix $A$ is normalized to obtain the final feature matrix

$$\text{feature-}A = \text{normalize}(A) = \begin{bmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,k-1} & f_{1,k} \\ f_{2,1} & f_{2,2} & \cdots & f_{2,k-1} & f_{2,k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ f_{j-win-1,1} & f_{j-win-1,2} & \cdots & f_{j-win-1,k-1} & f_{j-win-1,k} \\ f_{j-win,1} & f_{j-win,2} & \cdots & f_{j-win,k-1} & f_{j-win,k} \end{bmatrix} \quad (6)$$

**2.4. Reference Cell Selection Method Based on K-Means Clustering.** The voltage data of electric vehicles in actual operation generally consist of batteries of the same type connected in series and have similar physical properties at the factory. However, in an actual application, every single cell exhibits inconsistent performance, such as inconsistent voltage, inconsistent SOH, etc., due to external temperature, internal temperature, internal resistance, material properties, process flow, and other factors.[36−38] As a result, the voltage data gathered from every single cell may also behave differently under normal conditions. Therefore, the inconsistency of each cell within the battery pack needs to be reduced before fault detection. Otherwise, the algorithm will show some false alarms due to inconsistency.



https://doi.org/10.1021/acsomega.2c04991
ACS Omega 2022, 7, 40145−40162

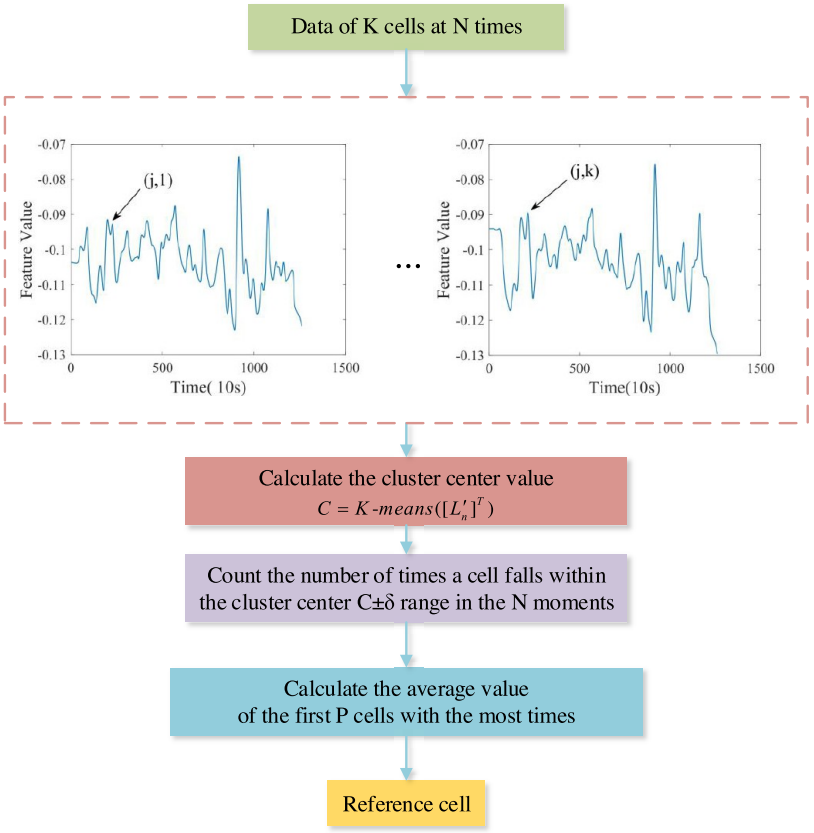



**Figure 5.** Flow chart of reference cell selection based on K-means.

To address this problem, this paper proposes a K-means clustering-based method to select reference cells in a battery pack to represent the current operation of the battery pack in the process. The specific process is as follows. The reference cell selection process is shown in Figure 5.

First, the feature data extracted by the current process are obtained, either for the first $n$ moments or for all sampling points of the current process. Now, we assume that with $k$ cells composing the power pack of an electric vehicle, a feature matrix is constructed using the feature data

$$L_n' = [l_1', l_2', \ldots, l_{k-1}', l_k']$$

where $l_k'$ denotes the feature data of the $k$th cell and $L_n'$ denotes the feature matrix consisting of $k$ cells with n sampling moments.

Second, the cluster center point $c_j$ among $k$ cells at the $j$th moment is found using the K-means clustering method

$$c_j = \mathrm{K} - \mathrm{means}([L_n']^T) \tag{7}$$

where $T$ denotes transposition.

The computed results of K-means clustering are constructed into a cluster center matrix $C$.

$$C = \begin{bmatrix} c_1 \\ c_2 \\ \cdots \\ c_{j-1} \\ c_j \end{bmatrix}$$

where $c_j$ denotes the cluster center value of the $k$th cell at the $j$th moment.

Finally, the cell that is closest to the cluster center value at each moment is calculated; so, first, it is necessary to set an interval radius $e$ and let the cluster center range matrix $C'$ be

$$C' = [C-e, C+e] = \begin{bmatrix} c_1 - e & c_1 + e \\ c_2 - e & c_2 + e \\ \cdots & \cdots \\ c_{j-1} - e & c_{j-1} + e \\ c_j - e & c_j + e \end{bmatrix}$$

In the feature matrix $L_n'$, the cells belonging to $C'$ at moment $j$ are found

$$L_n' \in C'$$

All eligible cell numbers $N$ at moment $j$ are recorded

$$N = [n_1, n_2, \cdots]$$

$$M = [m_1, m_2, \cdots, m_{k-1}, m_k]$$

where $M$ denotes the number of times the cell is counted in moment $j$. The cell number with the highest number of times in $M$ is selected and averaged, which is the reference cell $U_{\mathrm{re}}$ for this process. The formula is as follows

$$U_{\mathrm{re}} = \frac{\sum_{i=\max_i}^{p} l_i'}{p} \tag{8}$$

### 2.5. Fault Detection Based on the Fréchet Algorithm.

Fréchet's algorithm is a metric for path space similarity proposed by Maurice René Fréchet, a French mathematician, in 1906.[39] It is now commonly used in other fields such as railroad turnout



https://doi.org/10.1021/acsomega.2c04991
ACS Omega 2022, 7, 40145−40162

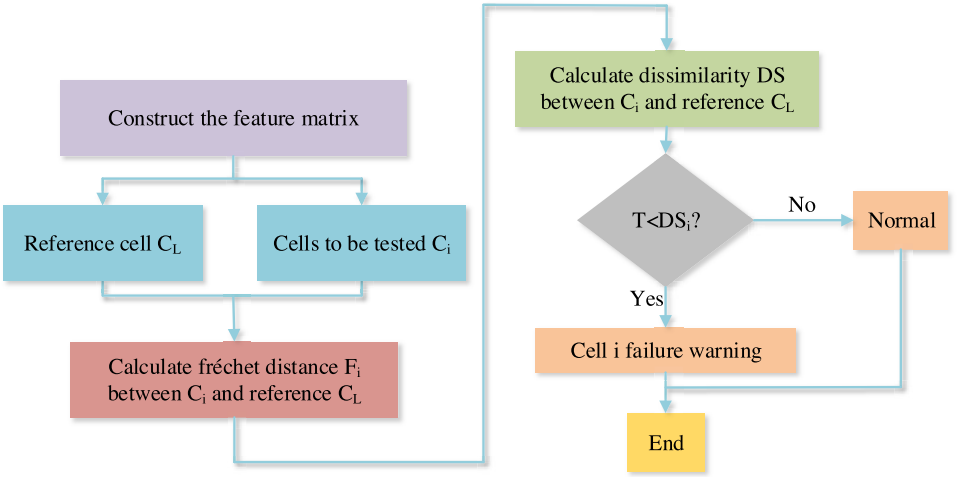



**Figure 6.** Flow chart of fault diagnosis based on the Fréchet algorithm.

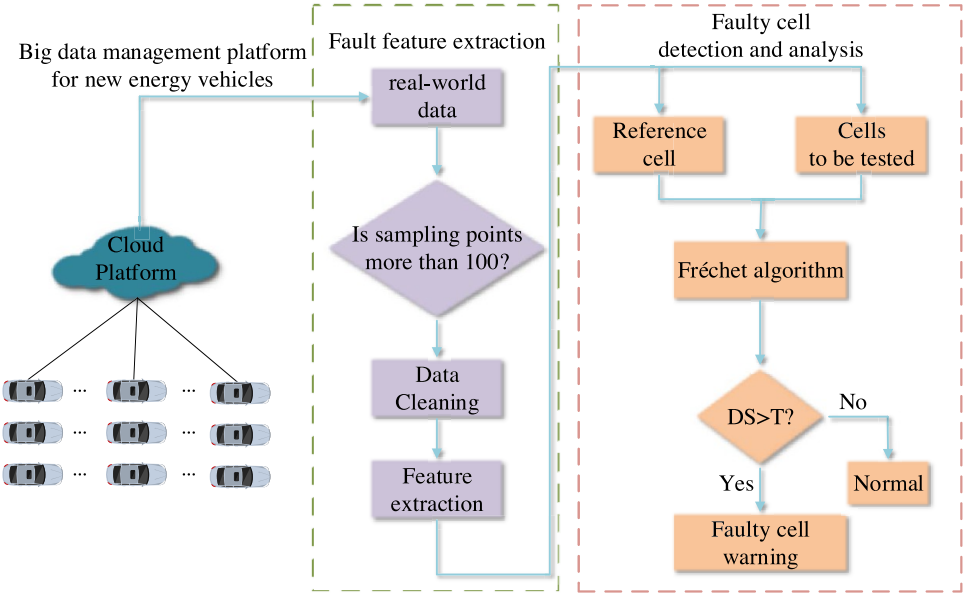



**Figure 7.** Overall flow chart of fault diagnosis.

fault detection and transformer fault detection.[40−42] Fréchet's algorithm extends the realistic distance problem in general sets and especially provides a good calculation method for similarity measures between two variables. After the feature extraction in Section 2.1−2.4, each normal cell of EV battery pack has similar changes in general. If there is a cell with malfunction, it will show different change features from other cells. Therefore, the introduction of the Fréchet algorithm can identify the cell well.

The flow of calculating the dissimilarity is shown in Figure 6.

Fréchet's mathematical definition is as follows:

Supposing that $L_1$ and $L_2$ are two continuous curves on the space $S$, i.e., $L_1$:[0,1] → $S$ and $L_2$:[0,1] → $S$; and set reparameterized functions $\alpha$ and $\beta$ in the unit interval, and $\alpha$: [0,1] → [0,1], $\beta$:[0,1] → [0,1]; then, the Fréchet distance $F(L_1, L_2)$ between two continuous curves $L_1$ and $L_2$ is defined as

$$F(L_1, L_2) = \inf_{\alpha,\beta} \max\{d(L_1(\alpha(t)), L_2(\beta(t)))\}, \quad t \in [0, 1] \tag{9}$$

where $d$ is the metric function, and the metric functions are Gaussian affiliation function, weighted Euclidean distance, etc.

According to the above definition, this fault diagnosis method uses the discrete Fréchet algorithm. This method is similar to the continuous Fréchet distance calculation method and can be used to calculate the dissimilarity between discrete reference cells and discrete cells to be detected. And the method is easy to be implemented on a computer. The specific procedure is as follows.

Supposing the reference cell $C_L$ for this discharge process is found to be

$$C_L = U_{re} = \begin{bmatrix} u_1 \\ u_2 \\ \cdots \\ u_{j-1} \\ u_{j-1} \end{bmatrix}$$

The cell to be detected is

$$C_2 = U_{detect} = [v_1, v_2, \cdots, v_{k-1}, v_k]$$



https://doi.org/10.1021/acsomega.2c04991
ACS Omega 2022, 7, 40145−40162

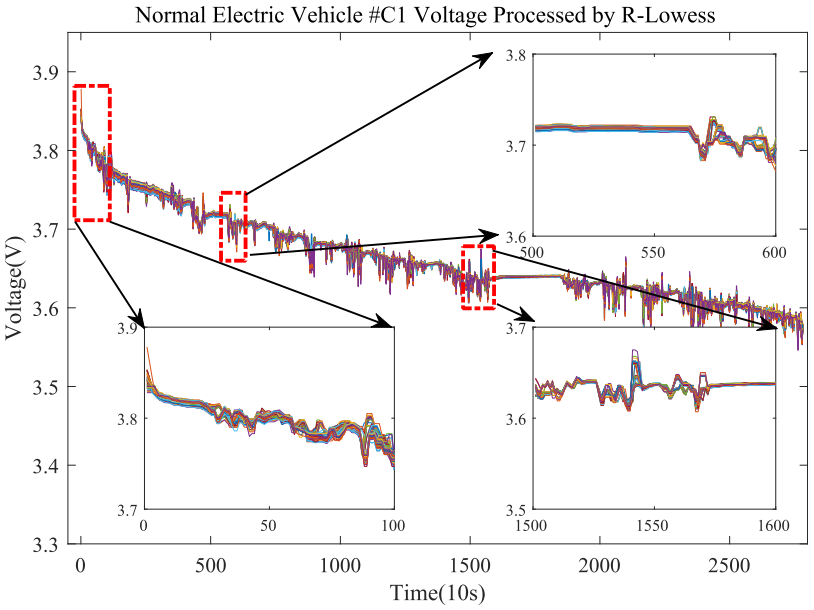



**Figure 8.** Normal vehicle #C1 data cleaning result by R-Lowess.

where $v_k$ denotes the feature value of the $k$th cell and the sampling point is all of the data of the current process.

Using the definition of the Fréchet distance, the Fréchet distance $F(C_L, C_2)$ can be calculated for the reference cell $C_L$ and the detection cell $C_2$

Then, we calculate the dissimilarity

$$\mathrm{DS} = F(C_L, C_2) \tag{10}$$

We set the threshold $T$. If the dissimilarity $\mathrm{DS}_i$ is greater than the threshold $T$, the cell is considered as a faulty cell. On the contrary, the current cell is considered as a normal cell

$$T = 1.3^{*}(\mathrm{mean}(\mathrm{DS}) + 3^{*}\mathrm{std}(\mathrm{DS})) \tag{11}$$

where mean is the mean of DS of all cells at the current moment and std is the standard deviation of DS of all cells at the current moment. That is, $T$ is 1.3 times the $3\sigma$ principle.

Typically, training data from actual vehicle operations are used to threshold set, which have been tagged in detail with the time of failure and the number of the faulty unit. Adaptive thresholds are then set based on 3-$\sigma$, and margins are considered so that faulty units are alarmed and normal units are not. Based on the adjusted thresholds, it is applied to another vehicle of the same model for testing to determine if it is effective.

**2.6. General Flow Chart of Fault Diagnosis.** Battery failure of electric vehicles still affects users. And serious battery failure can lead to thermal runaway, which eventually triggers spontaneous combustion and brings incalculable harm to people. This fault diagnosis method is implemented based on K-means clustering and the Fréchet algorithm. And the overall flow of fault diagnosis is shown in Figure 7.

The purpose is tantamount to diagnosing the battery of tiny faults and achieving early warning in the current calculation process. The overall steps are presented as follows.

**Step 1:** To judge the calculation conditions of the discharge process data collected in the big data platform for electric vehicles. If the data points collected in the current discharge process exceed 100, it meets the calculation requirements and goes to the next step. If the data points collected in the current discharge process are less than 100 points, the discharge process sampling is invalid and cannot be used for fault diagnosis.

**Step 2:** The data of the discharge process is cleaned. As can be seen from Section 2.2, the voltage data collected by the sensor have a large amount of noise, which can greatly affect the judgment result of the fault diagnosis algorithm. Therefore, this method introduces an R-Lowess algorithm for noise reduction of the collected voltage data. R-Lowess can effectively remove the anomalies caused by sensor errors and can retain the fault variation characteristics of the battery.

**Step 3:** In the early stage, it is difficult to achieve an early warning and easy to miss the alarm if using the collected raw voltage data for fault diagnosis because of the small change characteristics of the battery voltage where minor faults occur. Therefore, this method introduces an OLS-based feature extraction method to extract the potential features of the voltage data. The extracted features can effectively highlight the evolution of battery faults compared to other normal battery variation features. And it can achieve early warning in the current calculation process.

**Step 4:** A reference cell selection method based on K-means clustering is developed to reduce the false positives of the fault diagnosis algorithm. It is able to effectively reduce the impact of inconsistency between cells and can decrease the false alarm rate of the algorithm. This method needs to be able to reduce false positives compared to other literature with neighboring cells.

**Step 5:** A fault diagnosis method is proposed based on the Fréchet algorithm. Utilizing the extracted features, the dissimilarity of each cell to the reference cell is calculated. The result is constructed as the dissimilarity matrix of each cell at each moment. Threshold value $T$ is set based on the $3\delta$ principle. If a cell in the dissimilarity matrix exceeds the threshold value $T$ at a certain moment, the algorithm alarms. A fault occurs in that cell. If no cell in the dissimilarity matrix of the current process exceeds the threshold $T$, the algorithm returns and proceeds to the calculation of the next discharge process.



https://doi.org/10.1021/acsomega.2c04991
ACS Omega 2022, 7, 40145−40162

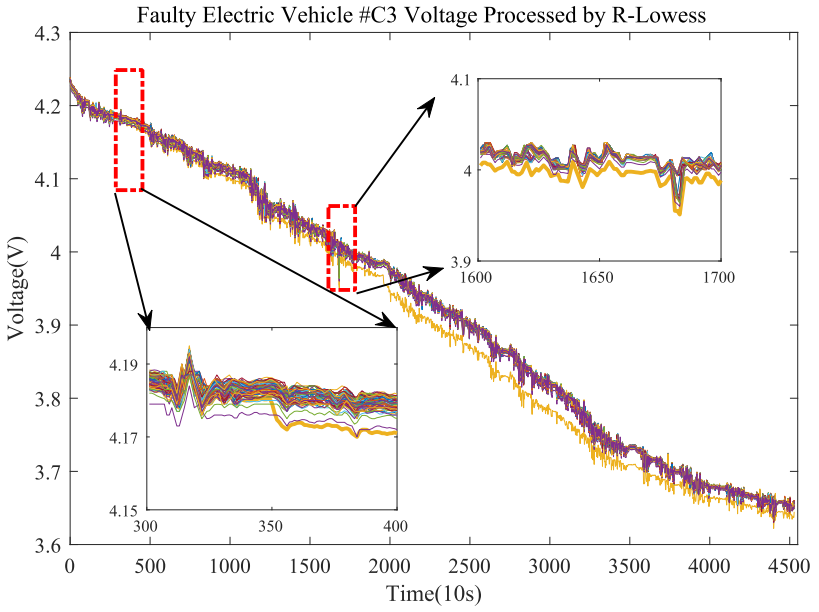



**Figure 9.** Faulty vehicle #C3 data cleaning result by R-Lowess.

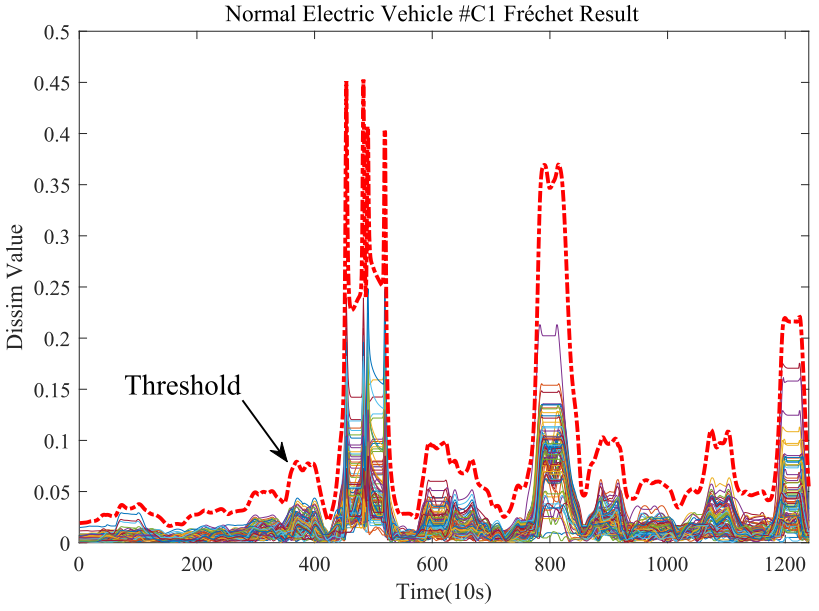



**Figure 10.** Normal vehicle #C1 dissimilarity result by the Fréchet algorithm.

## 3. RESULTS AND DISCUSSION

Faults such as extrusion, loose connection, internal short circuit, etc. generally exist in the battery pack. And the battery fault diagnosis contains fault cell number, fault type, fault cause, etc. However, more accurate models and more specialized technical support are needed for the analysis of the specific causes of battery failure. In this paper, only voltage data are considered, so it is not the scope of this paper to determine the cause of the fault. The focus of fault diagnosis in this paper is to achieve early warning of a minor fault in the discharge process. To verify the effectiveness of the method proposed in this paper, three real-world vehicles are used for verification illustration.

**3.1. Data Cleaning Analysis.** Figure 8 shows the voltage data graph of each cell of normal vehicle #C1 after data cleaning by R-Lowess. According to the original voltage graph, when the data of this discharge process began to be collected, a voltage drop occurred at #Cell 1−23 due to the abnormality of the sensor, which greatly affected the data quality of the voltage. After the R-Lowess process, the voltage drop at the 10th sampling point has been removed. And R-Lowess also removes some burrs in the data collection during the discharge process, which improves the data quality and is useful to the fault diagnosis results. Also, it is found from Figure 8 that subsequent points of normal voltage fluctuations, such as the 560−600th and 1500−1580th sampling points, are able to be retained intact. These sampling points belong to normal fluctuations and are not abnormal data. Similarly, in other parts of the voltage data, the normal voltage fluctuation feature is retained by the R-Lowess method, and no further examples will be given.

As shown in Figure 9, the fault occurred in vehicle #C3. The vehicle is used as an example in this section to illustrate that the noise reduction algorithm is not very sensitive to the changing



https://doi.org/10.1021/acsomega.2c04991
ACS Omega 2022, 7, 40145−40162

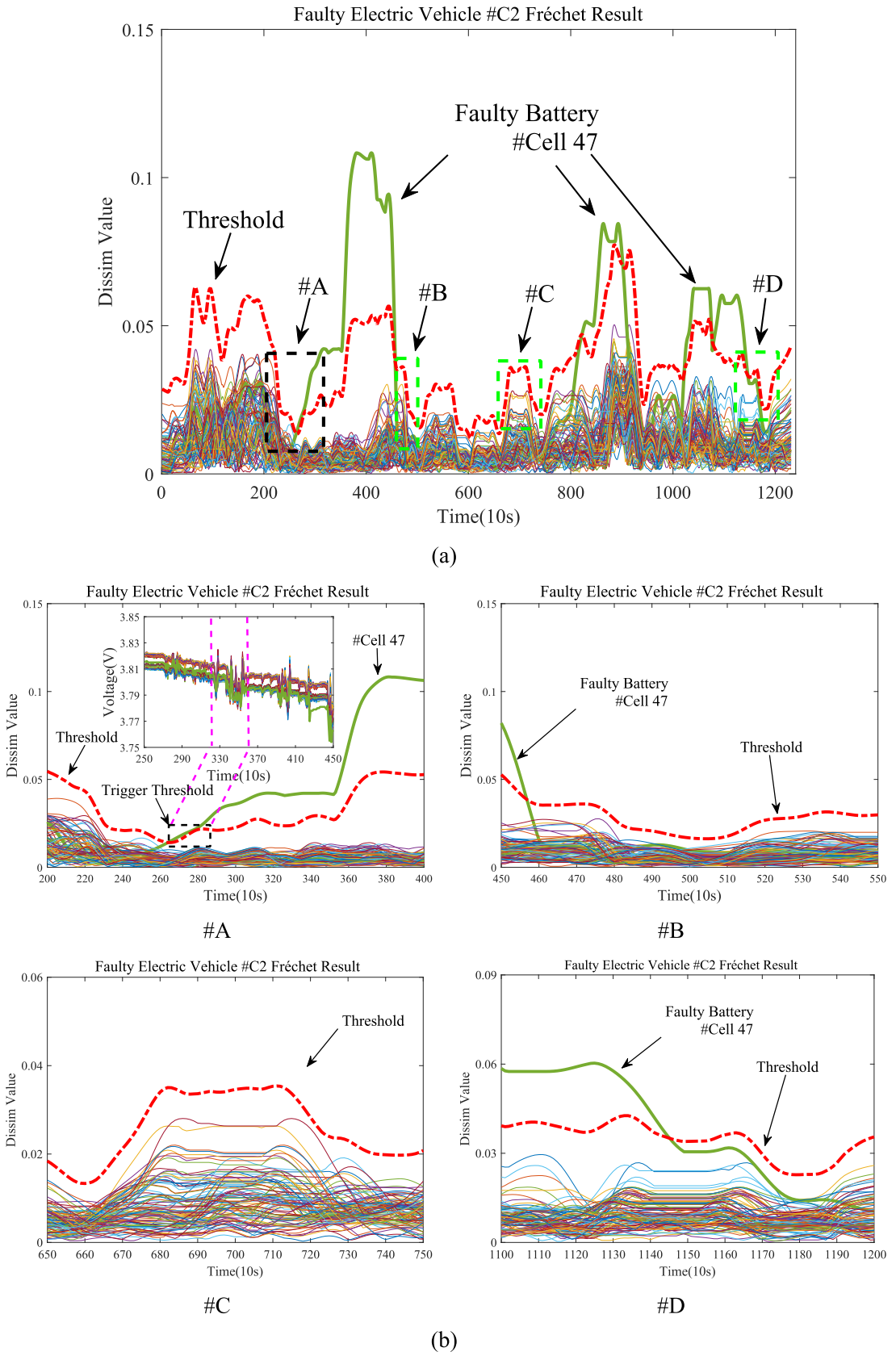



**Figure 11.** Faulty vehicle #C2 dissimilarity result by the Fréchet algorithm: (a) calculation result of the whole process of the faulty vehicle #C2 and (b) detailed descriptions at #A, #B, #C, and #D.

features of the fault monolith. The R-Lowess method is able to retain the slight fault features intact and will only remove outliers that are anomalous due to the sensors. In the faulty vehicle #C3, #Cell 31 shows a weak internal short-circuit fault at the 351st



https://doi.org/10.1021/acsomega.2c04991
ACS Omega 2022, 7, 40145−40162

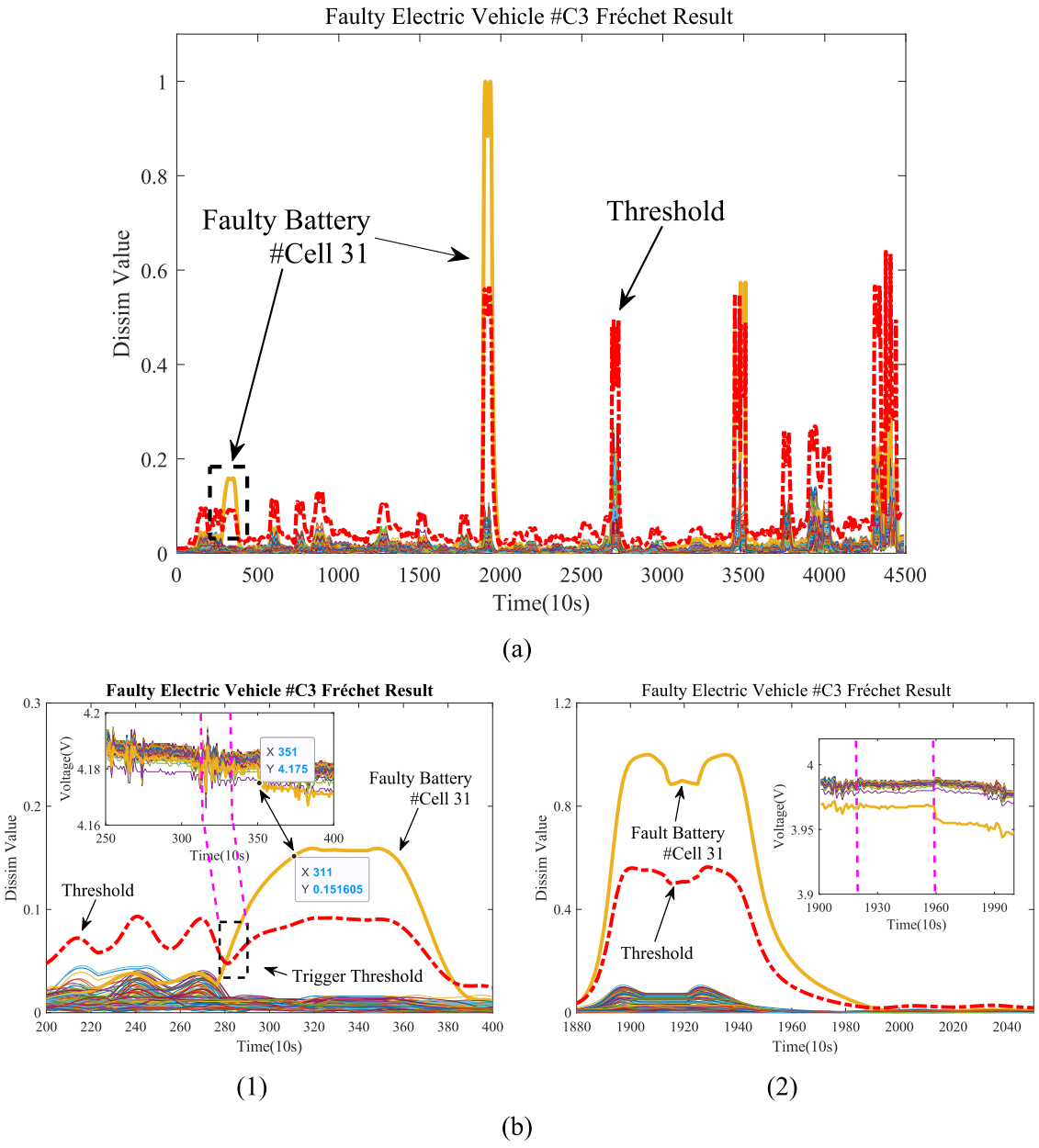



**Figure 12.** Faulty vehicle #C3 dissimilarity result by the Fréchet algorithm: (a) calculation result of the whole process of the faulty vehicle #C3 and (b) detailed description of sampling points 200−400 and 1880−2450 in panel (a).

sampling point and the voltage shows a decreasing trend. However, due to the cell equalization mechanism, this cell voltage drops by 0.01 V before the cell voltage stabilizes again. As seen at the 351st sampling point, this fault feature was not removed by the R-Lowess algorithm. Also, at the 1680th sampling point, some cells in the battery pack show a voltage fluctuation at the same time. This noise reduction algorithm is equally insensitive to this normal fluctuation feature and retains the fluctuation feature. And the fault feature is also retained for #Cell 31 where the fault occurred earlier. During subsequent #Cell 31 monomer discharge, the evolution of the internal short circuit intensifies and R-Lowess is able to retain the fault signature, which is no longer described here as an example. Furthermore, it is well known from the original voltage plot that at the 3600th sampling point, there is an abrupt rise in the monomer voltage due to a sensor acquisition error, a feature that is removed from the voltage data after R-Lowess processing. After the R-Lowess processing, the data quality of the discharge process was considerably improved.

In summary, through the analysis of one normal vehicle and one faulty vehicle, it can be proved that the noise reduction algorithm proposed in this paper can effectively remove the data collection errors caused by sensors, while preserving the evolution features of the battery unit when it happens to be faulty.

**3.2. Fault Diagnosis Analysis.** The dissimilarity detection results of the Fréchet algorithm for normal vehicle #C1 are shown in Figure 10. From Figure 10, it can be seen that the feature data extracted by the process are input into the Fréchet algorithm, and no false alarms or missed alarms occur in all sampling points of normal vehicle #C1. In Figure 10, the red dashed line is the threshold $T$. Figure 10 also shows that normal cells are basically concentrated in a small range at the same sampling point, and individual cells may deviate from the


https://doi.org/10.1021/acsomega.2c04991
ACS Omega 2022, 7, 40145−40162

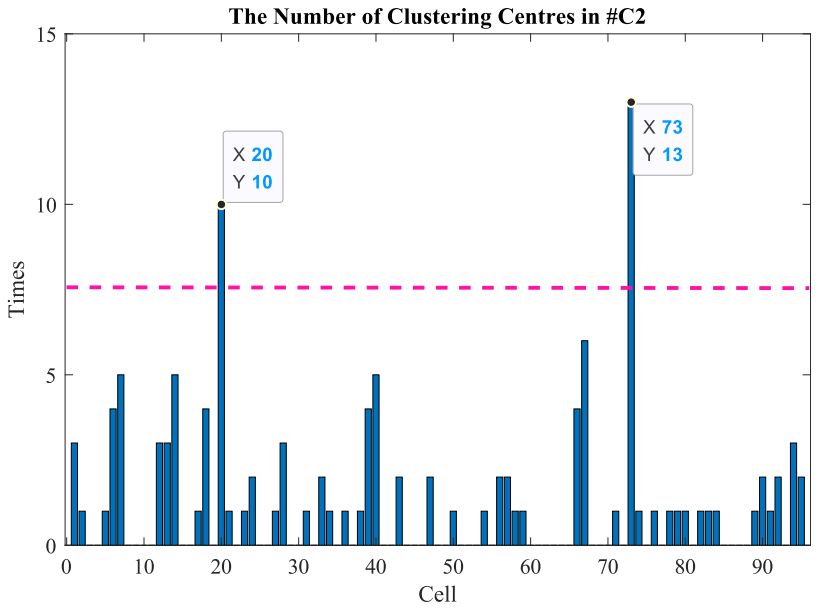



**Figure 13.** Cluster center distribution of faulty vehicle #C2 in the first 100 sampling points.

concentration range of other cells due to factors such as poor consistency. But the cell does not exceed the threshold $T$ set in this paper. It also proves that the threshold setting with a certain margin considered in this method is reasonable.

Figure 11 shows the result of the dissimilarity calculation of the Fréchet algorithm for the faulty vehicle #C2. The voltage data of this discharge process is input into this method, and #Cell 47 shows a slight voltage drop at the 425th sampling point. This method provides an early warning for this single unit in the calculation of this process. In Figure 11(a), the red dashed line is the threshold line, and the green bolded solid line is the result of the dissimilarity calculation of faulty #Cell 47 in this method. After the extraction process of the new features, Figure 11(b) #A shows that #Cell 47 has exceeded the threshold at the 280th (corresponds to the 320th sampling point of the voltage data due to the sliding window setting as 40) sampling point and the algorithm warns. In contrast, in the original voltage data, this sudden voltage drop feature occurs at the 425th sampling moment, and the algorithm in this paper warns almost 100 sampling points earlier. For the other normal cells, no false alarm occurred during the whole calculation process. Figure 11(b) #B−#D also shows that the 460th, 720th, and 1170th sampling points were also just close to the threshold value and did not exceed it, which were normal cells. The focus of this paper is to achieve early warning of minor faults, and the diagnosis of faults for which the voltage difference has become very obvious at a later stage is not a focus of this paper. To distinguish the fault becoming serious, this algorithm does not alarm the voltage difference becoming largely due to short-circuit fault at a later stage, but only the sudden voltage change due to the intensification of the fault.

Figure 12 shows the result of the dissimilarity calculation of the Fréchet algorithm for the faulty vehicle #C3. The evolution of the fault in the early stage of this vehicle is more insignificant compared to the faulty vehicle #C2. The innovative features of the faulty vehicle #C3 are input into this diagnostic method, and the weak voltage drop due to the faulty cell can also be warned in advance after the calculation of this method. In Figure 12, the red dashed line is the threshold line of the process and the thickened yellow line is the fault #Cell 31 of this faulty vehicle. It can be concluded from Figure 12(b)(1) that this method has achieved the warning at the 290th sampling point (corresponding voltage value of 330th sampling point) before the cell voltage surge point occurs, while the voltage surge point is at the 351st moment and this method has advanced the warning by 31 sampling points. After the 320th sampling point, a weak internal short circuit occurs in the cell causing a voltage drop. Since there is an equalization mechanism within the battery pack, it slows down the voltage drop trend of the single faulty cell #Cell 31. For this phenomenon, the algorithm does not warn about this tiny internal short circuit anymore since it was already warned at the 320th sampling moment earlier. However, #Cell 31 shows a sudden voltage drop again at the subsequent 1960th sample point, and the method in this paper performs the fault warning again at the 1930th moment and 30 samples earlier with this sudden drop point in Figure 12(b)(2). This is due to the fact that #Cell 31 had already experienced a fault in the previous period, and fault evolution became severe with that sampling point. Fault expansion makes the dissimilarity results calculated by this algorithm very prominent, which also verifies the effectiveness of this method for detecting sudden voltage changes.

In summary, the effectiveness and feasibility of the fault diagnosis algorithm proposed in this paper are verified using one normal vehicle and two faulty vehicles. After the current process of discharge data is subjected to data noise reduction, new feature extraction, and dissimilarity calculation of the proposed method in this paper, it can amplify the minor fault features and can give early warning to the battery cells long before the sudden voltage drop.

**3.3. Comparative Experimental Analysis Based on Reference Cells.** When calculating the dissimilarity of adjacent cells, it was found that individual cells would be affected by inconsistency and the algorithm would calculate abnormal dissimilarity, which is prone to false positives. Therefore, in this paper, K-means clustering is used to calculate the ideal cell that is an equivalent substitute for this process as a reference cell.

Figure 13 shows the cell calculation for the faulty vehicle #C2. In the algorithm setup, the average of the cell voltage of the first two highest counts is chosen as the ideal cell for this process. As



https://doi.org/10.1021/acsomega.2c04991
ACS Omega 2022, 7, 40145−40162

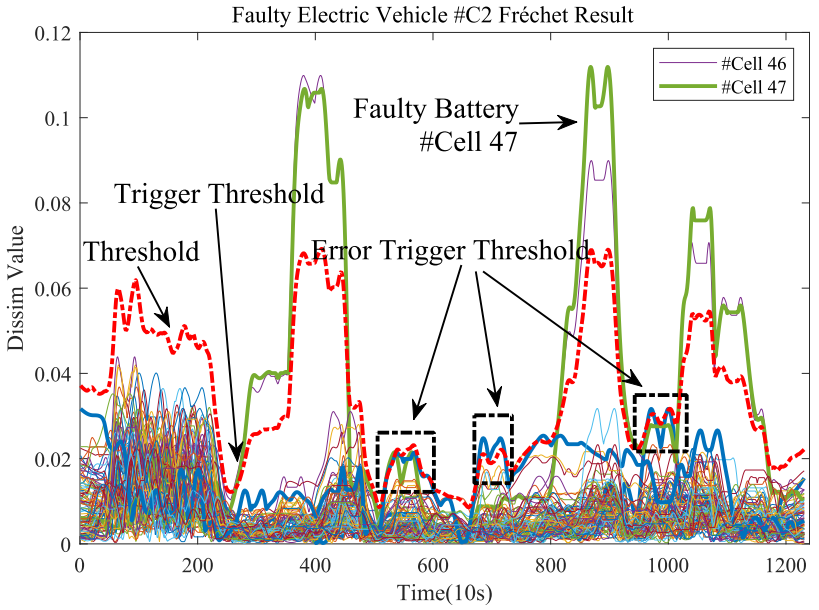



**Figure 14.** Faulty vehicle #C2 using adjacent cell dissimilarity results by the Fréchet algorithm.

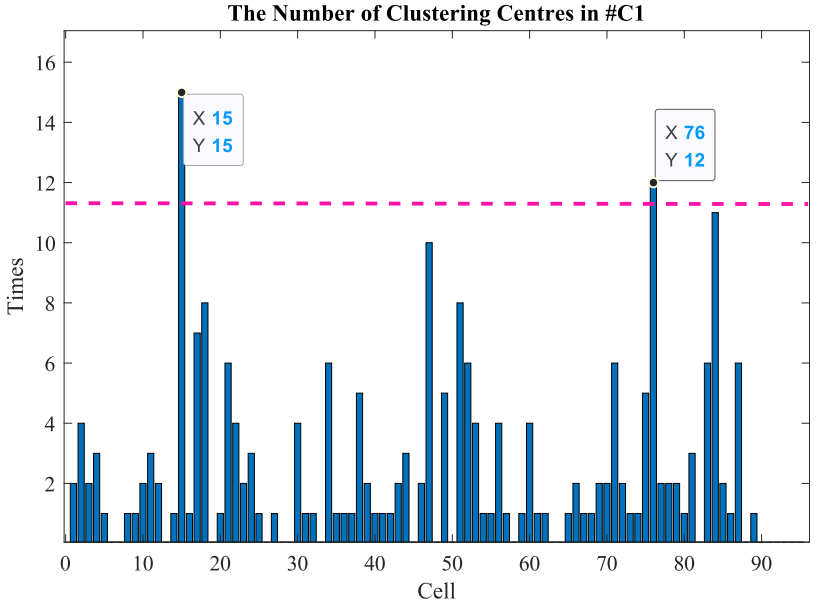



**Figure 15.** Cluster center distribution of normal vehicle #C1 in the first 100 sampling points.

can be seen from Figure 13, the first two highest counts of cells for the faulty vehicle #C2 are #Cell 73, which appears within the cluster center 13 times in the first hundred moments, and #Cell 20, which appears within the cluster center 10 times in the first hundred moments. Therefore, the reference cell voltage value for the faulty vehicle #C2 during this discharge is the average voltage of #Cell 73 and #Cell 20. This algorithm is able to reduce false alarms by selecting only the first 100 moments and the first two highest counts of cells. In practical applications, this parameter can be adjusted according to the actual situation.

Figure 14 shows the results calculated for the faulty vehicle #C2 without the use of a reference cell. For comparison with the results using the reference cell, the input is also the features extracted for this discharge process, but the dissimilarity results are between neighboring cells. In Figure 14, the red dashed line is the threshold $T$. The formula for setting the $T$ value is the same as before. The bold solid green line is fault #Cell 47. And the solid purple line is the #Cell 46 adjacent to fault #Cell 47. The purple curve is the result of dissimilarity between the single #Cell 46 and #Cell 47, and the bold solid green line is the result of dissimilarity between the fault #Cell 47 and #Cell 48. Therefore, according to the calculation logic, the faulty cell can be located at #Cell 47. Although calculating the dissimilarity between neighboring cells is able to detect the faulty cells when a fault occurs, it can also be noticed from Figure 14 that at the 550th, 700th, and 950th moments of the process, some cells exceeded the threshold $T$ and the algorithm had a false alarm.

Compared with the calculation results of the Fréchet dissimilarity without using the reference cells, the fault detection processed by the K-means clustering method, in Figure 11, proposed in this paper, can reduce the influence of inconsistency of each cell. The fault diagnosis algorithm did not have fault false



https://doi.org/10.1021/acsomega.2c04991
ACS Omega 2022, 7, 40145−40162

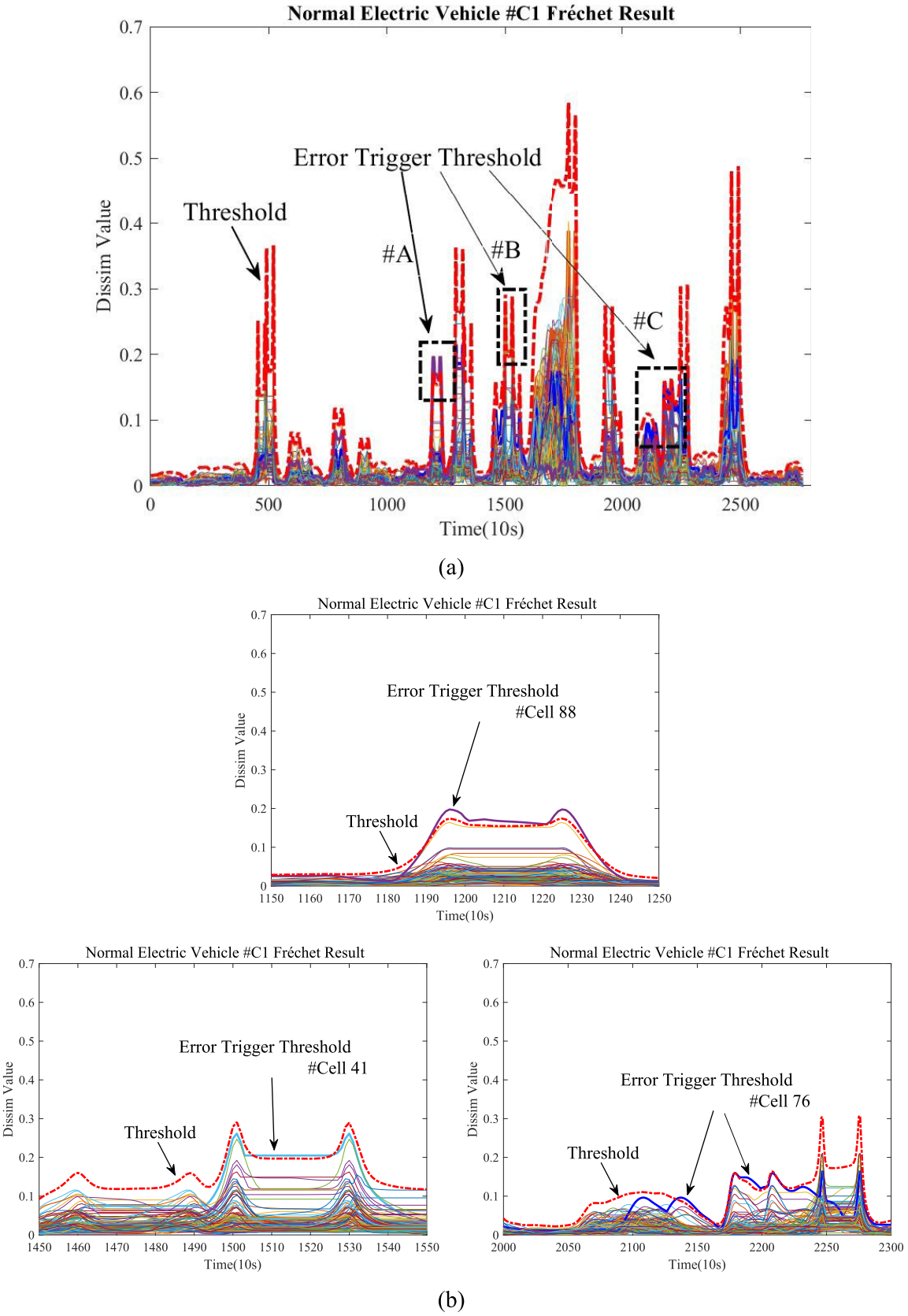



**Figure 16.** Normal vehicle #C1 using adjacent cell dissimilarity results by the Fréchet algorithm: (a) calculation result of the whole process of normal vehicle #C1 and (b) detailed description of #A, #B, and #C in panel (a).



https://doi.org/10.1021/acsomega.2c04991
*ACS Omega* 2022, 7, 40145−40162

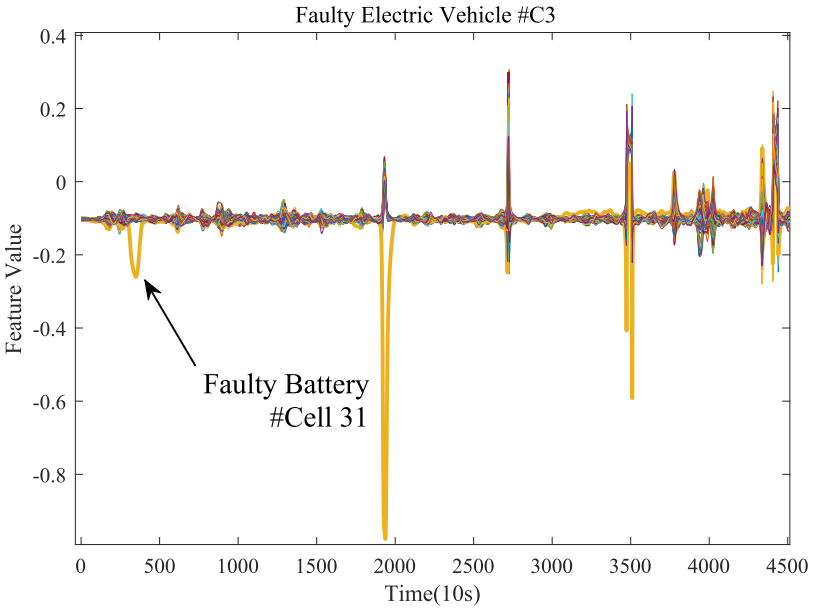



**Figure 17.** Faulty vehicle #C3 features extracted based on the OLS method.

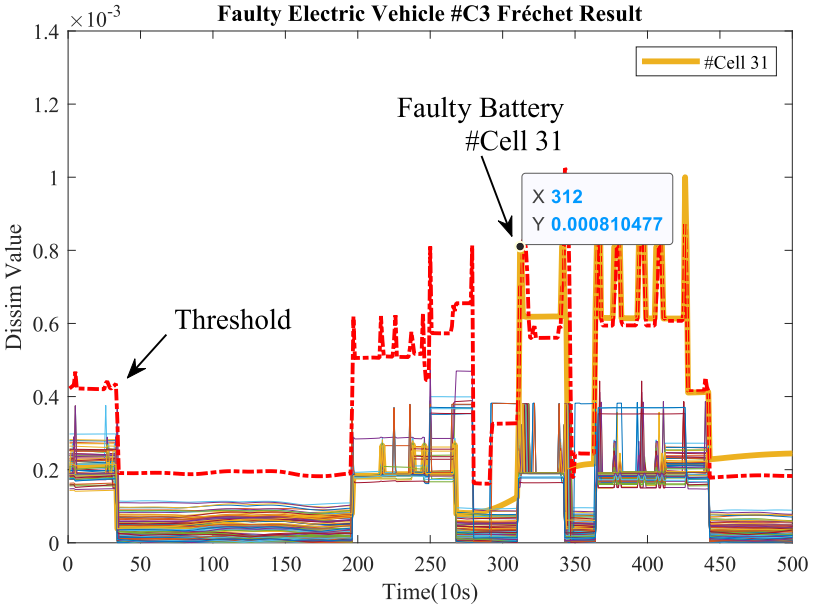



**Figure 18.** Faulty vehicle #C3 dissimilarity results using literature features.

alarms throughout the calculation process, and only the faulty cells were warned, which proved the effectiveness of this method.

In a similar way, Figure 15 shows the distribution of the number of times each cell in the normal vehicle #C1 falls within the cluster center. The two cells with the highest number of occurrences in the first hundred moments were also selected for this vehicle. As seen in Figure 15, 95 cells of the process fell within the cluster center range for 15 of the first hundred moments for #Cell 15 and for 12 of the first hundred moments for #Cell 76. Therefore, the reference cell for normal vehicle #C1 during this discharge is the average of the voltage between #Cell 15 and #Cell 76.

Figure 16(a) shows the dissimilarity results calculated for normal vehicle #C1 using neighboring cells. The red dashed line is the threshold $T$. The threshold $T$ is established in the same way as in the previous section. In Figure 16b, without using the reference cells selected by our method, the algorithm shows a false alarm for #Cell 88 at the sampling point around 1200, a false alarm for #Cell 41 at the sampling point around 1500, and a false alarm for #Cell 41 at the sampling point around 2200. And from Figure 10, it can be found that the results of Fréchet dissimilarity calculated using the reference cells selected by our method are more satisfactory. Our proposed fault diagnosis algorithm does not show any false alarms while calculating the whole process.

In summary, it can be demonstrated by a faulty vehicle and a normal vehicle that using the adjacent cell calculation method will lead to many false alarms due to the inconsistency between cells and difficulty in setting the threshold value. However, after selecting the reference cells using our method, the performance of the algorithm is improved, the false alarm rate is greatly



https://doi.org/10.1021/acsomega.2c04991
ACS Omega 2022, 7, 40145−40162

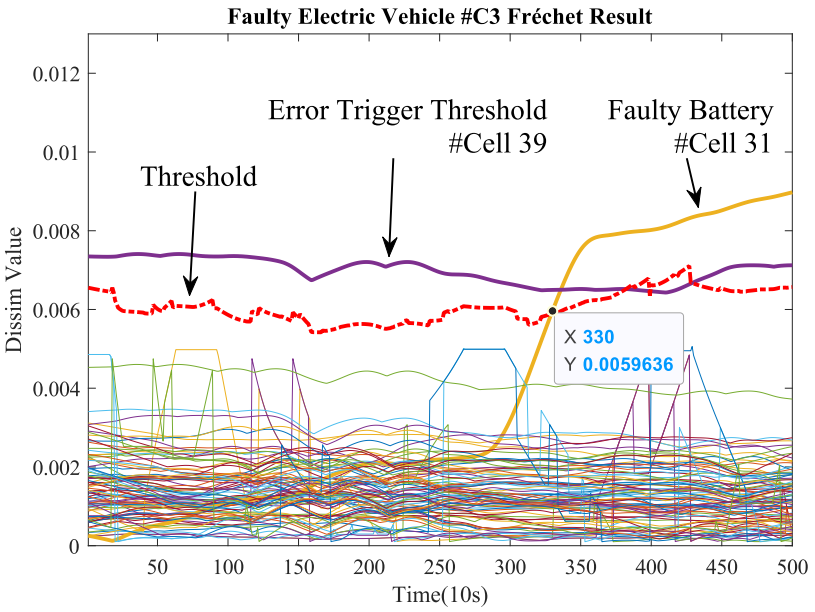



**Figure 19.** Faulty vehicle #C3 dissimilarity results using original voltage data.

reduced, and it is able to warn accurately. In this calculation, the selection of two cells is generally sufficient to achieve the effect in reducing the inconsistent influence. In other cases with different applications, voltage averages for more cells at more moments can be selected depending on data quality, etc. In addition, the use of this method mitigates the effect of severe deviations in the voltage of the faulty cell on the reference cell when averaging the voltages on all cells.

**3.4. Comparative Experimental Analysis Based on Feature Extraction.** Figure 17 shows the feature values extracted by the faulty vehicle #C3 through the method of this paper. It can be observed that the feature extraction through the method in this paper can amplify the small voltage drop feature, which is beneficial to the detection of the fault diagnosis algorithm. In addition, the study in this paper focuses on the prewarning of minor faults, but the characteristics of the late faults where the voltage difference becomes large are not studied. Therefore, this method does not diagnose for the late voltage value that continues to deviate significantly from the normal cell voltage. The sampling point around 1900 in Figure 17, where the fault characteristics appear to change markedly, is the result of the short-circuit fault aggravation of this single cell and the coupling effect of the cell equalization mechanism. This phenomenon has been analyzed and described in Section 3.2. Further, in the actual vehicle, we found that cells with large voltage differences may sometimes be caused by cellular inconsistency and do not form part of the cell generation fault. Subsequent experiments will be illustrated.

Figure 18 shows the results of fault diagnosis using the normalized cell voltage features from ref 13 and only the first five hundred moments are plotted to facilitate a comparison of the detection in the initial occurrence of a fault. For the fairness of the comparison, only the feature values are swapped in the comparison test. The other conditions are the same, such as data cleaning, sliding window size, threshold setting, etc. Figure 18 shows this diagnosis results in an alarm for the single #Cell 31 only at the 312th moment (corresponding to the 352nd moment of voltage data), while the feature extracted in this paper is a warning signal at the 280th moment (for the 320th moment of voltage data), 32 sampling points ahead. By comparing with the normalized cell voltage features in ref 13, the effectiveness of the proposed features in this paper for the amplification and early warning of minor fault features is shown by Figure 12 in Section 3.2.

Figure 19 shows the results of the dissimilarity calculated from the voltage values after data noise reduction. For ease of comparison, only the first five hundred moments containing the fault features are plotted. Figure 19 shows that in the first five hundred of the faulty vehicle #C3, although it was possible to locate the fault at the 330th sampling point (corresponding to the 370th moment of the voltage value), the sudden voltage drop was at the 351st moment and did not complete the early warning. Compared with the feature extraction method set out in the present paper, our method warned at the 280th sampling point (corresponding to the 320th moment of voltage data), which was 31 points faster than using the original voltage data. Also, the solid purple line #Cell 39 appears to be a false alarm. This is due to the poor consistency of this cell compared to other cells, and there is always a certain voltage difference with the voltage values of other cells, which can be clearly found in the purple curve cell from the original voltage data in Figure 3. However, the features proposed in this paper did not produce false alarms for this cell in Figure 12 and the OLS-based features extracted in this paper were effective in highlighting the fault features for early warning.

In summary, compared with using the normalized cell voltage features of ref 13 and using only the original voltage, the features extracted in this paper can be more effective in reducing false alarms, amplifying minor fault features, and achieving early warning.

## 4. CONCLUSIONS

This paper proposes a fault diagnosis method based on K-means and the Fréchet algorithm, which can give early warning to cells where weak faults occur through potential feature extraction of voltage data in a discharge process. The data smoothing method based on R-Lowess proposed in this paper can effectively remove the noisy data caused by sensors, while it is insensitive to the general weak fault features. Compared with the input



https://doi.org/10.1021/acsomega.2c04991
ACS Omega 2022, 7, 40145−40162

original voltage data, the fault feature extraction method based on the OLS method proposed in this paper can achieve early fault diagnosis and can amplify the small fault features for easy detection. A reference cell selection method based on K-means clustering is proposed to decrease the influence of inconsistency between cells on the algorithm results. Compared with the general calculation between adjacent cells, the reference cells selected in this paper can effectively mitigate the algorithm false positives caused by cell inconsistency and greatly improve the accuracy of the fault diagnosis algorithm. And based on the extracted new fault features, the Fréchet algorithm is proposed for fault diagnosis and localization. Through the actual operation of electric vehicle data verification, the method can also be applied in the field of electric vehicle battery fault diagnosis and can effectively identify the fault cells with high accuracy.

In addition, due to the lack of fault data, the stability of this algorithm can be verified on more faulty vehicles in the ensuing research. Or some data enhancement methods are introduced to increase the faulty vehicle data.

## AUTHOR INFORMATION


### Corresponding Author

**Fan Zhang** − *Hubei Collaborative Innovation Center for High-efficiency Utilization of Solar Energy, Hubei University of Technology, Wuhan 430068, P. R. China; Hubei Engineering Research Center for Safety Monitoring of New Energy and Power Grid Equipment, Hubei University of Technology, Wuhan 430068, P. R. China*; Email: joyce_zhang@hbut.edu.cn

### Authors

**Minghu Wu** − *Hubei Collaborative Innovation Center for High-efficiency Utilization of Solar Energy, Hubei University of Technology, Wuhan 430068, P. R. China; Hubei Engineering Research Center for Safety Monitoring of New Energy and Power Grid Equipment, Hubei University of Technology, Wuhan 430068, P. R. China*

**Wanyin Du** − *Hubei Collaborative Innovation Center for High-efficiency Utilization of Solar Energy, Hubei University of Technology, Wuhan 430068, P. R. China*; orcid.org/0000-0002-5762-7790

**Nan Zhao** − *Hubei Collaborative Innovation Center for High-efficiency Utilization of Solar Energy, Hubei University of Technology, Wuhan 430068, P. R. China; Hubei Engineering Research Center for Safety Monitoring of New Energy and Power Grid Equipment, Hubei University of Technology, Wuhan 430068, P. R. China*

**Juan Wang** − *Hubei Collaborative Innovation Center for High-efficiency Utilization of Solar Energy, Hubei University of Technology, Wuhan 430068, P. R. China*

**Lujun Wang** − *Hubei Collaborative Innovation Center for High-efficiency Utilization of Solar Energy, Hubei University of Technology, Wuhan 430068, P. R. China*

**Wei Huang** − *Hubei Collaborative Innovation Center for High-efficiency Utilization of Solar Energy, Hubei University of Technology, Wuhan 430068, P. R. China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c04991


### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was supported by the Key Research and Development Plan of Hubei Province (2021BGD013), the Open Foundation of Hubei Key Laboratory for High-efficiency Utilization of Solar Energy and Operation Control of Energy Storage System (HBSEES202214), the Special Project of Central Government for local Science and Technology Development of Hubei Province (2019ZYYD020), the Hubei University of Technology Ph.D. Research Startup Fund Project (BSQD2020015), the Science and Technology Research Program of Hubei Provincial Department of Education (T201805), and the subsidy project after demonstration project of the Intellectual Property Application of Hubei Province (2019ZSCQ01).

## REFERENCES


(1) Lai, X.; Jin, C.; Yi, W.; Han, X.; Feng, X.; Zheng, Y.; Ouyang, M. Mechanism, modeling, detection, and prevention of the internal short circuit in lithium-ion batteries: recent advances and perspectives. *Energy Storage Mater.* **2021**, *35*, 470−499.

(2) Cao, Y.; Liu, C.; Jiang, J.; Zhu, X.; Zhou, J.; Ni, J.; Zhang, J.; Pang, J.; Rummeli, M. H.; Zhou, W.; et al. Theoretical insight into high efficiency triple junction tandem solar cells via the band engineering of antimony chalcogenides. *Sol. RRL* **2021**, *5*, No. 2000800.

(3) Mei, N.; Xu, X.; Li, R. Heat dissipation analysis on the liquid cooling system coupled with a flat heat pipe of a lithium-ion battery. *ACS Omega* **2020**, *5*, 17431−17441.

(4) Xiong, R.; Sun, W.; Yu, Q.; Sun, F. Research progress, challenges and prospects of fault diagnosis on battery system of electric vehicles. *Appl. Energy* **2020**, *279*, No. 115855.

(5) Huang, L.; Liu, L.; Lu, L.; Feng, X.; Han, X.; Li, W.; Ouyang, M.; et al. A review of the internal short circuit mechanism in lithium-ion batteries: Inducement, detection and prevention. *Int. J. Energy Res.* **2021**, *45*, 15797−15831.

(6) Wang, J.; Zhang, S.; Hu, X. A Fault Diagnosis Method for Lithium-Ion Battery Packs Using Improved RBF Neural Network. *Front. Energy Res.* **2021**, *9*, No. 702139.

(7) Zheng, C.; Chen, Z.; Huang, D. Fault diagnosis of voltage sensor and current sensor for lithium-ion battery pack using hybrid system modeling and unscented particle filter. *Energy* **2020**, *191*, No. 116504.

(8) Ren, L.; Dong, J.; Wang, X.; Meng, Z.; Deen, J. A data-driven auto-cnn-lstm prediction model for lithium-ion battery remaining useful life. *IEEE Trans. Ind. Inf.* **2020**, *17*, 3478−3487.

(9) Zhou, Y.; Huang, Y.; Pang, J.; Wang, K. Remaining useful life prediction for supercapacitor based on long short-term memory neural network. *J. Power Sources* **2019**, *440*, No. 227149.

(10) Chao, W.; Zhu, C.; Sun, J.; Ge, Y. A synthesized diagnosis approach for lithium-ion battery in hybrid electric vehicle. *IEEE Trans. Veh. Technol.* **2017**, *66*, 5595−5603.

(11) Shang, Y.; Lu, G.; Kang, Y.; Zhou, Z.; Duan, B.; Zhang, C. A multi-fault diagnosis method based on modified Sample Entropy for lithium-ion battery strings. *J. Power Sources* **2020**, *446*, No. 227275.

(12) Jiang, J.; Li, T.; Chang, C.; Yang, C.; Liao, L. Fault diagnosis method for lithium-ion batteries in electric vehicles based on isolated forest algorithm. *J. Energy Storage* **2022**, *50*, No. 104177.

(13) Jiang, L.; Deng, Z.; Tang, X.; Hu, L.; Lin, X.; Hu, X. Data-driven fault diagnosis and thermal runaway warning for battery packs using real-world vehicle data. *Energy* **2021**, *234*, No. 121266.

(14) Bharathraj, S.; Adiga, S. P.; Kaushik, A.; Mayya, K. S.; Lee, M.; Sung, Y. Towards in-situ detection of nascent short circuits and accurate estimation of state of short in Lithium-Ion Batteries. *J. Power Sources* **2022**, *520*, No. 230830.

(15) Ma, G.; Xu, S.; Cheng, C. Fault detection of lithium-ion battery packs with a graph-based method. *J. Energy Storage* **2021**, *43*, No. 103209.

(16) Xue, Q.; Li, G.; Zhang, Y.; Shen, S.; Chen, Z.; Liu, Y. Fault diagnosis and abnormality detection of lithium-ion battery packs based on statistical distribution. *J. Power Sources* **2021**, *482*, No. 228964.

(17) Ma, M.; Duan, Q.; Li, X.; Liu, J.; Zhao, C.; Sun, J.; Wang, Q. Fault diagnosis of external soft-short circuit for series connected lithium-ion battery pack based on modified dual extended Kalman filter. *J. Energy Storage* **2021**, *41*, No. 102902.
(18) Yang, Z.; Li, J.; Jiang, H.; Liu, Z. A novel model-based damage detection method for lithium-ion batteries. *J. Energy Storage* **2021**, *42*, No. 102970.
(19) Gao, Z.; Cecati, C.; Ding, S. X. A Survey of Fault Diagnosis and Fault-Tolerant Techniques—Part II: Fault Diagnosis With Knowledge-Based and Hybrid/Active Approaches. *IEEE Trans. Ind. Electron.* **2015**, *626*, 3768−3774.
(20) Feng, X.; Pan, Y.; He, X.; Wang, L.; Ouyang, M. Detecting the internal short circuit in large-format lithium-ion battery using model-based fault-diagnosis algorithm. *J. Energy Storage* **2018**, *18*, 26−39.
(21) Sazhin, S. V.; Dufek, E. J.; Gering, K. L. Enhancing Li-ion battery safety by early detection of nascent internal shorts. *J. Electrochem. Soc.* **2016**, *164*, A6281−A6287.
(22) Dey, S.; Perez, H. E.; Moura, S. J. Model-based battery thermal fault diagnostics: Algorithms, analysis, and experiments. *IEEE Trans. Control Syst. Technol.* **2019**, *27*, 576−587.
(23) Muddappa, V. K. S.; Anwar, S. Electrochemical Model Based Fault Diagnosis of Li-Ion Battery Using Fuzzy Logic. *Am. Soc. Mech. Eng.* **2014**, 46483.
(24) Li, D.; Zhang, Z.; Liu, P.; Wang, Z.; Zhang, L. Battery fault diagnosis for electric vehicles based on voltage abnormality by combining the long short-term memory neural network and the equivalent circuit model. *IEEE Trans. Power Electron.* **2021**, *36*, 1303−1315.
(25) Pan, Y.; Feng, X.; Zhang, M.; Han, X.; Lu, L.; Ouyang, M. Internal short circuit detection for lithium-ion battery pack with parallel-series hybrid connections. *J. Cleaner Prod.* **2020**, *255*, No. 120277.
(26) Schmid, M.; Kleiner, J.; Endisch, C. Early detection of internal short circuits in series-connected battery packs based on nonlinear process monitoring. *J. Energy Storage* **2022**, *48*, No. 103732.
(27) Xia, B.; Shang, Y.; Nguyen, T.; Mi, C. A correlation based fault detection method for short circuits in battery packs. *J. Power Sources* **2017**, *337*, 1−10.
(28) Lai, X.; Yi, W.; Kong, X.; Han, X.; Zhou, L.; Sun, T.; Zheng, Y. Online detection of early stage internal short circuits in series-connected lithium-ion battery packs based on state-of-charge correlation. *J. Energy Storage* **2020**, *30*, No. 101514.
(29) Zheng, Y.; Lu, Y.; Gao, W.; Han, X.; Feng, X.; Ouyang, M. Micro-Short-Circuit Cell Fault Identification Method for Lithium-Ion Battery Packs Based on Mutual Information. *IEEE Trans. Ind. Electron* **2021**, *68*, 4373−4381.
(30) Schmid, M.; Gebauer, E.; Hanzl, C.; Endisch, C. Active model-based fault diagnosis in reconfigurable battery systems. *IEEE Trans. Power Electron.* **2021**, *36*, 2584−2597.
(31) Cleveland, W. S. Robust locally weighted regression and smoothing scatterplots. *J. Am. Stat. Assoc.* **1979**, *74*, 829−836. 1979
(32) Guo, Y.; Zhang, Z.; Chen, Y.; Li, H.; Liu, C.; Lu, J.; Li, R. Sensor Fault Detection Combined Data Quality Optimization of Energy System for Energy Saving and Emission Reduction. *Processes* **2022**, *10*, 347.
(33) Jing Song, G.; Wen Wang, Q. On the weighted least-squares, the ordinary least-squares and the best linear unbiased estimators under a restricted growth curve model. *Stat. Pap.* **2014**, *55*, 375−392.
(34) Mirezi, B.; Kaçıranlar, S.; Özbay, N. A minimum matrix valued risk estimator combining restricted and ordinary least squares estimators. *Commun. Stat. Theory Methods* **2021**, 1−11.
(35) Chen, G. Y.; Wang, S. Q.; Wang, D. Q.; Gan, M. Regularization methods for separable nonlinear models. *Nonlinear Dyn.* **2019**, *98*, 1287−1298.
(36) Yang, S.; Zhang, C.; Jiang, J.; Zhang, W.; Zhang, L.; Wang, Y. Review on state-of-health of lithium-ion batteries: Characterizations, estimations and applications. *J. Cleaner Prod.* **2021**, *314*, No. 128015.
(37) Sarmah, S. B.; Kalita, P.; Garg, A.; Niu, X. D.; Zhang, X. W.; Peng, X.; Bhattacharjee, D. A review of state of health estimation of energy storage systems: Challenges and possible solutions for futuristic applications of li-ion battery packs in electric vehicles. *J. Electrochem. Energy* **2019**, *16*, No. 040801.
(38) Chang, C.; Wu, Y.; Jiang, J.; Jiang, Y.; Tian, A.; Li, T.; Gao, Y. Prognostics of the state of health for lithium-ion battery packs in energy storage applications. *Energy* **2022**, *239*, No. 122189.
(39) Huang, S.; Yang, X.; Wang, L.; Chen, W.; Zhang, F.; Dong, D. Two-stage turnout fault diagnosis based on similarity function and fuzzy c-means. *Adv. Mech. Eng.* **2018**, *10*, No. 1687814018811402.
(40) Chan, T. M.; Rahmati, Z. An improved approximation algorithm for the discrete Fréchet distance. *Inf. Process Lett.* **2018**, *138*, 72−74.
(41) Jekel, C. F.; Venter, G.; Venter, M. P.; Stander, N.; Haftka, R. T. Similarity measures for identifying material parameters from hysteresis loops using inverse analysis. *Int. J. Mater. Form.* **2019**, *12*, 355−378.
(42) Weng, H.; Wang, S.; Wan, Y.; Lin, X.; Li, Z.; Huang, J. Discrete Fréchet distance algorithm based criterion of transformer differential protection with the immunity to saturation of current transformer. *Int. J. Electr. Power* **2020**, *115*, No. 105449.


https://doi.org/10.1021/acsomega.2c04991
*ACS Omega* 2022, 7, 40145−40162