International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 7914-7925

https://doi.org/10.70517/ijhsa463677

Intelligent Prediction Modeling of Battery Performance Decline Trend Driven by Real Vehicle Data

Lu He[1,*] and Wei Wei[1]
[1] School of Computer and Electronic Information, Guangxi University, Nanning, Guangxi, 530000, China
Corresponding authors: (e-mail: frauhelu@163.com).

Abstract The rapid development of the electric vehicle industry has prompted much attention to the assessment of power battery health status. In this study, an intelligent prediction model of battery performance decline trend is proposed based on real vehicle data. Firstly, the capacity increment analysis is used to extract health features from the battery charging process, which is processed by a double filter algorithm to obtain a smooth capacity increment curve, and ten key health features are extracted. Subsequently, the correlation between the features and the battery health state is evaluated by Pearson correlation analysis, and the study shows that the correlation coefficient of most of the health features is greater than 0.85, which verifies the effectiveness of the feature extraction. Based on this, a GRA-EMD-BILSTM prediction model incorporating the attention mechanism was constructed, which utilized empirical mode decomposition to decompose the non-smooth differential pressure sequence into multiple smooth components, and screened the associated features by gray correlation analysis, and combined with a bidirectional long- and short-term memory network to achieve high-precision prediction. The experimental results show that the prediction error of this method for B5 batteries is controlled in the range of -1.87% to 1.43%, and the MAE, RMSE and MAPE indexes are reduced by 0.0081, 0.011, and 0.0122, respectively, compared with the LSTM method alone. This study provides a reliable health state monitoring technology for battery management systems, which is of great significance for extending the service life of the batteries and guaranteeing the safe operation of electric vehicles.

Index Terms Capacity incremental analysis, Health feature extraction, Gray correlation analysis, Empirical modal decomposition, Bidirectional long- and short-term memory network, Intelligent prediction model

I. Introduction

Against the background of global energy crisis and increasingly severe environmental problems, the development of new energy vehicles has become the focus of attention of governments and research institutions. With the continuous development and progress of the automobile industry, the problems of environmental pollution and energy scarcity have become more and more serious, so all countries have begun to formulate the development plan of new energy vehicles [1]-[3]. Among them, a series of battery components, such as fuel cells and lithium batteries, serve as the key power sources of automobiles, determining their overall performance, production costs and revenues [4], [5].

In recent years, the durability study of automotive batteries has become has become an important and hot issue with practical application value. In a stable and reliable operating environment, all components of the battery will undergo a certain degree of irreversible degradation with the increase of the usage time, which will result in the natural decline of the battery performance, and the rate of the battery performance decline will also be accelerated when it is operated under complex operating conditions, such as start-stop cycles, load cycles, thermal cycles, etc. [6]-[9]. In practice, ignoring the battery performance degradation problem may lead to misunderstanding of the tolerability of internal battery components or battery serviceable conditions, overconfidence in the health of the system, and thus failure of early health management or timely system maintenance, and ultimately lead to premature battery damage or more serious safety accidents [10]-[13]. In order to improve the durability of the battery and ensure its long-term safe and reliable operation, it is necessary to make the necessary predictions about the fuel cell during its use. The prediction of battery degradation trend can help drivers to prolong the performance of battery packs by adjusting their driving habits, and can effectively avoid battery pack failures caused by battery degradation to protect the lives and properties of drivers [14]-[17].

As the core power source of electric vehicles (EVs), the health status of lithium-ion (Li-ion) batteries has a direct impact on the range capability, safety performance and user experience of EVs. During the actual operation of electric vehicles, Li-ion batteries experience complex and variable operating conditions, such as diverse charging and discharging modes, different ambient temperatures, and various load states, which together lead to irreversible degradation of battery performance. Currently, the industry relies on parameters such as residual capacity and internal resistance to assess the health status of the battery, but these parameters usually require a complete charge/discharge test to obtain, making it difficult to meet the demand for real-time monitoring. Battery management systems need to accurately predict battery performance decline trends based on partially accessible data without interrupting normal use, so that timely maintenance measures can be taken. Traditional battery health assessment methods mainly include electrochemical modeling, equivalent circuit modeling, and data-driven methods, in which electrochemical modeling and equivalent circuit modeling can reflect the internal mechanism of the battery, but the parameters are difficult to calibrate and complicated to calculate, and it is difficult to adapt to the changing conditions of the real vehicle environment. On the other hand, data-driven methods have gradually become the focus of battery health assessment research due to the advantages of no need to understand the internal mechanism of the battery and flexible model construction. However, the existing data-driven methods still face challenges in how to extract effective health features from limited battery operation data and how to construct prediction models that can accurately describe the nonlinear battery decline process.

Based on these issues, this study proposes an intelligent prediction model for battery performance decline trend based on real vehicle data. First, health features are extracted from the battery charging process using capacity incremental analysis, and smooth feature curves are obtained by processing through a filtering algorithm; second, correlation analysis is used to assess the validity of the health features, and features that are highly correlated with the battery health state are screened out; finally, a GRA-EMD-BILSTM prediction model incorporating the attention mechanism is designed, and non-smoothness is handled through empirical mode decomposition sequences, screening associated features using gray correlation analysis, and combining with bi-directional long and short-term memory networks to achieve high-precision prediction. The aim of this study is to provide a practical and effective intelligent assessment method for battery health status monitoring, which in turn supports the optimization of battery management system and safe operation of electric vehicles.

II. Health state feature extraction of lithium battery driven by real vehicle data

II. A. Principle of lithium-ion battery

II. A. 1) Composition of lithium-ion batteries

Lithium-ion battery specifics are as follows:

(1) Shell: the shell consists of external and internal parts, the internal increase in the vent and other designs, the external can be encapsulated with two kinds of materials, one is steel material and the other is aluminum material. For some need to provide short-circuit protection, researchers can also have the protection function of the integrated circuit board designed in the external.

(2) Diaphragm: This part plays the role of isolation, isolating the positive and negative poles, avoiding the formation of a circuit between the two, thereby protecting the internal stability.

(3) electrolyte solvent: the requirement itself is the need to have good electrical conductivity, on the basis of which also requires the internal resistance as small as possible, earlier used ether and so on as a solvent, but this kind of solvent itself with toxicity, under certain conditions, will also trigger an explosion and other phenomena, with the progress of science and technology, people have found a better material, such as potassium perchlorate, etc., and gradually replace the original material, and the use of more stable Gel, to ensure the safety and reliability of lithium-ion batteries in the process of use.

(4) Negative electrode material: the most commonly used material is graphite carbon material, with low embedding potential and huge resource storage of carbon material, which is one of the reasons why it can be widely used, and secondly, it is less polluting to the ecological environment, which is in line with the requirements of sustainable development, and its electrochemical properties are stable, which can ensure the safety in the process of use.

(5) Positive material: In order to realize higher output voltage, this part needs higher redox potential. The more common ones are lithium manganate and other materials, which use a layered lattice structure, and can be thickened or thinned out for different use cases, which increases its scope of application and is more convenient for promoting its use.

II. A. 2) Principle of operation of lithium-ion batteries

The cathode and anode, which consist of positive and negative particles respectively, the diaphragm, which allows only the electrolyte to pass through, the electrolyte, which can generate an electric charge by transporting lithium ions, and the battery casing together make up the lithium ion battery. As can be seen in the figure below, a collector capable of collecting charge is connected externally to simulate the general situation during charging and discharging. The current that is continuously released outward from the positive electrode will have an effect on the lithium ions within the electrolyte, and the ions in the negative electrode portion will be mentioned within the liquid, and due to the increase in the number of lithium ions in the negative electrode portion, it will promote their movement towards the positive electrode, and the collector in the external circuit will pick up the electrical energy that is coming from the negative electrode, and transmit it to the positive particles, through its positive electrode.

Charging is carried out by an external power source, and the general situation inside the battery is that the lithium ions, assisted by the electrolyte, pass through the intermediate barrier from left to right into the negative electrode, and the general situation outside is that the electrons pass through the circuit and enter the negative electrode of the battery, and finally combine with the ions at the negative electrode. The relevant changes that occur in the positive part of the battery during charging are:

$$\mathrm{LiMn_2O_4} \rightarrow \mathrm{Li_{1-x}Mn_2O_4} + x\mathrm{Li}^+ + xe^- \tag{1}$$

The relevant changes that occur in the negative portion during charging are:

$$6\mathrm{C} + x\mathrm{Li}^+ + xe^- \rightarrow \mathrm{Li}_x\mathrm{C}_6 \tag{2}$$

When the battery is discharged, what happens is completely reversed from when it is charged; the lithium ions that were moving toward the negative terminal, now with the help of the electrolyte, move once again across the diaphragm toward the positive terminal, and the electrons from the outside will move from right to left via an external circuit to the positive terminal, where they will combine with the ions and the following reaction will take place:

$$\mathrm{Li_{1-x}Mn_2O_4} + x\mathrm{Li}^+ + xe^- \rightarrow \mathrm{LiMn_2O_4} \tag{3}$$

The reactions that will occur at the negative electrode are as follows:

$$Li_xC_6 \rightarrow 6C + xLi^+ + xe^- \tag{4}$$

With the advancement of battery-related technology, its performance has been greatly improved compared with the original battery. The advantages are as follows:

1. When the battery is charging and discharging, its requirement for the ambient temperature has been reduced.
2. Higher voltage can be transmitted outward due to improved materials.
3. The phenomenon of self-discharge in batteries has been improved.
4. The amount of energy that can be released has been increasing.
5. After discharging, it is easier to obtain the remaining internal capacity.
6. It has become safer and more reliable during use.
7. Due to the improvement of the production material, the waste battery after use is almost harmless to the environment.
8. Its reusability has been improved, greatly increasing the number of times and duration of use.

Lithium-ion batteries also have their own aspects that need to be improved:

1. In order to increase safety, there needs to be a circuit used to play a protective role.
2. As the discharge voltage increases, the range of variation also increases, making it difficult to control.
3. The price of the materials used for manufacturing has increased due to the use of new materials.
4. Because of the difference between them and ordinary batteries, they cannot be mixed with ordinary batteries.

After the researchers' continuous exploration, the above mentioned situations are gradually solved to a certain extent, and it is believed that they will be completely overcome in the near future.

II. B.Power lithium battery health feature extraction

II. B. 1) Health Feature Extraction Based on Capacity Incremental Approach

(1) Capacity increment analysis method

Capacity incremental [18], [19] through the online measurement of the voltage, current data, equal interval calculation to obtain a set of voltage ΔV and capacity ΔQ, and by differential analysis methods, the traditional charge and discharge voltage curves involving the battery first-order phase change of the voltage plateau into the capacity incremental curves on the capacity incremental curve can be clearly identified capacity incremental $(\Delta Q / \Delta V)$ peaks, which have higher sensitivity to data changes compared to conventional charge/discharge curves. By analyzing the IC curve, the degradation process, aging mechanism characteristics and health state changes of Li-ion batteries can be explored by linking the electrochemical reaction changes to the peak positions, heights, areas and distances between different peaks on the IC curve. The capacity increment IC is:

$$IC = \frac{dQ}{dV} \approx \frac{\Delta Q}{\Delta V} = \frac{Q_k - Q_{k-1}}{V_k - V_{k-1}} \tag{5}$$

where IC is the capacity increment; ΔQ is the capacity change; ΔV is the voltage change; Q_k and Q_{k-1} are the capacity of the battery at the time of k and $k-1$; V_k and V_{k-1} are the voltage of the battery at the time of k and $k-1$.

Since the cyclic aging test experiment uses constant current to charge and discharge, ΔQ in the above equation is converted to the product of constant current I_k and time interval:

$$\Delta Q = I_k (t_k - t_{k-1}) \tag{6}$$

where I_k is the charging current at k moment; t_k and t_{k-1} are k and $k-1$ moments.

According to Eq. (5) and Eq. (6) can be obtained:

$$\frac{dQ}{dV} = \frac{I_k (t_k - t_{k-1})}{V_k - V_{k-1}} \tag{7}$$

Capacity increment analysis converts the battery voltage platform into the capacity increment peak in the IC curve that is easier to analyze, and analyzes the performance of lithium-ion batteries through the trend of capacity change, which can powerfully capture the performance decline process of power battery.

(2) IC curve denoising processing

This section extracts and processes the B5, B6, B7 and B8 charging processes. Since there is no real-time capacity data in the NASA raw data set and the sampling interval of battery voltage is large, extracting the IC curve will lead to uncertainty in the value of the peaks and valleys. Therefore, the battery voltage sampling interval is reduced by linear interpolation, which in turn captures the changing characteristics of the IC curve. However, since the voltage interval after interpolation becomes smaller, resulting in too many "burrs" in the IC curve extraction, this paper uses the moving window smoothing method to denoise the IC curve.

Unlike the NASA dataset, real-time capacity was recorded in the Oxford battery aging test. Therefore traversing each cycle to obtain voltage and capacity data, dQ/dV was calculated by equation (5). Kalman filtering algorithm is used to set parameters such as process noise covariance, observation noise covariance, initial error covariance, etc., and the uncertain information on the curve is fused to remove the interfering data in order to smooth the curve.

(3) IC curve-based health feature extraction

In addition, only considering the effect of IHF7-IHF9 on SOH and ignoring the potential relationship among the three IHFs will lead to feature redundancy and low model generalization ability. Therefore, IHF1~IHF3 are downscaled to one-dimensional IHF4 to ensure the mutual independence between feature attributes.

In the MDS algorithm, let the distance matrix of n samples $x_1, x_2, x_3 \cdots x_n$ under the high-dimensional space be D, where any two samples x_i, the distance between any two samples x_j is d_{ij}. The sample matrix after dimensionality reduction is Z, where z_i and z_j follow $\left\| z_i - z_j \right\| = d_{ij}$:

$$d_{ij}^2 = \left\| z_i - z_j \right\|^2 = \left\| z_i \right\|^2 + \left\| z_j \right\|^2 - 2 z_i^T z_j \tag{8}$$

Centering the sample matrix Z yields Eq:

$$\sum_{i=1}^{n} z_i = 0 \tag{9}$$

$$\sum_{i=1}^{n} d_{ij}^2 = \sum_{i=1}^{n} \left\| z_j \right\|^2 + n \left\| z_j \right\|^2 \tag{10}$$

$$\sum_{j=1}^{n} d_{ij}^2 = \sum_{j=1}^{n} \left\| z_j \right\|^2 + n \left\| z_i \right\|^2 \tag{11}$$

Summing both sides of equation (11) again gives equation (12):

$$\sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \left\| z_j \right\|^2 + n \sum_{i=1}^{n} \left\| z_i \right\|^2 = 2n \sum_{i=1}^{n} \left\| z_i \right\|^2 \tag{12}$$

Let the one-dimensional inner product matrix $B = Z^T Z$, then:

$$b_{ij} = -\frac{1}{2}\left(\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}^2 - \frac{1}{n}\sum_{i=1}^{n} d_{ij}^2 - \frac{1}{n}\sum_{j=1}^{n} d_{ij}^2 + d_{ij}^2 \right) \tag{13}$$

An eigen-decomposition of matrix B is obtained:

$$B = V\Lambda V^T \tag{14}$$

where, Λ is the eigenmatrix; V is the eigenvector matrix. The first largest eigenvalue and eigenvector are selected when downscaling the data to one-dimensional space. The data points after dimensionality reduction are represented as:

$$Z = V_1\Lambda_1^{1/2} \tag{15}$$

where, Λ_1 is the reduced eigenmatrix, $\Lambda_1^{1/2}$ is the new diagonal matrix consisting of the $1/2$ th power of the eigenvalues; and V_1 is the one-dimensional eigenvector matrix.

II. B. 2) Correlation analysis of battery health characteristics

Pearson and Spearman correlation coefficients are often used to characterize the degree of correlation between two variables, the closer the absolute value is to 1, the stronger the correlation is, and when the absolute value is 1, the two variables are completely correlated. The Pearson correlation coefficient is most appropriate when continuous data, normal distribution, and linear relationship are met, and the Spearman correlation coefficient is more reliable if any of the above conditions are not met, so this paper is based on the two for a comprehensive examination of health characteristics.

Pearson correlation coefficient can measure the linear relationship between two variables, and its calculation formula is as follows:

$$\rho = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{16}$$

where ρ is the Pearson correlation result; X_i is the health characteristic; $\bar{X}$ is the mean of the health characteristic; Y_i is the SOH; $\bar{Y}$ is the mean of the SOH.

Spearman's correlation coefficient refers to the method of obtaining correlation based on the rank order of two variables, which has no specific need for the form of the initial variable distribution and is widely used. The formula is:

$$\rho = \frac{\left(\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})\right)}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2 \sum_{i=1}^{N}(y_i - \bar{y})^2}} \tag{17}$$

where x_i represents the i th value of the battery feature data, $\bar{x}$ represents the average value of the feature data, y_i is the i th value of the battery capacity, and $\bar{y}$ is the average value of the battery capacity. In this paper, we will use the above equation to calculate the correlation coefficients of the health features IHF1~IHF4 extracted above.

II. C.Battery performance degradation feature extraction

II. C. 1) Health feature extraction based on capacity incremental analysis method

Here, in order to be able to further mine more information from the local data of the state of charge of lithium batteries, this paper continues to use the incremental capacity analysis (ICA) method for the mining of battery health characteristic information.ICA is a method used to analyze the performance of lithium-ion batteries. The method converts the complex characteristics of the internal electrochemical reflection of the battery into the characteristic parameters on the IC curve by means of the incremental capacity curve (IC), which demonstrates the performance changes of the battery in a more intuitive way, thus providing an effective research tool for analyzing the cyclic aging as well as the degradation mechanism of lithium-ion batteries.

Figure 1 shows the voltage-capacity (V-Q) curve. It can be found that in the middle of the charging process, there is a long period of time, the voltage changes more slowly, but in the voltage platform range is the main range of the charge amount, the battery charge throughput overall amount is larger, the practical application of the platform range is difficult to extract directly from the effective and easy to recognize the characteristics of the battery used to estimate the battery state of health SOH. The capacity increment curve is obtained by first-order derivation of the V-Q curve to obtain the capacity increment of the battery within one unit voltage, i.e., the change of dQ/dV, which transforms the flat voltage plateau region into the trough value in the IC curve that is easier to observe and analyze. In the capacity increment curve, the flatter the voltage plateau is, the more prominent the peak of the corresponding curve is, i.e., the change of the peak value can be used to reflect the impact due to battery aging during the battery charging process.

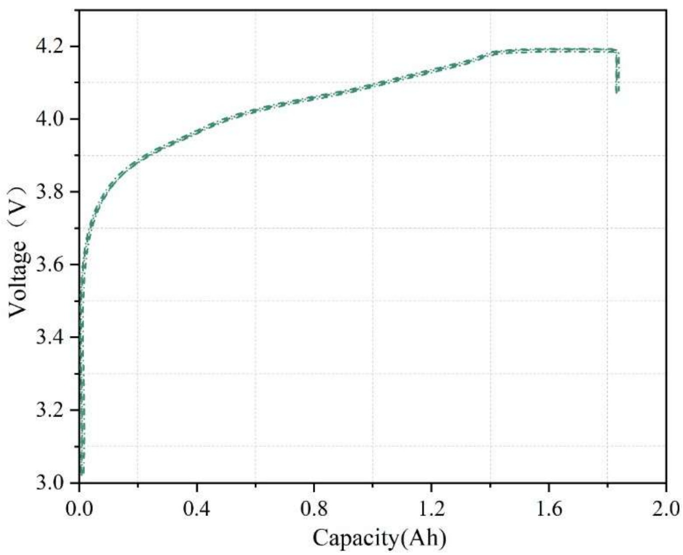


Figure 1: Voltage - capacity V-Q curve

The raw capacity increment IC curve is shown in Fig. 2. The sampling frequency is 0.2 Hz, and it can be seen that the original IC curve still has more noise signals, which cannot be directly used to extract the battery health characteristics, so the IC curve needs to be processed with noise reduction.

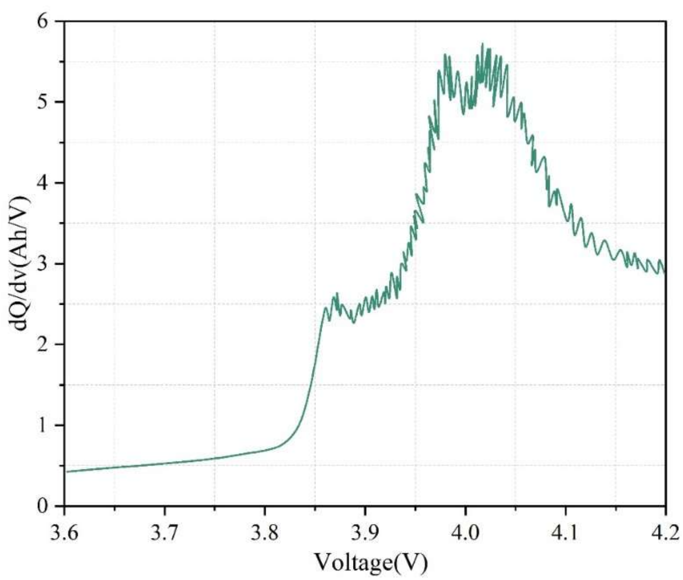


Figure 2: Original capacity incremental IC curve

From the curve smoothing effect and the retention of curve features, this paper firstly applies the MASF filtering method, sets the small moving window width M=25, i.e., takes the neighboring 25 data points to perform the mean value filtering calculation, and then generates the Gaussian window with σ 6 and the number of output points of the window 36, and then does the convolution operation with the curves after the MASF filtering, to get the filtered and noiseless IC curves after noise reduction. IC curve after filtering and noise reduction. Figure 3 shows the IC curve of the capacity increment after the filtering process. It can be seen that after the MASF-GSF double filtering algorithm, a more satisfactory curve smoothing effect is obtained, and the characteristic information such as the peaks and valleys of the curve is retained to a larger extent.

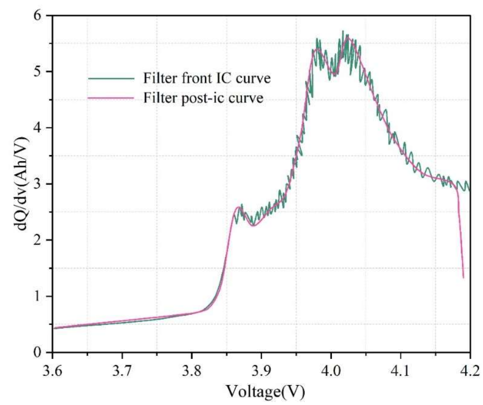


Figure 3: The volume incremental IC curve after filtering processing

Figure 4 shows the capacity increment curve of lithium battery under different cycle times. It can be seen that with the increase in the number of battery cycle charging and discharging, the battery gradually aging, the corresponding IC curve has also undergone obvious changes, the peak and valley area of the curve is gradually moving to the lower right, which is specifically manifested as a gradual decrease in the peak intensity of the peak of the peak of the IC curve, but the peak voltage corresponding to the peak is gradually increasing, that is, it indicates that the voltage plateau area of the battery charging process is changing towards a larger value of the voltage. The main reason for this phenomenon is related to the change of electrochemical information inside the battery aging, the battery positive and negative electrode active material in the cycle aging process there is an unavoidable loss, at the same time, the negative electrode surface of the battery SEI passivation film thickening makes the internal resistance increase, resulting in the battery in the constant-current charging stage of the battery charge to reduce the amount of energy, and will be faster to reach the cut-off voltage of the constant-current charging into the constant-voltage charging. In the figure, three IC curve spikes are marked with arrows, which are No. 1, No. 2 and No. 3, in which it can be seen that with the cyclic aging of the battery, the No. 1 spike gradually disappears, and the curve in this region becomes more and more gentle, compared with the No. 2 spike and No. 3 spike can still be clearly identified, and the peak intensity shows a decreasing trend, and the peak voltage is gradually shifted to the right, and therefore can be used as a characteristic to characterize the health state of the battery. Therefore, it can be used to characterize the health state of the battery.

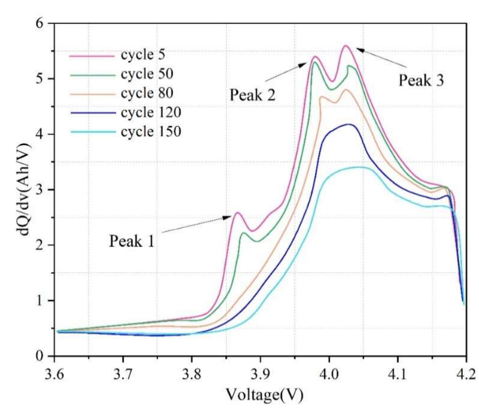


Figure 4: The incremental curve of the lithium battery in different cycles

Through the above analysis, the health features of the battery are extracted based on the ICA method, and the health feature extraction based on the ICA method is shown in Table 1. All six of the feature values are roughly distributed in the range of charging voltage interval from 3.8V to 4.0V in the constant current charging stage of the battery, which is in line with the demand for feature extraction of some of the charging data of the battery, so that the feature extraction of the battery based on the ICA method is theoretically feasible.

Table 1: The health feature extraction based on the ICA method

IC characteristic information	ICA battery health feature extraction	
	Capacity value-added (dQ/Dv)	Voltage (V)
Peak 2	HF5	HF6
Peak 3	HF7	HF8
Wave valley	HF9	HF10

II. C. 2) Correlation analysis of health characteristics of lithium-ion batteries

Pearson correlation analysis is a mathematical statistical method used to measure the degree of correlation between two continuous variables. The basic idea of this method is to calculate the Pearson correlation coefficient based on the covariance and standard deviation of the two variables in order to measure the correlation between the variables. The value of Pearson correlation coefficient ranges from [-1,1], -1 means that the two variables are completely negatively correlated, 0 means that the two variables are completely uncorrelated, and 1 means that the variables are completely positively correlated.

When the absolute value of the Pearson correlation coefficient is greater than 0.85, it indicates that the two variables are extremely strongly correlated. The Pearson correlation coefficients between the state of health SOH of the four batteries, B5, B6, B7, and B8, in NASA's publicly available lithium-ion battery dataset and the individual health characteristics were calculated by Pearson correlation analysis.

The Pearson correlation coefficients between the battery health features and battery SOH are shown in Table 2. The results show that on four different lithium batteries, most of the extracted battery health features have their correlations greater than 0.85, and some of them have correlations near 0.8, but in the analyzed results of lithium battery No. B8, very few outliers appeared, and the results of HF9 and HF10 are extremely low, which are only 0.2107 and 0.1924. Since Pearson correlation analysis only measures linear correlation between variables, the calculation of correlation in the case of nonlinearity has some limitations. In view of this, gray correlation analysis was continued to do quantitative analysis of the correlation between battery SOH and each health characteristic.

Table 2: The Pearson correlation coefficient of the battery SOH

Battery number	B5	B6	B7	B8
HF1	0.9441	0.8987	0.9531	0.7214
HF2	0.9857	0.9599	0.9839	0.8256
HF3	0.9892	0.9693	0.9837	0.8789
HF4	0.9932	0.9701	0.9791	0.8677
HF5	0.9932	0.9763	0.9824	0.9087
HF6	0.9947	0.9832	0.9839	0.9637
HF7	0.9908	0.9768	0.9802	0.9224
HF8	0.9952	0.9847	0.9851	0.9767
HF9	0.995	0.9913	0.9861	0.9859
HF10	0.9901	0.9862	0.9802	0.9735

III. Intelligent prediction model of battery performance decline trend and analysis of prediction results

III. A. Intelligent Prediction Model Design for Battery Performance Decline Trend

III. A. 1) BILSTM incorporating attention mechanisms

In this paper, BILSTM [20] is used as the basic network framework, in addition, this network can avoid long sequences from gradient vanishing or exploding during the training process.The neuron interior of BILSTM is the same as that of LSTM, which contains three gating modules, namely, forgetting gate, remembering gate, and output gate.X denotes inputs at the moment of t, C denotes cellular state, and H denotes the hidden layer state. The input of the network includes the differential pressure component obtained after processing and the correlation features affecting the change of differential pressure, which still retains the temporal information of differential pressure and is affected by the forward correlation features.The neurons of the BILSTM can selectively forget and memorize the new information about the cell state through the forgetting gate and the memory gate to save this useful information and pass it on to the subsequent neurons. The forgetting, memorizing, and outputting of information are controlled by the forgetting weight, memorizing weight, and outputting weight obtained from the last moment H and the current X through the internal weighting calculation of the neuron. Through the three gates of LSTM neurons, the feature information and temporal information hidden in the correlation features and pressure difference components can be effectively captured. Compared with LSTM, BILSTM can encode the information from back to front, which improves the ability to obtain the hidden information.

After the input information is processed by the BILSTM neurons, the output y is the H^t subjected to Softmax processing. I.e:

$$C^t = W^f C^{t-1} + W^i W \tag{18}$$

$$H^t = W^0 \tanh\left(C^t\right) \tag{19}$$

where, W^0 is the output weight; W^i is the memory weight; W^f is the forgetting weight.

In this paper, the attention mechanism is added to the BILSTM network framework to reasonably assign the weights according to the size of the error between the result and the point, which replaces the way of randomly assigning the weights by the neurons.

The attention mechanism can give enough attention to the temporal features extracted by BILSTM and ignore the unimportant information by reasonably assigning the weights. The attention weights are calculated using Softmax function, and after obtaining the attention weight W, it is weighted and integrated with the output of the neuron and mapped through the fully connected layer to obtain the predicted value. I.e:

$$e_i = \frac{h'h}{\|h'\|\|h\|} \tag{20}$$

$$W = \frac{e^i}{\sum_{j=1}^{m} e^j} \tag{21}$$

III. A. 2) GRA-EMD-BILSTM network

The parameters that can reflect the car condition information in the open source/non-open source data are total current, total voltage, state of charge (SOC), temperature and other parameters, these features reflect the external environment and driving behavior of the car, but it is not possible to judge whether they are correlated with the differential pressure or not, and if the poor correlation is made as a redundant feature inputted into the network, it will reduce the prediction accuracy of the model, so GRA-EMD-BILSTM is used to calculate the the correlation between the above parameters and the differential pressure to filter out the non-redundant features. Table 3 shows the correlation analysis table, the parameters with correlation greater than 0.6 are selected and considered as correlated features with higher correlation with differential pressure.

Table 3: Correlation analysis

Correlation degree	Cooperative hierarchy
$-1 \le G < -0.5$	Severe disorder
$-0.5 \le G < 0$	Moderate disorder
$0 \le G < 0.4$	Mild disorder
$0.4 \le G < 0.6$	Basic synergy
$0.6 \le G < 0.8$	Moderate synergy
$0.8 \le G < 0.9$	Good collaboration
$0.9 \le G < 1$	High quality synergy

The specific method of GRA-EMD-BILSTM is to construct each feature into a complete matrix $\begin{bmatrix} ri_1(1) & ri_1(2) & \cdots & ri_1(m) \\ \vdots & \vdots & \vdots & \vdots \\ ri_n(1) & ri_n(2) & \cdots & ri_n(m) \end{bmatrix}$ in the form of $\begin{bmatrix} ri_1(1) \\ \vdots \\ ri_n(1) \end{bmatrix}$, determine the pressure difference sequence

$[ri_0(1)\ ri_0(2) \cdots ri_0(m)]$, and after calculating the difference between each feature sequence and the corresponding positional element of the pressure difference sequence, determine the maximum difference and minimum difference between each feature sequence and the pressure difference sequence, respectively, and then determine the two levels of maximum difference and the two levels of minimum difference from these differences, which are $\Delta(Max)$ and $\Delta(Min)$, respectively, and finally Calculate the correlation coefficient, take the average value after getting the correlation coefficient between each feature sequence and the differential pressure sequence, and the result obtained is the correlation degree. That is:

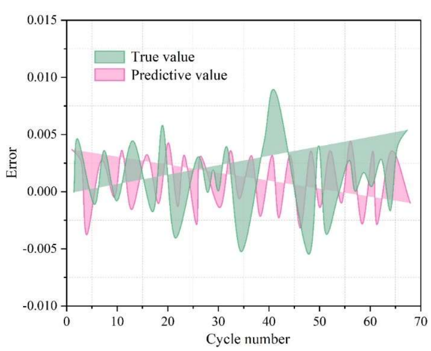
$$\xi_j(s) = \frac{\Delta(\min) + \rho\Delta(\max)}{\left|ri_0(s) - ri_j(s)\right| + \rho\Delta(\max)} \tag{22}$$

where, ρ is the resolution factor, which is taken as 0.5.

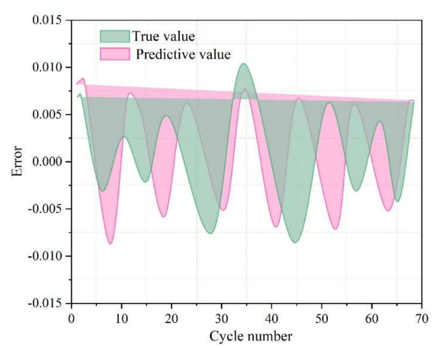
Although BILSTM is able to capture the nonlinearity and temporality of the differential pressure sequence, the nonsmooth part of this type of sequence with a high frequency of change leads to a decrease in the prediction accuracy, therefore, in this paper, EMD is used to decompose the differential pressure sequence into multiple smooth, multiscale components, which provides a better signal-to-noise ratio compared to the Fourier variation and the wavelet variation. The differential pressure sequence $V(t)$ is decomposed by EMD to obtain multiple intrinsic mode functions (IMFs) and residual signals Res. Namely:

$$V(t) = \sum_{i=1}^{n} IMF(t) + \mathrm{Re}\,s \tag{23}$$

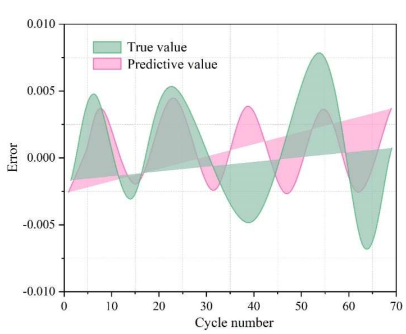
The GRA-EMD-BILSTM model incorporates the attention mechanism, the differential pressure sequence is decomposed by EMD to obtain multiple components, in the construction of feature engineering, these components are used as the target labels, combined with the corresponding GRA screened correlation features together with the input to the BILSTM network, the attention mechanism is introduced into the output portion of the network, and the predicted value of each component is obtained by the fitting of the fully connected layer, and the predicted value of each component is superimposed to be the final prediction result. The predicted values of each component are superimposed as the final prediction result.



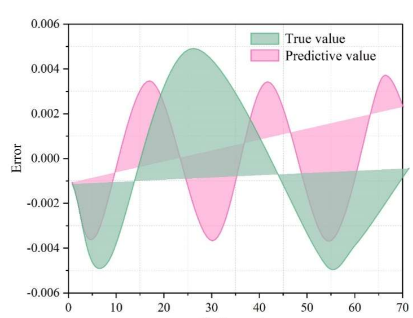
(a)IMF1



(b)IMF2



(c)IMF3



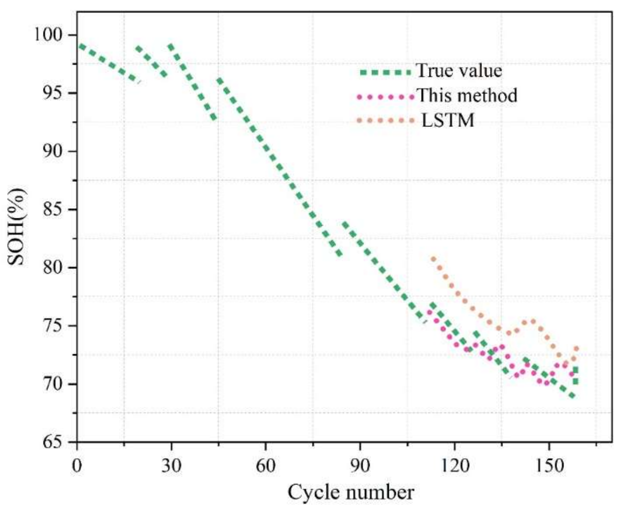
(d)IMF4

Figure 5: B5 lithium battery model component prediction results

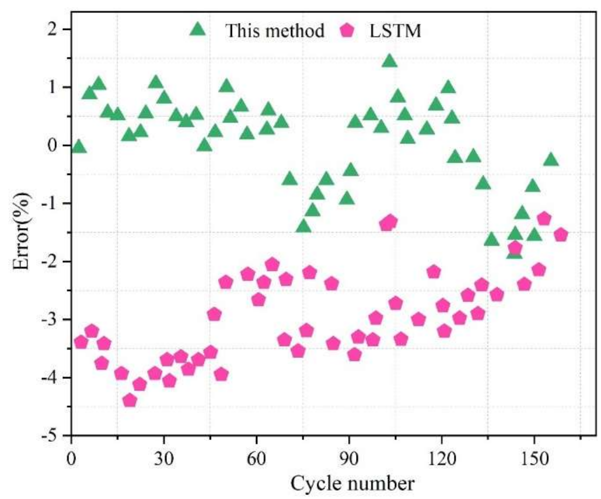
III. B. Analysis of forecast results

The prediction results of each modal component of the B5 lithium battery are shown in Fig. 5, where (a)~(d) represent IMF1-IMF4, respectively; the B5 prediction results and errors are shown in Fig. 6, where (a) and (b) represent the B5 prediction results and the errors of the B5 prediction results, respectively; and the RMSE indexes of the prediction errors of the battery for each modal are shown in Table 4.

It can be seen that the original sequence is first processed by using EMD method, which can well deal with the partial capacity augmentation phenomenon that exists in lithium-ion batteries, and then each IMF component and residual term is predicted separately by using LSTM method, and then the prediction results are summed up and reconstructed to output the prediction results, and it can be seen that compared with the separate LSTM prediction results, the prediction results of GRA-EMD-BILSTM The accuracy of the results is greatly improved, and the error is less than -1.87% to 1.43%, and according to the computational evaluation indexes, it can be seen that the MAE, RMSE, and MAPE of the B5 cell of the GRA-EMD-BILSTM method used in this paper are reduced by 0.0081, 0.011, and 0.0122, respectively, compared with that of the LSTM method alone, which indicates that the adopted GRA-EMD-BILSTM method not only can take into account the effect of capacity augmentation, but also ensures a good prediction accuracy.



(a) B5 prediction results



(b)B5 prediction error

Figure 6: B5 prediction results and errors

Table 4: The battery model prediction error RMSE index

Method	MAE	RMSE	MAPE
LSTM	0.0139	0.0167	0.0195
This method	0.0058	0.0057	0.0073

IV. Conclusion

The intelligent prediction model in this paper is able to effectively capture the battery performance degradation trend driven by real vehicle data. Ten health features are extracted from the battery charging process by capacity incremental analysis, and it is verified that these features are highly correlated with the battery health state, with the Pearson correlation coefficients of most of the features exceeding 0.85. The GRA-EMD-BILSTM prediction model incorporating the attention mechanism successfully solves the problem of non-smoothness in the battery sequence and significantly improves the prediction accuracy. Experiments show that the prediction error of this model for B5 batteries is controlled between -1.87% and 1.43%, and the prediction evaluation indexes MAE is reduced by 0.0081, RMSE is reduced by 0.0057, and MAPE is reduced by 0.0073 compared with the traditional LSTM method. The empirical mode decomposition method effectively handles the capacity augmentation phenomenon that exists in the lithium-ion batteries, and it can be used to predict the capacity of lithium-ion batteries by decomposing the nonstationary sequences into multiple smooth components, combined with gray correlation analysis to screen the correlation features, which substantially improves the model's ability to predict the battery decline trend. This health state assessment method based on partial charging data provides a reliable monitoring means for the battery management system, which is of great value for extending the service life of batteries and ensuring the safe operation of electric vehicles.

References

[1] Wang, X., Song, Y., Xia, W., Liu, H., & Yang, S. (2018). Promoting the development of the new energy automobile industry in China: technology selection and evaluation perspective. Journal of Renewable and Sustainable Energy, 10(4).

[2] Yuan, K., Wang, C., & Xiao, F. (2025). Navigating the energy crisis: automakers' coopetition strategies under dual credit policy. Science and Technology for Energy Transition, 80, 11.

[3] Su, C. W., Yuan, X., Shao, X., & Moldovan, N. C. (2023). Explore the environmental benefits of new energy vehicles: evidence from China. Annals of Operations Research, 1-20.

[4] Kendall, M. (2018). Fuel cell development for New Energy Vehicles (NEVs) and clean air in China. Progress in Natural Science: Materials International, 28(2), 113-120.

[5] Zhao, Q., Hu, L., Li, W., Liu, C., Jiang, M., & Shi, J. (2020). Recovery and regeneration of spent lithium-ion batteries from new energy vehicles. Frontiers in Chemistry, 8, 807.

[6] Kabir, M. M., & Demirocak, D. E. (2017). Degradation mechanisms in Li-ion batteries: a state-of-the-art review. International Journal of Energy Research, 41(14), 1963-1986.

[7] Zhang, T., Wang, P., Chen, H., & Pei, P. (2018). A review of automotive proton exchange membrane fuel cell degradation under start-stop operating condition. Applied energy, 223, 249-262.

[8] Venugopal, P. (2019). State-of-health estimation of li-ion batteries in electric vehicle using indrnn under variable load condition. Energies, 12(22), 4338.

[9] Stadler, R., & Maurer, A. (2019). Methods for durability testing and lifetime estimation of thermal interface materials in batteries. Batteries, 5(1), 34.

[10] Zhang, G., Wei, X., Wang, X., Zhu, J., Chen, S., Wei, G., ... & Dai, H. (2024). Lithium-ion battery sudden death: Safety degradation and failure mechanism. ETransportation, 20, 100333.

[11] Etxandi-Santolaya, M., Mora-Pous, A., Canals Casals, L., Corchero, C., & Eichman, J. (2024). Quantifying the impact of battery degradation in electric vehicle driving through key performance indicators. Batteries, 10(3), 103.

[12] Arun, M., Giddey, S., Joseph, P., & Dhawale, D. S. (2025). Challenges and mitigation strategies for general failure and degradation in polymer electrolyte membrane-based fuel cells and electrolysers. Journal of Materials Chemistry A, 13(16), 11236-11263.

[13] Rauf, H., Khalid, M., & Arshad, N. (2022). Machine learning in state of health and remaining useful life estimation: Theoretical and technological development in battery degradation modelling. Renewable and Sustainable Energy Reviews, 156, 111903.

[14] Zhou, J., Shu, X., Zhang, J., Yi, F., Jia, C., Zhang, C., ... & Wu, G. (2024). A deep learning method based on CNN-BiGRU and attention mechanism for proton exchange membrane fuel cell performance degradation prediction. International Journal of Hydrogen Energy, 94, 394-405.

[15] Nuhic, A., Bergdolt, J., Spier, B., Buchholz, M., & Dietmayer, K. (2018). Battery health monitoring and degradation prognosis in fleet management systems. World Electric Vehicle Journal, 9(3), 39.

[16] Russell, M. B., King, E. M., Parrish, C. A., & Wang, P. (2021). Stochastic modeling for tracking and prediction of gradual and transient battery performance degradation. Journal of Manufacturing Systems, 59, 663-674.

[17] He, H., Zhang, W., Long, Z., Xie, Y., Zhou, J., & Zhuang, L. (2025). Prediction of lithium-ion batteries health decline trajectories based on early image features. Journal of Energy Storage, 124, 116820.

[18] Hongxing Chen,Wenhui Yue,Guangfu Bin,Qi Jiang,Wei Shao & Chengqi She. (2024). Filter methods comparation for incremental capacity analysis in lithium-ion batteries health prediction. Journal of Energy Storage,101(PB),113878-113878.

[19] Ehsan Maleki Behzad,Ghafi Hamid,Azimi Nasab Morteza,Zand Mohammad,Sanjeevikumar P. & Khan Baseem. (2023). Statistical analysis of storage capacity increment effect in micro-grid management with simultaneous use of reconfiguration and unit commitment. Cogent Engineering,10(2),

[20] Mayank Kumar Jain,Dinesh Gopalani & Yogesh Kumar Meena. (2025). Hybrid CNN-BiLSTM model with HHO feature selection for enhanced fake news detection. Social Network Analysis and Mining,15(1),43-43.