Comparative Study of Incremental Capacity Curve Determination Methods for Lithium-Ion Batteries Considering the Real-World Situation

Peng Liu, Yizhong Wu, Chengqi She, Zhenpo Wang, and Zhaosheng Zhang

***Abstract*—The incremental capacity analysis (ICA) method is a promising method in battery state of health (SOH) estimation studies. The incremental capacity (IC) curve determination is one of the critical parts of the ICA method. However, the uncertain and incomplete charging conditions of real-world electric vehicles (EVs) significantly limit the IC curve determination. This article provides a thorough analysis of the practicality and limitations of four IC curve determination methods based on the datasets collected from real-world operating EVs with a comprehensive comparison scheme. The Lorentz function fitting method is improved by the differential evolution algorithm, breaking the limitation of fixed parameter boundary constraints. A novel PCHIP method is further proposed to determine the IC curve, with higher robustness to realistic uncertain and incomplete charging conditions. The proposed method is validated by real-world data from 40 EVs with low sampling frequency. The results show the extracted features from the IC curves, determined by the proposed method have a stronger correlation with the SOH, allowing the accurate SOH estimation with a 2% error. With less computational resources and sampling frequency requirements, this method shows great potential for the realistic battery management system implementation.**

***Index Terms*—Incremental capacity analysis (ICA), lithium-ion battery, real-world data, state of health (SOH).**

I. Introduction

With the aggravation of the energy crisis and global warming, the transition to electrification of the transportation industry has become a worldwide consensus. The development of electric vehicles (EVs) has been widely recognized as one of the effective ways to achieve energy saving and emission reduction in the transportation sector [1], [2]. Due to the superiority in capacity density and cycle life, lithium-ion batteries have become a preferable choice over other battery chemistries in various aspects and are being dominantly applied as storage devices for EVs [3]. However, the performance of lithium-ion batteries will inevitably deteriorate with time and cycling, including energy and power degradation [4]. The aging of the battery will gradually reduce the remainder driving range and dynamic performance of the EVs [5]. Thus, the precise estimation of the battery state of health (SOH) is crucial for the EVs' efficient and safe operation.

The currently proposed methods for SOH estimation are generally categorized into model-based and data-driven methods [6]. Model-based methods are usually based on an electrochemical model or an equivalent circuit model, using various optimization algorithms and observers to identify the internal parameters and estimate SOH [7]. Electrochemical model-based methods require detailed cell specification information and complex partial differential equations, which requires high computational resources for the battery management system (BMS) and is limited in practical applications [8]. The equivalent circuit model-based methods use circuit elements with nonlinear parameters to simulate the behavior of the battery. However, these methods have limited accuracy and robustness in modeling complex battery systems [9]. Data-driven methods are gradually becoming a popular direction for SOH estimation, due to their flexibility and nonlinear matching ability [10]. These methods estimate SOH through mapping externally measurable features to the internal degradation of the battery. Machine learning methods, such as support vector regression (SVR) [11], Gaussian process regression [12], random forest (RF) [13], and neural network methods [14], have been broadly employed in SOH estimation. It is crucial and challenging for data-driven methods to select the appropriate health features. Existing health features are mainly extracted from current, voltage, and temperature data during battery charging and discharging [15]. Among them, features, such as peak height and peak position, of IC curves are proved to be highly correlated features with the SOH and show an excellent performance [16].

The incremental capacity analysis (ICA) method is considered an effective and noninvasive technique to identify lithium-ion battery (LIB) degradation modes, which can be implemented in BMS [17]–[19]. The ICA method converts the hardly extracted plateau region on the voltage curve into clearly identifiable

Manuscript received December 24, 2021; revised March 31, 2022; accepted April 29, 2022. Date of publication May 10, 2022; date of current version June 24, 2022. This work was supported by the Ministry of Science and Technology of the People's Republic of China under Grant 2019YFE0107900. Recommended for publication by Associate Editor S. Williamson. *(Corresponding authors: Chengqi She; Zhenpo Wang.)*

Peng Liu, Yizhong Wu, Zhenpo Wang, and Zhaosheng Zhang are with the National Engineering Laboratory for Electric Vehicles, Beijing Institute of Technology, Beijing 100081, China (e-mail: bitliupeng@bit.edu.cn; 3120200402@bit.edu.cn; wangzhenpo@bit.edu.cn; zhangzhaosheng@bit.edu.cn).

Chengqi She is with the National Engineering Laboratory for Electric Vehicles, Beijing Institute of Technology, Beijing 100081, China, and also with the Hunan Provincial Key Laboratory of Health Maintenance for Mechanical Equipment, Hunan University of Science, Technology, Xiangtan 411201, China (e-mail: shechengqi@bit.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPEL.2022.3173464.

Digital Object Identifier 10.1109/TPEL.2022.3173464

peaks. Each peak has multiple features, such as peak height and position, describing the internal chemical reactions and battery degradation. To accurately and conveniently identify the features of the IC curve, researchers have proposed numerous methods. Li *et al.* [20] developed a method for determining the IC curve based on a Gaussian filter and used the peak position of the IC curve to construct a linear regression (LR) model for SOH estimation. Jiang *et al.* [21] used the Kalman filter to obtain a smooth IC curve and established a capacity estimation framework for ICA considering charging conditions. Agudelo *et al.* [22] applied the Savitzky–Golay filter to smooth the IC curve and used peak area as the health feature. In addition, some researchers have also considered determining the IC curve by building a capacity–voltage (CV) model. Han *et al.* [23] constructed a neural network to fit the charging voltage curve and quantitatively analyzed the relationship between the neural network parameters and the battery SOH. Li *et al.* [24] used a Lorentzian function to build a (CV) model, compared to the work in [25], which reduces the number of parameters of the model. Weng *et al.* [26] proposed an SVR method to fit the charging curve, achieving 1% SOH error prediction.

In summary, the current IC curve determination methods are mainly divided into numerical differentiation methods with filtering and fitting methods based on CV models. Although these methods have already shown expected performance in the laboratory environment, some drawbacks that need further improvement in the realistic situation with a partial charging process and low sampling frequency. Filtering methods can obtain smooth IC curves, but suffer from insufficient data and numerical problems at low sampling frequency realistic data, which disturb the extraction of IC features [14], [20]. Fitting methods have the ability to extract IC features by building function expressions conveniently. However, the accurate fitted CV models depend on sufficient data and reasonable function parameter boundary settings derived from experiments [24]. The complex models and computationally intensive parameter identification algorithm limit the implementation of the fitting methods on BMS. Moreover, most of the existing IC curve determination methods are based on experimental data of low current complete charging without considering the variable operating conditions of realistic EVs, which are intuitively different from the well-controlled laboratory environment [27]. In real-world situations, the charging process is uncertain and rarely follows a 0%–100% complete constant-current (CC)–constant-voltage charging pattern due to the range anxiety of the driver [28]. Considering that the fitting methods require identifying the number of IC peaks in the charging process before determining the model structure [24], [28], the incomplete and uncertain charging process in the realistic situation restricts the determination of the function structure of the fitting methods.

To explore the practicability and limitations of the IC curve determination methods under realistic charging conditions and low sampling frequency, this article attempts to comprehensively compare the proposed IC curve determination methods based on the real-world data of EVs and propose a novel method with higher robustness to the realistic uncertain and incomplete charging process. The main contributions of this article are as follows.

1) First, a comprehensive comparison scheme of IC curve determination methods considering the real-world situation is proposed, including computational efficiency, robustness, and IC features comparisons.
2) Second, the Lorentz function fitting method is improved based on the differential evolutionary (DE) algorithm with limitation breaking through of parameter boundary constraints and improvement of universality.
3) Finally, a novel IC curve determination method is proposed based on the Hermite interpolation (HI) method, which is more applicable in incomplete charging processes and shows better potential for BMS implementation due to the low demand for computational resources and sampling frequency.

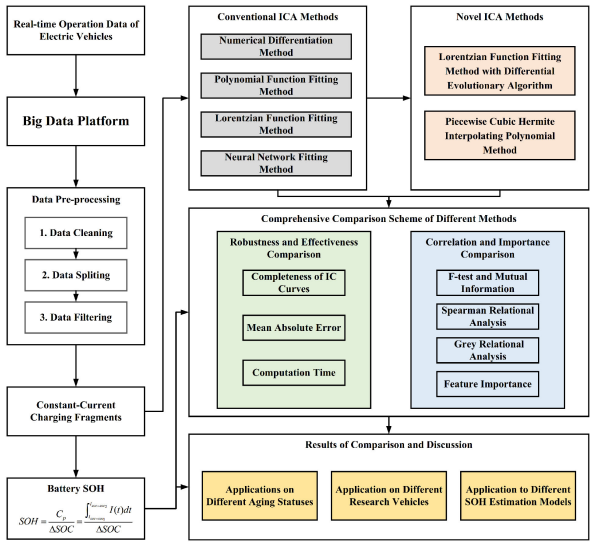


Fig. 1. Logical diagram of research.

By using real-world data from 40 EVs, the proposed method has proved to be more effective than the previous methods for extracting health features and allows for a more accurate estimation of SOH for batteries with varying trajectories of degradation.

The rest of this article is organized as follows. Section II describes the data source and preprocessing process. Section III introduces the conventional methods. Section IV demonstrates the algorithm to improve the existing method and describes a new method to generate IC curves. Section V proposes a comprehensive comparison scheme. Section VI discusses the comparison results. Finally, Section VII concludes this article. The flowchart of the study is shown in Fig 1.

II. Data Introduction

A. Data Source

The data used in this article are real-world operation data of EVs collected by the Open Lab of the National Big Data Alliance of New Energy Vehicle of China. The data collection

TABLE I
SPECIFICATIONS OF THE STUDIED EVS

Parameters	Type A	Type B
Battery material	$LiFeO_4$	$LiFeO_4$
Nominal capacity	240 Ah	172Ah
Nominal voltage	537.6 V	518.4V

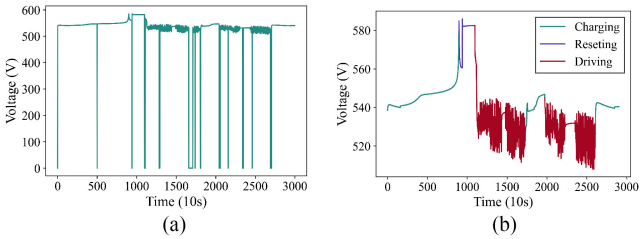


Fig. 2. Data cleaning of abnormal data. (a) Before data cleaning. (b) After data cleaning.

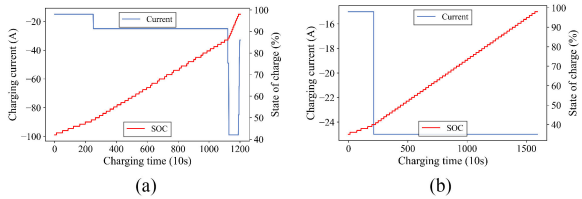


Fig. 3. Uncertain and partial realistic charging processes. (a) Charging fragment A. (b) Charging fragment B.

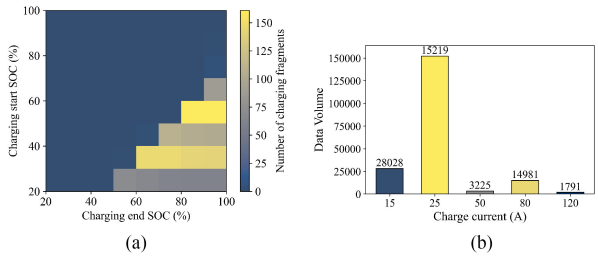


Fig. 4. Charging fragment SOC and current statistics. (a) Starting and ending SOC. (b) Charging current.

frequency is 0.1 Hz, covering 73 data items, including time, speed, mileage, total voltage, total current, state of charge (SOC), and temperature. A detailed introduction to the data can be found in [14]. The real-world data of 40 vehicles of two types operating in different cities are collected for the study, covering the historical data from February 2017 to January 2020, and Table I gives the specifications of the studied electric buses.

B. Data Preprocessing

The real-world data of EVs are collected through onboard sensors and transferred to the Big Data platform. Due to the inherent sensor limitations and transmission capability, missing value and noisy data are inevitably introduced during data collection. In order to obtain appropriate data for ICA, the raw data need to be preprocessed, including data cleaning, data slicing, and data filtering.

1) Data Cleaning: Data cleaning aims to remove abnormal data (e.g., the voltage of 0 V). For illustration, Fig. 2 visualizes the comparison of the data quality before and after cleaning. It can be seen that the data with the voltage of 0 V are removed, and the cleaned data can be divided into charging, driving, and resting states.

2) Data Splitting: ICA is achieved through charging or discharging the battery at a small CC rate [29]. However, there are rarely CC discharge data in realistic driving situations. By contrast, there are enough CC charging data accumulated during the charging period of the EVs. Therefore, the charging data are extracted from the cleaned data to perform ICA. Table II presents a sample of a typical charging fragment.

3) Data Filtering: Furthermore, it is worth noting that the charging process under realistic conditions is incomplete and uncertain. For the demonstrative purpose, Fig. 3 shows the realistic charging processes for two different charging fragments. It can be seen that the charging strategies and ranges are different between different charging fragments. Since the number of peaks in the IC curve varies in different SOC ranges [27], it is necessary to extract data in the same SOC range for each charging fragment to determine the model structure of the CV model-based fitting methods [28].

The principle of data filtering is to select data with a small charging current while ensuring that a sufficient amount of data is available for analysis. Fig. 4(a) shows the statistics of charging start and end SOC for all charging fragments of one studied vehicle, and it can be found that the majority of the charging fragments cover an SOC range of 50%–80%.

Fig. 4(b) shows the statistical results of the charging current in the 50%–80% charging range. It can be found that 76% of the charging data have a current of 25 A. Thus, considering sensor error and the fluctuation of the charging current, the data of the continuous charging fragments containing the SOC range of 50%–80% with a current of 24–26 A are retained for this studied vehicle.

III. CONVENTIONAL IC CURVE DETERMINATION METHODS

A. Numerical Differentiation Method With Gaussian Filtering

IC value is calculated through the charging capacity divided by voltage change within the specific interval, and the IC curve is depicted as the IC value versus voltage. Since current and voltage can be collected directly, the numerical differentiation method is the most straightforward way to generate the IC curve. The charging capacity can be acquired by ampere–hour integration as

$$Q = \int_0^T I(t)\mathrm{d}t = IT \tag{1}$$

where I is the charging current and T is the charging time. Then, the IC value can be calculated by

$$\mathrm{IC} = \frac{\mathrm{d}Q}{\mathrm{d}V} = \frac{\Delta Q}{\Delta V} = \frac{Q_j - Q_{j-1}}{V_j - V_{j-1}} = \frac{I\Delta t}{\Delta V} \tag{2}$$

TABLE II
SAMPLE OF CHARGING FRAGMENT

Timestamp	State	Speed (km/h)	Mileage (km)	Voltage (V)	Current (A)	State of charge (%)	Maximum temperature (°C)	Minimum temperature (°C)
20180226125849	1	0	42938.8	542.1	−78	44	19	15
20180226125859	1	0	42938.8	543.6	−78	44	19	15
20180226125909	1	0	42938.8	544.4	−78	44	19	15
...	...	...	...	...	...	...	...	...
20180226140159	1	0	42938.8	576.0	−32	98	30	25
20180226140209	1	0	42938.8	582.0	−32	98	30	25

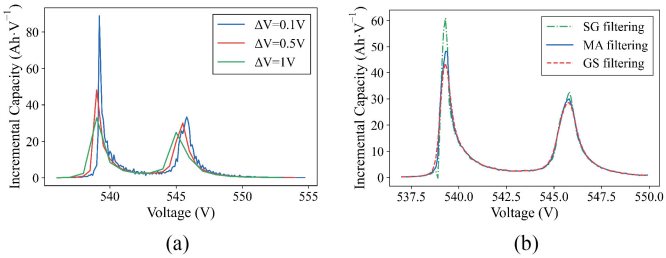


Fig. 5. IC curves at different voltage intervals and different filtering methods. (a) Different voltage intervals. (b) Different filtering methods.

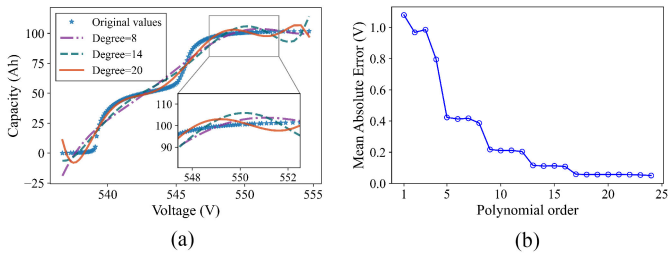


Fig. 6. Results of the PFF method. (a) Results of fit with voltage as the independent variable. (b) Mean absolute error (MAE) of different orders.

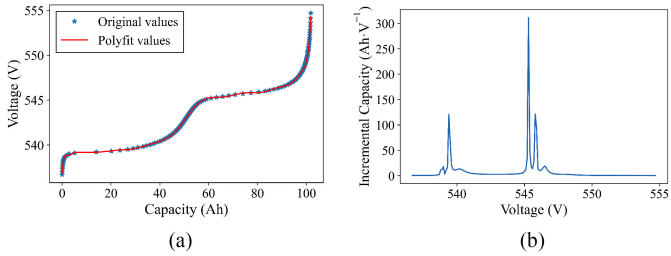


Fig. 7. CV and IC curves determined by the PFF method. (a) CV curve. (b) IC curve.

where ΔQ denotes the charged capacity in time Δt and ΔV is the increment of the voltage in time Δt.

Although IC value can be conveniently derived using the numerical differentiation method, the non-negligible noise must be smoothed. A rough smoothing can be achieved by selecting different ΔV. As shown in Fig. 5(a), despite the IC curve can be smoother as the ΔV increases, it will also be flattened and result in the peak dropping or even disappearing. In order to reserve adequate information of the IC curve, the voltage interval of 0.1 V is chosen.

To obtain a smooth IC curve, the IC values obtained after numerical differentiation need to be smoothed. The comparison of the three commonly used filtering algorithms: 1) the Savitzky–Golay filtering; 2) the moving average filtering; and 3) the Gaussian filtering, is shown in Fig. 5(b). It can be found that the IC curve is further smoothed by the Gaussian filtering, and Li *et al.* [20] concluded that the Gaussian filtering is more robust than other methods. Therefore, the IC curves determined by the numerical differentiation method with Gaussian filtering (NDGF method) are used in the subsequent comparative study.

B. Polynomial Function Fitting Method

Essentially, the IC value is the derivative of the charged capacity with respect to the voltage. Therefore, a polynomial function fitting (PFF) method was proposed to acquire IC value by deriving the CV function. For calculation purposes, the independent variable of the polynomial function was generally set to voltage. However, when applied to real-world charging data, the polynomial function with voltage as the independent variable fails to fit the CV function accurately. As shown in Fig. 6(a), three polynomial functions of degrees 8, 14, and 20 are used to fit the CV function, but none can accurately approximate the original CV curve. Therefore, the independent variable of the polynomial function is set to the capacity here as

$$V = f_n(Q) = p_n Q^n + p_{n-1} Q^{n-1} + \cdots + p_1 Q + p_0 \quad (3)$$

where n is the polynomial degree, Q denotes the capacity, V represents the voltage, and p_n is the polynomial coefficient to be derived.

The polynomial degree n is the main factor affecting the accuracy of the fitted results. Fig. 6(b) shows the fitting error of the polynomial function with different degrees. $n = 13$ is chosen to make a tradeoff between fitting accuracy and time cost. As described in Fig. 7(a), the fitted voltages closely follow the actual voltage, and two voltage plateaus can be observed in the charging curve. The IC value can be obtained by taking the derivative of (3). Fig. 7(b) shows the IC curve. It can be seen that although the PFF method fits the CV curve accurately, some undesirable peaks do not correspond to the voltage plateau in the IC curve, which is the result of the overfitting phenomenon that occurs in higher order PFF.

C. Lorentzian Function Fitting Method

The PFF method shows an accurate CV curve fitting, but unsatisfactorily fits the IC curve after differentiation. Intending to find an appropriate fitting function for the CV curve, another

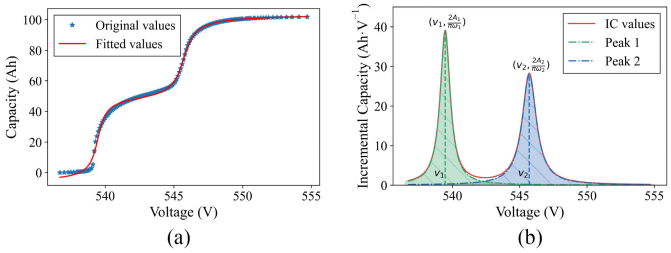


Fig. 8. CV and IC curves determined by the LFF method. (a) CV curve. (b) IC curve.

reasonable approach is to construct a multipeaked function related to the IC curve and define the CV function by integration. The Lorentzian function is a typical single-peaked function, given as

$$L(x) = \frac{2A}{\pi} \frac{\omega}{4(x - x_0)^2 + (\omega)^2} \tag{4}$$

where x_0 is the symmetry center of the peak, A represents the area under the peak, and ω is a parameter specifying the width of the peak.

Furthermore, Li *et al.* [24] used the Lorentzian function fitting (LFF) method to generate IC curve. The LFF method approximates the CV curve using the function derived by integrating the Lorentzian function, given as

$$Q = \sum_{i=1}^{n} \frac{A_i}{\pi} \arctan\left(\frac{2(v - v_i)}{\omega_i}\right) + C \tag{5}$$

where Q is defined as the charged capacity, n is the number of peaks, which varies for different materials of the batteries and charging SOC ranges, and for this study vehicle $n = 2$, A_i denotes the area corresponding to the ith peak, ω_i is the half-peak width of the ith peak, v_i represents the center of symmetry of the ith peak, and C is a constant produced by integration.

After the parameters are identified by the nonlinear least-squares (NLS) algorithm, the fitting results and the IC curve are shown in Fig. 8. The results exhibit that the two voltage plateaus in the CV curve can be transformed into two explicit observed peaks in the IC curve smoothly. Another advantage of the LFF method is that the parameters are associated with the battery phase transition. Specifically, the v_i represents the peak potential of the redox reaction, and the A_i is the total charged capacity in the phase transition derived from the usable active material [24]. Therefore, features of the IC curve relevant to battery aging can be derived directly from the parameters without calculating the complete IC curve and avoiding the introduction of calculation errors.

D. Neural Network Fitting Method

As discussed before, the parameters in the LFF method can represent the voltage plateau and corresponding capacity during the phase transition. Similarly, Han *et al.* [23] proposed a neural network fitting (NNF) method, in which the parameters of the neurons can directly reflect the capacity of the voltage plateau.

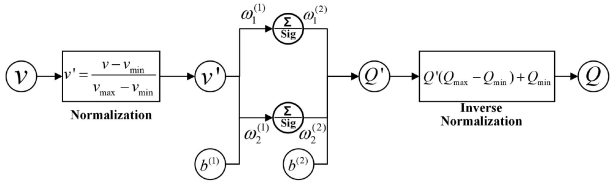


Fig. 9. Schematic diagram of the neural network structure.

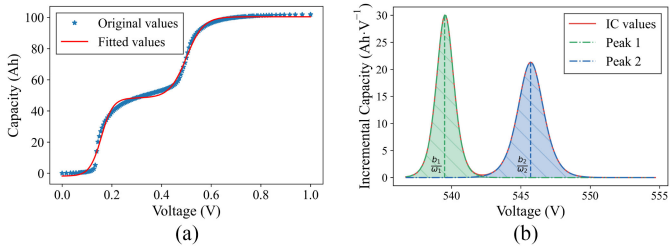


Fig. 10. CV and IC curves determined by the NNF method. (a) CV curve. (b) IC curve.

The critical part of the NNF method is the activation function in the hidden layers. Since the derivative function of the sigmoid function is also a single-peaked function, it is chosen as the activation function in this article. A shallow layer neural network with a single input, single output, and single hidden layer is further constructed to fit the CV curve. The neural network takes the normalized voltage as input, the normalized capacity as output, and has two sigmoid neurons in the hidden layer. Fig. 9 shows the detailed description of the model. According to the constructed neural network model, the CV function can be expressed as

$$Q' = \sum_{i=1}^{2} \omega_i^{(2)} \left(\frac{1}{1 + e^{-\left(\omega_i^{(1)} v' + b_i^{(1)}\right)}}\right) + b^{(2)} \tag{6}$$

where Q' is the normalized capacity, v' is the normalized voltage, $\omega_i^{(1)}$ represent the weights from the input layer to the hidden layer, $b_i^{(1)}$ are the biased items for the ith neuron, $\omega_i^{(2)}$ represent the weights from the hidden layer to the output layer, and $b^{(2)}$ is the bias item for the output layer.

The IC value can be obtained by taking the derivative of (6) after the neural network model is trained. Fig. 10 shows the fitted CV curve and the determined IC curve by the NNF method. Similar to the LFF method, it can be seen that the IC curve generated by the NNF method can be separated into two separate curves. Each curve reflects the corresponding voltage plateau, and the capacity related to the voltage plateau can be calculated by

$$A_i = \omega_i^{(2)}. \tag{7}$$

IV. Novel IC Curve Determination Methods

A. Improved LFF Method

Attributed to the characteristic of the LFF method, the peaks height and position of the IC curve can be directly calculated

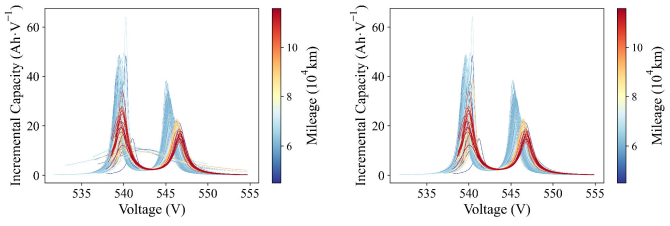


Fig. 11. Conventional LFF method and improved LFF method. (a) Conventional LFF method. (b) Improved LFF method.

from the function parameters. The NLS algorithm is widely used to identify these parameters. The reasonable bounds on the parameters need to be constrained to find the optimal solution to avoid local convergence in the NLS algorithm identification process [28]. In previous studies, the fixed parameter boundaries were determined based on detailed empirical knowledge from full charging cycle experiments [24]. However, it is time-consuming and uneconomical to conduct experiments on the on-board battery to obtain reasonable parameter boundaries. Moreover, the fixed parameter boundaries could not be adapted to all charging fragments due to the partial and diverse charging patterns of the real vehicle, as shown in Fig. 11(a).

The DE algorithm is applied to address this problem. The DE algorithm is a population-based adaptive global optimization algorithm that decides the direction of the optimization search through cooperation and competition between individuals within a population [30]. Similar to the genetic algorithm, the DE algorithm involves initializing populations, mutation, crossover, and selection. The detailed algorithm steps are as follows.

1) *Define the optimization objective function f:*

$$\min f\left(x_1, x_2, \ldots, x_D\right) = \min \sum \left(y_i - \hat{y}_i\right)^2 \quad \text{s.t. } x_j^L \le x_j \le x_j^U \tag{8}$$

where y_i is the original capacity value, $\hat{y}_i$ is the fitted capacity calculated by (5), x_j is the jth parameter in the LFF method, and D is the number of parameters.

2) *Initialize the population:* As with the genetic algorithm, the initial population is randomly generated.

$$\left\{X_i(0) \mid x_{i,j}^L \le x_{i,j}(0) \le x_{i,j}^U \quad i = 1, 2, \ldots, \mathrm{NP}; j = 1, 2, \ldots, D\right\} \tag{9}$$

where $X_i(0)$ is the ith individual of the initial population, $x_{i,j}(0)$ denotes the jth parameter in $X_i(0)$, and NP indicates the population size.

$$x_{i,j}(0) = x_{i,j}^L + \mathrm{rand}(0,1)\left(x_{i,j}^U - x_{i,j}^L\right) \tag{10}$$

where $x_{i,j}^U$ and $x_{i,j}^L$ are the lower and upper bounds of the jth dimension and $\mathrm{rand}(0,1)$ denotes the random number in the interval $[0,1]$.

3) *Mutation:* The DE algorithm achieves variation of individuals through the mutation strategy. The strategy *DE/rand/1* is the most widely used mutation strategy in the proposed papers [31], which uses the scaled differences of parent solutions as a mutation operator to generate candidates for the next generation population, given as

$$V_i(g+1) = X_{r1}(g) + F\left(X_{r2}(g) - X_{r3}(g)\right) \tag{11}$$

where $r1$, $r2$, and $r3$ are the mutually exclusive integers randomly generated within the range $[1, NP]$, $X_{ri}(g)$ denotes the rith individual of the gth generation population, and F is the difference coefficient, a real number for scaling the difference vector.

4) *Crossover:* The purpose of the crossover is to select individuals randomly, and the crossover operates as

$$U_{i,j}(g+1) = \begin{cases} V_{i,j}(g+1) & \text{if rand}\,(0,1) \le \mathrm{CR} \\ x_{i,j}(g) & \text{otherwise} \end{cases} \tag{12}$$

where CR is the crossover probability, $V_{i,j}(g+1)$ denotes the jth parameter in the ith individual generated in the previous step, and $U_{i,j}(g+1)$ is the jth parameter of the ith individual after crossover.

5) *Selection:* The DE algorithm uses a greedy selection strategy to select a better individual by comparing the fitness function value of the current individual with that of the incoming individual.

$$X_i(g+1) = \begin{cases} U_i(g+1), & \text{if } f\left(U_i(g+1)\right) \le f\left(X_i(g)\right) \\ X_i(g), & \text{otherwise} \end{cases} \tag{13}$$

where $X_i(g+1)$ is the ith individual of the $(g+1)$th generation population.

With the DE algorithm applied to each charging fragment, the parameter boundaries can be adaptively adjusted to accommodate the charging fragments at different aging statuses. The IC curves determined by the improved LFF method are shown in Fig. 11(b). It can be seen that the improved LFF method fits the CV curves of all charging fragments successfully and improves the stability of the IC curves determined by the LFF method.

B. Hermite Interpolating Polynomial Method

The IC curve determination methods discussed in the previous section have been broadly used under well-controlled laboratory conditions. However, some drawbacks are awaiting further improvement for better application of the ICA method under complex real-world conditions. Specifically, the NDGF method can only calculate the discrete IC value, which causes difficulty in obtaining accurate quantitative information of the IC curves. Moreover, the parameter identification algorithm of the fitting method complicates the IC curve determination process and reduces flexibility and versatility. To overcome these drawbacks, an interpolation method is proposed in this article.

The HI is one of the common interpolation methods in mathematical modeling, which can effectively tackle the problem of insufficient data volume [32]. Considering the insufficient amount of low-current CC charging data under realistic conditions, the HI methods can be applied to generate sufficient charging data for ICA. Therefore, an HI method is used here to rebuild the mapping relationship between charged capacity and voltage. In addition, to prevent the Runge phenomenon of the HI method at

higher orders [33], the piecewise cubic HI (PCHIP) method is applied.

The process of the PCHIP method can be described as follows. For each charging fragment, a set of ordered voltage $v_0 < v_1 \cdots v_k < v_{k+1} \cdots < v_{n-1} < v_n$ has the corresponding charging capacity $q_0 < q_1 \cdots q_k < q_{k+1} \cdots < q_{n-1} < q_n$. Based on the n data points (v_i, q_i), the interpolating polynomial function is constructed as

$$Q(v) = \begin{cases} Q_0(v), v \in [v_0, v_1] \\ Q_1(v), v \in [v_1, v_2] \\ \vdots \\ Q_{n-1}(v), v \in [v_{n-1}, v_n] \end{cases} \tag{14}$$

$$\begin{aligned} Q_i(x) = &\left(1 + 2\frac{v - v_i}{v_{i+1} - v_i} + (v - v_i)\, q_i'\right)\left(\frac{v - v_{i+1}}{v_i - v_{i+1}}\right)^2 q_i \\ &+ \left(1 + 2\frac{v - v_{i+1}}{v_i - v_{i+1}} + (v - v_{i+1})\, q_{i+1}'\right) \\ &\times \left(\frac{v - v_i}{v_{i+1} - v_i}\right)^2 q_{i+1} \end{aligned}$$

$$v \in [v_i, v_{i+1}]\, (i = 0, 1, 2, \ldots, n-1) \tag{15}$$

where n denotes the number of the data in the charging fragment, q_i and v_i are the charged capacity and voltage of the ith data point, respectively, and q_i' is the derivative of the q_i with respect to the v_i.

It can be seen that all the parameters in the polynomial $Q_i(v)$ except q_i' and q_{i-1}' can be obtained directly from the original data. Aiming to obtain the derivatives q_i' and q_{i-1}', a new strategy is proposed to solve these problems.

1) For interior data points v_k, q_k' can be calculated as follows:

$$\delta_k = \frac{q_{k+1} - q_k}{h_k} \tag{16}$$

where δ_k denotes the difference quotients at v_k and $h_k = v_{k+1} - v_k$.

If the δ_{k-1} and δ_k have opposite signs, then the q_k' at that point is set to be 0. Otherwise, q_k' is equal to the weighted average of δ_k and δ_{k-1}, given as

$$q_k' = \frac{\delta_{k-1} h_{k-1} + \delta_k h_k}{h_{k-1} + h_k}. \tag{17}$$

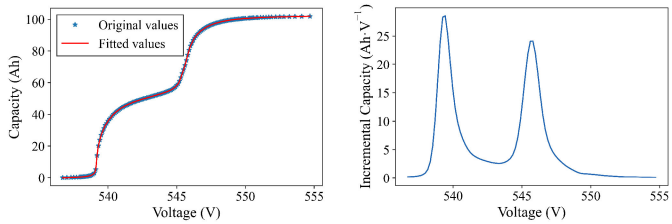
2) For the boundary data points v_0 and v_n

$$q_0' = \delta_0, q_n' = \delta_{n-1} \tag{18}$$

where δ_0 and δ_n are the difference quotients at v_0 and v_n, respectively, as defined in (16).

Once the q_i' is obtained, the interpolating polynomial function is determined. As shown in (19), the constructed polynomial Q_i, equals to the charging capacity q_i at the corresponding voltage v_i and the derivatives, is also constrained by the q_i'. Combined with the derivatives, more information on IC peaks can be retained from the charging data. Meanwhile, the parameters can be derived by simple numerical calculation, eliminating the tedious parameter identification process and increasing the possibility of the application on BMS.

$$Q_i(v_{i+1}) = q_{i+1}, Q_i(v_i) = q_i$$



(a) (b)

Fig. 12. Fitting results of CV curves and IC curves determined by the PCHIP method. (a) CV curve. (b) IC curve.

$$Q_i'(v_{i+1}) = q_{i+1}', Q_i'(v_i) = q_i'. \tag{19}$$

Based on the modified PCHIP method, the IC curves are obtained smoothly, as shown in Fig. 12(a). The IC curve generated by further differentiation and Gaussian filtering is shown in Fig. 12(b). It substantiates that the proposed method has a good performance in describing the relationship between voltage and capacity and identifying the peaks of the IC curve.

V. Comprehensive Comparison Scheme of Different Methods

A comprehensive comparison scheme of different IC curve determination methods is proposed in this section. The PFF method is out of consideration here due to the overfitting phenomenon.

A. Robustness and Effectiveness Comparison

The feasibility of the NDGF, LFF, NNF, and PCHIP methods is demonstrated when they are applied to one charging fragment, as described above. Considering that the IC curves will be diverse at different aging statuses of battery, and the robustness and effectiveness of four IC curve determination methods should be verified further on all historical charging fragments of real-world vehicles.

First, for the LFF and NNF methods, the accuracy of the fitted CV curve has a significant impact on the determination of the IC curve. Therefore, a comparison of the goodness of fitting is required. The goodness of fitting is evaluated quantitatively using the MAE here. Then, the four IC curve determination methods are applied to all charging fragments, and the robustness of these methods is compared based on the smoothness of the determined IC curves.

On the other hand, the high demand for computation resources and calculation time restrict the application of current SOH estimation techniques in the real-world industry. Considering the complexity of the parameter identification algorithms of the LFF and NNF methods, it is necessary to compare the computational time of different methods to evaluate their effectiveness.

B. Feature Correlation and Importance Comparison

Benefiting from the evolution of IC curves with the battery aging process, the features extracted from IC curves can be conveniently used as input features for data-driven SOH estimation

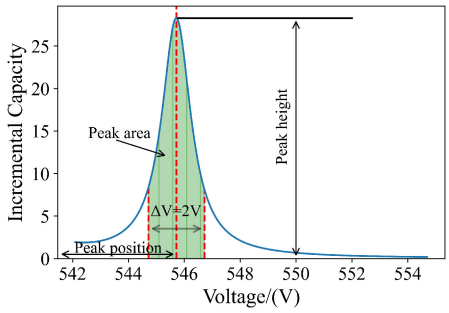


Fig. 13. Quantitative characteristics of IC curves for further analysis.

models. For a data-driven model, it is critical to select appropriate input features in consideration of efficiency and reliability. Therefore, in order to evaluate different IC curve determination methods, it is necessary to compare the IC features extracted from different IC curve determination methods. Although several previous studies have explored the relationship between IC features and SOH based on the correlation analysis [15], the influence and contribution of a given feature on the SOH estimation model accuracy are seldom quantified. To bridge this deficiency, the feature importance (FI) analysis is also applied here for comparison.

To conduct feature correlation and importance analysis, the multiple IC features H_i, P_i, and A_i are extracted from the second peak of the IC curves determined by the four IC methods, because the peak has a closer relationship with the loss of lithium inventory (LLI) [34]. The H_i and the P_i are the vertical and horizontal coordinates of the second peak, respectively. The A_i is obtained by taking the neighborhood δ at the second peak and integrating over the interval $(P_i - \delta, P_i + \delta)$. For illustrative purposes, the extracted IC features are shown in Fig. 13. Since the NDGF method cannot obtain the functional expression of the IC curve, the A_i cannot be extracted precisely. Thus, excluding the A_i feature from the NDGF method, a total of 11 IC features are extracted.

Another critical issue in the feature analysis is the definition of the SOH in real-world conditions. By convention, SOH is quantified by the capacity loss and the increase in internal resistance [35]. To quantitatively assess the SOH, the reference performance test (RPT) is usually conducted to measure the capacity and resistance of batteries [4]. However, the RPT needs to be regularly performed at a constant temperature. While conducting RPT is time-consuming and expensive in practical application, the concept of regional capacity was proposed and verified in [36] as an on-board indicator for battery SOH. Similarly, considering the high coulomb efficiency of commercial lithium-ion batteries and the convenience of obtaining charging data in practical applications, the capacity per 1% SOC for a fixed SOC range is defined as the SOH in this article, as shown in the following equation. The common SOC range and current of studied EVs are determined from the results in Fig. 4.

$$\mathrm{SOH} = \frac{C_p}{\Delta \mathrm{SOC}} = \frac{\int_{t_{\mathrm{soc_{start}}}}^{t_{\mathrm{soc_{end}}}} I(t)\mathrm{d}t}{\mathrm{SOC_{end}} - SOC_{\mathrm{start}}} \tag{20}$$

TABLE III
CALCULATION STEPS OF FI

Algorithm 1

step 1: Randomly split the training and test set in the ratio of 7:3.
step 2: Train the model using the training set and calculate the rmse of the model:
$\mathrm{RMSE}_{i0} = \sqrt{\sum_{k=1}^{n} \frac{(\hat{y}_k - y_k)^2}{n}}$
where $\hat{y}_k$ is the estiamted SOH, y_k is the actual SOH.
step 3: Randomly shuffler feature x_i in turn to retrain the model, and calculate RMSE_{ij} for the jth training.
step 4: Calculate the percentage decrease of model RMSE as FI.
$\mathrm{FI}_{ij} = \mathrm{RMSE}_{ij} - \mathrm{RMSE}_{i0}$
step 5: Repeat step1-step4 10 times and average the FI_{ij}.
$\mathrm{FI}_i = \frac{\sum_{j=1}^{10}(\mathrm{RMSE}_{ij} - \mathrm{RMSE}_{i0})}{10}$
step 6: Process the maximum–minimum normalization to FI_i

Considering the complex and variable operational conditions in the real world, batteries may not follow a simple linear degradation trajectory [37]. The commonly used Pearson correlation coefficient can only measure the degree of linear correlation between variables, so it has limited effectiveness in practical situations. To break this limitation, this article investigates the relationships between the 11 IC features and SOH using F-test, Spearman relational analysis (SRA), grey relational analysis (GRA), and mutual information (MI).

F-test and SRA are employed to deduce the linear correlation. F-test evaluates the correlation between the two variables by constructing the statistic F. With the increase of F, the correlation between x and y is stronger. SRA evaluates the correlation of two variables using a monotonic equation. The coefficient ρ is +1 or -1 when the two variables are fully monotonically correlated. The detailed procedures for F-test and SRA were introduced in [38] and [39].

GRA and MI are used to explore the nonlinear correlation between IC features and SOH. GRA quantitatively describes the degree of association by comparing the similarity of the geometric curves between two series. The stronger is the association of the two series, the closer to 1 the GRA coefficient will be. MI is the maximum information-based method, describing the strength of the relationship between two variables. The detailed procedures for GRA and MI were introduced in [29] and [40].

After the correlation analysis is completed, the FI analysis is performed to quantify further how useful the IC features are in the SOH estimation models. Since the same feature contributes differently to different models, it is necessary to construct a specific model to compare the FI. In this study, four commonly used models are developed: 1) SVR; 2) RF; 3) LR; and 4) classification and regression tree (CART). In each model, the 11 IC features are input to the models for the training. In addition, temperature and mileage have a significant influence on SOH [4]. Thus, they are also included in the SOH estimation. After the training, the root mean square error (rmse of the model can be obtained. Then, the FI of feature X is defined as the reduction in rmse when feature X is randomly shuffled. Table III shows the detailed calculation algorithm.

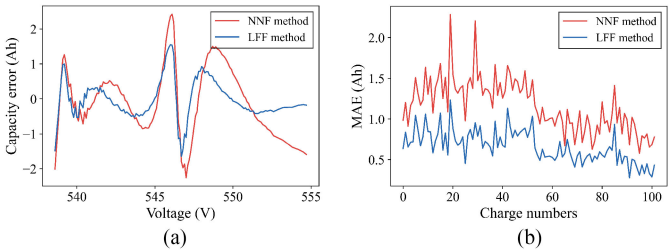


Fig. 14. Comparison of the LFF and NNF methods fitted results. (a) Fitted error in one charging fragment. (b) MAE for all charging fragments.

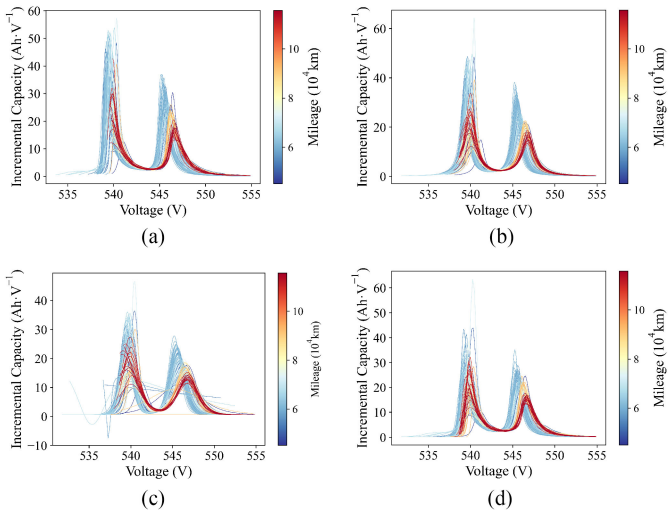


Fig. 15. IC curves during the degradation of a research vehicle. (a) NDGF method. (b) LFF method. (c) NNF method. (d) PCHIP method.

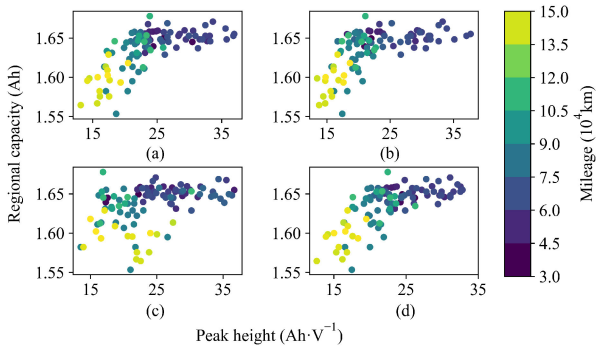


Fig. 16. Relationship between SOH and the peak height of the IC curve.

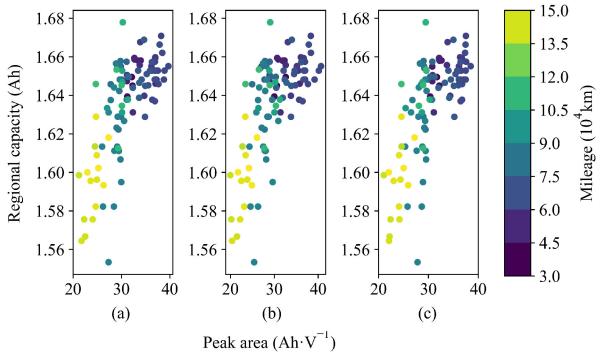


Fig. 17. Relationship between SOH and the peak area of the IC curve.

VI. Comparison Results and Discussion

A. Applications on Different Aging Statuses

1) Robustness and Effectiveness Comparison Results: The fitted accuracy comparison results for LFF and NNF methods are illustrated in Fig. 14. The MAE for the LFF method is 0.43Ah, 44% less than the NNF method of 0.77Ah in one charging fragment, as shown in Fig. 14(a). The maximum MAE of the two methods is 2.2 Ah, and the MAE of the NNF method is higher than the LFF method on all charging fragments, as shown in Fig. 14(b). It can be concluded that the LFF method provides a more accurate description of charging capacity versus voltage. It is worth noting that different from the fitting method, the interpolation function has the same charged capacity Q_i as the original data q_i, as proved in (19).

Fig. 15 demonstrates the evolution of IC curves determined by the abovementioned methods. It can be observed that all the IC curves show similar shapes with two distinct peaks centered at around 540 and 545.5 V. The evolution of the IC peaks with mileage increasing can be clearly observed, which contains insightful information associated with battery aging. Specifically, the reduction in the height of the second IC peak results from the loss of active material (LAM) and LLI, which is the main degradation mode during the early period [34]. And, the position of the second peak has a moving tendency to the high voltage, which is the result of increasing internal resistance [41]. In addition, it is worth noting that there are some unreasonable curves in the IC curves obtained by the NNF method, which are the result of underfitting occurring on the charging fragments with insufficient data [42].

The times taken by the four methods to generate all IC curves for one studied vehicle are 34, 105, 59, and 38 s, respectively. The results show that generating all the IC curves for one vehicle using the PCHIP saves 67% of the time compared to the LFF and NNF methods.

In summary, compared with the NDGF method, the PCHIP method has the same low computational demand while obtaining quantitative IC features accurately. Compared with the PFF method, the PCHIP method retains the ability to fit the CV curve with high accuracy while avoiding overfitting. Compared with the LFF and NNF methods, the PCHIP method can precisely capture the IC peaks while eliminating the tedious parameter identification process.

2) Feature Correlation and Importance Comparison Results: Figs. 16–18 show the correlations between IC features and SOH for one studied vehicle, and the SRA coefficient is attached in the figures. Besides, the mileage of the vehicle is indicated by color. For this study vehicle, the mileage increased from 30 000 to 150 000 km over the time covered by the dataset. It can be seen that as the mileage increases, the regional capacity shows an overall decreasing trend, indicating that the SOH decreases as the battery ages. In addition, the three kinds of IC features

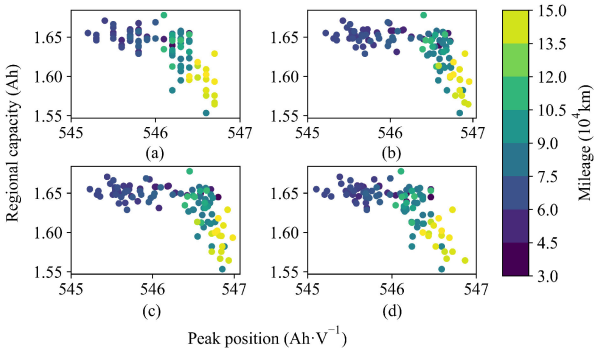


Fig. 18. Relationship between SOH and the peak position of the IC curve.

TABLE IV
CORRELATION OF IC FEATURES OBTAINED BY DIFFERENT METHODS WITH SOH

Methods	Features	F-test	ρ	GRA	MI
NDGF method	Height	0.436	0.635	0.809	0.783
	Position	0.748	−0.645	0.669	0.731
LFF method	Height	0.414	0.588	0.796	0.632
	Area	0.869	0.659	0.800	0.984
	Position	0.668	−0.603	0.653	0.690
NNF method	Height	0.218	0.614	0.802	0.467
	Area	0.623	0.689	0.804	0.713
	Position	0.655	−0.610	0.651	0.709
PCHIP method	Height	0.386	0.641	0.798	0.642
	Area	1.000	0.701	0.799	1.000
	Position	0.688	−0.623	0.675	0.705

have a clear tendency with the SOH decrease. Generally, the peak height and area decrease with the SOH decrease, while the peak position increases with the SOH decrease.

Table IV gives the correlation coefficients between SOH and the 11 IC features. For direct comparison, the computational results of the F-test and MI are normalized. It can be concluded that all the IC features show a good correlation with the SOH. The results explicitly show that the IC features extracted from the LFF and PCHIP methods have a stronger correlation with SOH.

On the other hand, it is seen clearly in Table IV that all four correlation coefficients between peak area and SOH are larger than the other two kinds of IC features, suggesting the peak area has a closer relationship with SOH. This is because the peak area represents the charged capacity in the phase transition over a specific range of voltages [24], which is conceptually closer to the SOH, as defined above. Thanks to the PCHIP method having the minimum error in fitting the charged capacity, the peak area extracted from the PCHIP method has the strongest correlation with SOH.

The FIs of all the 11 IC features in all models are given in Table V. It can be observed that the FI summation of the IC feature extracted from the PCHIP method is larger than other methods in SVR, RF, CART, and LR models, indicating that the IC features extracted by the PCHIP method have a more substantial influence on the performance of these models. In

TABLE V
FIS OF ALL FEATURES IN THE FOUR MODELS

Methods	Features	SVR	RF	CART	LR
NDGF method	Height	0.368	0.775	0.368	0.209
	Positon	0.302	0.054	0.061	0.220
	Total	0.670	0.829	0.429	0.429
LFF method	Height	0.181	0.083	0	0
	Area	0.474	0.133	0.010	0.478
	Positon	0.098	0.087	0.111	0.082
	Total	0.753	0.304	0.122	0.560
NNF method	Height	0.235	0.110	0.128	0.040
	Area	0.308	0.083	0.248	0.095
	Positon	0	0.158	0.018	0.036
	Total	0.542	0.351	0.394	0.179
PCHIP method	Height	0.184	0	0.052	0.233
	Area	1.000	1.000	1.000	0.353
	Positon	0.328	0.019	0.028	0.136
	Total	1.512	1.019	1.080	0.722

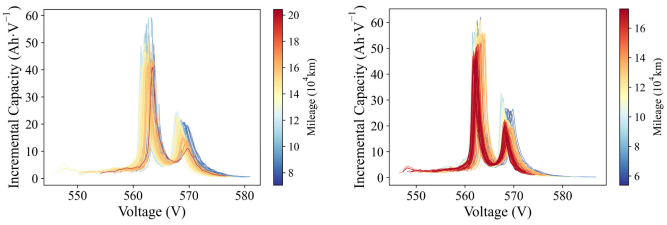


Fig. 19. IC curves for two Type B vehicles.

addition, consistent with the results of the correlation analysis, the peak area also shows the highest importance compared to peak height and position, suggesting that the peak area is a promising candidate for SOH prediction.

B. Application on Different Research Vehicles

With the abovementioned comprehensive investigation, the robustness and feasibility of the different methods have been verified and compared when they are applied to one studied vehicle. However, the batteries of EVs have a variety of degradation trajectories due to the differences in driver driving habits, traffic conditions, and environmental climate. It is worth investigating and comparing the universality of different methods based on different vehicles to draw solid conclusions. In this section, four IC curve determination methods are applied to 40 Evs in two types with various charging strategies. It is worth noting that the CC charging current is not fixed for different vehicles in practical applications. Therefore, the same data preprocessing procedure shown in Section II is applied to obtain the commonly charging current and SOC range of different EVs.

Fig. 19 shows the IC curves for two type B vehicles at different aging statuses. It is depicted that the IC curves generated by the PCHIP method are smooth along with the degradation process and the two IC peaks are clearly identifiable.

Furthermore, the computational efficiency and accuracy of the four IC determination methods based on all 40 vehicles are evaluated. The methods are all implemented on a personal computer with an Intel Core i7-6700HQ CPU running at 2.60 GHz

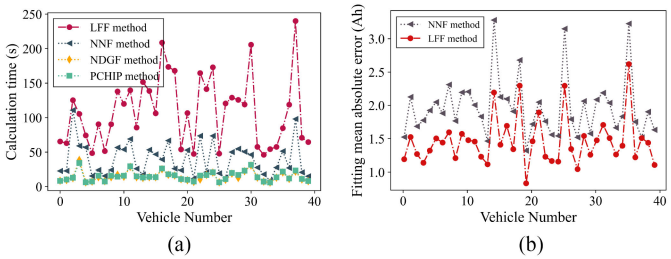


Fig. 20. Calculation time and fitting error for all vehicles. (a) Calculation time. (b) Fitting error.

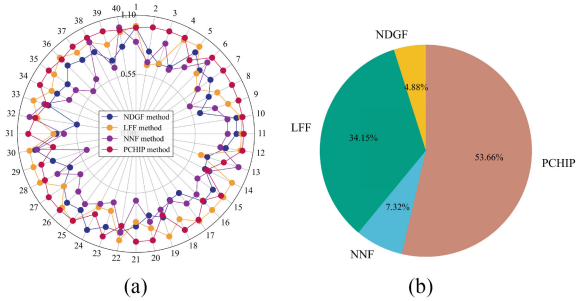


Fig. 21. Correlation comparison results for all vehicles. (a) Mean value of correlation coefficients (b) Percentage of method with the highest correlation.

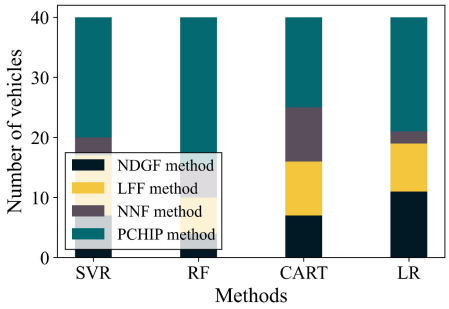


Fig. 22. FI analysis results for all vehicles.

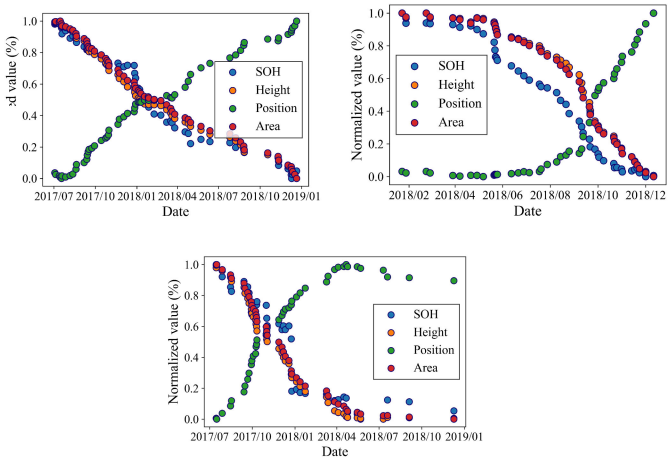


Fig. 23. SOH and IC features for different degradation trajectories.

and 16 GB of RAM. The results in Fig. 20 show that the PCHIP method and the NDGF method have a considerable advantage in computation efficiency, where the PCHIP method compared to the LFF method saving 88% of the time. Apart from this, the computation time of the LFF method applied to different EVs fluctuates drastically. In contrast, the computation time of the PCHIP method stably maintains around 10 s for all EVs, which demonstrates the high robustness of the PCHIP method to different degradation modes.

Finally, the correlation and FI analysis on all studied vehicles are performed. Fig. 21 shows the results of correlation analysis for 40 EVes. For each vehicle, the four points highlighted by different colors represent the average of four correlation coefficients between IC features and SOH. For 53.66% studied vehicles, it can be seen that the IC features extracted from the PCHIP method have the strongest correlation with SOH.

Similarly, following the same procedure suggested in Table III, the FI of the IC features extracted from different methods in the four models is analyzed. The average FI of the features extracted from the same IC curve is used to evaluate the performance of different IC curve determination methods. The statistical results are shown in Fig. 22. It can be seen that for the SVR, RF, CART, and LR models, the PCHIP method has the highest FI on 20, 25, 15, and 19 vehicles, respectively. On average, the IC features extracted by the PCHIP method slightly outperformed the other three methods.

The results of the comparison highlight that the features extracted from the PCHIP method have a stronger correlation with the SOH on most vehicles and exhibit higher importance in different models. Furthermore, the results also reveal the challenge of constructing a model with good generalizability to SOH estimates of a large scale of EVs, demonstrating the significance of correlation analysis and FI analysis in the data-driven approach.

C. Application to Different SOH Estimation Models

The degradation modes of Li-ion batteries are commonly identified as LLI, LAM, ohmic resistance increase (ORI), and lithium plating [19], [34]. Due to the various operational conditions, the battery degradation modes are different in real-world situations, resulting in different capacity degradation trajectories. Fig. 23 shows the normalized SOH and IC features of three typical degradation trajectories. It can be seen that the three IC features evolute regularly as the battery aging, and the evolutions of IC features are different for vehicles with different degradation trajectories. The degradation mode of the battery can be inferred by analyzing the evolution of IC features. Specifically, the increase of peak position is the result of ORI; the reduction of peak area and height indicates LLI. Initial peak area and height increase, or initial steady evolution represents LAM [34].

To further validate the performance of the PCHIP method, the same machine learning algorithms used in the FI analysis are applied to estimate the SOH of batteries with different capacity degradation trajectories. The input features are the peak height,

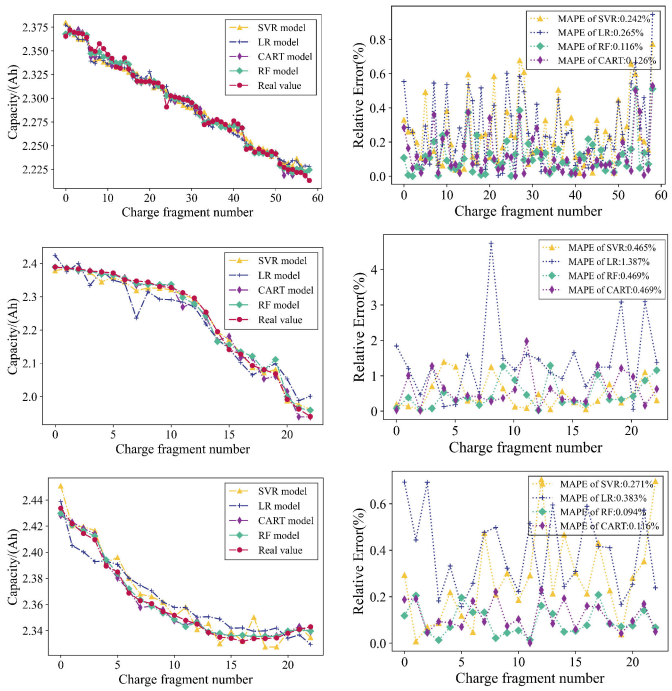


Fig. 24. SOH estimation results for different degradation trajectories.

TABLE VI
RANGE OF PARAMETERS FOR ONE STUDIED VEHICLE

Vehicle 1	RMSE	Vehicle 2	RMSE	Vehicle 3	RMSE
SVR	0.26%	SVR	0.57%	SVR	0.32%
LR	0.15%	LR	1.61%	LR	0.41%
RF	0.10%	RF	0.65%	RF	0.11%
CART	0.08%	CART	0.77%	CART	0.14%

position, and area extracted from IC curves determined by the PCHIP method. Simultaneously, the temperature and mileage factors are carefully considered. As with most data-driven research, 70% of data are used as the training set for each vehicle, while the remaining data are used as the testing set [43]. The parameters of each model are determined using a grid search method with cross-validation. The detailed principles and procedures of these algorithms can be found in [43]. In addition, rmse and mean absolute percentage error (MAPE) are introduced as the metrics to evaluate the performance of the models.

The SOH estimation results obtained on three vehicles with different degradation trajectories are shown in Fig. 24. It can be seen that the estimated SOH is very approximate to the actual SOH. The results in Table VI give that the rmses of the estimated SOHs are well-confined between 0.08% and 1.61%. Specifically, for linear degradation, all algorithms can estimate SOH accurately with MAPE within 1%, and the maximum rmse is 0.26%. For nonlinear degradation, the SVR model and the LR model estimation accuracy is inferior to the other models, with the maximum rmse of the LR model being 1.61% and the MAPE of the LR model being 1.38%, but the error is still maintained within an acceptable range. Hence, it is proof that the PCHIP method has promising accuracy and good robustness in estimating the SOH of batteries with different degradation trajectories.

VII. CONCLUSION

Based on partial charging data of 40 EVs operating in the real world, the practicability and limitations of different IC curve determination methods were comprehensively demonstrated in this article. To solve the problems of the existing methods, the PCHIP method was proposed. The results validated the feasibility and generalizability of this method for extracting aging features and estimating SOH under partial and uncertain charging conditions. The main conclusions of this article are as follows.

1) The NDGF method can calculate the discrete IC values, but with realistic data sampled at low frequencies, the IC features could not be accurately obtained due to insufficient data and lack of functional expressions, which affects the further application of the NDGF method for SOH estimation. In addition, the PFF method suffers from an overfitting phenomenon and is challenging to apply in real-world operating conditions.
2) The NNF method can obtain the IC features accurately, but requires sufficient training data to avoid the underfitting problem. In real-world data with lower sampling frequency, there may not be sufficient data for neural network training, resulting in a lack of robustness of the NNF method.
3) The improved LFF method eliminates the need for prior knowledge of batteries obtained from experiments in identifying parameters. However, it is still necessary to determine the function structure based on the number of IC peaks during the charging process. In the real-world partial and uncertain charging process, it is challenging to determine a constant Lorentz function structure. Meanwhile, the computationally intensive parameter identification process disturbs the application of LFF methods.
4) According to the compared results in this article, the PCHIP method outperformed the fitting methods by eliminating the process of determining the function structure for different charging SOC ranges, thus making it more applicable in realistic uncertain and partial charging situations. In addition, this method provides a simple computational process without parameter identification and low sampling frequency requirements, thus improving the efficiency and making it easier to implement in the on-board BMS.

REFERENCES

[1] Z. Lyu, R. Gao, and L. Chen, "Li-ion battery state of health estimation and remaining useful life prediction through a model-data-fusion method," *IEEE Trans. Power Electron.*, vol. 36, no. 6, pp. 6228–6240, Jun. 2021.
[2] S. Wang, C. Fernandez, W. Cao, C. Zou, C. Yu, and X. Li, "An adaptive working state iterative calculation method of the power battery by using the improved Kalman filtering algorithm and considering the relaxation effect," *J. Power Sources*, vol. 428, pp. 67–75, Jul. 2019.
[3] Y. Li *et al.*, "Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review," *Renewable Sustain. Energy Rev.*, vol. 113, Oct. 2019, Art. no. 109254.

[4] L. Lam and P. Bauer, "Practical capacity fading model for Li-ion battery cells in electric vehicles," *IEEE Trans. Power Electron.*, vol. 28, no. 12, pp. 5910–5918, Dec. 2013.
[5] X. Li, L. Zhang, Z. Wang, and P. Dong, "Remaining useful life prediction for lithium-ion batteries based on a hybrid model combining the long short-term memory and Elman neural networks," *J. Energy Storage*, vol. 21, pp. 510–518, Feb. 2019.
[6] S. Wang, S. Jin, D. Deng, and C. Fernandez, "A critical review of online battery remaining useful lifetime prediction methods," *Front. Mech. Eng.*, vol. 7, p. 719718, 2021.
[7] W. Cao, S. Wang, and C. Fernandez, "A novel adaptive state of charge estimation method of full life cycling lithium–ion batteries based on the multiple parameter optimization," *Energy Sci. Eng.*, vol. 7, no. 5, pp. 1544–1566, May 2019.
[8] S. Wang *et al.*, "A novel safety assurance method based on the compound equivalent modeling and iterate reduce particle-adaptive Kalman filtering for the unmanned aerial vehicle lithium ion batteries," *Energy Sci. Eng.*, vol. 8, no. 5, pp. 1484–1500, 2020.
[9] Q. Lin, J. Wang, R. Xiong, W. Shen, and H. He, "Towards a smarter battery management system: A critical review on optimal charging methods of lithium ion batteries," *Energy*, vol. 183, pp. 220–234, Sep. 2019.
[10] X. Hu, J. Jiang, D. Cao, and B. Egardt, "Battery health prognosis for electric vehicles using sample entropy and sparse Bayesian predictive modeling," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2645–2656, Apr. 2016.
[11] Y. Guo, K. Huang, and X. Hu, "A state-of-health estimation method of lithium-ion batteries based on multi-feature extracted from constant current charging curve," *J. Energy Storage*, vol. 36, Apr. 2021, Art. no. 102372.
[12] K. Liu, Y. Shang, Q. Ouyang, and W. D. Widanage, "A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery," *IEEE Trans. Ind. Electron.*, vol. 68, no. 4, pp. 3170–3180, Apr. 2021.
[13] Y. Li *et al.*, "Random forest regression for online capacity estimation of lithium-ion batteries," *Appl. Energy*, vol. 232, pp. 197–210, Dec. 2018.
[14] C. She, Z. Wang, F. Sun, P. Liu, and L. Zhang, "Battery aging assessment for real-world electric buses based on incremental capacity analysis and radial basis function neural network," *IEEE Trans. Ind. Infor.*, vol. 16, no. 5, pp. 3345–3354, May 2020.
[15] Y. Li, D.-I. Stroe, Y. Cheng, H. Sheng, X. Sui, and R. Teodorescu, "On the feature selection for battery state of health estimation based on charging–discharging profiles," *J. Energy Storage*, vol. 33, Jan. 2021, Art. no. 102122.
[16] L. Zheng, J. Zhu, D. D.-C. Lu, G. Wang, and T. He, "Incremental capacity analysis and differential voltage analysis based state of charge and capacity estimation for lithium-ion batteries," *Energy*, vol. 150, pp. 759–769, May 2018.
[17] M. Dubarry and B. Liaw, "Identify capacity fading mechanism in a commercial LiFePO4 cell," *J. Power Sources*, vol. 194, no. 1, pp. 541–549, 2009.
[18] M. Dubarry, C. Truchot, and B. Y. Liaw, "Synthesize battery degradation modes via a diagnostic and prognostic model," *J. Power Sources*, vol. 219, pp. 204–216, Dec. 2012.
[19] M. Dubarry, M. Berecibar, A. Devie, D. Anseán, N. Omar, and I. Villarreal, "State of health battery estimator enabling degradation diagnosis: Model and algorithm description," *J. Power Sources*, vol. 360, pp. 59–69, Aug. 2017.
[20] Y. Li *et al.*, "A quick on-line state of health estimation method for li-ion battery with incremental capacity curves processed by gaussian filter," *J. Power Sources*, vol. 373, pp. 40–53, Jan. 2018.
[21] B. Jiang, H. Dai, and X. Wei, "Incremental capacity analysis based adaptive capacity estimation for lithium-ion battery considering charging condition," *Appl. Energy*, vol. 269, Jul. 2020, Art. no. 115074.
[22] B. O. Agudelo, W. Zamboni, and E. Monmasson, "Application domain extension of incremental capacity-based battery SoH indicators," *Energy*, vol. 234, Nov. 2021, Art. no. 121224.
[23] X. Han *et al.*, "A comparative study of charging voltage curve analysis and state of health estimation of lithium-ion batteries in electric vehicle," *Automot. Innov.*, vol. 2, no. 4, pp. 263–275, Dec. 2019.
[24] X. Li, J. Jiang, L. Y. Wang, D. Chen, Y. Zhang, and C. Zhang, "A capacity model based on charging process for state of health estimation of lithium ion batteries," *Appl. Energy*, vol. 177, pp. 537–543, Sep. 2016.
[25] S. Torai, M. Nakagomi, S. Yoshitake, S. Yamaguchi, and N. Oyama, "State-of-health estimation of LiFePO4/graphite batteries based on a model using differential capacity," *J. Power Sources*, vol. 306, pp. 62–69, Feb. 2016.
[26] C. Weng, Y. Cui, J. Sun, and H. Peng, "On-board state of health monitoring of lithium-ion batteries using incremental capacity analysis with support vector regression," *J. Power Sources*, vol. 235, pp. 36–44, Aug. 2013.
[27] X. Bian, Z. Wei, W. Li, J. Pou, D. U. Sauer, and L. Liu, "State-of-health estimation of lithium-ion batteries by fusing an open circuit voltage model and incremental capacity analysis," *IEEE Trans. Power Electron.*, vol. 37, no. 2, pp. 2226–2236, Feb. 2022.
[28] J. He, S. Meng, X. Li, and F. Yan, "Partial charging-based health feature extraction and state of health estimation of lithium-ion batteries," *IEEE J. Emerg. Sel. Topics Power Electron.*, early access, Jan. 20, 2022, doi: 10.1109/JESTPE.2022.3143831.
[29] X. Li, Z. Wang, L. Zhang, C. Zou, and D. D. Dorrell, "State-of-health estimation for li-ion batteries by combing the incremental capacity analysis method with grey relational analysis," *J. Power Sources*, vol. 410–411, pp. 106–114, Jan. 2019.
[30] A. K. Qin, V. L. Huang, and P. N. Suganthan, "Differential evolution algorithm with strategy adaptation for global numerical optimization," *IEEE Trans. Evol. Comput.*, vol. 13, no. 2, pp. 398–417, Apr. 2009.
[31] J. I. Qiang and C. Mitchell, "A unified differential evolution algorithm for global optimization," United States: N. p., 2014. [Online]. Available: https://www.osti.gov/servlets/purl/1163659
[32] C. de Boor, K. Höllig, and M. Sabin, "High accuracy geometric hermite interpolation," *Comput. Aided Geometric Des.*, vol. 4, no. 4, pp. 269–278, Dec. 1987.
[33] J. Wang, Y. Li, H. Zhu, and T. Ma, "Interpolation method research and precision analysis of GPS satellite position," *J. Syst. Sci. Inf.*, vol. 6, no. 3, pp. 277–288, Jun. 2018.
[34] D. Anseán *et al.*, "Lithium-ion battery degradation indicators via incremental capacity analysis," *IEEE Trans. Ind. Appl.*, vol. 55, no. 3, pp. 2992–3002, May/Jun. 2019.
[35] A. El Mejdoubi, H. Chaoui, H. Gualous, P. van den Bossche, N. Omar, and J. van Mierlo, "Lithium-ion batteries health prognosis considering aging conditions," *IEEE Trans. Power Electron.*, vol. 34, no. 7, pp. 6834–6844, Jul. 2019.
[36] L. Song, K. Zhang, T. Liang, X. Han, and Y. Zhang, "Intelligent state of health estimation for lithium-ion battery pack based on Big Data analysis," *J. Energy Storage*, vol. 32, Dec. 2020, Art. no. 101836.
[37] P. M. Attia *et al.*, ""Knees" in lithium-Ion battery aging trajectories," *J. Electrochem. Soc.*, 2022, *arXiv:2201.02891*.
[38] M. L. Tiku, "Tables of the power of the F-test," *J. Amer. Stat. Assoc.*, vol. 62, no. 318, pp. 525–539, Jun. 1967.
[39] S. Zhang, B. Zhai, X. Guo, K. Wang, N. Peng, and X. Zhang, "Synchronous estimation of state of health and remaining useful lifetime for lithium-ion battery using the incremental capacity and artificial neural networks," *J. Energy Storage*, vol. 26, Dec. 2019, Art. no. 100951.
[40] Y. Zhao, Z. Wang, Z.-J. M. Shen, and F. Sun, "Data-driven framework for large-scale prediction of charging energy in electric vehicles," *Appl. Energy*, vol. 282, Jan. 2021, Art. no. 116175.
[41] W. Pan, X. Luo, M. Zhu, J. Ye, L. Gong, and H. Qu, "A health indicator extraction and optimization for capacity estimation of li-ion battery using incremental capacity curves," *J. Energy Storage*, vol. 42, Oct. 2021, Art. no. 103072.
[42] S. Wang, S. Jin, D. Bai, Y. Fan, H. Shi, and C. Fernandez, "A critical review of improved deep learning methods for the remaining useful life prediction of lithium-ion batteries," *Energy Rep.*, vol. 7, pp. 5562–5574, Nov. 2021.
[43] M.-F. Ng, J. Zhao, Q. Yan, G. J. Conduit, and Z. W. Seh, "Predicting the state of charge and health of batteries using data-driven machine learning," *Nat. Mach. Intell.*, vol. 2, no. 3, pp. 161–170, Mar. 2020.

Peng Liu received the Ph.D. degree in mechanical engineering from the Beijing Institute of Technology, Beijing, China, in 2011.

He is currently an Associate Professor with the School of Mechanical Engineering, Beijing Institute of Technology. His research interests include Big Data analysis for power battery systems and real-world electrical vehicles.

Yizhong Wu received the B.S. degree in vehicle engineering in 2020 from the Beijing Institute of Technology, Beijing, China, where he is currently working toward the master's degree in mechanical engineering with the National Engineering Laboratory for Electric Vehicles.

His research focuses on state of health estimation of lithium-ion batteries based on real-world Big Data.

Chengqi She received the Ph.D. degree in mechanical engineering from the Beijing Institute of Technology, Beijing, China, in 2022.

He is currently an Assistant Professor with the Hunan Provincial Key Laboratory of Health Maintenance for Mechanical Equipment, Hunan University of Science and Technology, Xiangtan, China. His research interests include modeling and state of health estimation of lithium-ion batteries based on real-world Big Data.

Zhenpo Wang received the Ph.D. degree in automotive engineering from the Beijing Institute of Technology, Beijing, China, in 2005.

He is currently a Professor with the Beijing Institute of Technology, where he is also the Director of the National Engineering Laboratory for Electric Vehicles. He has authored or coauthored four monographs and translated books as well as more than 80 technical papers. He also holds more than 60 patents. His research interests include pure electric vehicle integration, packaging and energy management of battery systems, and charging station design.

Prof. Wang was the recipient of numerous awards, including the second National Prize for Progress in Science and Technology and the first Prize for Progress in Science and Technology from the Ministry of Education, China, and the second Prize for Progress in Science and Technology from Beijing Municipal, China.

Zhaosheng Zhang received the Ph.D. degree in automotive engineering from Tsinghua University, Beijing, China, in 2013.

He is currently a Lecturer with the School of Mechanical Engineering, Beijing Institute of Technology, Beijing, China. He has authored or coauthored two monographs and more than 20 technical papers. He holds more than 10 patents. His research interests include intelligent transportation and Big Data analysis.