Feature selection strategy optimization for lithium-ion battery state of health estimation under impedance uncertainties

Xinghao Du, Jinhao Meng, Yassine Amirat, Fei Gao, Mohamed Benbouzid

Journal of Energy Chemistry 101 (2025) 87–98

Contents lists available at ScienceDirect

Journal of Energy Chemistry

journal homepage: www.elsevier.com/locate/jechem

http://www.journals.elsevier.com/journal-of-energy-chemistry/

Feature selection strategy optimization for lithium-ion battery state of health estimation under impedance uncertainties

Xinghao Du [a], Jinhao Meng [b,*], Yassine Amirat [c], Fei Gao [d], Mohamed Benbouzid [a,*]

[a] UMR CNRS 6027 IRDL, University of Brest, Brest 29238, France
[b] School of Electrical Engineering, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, China
[c] L@bISEN, ISEN Yncréa Ouest, Brest 29200, France
[d] School of Energy and Computer Science, University of Technology of Belfort-Montbeliard, Belfort 90000, France

ARTICLE INFO

Article history:
Received 27 July 2024
Revised 31 August 2024
Accepted 2 September 2024
Available online 27 September 2024

Keywords:
Lithium-ion battery
State of health
Electrochemical impedance spectroscopy
Extreme gradient boosting

ABSTRACT

Battery health evaluation and management are vital for the long-term reliability and optimal performance of lithium-ion batteries in electric vehicles. Electrochemical impedance spectroscopy (EIS) offers valuable insights into battery degradation analysis and modeling. However, previous studies have not adequately addressed the impedance uncertainties, particularly during battery operating conditions, which can substantially impact the robustness and accuracy of state of health (SOH) estimation. Motivated by this, this paper proposes a comprehensive feature optimization scheme that integrates impedance validity assessment with correlation analysis. By utilizing metrics such as impedance residuals and correlation coefficients, the proposed method effectively filters out invalid and insignificant impedance data, thereby enhancing the reliability of the input features. Subsequently, the extreme gradient boosting (XGBoost) modeling framework is constructed for estimating the battery degradation trajectories. The XGBoost model incorporates a diverse range of hyperparameters, optimized by a genetic algorithm to improve its adaptability and generalization performance. Experimental validation confirms the effectiveness of the proposed feature optimization scheme, demonstrating the superior estimation performance of the proposed method in comparison with four baseline techniques.

1. Introduction

The rapid development of battery technology has deeply impacted numerous sectors, including electric vehicles (EVs), energy storage systems (ESSs), and portable electronics. Lithium-ion batteries (LIBs), serving as the primary power source in these domains, are characterized by their high energy and power density, extended lifespan, and absence of memory effects [1]. Nonetheless, LIBs are prone to performance degradation with battery aging, resulting in unforeseen malfunctions and diminished operational efficiency. To mitigate such risks and ensure consistent performance, battery management systems (BMSs) have been widely implemented. A crucial aspect of effective BMS implementation lies in accurately assessing the battery's state of health (SOH), enabling proactive maintenance, optimizing energy utilization, and preventing potential safety risks [2].

1.1. Literature review

Battery SOH estimation methods prevailing in the literature can be broadly classified into model-based and data-driven methods, detailed as follows.

Model-based methods for estimating battery SOH mainly include empirical models [3], electrochemical models [4], and equivalent circuit models (ECMs) [5]. Empirical models use mathematical functions, such as polynomials or exponential equations, to describe the quantitative relationship between various stress factors and the battery aging trajectories. These models can offer straightforward predictions of battery lifetime and performance, making them useful for initial assessments. However, their predictions are highly specific to the battery's chemistry and operating conditions, leading to poor accuracy when applied to different battery types and testing scenarios [6]. Consequently, empirical models cannot generalize well across diverse battery systems, limiting their broader applicability. Electrochemical models utilize partial differential equations (PDEs) to capture the complex dynamics within a battery, encompassing electrode reactions, lithium-ion

* Corresponding authors.
E-mail addresses: jinhao@xjtu.edu.cn (J. Meng), Mohamed.Benbouzid@univ-brest.fr (M. Benbouzid).

https://doi.org/10.1016/j.jechem.2024.09.032

transport, and the formation and decomposition of the solid electrolyte interphase (SEI) film. These models offer profound physical insights into battery behavior and degradation mechanisms. However, they demand comprehensive information on microscopic parameters and substantial computational resources to simulate these processes accurately. This high level of detail and computational intensity can be a significant barrier, making it challenging to apply these models practically in real-time BMSs [7]. ECMs, on the other hand, represent the battery with an electrical circuit composed of resistors, capacitors, and voltage sources to characterize the battery's dynamic behavior. These models use state-space equations to estimate internal resistance and battery states, balancing complexity and computational efficiency. However, ECM-based methods are sensitive to noise disturbances and require prior knowledge of the battery's state of charge (SOC), limiting their accuracy and reliability [8].

In recent studies, data-driven approaches utilizing machine learning algorithms have significantly progressed and found extensive applications in battery health evaluation. Numerous estimation models have been developed using historical data to characterize battery degradation patterns and predict aging status. Techniques such as Gaussian process regression (GPR) [9], support vector regression (SVR) [10], random forest (RF) [11], extreme learning machine (ELM) [12], long short-term memory (LSTM) networks [13,14], and deep neural networks (DNN) [15] have been employed for battery health assessment. These methods rely heavily on well-extracted features that correlate strongly with battery health to ensure accurate modeling [16]. Depending on the types of features selected, prevalent data-driven estimation approaches can be classified into three subgroups.

- *Operational data-based feature:* Features derived from operational data, such as charging and discharging curves, have been widely utilized as inputs for battery SOH estimation. Wang *et al.* [17] propose the use of LSTM networks combined with an adaptive state parameter feedback correction strategy to predict the battery remaining useful life (RUL), which accounts for variations in current rates and ambient temperature changes, thereby improving prediction accuracy. Furthermore, identifying peak shifts and intensities in incremental capacity (IC) and differential voltage (DV) curves facilitates pinpointing battery degradation mechanisms and estimating the battery's capacity loss. For instance, Li *et al.* [18] utilize a grey forecasting model to extract health indicators from IC analysis and propose an iterative and probabilistic prediction method combined with GPR for battery health prognosis. Xia *et al.* [19] combine IC and DV analysis with a second-order ECM to enhance the accuracy of battery SOH and RUL predictions. However, gathering reliable IC and DV curves necessitates charging or discharging at very low current rates, which limits their practical application at higher current rates.
- *Statistical characteristics-based feature:* These methods generally involve extracting statistical metrics from operational data and using these metrics to model and predict battery degradation. They are particularly effective at capturing the nonlinear relationships between statistical features and battery capacity loss, providing a nuanced understanding of battery degradation mechanisms. For example, Deng *et al.* [20] utilize the average and standard deviation of the random capacity increment sequence extracted from partial charging processes as inputs to improve the modeling accuracy and reduce uncertainties. Similarly, Hu *et al.* [21] use the sample entropy of short voltage sequences to reflect the capacity loss and employ sparse Bayesian predictive modeling to develop a comprehensive data-driven battery SOH estimator.
- *Impedance-based feature*: Compared to features obtained directly from voltage and current measurements, electrochemical impedance spectroscopy (EIS) offers comprehensive insights into electrochemical processes and material properties over a wide range of frequencies, thereby uncovering the battery aging mechanisms [22,23]. To illustrate, Zhang *et al.* [9] utilize impedance spectra collected at various temperatures as input features for a GPR model to characterize battery degradation patterns. Likewise, Xia *et al.* [24] extract geometric impedance health indicators from the Nyquist plot of EIS data collected in the high- and medium-frequency ranges and develop a recurrent GPR method to estimate the battery SOH. An analytical calculation model is proposed in [25] to convert the charge transfer resistance at various SOC and temperatures to a standard state, making it more comparable and reliable for SOH estimation.

To construct a comprehensive degradation model, extracting features from impedance spectra is highly desirable. Extensive research has been conducted on designing perturbation signal injection devices that facilitate impedance measurements in onboard applications [26,27], with significant progress made in enhancing the efficacy of impedance acquisition [22,28]. These advancements have streamlined the process of impedance feature extraction, making it more practical for real-time implementation. However, employing EIS data for battery SOH estimation presents significant challenges due to the spectrum's high dimensionality, capturing both real and imaginary impedance parts over a broad frequency range. This necessitates advanced data-processing techniques to effectively distill meaningful information. Zhang *et al.* [29] and Messing *et al.* [30] convert the impedance spectrum into lower-dimensional features by fitting through ECMs. However, ECMs are often tailored to specific battery chemistries, which can limit their applicability across diverse types of batteries. Additionally, the fitting processes for ECMs frequently yield non-unique results, introducing uncertainties that affect the SOH estimation reliability. Obregon *et al.* [31] propose the use of a convolutional autoencoder to extract features from EIS data, and a DNN is further developed for battery SOH estimation. Despite their effectiveness, DNNs are susceptible to overfitting, potentially compromising their reliability in practical applications. Xiong *et al.* [32] propose a convolutional neural network (CNN) method for battery capacity estimation by taking raw impedance measurements. An input reconstruction module is developed to exploit unlabeled impedance spectra, thus reducing the costs associated with training data collection. However, it is crucial to ensure that EIS data is validated before interpretation and application. Without proper validation, the reliability of battery health evaluations remains questionable, as inaccuracies in impedance data can lead to erroneous assessments and decisions [33].

1.2. Research gaps and contributions

It can be deduced from preceding studies that EIS contains valuable information reflecting battery aging dynamics over various time scales. Extensive research efforts have been dedicated to interpreting battery impedance spectra and extracting informative features to identify battery degradation patterns. However, current studies inadequately address the uncertainties inherent in impedance measurements. Factors such as temperature fluctuations, noise disturbances, and variations in operating conditions introduce significant variability in impedance data [22], compromising the reliability of battery health assessments. Currently, there is a notable research gap in the development of effective methods to counteract these uncertainties and construct a robust SOH estimator.

Motivated by this, this paper proposes an innovative and comprehensive feature optimization scheme integrating impedance validity and correlation analysis. Subsequently, the battery SOH estimation framework is constructed by proposing a genetic algorithm-based extreme gradient boosting (GA-XGBoost). The main contributions of this paper are as follows:

1. The feature optimization scheme effectively addresses impedance measurement uncertainties. By applying metrics such as impedance residuals and correlation coefficients, invalid and insignificant impedance data are systematically filtered out of the training dataset. This selective process not only elevates the quality of the input data but also substantially increases the robustness of the overall modeling framework, leading to more accurate and reliable SOH estimation results.
2. To capture the intricate battery degradation dynamics across various operating conditions, the XGBoost modeling framework is developed for battery SOH estimation. The genetic algorithm is applied to iteratively search for optimal hyperparameters with five-fold cross-validation, significantly improving the model's adaptability and generalization performance.

The remainder of this paper is organized as follows: The impedance dependencies and uncertainties are analyzed in Section 2. Section 3 details the feature optimization and SOH estimation framework. Section 4 presents the experimental studies. Section 5 summarizes this paper.

2. Electrochemical impedance spectroscopy

EIS encompasses valuable information that can be used for analyzing and interpreting complex battery degradation patterns. To establish a comprehensive and reliable battery degradation model using impedance measurements, it is essential to examine impedance dependencies and uncertainties.

2.1. Impedance dependencies analysis

The schematic of impedance dependencies on battery degradation is presented in Fig. 1. In the kilohertz region, the battery impedance variation is correlated with electrolyte decomposition and solid electrolyte interface (SEI) film formation. The impedance spectrum measured in the Hertz region reflects the deterioration of the active material in the cathode. Furthermore, structural irregularities in the electrode are typically observed in the millihertz region [34].

As shown in Fig. 1, during the battery degradation process, the impedance spectrum exhibits a rightward shift due to the decrease in electrolyte conductivity. As charge-transfer reactions slow down, the medium-frequency semicircle enlarges, with the Warburg straight line moving upward and to the right. The impedance data measured in different frequency regions exhibit distinctive variation trends with the deterioration of battery capacity. By analyzing the frequency response characteristics, it is feasible to gain valuable insights into the underlying degradation mechanisms for batteries.

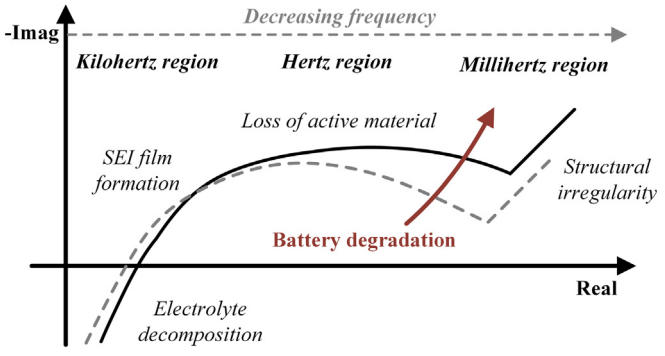

Fig. 1. Schematic of impedance dependencies on battery degradation.

2.2. Impedance uncertainties analysis

The effectiveness of utilizing impedance data for analyzing battery degradation heavily relies on the accuracy and reliability of impedance measurements. Therefore, it is crucial to evaluate and address uncertainties associated with impedance measurement data thoroughly. Suppose a sinusoidal perturbation signal, denoted as $i(t)$, is superimposed onto the load current during the battery charging or discharging. The corresponding voltage response signal, expressed as $u(t)$, can be obtained. $i(t)$ and $u(t)$ are expressed as:

$$\begin{cases} i(t) = |I(\omega_k)| \sin(\omega_k t + \psi) \\ u(t) = |U(\omega_k)| \sin(\omega_k t + \varphi) \end{cases} \tag{1}$$

where ω_k represents the angular frequency; $|I(\omega_k)|$ and ψ represent the magnitude and phase angle of the current perturbation signal, respectively; $|U(\omega_k)|$ and φ are the magnitude and phase angle of the voltage response signal, respectively.

By applying Ohm's law, the actual impedance $\boldsymbol{Z}(\omega_k)$ and measured one $\hat{\boldsymbol{Z}}(\omega_k)$ are calculated as:

$$\begin{cases} \boldsymbol{Z}(\omega_k) = \frac{\boldsymbol{U}}{\boldsymbol{I}} = \frac{|U(\omega_k)|e^{j\varphi}}{|I(\omega_k)|e^{j\psi}} \\ \hat{\boldsymbol{Z}}(\omega_k) = \frac{\boldsymbol{U}+\boldsymbol{U}_\mathrm{p}+\boldsymbol{U}_\mathrm{n}}{\boldsymbol{I}+\boldsymbol{I}_\mathrm{p}+\boldsymbol{I}_\mathrm{n}} \end{cases} \tag{2}$$

where $\boldsymbol{I}$ and $\boldsymbol{U}$ represent the complex perturbation current and response voltage signals, respectively; $\boldsymbol{I}_\mathrm{p}$ and $\boldsymbol{U}_\mathrm{p}$ represent the load current and polarization voltage signals; $\boldsymbol{I}_\mathrm{n}$ and $\boldsymbol{U}_\mathrm{n}$ represent the noise current and voltage signals.

To quantify impedance uncertainties, the normalized impedance deviation (NID) between $\boldsymbol{Z}(\omega_k)$ and $\hat{\boldsymbol{Z}}(\omega_k)$ is calculated by:

$$|\Delta \boldsymbol{Z}| = \left|\hat{\boldsymbol{Z}}(\omega_k) - \boldsymbol{Z}(\omega_k)\right| / |\boldsymbol{Z}(\omega_k)| \tag{3}$$

Suppose the battery is operated in the constant current (CC) charging or discharging mode, $\boldsymbol{I}_\mathrm{p}$ can be removed from Eq. (2). By substituting Eq. (2) into Eq. (3), the calculation of $|\Delta \boldsymbol{Z}|$ is reorganized as follows:

$$|\Delta \boldsymbol{Z}| = \left| \underbrace{\frac{\frac{\boldsymbol{I}}{\boldsymbol{I}_\mathrm{n}} - \frac{\boldsymbol{U}}{\boldsymbol{U}_\mathrm{n}}}{\frac{\boldsymbol{U}}{\boldsymbol{U}_\mathrm{n}}\left(\frac{\boldsymbol{I}}{\boldsymbol{I}_\mathrm{n}} + 1\right)}}_{\text{Noise - induced}} + \underbrace{\frac{\boldsymbol{U}_\mathrm{p}}{\boldsymbol{U}} \frac{\boldsymbol{I}}{\boldsymbol{I}+\boldsymbol{I}_\mathrm{n}}}_{\text{Polarization - induced}} \right| \tag{4}$$

As seen in Eq. (4), the impedance deviation is segmented into two components. The noise-induced impedance deviation is contingent upon the signal-to-noise ratio (SNR) of the perturbation and response signals, denoted as $\boldsymbol{I}/\boldsymbol{I}_\mathrm{n}$ and $\boldsymbol{U}/\boldsymbol{U}_\mathrm{n}$, respectively. The polarization effects, on the other hand, are positively correlated to the ratio between $\boldsymbol{U}_\mathrm{p}$ and $\boldsymbol{U}$. Specifically, the polarization voltage hysteresis components are pronounced in the low-frequency region [35], which inevitably introduces uncertainties in impedance measurement results. To ensure an accurate representation of battery degradation, it is crucial to address impedance uncertainties before establishing the estimation model.

3. Proposed SOH estimation method

The schematic framework of the proposed method is presented in Fig. 2. Initially, impedance data across various frequency ranges are collected during the battery degradation processes. Subsequently, a feature optimization scheme is implemented by integrating impedance validity and correlation analysis. Metrics such as impedance residuals and correlation coefficients are utilized to identify and exclude invalid and insignificant impedance features. Finally, the XGBoost modeling framework is constructed for battery SOH estimation to capture the intricate battery degradation dynamics across diverse operating conditions. The hyperparameters are optimized using a genetic algorithm with five-fold cross-validation to enhance the modeling performance. Details regarding the feature optimization scheme and the SOH estimator will be introduced subsequently.

3.1. Impedance data optimization

To enhance the impedance modeling performance and training efficiency, the optimization of impedance data is detailed as follows. Defining $D = \{(X, Y)\}$ as the training dataset, where X denotes the fully observed impedance features and Y denotes the battery SOH, which is defined as the ratio between the battery's present capacity and its nominal capacity. In this study, the nominal capacity is the battery capacity measured during the first charge–discharge cycle, as in [31]. The input vector $X_{l,:}$ is expressed as:

$$X_{l,:} = \left[\hat{\boldsymbol{Z}}'(l,\omega_1),\ldots,\hat{\boldsymbol{Z}}'(l,\omega_K),\hat{\boldsymbol{Z}}''(l,\omega_1),\ldots,\hat{\boldsymbol{Z}}''(l,\omega_K)\right] \tag{5}$$

where l represents the row index of the input vector. $\hat{\boldsymbol{Z}}'(l,\omega_k)$ and $\hat{\boldsymbol{Z}}''(l,\omega_k)$ (k = 1, 2, …, K) represent the real and imaginary impedance data measured at frequency ω_k. K represents the number of impedance measurement data.

3.1.1. Impedance validity analysis

The impedance validity analysis is conducted by analyzing the linear Kramers-Kronig (K-K) relations [33], which quantitively evaluates the relationship between the real and imaginary components of the impedance data. In this work, the impedance data are fitted with a Voigt circuit model (VCM) consisting of n serially connected RC circuits, expressed as $R_i//C_i$, an inductance L, and a resistance R_s, as presented in Fig. 2. The real and imaginary parts of the VCM impedance are calculated by:

$$\begin{cases} \tilde{\boldsymbol{Z}}'(\omega_k) = R_s + \sum_i^n \frac{R_i}{1+(\omega_k\tau_i)^2} \\ \tilde{\boldsymbol{Z}}''(\omega_k) = \sum_i^n \frac{-\omega_k R_i\tau_i}{1+(\omega_k\tau_i)^2} + \omega_k L_0 \end{cases} \tag{6}$$

where ω_k denotes the angular frequencies of the experimental impedance measurement data; $\tilde{\boldsymbol{Z}}'(\omega_k)$ and $\tilde{\boldsymbol{Z}}''(\omega_k)$ denote the real and imaginary parts of the fitted impedance, respectively; τ_i (i = 1, 2, …, n) denotes the time-constants of the i-th RC circuit, which are log-

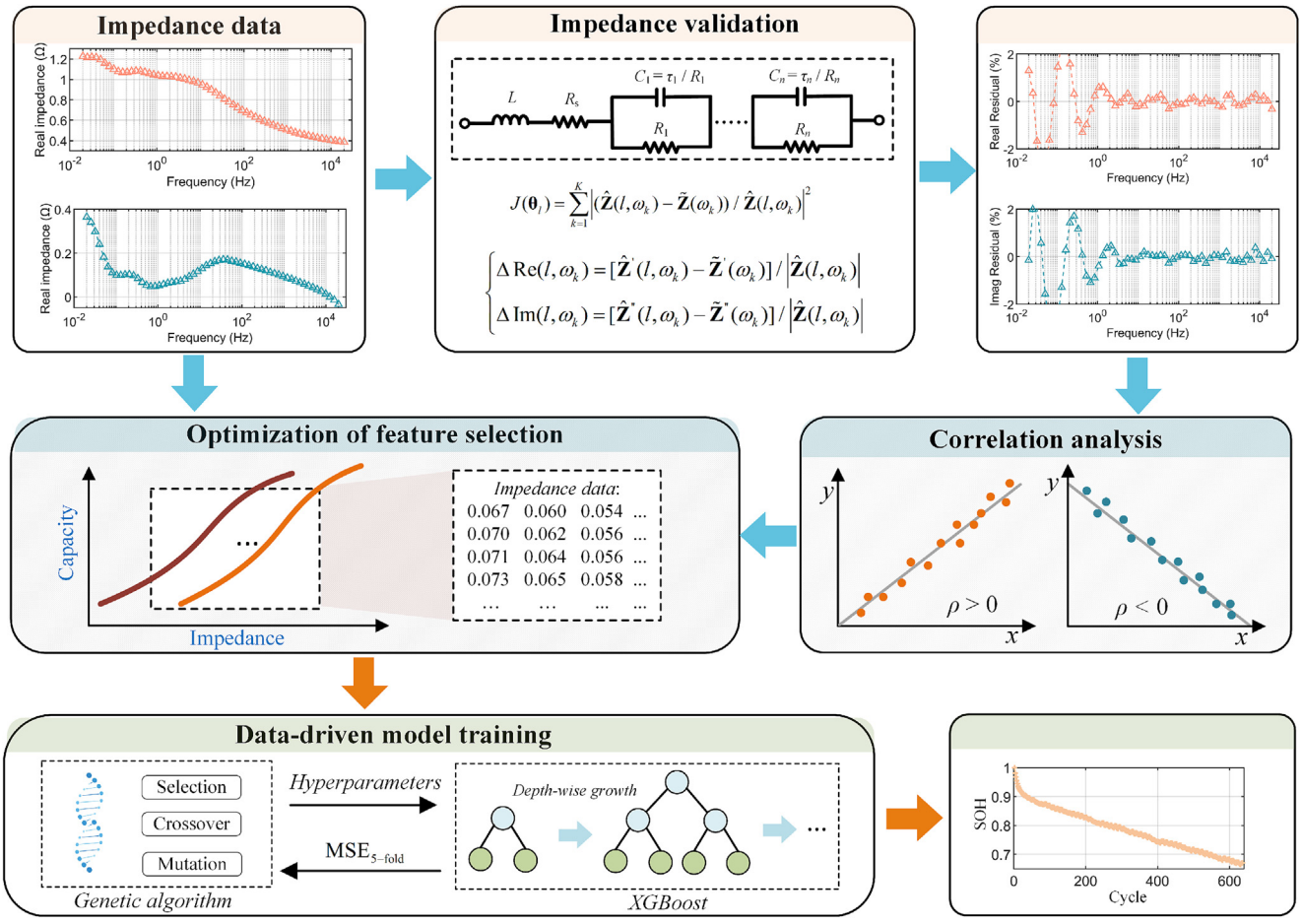

Fig. 2. Schematic framework of the proposed method.

arithmically distributed over the inverse range of angular frequencies of the impedance measurements. τ_i is calculated by:

$$\tau_i = \exp\left[\ln \tau_{\min} + \frac{i-1}{n-1} \ln(\tau_{\max}/\tau_{\min})\right] \tag{7}$$

where $\tau_{\min}$ and $\tau_{\max}$ denote the minimum and maximum time-constants calculated by,

$$\begin{cases} \tau_{\max} = 1/\min(\omega_k) \\ \tau_{\min} = 1/\max(\omega_k) \end{cases} \tag{8}$$

In this work, we employ the Nelder-Mead algorithm [36] to minimize the objective function presented in Eq. (9) and thereby determine the optimal parameter vector. The Nelder-Mead algorithm is a robust and well-established optimization technique, particularly advantageous for addressing complex, non-linear problems. The objective function is expressed as the normalized square deviation between the measured impedance $\hat{\boldsymbol{Z}}(l,\omega_k)$ and the fitted one $\tilde{\boldsymbol{Z}}(\omega_k)$ as follows:

$$\begin{cases} J(\boldsymbol{\theta}_l) = \sum_{k=1}^{K} \left| \left(\hat{\boldsymbol{Z}}(l,\omega_k) - \tilde{\boldsymbol{Z}}(\omega_k)\right)/\hat{\boldsymbol{Z}}(l,\omega_k) \right|^2 \\ \boldsymbol{\theta}_l = [L, C_0, R_s, R_1, ..., R_n] \end{cases} \tag{9}$$

It should be noted that the parameter identification process of the VCM often leads to non-unique solutions, where multiple sets of parameters can adequately describe the same impedance spectrum. To effectively analyze the reproductivity and validity of the impedance measurement data, the residuals between the measured and fitted impedance are calculated as follows:

$$\begin{cases} \Delta\mathrm{Re}(l,k) = \left[\hat{\boldsymbol{Z}}'(l,\omega_k) - \tilde{\boldsymbol{Z}}'(\omega_k)\right]/\hat{\boldsymbol{Z}}(l,\omega_k) \\ \Delta\mathrm{Im}(l,k) = \left[\hat{\boldsymbol{Z}}''(l,\omega_k) - \tilde{\boldsymbol{Z}}''(\omega_k)\right]/\hat{\boldsymbol{Z}}(l,\omega_k) \end{cases} \tag{10}$$

where $\Delta\mathrm{Re}(l,\omega_k)$ and $\Delta\mathrm{Im}(l,\omega_k)$ represent the real and imaginary impedance residuals, respectively. The root mean square (RMS) residuals, serving as the primary metric for impedance feature selection, are determined as follows:

$$\begin{cases} \xi_k = \sqrt{\frac{1}{M}\sum_{l=1}^{M} (\Delta\mathrm{Re}[l,k])^2} \\ \xi_{60+k} = \sqrt{\frac{1}{M}\sum_{l=1}^{M} (\Delta\mathrm{Im}[l,k])^2} \end{cases} \tag{11}$$

where M represents the number of samples. It is important to note that conventional K-K validation tests rely on the assumptions of stationarity and linearity in the system under test [33]. However, these assumptions would not hold under the dynamic conditions typically encountered in battery systems, potentially resulting in the exclusion of valuable data. In contrast, our proposed method focuses on the statistical analysis of impedance residuals across different frequencies, offering a more comprehensive and practical assessment of EIS measurements.

3.1.2. Correlation analysis

The Pearson correlation analysis has been widely recognized as a promising tool for quantifying the intensity of the linear relationship between variables and aiding in feature selection by identifying which features have a strong linear relationship with the objective variable [37]. To select the impedance feature with a strong linear correlation with battery degradation, the Pearson correlation coefficient is calculated as follows:

$$\rho_k = \frac{\sum \left(X_{l,k} - \bar{X_{:,k}}\right)\left(Y_l - \bar{Y}\right)}{\sqrt{\sum \left(X_{l,k} - \bar{X_{:,k}}\right)^2}\sqrt{\sum \left(Y_l - \bar{Y}\right)^2}} \tag{12}$$

where $\bar{X_{:,k}}$ and $\bar{Y}$ denote the average value of $X_{l,k}$ and Y_l, respectively. When ρ_k approaches 1 or -1, it indicates a significant positive or negative correlation between the impedance data and battery SOH. To filter out the insignificant data and improve the training efficiency, ρ_k serves as the secondary metric for impedance feature selection.

The proposed impedance feature selection scheme is summarized in Algorithm 1, where ξ_0 and ρ_0 are predefined thresholds for impedance data dropout. As a first step, the real and imaginary impedance residuals are calculated by fitting the impedance measurement data with the VCM. Further, ξ_k and ρ_k are calculated by applying Eqs. (11) and (12), respectively. The invalid or insignificant data are filtered out of the training dataset. The remaining impedance data are then selected as the final inputs to construct the battery degradation model.

Algorithm 1 (*The proposed impedance feature selection scheme.*).

Input: Training dataset $D = \{(X, Y)\}$

Initialization: ξ_0,ρ_0
for $l \leftarrow 1$ to M **do**
  Compute τ_i using Eqs. (7) and (8)
  Compute $\boldsymbol{\theta}_l$ by solving Eq. (9)
  Compute $\tilde{\boldsymbol{Z}}'(\omega_k)$ and $\tilde{\boldsymbol{Z}}''(\omega_k)$ using Eq. (6)
  Compute $\Delta\mathrm{Re}(l,k)$ and $\Delta\mathrm{Im}(l,k)$ using Eq. (10)
end for
for $k \leftarrow 1$ to K **do**
  Compute ξ_k using Eq. (11)
  Compute ρ_k using Eq. (12)
  if $\xi_k > \xi_0$ **or** $|\rho_k| < \rho_0$ **then**
    $X = Xleft\{X_{:,k}\}$
  end if
end for

3.2. SOH estimation

3.2.1. XGBoost algorithm

Upon obtaining the designated impedance features, a data-driven model can be constructed for estimating the battery degradation trajectories. Decision tree methods are particularly well-suited for this task due to their transparent interpretations and effective handling of non-linear relationships within the data [38]. Among these methods, XGBoost is selected for its superior precision and reduced risk of overfitting. Meanwhile, XGBoost stands out due to its capability to incorporate a second-order Taylor expansion of the target function, resulting in a more accurate loss-function [39]. This capability enhances the modeling performance by capturing the complex patterns of battery degradation. The specifications and configuration of the XGBoost algorithm are described in the following.

Given that X and Y are expressed as the input and target data, respectively, an XGBoost regression model with N trees is expressed as follows:

$$\hat{Y} = \sum_{i=n}^{N} f_n(X), f_n \in \mathscr{F} \tag{13}$$

where $\hat{Y}$ denotes the model output, f_n denotes a scoring function to estimate output from a regression tree, and $\mathscr{F}$ denotes a set of all possible regression trees.

The XGBoost algorithm runs iterative steps, updating the model outputs. The model output at the t-th iteration is expressed as follows:

$$\hat{Y}^t = \hat{Y}^{t-1} + f_t(X) \tag{14}$$

where $\hat{Y}^t$ and $\hat{Y}^{t-1}$ represent the updated model output and the model output from the previous iteration, respectively. f_t is the scoring function from the present regression tree.

The objective function of XGBoost is expressed as:

$$\begin{aligned} J(f_t) &= \sum_{l=1}^{L} \mathscr{L}\left(y_l, \hat{y}_l^{t-1} + f_t(X_l)\right) + \sum_{n=1}^{t} \Omega(f_n) \\ &= \sum_{l=1}^{L} \mathscr{L}\left(y_l, \hat{y}_l^{t-1} + f_t(X_l)\right) + \Omega(f_t) + Const \end{aligned} \tag{15}$$

where $Const$ denotes the constant term; $\mathscr{L}$ denotes the loss-function, which is expressed as mean square error (MSE) in a regression model; Ω denotes the regularization term.

By using the second-order Taylor expansion on the loss-function in XGBoost, Eq. (15) can be rewritten as:

$$J(f_t) \approx \sum_{l=1}^{L} \left[\mathscr{L}\left(y_l, \hat{y}_l^{t-1}\right) + g_l f_t(X_l) + \frac{1}{2} h_l f_t^2(X_l)\right] + \Omega(f_t) + Const \tag{16}$$

where g_l and h_l are the first and second derivatives, which are calculated by:

$$\begin{cases} g_l = \partial \mathscr{L}\left(y_l, \hat{y}_l^{t-1}\right) / \partial \left(\hat{y}_l^{t-1}\right) \\ h_l = \partial^2 \mathscr{L}\left(y_l, \hat{y}_l^{t-1}\right) / \partial \left(\hat{y}_l^{t-1}\right)^2 \end{cases} \tag{17}$$

After removing all the constant terms, the cost-function in XGBoost at the t-th step can be simplified as:

$$J(f_t) \approx \sum_{l=1}^{L} \left[g_l f_t(X_l) + \frac{1}{2} h_l f_t^2(X_l)\right] + \Omega(f_t) \tag{18}$$

3.2.2. Proposed GA-XGBoost method

XGBoost offers a diverse set of hyperparameters that significantly influence modeling performance. For instance, fine-tuning the maximum tree depth and subsample ratio can mitigate overfitting while accelerating training speed. Adjusting the learning rate enhances the model's resilience to noisy data and improves overall robustness, ensuring more accurate predictions in complex scenarios [40]. This study investigates six key hyperparameters, detailed in Table 1, aiming to optimize their settings for improving modeling performance.

Table 1
Specifications of the hyperparameters in XGBoost modeling.

Hyperparameter	Description	Range
p_1	Number of trees	[10, 1000]
p_2	Maximum depth of trees	[1,30]
p_3	Minimum sum of the instance weight in a child	[1,10]
p_4	Subsample ratio of the training instance	[0.8, 1]
p_5	Column sampling ratio by tree	[0.8, 1]
p_6	Learning rate	[0, 0.3]

Optimizing hyperparameters in XGBoost is a complex and computationally intensive process, often requiring navigation through a vast, non-linear search space filled with numerous local minima. To address this, the study employs a genetic algorithm chosen for its fast convergence speed and exceptional capability in dealing with complex and high-dimensional search spaces. Furthermore, the genetic algorithm's excellent ability to balance exploration and exploitation ensures a more comprehensive search across the solution space, making it particularly advantageous for hyperparameter optimization [41,42].

In constructing the GA-XGBoost modeling approach, several key steps are followed:

Step 1: Obtain objective function values. Firstly, the population of the hyperparameters is initialized within the specified ranges listed in Table I. The XGBoost models are then constructed by the population using Eqs. (13)–(18). For each individual, the objective function is determined via a five-fold cross-validation process on the XGBoost model for battery SOH estimation. Specifically, the objective function is as follows:

$$f_{GA} = \mathrm{MSE}_{5-\mathrm{fold}}(\mathrm{XGBoost}) \tag{19}$$

where $\mathrm{MSE}_{5-\mathrm{fold}}$ denotes the mean square error (MSE) of XGBoost under five-fold cross-validation.

Step 2: Hyperparameter optimization. In the genetic algorithm, favorable individuals are selected for crossover and mutation operations. Crossover entails merging genetic information from parents to generate offspring individuals, while mutation introduces random alterations, fostering population diversity. Throughout iterations, the optimized hyperparameters and the objective function values are stored. The optimization concludes when the maximum iteration limit is attained, and the hyperparameters corresponding to the lowest objective function value are utilized to train the XGBoost model for battery SOH estimation.

3.3. Performance metrics

To quantitatively assess the battery SOH estimation performance, several metrics such as the root mean squared error (RMSE), mean absolute percentage error (MAPE), and R-square are introduced as follows:

$$RMSE = \sqrt{\frac{1}{L}\sum_{l=1}^{L}\left(Y_l - \hat{Y}_l\right)^2} RMSE = \sqrt{\frac{1}{M}\sum_{l=1}^{M}\left(Y_l - \hat{Y}_l\right)^2} \tag{20}$$

$$RMSE = \sqrt{\frac{1}{L}\sum_{l=1}^{L}\left(Y_l - \hat{Y}_l\right)^2} MAPE = \frac{1}{M}\sum_{l=1}^{M}\left|\frac{Y_l - \hat{Y}_l}{Y_l}\right| \times 100\% \tag{21}$$

$$RMSE = \sqrt{\frac{1}{L}\sum_{l=1}^{L}\left(Y_l - \hat{Y}_l\right)^2} R^2 = 1 - \frac{\sum_{l=1}^{M}\left(Y_l - \hat{Y}_l\right)^2}{\sum_{l=1}^{M}\left(Y_l - \bar{Y}\right)^2} \tag{22}$$

where Y_l and $\hat{Y}_l$ are the measured and estimated SOH values, respectively; M is the number of samples.

4. Results and discussions

In this section, the proposed impedance feature optimization scheme and GA-XGBoost model for battery SOH estimation are verified. For starters, the impedance validity and correlation analysis are presented. Further, the proposed SOH estimator is validated using different testing datasets.

4.1. Impedance validity and correlation analysis

The impedance data investigated in this work are available from the public dataset as detailed in [9], where 12 LIBs are cycled at 25 °C (25C01-25C08), 35 °C (35C01 and 35C02), and 45 °C (45C01 and 45C02), respectively. Each cycle comprises a 1C-rate constant current–constant voltage (CC-CV) charge followed by a 2C-rate discharge process. The impedance measurements were taken at nine different stages during the battery cycling process, spanning a frequency range of 20 mHz to 20 kHz. These measurements were performed using an electrochemical workstation that generated an excitation current of 5 mA to perturb the system and record the corresponding impedance responses.

In this study, the training dataset comprises four batteries: 25C01, 25C02, 35C01, and 45C01. By adapting the model to data from multiple temperature environments during the training phase, it can learn to identify the specific impedance characteristics associated with different temperature conditions. The features extracted from the EIS data are the real and imaginary parts of the impedance collected at 60 different frequencies, logarithmically ranging from 20 mHz to 20 kHz. The input dataset has a dimensionality of 1,027 rows (corresponding to the cycle numbers) and 120 columns (corresponding to the impedance features). The output dataset, which contains the SOH values, has a dimensionality of 1,027 rows and 1 column. The testing dataset includes two batteries: 35C02 and 45C02. The impedance data collected at Stage III (charging after 20 min) and Stage IV (end of charging) are specifically examined in this study due to the practical convenience of obtaining impedance data during the battery charging process [43]. During Stage III, impedance measurements are conducted while the batteries are actively charging, which introduces significant polarization effects and changes in electrochemical behaviors, contributing to measurement uncertainties. Similarly, in Stage IV, the batteries have not yet fully relaxed after charging, leading to residual polarization effects that can influence the accuracy of the impedance measurements. The evolution of battery impedance spectra during the degradation process is illustrated in Fig. 3, showing that the impedance spectrum expands and shifts to the right as the battery SOH decreases.

The validity of impedance measurements is analyzed using Eq. (6)–Eq. (12). As shown in Fig. 4, the presence of impedance measurement uncertainties is the most pronounced in the millihertz region, where the fitted impedance significantly deviates from the actual impedance measurements. The gradual increase in impedance residuals with decreasing frequency is attributed to the low-frequency polarization voltage components during battery charging, according to Eq. (4). It should be noted that the unreliable impedance measurements caused by polarization effects are less pronounced under higher SOC conditions, as shown in Fig. 4(b). As battery charging progresses, fluctuations in ionic concentration diminish, resulting in reduced interference with impedance measurements [44]. Additionally, the impedance uncertainties persist regardless of the testing temperature and battery aging status. These factors can exacerbate the challenges in achieving accurate assessments of battery degradation, highlighting the necessity of performing the impedance validity analysis before constructing the SOH estimation model. To prove the effectiveness of the proposed method in addressing and mitigating impedance uncertainties, the impedance data measured at Stage III are utilized to train the SOH estimation model.

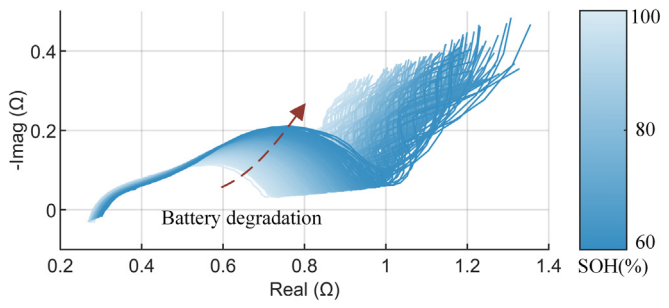

Fig. 3. Nyquist plot of impedance spectra for cell 35C01 at Stage IV during the degradation process.

The RMS residuals between the measured and fitted impedance across different frequency ranges, denoted as ξ_k, are depicted in Fig. 5(a). The analysis indicates that the maximum ξ_k is limited to 1% in the kilohertz region. The high accuracy in this frequency range can be attributed to the high-frequency measurements primarily reflecting the intrinsic properties of the electrolyte and the electrode surfaces, which are less variable and more stable over time. In the hertz range, impedance residuals can still peak at 2.33%, as the medium-frequency impedance measurements are particularly sensitive to the charge-transfer processes and the double-layer capacitance effects. Despite higher impedance residuals, the information gathered from this range is crucial for understanding electrochemical dynamics and aging processes. These measurements often reflect changes in the electrochemical interfaces, which are directly related to the battery's aging and degradation mechanisms [34]. The low-frequency impedance measurements are dominated by diffusion processes and are more susceptible to noise and external disturbances. These factors contribute to higher variability and less reliable data, which is evident from the higher residuals observed. According to [33,45], impedance data is deemed valid and reliable when the deviation between measured and fitted impedance is well below 1%. To enhance the SOH estimation performance, invalid impedance data are systematically filtered out from the training dataset when $\xi_k > \xi_0$, where ξ_0 is set to 0.5%, ensuring that only the most reliable impedance data are retained.

The correlation analysis of impedance data and battery SOH for different batteries is illustrated in Fig. 5(b). The analysis reveals that the relationship between battery degradation and impedance data varies significantly across different frequency ranges. Notably, the medium-frequency impedance measurements exhibit a strong correlation with battery SOH, suggesting that these frequencies are particularly sensitive to the electrochemical processes associated with battery aging. This heightened sensitivity can be attributed to the charge-transfer resistance and double-layer capacitance evolving with battery degradation, emphasizing the importance of medium-frequency impedance measurements for accurate SOH estimation. By contrast, the imaginary impedance measured at low frequencies shows a weak correlation with battery health, indicating that these measurements are less reliable for evaluating battery degradation. To reduce computational complexity and optimize the training process, the insignificant impedance data are removed from the training dataset once $|\rho_k|$ is lower than ρ_0, which is set to 0.6 in this study. The final input variables are selected from the remaining impedance data, distinguished by their high validity and importance, as listed in Table 2.

Before implementing the XGBoost method for battery SOH estimation, the genetic algorithm is employed to optimize the XGBoost hyperparameters, aiming to enhance the modeling accuracy and

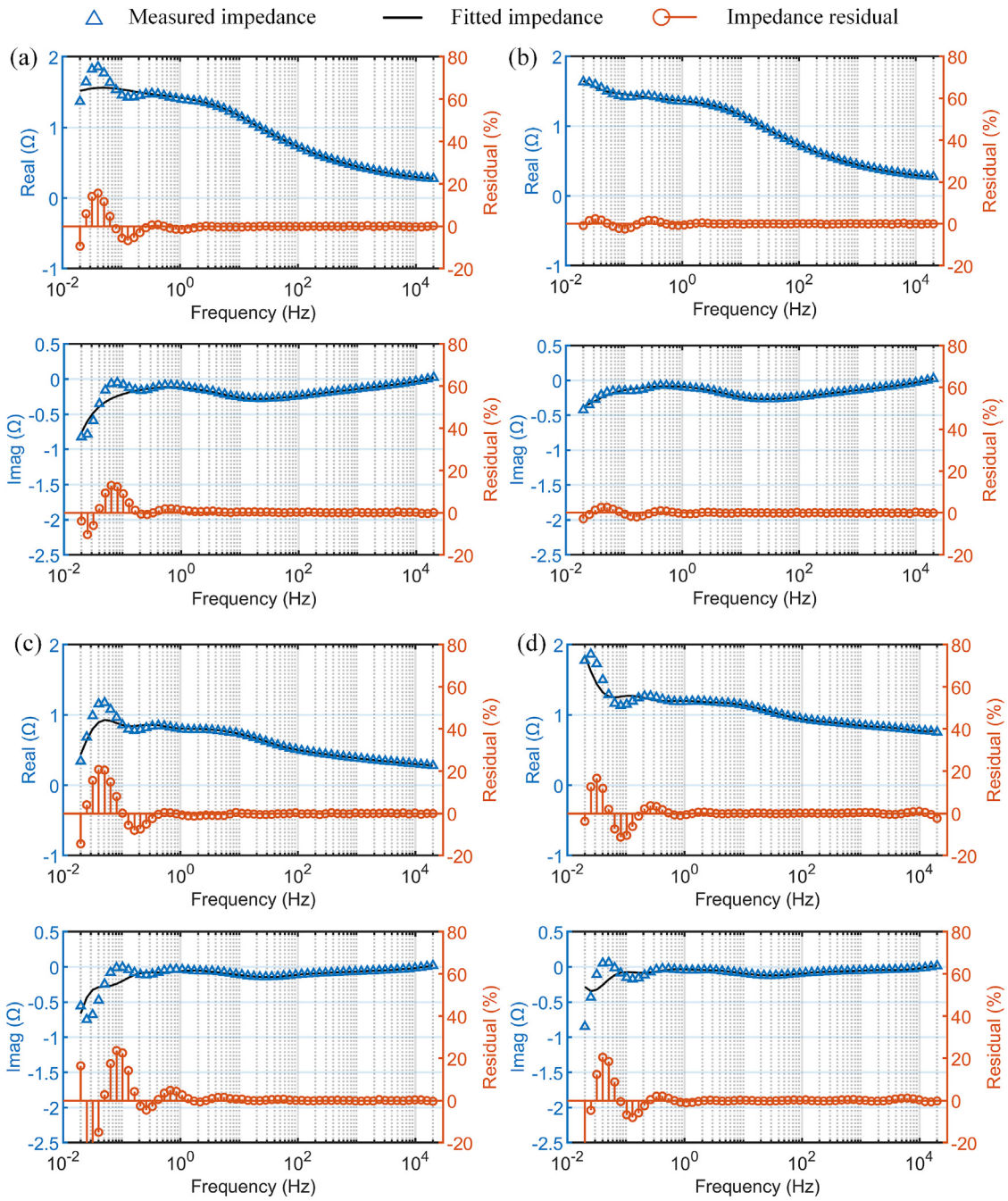

Fig. 4. Impedance validity analysis results: (a) 25C01-III, 100% SOH; (b) 25C01-IV, 100% SOH; (c) 35C01-III, 80% SOH; (d) 45C01-III, 80% SOH.

improve generalization ability. The maximum iteration for the genetic algorithm is set to 50, with a population size of 10. Using five-fold cross-validation on the training set, the optimization process yields the fitness values depicted in Fig. 6. The results show that the MSE of the cross-validation results decreases from 1.38×10^{-4} and converges to 1.24×10^{-4} from the 33-rd iteration. Finally, the values of the optimized hyperparameters are 338 for the number of trees, 30 for the maximum depth, 7 for the minimum child weight, 0.86 for the column sample ratio by tree, 0.99 for the subsample ratio, and 0.05 for the learning rate.

4.2. Estimation model verification

To evaluate the performance of the proposed method, four datasets from two individual cells (35C02 and 45C02), with impedance data collected at different SOCs (Stages III and IV), are utilized to test the SOH estimation model. The estimation performance of the proposed method is compared using raw EIS data as inputs to the GA-XGBoost and CNN models, with the CNN's hyperparameters obtained from [31,46]. The SOH estimation results, depicted in Fig. 7, demonstrate that the model incorporating optimized features delivers significantly more reliable performance and higher accuracy. The estimation results are closely aligned with the baseline across various aging conditions and testing scenarios, achieving an average MAPE of merely 1.29%. In contrast, the use of raw impedance data, which includes numerous invalid and irrelevant features, proves unsuitable for direct model input and often results in unreliable estimation performance. Specifically, the GA-XGBoost and CNN models, which utilize raw EIS data as model inputs, yield considerably higher average MAPEs of 2.80% and 3.95%, respec-

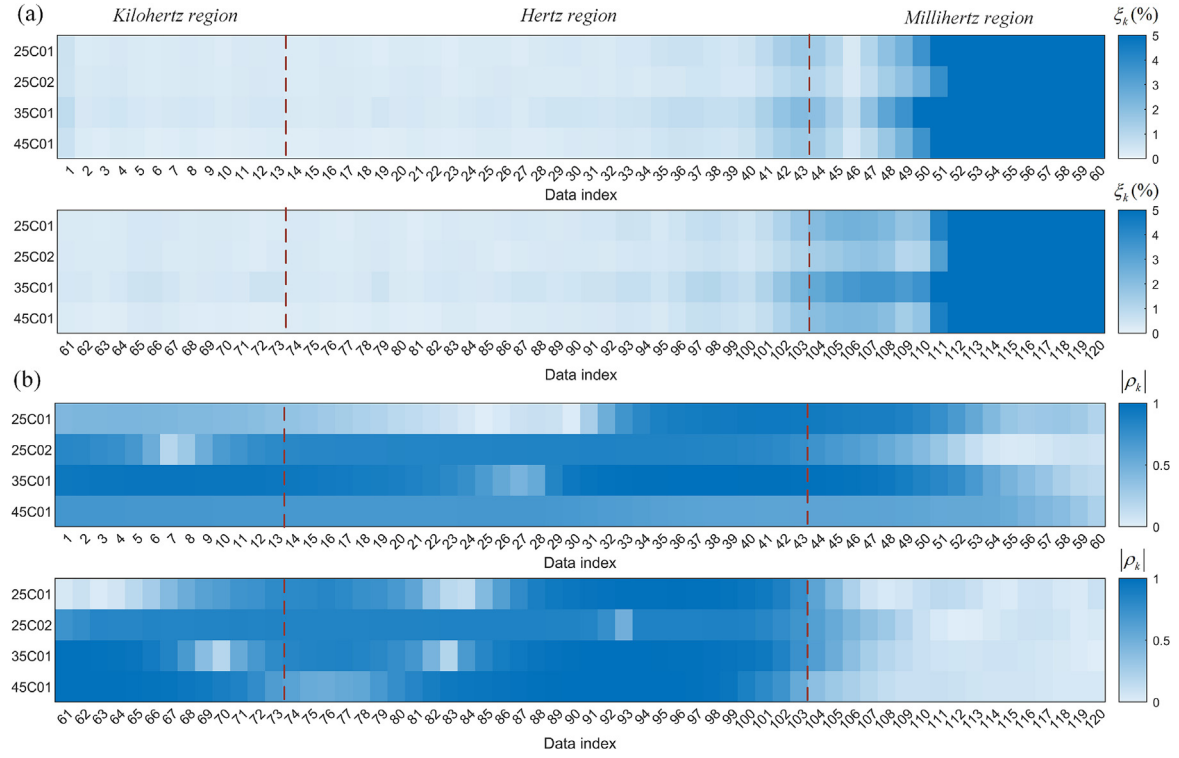

Fig. 5. (a, b) RMS residuals and correlation coefficients of impedance measurement data collected at Stage III.

Table 2
RMS residuals and correlation coefficients of the selected impedance data.

		$X_{:,80}$	$X_{:,86}$	$X_{:,88}$	$X_{:,89}$	$X_{:,90}$	$X_{:,95}$
ξ_k	25C01	0.24%	0.28%	0.34%	0.30%	0.33%	0.34%
	25C02	0.26%	0.21%	0.28%	0.33%	0.33%	0.37%
	35C01	0.29%	0.41%	0.47%	0.41%	0.46%	0.49%
	45C01	0.19%	0.20%	0.27%	0.28%	0.31%	0.31%
$\|\rho_k\|$	25C01	0.67	0.63	0.88	0.92	0.95	0.99
	25C02	0.85	0.86	0.86	0.85	0.85	0.86
	35C01	0.76	0.90	0.96	0.97	0.99	0.99
	45C01	0.76	0.96	0.98	0.99	0.99	0.99

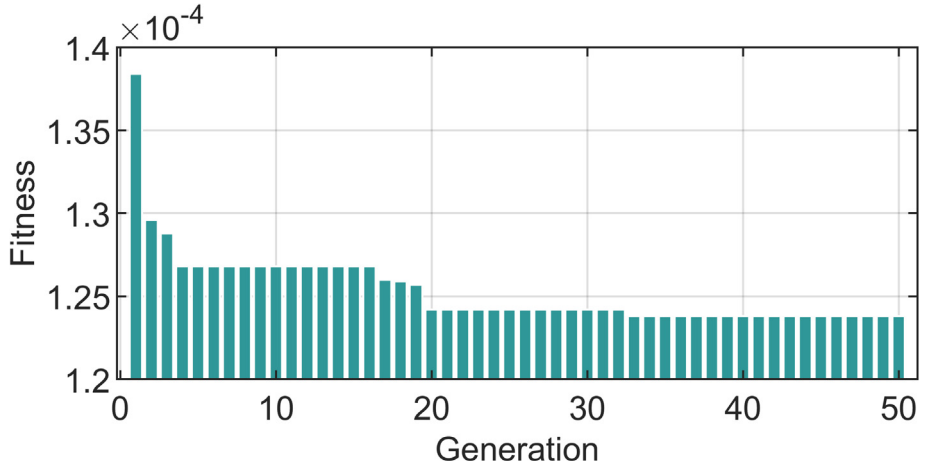

Fig. 6. Fitness values during the optimization process.

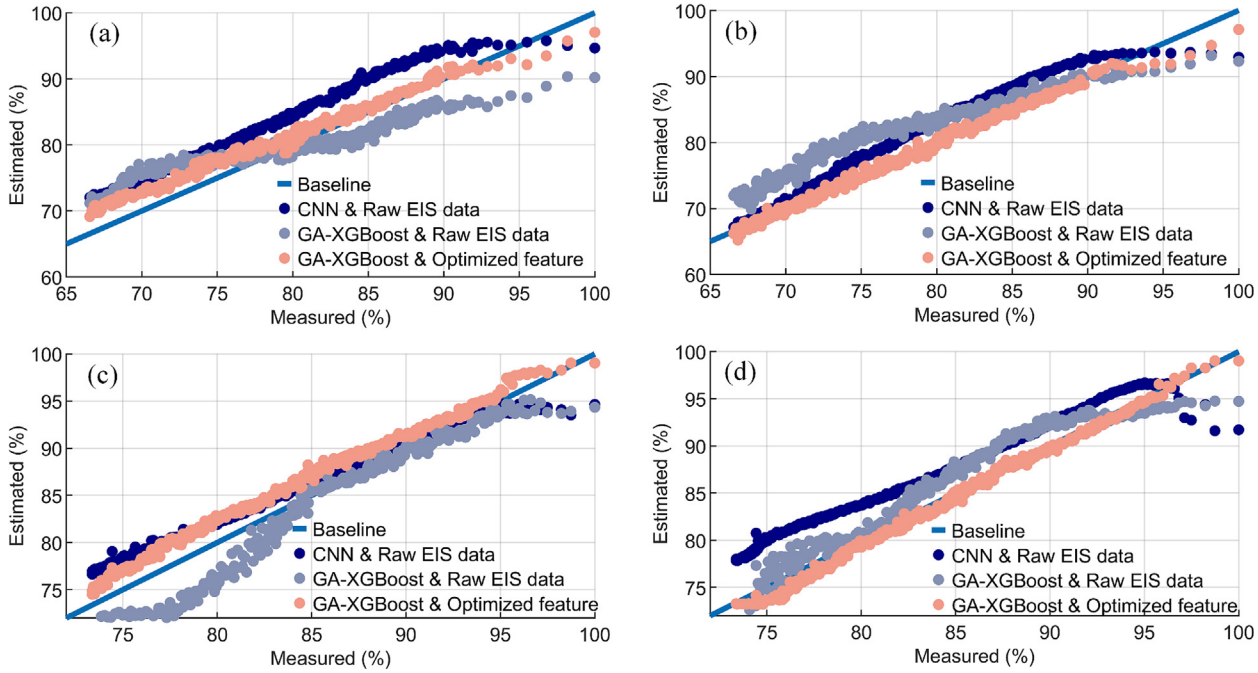

Fig. 7. Scatter diagram of SOH estimation results: (a) 35C02-III; (b) 35C02-IV; (c) 45C02-III; (d) 45C02-IV.

tively. These results highlight the critical importance of feature optimization in enhancing the accuracy, robustness, and overall effectiveness of the SOH estimation model, confirming that properly selected and processed features are essential for reliable battery health assessment.

Fig. 8 illustrates the GPR, DNN, SVR, and regular XGBoost models as benchmarks for comparative analysis. These models utilize optimized impedance data as inputs, known for their superior modeling performance. The hyperparameters of the GPR model are obtained from [9], while the DNN architecture comprises four hidden layers and four dropout layers, as detailed in [31]. The SVR uses a Gaussian kernel while the normal XGBoost model is built using default hyperparameters. The estimation results demonstrate the good generalization of all estimation models for accurately tracking reference SOH values, validating the reliability of optimized impedance features. Notably, the proposed method not only shows consistent SOH estimation performance but also yields a smooth and precise estimation curve. Conversely, the regular XGBoost model, with default hyperparameters, exhibits pronounced fluctuations in estimation results. It is noticed that while the proposed GA-XGBoost model performs well overall, it does not outperform the GPR model in certain instances as in Fig. 8(c). This can be attributed to the strengths of the GPR model in identifying underlying trends and managing uncertainties in datasets characterized by gradual changes and less complexity [39]. On the other hand, the GA-XGBoost model is a robust and highly adaptable approach specifically designed to capture complex, non-linear relationships. Consequently, in cases where the

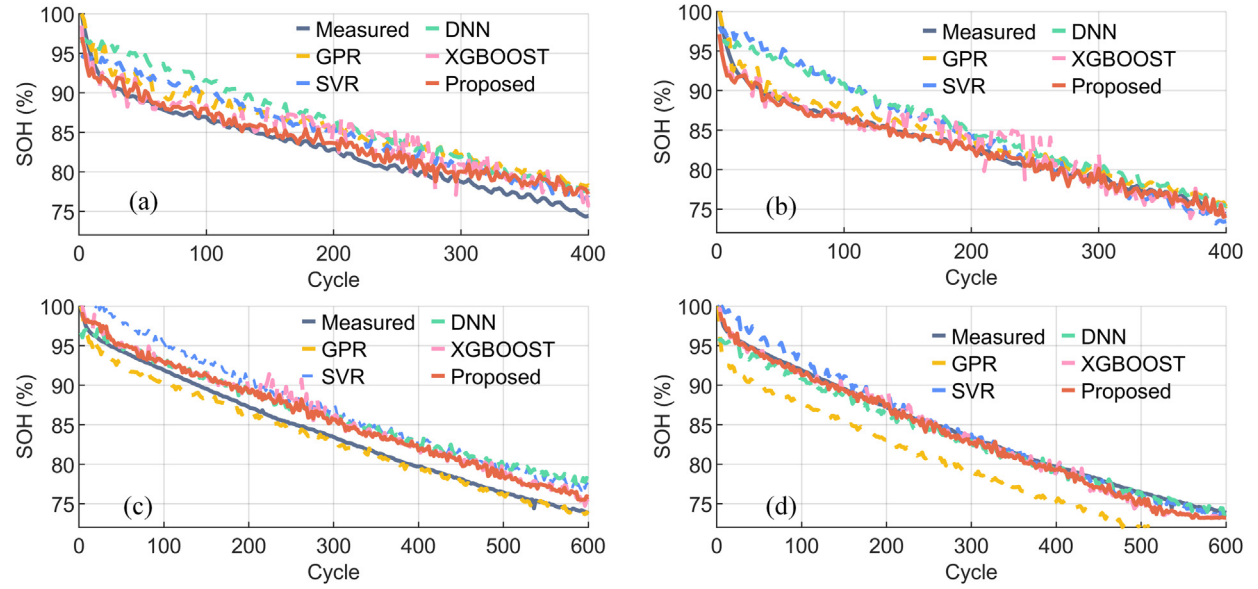

Fig. 8. SOH estimation results with different data-driven models: (a) 35C02-III; (b) 35C02-IV; (c) 45C02-III; (d) 45C02-IV.

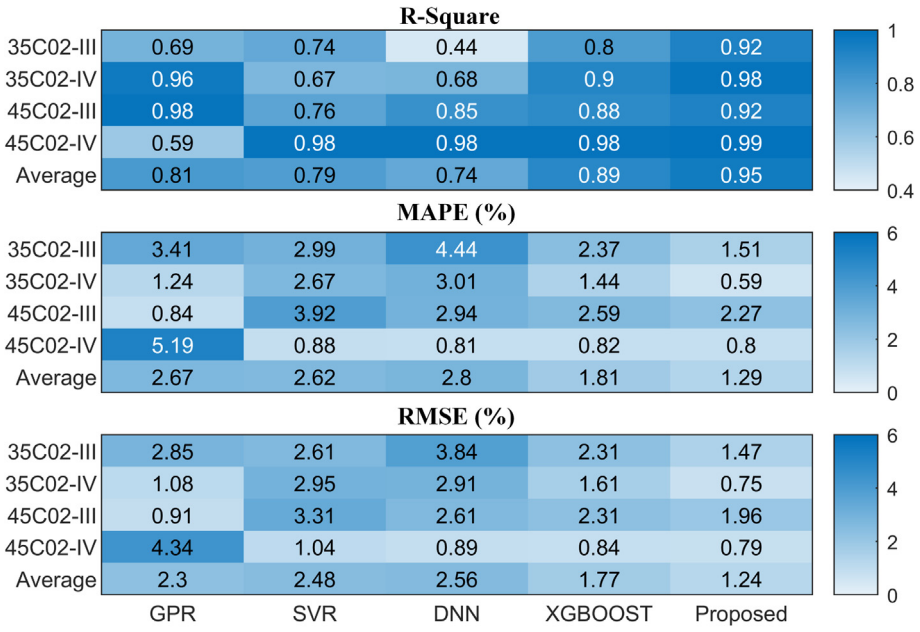

Fig. 9. Numerical evaluation of SOH estimation results.

data does not exhibit strong non-linearities, GPR's effectiveness in minimizing errors can lead to superior performance compared to GA-XGBoost.

The numerical assessment of the estimation outcomes is graphically outlined in Fig. 9. It is shown that the proposed method demonstrates good generalization ability and robustness, where R^2 is higher than 0.92 in all testing datasets. In contrast, the baseline methods exhibit inconsistent performance, with the GPR and DNN methods exhibiting the lowest R^2 values at 0.59 and 0.44, respectively. To provide a comprehensive perspective on estimation performance, the average metrics of each evaluation index across diverse testing datasets are presented. The average MAPE and RMSE of the proposed method are only 1.29% and 1.24%, respectively, while those of the baseline methods are larger than 1.81% and 1.77%, respectively. The superior performance of the proposed method highlights the effectiveness of the GA-XGBoost modeling framework for battery SOH estimation.

5. Conclusion

Evaluating battery health is crucial for ensuring the reliable operation, effective management, and timely maintenance of LIBs. This paper introduces an advanced feature optimization scheme in combination with a GA-XGBoost modeling approach for accurate and robust SOH estimation. Given the uncertainties in impedance measurements during battery operation, the study first performs an impedance validity and correlation analysis to filter out invalid and insignificant impedance data, ensuring only high-quality data is used for constructing the estimation model. Following this, an XGBoost modeling framework is developed to identify battery degradation patterns and estimate battery SOH values. The genetic algorithm is employed to optimize the hyperparameters within the XGBoost model, thereby significantly enhancing its adaptability and generalization performance across various operating conditions.

The experimental results demonstrate that the proposed method can accurately estimate battery degradation trajectories across various testing datasets, achieving an average MAPE of 1.29% and RMSE of 1.24%, respectively. The GA-XGBoost and CNN models utilizing raw EIS data as model inputs yield considerably higher MAPEs of 2.80% and 3.95%, respectively. Furthermore, all baseline methods, even when using optimized features, resulted in an average MAPE and RMSE exceeding 1.81% and 1.77%, respectively. These results highlight the advantages of the proposed feature optimization scheme and the GA-XGBoost modeling framework.

Future research will explore the feasibility and effectiveness of the proposed method under more severe noise conditions. Additionally, the proposed method can be further refined by integrating the feature engineering framework with the underlying battery degradation mechanisms, thereby enhancing the predictive accuracy and reliability of SOH estimation.

CRediT authorship contribution statement

Xinghao Du: Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Conceptualization. **Jinhao Meng:** Writing – review & editing, Visualization, Validation, Formal analysis. **Yassine Amirat:** Writing – review & editing, Visualization, Validation, Supervision, Formal analysis. **Fei Gao:** Writing – review & editing, Visualization, Validation, Supervision, Formal analysis. **Mohamed Benbouzid:** Writing – review & editing, Visualization, Validation, Supervision, Methodology, Formal analysis.

Declaration of competing interest

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

References

[1] X. Huang, J. Meng, W. Jiang, W. Liu, K. Liu, Y. Zhang, D.I. Stroe, R. Teodorescu, J. Energy Chem. 96 (2024) 679–697.
[2] J. Zhao, X. Han, M. Ouyang, A.F. Burke, J. Energy Chem. 87 (2023) 416–438.
[3] D. Chen, J. Meng, H. Huang, J. Wu, P. Liu, J. Lu, T. Liu, Energy 245 (2022) 123222.
[4] K. Khodadadi Sadabadi, X. Jin, G. Rizzoni, J. Power Sources 481 (2021) 228861.
[5] Y. Cui, P. Zuo, C. Du, Y. Gao, J. Yang, X. Cheng, Y. Ma, G. Yin, Energy 144 (2018) 647–656.

[6] J. Meng, L. Cai, D.I. Stroe, X. Huang, J. Peng, T. Liu, R. Teodorescu, IEEE Trans. Ind. Electron. 69 (2022) 2659–2668.
[7] E. Miguel, G.L. Plett, M.S. Trimboli, L. Oca, U. Iraola, E. Bekaert, J. Energy Storage 44 (2021) 103388.
[8] X. Du, J. Meng, Y. Zhang, X. Huang, S. Wang, P. Liu, T. Liu, IEEE Trans. Ind. Electron. 69 (2022) 5889–5899.
[9] Y. Zhang, Q. Tang, Y. Zhang, J. Wang, U. Stimming, A.A. Lee, Nat. Commun. 11 (2020) 6–11.
[10] Q. Li, D. Li, K. Zhao, L. Wang, K. Wang, J. Energy Storage 50 (2022) 104215.
[11] M. Lin, D. Wu, J. Meng, J. Wu, H. Wu, J. Power Sources 518 (2022) 230774.
[12] Y. Ma, L. Wu, Y. Guan, Z. Peng, J. Power Sources 476 (2020) 228581.
[13] M. Lin, J. Wu, J. Meng, W. Wang, J. Wu, Energy 268 (2023) 126706.
[14] S. Wang, P. Takyi-Aninakwa, S. Jin, C. Yu, C. Fernandez, D.I. Stroe, Energy 254 (2022) 124224.
[15] Z. Deng, L. Xu, H. Liu, X. Hu, B. Wang, J. Zhou, J. Energy Chem. 89 (2024) 345–354.
[16] Y. Li, D.I. Stroe, Y. Cheng, H. Sheng, X. Sui, R. Teodorescu, J. Energy Storage 33 (2021) 102122.
[17] S. Wang, Y. Fan, S. Jin, P. Takyi-Aninakwa, C. Fernandez, Reliab. Eng. Syst. Saf. 230 (2023) 108920.
[18] K. Li, N. Xie, Energy 303 (2024) 131888.
[19] F. Xia, K. Wang, J. Chen, J. Energy Storage 64 (2023) 107161.
[20] Z. Deng, X. Hu, P. Li, X. Lin, X. Bian, IEEE Trans. Power Electron. 37 (2022) 5021–5031.
[21] X. Hu, J. Jiang, D. Cao, B. Egardt, IEEE Trans. Ind. Electron. 63 (2016) 2645–2656.
[22] X. Du, J. Meng, Y. Amirat, F. Gao, M. Benbouzid, J. Energy Chem. 95 (2024) 464–483.
[23] K. Ando, T. Matsuda, D. Imamura, J. Energy Chem. 53 (2020) 285–289.
[24] Y. Zhou, G. Dong, Q. Tan, X. Han, C. Chen, J. Wei, Energy 262 (2023) 125514.
[25] X. Wang, X. Wei, H. Dai, J. Energy Storage 21 (2019) 618–631.
[26] R. Koch, R. Kuhn, I. Zilberman, A. Jossen, Electrochemical impedance spectroscopy for online battery monitoring-power electronics control, 2014 16th European Conference on Power Electronics and Applications, IEEE, 2014.
[27] J.A.A. Qahouq, Z. Xia, IEEE Trans. Ind. Electron. 64 (2017) 7019–7029.
[28] J. Guo, Y. Che, K. Pedersen, D.I. Stroe, J. Energy Chem. 79 (2023) 211–221.
[29] Q. Zhang, C.G. Huang, H. Li, G. Feng, W. Peng, IEEE Trans. Transp. Electrification 8 (2022) 4633–4645.
[30] M. Messing, T. Shoa, S. Habibi, J. Energy Storage 43 (2021) 103210.
[31] J. Obregon, Y.R. Han, C.W. Ho, D. Mouraliraman, C.W. Lee, J.Y. Jung, J. Energy Storage 60 (2023) 106680.
[32] R. Xiong, J. Tian, W. Shen, J. Lu, F. Sun, J. Energy Chem. 76 (2023) 404–413.
[33] N. Meddings, M. Heinrich, F. Overney, J.S. Lee, V. Ruiz, E. Napolitano, S. Seitz, G. Hinds, R. Raccichini, M. Gaberšček, J. Park, J. Power Sources 480 (2020) 228742.
[34] P. Iurilli, C. Brivio, V. Wood, J. Power Sources 505 (2021) 229860.
[35] X. Du, J. Meng, Y. Amirat, F. Gao, M. Benbouzid, IEEE Trans. Transp. Electrification (2024). https://doi.org/10.1109/TTE.2024.3399051.
[36] J.C. Lagarias, J.A. Reeds, M.H. Wright, P.E. Wright, SIAM J. Optim. 9 (1998) 112–147.
[37] Y. Zhou, M. Huang, Y. Chen, Y. Tao, J. Power Sources 321 (2016) 1–10.
[38] X. Sui, S. He, S.B. Vilsen, J. Meng, R. Teodorescu, D.I. Stroe, Appl. Energy 300 (2021) 117346.
[39] T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016.
[40] R. Shi, X. Xu, J. Li, Y. Li, Appl. Soft Comput. 109 (2021) 107538.
[41] C. Chang, Q. Wang, J. Jiang, T. Wu, J. Energy Storage 38 (2021) 102570.
[42] H. Alibrahim, S.A. Ludwig, Hyperparameter optimization: Comparing genetic algorithm against grid search and bayesian optimization, 2021 IEEE Congress on Evolutionary Computation (CEC), 2021.
[43] X. Huang, J. Meng, W. Liu, F. Ru, C. Duan, X. Xu, D.I. Stroe, R. Teodorescu, IEEE Trans. Ind. Electron. 71 (2024) 484–492.
[44] Q. Li, D. Yi, G. Dang, H. Zhao, T. Lu, Q. Wang, C. Lai, J. Xie, World Electr. Veh. J. 14 (2023) 321.
[45] M. Hahn, S. Schindler, L.C. Triebs, M.A. Danzer, Batteries 5 (2019) 43.
[46] T.K. Pradyumna, K. Cho, M. Kim, W. Choi, J. Power Electron. 22 (2022) 850–858.