Energy 304 (2024) 132077

Contents lists available at ScienceDirect

Energy

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

Comparison of techniques based on frequency response analysis for state of health estimation in lithium-ion batteries

Shaojin Wang [a], Jinrui Tang [a,*], Binyu Xiong [a,**], Junqiu Fan [b,c,***], Yang Li [d], Qihong Chen [a], Changjun Xie [a], Zhongbao Wei [e]

[a] School of Automation, Wuhan University of Technology, Wuhan, China
[b] The Electrical Engineering College, Guizhou University, Guiyang, Guizhou, China
[c] Guizhou Power Grid Co., Ltd.,Guiyang, 550002, Guizhou, China
[d] Department of Electrical Engineering, Chalmers University of Technology, Sweden
[e] National Engineering Laboratory for Electric Vehicles, School of Mechanical Engineering, Beijing Institute of Technology, Beijing, China

ARTICLE INFO

Handling editor: X Ou

Keywords:
State of health estimation
Electrochemical impedance spectroscopy
Equivalent circuit model
Distribution of relaxation times
Frequency response analysis

ABSTRACT

Frequency response analysis (FRA) methods are commonly used in the field of State of Health (SOH) estimation for Lithium-ion batteries (Libs). However, identifying their appropriate application scenarios can be challenging. This paper presents four FRA techniques, including electrochemical impedance spectra (EIS), mid-frequency and low-frequency domain equivalent circuit model (MLECM), distribution of relaxation time (DRT) and non-linear FRA (NFRA) technique. This paper proposes two estimation frameworks, machine learning and curve fitting, to be applied to each of the four techniques. Eight SOH estimation models are developed by linking the extracted feature parameters to the battery capacity variations. The paper compares the accuracy of estimation, estimation range, and other properties of the eight models. Application scenarios are identified for the techniques by using three classification methods: different estimation frameworks, frequency response linearity, and impedance technique. The results demonstrate that MLF is recommended for scenarios with a large amount of battery data, while CFF is recommended for scenarios with a small amount of data. NFRA could be applied to electric vehicle power batteries, while LFRA is recommended to be used for retired batteries. EIS method is recommended for complex and dynamic scenarios, while non-EIS method is recommended for scenarios that require high accuracy.

Nomenclature

		N	number of samples
		K	number of model parameters
C	charge/discharge ratio	n	the sample length
I_{AC}	applied current	S	RMSE
f	sampling frequency	**I**	variable exponential model
Hz	unit of frequency	**II**	exponential model
Z	impedance	**III**	power model
ω	angular velocity	**IV**	variable power model
ϕ	phase	**V**	rational model
j	imaginary unit		
$V(t)$	voltage signal	**Abbreviation**	
$I(t)$	current signal		
V_A	voltage amplitude	Libs	lithium-ion batteries

(*continued on next column*)

(*continued*)

I_A	current amplitude	FRA	frequency response analysis
Z_A	impedance amplitude	NFRA	nonlinear FRA
Z_{Re}	real part of impedance	LFRA	linear FRA
Z_{Im}	imagine part of impedance	SOH	state of health
H_i	higher harmonic response	SOC	state of charge
L	inductor	P2D	pseudo-two-dimensional
R_0	ohmic resistor	ECM	equivalent circuit model
R_{CT}	charge transfer resistor	TECM	traditional ECM
R_{SEI}	SEI film resistor	DRT	distribution of relaxation time
Z_W	warburg impedance	MLF	machine learning framework
R_{SC}	incorporating R_{SEI} and R_{CT}	CFF	curve fitting framework
Z_{MLECM}	MLECM impedance	RMSE	root mean square error

(*continued on next page*)

* Corresponding author.
** Corresponding author.
*** Corresponding author. Guizhou Power Grid Co., Ltd., Guiyang, China.
E-mail addresses: tangjinrui@whut.edu.cn (J. Tang), bxiong2@whut.edu.cn (B. Xiong), fans_edu_gzdx@yeah.net (J. Fan).

https://doi.org/10.1016/j.energy.2024.132077
Received 22 January 2024; Received in revised form 15 April 2024; Accepted 14 June 2024
Available online 19 June 2024

(*continued*)

Z_{CPE}	CPE impedance	BMS	battery management system
T_{CPE}	Parameters of the CPE	KK	Kramers-Kronig
P_{CPE}	Parameters of the CPE	HF	health factor
τ	time constant	GPR	gaussian process regression
Z_{pol}	polarization impedance	RLS	recursive least squares
R^2	goodness of fit value	PF	particle filter
y_{mea}	the measured value of SOH	SEI	solid electrolyte interface
y_{pre}	the estimation of SOH	CPE	constant phase element
$\overline{y}_{pre}$	the mean of the y_{pre}	ZARC	CPE and resistance in parallel
n_p	number of samples between two adjacent orders of magnitude	HQ	Hannan Quinn
$\overline{H}$	root mean square of 2nd and 3rd harmonics	EIS	electrochemical impedance spectroscopy
Z_{ZARC}	impedance of CPE and resistor in parallel	MLECM	mid-frequency and low-frequency domain ECM

1. Introduction

Lithium-ion batteries show promise for energy-efficient renewable and electric vehicle applications. The battery management system (BMS) is responsible for detecting important parameters such as battery current, voltage, and temperatures [1,2]. It is crucial for the proper functioning of the battery. The detection of state of health (SOH) of the battery is a crucial aspect of the BMS. It is calculated to enhance energy management and prevent potential hazards [3–5].

The estimation of SOH is commonly achieved through direct measurement, model-based, or data-driven based methods [6]. The direct measurement methods are rely heavily on sensor accuracy [7]. Model-based approaches mainly include those based on equivalent circuit models (ECM) [8,9], pseudo-two-dimensional (P2D) models [10,11], single-particle (SP) models [12], Multiphysics field model [13] and so on. The methods generally result in higher estimation accuracy, but constructing the models is more challenging and requires more effort. Data-driven approaches are easier to implement than the former and show great potential [3]. Examples include feedforward neural network algorithms [14–19], classification and regression algorithms [20,21], probabilistic algorithms [22,23], recurrent neural networks [24–26], rule-based algorithms [27,28], and hybrid algorithms [29,30].

Electrochemical impedance spectroscopy (EIS) is a crucial method for battery diagnosis due to its high accuracy, non-invasive nature, and lack of requirement for a complete charge/discharge curve [22,31]. Meddings et al. [32] discuss the practical process of applying EIS to commercial Libs, focusing on advanced interpretation of EIS and validation methods. Machine learning prioritises workload reduction by automatically selecting the most relevant features. Zhang et al. [22] find a strong relation between the impedance at 17.8 Hz and 2.16 Hz. Jones et al. [33] use a probabilistic machine learning (PML) approach in combination with EIS measurement to predict battery SOH while accounting for calibration uncertainty. However, if the amount of data is too large, it can result in excessively long training, times and increased dependence on the accuracy of the algorithm.

Current approaches describe each electrochemical reaction process in terms of a specific circuit element or combination of elements. ECM is a commonly used model in the electrical field for explaining and simulating EIS. Pietro et al. [31] provide a systematic review of the current research on characterising Lib degradation or modelling circuits with EIS. They also discuss the existing links between degradation mechanisms and the most reliable modelling approaches. Most of the current approaches [34–37] are based on the EIS to construct a traditional ECM (TECM). Some scholars have developed alternative ECMs, such as second-order Davening ECMs [38], simplified ECMs [39], low-frequency domain ECMs [40], and temperature-compensated fractional order models [41]. To ensure accurate analysis of the battery and avoid misinterpretation due to over-reliance on the ECM, it is proposed to use distribution of relaxation time (DRT) to verify all constants for the entire electrochemical system. This technique directly distinguishes the time constants of the major electrochemical processes directly. It simplifies impedance analysis, and significantly improves the accuracy of kinetic interpretation on time scales [42,43]. He et al. [44] conducted a DRT analysis of impedance spectra to determine the impedance distribution characteristics of batteries with different materials during cycling. This analysis helped to quickly identify the aging mechanism and predict the battery's remaining life. Zhu et al. [45] proposed a method for interpreting and separating battery interfacial processes. The method is based on the temperature dependence of battery impedance, which is found by the DRT technique. The technology is capable of meeting the challenge with ease not only in conventional Libs but also in systems with complex battery chemistries [46].

Nonlinear frequency response analysis (NFRA) is found to be an effective method for the dynamic analysis of Libs. In contrast to EIS, NFRA is not restricted to the linear response of the system. Instead, it employs high current excitation signals greater than 1.5C to obtain and analyze the higher harmonic response signals represented by $Y_n(n>2)$ [47,48]. The NFRA can be seen as an extension of the EIS. It provides a comprehensive and detailed reflection of the system response and establishes a link between the model parameters and the experimentally observed phenomena [47]. Impedance is only meaningful for fundamental frequencies. Higher-order response functions are not defined in the same way. Therefore, proposing a highly feasible method for the contributing to higher-order responses is a problem to be solved in this field. Rafael et al. [49] proposed a method for frequency domain analysis of nonlinear systems using NFRA functions. This method can be applied to nonlinear modal analysis.

Previous studies have typically validated the feasibility of their methods using separate datasets, resulting in a lack of comparability. Two frameworks for estimating SOH based on machine learning (ML) and curve fitting (CF) are developed. Within these frameworks, four technologies for processing fresh battery data and performing estimation are compared: Impedance-based, mid-frequency and low-frequency domain ECM (MLECM)-based, DRT-based, and NFRA-based.

2. Dataset and test conditions

The paper presents the aging and testing results of commercial ternary 18,650 Libs with a nominal capacity of 3Ah from four different testing institutions (a total of 12 batteries) [5]. Table 1 shows the battery parameters, and Fig. 1 shows the battery capacity degradation curve. All batteries are subjected to an aging cycle between 3.0 V and 4.2 V at 45 °C. The battery charging protocol follow a CCCV pattern: 1.33C (4 A) CC charging until the cutoff voltage of 4.2 V is reached, followed by 4.2 V CV charging until 0.1C (0.3 A) is reached. The discharge protocol involves 1.33C (4 A) CC discharging until the cutoff voltage of 3.0 V is reached. Perform a capacity calibration test at regular cycles intervals, such as every 50 or 100 cycles at room temperature (23 °C). Use CCCV charge and 0.4C (1.25 A) discharge to cutoff voltage. Record the discharged capacity as the current capacity of the battery. Electrochemical workstations from four different institutions are used to carry out the EIS or NFR tests. The battery is fully charged after the capacity calibration test and the state of charge (SOC) is adjusted using a discharge current of

Table 1
Battery parameters.

Category	Parameter
Type	18,650 (18 mm*65 mm)
Anode	$LiNi_{0.8}Co_{0.1}Mn_{0.1}O_2$
Cathode	Silicon-Graphite
Nominal discharge capacity (m·Ah)	3000
Charge cut-off voltage (V)	3.0
Discharge cut-off voltage (V)	4.2

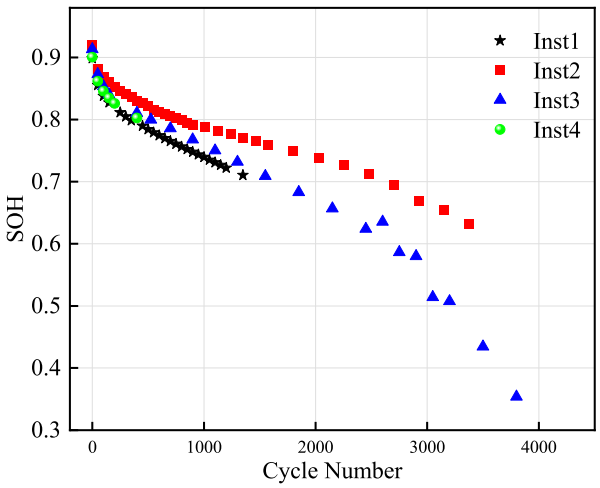

Fig. 1. Battery capacity decline in four institutions (for the sake of simplicity, the average decline curves of the four institutions are used as the unit of calculation, rather than individual batteries).

1.25 A to obtain data at different SOCs. EIS measurements should only be conducted when the battery's polarization reaction of the battery is relatively weakened. Therefore, the battery should rest for at least 30 mins after each SOC adjustment before testing.

Table 2 displays the parameter settings for the EIS and NFR measurements. As previously mentioned, NFR measurements can begin before the polarization fully subsides. To obtain harmonic signals with a good signal-to-noise ratio for extracting battery aging characteristics, a large amplitude current of 5 A is used, which is much larger than the 0.5 A excitation current required for EIS measurements. The maximum SOC parameter for NFR measurement is set at 80 % to prevent overcharging of the battery due to high currents.

3. Methods of analysis

3.1. Methodology introduction

A summary of four FRA techniques is shown in Fig. 2. EIS is a non-invasive technique used to characterize Libs. It involves applying a small-amplitude current or voltage excitation signal to the battery system and combining the frequency-domain characteristics to obtain the impedance over a wide range of frequencies. This provides information on material properties, interfacial phenomena, and electrochemical reactions. Furthermore, it is directly linked to the potential battery degradation, allowing for the mapping of the battery's state change process. Its broad range of applications enables the identification and tracking the evolution of the battery's decline process in a shorter timeframe [31]. The impedance expression for the given excitation, $I(t) = I_A \sin(\omega t)$, is:

$$Z = \frac{V(t)}{I(t)} = \frac{V_A \sin(\omega t + \phi)}{I_A \sin(\omega t)} = Z_A \frac{\sin(\omega t + \phi)}{\sin(\omega t)} \tag{1}$$

wherein the voltage signals $V(t)$ are current-excited response signals, shifted in phase ϕ, and having different amplitudes V_A.

The function expressing of impedance is presented in a complex form using Euler's formula. The expression for $I(t)$ is provided below:

$$I(t) = I_A e^{j\omega t} \tag{2}$$

The expression for the response of the system voltage $V(t)$ is:

$$V(t) = V_A e^{j\omega t - j\phi} \tag{3}$$

The complex expression for impedance is:

$$Z = \frac{V}{I} = Z_A e^{j\phi} = Z_A(\cos\phi + j\sin\phi) = Z_{\mathrm{Re}} + jZ_{\mathrm{Im}} \tag{4}$$

Fig. 3(a) illustrates the changes in EIS as the SOC varies. The impedance spectrum between the mid-frequency and high-frequency regions (0.187 Hz–10 kHz) is not affected by changes in SOC, but once it falls below a critical value, it changes significantly with SOC. To ensure minimal impact of SOC, it is important to acquire features within this range. Fig. 3(b) illustrates the trend of EIS as the number of cycles increases. The mid-frequency and high-frequency regions appear to be more sensitive to battery aging. The combined analysis indicates that the charge transfer process in the mid-frequency region occurs at the intersection favoured by both. The feature parameters representing this process can be prioritised. To balance the negative effects of SOC and positive effects of SOH on the impedance spectrum, it is necessary to build an effective estimation model. This can be achieved by analyzing the correlation between changes in the impedance spectrum and potential degradation mechanisms, and extracting characteristic quantities that can accurately characterize battery aging. The focus of this work is to improve the generalizability and robustness of the estimated model while maintaining accuracy.

The NFRA method is similar to the EIS acquisition method. However, it differs in that it applies a large current amplitude I_{AC} at a certain frequency range to the battery and detects the higher harmonic response H_i with $i > 1$, in addition to acquiring a voltage signal at the same frequency as the input current. Although the intensity of the harmonics decreases as the number of times i increases, typically only the first two harmonics are considered to reduce the workload and computational load. Therefore, this work focuses solely on the root mean square of the first two harmonics rather than the harmonics themselves. Research has demonstrated that the root mean square is commonly utilised in electrical and energy storage applications. It is considered more reasonable than the sum or square root of the higher harmonics in terms of the overall nonlinearity of the system [48].

$$\overline{H} = \sqrt{\frac{\sum_{i=2}^{3} H_i^2}{2}} \tag{5}$$

Fig. 3(c) and (d) illustrate the variation of $\overline{H}$ with SOC and SOH, respectively. As the cell ages, $\overline{H}$ exhibits a strong regularity in the low-frequency range (0.1Hz–2.03 Hz), consistent with the patterns of change observed in the second and third harmonics [5]. This passage discusses the validation of the feasibility of $\overline{H}$. The nonlinear response is limited to nonlinear processes of the system, such as the charge transfer process. The high-frequency counterpart of the battery impedance is transferred from inductive to capacitive behavior [31]. This can result in additional nonlinearities, which can negatively impact the study. Designing a high-precision and strong applicability model based on NFRA data to estimate SOH and reduce the influence of SOC on the selected features is a focus of this paper.

The traditional analysis approach involves incorporating both EIS and NFR full frequency band data into the model for training and testing. However, this can overload the workload and impede the progress of machine learning in the field of Libs. The variability of information across frequency points differs significantly, and some frequency points may contain duplicated information. To accelerate industry

Table 2
Measurement details for EIS and NFR.

Parameters	EIS	NFR
Temperature (°C)	23	23
I_{AC} (A)	0.5 (0.17C)	5 (1.7C)
f-rage (Hz)	0.01–10 k	0.1–337
n_p for f >66 Hz	10	10
n_p for f < 66 Hz	10	5
SOC (%)	20,35,50,65,80,100	20,35,50,65,80

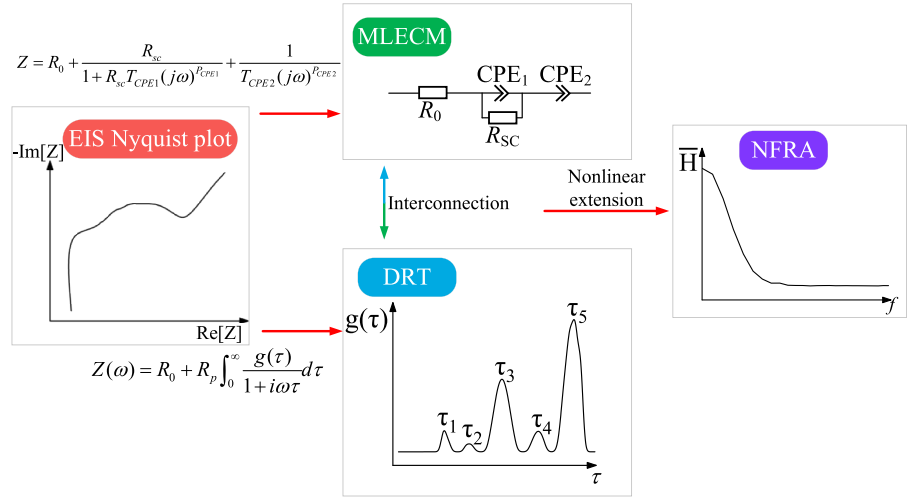

Fig. 2. Summary of FRA techniques.

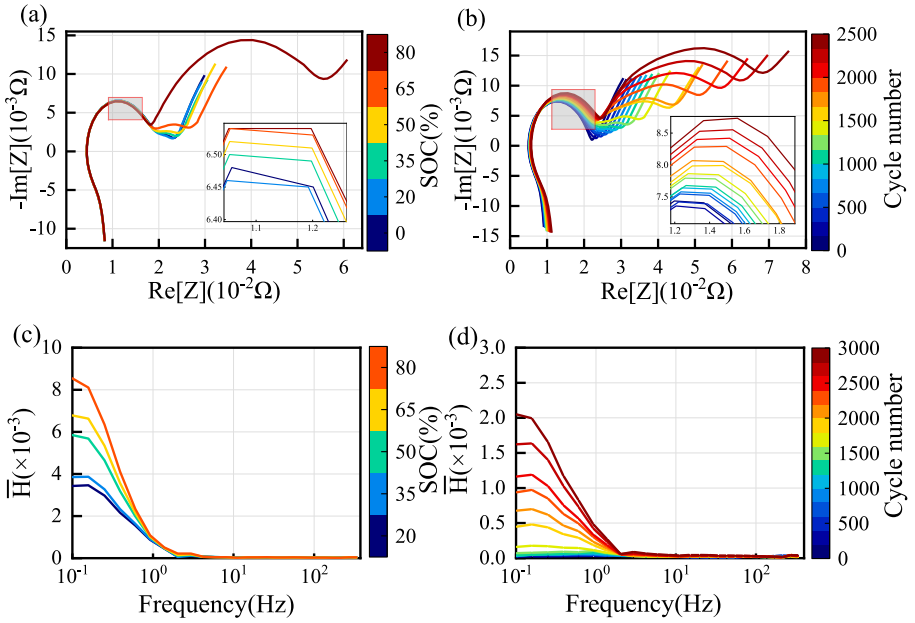

Fig. 3. EIS curve under different variables: (a) SOC; (b) Cycle number. $\overline{H}$ curve under different variables: (c) SOC; (d) Cycle number.

development, this paper proposes extracting a feature to characterize the richness of information across the entire frequency band or accurately within a small frequency range. The method used in this paper streamlines the data to some extent.

3.2. Validity of impedance data

To demonstrate the validity of impedance spectral data, it is crucial to verify that the source of the data, namely the battery system, is linear, stable, and provides finite impedance values within a specific frequency range [50]. This section introduces the Kramers-Kronig (KK) relation, which is used to assess the validity of impedance data by comparing the deviation of the real and imaginary parts of the obtained impedance data to the measured values [51]. The KK relation calculates the real and imaginary parts of the impedance spectrum of a linear, time-invariant, and causal system using the following equation [52]:

$$Z_{Re}(\omega) = \frac{2}{\pi} \cdot \int_0^{\infty} \frac{\omega' \cdot Z_{Im}(\omega')}{\omega^2 - \omega'^2} d\omega' \tag{6}$$

$$Z_{Im}(\omega) = \frac{-2}{\pi} \cdot \int_0^{\infty} \frac{\omega' \cdot Z_{Re}(\omega')}{\omega^2 - \omega'^2} d\omega' \tag{7}$$

where, $\omega = 2\pi f$, f is frequency. Therefore, the frequency f can be used to express the value of KK.

The KK residual is the difference between the computed real and imaginary parts' KK values and the measured KK value. Fig. 4(a) and (b) display the KK residuals of the real and imaginary components of the impedance of a lithium-ion battery. The KK residuals are generally accepted as valid impedance data within a 3 % margin. The plots of both are enlarged and found to have KK residuals of less than 3 % at frequencies greater than or equal to 1.4744 Hz for both the imaginary and real parts. The impedance data for the mid-frequency and high-frequency range are suitable for further calculations.

3.3. Modeling MLECM

To deeply investigate the differences in the impedance spectra of different SOC and SOH, an equivalent circuit model is used to fit the EIS.

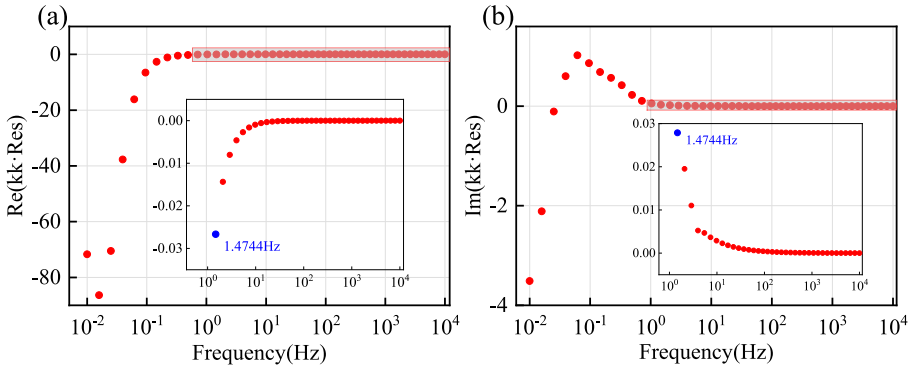

Fig. 4. Relative Kramers-Kronig residuals of two impedance coordinates: (a) real part; (b) imaginary part.

The fitted parameters' values are obtained using ZView software. The original Randles model comprised four components: inductor L, ohmic resistor R_0, charge transfer resistor R_{CT} and a double layer capacitor in parallel, as well as Warburg impedance Z_W [53]. TECM, which has a wide range of applications, has evolved through development, as shown in Fig. 5(a). In a previous study [54], MLECM is proposed as shown in Fig. 5 (b). In contrast to TECM, MLECM eliminates the stray inductance caused by analogue collectors and conductors, which affects the high-frequency region. According to literature [35], TECM typically selects the charge transfer resistance R_{CT} as the aging characteristic, which represents the mid-frequency. This verifies the inference made in Section 3.1. However, the contribution of the SEI film resistance R_{SEI} to battery capacity degradation is often overlooked by R_{CT}. Therefore, the two ZARC elements in the mid-frequency region have been optimized into one ZARC. Additionally, an R_{SC} that incorporates both R_{SEI} and R_{CT} has been proposed. In the mid-frequency region of the EIS, the main processes represented are the transfer of battery charge and the diffusive migration of lithium ions [55]. To summarize, R_{SC} is the circuit element parameter with the greatest potential for characterising battery aging.

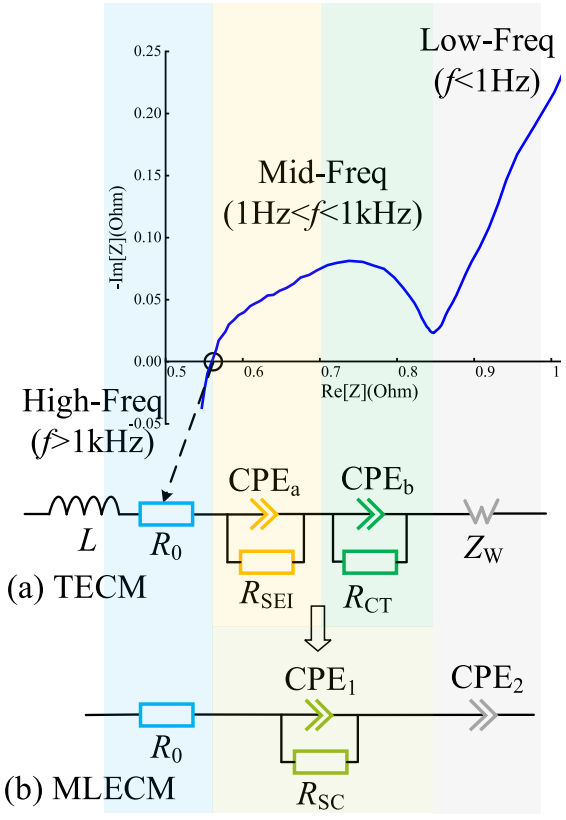

Fig. 5. Typical EIS and the correspondence to ECMs: (a)TECM; (b)MLECM.

ECMs are always generated theoretically, so deviations from the actual operating performance of Libs are unavoidable. This is especially true at low frequency, where accuracy is more difficult to achieve. Although the electrochemical significance of the ECM is not fully characterized, a more complete ECM can still provide sufficient state estimation information [43]. Thus, a relatively complete model is developed in this section.

In Fig. 5(b), battery impedance of MLECM expression is given by:

$$Z_{MLECM} = R_0 + Z_{ZARC} + Z_{CPE2} = R_0 + \frac{R_{sc}}{1 + R_{sc}T_{CPE1}(j\omega)^{P_{CPE1}}} + \frac{1}{T_{CPE2}(j\omega)^{P_{CPE2}}} \tag{8}$$

where T and P are parameters of the CPE, respectively. j is the imaginary units, and ω is the angular frequency.

3.4. DRT fitting

Due to the aggregate nature of the battery system, the ECM approach may be somewhat dependent and can result in analytical errors on the battery. The time constants can be effectively distinguished by the DRT obtained through quantitative back-convolution of the EIS. Different time constants represent various relaxation properties that distinguish between kinetic processes in the battery. These processes include lithium-ion conduction, adsorption and release at the interface. This property can effectively improve the accuracy of kinetic interpretations on various time scales [42].

The DRT plot displays five local maxima (i.e., peaks), each representing the resistance contribution of one of the battery's internal kinetic processes to the total polarization resistance of the battery [46]. The time constant τ characterises each polarization process horizontally, with the relationship between τ and frequency f shown as follows [42]:

$$\tau = RC = \frac{1}{2\pi f} \tag{9}$$

During the construction of the equivalent circuit, it is commonly assumed that the ohmic impedance R_0 and the polarization impedance R_{pol} are connected in series to simulate the convergence impedance of the electrochemical system [42]. In DRT technology, R_{pol} is commonly interpreted as the series form of several RC parallel circuits. When considering the inductance L, the total circuit impedance can be expressed as follows:

$$Z(f) = R_0 + 2\pi fLi + Z_{pol}(f) \tag{10}$$

The impedance of each RC parallel circuit can be expressed as $\frac{R}{1+j\omega\tau}$. If each R is described by a relaxation distribution function $g(\tau)$, after extracting the polarization impedance R_{pol}, the total circuit impedance is denoted as:

$$Z(f) = R_0 + 2\pi fLi + R_p \int_0^{\infty} \frac{g(\tau)}{1 + 2\pi f\tau i} d\tau \tag{11}$$

$$\int_0^{\infty} g(\tau) d\tau = 1 \tag{12}$$

Ignoring the impurity inductance L, the final expression is:

$$Z(f) = R_0 + R_p \int_0^{\infty} \frac{g(\tau)}{1 + 2\pi f\tau i} d\tau \tag{13}$$

3.5. Machine learning frameworks (MLF)

The Impedance HFs-based model's aging parameters is the certain frequency point or range. MLECM HFs-based model's parameter is the certain circuit element. DRT HFs-based model's parameter for the peak parameter at some time constant. NFRA HFs-based model's parameter is the $\overline{H}$. The Gaussian Process Regression (GPR) algorithm is used to build an MLF for estimating capacity. The corresponding flowchart is shown in light red in Fig. 6.

First, given n pairs of training sets $D = \{(x_i, y_i), i = 1, 2, \ldots, n\}$ consisting of input x_i and output y_i. The predictive distribution for calculating the unknown observation $\hat{y}$ is provided based on the test metric $\hat{x}$. $X = [x_1, x_2, \ldots, x_n]$ is defined as the training set and $Y = [y_1, y_2, \ldots, y_n]$ as the test set. Both methods use x_i as the aging parameter, while y_i represents the SOH of the battery in its current aging state. The model is trained using normalized inputs, which are adjusted for mean and standard deviation. This eliminates the potential impact of magnitude on the results.

Second, the GPR model performs nonparametric regression using a gaussian process. It defines an independent and identically distributed gaussian noise $y_i = f(x_i + \varepsilon_i,\ where\ \varepsilon_i \sim \boldsymbol{N}(0, \sigma^2)$. The output $F = (f(x_1), f(x_2), \ldots, f(x_N))$ is defined as a gaussian random field $F \sim \boldsymbol{N}(0, K)$, where $K_{ij} = k(x_i, x_j)$ is the covariance kernel. The kernel indicates the proximity between the points x_i and x_j.

The joint distribution of the training set $\{(x_i, y_i), i = 1, 2, .., n\}$ and the output of the test set is:

$$\begin{bmatrix} Y \\ \hat{y} \end{bmatrix} = \boldsymbol{N}\left(0, \begin{bmatrix} K(X,X) + \sigma^2 I & K(X,\hat{x}) \\ K(\hat{x},X) & K(\hat{x},\hat{x}) \end{bmatrix}\right) \tag{14}$$

Finally, the predicted mean is generated by adjusting the training set over $\hat{x}$:

$$\hat{y}\_mean = K(\hat{x},X)\left(K(X,X) + \sigma^2 I\right)^{-1} Y \tag{15}$$

Meanwhile, its predicted variance is:

$$\Delta^2 = K(\hat{x},\hat{x}) - K(\hat{x},X)\left(K(X,X) + \sigma^2 I\right)^{-1} K(X,\hat{x}) \tag{16}$$

This is a unit of measurement that includes uncertainty. A confidence range is provided based on the prediction variance, typically with a 95 % range for guaranteed accuracy.

After obtaining the test results, the training dataset is used to develop an empirical model. A suitable polynomial model is then selected based on the principle of minimizing the fitting error [56,57] :

$$SOH = a + \sum_{i=1}^{n} bx_i + \sum_{i=1}^{n} cx_i^2 + \ldots \tag{17}$$

To improve enhance the credibility of the results, we applied the "1/5″ principle during the model building and validation phase. The dataset is divided randomly into five parts. One part is used as a test set, and the remaining four parts are used as a training set to train the model. This ensures that all the data are used as an over-test set. The four models discussed in this section adhere to this principle. The paper provides the mean of the estimation result and RMSE for a randomized set of 5 testing data.

3.6. Curve fitting framework (CFF)

This section describes the methodology for extracting aging parameters using the four techniques under CFF. The methodology involves using the CF of EIS and NFR work in MATLAB R2020b and the CF toolbox.

The aging characteristics selected for each of the four CF-based methods are identical to the aging parameters mentioned in the previous section. The empirical model is determined based on the high sim-

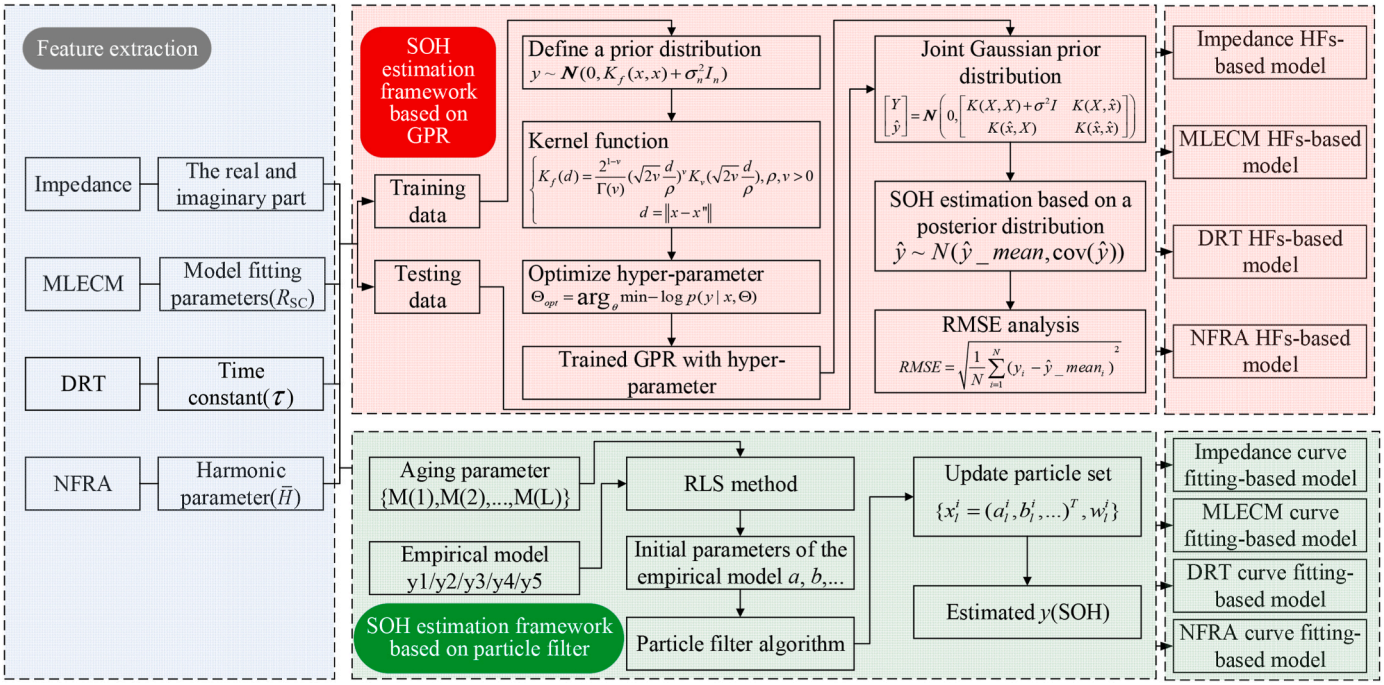

Fig. 6. SOH estimation framework based on FRA.

ilarity of the normal aging of the battery, according to the principle of minimizing the fitting error of the aging characteristics with the life decay rule [56,57]. Overfitting should be avoided. The following equation shows commonly used empirical models, including the variable exponential model, exponential model, power model, variable power model and rational model [58].

$$\begin{cases} y_1 = ae^{-bx} + c \\ y_2 = ae^{bx} + ce^{dx} \\ y_3 = ax^b \\ y_4 = ax^b + c \\ y_5 = \dfrac{a}{x+b} \end{cases} \tag{18}$$

where x is the battery capacity decline characterization quantity, y is the battery SOH, and a, b, … are the parameters of the empirical model. The empirical model characterises the variation of aging parameters and validates its applicability.

Battery capacity changes dynamically and can impact the system, leading to noise interference and other issues [59–62]. The introduction of the particle filter (PF) in this section reduces the overall impact of large deviations of individual particles. If a small range of particles exhibits dispersion when in the dataset, the PF will not have much impact on the results. The desired result is achieved only when there is an excess of dispersed particles, which is nearly impossible to achieve in practice [9]. Introducing PF in the SOH estimation framework based on the CF can effectively improve the stability and robustness of the system.

The algorithm for estimating the SOH of the battery using the aging parameters and the PF algorithm consists of the following steps. The corresponding flowchart is shown in the light green part of Fig. 6. The first step is to obtain battery aging parameters. Then, a recursive least squares (RLS) algorithm is used to fit the data from the initial cycle and the Lth cycle to determine the initial parameters of the model, such as a, b, and so on. Next, the PF algorithm incorporates the initial parameters of the empirical model and aging parameters from the initial cycle to the Lth cycle through the particle set $\left\{x_l^j = \left(a_l^i, b_l^i, \ldots\right)^T, w_l^j\right\}$ and the output-filtered aging parameters $M(l), l = 1, 2, 3, \ldots, L$ for real-time updating and adjustment. Finally, the SOH value, also known as the individual particle y values, is calculated iteratively from the particle set of the Lth cycle using the above equation.

Battery data from various institutions exhibit variability depending on the cycling equipment, EIS and NFRA measurement tools used. The inconsistency in battery aging rates and measurement data between institutions is evident. Thus, the CFF used to estimate SOH in this section is categorized by institution in the results, which differs from the presentation of the results in the previous section.

4. SOH estimation model based on FRA

Using on the data presented in Chapter 2, two SOH estimation frameworks based on ML and CF are developed. Eight estimation models are evaluated, and their respective strengths and weaknesses are compared.

4.1. ML-based SOH estimation framework

4.1.1. Impedance HFs-based model

This section presents the development of the Impedance HFs-based model using EIS raw data. To reduce the workload, the HFs are extracted from the EIS diagram to represent the impedance spectrum information across the full frequency band. Section 3.1 summarises that the variation in SOC does not significantly affect the impedance at mid-frequency and high-frequency. At 15.8489 Hz, significant differences in EIS occur at different SOCs. This section extracts HFs from EIS regions with frequencies greater than the threshold, in combination with the KK relationship. A schematic diagram of the HFs as shown in Fig. 7 is made based on past contributions [63].

To identify the seven HFs with the strongest correlation to the battery. Fig. 8 displays the Spearman rank correlation coefficients for each of the seven battery aging factors at various SOCs. The shade of colour indicates the absolute value of the correlation coefficient's proximity to 1, implying a close relationship between the health factor in the area and battery capacity degradation. Conversely, the lighter the colour, the weaker the correlation. The results indicate that the negative imaginary part value of the critical point corresponding to HF_2 has the highest correlation of 0.99133 at SOC = 20 %. This is in line with the conclusion in section 3.1 that the features extracted from the mid-frequency region are assumed to effectively balance SOC and SOH. The impedance HFs-based model uses HF_2 at 20 % SOC as the characteristic parameter due to workload reduction.

The MLF is trained and tested using HFs. The results are presented in Fig. 9(a), which shows that the model's RMSE is 1.25 % under the principle of "1/5″, meeting the accuracy requirement for real vehicle operation.

4.1.2. MLECM HFs-based model

The circuit elements in the MLECM presented in Section 3.3 are chosen as the parameters to characterize aging in this section. To prevent any impact on the overall estimation results, we will exclude data with a normalized residual greater than 5×10^{-3} during the fitting process as well as data from cycles where the fit fails. Other techniques also follow a similar approach to data handling, but will not be discussed further. Fig. 10(a) shows Spearman's rank correlation coefficient, which characterises the relationship between six circuit element parameters and battery ageing. The absolute value of the coefficient is used for the vertical axis. Among the different institutions, only the R_{SC} coefficient consistently maintains a high level, while the other circuit elements show mixed performance and cannot be used to characterize the aging of the batteries. The coefficient for the R_{SC} as a whole is 95.87 %, which is higher than the coefficients for any other circuit element parameters. The calculated value is not a simple average of the four institutions, but rather the result of an overall analysis. The performance aligns with the expected results. Section 3.3 presents the R_{SC}, which combines SEI film resistance and charge transfer resistance as characterization parameters for battery aging [31,34]. It represents the process of charge transfer from the electrolyte to the electrode surface and from the electrode

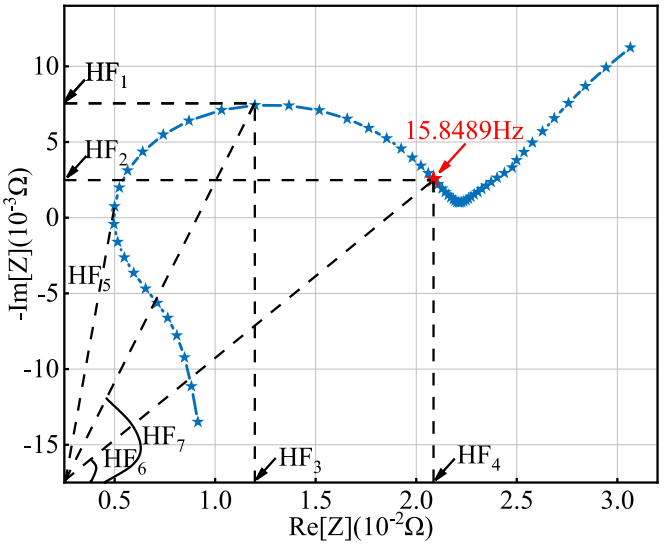

Fig. 7. Schematic diagram of the selected HFs for the impedance HFs-based model.

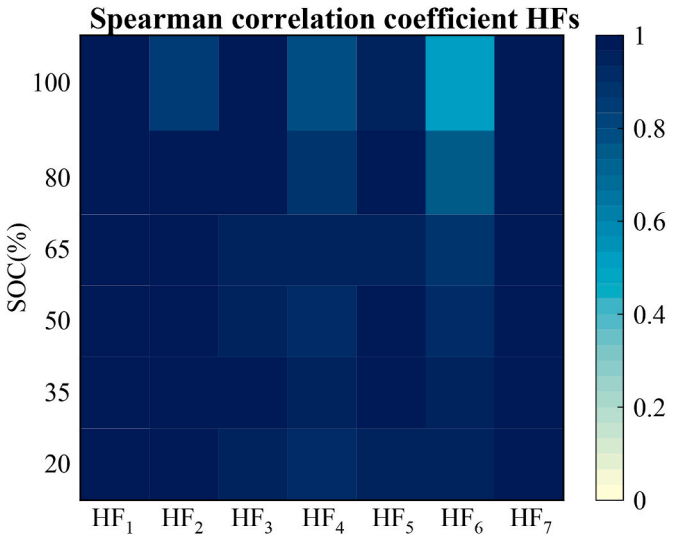

Fig. 8. The absolute value of the Spearman rank correlation coefficient between battery aging and Impedance HFs.

surface to the active substance of the electrode body [64].

The SOH is estimated by the MLF and the results are presented in Fig. 9(b). Although the SOH estimation cannot be lower than 0.6, the estimation error is significantly reduced, and the RMSE is always less than 0.85 %. These results demonstrate that proposed MLECM HFs-based model has consistently good prediction performance.

4.1.3. DRT HFs-based model

DRT is a novel technique based on impedance spectroscopy that characterises the chemical reaction process within a battery. Fig. 11 shows the relationship between the DRT diagram and the EIS diagram. The DRT diagram is divided into high-frequency, mid-frequency and low-frequency regions by the red dashed line, which corresponds to the Nyquist diagram. Peak P1 represents the polarization in the high-frequency region. Peaks P2 and P3 represent the reaction in the mid-frequency double arc. The remaining peaks represent the diffusion process [46]. However, the correspondence between regions is not exact with the frequency of the impedance spectrum. As per the *KK* relation in Section 3.2, only the mid-frequency and high-frequency regions are considered, where $\tau < 1$ ($f > 1/2\,\pi$).

The horizontal and vertical coordinates of P1, P2 and P3 in the DRT diagrams serve as aging characteristics for Spearman correlation analysis. The results are presented in Fig. 10(b). The Spearman rank correlation coefficients for y (P3) have values greater than 90 % at all four institutions, even reaching 99 % at institutions 1 and 2. The coefficient of y (P3) is 91.87 %, which is significantly higher than the remaining five features. It has been demonstrated that the vertical coordinate of point P3 in the DRT plot is closely linked to battery aging. Additionally, since P3 corresponds to R_{CT}, which is highly correlated with battery aging, it can be concluded battery aging is indeed related to RCT, which is consistent with the findings of ref [35].

The feature parameter y (P3) is selected and inputted into the MLF for training, testing the data, and drawing conclusions, as illustrated in Fig. 9(c). The DRT HFs-based model shows high feasibility, with an average RMSE of 1.04 % across the four institutions.

4.1.4. NFRA HFs-based model

The calculations that follow are based on the $\overline{H}$ parameter proposed in Section 3.1 for characterization. Fig. 12 presents the Spearman's rank correlation coefficients of $\overline{H}$ at different SOCs. The maximum value of 0.98411 is observed at a frequency of 0.604 Hz and 50 % SOC. This frequency point also happens to be in the low-frequency region (0.1 Hz–2.03 Hz) where it was concluded in section 3.1 that $\overline{H}$ exhibits a strong regularity. Thus, in this section, we have chosen $\overline{H}$ at 50 % SOC at

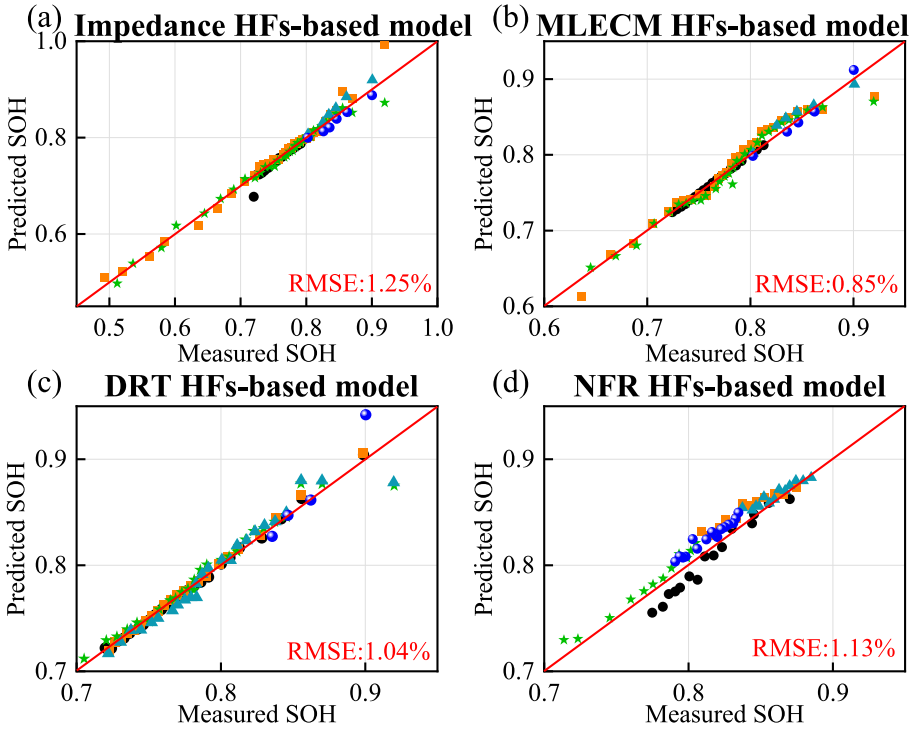

Fig. 9. Average RMSE under the ML-based SOH estimation framework: (a) Impedance HFs-based model; (b) MLECM HFs-based model; (c) DRT HFs-based model; (d) NFR HFs-based model. Each model includes five shapes (circle, square, star, triangle, ball) and five colours (black, orange, green, light blue, blue) that represent the RMSEs of a random 20 % of the dataset as the testing data. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

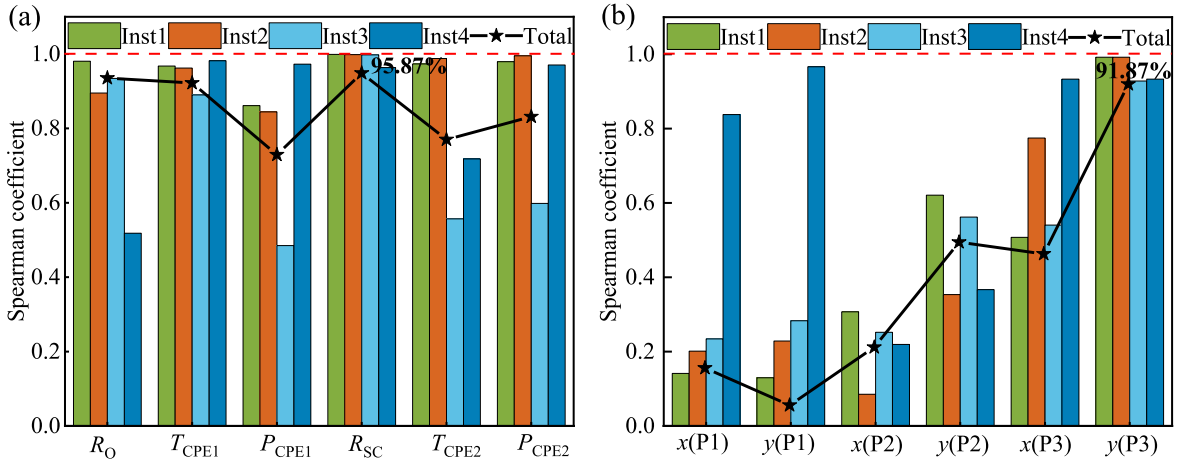

Fig. 10. Spearman's rank correlation coefficients between different parameters and battery aging: (a) MLECM elements; (b) coordinates of the three characteristic points of the DRT. *x* represents the horizontal coordinates of certain points, *y* represents the vertical coordinates; bar represents the institution, and line analyzes them as a whole.

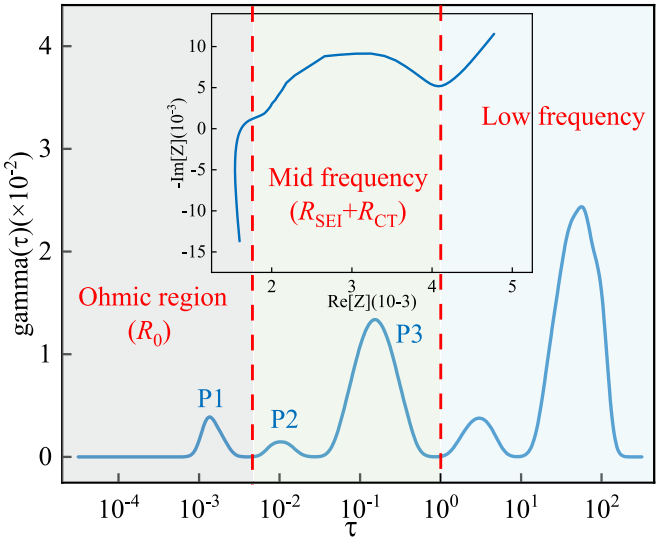

Fig. 11. Correspondence between DRT and EIS diagrams.

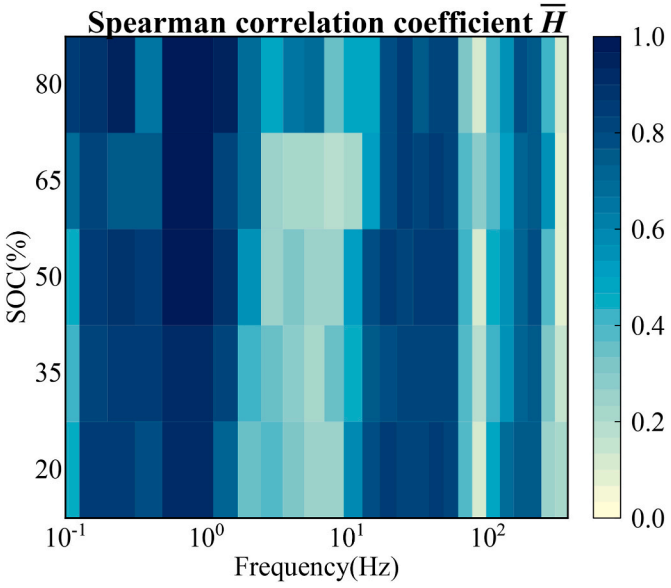

Fig. 12. The absolute value of the Spearman rank correlation coefficient between battery aging and $\overline{H}$.

this frequency point as the characteristic vector.

The MLF is used to train and test the data, as shown in Fig. 9(d). The test set consisted of 20 % of the data randomly selected on five occasions. The average RMSE of the NFRA HFs-based model estimation results is 1.13 %, indicating that this method is feasible for estimating SOH.

4.2. CF-based SOH estimation framework

4.2.1. Impedance CF-based model

This section includes the same features (HF_2 at 20 % SOC) as those in section 4.1.1 to fit the objective function. CF is performed using each empirical model presented in Equation (18). The goodness of fit value R^2 and RMSE are used to evaluate the accuracy of the empirical model in fitting the experimental data. The value of R^2 ranges from 0 to 1. A value closer to 1 indicates a better fitting effect, while a value closer to 0 indicates a worse fitting effect. Similarly, a smaller RMSE value indicates a better model fitting effect. The formulas for R^2 and RMSE are as follows:

$$RMSE = \sqrt{\sum_{i=1}^{N}\left(y_{mea} - y_{pre}(i)\right)^2 / N} \tag{19}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left[y_{mea}(i) - y_{pre}(i)\right]^2}{\sum_{i=1}^{N}\left[y_{mea}(i) - \overline{y}_{mea}(i)\right]^2} \tag{20}$$

where $y_{mea}(i)$ represents the measured value of SOH, $y_{pre}(i)$ represents the model estimation of SOH, $\overline{y}_{pre}(i)$ represents the mean of the model estimation of SOH, and N represents the number of samples.

To avoid function fitting failures and unify the magnitude of feature parameters, we normalize them within the range of 1–10. Fig. 13 shows the fitting results for a certain institution under different empirical models.

Fig. 14(a) shows the distribution of the average fitting result index for the Impedance CF-based model. It is evident that all models, except for model **I,** fit the SOH decline track well. The model **II,** with the most

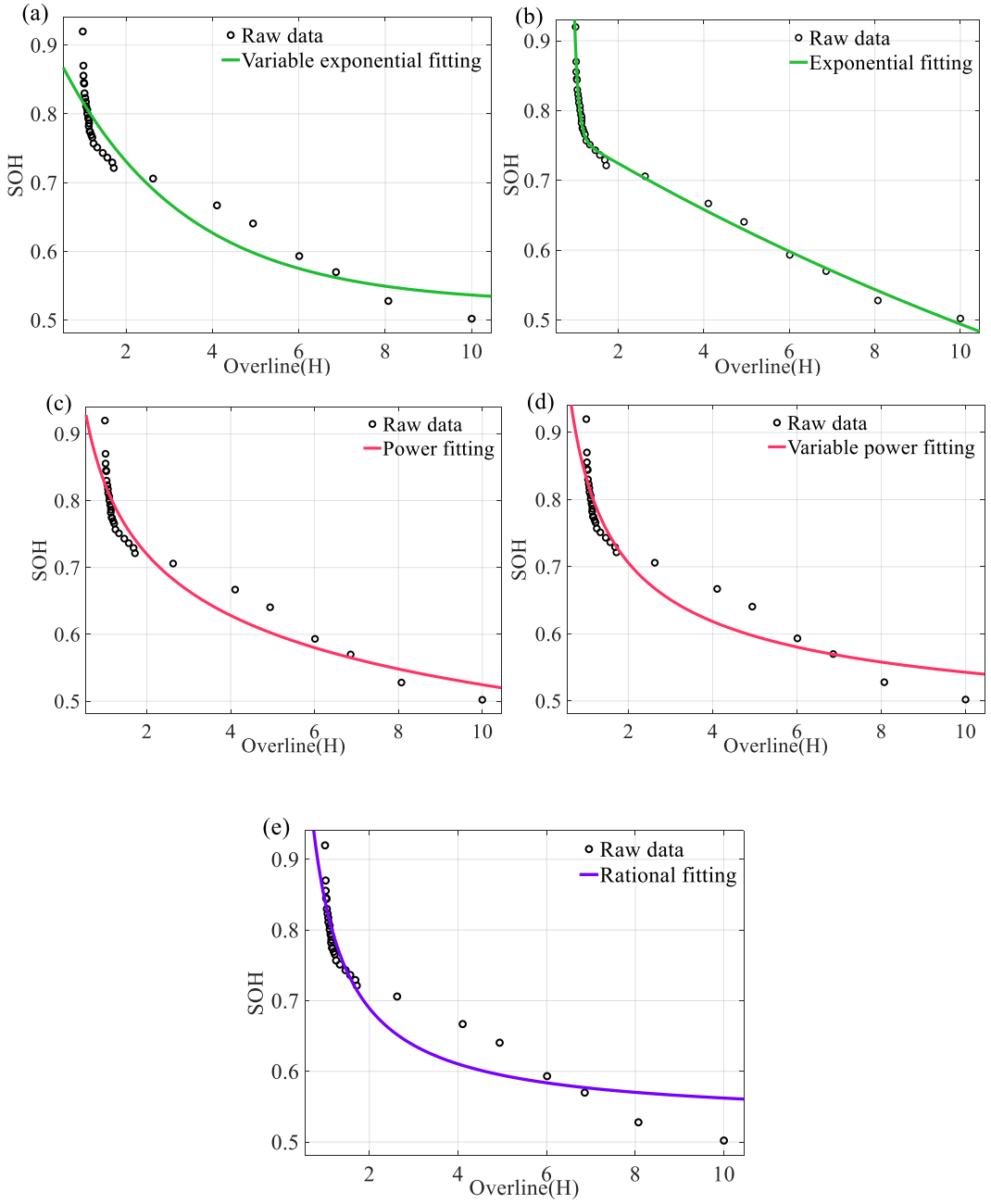

Fig. 13. Fitting results for a certain institution under 5 empirical models: (a) Variable exponential model; (b) Exponential model; (c) Power model; (d) Variable power model; (e) Rational model. The horizontal coordinate 'Overline(H)' represents $\overline{H}$.

parameters, fits the track the best. However, if there are too many parameters overfitting may occur, which can result in the noise characteristics of the measurement data being display. Additionally, an increase in parameters can lead to a significant computational load, which is undesirable. Thus, it is important to balance and quantify the relationship between model complexity and fitting accuracy. This can be achieved by incorporating the Hannan-Quinn (HQ) criterion for a comprehensive analysis of each empirical model.

The HQ criterion considers the impact of sample data length based on the Akaike Information Criterion. This prevents the issue of disregarding model complexity in the pursuit of model fitting accuracy when the data amount is excessive. Based on the Bayesian Information Criterion, this phenomenon of poor fitting accuracy but optimal results can be avoided by reducing the length of sample data. The formula for the HQ calculation is as follows :

$$HQ = K\ln(\ln(n)) + 2\ln(S) \tag{21}$$

where K represents the number of model parameters and n represents the sample length and S=RMSE. The equation's right-hand side has two terms. The first term is a function of the number of model parameters and sample length, while the second term is a function of RMSE. The objective is to obtain an empirical model with high fitting accuracy and low parameters. The HQ value is an indicator of the model's quality, with lower value indicating better performance. Fig. 14 (b), (d), (f) and (h) display the average HQ values for four institutions.

Fig. 14(b) displays the high-quality performance of the Impedance CF-based model. The minimum HQ is achieved by the model **IV**, while the model **I** has the maximum value. The HQ value is determined by the number of model parameters and the RMSE when the sample lengths are all the same. Despite having four parameters, the model **II** performs

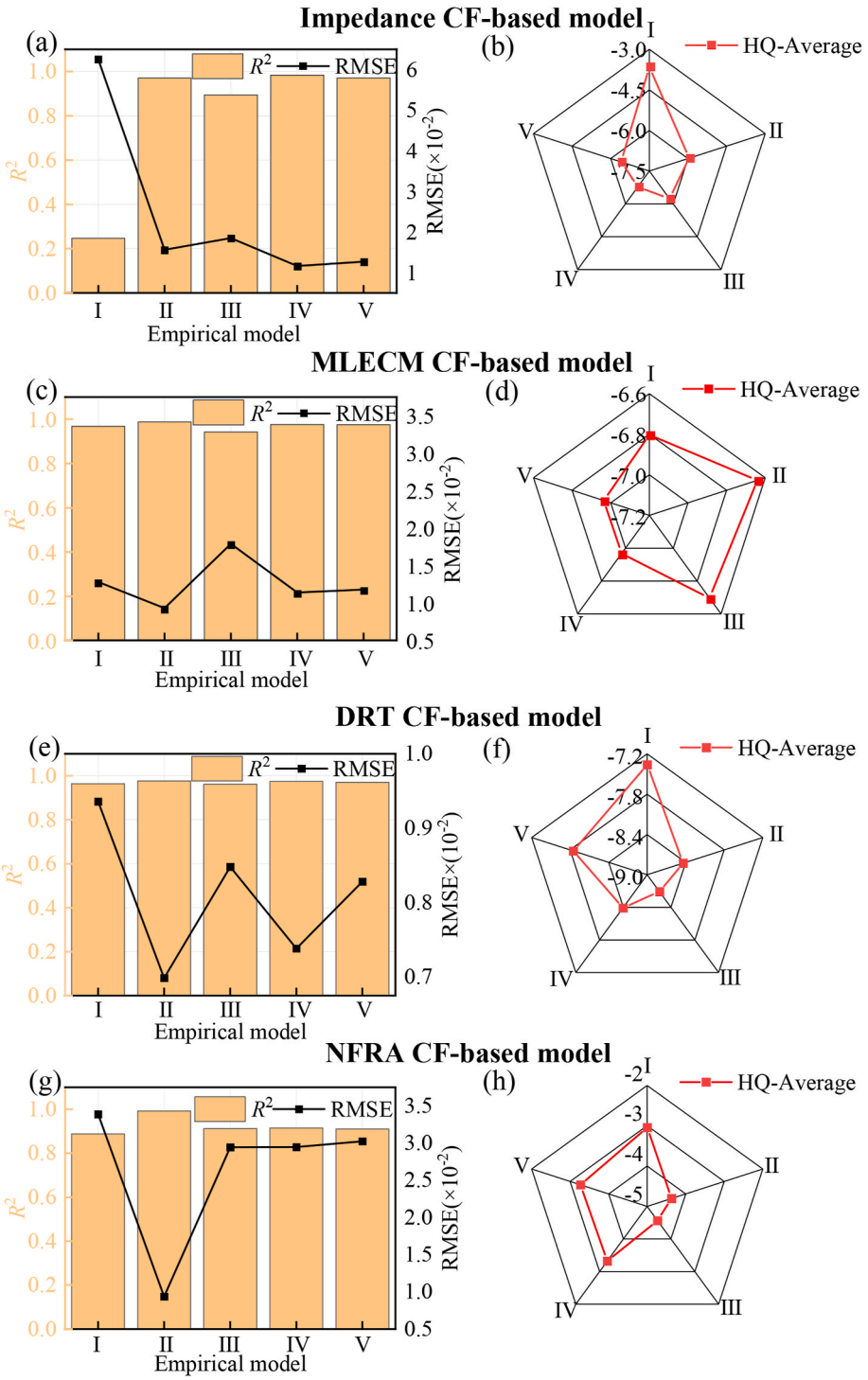

Fig. 14. RMSE, R^2 and HQ values for the four models under different empirical models. The four left-hand side plots show the RMSE and R^2, and the four right-hand side plots show the average HQ values. Empirical models **I**: Variable exponential model; **II**: Exponential model; **III**: Power model; **IV**: Variable power model; **V**: Rational model.

poorly on a combination of different institutions. The prominence of the HQ values is reduced because four parameters excessively weight HQ. Although the model **III** has only two model parameters, its RMSE is also larger. Therefore, its HQ value be minimized. The model **I** has a significantly larger RMSE than the other models, despite having only 3 parameters. This result in it has the highest HQ values. The model **IV** has the smallest RMSE due to three model parameters, resulting in the smallest HQ value. Thus, the Impedance CF-based model is fitted through model **IV** using HF_2 at 20 % SOC.

Impedance CF-based model begin with PF to estimate the parameters of IV in real-time. SOH estimation is then performed, and the results are displayed in Fig. 15(a). The average RMSE of 1.14 % indicates that the model can effectively estimate the SOH of batteries.

4.2.2. *MLECM CF-based SOH model*

The characterization parameter used is the same as in Section 4.1.2: the R_{SC} at 20 % SOC. Following In the same steps as outlined in Section 4.2.1, the R_{SC} curves are fitted using five empirical models. The fit

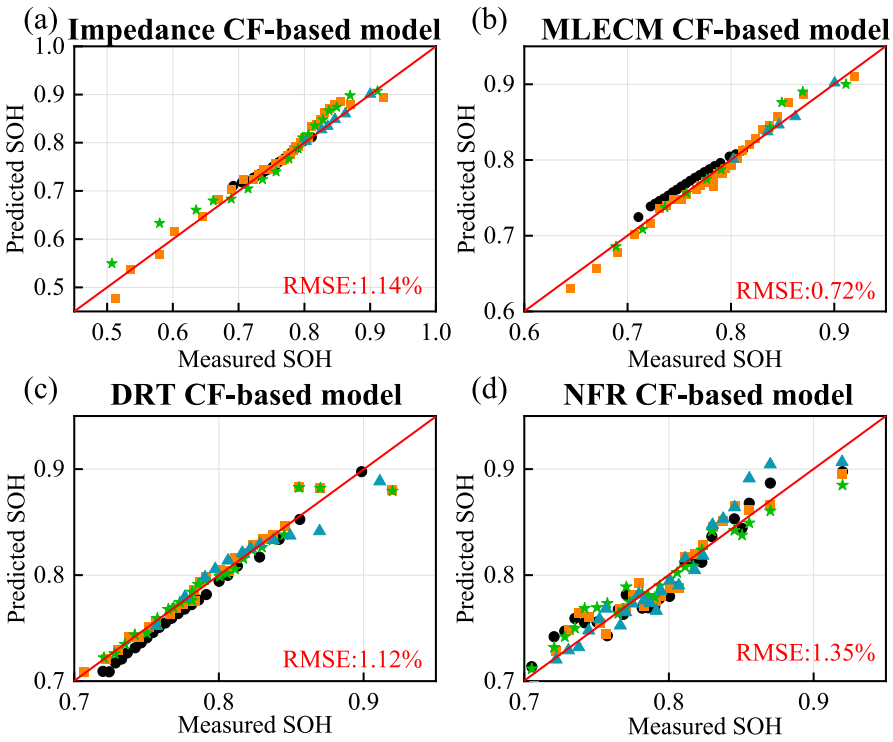

Fig. 15. Average RMSE under the CF-based SOH estimation framework: (a) Impedance CF-based model; (b) MLECM CF-based model; (c) DRT CF-based model; (d) NFR CF-based model. Each model includes four shapes (circle, square, star, triangle) and four colours (black, orange, green, light blue) that represent the RMSE for each of the four institutions. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

performance is displayed in Fig. 14(c). All five empirical models are found to fit the R_{SC} trend better, indicating that the R_{SC} trend aligns more closely with the characteristics of the empirical models. When examining Fig. 14(d), it is evident that the lowest HQ value of −6.9668 is associated with IV. The reason for the model **II** and **III** struggling to show good results in HQ is due to having too many parameters and a slightly higher RMSE, respectively. On the other hand, the model **IV** achieves the smallest HQ value with a slightly smaller RMSE compared to the remaining three models, all of which have three parameters.

Fig. 15(b) displays the estimation results of the MLECM CF-based model under the MLF, which maintains the same SOH estimation range as Fig. 9(b). However, the estimation error is only 0.72 %, which is 0.13 % smaller than the previous estimate. This is because R_{SC} performed better than MLF on the CFF. This is further demonstrated by the fit of the R_{SC} to the five empirical models. In cases where there is insufficient data to support the training of models in ML, CFF is often a better choice. This greatly advances the development of research in this field.

4.2.3. DRT CF-based model

This section uses the same feature parameters as described in Section 4.1.3. Specifically, it focuses on the vertical coordinates of the P3 point within the DRT plot. The feature is fitted using various empirical models, and the performance index of the fit is displayed in Fig. 14(e). Although The model **II** has the most parameters, it provides the best fit for this section. Fig. 14(f) displays the HQ values for all empirical models. The value of −8.44 for model **II** is inferior to that of model **III**, which is −8.69. Additionally, the excess parameters of model **II** lead to increased complexity. Therefore, in this section, model **III** is selected to fit the feature parameter.

The DRT CF-based model uses the PF algorithm to estimate the parameters of the model **III** in real-time for SOH estimation. The results are presented in Fig. 15(c). The RMSE average across the four institutions is 1.12 %, which is inferior to the estimation of the DRT HFs-based SOH model in section 4.1.3. The results indicate that the features extracted using DRT in this section are not as effective on CFF as on MLF.

4.2.4. NFRA CF-based model

The same characterization parameter as described in Section 4.1.4 has been chosen for this section. Specifically, $\overline{H}$ at 50 % SOC in 0.604 Hz is used. Fig. 14(g) shows the fitting results of the R_{SC} curves using five empirical models. From Fig. 14(h), it is evident model **III** with two parameters has the smallest HQ value of −4.565. However, model **II** has the smallest RMSE but a combined HQ value of −4.3696. The reason for the difference in the percentage of the model parameter K in HQ is due to the difference in data length, specifically the value of n in Eq. (14).

The NFRA CF-based model uses the PF algorithm to estimate the parameters of the model **III** in real-time, which in turn performs SOH estimation. The results are displayed in Fig. 15(d). The estimated SOH range from 0.7 to 0.95, with an average RMSE of 1.35 % across the four institutions. This is higher than the NFRA HFs-based model, as the feature parameters used in this section do not perform as well as the MLF in the CFF.

4.3. Comparison of methods

Table 3 shows a comparison of the performance of eight models. The RMSE performance of all eight models, consisting of two frameworks summarized concerning FRA, on the dataset is less than 1.35 %. This suggests that the models have good estimation performance. The selection of SOC for EIS and NFR measurements is based on the Spearman rank correlation coefficient. The battery can be measured with 20 % and 50 % SOC, which means that a full charge or discharge is not necessary. This is in line with the driver's charging and discharging habits.

Table 3 shows that Impedance had the highest average RMSE among the linear FRA techniques, with a value of 1.195 %. However, when comparing the estimation range of SOH% under global conditions, it is evident that Impedance has the largest range, reaching as low as 45 %.

Table 3
Performance comparison of the eight methods.

Method	Impedance		MLECM		DRT		NFRA	
	HFs-based	CF-based	HFs-based	CF-based	HFs-based	CF-based	HFs-based	CF-based
Measurement SOC%	20	20	20	20	20	20	50	50
SOH% estimation range	45–100	45–100	60–95	60–95	70–95	70–95	70–95	70–95
SOH% estimation accuracy	1.25	1.14	0.85	0.72	1.04	1.12	1.13	1.35
Fresh data required	✓	✓	×	×	×	×	✓	✓
Data preprocessing required	×	×	✓	✓	✓	✓	×	×
Full frequency spectrum required	×	×	×	×	✓	✓	×	×
Empirical model	–	Variable power model	–	Variable power model	–	Power model	–	Power model

MLECM has the next largest range while DRT and NFRA have the narrowest range. The combination of the above findings leads to the conclusion that both MLECM and DRT, which are secondary technologies based on EIS, have experienced some degree of information loss during their development. NFRA, as an extension of EIS to the nonlinear domain, provides less information about battery degradation compared to the impedance spectrum. Impedance and MLECM accuracy perform worse when using MLF compared to CFF. DRT and NFRA follow the opposite. The results indicate that the parameter choices of DRT and NFRA are better suited to the properties of the CFF. The selection of estimation methods is not completely fixed. Combining all methods with good estimation accuracy leads to the conclusion that it is difficult to recommend any one technology or estimation framework in isolation due to a certain performance. This finding aligns Chan et al.'s research [5].

The Impedance and NFRA can be estimated solely from raw data. The primary advantage of utilising raw data is that it enables early analysis of frequency points with high feature correlation. This allows for later stage analysis to focus solely on these specific frequency points, thereby reducing workload and improving efficiency. There is no need for preprocessing. The four methods corresponding to MLECM and DRT require transforming the raw data into their respective feature parameters before subsequent calculations can be performed. This reduces the amount of raw data to a certain extent. However, the small improvement in accuracy at the expense of the estimation range does not provide a significant advantage for such methods in this study.

Among the four models corresponding to CFF, only two empirical models are selected. This indicates that fitting the trend of characteristic parameters for different methods with a fixed curve is difficult due to the length of the data and the degree of CF fitting.

Table 4 presents a comparison of the above models using to three different classification methods. For various frameworks, there is not a significant difference in estimation accuracy between MLF and CCF. The former has an average accuracy of 0.015 % higher than the latter Both frameworks are capable of estimating SOH accurately. Combined with Table 3, the CFF controls the RMSE from 0.72 % to 1.35 %. The CFF is largely influenced by how well the empirical model fits the battery characteristic parameters fit. Feature parameters that fit the trend of the empirical model perform better on the CFF, while those that do not perform better on the MLF. It is important to note that this is an objective evaluation based on empirical data. The RMSE for the MLF has a high lower bound and a low upper bound when compared to the CFF. In summary, the MLF is generally more accurate and stable; while the CFF may be less stable but can provide higher estimation accuracy in specific cases. Additionally, the CFF has a lower computational load and requires less data. It is important to note that these evaluations are objective and based on specific criteria.

Table 4
Comparison of methods under three classification conditions for estimation accuracy and application scenarios.

Categorization	Method	Average-RMSE (%)	Application Scenarios
Framework (Classification 1)	MLF	1.0675	a large amount of data
	CFF	1.0825	a small amount of data
Frequency response analysis (Classification 2)	LFRA	1.02	electric vehicle power battery
	NFRA	1.24	retired battery
Impedance (Classification 3)	EIS	1.195	scenes with complex and dynamic
	Non-EIS	1.035	scenes with high accuracy requirements

Table 4 presents the mean values of the estimated RMSE for both linear and nonlinear methods. The study find that the linear technique had an average RMSE 0.22 % lower than the nonlinear technique, indicating slightly better accuracy. Since the former is measured by small amplitude signals, it presupposes the need for sufficient rest periods. Due to the limitations of space, weight, and volume in electric vehicles, it is challenging to conduct online measurements using LRFA methods (i.e., the first three techniques) from the integrated and systematic BMS. Therefore, there is limited potential for the development of LRFA-based techniques in the electric vehicle field. The NFRA technology bridges the industry gap by measuring signals emitted by BMS with a much larger amplitude than before (10 times larger in this paper). This drastically shortens rest time, or even allows for direct discarding, enabling online measurements. In practical applications, it is necessary to achieve a higher level of estimation accuracy, while considering the costs of safety and economy. The technique's estimation accuracy is within the acceptable range, with a deviation of only 1.35 %. Electric vehicle power batteries typically use 80 % SOH as the threshold for battery safety, meaning that the battery needs to be replaced at this point. This technology happens to cover the vast majority of power battery health variations. Batteries that fall below this threshold are referred to as retired batteries. They can be utilised in scenarios where the demand for electricity is not high, such as home energy management, urban clean vehicles, urban traffic lighting systems, and even energy storage power stations. Ensuring a long battery life under the given scenarios places high demands on the SOH estimation range based on economic considerations. The Impedance technology SOH estimation range threshold can be as low as 45 %. In summary, NFRA technology is chosen for areas such as electric vehicles, which are difficult to measure online and have high power requirements. LFR technologies, such as Impedance, are considered for areas such as home energy management, which require long-term operation.

According to Table 3, the impedance technique has the lowest accuracy among the LFRA techniques. However, the estimation range for its SOH is the widest, reaching as low as 45 %, whereas other methods only reach as low as 60 % or 70 %. Impedance can achieve heights that are difficult to reach with other methods, even when accuracy is limited. Therefore, to achieve greater accuracy in estimation, methods other than Impedance should be considered. If the goal is to widen the range of SOH estimation, the Impedance technique can be the best choice.

5. Conclusions and outlook

Spearman's rank correlation coefficient is used to screen potential characterization parameters with the highest correlation between SOC

and battery aging in order to validate the feasibility of all SOH models for predicting battery aging. The SOH estimation framework based on ML is formed by placing the battery data in the GPR model for data training and testing, according to the "1/5" principle in the model training and testing process. The selected empirical model balances model complexity and fitting accuracy by using the smallest HQ value. The initial model parameters are determined by RLS. Then PF is used for real-time updating and adjusting, and the final SOH estimate is output to form a SOH estimation model for lithium-ion batteries based on CF.

The main contributions of this paper are as follows:

1. Impedance-based, MLECM-based and DRT-based methods in the linear range have been integrated, improved, and extended to the nonlinear domain. Additionally, an NFRA-based method has been developed.
2. Four technologies are used to form eight models, which were developed and applied using MLF and CFF. The feasibility of these model is verified using data from four institutions.
3. By applying the same data, the four techniques are effectively analysed for their advantages and disadvantages through side-by-side comparisons.
4. Application scenarios are identified for the techniques by using three classification methods: different estimation frameworks, frequency response linearity, and impedance technique.

The paper's insights based on the existing dataset are limited. In the future development of the industry, a more practical SOH model that applies to a wider range of conditions will be necessary. This will require a higher level of accuracy and robustness from the model.

CRediT authorship contribution statement

Shaojin Wang: Writing – original draft, Methodology. **Jinrui Tang:** Writing – review & editing. **Binyu Xiong:** Investigation, Methodology, Supervision. **Junqiu Fan:** Writing – review & editing. **Yang Li:** Supervision, Investigation. **Qihong Chen:** Writing – review & editing. **Changjun Xie:** Software, Data curation. **Zhongbao Wei:** Data curation, Supervision.

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.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant #52177221.

References

[1] Wang Y, Ni R, Jiang X, Yin M, Zhang D, Xie Z. An electrochemical-mechanical coupled multi-scale modeling method and full-field stress distribution of lithium-ion battery. Appl Energy 2023;347:121444. https://doi.org/10.1016/j.apenergy.2023.121444.
[2] Abu SM, Hannan MA, Hossain Lipu MS, Mannan M, Ker PJ, Hossain MJ, Mahlia TMI. State of the art of lithium-ion battery material potentials: an analytical evaluations, issues and future research directions. J Clean Prod 2023;394:136246. https://doi.org/10.1016/j.jclepro.2023.136246.
[3] Xu C, Zhang F, Feng X, Jiang F, Ren D, Lu L, Yang Y, Liu G, Han X, Friess B, Ouyang M. Experimental study on thermal runaway propagation of lithium-ion battery modules with different parallel-series hybrid connections. J Clean Prod 2021;284:124749. https://doi.org/10.1016/j.jclepro.2020.124749.
[4] Rostami F, Kis Z, Koppelaar R, Jiménez L, Pozo C. Comparative sustainability study of energy storage technologies using data envelopment analysis. Energy Storage Mater 2022;48:412–38. https://doi.org/10.1016/j.ensm.2022.03.026.
[5] Chan HS, Dickinson EJF, Heins TP, Park J, Gaberšček M, Lee YY, Heinrich M, Ruiz V, Napolitano E, Kauranen P, Fedorovskaya E, Moškon J, Kallio T, Mousavihashemi S, Krewer U, Hinds G, Seitz S. Comparison of methodologies to estimate state-of-health of commercial Li-ion cells from electrochemical frequency response data. J Power Sources 2022;542:231814. https://doi.org/10.1016/j.jpowsour.2022.231814.
[6] Zhang C, Kang Y, Duan B, Zhou Z, Zhang Q, Shang Y, Chen A. An adaptive battery capacity estimation method suitable for random charging voltage range in electric vehicles. IEEE Trans Ind Electron 2022;69:9121–32. https://doi.org/10.1109/TIE.2021.3111585.
[7] Gong L, Zhang Z, Li Y, Li X, Sun K, Tan P. Voltage-stress-based state of charge estimation of pouch lithium-ion batteries using a long short-term memory network. J Energy Storage 2022;55:105720. https://doi.org/10.1016/j.est.2022.105720.
[8] Song Y, Peng Y, Liu D. Model-based health diagnosis for lithium-ion battery pack in space applications. IEEE Trans Ind Electron 2021;68:12375–84. https://doi.org/10.1109/TIE.2020.3045745.
[9] Zhengxin J, Qin S, Yujiang W, Hanlin W, Bingzhao G, Lin H. An Immune Genetic Extended Kalman Particle Filter approach on state of charge estimation for lithium-ion battery. Energy 2021;230:120805. https://doi.org/10.1016/j.energy.2021.120805.
[10] Gao Y, Liu K, Zhu C, Zhang X, Zhang D. Co-estimation of state-of-charge and state-of- health for lithium-ion batteries using an enhanced electrochemical model. IEEE Trans Ind Electron 2022;69:2684–96. https://doi.org/10.1109/TIE.2021.3066946.
[11] Wang Q-K, Shen J-N, Ma Z-F, He Y-J. Decoupling parameter estimation strategy based electrochemical-thermal coupled modeling method for large format lithium-ion batteries with internal temperature experimental validation. Chem. Eng. J. 2021;424:130308. https://doi.org/10.1016/j.cej.2021.130308.
[12] Wang P, Yang L, Wang H, Tartakovsky DM, Onori S. Temperature estimation from current and voltage measurements in lithium-ion battery systems. J Energy Storage 2021;34:102133. https://doi.org/10.1016/j.est.2020.102133.
[13] Ragone M, Yurkiv V, Ramasubramanian A, Kashir B, Mashayek F. Data driven estimation of electric vehicle battery state-of-charge informed by automotive simulations and multi-physics modeling. J Power Sources 2021;483:229108. https://doi.org/10.1016/j.jpowsour.2020.229108.
[14] Chen C, Xiong R, Yang R, Shen W, Sun F. State-of-charge estimation of lithium-ion battery using an improved neural network model and extended Kalman filter. J Clean Prod 2019;234:1153–64. https://doi.org/10.1016/j.jclepro.2019.06.273.
[15] Mao L, Hu H, Chen J, Zhao J, Qu K, Jiang L. Online state-of-health estimation method for lithium-ion battery based on CEEMDAN for feature analysis and RBF neural network. IEEE J. Emerg. Sel. Top. Power Electron. 2023;11:187–200. https://doi.org/10.1109/JESTPE.2021.3106708.
[16] Hossain Lipu MS, Hannan MA, Hussain A, Saad MH, Ayob A, Uddin MN. Extreme learning machine model for state-of-charge estimation of lithium-ion battery using gravitational search algorithm. IEEE Trans Ind Appl 2019;55:4225–34. https://doi.org/10.1109/TIA.2019.2902532.
[17] Xia B, Cui D, Sun Z, Lao Z, Zhang R, Wang W, Sun W, Lai Y, Wang M. State of charge estimation of lithium-ion batteries using optimized Levenberg-Marquardt wavelet neural network. Energy 2018;153:694–705. https://doi.org/10.1016/j.energy.2018.04.085.
[18] Xi Z, Wang R, Fu Y, Mi C. Accurate and reliable state of charge estimation of lithium ion batteries using time-delayed recurrent neural networks through the identification of overexcited neurons. Appl Energy 2022;305:117962. https://doi.org/10.1016/j.apenergy.2021.117962.
[19] Chemali E, Kollmeyer PJ, Preindl M, Emadi A. State-of-charge estimation of Li-ion batteries using deep neural networks: a machine learning approach. J Power Sources 2018;400:242–55. https://doi.org/10.1016/j.jpowsour.2018.06.104.
[20] Xu T, Peng Z, Liu D, Wu L. A hybrid drive method for capacity prediction of lithium-ion batteries. IEEE Trans. Transp. Electrification 2022;8:1000–12. https://doi.org/10.1109/TTE.2021.3118813.
[21] Shu X, Chen Z, Shen J, Shen S, Guo F, Zhang Y, Liu Y. Ensemble learning and voltage reconstruction based state of health estimation for lithium-ion batteries with twenty random samplings. IEEE Trans Power Electron 2023;38:5538–48. https://doi.org/10.1109/TPEL.2023.3235872.
[22] Zhang Y, Tang Q, Zhang Y, Wang J, Stimming U, Lee AA. Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning. Nat Commun 2020;11:1706. https://doi.org/10.1038/s41467-020-15235-7.
[23] Niu G, Wang X, Liu E, Zhang B. Lebesgue sampling based deep belief network for lithium-ion battery diagnosis and prognosis. IEEE Trans Ind Electron 2022;69:8481–90. https://doi.org/10.1109/TIE.2021.3109527.
[24] Cui Z, Wang C, Gao X, Tian S. State of health estimation for lithium-ion battery based on the coupling-loop nonlinear autoregressive with exogenous inputs neural network. Electrochim Acta 2021;393:139047. https://doi.org/10.1016/j.electacta.2021.139047.
[25] Chen Z, Zhao H, Shu X, Zhang Y, Shen J, Liu Y. Synthetic state of charge estimation for lithium-ion batteries based on long short-term memory network modeling and adaptive H-Infinity filter. Energy 2021;228:120630. https://doi.org/10.1016/j.energy.2021.120630.
[26] Ye Z, Yu J. State-of-Health estimation for lithium-ion batteries using domain adversarial transfer learning. IEEE Trans Power Electron 2022;37:3528–43. https://doi.org/10.1109/TPEL.2021.3117788.
[27] Liu D, Wang S, Fan Y, Liang Y, Fernandez C, Stroe D-I. State of energy estimation for lithium-ion batteries using adaptive fuzzy control and forgetting factor

recursive least squares combined with AEKF considering temperature. J Energy Storage 2023;70:108040. https://doi.org/10.1016/j.est.2023.108040.
[28] Zahid T, Xu K, Li W, Li C, Li H. State of charge estimation for electric vehicle power battery using advanced machine learning algorithm under diversified drive cycles. Energy 2018;162:871–82. https://doi.org/10.1016/j.energy.2018.08.071.
[29] Hannan MA, Lipu MSH, Hussain A, Ker PJ, Mahlia TMI, Mansor M, Ayob A, Saad MH, Dong ZY. Toward enhanced state of charge estimation of lithium-ion batteries using optimized machine learning techniques. Sci Rep 2020;10:4687. https://doi.org/10.1038/s41598-020-61464-7.
[30] Fan Y, Xiao F, Li C, Yang G, Tang X. A novel deep learning framework for state of health estimation of lithium-ion battery. J Energy Storage 2020;32:101741. https://doi.org/10.1016/j.est.2020.101741.
[31] Iurilli P, Brivio C, Wood V. On the use of electrochemical impedance spectroscopy to characterize and model the aging phenomena of lithium-ion batteries: a critical review. J Power Sources 2021;505:229860. https://doi.org/10.1016/j.jpowsour.2021.229860.
[32] Meddings N, Heinrich M, Overney F, Lee J-S, Ruiz V, Napolitano E, Seitz S, Hinds G, Raccichini R, Gaberšček M, Park J. Application of electrochemical impedance spectroscopy to commercial Li-ion cells: a review. J Power Sources 2020;480:228742. https://doi.org/10.1016/j.jpowsour.2020.228742.
[33] Jones PK, Stimming U, Lee AA. Impedance-based forecasting of lithium-ion battery performance amid uneven usage. Nat Commun 2022;13:4806. https://doi.org/10.1038/s41467-022-32422-w.
[34] Zhang X, Zhang X, Sun X, An Y, Song S, Li C, Wang K, Su F, Chen C-M, Liu F, Wu Z-S, Ma Y. Electrochemical impedance spectroscopy study of lithium-ion capacitors: modeling and capacity fading mechanism. J Power Sources 2021;488:229454. https://doi.org/10.1016/j.jpowsour.2021.229454.
[35] Wang X, Wei X, Dai H. Estimation of state of health of lithium-ion batteries based on charge transfer resistance considering different temperature and state of charge. J Energy Storage 2019;21:618–31. https://doi.org/10.1016/j.est.2018.11.020.
[36] Buchicchio E, De Angelis A, Santoni F, Carbone P, Bianconi F, Smeraldi F. Battery SOC estimation from EIS data based on machine learning and equivalent circuit model. Energy 2023;283:128461. https://doi.org/10.1016/j.energy.2023.128461.
[37] Zhang Q, Huang C-G, Li H, Feng G, Peng W. Electrochemical impedance spectroscopy based state-of-health estimation for lithium-ion battery considering temperature and state-of-charge effect. IEEE Trans. Transp. Electrification 2022;8:4633–45. https://doi.org/10.1109/TTE.2022.3160021.
[38] Esser M, Rohde G, Rehtanz C. Electrochemical impedance spectroscopy setup based on standard measurement equipment. J Power Sources 2022;544:231869. https://doi.org/10.1016/j.jpowsour.2022.231869.
[39] Feng F, Yang R, Meng J, Xie Y, Zhang Z, Chai Y, Mou L. Electrochemical impedance characteristics at various conditions for commercial solid–liquid electrolyte lithium-ion batteries: Part 1. experiment investigation and regression analysis. Energy 2022;242:122880. https://doi.org/10.1016/j.energy.2021.122880.
[40] Zheng Y, Shi Z, Guo D, Dai H, Han X. A simplification of the time-domain equivalent circuit model for lithium-ion batteries based on low-frequency electrochemical impedance spectra. J Power Sources 2021;489:229505. https://doi.org/10.1016/j.jpowsour.2021.229505.
[41] Wang Y, Li M, Chen Z. Experimental study of fractional-order models for lithium-ion battery and ultra-capacitor: modeling, system identification, and validation. Appl Energy 2020;278:115736. https://doi.org/10.1016/j.apenergy.2020.115736.
[42] Lu Y, Zhao C-Z, Huang J-Q, Zhang Q. The timescale identification decoupling complicated kinetic processes in lithium batteries. Joule 2022;6:1172–98. https://doi.org/10.1016/j.joule.2022.05.005.
[43] Yang B, Wang D, Sun X, Chen S, Wang X. Offline order recognition for state estimation of Lithium-ion battery using fractional order model. Appl Energy 2023;341:120977. https://doi.org/10.1016/j.apenergy.2023.120977.
[44] He R, He Y, Xie W, Guo B, Yang S. Comparative analysis for commercial li-ion batteries degradation using the distribution of relaxation time method based on electrochemical impedance spectroscopy. Energy 2023;263:125972. https://doi.org/10.1016/j.energy.2022.125972.
[45] Zhu J, Knapp M, Liu X, Yan P, Dai H, Wei X, Ehrenberg H. Low-temperature separating lithium-ion battery interfacial polarization based on distribution of relaxation times (DRT) of impedance. IEEE Trans. Transp. Electrification 2021;7:410–21. https://doi.org/10.1109/TTE.2020.3028475.
[46] Soni R, Robinson JB, Shearing PR, Brett DJL, Rettie AJE, Miller TS. Lithium-sulfur battery diagnostics through distribution of relaxation times analysis. Energy Storage Mater 2022;51:97–107. https://doi.org/10.1016/j.ensm.2022.06.016.
[47] Vidaković-Koch T, Miličić T, Živković LA, Chan HS, Krewer U, Petkovska M. Nonlinear frequency response analysis: a recent review and perspectives. Curr Opin Electrochem 2021;30:100851. https://doi.org/10.1016/j.coelec.2021.100851.
[48] Harting N, Wolff N, Krewer U. Identification of lithium plating in lithium-ion batteries using nonlinear frequency response analysis (NFRA). Electrochim Acta 2018;281:378–85. https://doi.org/10.1016/j.electacta.2018.05.139.
[49] Bayma RS, Zhu Y, Lang Z-Q. The analysis of nonlinear systems in the frequency domain using Nonlinear Output Frequency Response Functions. Automatica 2018;94:452–7. https://doi.org/10.1016/j.automatica.2018.04.030.
[50] Macdonald DD, Urquidi-Macdonald M. Application of kramers-kronig transforms in the analysis of electrochemical systems: I . Polarization resistance. J Electrochem Soc 1985;132:2316–9. https://doi.org/10.1149/1.2113570.
[51] Manikandan B, Ramar V, Yap C, Balaya P. Investigation of physico-chemical processes in lithium-ion batteries by deconvolution of electrochemical impedance spectra. J Power Sources 2017;361:300–9. https://doi.org/10.1016/j.jpowsour.2017.07.006.
[52] Schönleber M, Klotz D, Ivers-Tiffée E. A method for improving the robustness of linear kramers-kronig validity tests. Electrochim Acta 2014;131:20–7. https://doi.org/10.1016/j.electacta.2014.01.034.
[53] Thele M, Bohlen O, Sauer DU, Karden E. Development of a voltage-behavior model for NiMH batteries using an impedance-based modeling concept. J Power Sources 2008;175:635–43. https://doi.org/10.1016/j.jpowsour.2007.08.039.
[54] Chang C, Wang S, Tao C, Jiang J, Jiang Y, Wang L. An improvement of equivalent circuit model for state of health estimation of lithium-ion batteries based on mid-frequency and low-frequency electrochemical impedance spectroscopy. Measurement 2022;202:111795. https://doi.org/10.1016/j.measurement.2022.111795.
[55] Wang X, Wei X, Zhu J, Dai H, Zheng Y, Xu X, Chen Q. A review of modeling, acquisition, and application of lithium-ion battery impedance for onboard battery management. eTransportation 2021;7:100093. https://doi.org/10.1016/j.etran.2020.100093.
[56] Broussely M, Herreyre S, Biensan P, Kasztejna P, Nechev K, Staniewicz RJ. Aging mechanism in Li ion cells and calendar life predictions. J Power Sources 2001;97–98:13–21. https://doi.org/10.1016/S0378-7753(01)00722-4.
[57] Guha A, Patra A, Vaisakh KV. Remaining useful life estimation of lithium-ion batteries based on the internal resistance growth model. In: 2017 Indian control conf. ICC; 2017. p. 33–8. https://doi.org/10.1109/INDIANCC.2017.7846448.
[58] Saha B, Goebel K, Poll S, Christophersen J. Prognostics methods for battery health monitoring using a bayesian framework. IEEE Trans Instrum Meas 2009;58:291–6. https://doi.org/10.1109/TIM.2008.2005965.
[59] Xiong R, Zhang Y, He H, Zhou X, Pecht MG. A double-scale, particle-filtering, energy state prediction algorithm for lithium-ion batteries. IEEE Trans Ind Electron 2018;65:1526–38. https://doi.org/10.1109/TIE.2017.2733475.
[60] Lotfi F, Ziapour S, Faraji F, Taghirad HD. A switched SDRE filter for state of charge estimation of lithium-ion batteries. Int J Electr Power Energy Syst 2020;117:105666. https://doi.org/10.1016/j.ijepes.2019.105666.
[61] Xuan D-J, Shi Z, Chen J, Zhang C, Wang Y-X. Real-time estimation of state-of-charge in lithium-ion batteries using improved central difference transform method. J Clean Prod 2020;252:119787. https://doi.org/10.1016/j.jclepro.2019.119787.
[62] Haus B, Mercorelli P. Polynomial augmented extended kalman filter to estimate the state of charge of lithium-ion batteries. IEEE Trans Veh Technol 2020;69:1452–63. https://doi.org/10.1109/TVT.2019.2959720.
[63] Chang C, Wang S, Jiang J, Gao Y, Jiang Y, Liao L. Lithium-ion battery state of health estimation based on electrochemical impedance spectroscopy and cuckoo search algorithm optimized elman neural network. J. Electrochem. Energy Convers. Storage 2022;19:030912. https://doi.org/10.1115/1.4054128.
[64] Vivier V, Orazem ME. Impedance analysis of electrochemical systems. Chem. Rev. 2022;122:11131–68. https://doi.org/10.1021/acs.chemrev.1c00876.