

A lithium-ion battery remaining useful life prediction method based on the incremental capacity analysis and Gaussian process regression

Xiaoqiong Pang^{a,*}, Xiaoyan Liu^a, Jianfang Jia^b, Jie Wen^b, Yuanhao Shi^b, Jianchao Zeng^a, Zhen Zhao^a

^a School of Data Science and Technology, North University of China, No.3, XueYuan Road, JianCaoPing District, Taiyuan 030051, Shanxi, China

^b School of Electrical and Control Engineering, North University of China, No.3, XueYuan Road, JianCaoPing District, Taiyuan 030051, Shanxi, China

ARTICLE INFO

Keywords:

Lithium-ion batteries
Remaining useful life
Incremental capacity curve
Gaussian process regression

ABSTRACT

Remaining useful life (RUL) is a critical metric of lithium-ion battery prognostic and health management. Accurate prediction of RUL is of great significance to the safety and reliability of lithium-ion batteries, which is able to provide useful reference information for maintenance. In this work, a novel method fusing the incremental capacity analysis (ICA) and Gaussian process regression (GPR) for RUL prediction of lithium-ion batteries is proposed. Firstly, the IC curve, which has higher sensitivity than the traditional charge/discharge curve, is used to analyze the performance degradation process of the lithium-ion battery. Then the peak value of the IC curve and the regional area under the peak value of the IC curve are extracted as health indicators (HIs). Secondly, the RUL prediction framework of lithium-ion batteries based on ICA and GPR is established, and the uncertainty expression of the prediction results is given. Finally, the experimental results show that the HIs extracted in this paper can effectively reflect the degradation state of the battery, and the proposed method has high accuracy in predicting RUL.

1. Introduction

As a portable energy source, lithium-ion batteries are rapidly popularized with its advantages of lightweight, high energy density, high chemical reactivity and long life. It plays an important role in the fields of consumer electronics (mobile phones and laptops), hybrid and electric vehicles in the automotive industry and space exploration [1,2]. However, with the increase of charging and discharging times, lithium-ion batteries will inevitably degrade and even fail, if effective measures cannot be taken before the battery failure, the lithium-ion battery equipment will not be able to operate healthily, which may cause casualties in severe cases [3]. Battery management system (BMS) is widely used to solve the above safety problems, which uses the remaining useful life (RUL) prediction information of lithium-ion batteries [4,5]. Therefore, the prediction of RUL by calculating the time interval from starting time to threshold point has attracted extensive attention.

At present, many types of research on RUL prediction of lithium-ion batteries use capacity [6,7,8,9] or internal resistance [10,11] to reflect the degradation state of the battery. However, in practical application, the cost of monitoring or measuring the internal resistance of the

lithium-ion batteries is high and difficult. Moreover, it is difficult to achieve full charge or discharge conditions, and the maximum capacity of battery charge/discharge is generally difficult to test. Therefore, it is of great practical value to extract health indicators (HIs) that are easy to measure or calculate to evaluate the health state of the battery.

Besides, many studies extracted the features that can reflect the health status of lithium-ion batteries by analyzing the charge/discharge curves, such as the charge time interval in fixed voltage range [12], battery terminal voltage [13], voltage region during the charging process [14], the length of constant voltage stage of the current curve during the charging process [15], the area under the constant voltage curve in the charging processes [15] and so on. However, the above work did not make the qualitative and quantitative evaluations and analyses on the extracted HIs. Moreover, the traditional charge/discharge curves cannot sensitively identify the drastic chemical changes inside the battery, which may lead to the loss of some key information of battery behavior [16].

Incremental capacity analysis (ICA) allows the detection of gradual changes in battery behavior under the condition of incomplete charge or discharge [17,18,19,20,21,22]. Compared with the traditional method

* Corresponding author.

E-mail addresses: xqpang@nuc.edu.cn (X. Pang), jiajianfang@nuc.edu.cn (J. Jia), yhshi@nuc.edu.cn (Y. Shi).

<https://doi.org/10.1016/j.microrel.2021.114405>

Received 5 April 2021; Received in revised form 22 September 2021; Accepted 3 October 2021

Available online 11 October 2021

0026-2714/© 2021 Elsevier Ltd. All rights reserved.

based on charge/discharge curves, ICA has higher sensitivity and can obtain the key information of cell behavior related to its electrochemical performance [16,23,24,25,26]. Although ICA is a useful method to analyze battery degradation, most of the ICA on lithium-ion batteries mainly focused on the analysis of electrochemical aging mechanism, while there are only a handful of studies [27,28] on RUL prediction. Zhang et al. [27] selected ten fixed voltage positions on the local IC curve, and finally obtained five values as the health index series by calculating the difference; then they used the artificial neural network for RUL prediction of lithium-ion batteries. Li et al. [28] utilized the voltage positions of one valley and two peaks in the IC curve as features to estimate the battery health state. However, the above two researches all use voltage values to identify feature points. Thus consistent and accurate voltage values are required. When the external contact resistance changes, such as the loose contact point, the voltage value may be unreliable. Therefore, one motivation of this paper is to apply ICA to extract more sensitive characteristics of battery degradation than traditional charge/discharge curves. At the same time, to avoid the problem in the above two works [27,28], we focus on more reliable HIs and abandon the information of voltage shift.

In recent years, the existing RUL prediction methods for lithium-ion batteries can be roughly divided into two types: model-based [29,30,31,32,33,34,35] and data-driven [36,37,38,39,40]. The difficulty of the model-based prediction method is that the mechanism of lithium-ion batteries is complex, and it is difficult to build an accurate model. The data-driven prediction method relies on the existing historical monitoring data, which avoids the study of the complex internal mechanism and electrochemical reaction process of lithium-ion batteries. It has become a hotspot in the research of battery life. The data-driven method of lithium-ion battery health prediction usually uses machine learning models, such as support vector machines, logistic process regression, neural networks, to establish the relationship between monitoring data and health status, track the degradation status of the battery and predict its RUL. Nevertheless, there is a common disadvantage of these methods, which is the lack of uncertainty expression and management of prediction. In industrial applications, the uncertainty in models and data may lead to poor prediction reliability. Therefore, it is essential to consider the uncertainty of prediction results when studying the RUL prediction of lithium-ion batteries.

At present, the methods with uncertainty expression ability mainly include particle filter (PF), relevance vector machine (RVM) and Gaussian process regression (GPR). PF is a model-based method, which needs physical or electrochemical knowledge to model the degradation trend of the battery. RVM method has high sparsity, which may lead to less stability of prediction results. GPR is a probabilistic approach for modeling and forecasting based on statistical learning theory and Bayesian theory. It can model and predict the dynamic behavior characteristics of any linear or nonlinear system and explain the uncertainty of prediction results in the form of probability [41,42,43,44]. Considering that the degradation process of lithium-ion batteries is a complex, dynamic and nonlinear electrochemical process, GPR method is suitable for the establishment of the RUL prediction model of lithium-ion batteries.

Based on the above analysis, this paper develops a RUL prediction method of lithium-ion batteries based on ICA and GPR. In this work, the IC curve, which has greater sensitivity than the charge/discharge curve and does not need full charging or discharging data, is used to extract the features reflecting battery degradation state, and the extracted features are easy to calculate. Specifically, the peak value of the IC curve corresponding to voltage plateaus and the regional area under the peak value of the IC curve representing regional discharge capacity are selected as HIs. In addition, the qualitative and quantitative analysis of the extracted HIs is carried out, and the analysis results show that the extracted HIs can well reflect the degradation state of the battery. In the RUL prediction stage, a combined model based on ICA and GPR is proposed to predict the RUL of the lithium-ion batteries, and the

uncertainty expression of the prediction result is given.

The remainder of this paper is structured as follows. Section 2 introduces the ICA method and extracts the HIs. Section 3 describes the proposed ICA-GPR method for battery RUL prediction. Section 4 presents the RUL prediction results and discussions. Finally, the conclusion and future work are shown in Section 5.

2. Incremental capacity curve and health indicator extraction

In this section, we mainly introduce the ICA method and HIs extraction. Firstly, the battery aging experimental dataset from the NASA database is analyzed. Then, the initial IC curves are obtained by calculation and the filter method is used to smooth these curves. Finally, two HIs, namely the peak value and the regional area under the peak value of IC curves, are extracted, and the correlation between the proposed HIs and battery capacity degradation is evaluated qualitatively and quantitatively.

2.1. Experiment data

The test data in this work is from the battery data set provided by the NASA PCoE research center [45]. It contains a set of degradation state data of four 18,650 lithium-ion batteries (#5, #6, #7 and #18) collected in the same experimental environment. The detailed experimental conditions and parameters are given in Table 1. Fig. 1 shows the true capacity degradation curves of the four batteries. It should be noted that the four capacity curves do not decrease monotonously with the number of cycles due to the existence of some regeneration capacities.

2.2. Incremental capacity curve analysis

ICA is a significant method to analyze the degradation mechanism of batteries. ICA involves the derivative of capacity Q to voltage V , i.e., $\Delta Q/\Delta V$, and the calculation formula is shown in Eq. (1). In this work, we use the data during the constant-current discharge stage to calculate IC curves.

$$\frac{dQ}{dV} = \frac{I \times dt}{dV} = I \times \frac{dt}{dV} \quad (1)$$

where I , V and t are current, voltage and time in discharge phase, respectively.

It is a significant step to estimate the health state of the battery by using the characteristics of ICA. Due to the influence of noise or measurement error, the direct numerical derivative usually cannot give a smooth and accurate result. In order to overcome these problems, Gaussian filter (GF) is used to smooth the IC curves.

As an effective filtering tool, GF can capture the essence of the signal by separating low-frequency signals from high-frequency noise [27,28]. GF has the characteristic of Gaussian distribution, which is expressed as:

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (2)$$

where μ and σ are the mean value and standard deviation, respectively. When GF is applied, each data point can be replaced by a weighted

Table 1
Detailed experimental parameters of four batteries.

Battery label	Charging cut-off voltage (V)	Discharging cut-off voltage (V)	Discharging current (A)	Temperature (°C)
#5	4.2	2.7	2	24
#6	4.2	2.5	2	24
#7	4.2	2.2	2	24
#18	4.2	2.5	2	24

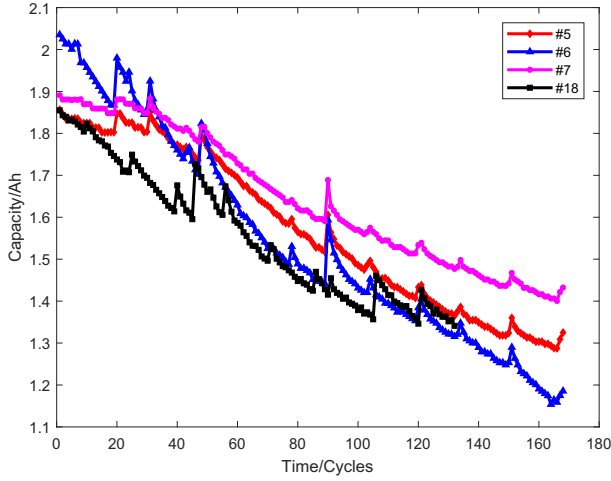


Fig. 1. Capacity degradation curves of four batteries (#5, #6, #7, and #18).

average of its neighbors. Hence, the nearest data have more influence on the average and the distant data play a smaller role. When GF is used for smoothing, μ is normally set to 0, because we want each data to be the one that has the biggest effect on its new smoothed value. σ will be used as the parameter to control the smoothness of the final curve, more specifically, how big the smoothing window we use for averaging. The bigger σ is, the smoother effect we can achieve, however, the curves may lose some important information if σ is too big. On the contrary, too small a value may lead to an undesirable smoothing effect. Thus, the value of σ should be kept in an appropriate range [27,28,46]. In our work, through comparative experiments, σ is chosen as 3, as it can provide good smoothing without losing or deforming the important features of interest on IC curves.

Fig. 2 shows the comparison of initial IC curve and GF smoothed IC curve of battery #5 in the first cycle. It can be seen that the blue curve is the initial IC curve, the voltage plateau region of the voltage capacity curve corresponds to the peak of the IC curve, and this region is very sensitive to trembling and measurement noise. Obviously, it is difficult to identify important information such as peak value and peak position from the initial IC curve. The purple curve is the curve processed by GF, it can be seen that GF can filter the noise well and the characteristics of the IC curve are well preserved.

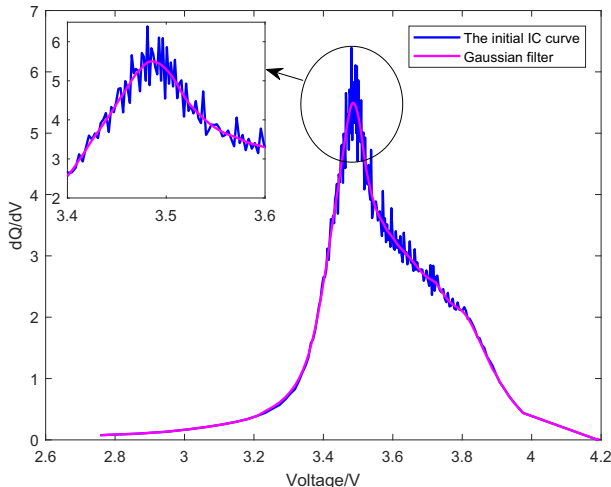


Fig. 2. Comparison of initial IC curve and GF smoothed IC curve of battery #5.

2.3. Extraction of health indicator

In this section, we analyze the IC curves of the discharge process under different cycles and extract the features which can reflect the battery health state. Taking the IC curves of battery #5 in seven different cycles as an example, as shown in Fig. 3. It is obvious that the IC peaks of different cycles have unique shapes, amplitudes and positions. With the increase of cycle times, the peak values of IC curves show an overall attenuation trend, and the areas under the IC peaks also show a gradually decreasing trend. Therefore, we consider selecting the peak value and the area under the peak value as the health factors reflecting the degradation state of the battery.

Obviously, it is easy to extract peak value from the IC curve smoothed by GF, but the determination of regional area under the peak value needs further analysis. Fig. 3 shows the complete IC curves that cover almost fully discharging process, in which the voltage decreases from 4.2 V to 2.7 V. It can be seen that the IC curves do not change significantly in the voltage range of 4.2 V–4 V and 3 V–2.7 V. While the IC curves change significantly in the voltage range of 4 V–3 V, which covers the primary discharge capacity in different cycles. Further observation and analysis of the slope change of IC curves, we can see that when the slope is greater than -2.4288 (the minimum slope of all curves at 4 V) and less than 1.7609 (the maximum slope of all curves at 3 V), that is, during discharge process when voltage drops from the initial value to 4 V and the voltage is lower than 3 V, the change of IC curves is not obvious. Correspondingly, the change of IC curves is obvious when voltage drops from 4 V to 3 V. Hence, this region is considered a significant area to identify the battery degradation process and the area between 4 V and 3 V below peak value is taken as the health index.

2.4. Health indicators assessment

In order to verify the expression ability of the extracted peak value (HI1) and regional area under the peak (HI2) on the degradation state of the battery, the correlation between the proposed HIs and capacity is evaluated through qualitative and quantitative analysis in this section.

2.4.1. Qualitative assessment

The qualitative assessment adopts a scatter diagram to analyze the consistency of change trend between both features and the capacity. Fig. 4 shows the normalized features and capacity of four batteries under all discharge cycles. It can be seen that except for the early cycles, the overall trend of both features and capacity of the four batteries is

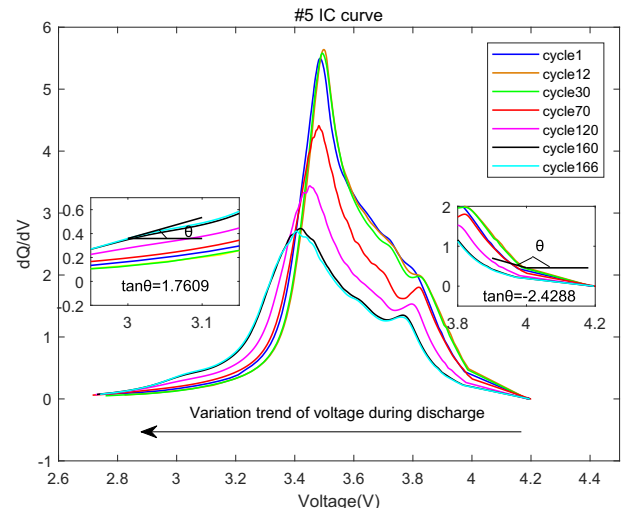


Fig. 3. IC curves of battery #5 under different cycles.

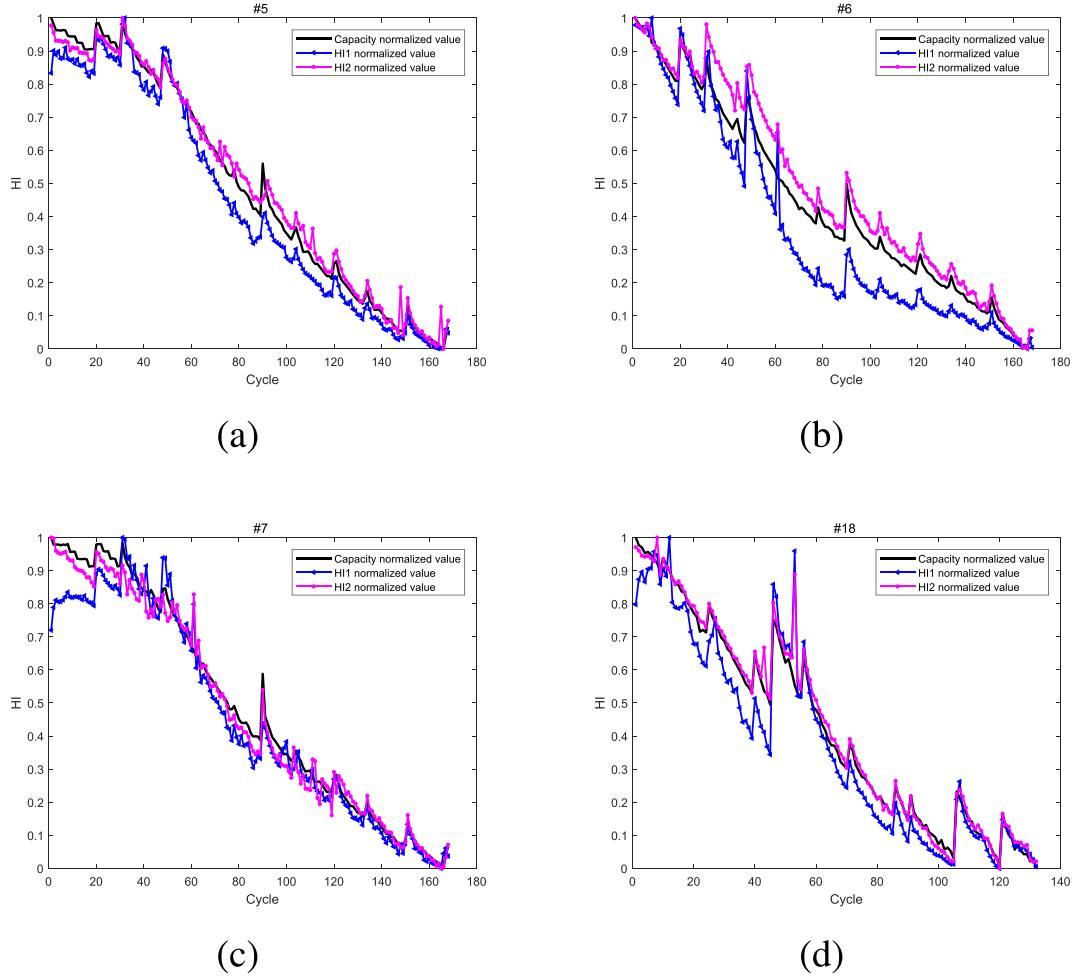


Fig. 4. Degradation curves of normalized capacity and HIs: (a) battery #5; (b) battery #6; (c) battery #7; (d) battery #18.

basically the same, and the short-term variation characteristics of the two features also well reflect the fluctuation of capacity, so these two features can be used as HIs to characterize battery degradation. In addition, the variation range of HI1 in partial cycles is large, which is due to the inconsistent loss proportion of active material and lithium-ion in different cycles, resulting in the difference of decline proportion of each peak.

2.4.2. Quantitative assessment

Pearson correlation coefficient and Spearman correlation coefficient are often used to evaluate the degree of correlation between variables in statistics. Normally, Pearson correlation analysis is more suitable for continuous variables with normal distribution. Spearman correlation coefficient is widely used without the above constraints. Through the analysis, it is found that the NASA battery data used in this work does not conform to the normal distribution. Hence, Spearman correlation analysis is selected to measure the correlation between HIs and capacity. Furthermore, to make the correlation coefficient between the sample variables statistically significant, the Spearman correlation coefficient is further tested for significance in our work. The calculation formula is expressed as:

$$\rho = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)} \quad (3)$$

$$d_i = x_i - y_i \quad (4)$$

where ρ is the Spearman correlation coefficient between x and y , and the

value of ρ is between -1 and 1 . The greater is the absolute value of ρ , the higher is the correlation degree between x and y . N is sample size, x_i, y_i are sample points indexed with i . It should be noted that the x and y sets should be ranked before calculating ρ . The correlation analysis between HIs and capacity is given in Table 2. It can be seen that the Spearman correlation coefficients between HIs and capacity are greater than 0.96 , which indicates that there is a strong correlation between battery capacity and HIs.

Furthermore, Spearman correlation coefficients are tested for significance. The four steps of significance test by p -value method are as follows:

- (1) Make assumptions on the population, such as:
 - Original hypothesis $H_0 : \rho = 0$ is that the two variables are not related;
 - Alternative hypothesis $H_1 : \rho \neq 0$ means that the two variables are related;
- (2) Find the statistics and calculate the minimum significance level of H_0 which can be rejected, namely p -value;

Table 2
Correlation analysis between four batteries' HIs and capacity.

Spearman correlation	#5	#6	#7	#18
HI1	0.9902	0.9956	0.9691	0.9801
(p-Value)	7.81E-144	9.92E-173	7.27E-103	3.41E-93
HI2	0.9955	0.9965	0.9901	0.9931
(p-Value)	6.80E-172	2.09E-180	1.71E-143	5.07E-123

- (3) Given the significance level α , the rejection region is determined;
 (4) If $p \leq \alpha$, H_0 is rejected at significance level α ; if $p > \alpha$, H_0 is accepted at significance level α .

If the significance test results reject the original hypothesis, it indicates that there is a significant correlation between the two variables. At this time, if the absolute value of the correlation coefficient between the two variables in the sample is closer to 1, it indicates that the correlation between the two variables is stronger. When the significance level α is set to 0.01, the p-value of the significance test is shown in Table 2. It can be seen that the p-value of HIs are far less than 0.01, and the result of the significance test is to reject the original hypothesis, indicating that the extracted HIs are significantly related to battery capacity.

To sum up, the results of qualitative and quantitative evaluation analysis show that the two HIs extracted from the IC curve have a strong correlation with battery capacity and can well characterize the deterioration state of the battery.

3. Method

The GPR method has been widely applied in the field of statistics and machine learning due to the advantages of convenient properties for building model without specific functional form. The GPR uses the probabilistic method to describe the regression function. In this section, a RUL prediction model for lithium-ion batteries is proposed based on the combination of ICA and GPR.

3.1. Gaussian process regression

GPR model is a nonparametric model. It obtains the prior distribution by setting the parameters of the model, and then deduces the posterior distribution by combining with the input of training data, so as to obtain the mean value and variance of the prediction, and expresses the uncertainty of prediction by 95% (or other values) confidence interval. The Gauss process (GP) can be written as follows:

$$f(\mathbf{x}) \sim GP(m(\mathbf{x}), k_f(\mathbf{x}, \mathbf{x}')) \quad (5)$$

The above function $f(\mathbf{x})$ constitutes a GP with the mean function $m(\mathbf{x})$ and the covariance function $k_f(\mathbf{x}, \mathbf{x}')$. Among them, $m(\mathbf{x}) = E[f(\mathbf{x})]$, $k_f(\mathbf{x}, \mathbf{x}') = E[(f(\mathbf{x}) - m(\mathbf{x})) \cdot (f(\mathbf{x}') - m(\mathbf{x}'))]$. The mean function is taken as 0; the square exponential covariance function is selected as the covariance function in this work, denoted by

$$m(\mathbf{x}) = 0 \quad (6)$$

$$k_f(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{(\mathbf{x} - \mathbf{x}')^2}{2l^2}\right) \quad (7)$$

where σ_f^2 is the signal variance, l is the variance scale.

In many scenes, the observed data can be expressed as an implicit function:

$$\mathbf{y} = f(\mathbf{x}) + \varepsilon \quad (8)$$

where \mathbf{y} is the observation vector $[y_1, y_2, \dots, y_n]$, ε is the Gaussian noise and $\varepsilon \sim N(0, \sigma_n^2)$. Hence, the prior distribution can be denoted as

$$\mathbf{y} \sim N(0, \mathbf{K}_f(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}_n) \quad (9)$$

where \mathbf{I}_n is a n-dimensional unit matrix and $\sigma_n^2 \mathbf{I}_n$ refers the noise covariance matrix. $\mathbf{K}_f(\mathbf{x}, \mathbf{x})$ is a n-dimensional symmetric positive definite matrix, as shown in Eq. (10).

$$\begin{cases} \mathbf{K}_f(\mathbf{x}, \mathbf{x}) = (k_{ij})_{n \times n} \\ k_{ij} = \sigma_f^2 \exp\left(-\frac{(x_i - x_j)^2}{2l^2}\right) \end{cases} \quad (10)$$

The hyper-parameter set $\theta = [\sigma_f, l, \sigma_n]$ can be determined by maximum likelihood method. The aim of maximum likelihood method is to maximize the log-likelihood function, as shown in Eq. (11).

$$\begin{aligned} L = \log p(\mathbf{y}|\mathbf{x}, \theta) &= -\frac{1}{2} \log(\det(\mathbf{K}_f(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}_n)) \\ &\quad - \frac{1}{2} \mathbf{y}^T [\mathbf{K}_f(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}_n]^{-1} \mathbf{y} - \frac{n}{2} \log 2\pi \end{aligned} \quad (11)$$

Conjugate gradient method is widely applied for finding parametric optimal solution. The basic idea is solving the maximum value of objective function by taking the derivative of the log-likelihood function. The partial derivative of Eq. (11) is shown as Eq. (12):

$$\begin{cases} \frac{\partial}{\partial \theta_i} \log p(\mathbf{y}|\mathbf{x}, \theta) = \\ \frac{1}{2} \text{tr} \left[\alpha \alpha^T - (\mathbf{K}_f(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}_n)^{-1} \frac{\partial (\mathbf{K}_f(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}_n)}{\partial \theta_i} \right] \\ \alpha = [\mathbf{K}_f(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}_n]^{-1} \mathbf{y} \end{cases} \quad (12)$$

where θ_i is an element in the hyper-parameter set θ .

The joint prior distribution of the observation value \mathbf{y} and the prediction value \mathbf{y}^* at a test point \mathbf{x}^* is denoted as

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}^* \end{bmatrix} \sim N\left(0, \begin{bmatrix} \mathbf{K}_f(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}_n & \mathbf{K}_f(\mathbf{x}, \mathbf{x}^*) \\ \mathbf{K}_f(\mathbf{x}, \mathbf{x}^*)^T & \mathbf{K}_f(\mathbf{x}^*, \mathbf{x}^*) \end{bmatrix}\right) \quad (13)$$

where $\mathbf{K}_f(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}_n$ is the covariance matrix formed by training data, $\mathbf{K}_f(\mathbf{x}, \mathbf{x}^*)$ is the covariance matrix formed by training data and test data, and $\mathbf{K}_f(\mathbf{x}^*, \mathbf{x}^*)$ is the covariance matrix formed by test data.

The posterior distribution of $p(\mathbf{y}^*|\mathbf{x}, \mathbf{y}, \mathbf{x}^*)$ is expressed as

$$p(\mathbf{y}^*|\mathbf{x}, \mathbf{y}, \mathbf{x}^*) = N(\mathbf{y}^*|\hat{\mathbf{y}}^*, \sigma^2(\mathbf{y}^*)) \quad (14)$$

where the prediction mean $\hat{\mathbf{y}}^*$ and the prediction covariance $\sigma^2(\mathbf{y}^*)$ are shown as follows

$$\hat{\mathbf{y}}^* = \mathbf{K}_f(\mathbf{x}, \mathbf{x}^*)^T [\mathbf{K}_f(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}_n]^{-1} \mathbf{y} \quad (15)$$

$$\begin{aligned} \sigma^2(\mathbf{y}^*) &= \mathbf{K}_f(\mathbf{x}^*, \mathbf{x}^*) - \mathbf{K}_f(\mathbf{x}, \mathbf{x}^*)^T \\ &\quad [\mathbf{K}_f(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}_n]^{-1} \mathbf{K}_f(\mathbf{x}, \mathbf{x}^*) \end{aligned} \quad (16)$$

The 95% confidence interval (CI) was used to evaluate the uncertainty of prediction results:

$$95\% \text{CI} = \hat{\mathbf{y}}^* \pm 1.96 \times \sigma^2(\mathbf{y}^*) \quad (17)$$

3.2. Modeling process based on the ICA and GPR

In this section, a novel ICA-GPR method for lithium-ion battery RUL prediction is proposed based on a combination of ICA and GPR. The main processes of the proposed method are shown in Fig. 5, which can be divided into the following four steps:

1. IC curves analysis. According to the measurement data of NASA battery database, such as current, time and voltage, the IC curve is calculated, and then the GF is used to smooth the IC curve.
2. Extraction of health indicator. The peak value and the regional area under the peak value are extracted from the IC curve obtained in the first step, and then the correlation analysis is carried out.

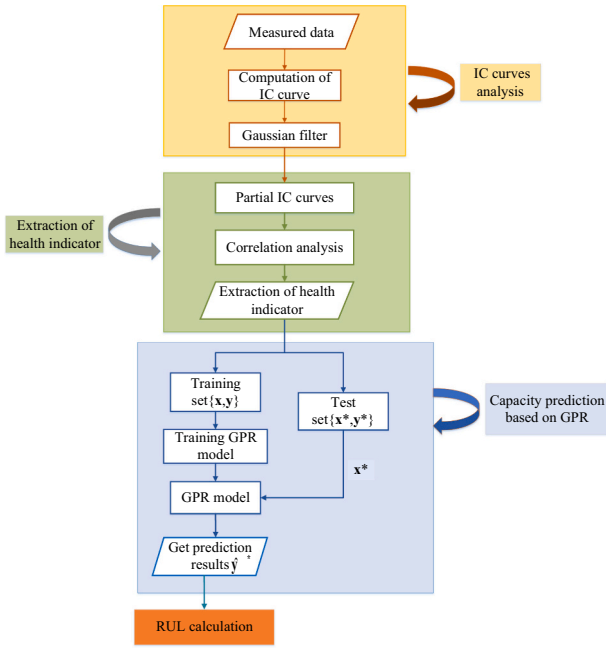


Fig. 5. Flowchart of RUL prediction for lithium-ion batteries based on ICA and GPR.

- Capacity prediction based on GPR. The two HIs $\{HI1(i), HI2(i)\}_{i=1}^n$ extracted from the IC curves are used as the input of the model and the battery capacity $\{C(i)\}_{i=1}^n$ as the output to build the capacity prediction model based on GPR. The training set is denoted as $\{x, y\}$, where $x = [HI1(1:m)^T, HI2(1:m)^T]^T$, $y = [C(1:m)^T]^T$, the cycle number corresponding to the prediction starting point is denoted as $m + 1$. The test set is denoted as $\{x^*, y^*\}$, where $x^* = [HI1(m + 1:n)^T, HI2(m + 1:n)^T]^T$, $y^* = [C(m + 1:n)^T]^T$. The prediction part takes x^* as the input to get the capacity prediction result \hat{y}^* .
- RUL calculation. The remaining useful life (N_{RUL}) is the difference between the total number of charge-discharge cycles when the actual capacity of the battery drops to the threshold value (N_{EOL}) and the number of charge-discharge cycles of the current battery (N_{ECL}). The calculation formula is as follows:

$$N_{RUL} = N_{EOL} - N_{ECL} \quad (18)$$

3.3. Performance evaluation

We use root mean square error (RMSE), R^2 and absolute error (AE) to measure the stability and accuracy of the proposed method.

$$RMSE = \sqrt{\frac{\sum_{i=m+1}^n (y_i - \hat{y}_i)^2}{n - m}} \quad (19)$$

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=m+1}^n (y_i - \hat{y}_i)^2}{\sum_{i=m+1}^n (y_i - \bar{y})^2} \quad (20)$$

$$AE = |RUL_{real} - RUL_{pre}| \quad (21)$$

where y_i represents the actual value of battery capacity, \hat{y}_i represents the predicted value of battery capacity, and \bar{y}_i represents the mean value of predicted battery capacity. RUL_{real} represents the real RUL, and RUL_{pre} represents the predicted RUL.

4. Results and discussion

In this section, the battery RUL prediction results are verified and analyzed. Firstly, in order to analyze the RUL prediction performance of the ICA-GPR framework proposed in this paper, we designed an experiment with different prediction starting points. Secondly, through the comparative experiments of different HIs, the effectiveness of the HIs extracted in this paper is verified. Finally, the accuracy of the proposed method is further verified by comparing it with the results of different methods in some existing literatures.

4.1. RUL prediction of lithium-ion batteries

In order to analyze the RUL prediction performance of the proposed model, we designed experiments with different prediction starting points. Experimental design: we select the data of the first 59 cycles, the first 79 cycles and the first 99 cycles as training sets, that is, the starting point of prediction is 60, 80 and 100, respectively. Besides, the failure threshold of the battery is 70%–80% of the rated capacity. In this paper, the capacity failure threshold for batteries #5, #6 and #18 are 1.38 Ah. In addition, because battery #7 does not reach the failure threshold, the failure threshold line is not given.

Fig. 6 shows the prediction results and their uncertainty expressions of four batteries at the prediction starting point 60, 80 and 100, respectively. It can be distinctly seen that the prediction curves are close to the real capacity degradation curves. Furthermore, the prediction curves of four batteries at prediction starting points 80 and 100 are closer to the actual capacity degradation curves than those at starting point 60.

Table 3 shows the prediction performance evaluation results of four batteries at the prediction starting point 60, 80 and 100, respectively. It can be seen that the RMSE of four batteries are less than 0.04, which indicates that the prediction performance of the proposed model is good; R^2 is greater than 0.9, which illustrates that the fitting degree between the prediction curve and the actual curve is high; the AE values of three batteries are less than ten at different prediction starting points, and the result of AE values shows that for each battery, the AE of RUL prediction result decreases with the prediction starting point moving backward.

4.2. Comparative experimental analysis of different HI

Considering that Ref. [27] also analyzes IC curve to extract HI, which can characterize the performance degradation of the same lithium-ion battery data to predict RUL. In order to verify the effectiveness of the proposed HIs in this work more comprehensively, we design a model M2 which is compared with the proposed model M1 in this paper, as shown in Table 4. Among them, M2 uses the HI extracted from Ref. [27] according to fixed voltage position, specifically, in the voltage range of 3.4 V–3.6 V, select ten values with voltage interval of 0.02 V (except 3.50 V) on the IC curve, then divide these ten values into 5 groups, and calculate the difference to get five HIs, finally select two HIs with high correlation for prediction. It should be noted that except for different health factors, the other steps of M2 are the same as model M1. Experimental design: we select the data of the first 79 cycles as the training set, and the predicted starting point is 80.

Table 5 gives the prediction performance of models M1 and M2 for four batteries (#5, #6, #7, and #18). It can be found that the RMSE of model M1 for each battery is less than that of model M2, and the AE of model M1 is evidently better than that of model M2, which indicates that model M1 has higher prediction accuracy. In addition, the R^2 of model M1 for each battery is greater than 0.94, while in the comparison model M2, the R^2 of other batteries are less than 0.9 except for battery #7, which indicates that model M1 has the higher fitting degree to the original curve. The above comparison results show that although the HIs of the two models are extracted based on the IC curve, the prediction effect of HIs extracted by model M1 is evidently better than that of

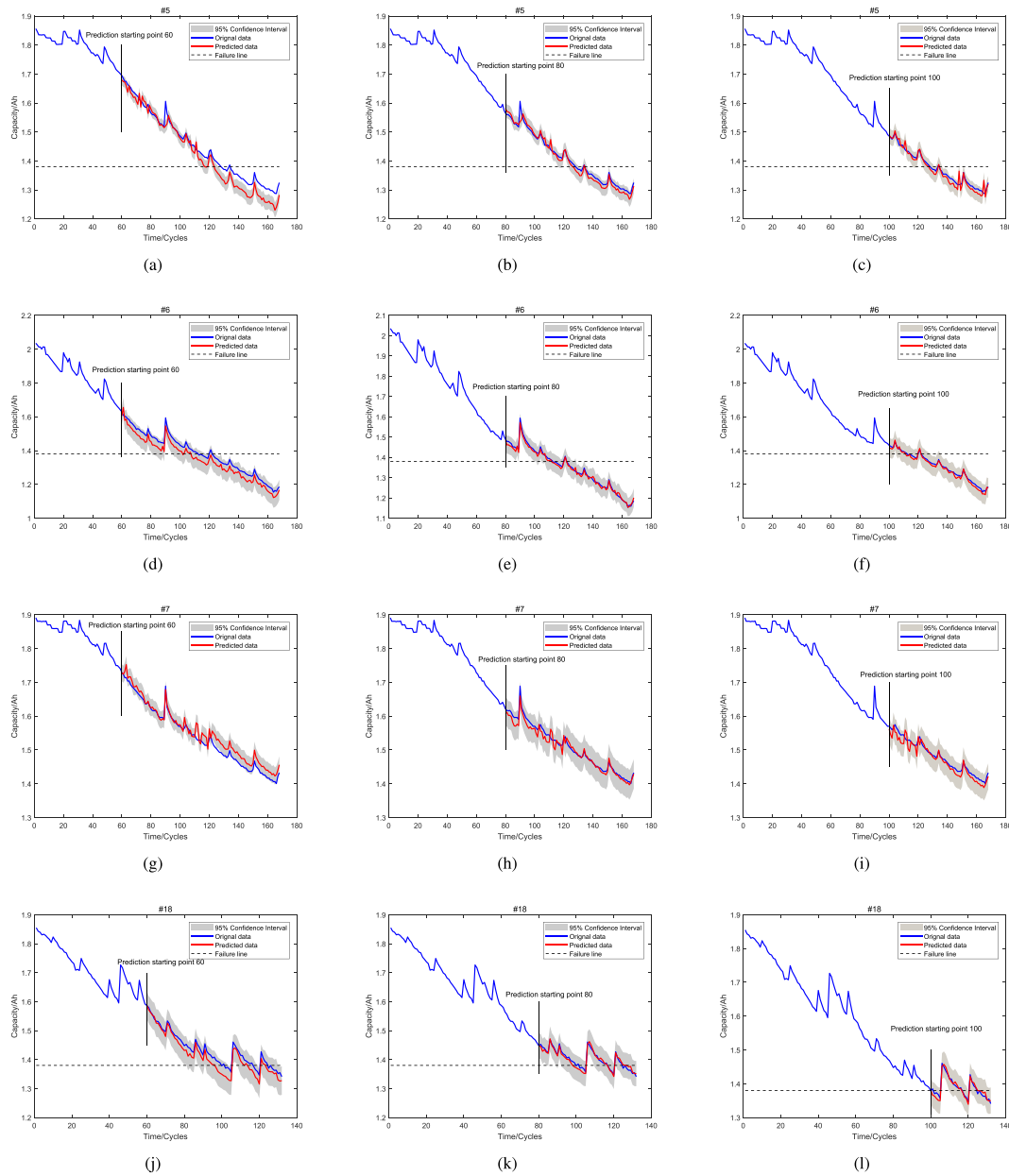


Fig. 6. Prediction results and uncertainty expression of four batteries at different prediction starting points. (a–c) the prediction result of battery #5 at the prediction starting point 60, 80 and 100, respectively; (d–f) the prediction result of battery #6 at the prediction starting point 60, 80 and 100, respectively; (g–i) the prediction result of battery #7 at the prediction starting point 60, 80 and 100, respectively; (j–l) the prediction result of battery #18 at the prediction starting point 60, 80 and 100, respectively.

Table 3
Prediction performance of four batteries at different prediction starting points.

Evaluate criteria	Prediction starting point	#5	#6	#7	#18
RMSE	60	0.0390	0.0355	0.0197	0.0205
	80	0.0175	0.0101	0.0128	0.0080
	100	0.0132	0.0106	0.0141	0.0083
R^2	60	0.9016	0.9209	0.9516	0.9018
	80	0.9586	0.9902	0.9640	0.9431
	100	0.9576	0.9840	0.9230	0.9285
AE	60	10	10	/	4
	80	2	3	/	2
	100	1	1	/	/

Table 4
Prediction models used in comparative experiments (M1 and M2).

Model	Model description
M1	Peak value and regional area under the peak in IC curve as HIs and combined with GPR
M2	Differences between IC curve values at fixed voltage positions as HIs and combined with GPR

model M2.

4.3. Comparisons with related works

The lithium-ion battery dataset used in this paper is popular in prognostic researches, and many results have been reported in recent

Table 5

Prediction performance of different models (M1 and M2) for four batteries.

Evaluate criteria	Model	#5	#6	#7	#18
RMSE	M1	0.0175	0.0101	0.0128	0.0080
	M2	0.0394	0.1120	0.0177	0.0313
R ²	M1	0.9586	0.9902	0.9640	0.9431
	M2	0.7963	0.3400	0.9308	0.5812
AE	M1	2	3	/	2
	M2	12	21	/	12

years. In order to further verify the RUL prediction accuracy of the proposed model, we compare the AE values of the proposed model with those of literature [47,48,49] at different prediction starting points. The specific results are shown in Table 6.

It can be observed that while the proposed method indeed does not significantly outperform the three references in all cases, promising results are still obtained. The predicted RUL AE values of the proposed method at different prediction starting points are relatively small, and no significant AE value is observed for all batteries. Specifically, compared with the work in [47], for all three batteries, the AE values of reference [47] are basically higher than those of the proposed model, and some difference even reaches 30 cycles. Compared with the work in [48], there are only two cases in which the AE value of our method is one more than that of work [48], that is when the prediction starting point for battery #6 and battery #18 is 80. In the rest cases, the AE values of our method are all no more than that of reference [48]. Compared with the work in [49], when the prediction starting point of battery #6 is cycle 60 and battery #18 is cycle 80, the proposed method is only a little worse where the gaps are 1 cycle. However, obvious improvements are achieved by the proposed method in the other cases, where the gaps are basically larger than 2 cycles and even reach 19 cycles. Therefore, the effectiveness of the proposed method on the RUL prediction is validated.

5. Conclusions

RUL prediction can reflect the health state of the battery, which is of great significance to BMS. In this paper, we use a combination of ICA and GPR to predict the RUL of lithium-ion batteries. The main contributions and innovations of this paper are as follows. Firstly, the IC curves are adopted instead of the conventional charge/discharge curves, where the more sensitive characteristics of the degradation process are gained for lithium-ion batteries. Hence, two HIs, namely the peak value and the regional area under the peak value are extracted from the elaborated curves. The validation has been conducted by both qualitative and quantitative evaluations where the proposed HIs have succeeded in reflecting the degradation trend of the battery. Secondly, a novel ICA-GPR method for lithium-ion battery RUL prediction is proposed based on a combination of ICA and GPR, and the uncertainty expression of the prediction result is given. To sum up, the proposed method has many advantages such as high accuracy, reliability, and output being probabilistic and so on. The drawback of this method may be not suitable for other types of batteries due to the difference in discharging IC curves. Thus the range selection procedure of the complete IC curves needs to change. In the future, we need to test the current model in other different aging conditions to study a more accurate and universal battery degradation model.

CRedit authorship contribution statement

Xiaoqiong Pang: Conceptualization, Methodology, Writing—Reviewing and Editing. **Xiaoyan Liu:** Validation, Writing—Original draft preparation, Visualization. **Jianfang Jia:** Supervision. **Jie Wen:** Supervision. **Yuanhao Shi:** Supervision. **Jianchao Zeng:** Supervision. **Zhen Zhao:** Visualization.

Table 6

Comparisons of the RUL prediction AE values with related studies on the same dataset.

Battery	Prediction starting point	Proposed	[47]	[48]	[49]
#5	60	10	24	14	29
	80	2	14	2	12
	100	1	10	12	8
#6	60	10	40	28	9
	80	3	16	2	10
	100	1	1	3	3
#18	60	4	8	10	6
	80	2	5	1	1
	100	/	/	/	/

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.

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China under Grant 7207011096, the Natural Science Foundation of Shanxi Province under Grant 201801D121159, 201801D221208, 201801D121188, 201901D111164, the Graduate Education Innovation Program of Shanxi Province under Grant 2019SY459 and the Research Project Supported by Shanxi Scholarship Council of China under Grant 2020-114.

References

- [1] D.T. Liu, J.B. Zhou, L.M. Guo, Y. Peng, Survey on lithium-ion battery health assessment and cycle life estimation, *Chin. J. Sci. Instrum.* 36 (2015) 116, <https://doi.org/10.19650/j.cnki.cjsi.2015.01.003>.
- [2] Q. Wang, B. Mao, S.I. Stolarov, J. Sun, A review of lithium ion battery failure mechanisms and fire prevention strategies, *Prog. Energy Combust. Sci.* 73 (2019) 95–131, <https://doi.org/10.1016/j.pecs.2019.03.002>.
- [3] L. Lu, X. Han, J. Li, J. Hua, M. Ouyang, A review on the key issues for lithium-ion battery management in electric vehicles, *J. Power Sources* 226 (2013) 272–288, <https://doi.org/10.1016/j.jpowsour.2012.10.060>.
- [4] N. Takami, H. Inagaki, Y. Tatebayashi, H. Saruwatari, K. Honda, S. Egusa, High-power and long-life lithium-ion batteries using lithium titanium oxide anode for automotive and stationary power applications, *J. Power Sources* 244 (2013) 469–475, <https://doi.org/10.1016/j.jpowsour.2012.11.055>.
- [5] N. Williard, W. He, C. Hendricks, M. Pecht, Lessons learned from the 787 dreamliner issue on lithium-ion battery reliability, *Energies* 6 (2013) 4682–4695, <https://doi.org/10.3390/en6094682>.
- [6] Y. Peng, Y. Hou, Y. Song, J. Pang, D. Liu, Lithium-ion battery prognostics with hybrid gaussian process function regression, *Energies* 11 (2018), <https://doi.org/10.3390/en11061420>.
- [7] B. Long, W. Xian, L. Jiang, Z. Liu, An improved autoregressive model by particle swarm optimization for prognostics of lithium-ion batteries, *Microelectron. Reliab.* 53 (2013) 821–831, <https://doi.org/10.1016/j.microrel.2013.01.006>.
- [8] B.E. Olivares, M.A. Cerda Munoz, M.E. Orchard, J.F. Silva, Particle-filtering-based prognosis framework for energy storage devices with a statistical characterization of state-of-health regeneration phenomena, *IEEE Trans. Instrum. Meas.* 62 (2013) 364–376, <https://doi.org/10.1109/tim.2012.2215142>.
- [9] X. Pang, R. Huang, J. Wen, Y. Shi, J. Jia, J. Zeng, A Lithium-ion battery RUL prediction method considering the capacity regeneration phenomenon, *Energies* 12 (2019), <https://doi.org/10.3390/en12122247>.
- [10] B. Pattipati, C. Sankavaram, K.R. Pattipati, System identification and estimation framework for pivotal automotive battery management system characteristics, *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* 41 (2011) 869–884, <https://doi.org/10.1109/tsmcc.2010.2089979>.
- [11] A. Eddahech, O. Briat, E. Woignard, J.M. Vinassa, Remaining useful life prediction of lithium batteries in calendar ageing for automotive applications, *Microelectron. Reliab.* 52 (2012) 2438–2442, <https://doi.org/10.1016/j.microrel.2012.06.085>.
- [12] J. Liu, Z. Chen, Remaining useful life prediction of lithium-ion batteries based on health indicator and gaussian process regression model, *IEEE Access* 7 (2019) 39474–39484, <https://doi.org/10.1109/access.2019.2905740>.
- [13] J. Wu, C. Zhang, Z. Chen, An online method for lithium-ion battery remaining useful life estimation using importance sampling and neural networks, *Appl. Energy* 173 (2016) 134–140, <https://doi.org/10.1016/j.apenergy.2016.04.057>.
- [14] Z. Chen, M. Sun, X. Shu, R. Xiao, J. Shen, Online state of health estimation for lithium-ion batteries based on support vector machine, *Appl. Sci.* 8 (2018), <https://doi.org/10.3390/app8060925>.

- [15] C. Lu, L. Tao, H. Fan, Li-ion battery capacity estimation: a geometrical approach, *J. Power Sources* 261 (2014) 141–147, <https://doi.org/10.1016/j.jpowsour.2014.03.058>.
- [16] C. Weng, Y. Cui, J. Sun, H. Peng, On-board state of health monitoring of lithium-ion batteries using incremental capacity analysis with support vector regression, *J. Power Sources* 235 (2013) 36–44, <https://doi.org/10.1016/j.jpowsour.2013.02.012>.
- [17] C.P. Lin, J. Cabrera, D.Y.W. Yu, F. Yang, K.L. Tsui, SOH estimation and SOC recalibration of lithium-ion battery with incremental capacity analysis & cubic smoothing spline, *J. Electrochem. Soc.* 167 (2020), <https://doi.org/10.1149/1945-7111/ab8f56>.
- [18] M. Dubarry, C. Truchot, B.Y. Liaw, Synthesize battery degradation modes via a diagnostic and prognostic model, *J. Power Sources* 219 (2012) 204–216, <https://doi.org/10.1016/j.jpowsour.2012.07.016>.
- [19] C. Weng, X. Feng, J. Sun, H. Peng, State-of-health monitoring of lithium-ion battery modules and packs via incremental capacity peak tracking, *Appl. Energy* 180 (2016) 360–368, <https://doi.org/10.1016/j.apenergy.2016.07.126>.
- [20] X. Tang, C. Zou, K. Yao, G. Chen, B. Liu, Z. He, F. Gao, A fast estimation algorithm for lithium-ion battery state of health, *J. Power Sources* 396 (2018) 453–458, <https://doi.org/10.1016/j.jpowsour.2018.06.036>.
- [21] Y. Jiang, J. Jiang, C. Zhang, W. Zhang, Y. Gao, Q. Guo, Recognition of battery aging variations for LiFePO₄ batteries in 2nd use applications combining incremental capacity analysis and statistical approaches, *J. Power Sources* 360 (2017) 180–188, <https://doi.org/10.1016/j.jpowsour.2017.06.007>.
- [22] F. Yang, D. Wang, Y. Zhao, K.-L. Tsui, S.J. Bae, A study of the relationship between coulombic efficiency and capacity degradation of commercial lithium-ion batteries, *Energy* 145 (2018) 486–495, <https://doi.org/10.1016/j.energy.2017.12.144>.
- [23] X. Li, C. Yuan, X. Li, Z. Wang, State of health estimation for Li-Ion battery using incremental capacity analysis and Gaussian process regression, *Energy* 190 (2020), <https://doi.org/10.1016/j.energy.2019.116467>.
- [24] H. Tian, P. Qin, K. Li, Z. Zhao, A review of the state of health for lithium-ion batteries: research status and suggestions, *J. Clean. Prod.* 261 (2020), <https://doi.org/10.1016/j.jclepro.2020.120813>.
- [25] M. Dubarry, B.Y. Liaw, Identify capacity fading mechanism in a commercial LiFePO₄ cell, *J. Power Sources* 194 (2009) 541–549, <https://doi.org/10.1016/j.jpowsour.2009.05.036>.
- [26] X. Li, Z. Wang, J. Yan, Prognostic health condition for lithium battery using the partial incremental capacity and Gaussian process regression, *J. Power Sources* 421 (2019) 56–67, <https://doi.org/10.1016/j.jpowsour.2019.03.008>.
- [27] S. Zhang, B. Zhai, X. Guo, K. Wang, N. Peng, X. Zhang, Synchronous estimation of state of health and remaining useful lifetime for lithium-ion battery using the incremental capacity and artificial neural networks, *J. Energy Storage* 26 (2019), <https://doi.org/10.1016/j.est.2019.100951>.
- [28] Y. Li, M. Abdel-Monem, R. Gopalakrishnan, M. Bercebar, E. Nanini-Maury, N. Omar, P. van den Bossche, J. Van Mierlo, A quick on-line state of health estimation method for Li-ion battery with incremental capacity curves processed by Gaussian filter, *J. Power Sources* 373 (2018) 40–53, <https://doi.org/10.1016/j.jpowsour.2017.10.092>.
- [29] A. Guha, A. Patra, Online estimation of the electrochemical impedance spectrum and remaining useful life of lithium-ion batteries, *IEEE Trans. Instrum. Meas.* 67 (2018) 1836–1849, <https://doi.org/10.1109/tim.2018.2809138>.
- [30] M.B. Pinson, M.Z. Bazant, Theory of SEI formation in rechargeable batteries: capacity fade, accelerated aging and lifetime prediction, *J. Electrochem. Soc.* 160 (2013) A243–A250, <https://doi.org/10.1149/2.044302jes>.
- [31] C. Lyu, Q.Z. Lai, T.F. Ge, H.H. Yu, L.X. Wang, N. Ma, A lead-acid battery's remaining useful life prediction by using elec-trochemical model in the Particle Filtering, framework, *Energy* 120 (2017) 975–984, <https://doi.org/10.1016/j.energy.2016.12.004>.
- [32] B.L. Xu, A. Oudalov, A. Ulbig, G. Andersson, D.S. Kirschen, Modeling of lithium-ion battery degradation for cell life assessment 9 (2018) 1131–1140, <https://doi.org/10.1109/tsg.2016.2578950>.
- [33] V. Ramadesigan, K. Chen, N.A. Burns, V. Boovaragavan, R.D. Braatz, V. R. Subramanian, Parameter estimation and capacity fade analysis of lithium-ion batteries using reformulated models, *J. Electrochem. Soc.* 158 (2011) A1048–A1054, <https://doi.org/10.1149/1.3609926>.
- [34] D.U. Sauer, H. Wenzl, Comparison of different approaches for lifetime prediction of electrochemical systems - using lead-acid batteries as example, *J. Power Sources* 176 (2008) 534–546, <https://doi.org/10.1016/j.jpowsour.2007.08.057>.
- [35] D. An, J.-H. Choi, N.H. Kim, Prognostics 101: a tutorial for particle filter-based prognostics algorithm using Matlab, *Reliability Engineering & System Safety* 115 (2013) 161–169, <https://doi.org/10.1016/j.res.2013.02.019>.
- [36] Y. Zhou, M. Huang, Lithium-ion batteries remaining useful life prediction based on a mixture of empirical mode decomposition and ARIMA model, *Microelectron. Reliab.* 65 (2016) 265–273, <https://doi.org/10.1016/j.microrel.2016.07.151>.
- [37] J. Qu, F. Liu, Y. Ma, J. Fan, A neural-network-based method for RUL prediction and SOH monitoring of lithium-ion battery, *IEEE Access* 7 (2019) 87178–87191, <https://doi.org/10.1109/access.2019.2925468>.
- [38] X. Pang, Z. Wang, J. Zeng, J. Jia, Y. Shi, J. Wen, Prediction for the remaining useful life of lithium-ion battery based on PCA-NARX, *Trans. Beijing Inst. Technol.* 39 (2019) 406–412, <https://doi.org/10.15918/j.tbit1001-0645.2019.04.012>.
- [39] D. Gao, M. Huang, Prediction of remaining useful life of lithium-ion battery based on multi-kernel support vector machine with particle swarm optimization, *J. Power Electron.* 17 (2017) 1288–1297, <https://doi.org/10.6113/jpe.2017.17.5.1288>.
- [40] X. Li, L. Zhang, Z. Wang, P. Dong, Remaining useful life prediction for lithium-ion batteries based on a hybrid model combining the long short-term memory and Elman neural networks, *J. Energy Storage* 21 (2019) 510–518, <https://doi.org/10.1016/j.est.2018.12.011>.
- [41] J. Jia, J. Liang, Y. Shi, J. Wen, X. Pang, J. Zeng, SOH and RUL prediction of lithium-ion batteries based on gaussian process regression with indirect health indicators, *Energies* 13 (2020), <https://doi.org/10.3390/en13020375>.
- [42] E. Schulz, M. Speekenbrink, A. Krause, A tutorial on Gaussian process regression: modelling, exploring, and exploiting functions, *J. Math. Psychol.* 85 (2018) 1–16, <https://doi.org/10.1016/j.jmp.2018.03.001>.
- [43] Z. Ge, Process data analytics via probabilistic latent variable models: a tutorial review, *Ind. Eng. Chem. Res.* 57 (2018) 12646–12661, <https://doi.org/10.1021/acs.iecr.8b02913>.
- [44] D. Liu, J. Pang, J. Zhou, Y. Peng, M. Pecht, Prognostics for state of health estimation of lithium-ion batteries based on combination Gaussian process functional regression, *Microelectron. Reliab.* 53 (2013) 832–839, <https://doi.org/10.1016/j.microrel.2013.03.010>.
- [45] B. Saha, K. Goebel, Battery data set, Available online: <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/> (accessed on 20 October 2020).
- [46] B. Jia, M. Xin, Data-driven enhanced nonlinear Gaussian filter 67 (2020) 1144–1148, <https://doi.org/10.1109/tcsii.2019.2926657>.
- [47] J. Yu, State of health prediction of lithium-ion batteries: multiscale logic regression and Gaussian process regression ensemble, *Reliab. Eng. Syst. Saf.* 174 (2018) 82–95, <https://doi.org/10.1016/j.res.2018.02.022>.
- [48] W. Zhang, X. Li, X. Li, Deep learning-based prognostic approach for lithium-ion batteries with adaptive time-series prediction and on-line validation, *Measurement* 164 (2020), <https://doi.org/10.1016/j.measurement.2020.108052>.
- [49] L. Zhang, Z. Mu, C. Sun, Remaining useful life prediction for lithium-ion batteries based on exponential model and particle filter, *IEEE Access* 6 (2018) 17729–17740, <https://doi.org/10.1109/access.2018.2816684>.