IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 2, 2019, accepted March 10, 2019, date of publication March 18, 2019, date of current version April 8, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2905740

Remaining Useful Life Prediction of Lithium-Ion Batteries Based on Health Indicator and Gaussian Process Regression Model

JIAN LIU AND ZIQIANG CHEN, (Senior Member, IEEE)

State Key Laboratory of Ocean Engineering, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai Jiao Tong University, Shanghai 200240, China

Corresponding author: Ziqiang Chen (chenziqiang@sjtu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 51677119.

ABSTRACT Achieving accurate and reliable remaining useful life (RUL) prediction of lithium-ion batteries is very vital for the normal operation of the battery system. The direct RUL prediction based on capacity largely depends on the laboratory condition. A novel method that combines indirect health indicator (HI) and multiple Gaussian process regression (GPR) model is presented for the RUL forecast to solve the capacity unmeasurable problem of operating battery in this paper. First, three measurable HIs are extracted in the constant-current and constant-voltage charge process. Both the Pearson and Spearman rank correlation analytical approaches show that the correlations between HIs and the capacity are good. Then, the GPR model is optimized with combined kernel functions to improve the ability to predict capacity regeneration. Next, based on the measurable HI versus cycle number data, three GPR models are built, and HIs prognosis results are achieved at a single point. The HIs prediction results are added in the multidimensional GPR model, which is accomplished by using HIs and capacity as input and output, respectively. The predicted capacity is used to compare with the threshold to acquire the RUL prediction result. The approach is validated by the two different life-cycle test datasets. The results indicate that an accurate and reliable RUL forecast of lithium-ion batteries can be realized by using the proposed approach.

INDEX TERMS Remaining useful life, lithium-ion battery, health indicator, Gaussian process regression.

I. INTRODUCTION

Owing to many advantages including high energy density, low self-discharge and long lifetime, lithium-ion batteries have been used and developed in a lot of fields, such as automobiles, ships, and satellites [1], [2]. However, during the continuous charge and discharge process of lithium-ion batteries, the performance of lithium-ion batteries will deteriorate with capacity decreasing and impedance increasing, which will cause equipment and system failures or even catastrophic loss [3]. It is vital to achieve accurate and reliable remaining useful life (RUL) prediction of lithium-ion batteries in scientific research and practical application. Through predicting the future variation trend of lithium-ion batteries' state and parameter, the accurate RUL prediction can be accomplished, which is essential for battery management system design, battery prognostics and health management [4], [5]. It is necessary to solve the problems in accuracy and reliability improvement of the prediction methods and on-line recognition of battery degradation state.

Both model-based methods and data-driven methods have been used in RUL forecast of lithium-ion batteries [6]–[8]. Bole *et al.* [9] built a Li-ion battery model by using the electrochemistry theory and used an unscented Kalman filtering algorithm to track parameters under randomized usage. Hu *et al.* [10] comparatively analyzed characteristics of twelve kinds of equivalent circuit models including complexity, accuracy and robustness. Owing to complexity and practical inconvenience of the electrochemical model and the equivalent circuit model, other model-based methods also adopt the empirical model to fit capacity degradation curve and extrapolate the model to forecast RUL with Kalman filter [11], particle filter (PF) [12], unscented particle filter [13] and interacting multiple model particle filter [14]. To develop

The associate editor coordinating the review of this manuscript and approving it for publication was Jiajie Fan.

a semi-empirical life model depending on aging mechanism, Wang *et al.* [15] studied performance degradation of the $LiFePO_4$ battery and established a generalized battery life model based on ampere-hour (Ah) throughput, charge-discharge rate, and temperature. Wang *et al.* [16] developed a conditional three-parameter capacity degradation model which is more suitable for estimating RUL than the sum of two exponential functions. Based on the exponential growth model, Saha *et al.* [1] examined several different PFs to complete the RUL prognosis task. To solve the sample degeneracy and impoverishment problem, Wang *et al.* [17] proposed the spherical cubature particle filter (SCPF) which could provide more accurate RUL predictions than the standard PF-based method. But one shortage of filter methods is that they are based on a system state equation, which may be impractical for battery RUL prediction during its lifetime due to the complexity of electrochemical reactions and degradation mechanism inside the battery [18]. Zhang *et al.* [19] established a linear model with the transformed capacities and cycles by using the Box-Cox transformation and the Monte Carlo simulation to achieve on-board prediction independent of offline training data.

The data-driven methods can mine deterioration information and the evolution law of lithium-ion battery's health state directly. To reduce hardware cost and overcome difficulty in getting an explicit quantitative formula, Wu *et al.* [20] used feed forward neural network (FFNN), a machine learning method in statistical model, to definite RUL reflected by voltage curves. Considering the capacity regeneration, Zhou and Huang [21] combined empirical mode decomposition (EMD) and autoregressive integrated moving average (ARIMA) model to predict the global attenuation of capacity accompanied by fluctuation. Li *et al.* [22] employed the support vector machine (SVM) algorithm to build a regression model by training the terminal voltage and the voltage derivative during the charging process to estimate battery RUL indirectly. These data-driven methods can give a satisfying and accurate prediction result. However, the estimation results of the above methods can only give the point prediction results without the capability to express indeterminacy. The Gaussian process regression (GPR) model can be used in solving regression problems with large dimensions, a few data, and nonlinearity [23]–[25]. Liu *et al.* [26] optimized GPR model with combined kernel functions to prognose state of health (SOH), which can realize the capacity regeneration phenomenon prediction. Richardson *et al.* [27] examined the ability of Gaussian processes for RUL prediction at different periods based on capacity vs. cycle data. He *et al.* [28] utilized wavelet analysis and multiscale GPR method to decouple complex SOH decline curve and accomplish precise SOH estimation. To reduce influence of local regeneration phenomenon, Yu [29] developed a novel approach combining multiscale logic regression model and GPR model, in which the local regenerations and fluctuations of capacity can be forecasted by using the GPR model with the lag vector.

Predicting battery RUL directly with capacity or impendence suffers from accumulated error and difficulty in measurement online [30]. Besides, the capacity measurement is influenced by the time-varying discharge rate and temperature. How to evaluate and describe the current SOH and RUL of battery is challenging. To address those problems, in addition to proposing a proper modeling algorithm, another potential approach is that we can estimate SOH and predict RUL by using the proper degradation features extracted from the measurable parameters [31]. Therefore, indirect approaches are applied gradually based on parameters which can be measured in real-time and online, including current, voltage, temperature, etc. Many scholars have built health indicator (HI) from discharge process such as discharging voltage difference of equal time interval [30], mean voltage falloff [32], sampling entropy of discharge voltage [33], and permutation entropy [34], etc. However, the methods based on the discharge process adopt the constant current discharge mode, which is limited in practical applications because of the varied working conditions and external environment. Williard *et al.* [35] measured four features and compared their effectiveness to estimate the SOH. The four features are capacity, resistance, length of the constant current (CC) charge time, and length of the constant voltage (CV) charge time. Compared with discharge period, the charge process is a relatively static state and the data is more stable, which is more convenient for analysis. Therefore, in this paper, we will combine measurable degradation features extracted from the CC and CV charge process and GPR model to solve the capacity unmeasurable problem and achieve reliable RUL prediction of lithium-ion batteries.

The main contribution in this study is that we propose a novel framework which combines indirect HI and multiple GPR model to achieve RUL prediction of lithium-ion batteries based on measurable degradation features. At first, the three measurable features are extracted as HIs during the CC and CV charge process. Then, with the ability of uncertainty expression, the GPR model is optimized with combined kernel functions to achieve more accurate and reliable prediction after considering the capacity regeneration phenomenon. The multidimensional GPR model can map the relationship between the HIs and the capacity which is established by using the HIs and capacity as the inputs and output. The HIs prediction results can be attained online through training the HIs vs. cycle datasets using single-input and single-output GPR model. Finally, RUL prediction of lithium-ion batteries can be realized by using HIs forecast results and multidimensional GPR model.

The rest of this paper is structured as follows. Firstly, Section 2 presents the datasets and the extraction process of HIs. The GPR model is established in Section 3. Then, the RUL prediction results is displayed and discussed in Section 4. Finally, the last part of this paper draws conclusions.

TABLE 1. The detailed parameters of batteries aging test.

Battery	CC charge current (A)	Cut off current of CV charge (mA)	Discharge current (A)	Cut off voltage of discharge (V)	Threshold of EOL (Ah)
B5	1.500	20	2.00	2.7	1.40
B6	1.500	20	2.00	2.5	1.40
B18	1.500	20	2.00	2.5	1.40
B33	1.500	20	2.00	2.0	1.40
CX36	0.675	50	1.35	2.7	1.08
CX37	0.675	50	1.35	2.7	1.08

II. CONSTRUCTION OF HEALTH INDICATOR

A. DATASETS OF LITHIUM-ION BATTERIES AGING TEST

Two different kinds of datasets of lithium-ion batteries aging test are used in this paper. The first datasets of battery life-cycle test are attained from the Ames Prognostics Center of Excellence in National Aeronautics and Space Administration (NASA) [1], [3]. The 18650 sized lithium-ion batteries with rated capacity of 2 Ah were carried through charge, discharge, and impedance experiments at room temperature (24°C) in NASA. Firstly, the batteries were in CC charge mode at 1.5 A until voltage reached 4.2 V. Then, the batteries were in CV charge mode until current dropped to 20 mA. Discharge process was carried out with constant current until voltage fell to cut off value. The impedance was measured between the charge and discharge process by using electrochemical impedance spectroscopy. The sample time interval varies from 2 seconds to 22 seconds during charge and discharge process. The second datasets are collected form the Center for Advanced Life Cycle Engineering in the University of Maryland [36]. By using the Arbin Battery Tester, the LiC_OO_2 cathode based cells with rated capacity of 1.35 Ah went through full charge and discharge test at about 25°C, which is similar to NASA. The sample time interval is 30 seconds. The end of life (EOL) criterion is defined at 20% or 30% fade in rated capacity. The capacity used in this paper is the full discharged electricity of the filled battery under a certain discharge condition. In this paper, four batteries are selected from datasets of NASA and two batteries are selected from datasets of Maryland, whose detailed test parameters are presented in Table 1.

Figure 1 presents the capacity attenuation curves and the threshold of EOL of the six selected batteries. With the side reaction going on between the electrode and the electrolyte, the lithium-ion is constantly consumed and the capacity presents a degradation trend. However, in the gap between battery charge and discharge process, the side reaction product is likely to dissipate. Therefore, the performance of battery in the next cycle will become better and the capacity will increase compared with the previous cycle. This phenomenon is known as capacity regeneration which results in a capacity declining tendency with the local dynamic fluctuation [37], [38].

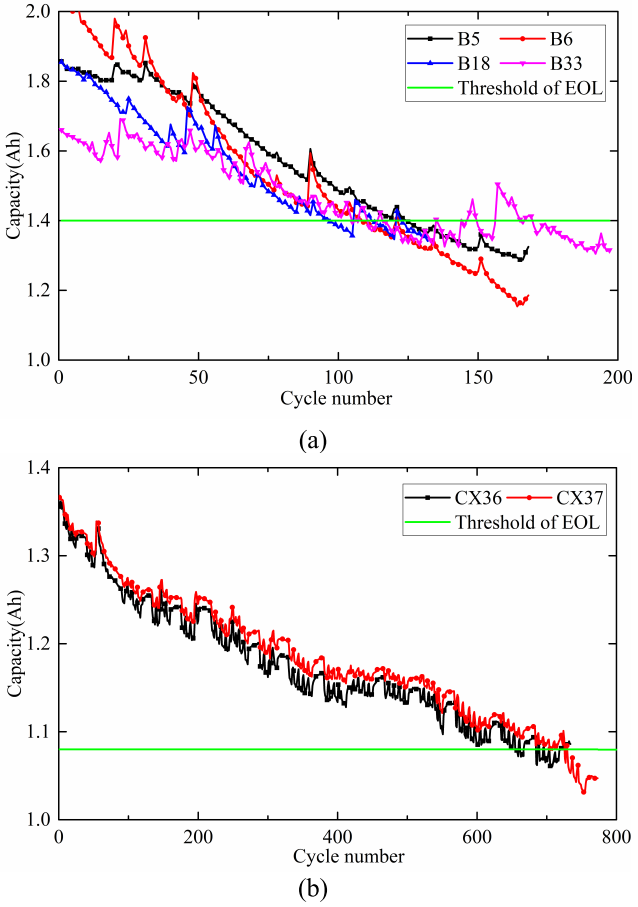

FIGURE 1. The capacity attenuation curves of batteries. (a) B5, B6, B18, and B33; (b) CX36 and CX37.

B. EXTRACTION OF HEALTH INDICATOR

As mentioned in Section 1, the CC and CV charge process are relatively stable and convenient for analysis. Therefore, the health indicator can be extracted from this process. In the operation process, the battery undergoes internal side reactions with lithium consumption and byproducts production. These reactions are accelerated by different usage modes and environmental conditions. Eventually, the charge storage capability of battery is lower than its required performance level and the battery can no longer achieve its intended functions. Although the capacity may be the direct indicator of battery performance, some features can also be extracted from the current, voltage, and temperature parameters to act as the HIs of battery performance.

As illustrated in Figure 2, the voltage and current variation curves of B5 during the CC and CV charge process are obtained at different cycles. As the cyclic charge and discharge process going on, the consuming electrolyte and the decreasing electrode will cause the impedance increasement and the capacity decline. Therefore, the battery will quickly reach the cut off voltage and the time of CC charge process will be reduced. With the cycle number increasing, the time of CV charge process will increase and the slope of the current

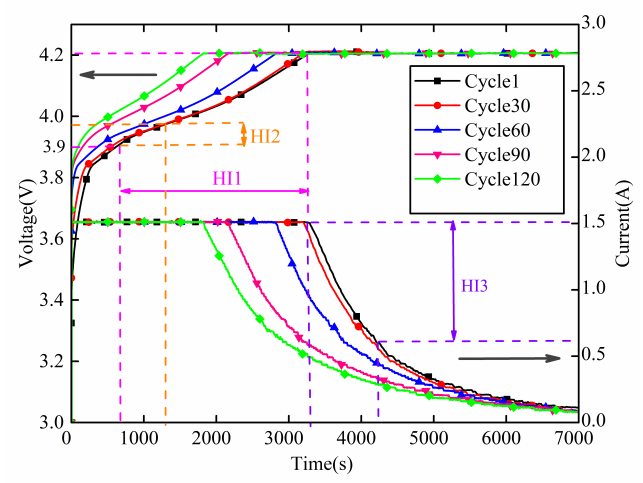

FIGURE 2. The voltage and current variation curves of B5 during CC and CV charge process at different cycles.

variation curves will decrease. There is a certain relationship between CC and CV charge period and the capacity of a battery. In practical applications, the battery will stop to use and start to charge before reaching discharge cut off voltage. Actually, the lithium-ion batteries are seldom fully discharged from the fully charged state. Three features are selected as HIs during the CC and CV charge process to characterize the variation of battery's health state and performance. HI1 is the charge time interval of voltage varying from 3.9 V to 4.2 V; the charge voltage varying from 3.9 V to the voltage after 500 seconds is used as HI2; HI3 is the CV charge current drop between 1.5 A (the CC charge current) and the current after 1000 s. Due to the difference in two datasets, it is worth noting that the HI2 and HI3 in Maryland is different from that in NASA. The HI2 is the charge voltage varying from 3.9 V to the voltage after 600 seconds and the HI3 is the CV charge current drop between the CC charge current and the current after 900 seconds in Maryland. The normalized HIs variation of six lithium-ion batteries are demonstrated in Figures 3-4.

The relationships between capacity and three normalized HIs are quantitatively analyzed by the Pearson and Spearman rank correlation coefficients, as shown in (1)-(2).

$$Pearson = \frac{E(a\beta) - E(\alpha)E(\beta)}{\sqrt{E(\alpha^2) - E^2(\alpha)}\sqrt{E(\beta^2) - E^2(\beta)}} \quad (1)$$

$$Spearman = \frac{\sum_i (\alpha_i - \overline{\alpha})(\beta_i - \overline{\beta})}{\sqrt{\sum_i (\alpha_i - \overline{\alpha})^2}\sqrt{\sum_i (\beta_i - \overline{\beta})^2}} \quad (2)$$

where α and β represents the capacity and the HI, respectively.

The absolute value of correlation coefficient is close to 1, which presents a good linear correlation. The correlation coefficient is equal to 0, which means no linear correlation. Seen from Table 2, most correlation coefficients are upper than 0.9 except for some values of HI3. Both the absolute

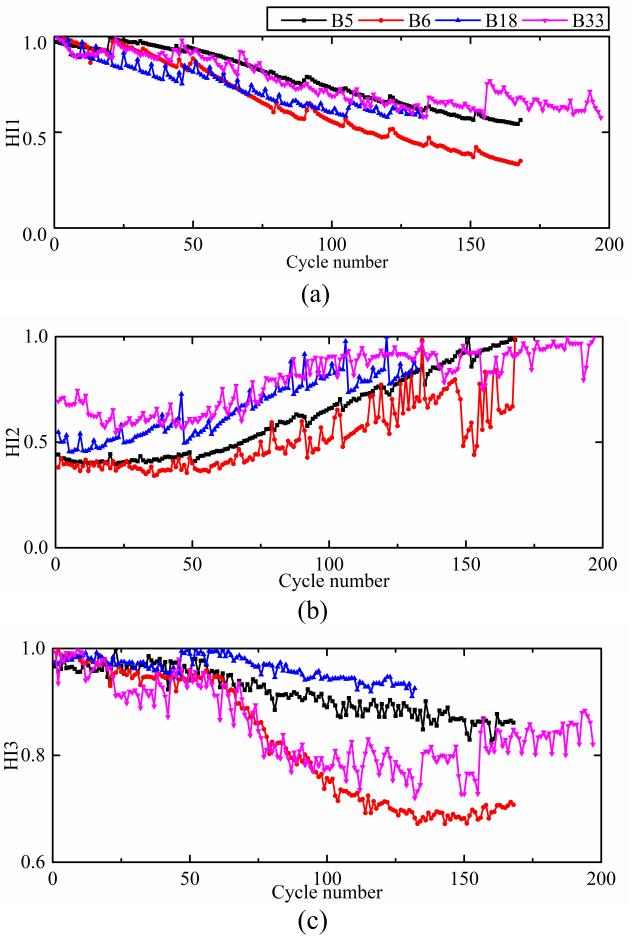

FIGURE 3. The normalized HIs of NASA batteries. (a) HI1; (b) HI2; (c) HI3.

TABLE 2. The correlation coefficient results.

Battery	Pearson			Spearman		
	HI1	HI2	HI3	HI1	HI2	HI3
B5	0.9931	-0.9708	0.9455	0.9909	-0.9753	0.9920
B6	0.9861	-0.8131	0.9374	0.9929	-0.9164	0.9489
B18	0.9822	-0.9282	0.7233	0.9826	-0.9438	0.7594
B33	0.9804	-0.9438	0.8216	0.9570	-0.9103	0.7164
CX36	0.9358	-0.9713	0.8509	0.9455	-0.9782	0.8544
CX37	0.9615	-0.9848	0.8660	0.9743	-0.9882	0.8669

values of Pearson and Spearman correlation coefficients are close to 1, which shows that there is a good correlation between the capacity and three HIs. Therefore, the HIs can be used to represent the capacity to predict the battery RUL.

III. OPTIMIZED GAUSSIAN PROCESS REGRESSION MODEL

A. GAUSSIAN PROCESS REGRESSION MODEL

Gaussian process regression model can achieve state prediction through prior knowledge in Bayesian framework and output prediction mean, variance, and confidence interval with the ability of uncertainty expression [24]–[26].

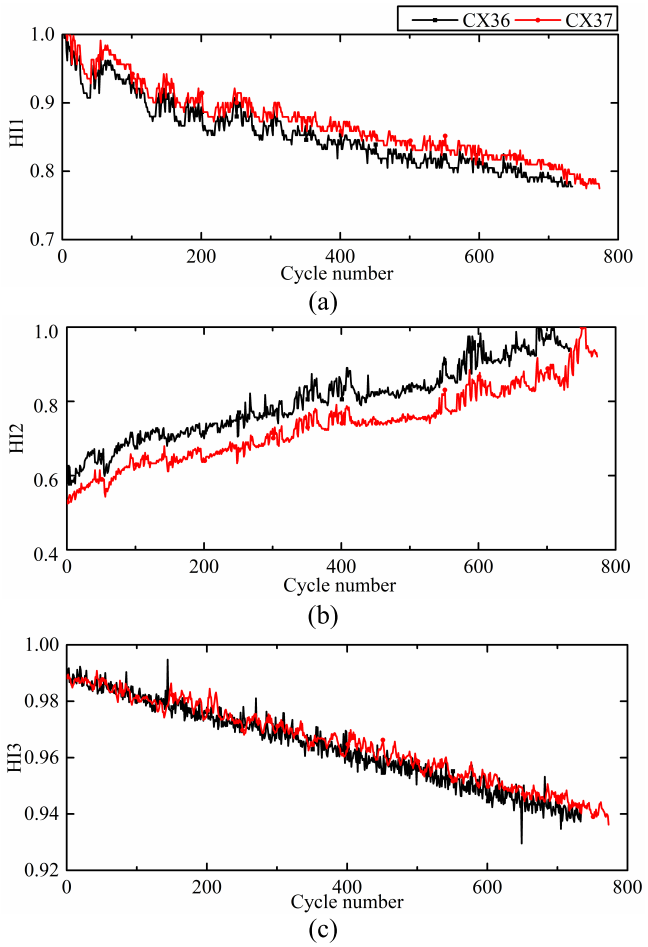

FIGURE 4. The normalized HIs of CX36 and CX37. (a) HI1; (b) HI2; (c) HI3.

The GPR model has been widely used in stock forecasting, fluid level of oil well prediction, time series analysis and prediction, etc.

Consider the following regression function,

$$y = f(x) + \varepsilon, \varepsilon \sim N(0, \sigma_n^2) \tag{3}$$

where x is the input vectors, n is the total number of input vectors, $f(x)$ is the function value, y is the measured value, ε is the white noise whose mean is 0 and variance is σ_n^2.

Then, the prior distribution can be obtained,

$$\mathbf{Y} \sim N(\mathbf{0}, \mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}) \tag{4}$$

The joint prior distribution of measured and predicted values is

$$\begin{bmatrix} \mathbf{Y} \\ \mathbf{f}_* \end{bmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} \mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I} & \mathbf{K}(\mathbf{X}, \mathbf{X}_*) \\ \mathbf{K}(\mathbf{X}_*, \mathbf{X}) & \mathbf{K}(\mathbf{X}_*, \mathbf{X}_*) \end{bmatrix}\right) \tag{5}$$

where $\mathbf{X} = [x_1, x_2, \cdots, x_n]$ is the training set; X_* is testing data; $\mathbf{Y} = [y_1, y_2, \cdots, y_n]$ is measured value set; $\mathbf{f}_* = [f(x_{*1}), f(x_{*2}), \ldots, f(x_{*n})]$ is predictive value set; $\mathbf{K}(\mathbf{X}, \mathbf{X}) = (K_{ij})$ is the symmetric positive definite covariance matrix of n dimensions, matrix elements $K_{ij} = k(x_i, x_j)$ is used to describe correlation between x_i and x_j; $\mathbf{K}(\mathbf{X}, \mathbf{X}_*) = \mathbf{K}(\mathbf{X}_*, \mathbf{X})^T$ is the covariance matrix between $\mathbf{X}$ and $\mathbf{X}_*$, $\mathbf{K}(\mathbf{X}_*, \mathbf{X}_*)$ is covariance matrix about $\mathbf{X}_*$; $\mathbf{I}$ is the identity matrix of n dimensions.

The posteriori distribution of the predicted value $\mathbf{f}_*$ is

$$\mathbf{f}_* \left| \mathbf{X}, \mathbf{Y}, \mathbf{X}_* \sim N(\overline{\mathbf{f}_*}, cov(\overline{\mathbf{f}_*})) \right. \tag{6}$$

It is the GPR model of the set of predicted value $\mathbf{f}_*$.

Its mean matrix is

$$\begin{aligned} \overline{\mathbf{f}_*} &= E[\mathbf{f}_* | \mathbf{X}, \mathbf{Y}, \mathbf{X}_*] \\ &= \mathbf{K}(\mathbf{X}_*, \mathbf{X})[\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]^{-1} \mathbf{Y} \end{aligned} \tag{7}$$

Its covariance matrix is

$$\begin{aligned} &cov(\mathbf{f}_*) \\ &= \mathbf{K}(\mathbf{X}_*, \mathbf{X}_*) - \mathbf{K}(\mathbf{X}_*, \mathbf{X})[\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]^{-1} \mathbf{K}(\mathbf{X}, \mathbf{X}_*) \end{aligned} \tag{8}$$

The 95% confidence interval of the model prediction results is

$$[\overline{\mathbf{f}_*} - 1.96\sqrt{cov(\mathbf{f}_*)}, \overline{\mathbf{f}_*} + 1.96 \times \sqrt{cov(\mathbf{f}_*)}] \tag{9}$$

After determining the mean matrix and covariance matrix, the GPR model can be obtained by optimizing the hyper-parameters during the training process. The GPR model generally adopts the zero mean function and the square exponential covariance function which are show in (10)-(11),

$$m(x) = 0 \tag{10}$$

$$k(x_i, x_j) = \sigma_{f_1}^2 \exp(-\frac{1}{2l_1^2}(x_i - x_j)^2) \tag{11}$$

where σ_{f1}^2 is the signal variance, l_1 is the length-scale, $\boldsymbol{\theta}_1 = [l_1, \sigma_{f1}^2]$ are the hyper-parameters.

The hyper-parameter optimization is generally achieved by using the maximum likelihood method. At first, the negative log-likelihood function of the training data under the conditional probability is obtained, as shown in (12); then, the partial derivative of (12) is shown in (13); finally, the partial derivative minimization is achieved through the conjugate gradient method and the optimum hyper-parameters can be obtained. Gaussian process regression algorithm involves inversion operation. In order to speed up the calculation process, the improved square root method (Cholesky decomposition method) is used to calculate $[\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]^{-1} = \mathbf{L}\mathbf{D}\mathbf{L}^T$ to decompose covariance matrix. By introducing the diagonal matrix instead of the square root operation, the problem of calculation instability can also be avoided.

$$\begin{aligned} L(\boldsymbol{\theta}) = &\frac{1}{2}\mathbf{Y}^T[\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]^{-1}\mathbf{Y} \\ &+ \frac{1}{2}\log\left|\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}\right| + \frac{n}{2}\log 2\pi \end{aligned} \tag{12}$$

$$\begin{aligned} \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_i} = &\frac{1}{2} tr((\boldsymbol{\omega}\boldsymbol{\omega}^T - [\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]^{-1}) \\ &- \frac{\partial[\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]}{\partial \boldsymbol{\theta}_i}) \end{aligned} \tag{13}$$

where $\boldsymbol{\omega} = [\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]^{-1}\mathbf{Y}$.

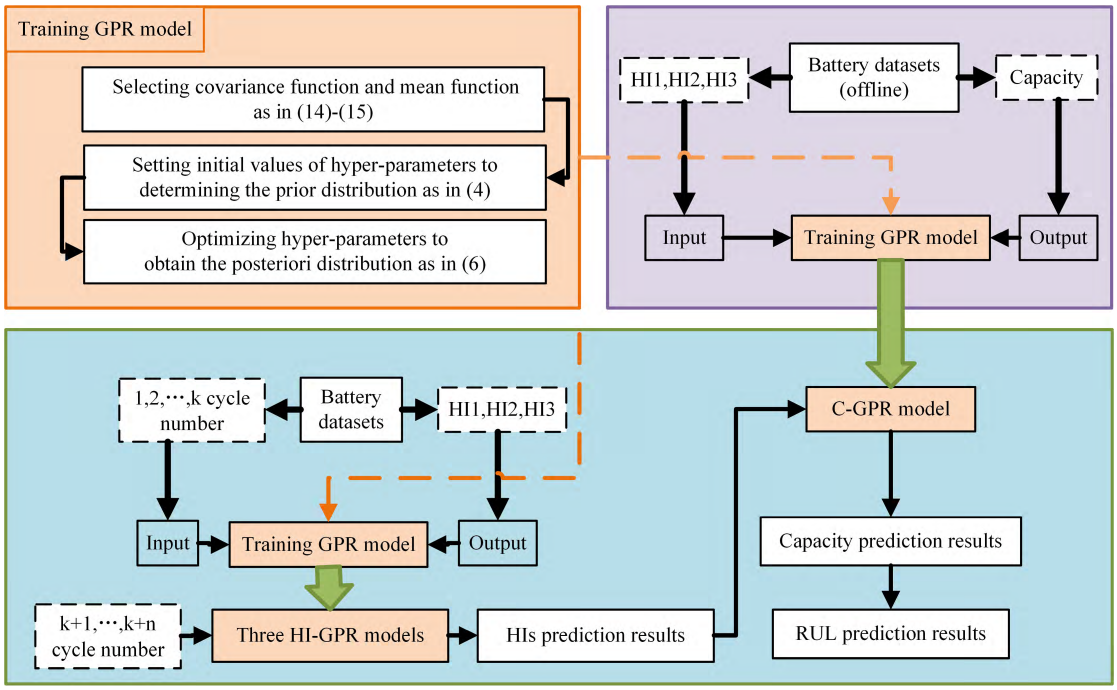

FIGURE 5. The procedure of battery RUL prediction.

B. COMBINED KERNEL FUNCTIONS

Affected by capacity regeneration phenomenon, the capacity degradation of batteries in two datasets shows declining tendency with the local dynamic fluctuation. In order to achieve more accurate capacity prediction, two different capacity variation of lithium-ion batteries should be considered. Therefore, the GPR model only with a single covariance function is incapable of meeting the prediction requirement. The covariance functions can be added to construct the complex covariance function to describe the complex problem [24], [26]. The linear function is selected as the mean function, as shown in (14). The linear mean function is used to improve the ability of multi-steps ahead prediction when the test data is distant from the training data. The local variation of regeneration phenomenon is approximately considered as periodic change in the degradation trend. Therefore, the square exponential covariance function can be used to describe the capacity degradation and the periodic covariance function can be used to reduce the impact of regeneration phenomenon. The square exponential covariance function and periodic covariance function are added as the combination of covariance function with local learning ability and generalization ability, as shown in (15), where p is a periodic parameter.

The hyper-parameters are $\theta_2 = [\mathrm{a}, b, l_1, \sigma_{f1}^2, l_2, \sigma_{f2}^2, p]$.

$$m(x) = ax + b \tag{14}$$

$$k(x_i, x_j) = \sigma_{f_1}^2 \exp(-\frac{1}{2l_1^2}(x_i - x_j)^2) + \sigma_{f2}^2 \exp(-\frac{2}{l_2^2}\sin^2(\frac{2\pi}{p}(x_i - x_j))) \tag{15}$$

C. RUL PREDICTION FRAMEWORK

After measurable HIs series data of previous k cycles is obtained, three one-dimensional GPR models, three HI-GPR models, can be established by training the cycle number and the HI as input and output, respectively. Based on the HIs prediction results of three HI-GPR models, the corresponding prediction results of capacity can be obtained by capacity prediction model which is also called C-GPR model. The C-GPR model, a multi-inputs and single-output GPR model, is trained offline by setting three normalized HIs and the capacity as inputs and output, respectively. Every time to next cycle, three HI-GPR models can be updated by adding the new acquired data into training process. The procedure of battery RUL prediction based on multiple GPR model and indirect HIs is presented as Figure 5.

IV. RESULTS AND DISCUSSION

A. THE EVALUATION CRITERIA

To analyze the prediction accuracy of the proposed method, the absolute error (AE), relative error (RE), mean absolute error (MAE), mean absolute percent error (MAPE) and root mean square error (RMSE) are used in this paper, which are explained in (16)-(20).

$$AE = \left|\phi_i - \hat{\phi}_i\right| \tag{16}$$

$$RE = \frac{\left|\phi_i - \hat{\phi}_i\right|}{\phi_i} \times 100\% \tag{17}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\phi_i - \hat{\phi}_i\right| \tag{18}$$

TABLE 3. The RMSE results of cross-validation.

Battery	B5	B6	B18	B33	CX36	CX37
RMSE	0.0153	0.0290	0.0289	0.0207	0.0891	0.0558

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\phi_i - \hat{\phi}_i}{\phi_i}\right| \tag{19}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\phi_i - \hat{\phi}_i)^2}{n}} \tag{20}$$

where ϕ is the actual or true value and $\hat{\phi}$ is the predicted or tested value.

B. TEN-FOLD CROSS-VALIDATION

In the establishment process of C-GPR model, the ten-fold cross-validation is used to evaluate the RMSE of the regression model. Taking the B5 battery as an example, the training data, three HIs vs. capacity series, are divided into ten subsets or folds. In each cross-validation trial, one subset is served as the test sample while the other nine folds are used as a training set. The cross-validation process is executed ten times until each of the ten folds is taken as the test set exactly once. Therefore, all the data points in the raw dataset are used in both training and testing process. When finishing the ten-fold cross-validation, the RMSE of cross-validation is calculated as the root square of the average error of all ten cross validation tests. The RMSE results are illustrated in Table 3. The RMSE values of six batteries in cross-validation process are lower than 0.1, which shows that the GPR model is well established.

C. RUL PREDICTION AT SINGLE POINT

The B5 battery is selected as the example to show the three HIs and capacity prediction results at single point. At kth cycle, three HIs prediction can be achieved by each HI-GPR model which is trained by measurable data. Based on the C-GPR model and three HIs prediction results, the corresponding capacity prediction is also accomplished. Then, compared with the threshold of EOL, we can get the RUL prediction result at kth cycle. At cycle 100, three HIs and capacity prognosis results are illustrated in Figure 6. As is shown in Fig. 6, we can see that the closer data point is to cycle 100, more accurate the prediction result is. With the cycle number increasement, the prediction results will increasingly deviate from the actual values. In addition to predicting the mean value, the GPR model is also able to provide the confidence interval, which can provide more reliable information for battery RUL forecast and equipment maintenance.

The zero mean function and the square exponential covariance function are taken as the kernel functions of basic GPR model. Compared with basic GPR model, the prediction curves of the combined GPR model using the combined

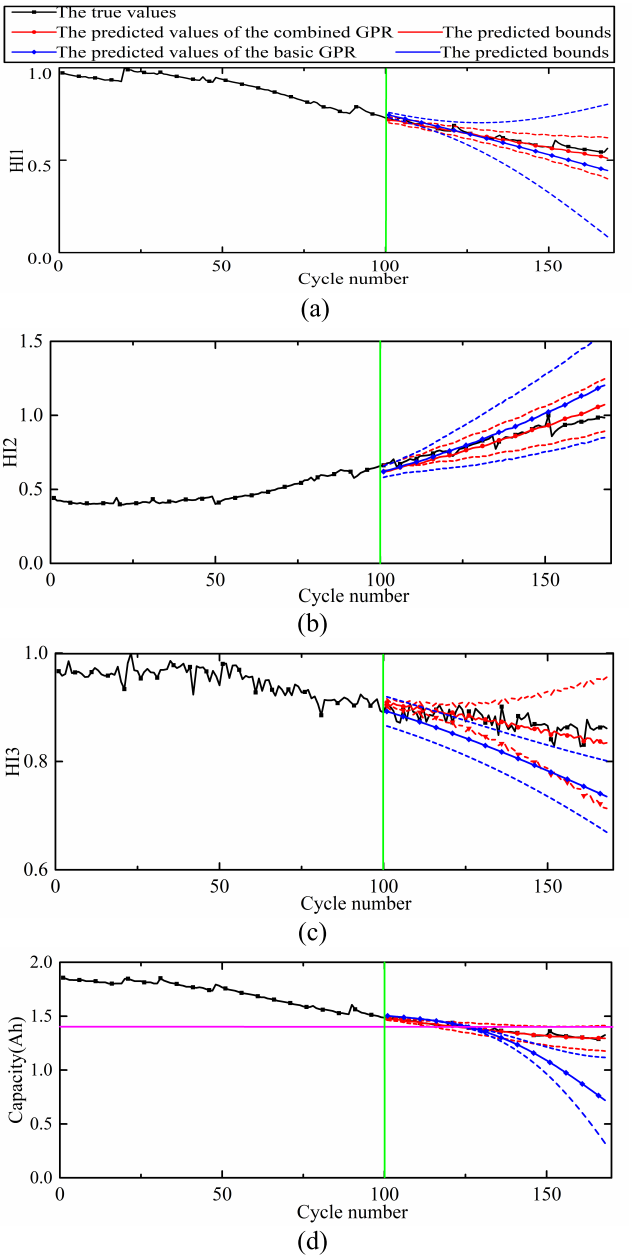

FIGURE 6. The prediction results of B5 at cycle 100. (a) HI1; (b) HI2; (c) HI3; (d) capacity.

kernel functions present degradation trend with local dynamic fluctuations. Using the same length of training data, the combined GPR model works batter in tracking the variation curve of actual value than the basic GPR model. The prediction error results of GPR model with two different kernel functions are presented in Table 4. Seen from Table 4, the RMSE and the MAPE of the combined GPR model are smaller, which suggests that this regression model can achieve more accurate and reliable forecast than the basic GPR model.

To evaluate the RUL prediction accuracy of the proposed method at single point, we compare it with the quadratic function, optimized relevance vector machine (RVM) [32],

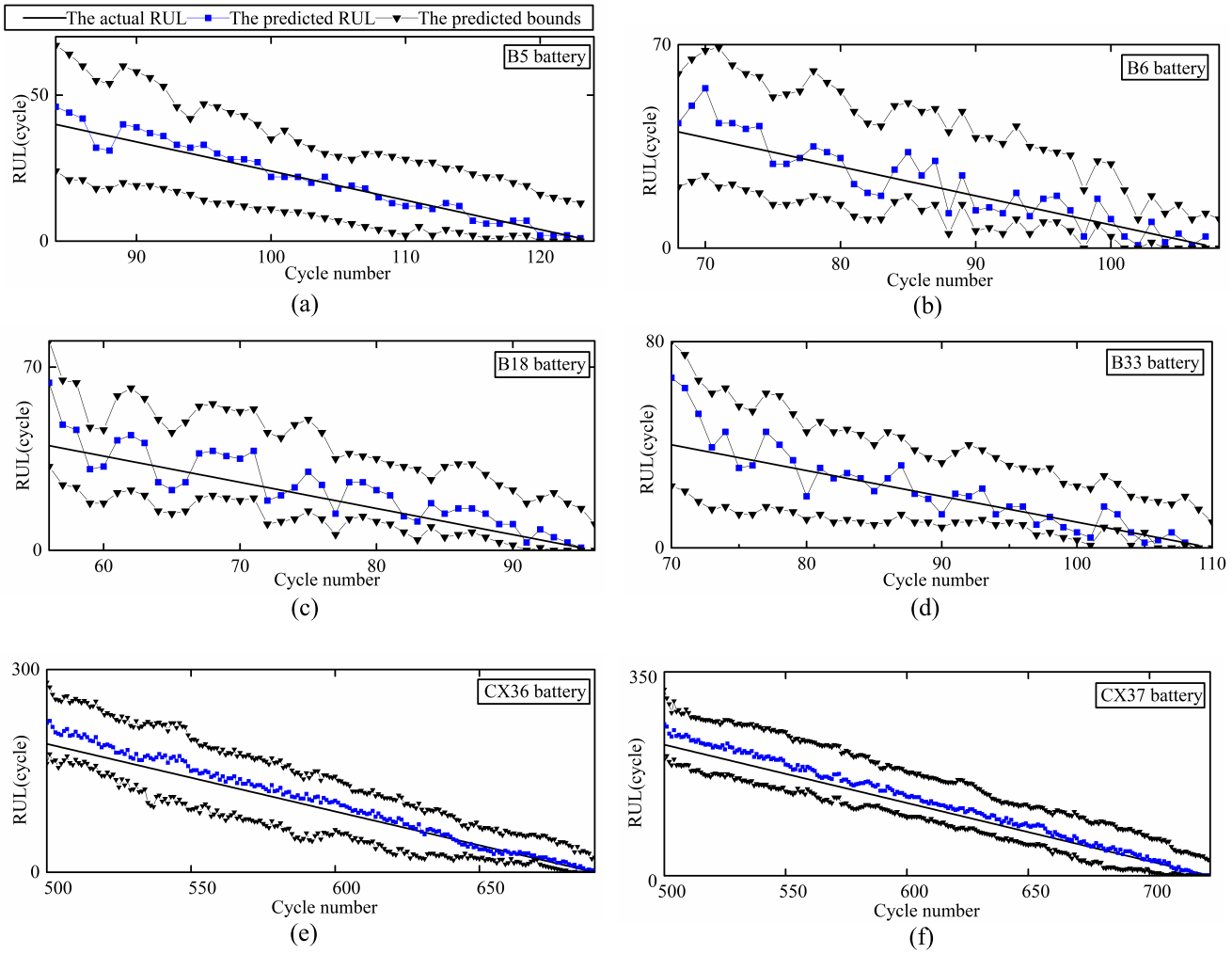

FIGURE 7. The RUL prediction results at different cycles. (a) B5; (b) B6; (c) B18; (d) B33; (e) CX36; (f) CX37.

TABLE 4. The prediction errors of B5 at cycle 100 with different GPR models.

Evaluation criterion	HI1	HI2	HI3	Capacity
RMSE of the proposed GPR	0.0165	0.0410	0.0166	0.0119
MAPE of the proposed GPR	0.0194	0.0402	0.0164	0.0060
RMSE of the basic GPR	0.0457	0.1871	0.0691	0.2289
MAPE of the basic GPR	0.0587	0.1826	0.0679	0.1141

and monotonic echo state network (MONSEN) [39]. The B5, B6, and B18 batteries are used to compare the RUL prediction results at two different staring points in Table 5. It is worth noting that the predicted results of MONSEN method are in the 90% confidence interval and the predicted results of the other two approaches are in the 95% confidence interval. The minimum and maximum AEs of RUL prediction with the proposed approach are 1 and 6 cycles, respectively. Except for the results at a few starting points, the AEs with the other methods are bigger than that with the proposed method. For B5 battery, the average RE of the proposed method is 8.9% and the average RE of the optimized RVM is 10.8%. At four different points, the average RE of the proposed method is 16.3% and the average RE of the optimized RVM is 22.2%, which indicates that the presented approach can achieve more accurate RUL prediction. Besides, the HI extracted from CC and CV charge process is more convenient and feasible than the HI extracted from CC discharge process and the capacity.

D. RUL PREDICTION AT DIFFERENT CYCLES

Based on the battery RUL prediction at the single point, we can achieve the long-term RUL prediction at different cycles. Figure 7 illustrates the RUL prediction results vs. cycle number of six batteries. For the B5, B6, B18, and B33 batteries, the 40 data points before EOL are selected as the analysis data. For the CX36 and CX37 batteries, the RUL prediction starts form cycle 500 to EOL. It can be found

TABLE 5. The comparison of RUL prediction results with different methods.

Battery	Starting point (cycle)	Actual RUL (cycle)	Method	Predicted RUL (cycle)	Predicted confidence interval	AE (cycle)	RE (%)
B5	80	44	The proposed method	50	[24,70]	6	13.6
			Optimized RVM	48	[38,82]	4	9.1
			Quadratic function	23	/	21	47.7
	100	24	The proposed method	23	[11.35]	1	4.2
			Optimized RVM	27	[18,24]	3	12.5
			Quadratic function	12	/	12	50.0
B6	60	48	The proposed method	45	[30,71]	3	6.3
			MONSEN	52	[26,77]	4	8.3
	80	28	The proposed method	31	[15,54]	3	10.7
			MONSEN	22	[11,32]	6	21.4
B18	60	36	The proposed method	42	[22,59]	6	16.7
			MONSEN	55	[30,77]	19	52.8
	80	16	The proposed method	21	[10,33]	5	31.3
			MONSEN	17	[9,24]	1	6.3

TABLE 6. The RUL prediction errors at different cycles.

Evaluation criterion	B5	B6	B18	B33	CX36	CX37
RMSE	3.2122	5.8100	6.9335	7.1908	12.7870	13.9798
MAE	2.5333	4.9268	5.8293	4.8780	10.4712	12.0929

from Fig. 7 that the predicted RUL results fluctuate around the real RUL values. Except for some predictive values with large deviation, most of the prognostic results are near to the real trajectory. With the training data increasement, the deviation between the predicted RUL and the true RUL becomes smaller. The RMSE and the MAE are used to evaluate the RUL forecasting results at different cycles in Table 6. For the B5, B6, B18, and B33 batteries, the RMSE is lower than 10 cycles and the MAE is no more than 6 cycles. For the CX36 and CX37 batteries, both the RMSE and MAE are below 15cycles. The prediction errors in Table 6 show that the presented approach can provide satisfactory RUL prediction performance for lithium-ion batteries at different cycles. In reference [29], for B5 battery, the MAE of RUL prediction results is 11.0 cycles at cycle 80, 90 100, and 110. For B6 battery, the MAE of RUL prediction results is 13.5 cycles at cycle 70, 80, 90, and 100. For B18 battery, the MAE of RUL prediction results is 6.0 cycles at cycle 60, 70, 80, and 90. At the same four cycles, the MAE of B5, B6 and B18 by using the proposed method are 3.8, 4.0 and 6.5 cycles, respectively, which indicates that the RUL prediction in this paper is more precise.

Through the results discussion and analysis, both the RUL prediction at single point and different cycles show satisfactory performances. And the two different kinds of batteries can also validate the adaptability and online application potential of the proposed approach based on the] measurable HIs.

V. CONCLUSIONS

In this paper, a novel RUL prediction framework of lithium-ion battery using indirect HIs and multiple GPR model is presented to achieve the RUL prediction at single point and different cycles. On the one hand, three HIs extracted from CC and CV charge process are convenient and reliable for analysis. The correlation analysis shows that the HIs can be able to characterize the battery health state. On the other hand, the GPR model is optimized by combined kernel functions and multidimensional inputs. The prediction results at single point illustrate that the optimized GPR model has better ability to track the real value change curve. For the long-term RUL prediction, the RMSE and the MAE of batteries in NANA are lower than 10 cycles and 6 cycles, respectively. The results show that the approach proposed in this paper is accurate and effective in RUL prediction. In the future, we will focus on optimizing the health indicators to improve the adaptability to different conditions and practical application. What's more, the further work is to optimize the GPR model for making the prediction of battery RUL more accurate and efficient.

REFERENCES

[1] B. Saha, K. Goebel, S. Poll, and J. Christophersen, "Prognostics methods for battery health monitoring using a Bayesian framework," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 2, pp. 291–296, Feb. 2009. doi: 10.1109/TIM.2008.2005965.

[2] I. S. Kim, "A technique for estimating the state of health of lithium batteries through a dual-sliding-mode observer," *IEEE Trans. Power Electron.*, vol. 25, no. 4, pp. 1013–1022, Apr. 2010. doi: 10.1109/TPEL.2009.2034966.

[3] B. Saha, K. Goebel, and J. Christophersen, "Comparison of prognostic algorithms for estimating remaining useful life of batteries," *Trans. Inst. Meas. Control*, vol. 31, nos. 3–4, pp. 293–308, Jun. 2009. doi: 10.1177/0142331208092030.

[4] R. Xiong, Y. Zhang, J. Wang, H. He, S. Peng, and M. Pecht, "Lithium-ion battery health prognosis based on a real battery management system used in electric vehicles," *IEEE Trans. Veh. Technol*, to be published. doi: 10.1109/TVT.2018.2864688.

[5] K. Goebel, B. Saha, A. Saxena, J. Celaya, and J. Christophersen, "Prognostics in battery health management," *IEEE Instrum. Meas Mag.*, vol. 11, no. 4, pp. 33–40, Aug. 2008. doi: 10.1109/MIM.2008.4579269.

[6] J. Zhang and J. Lee, "A review on prognostics and health monitoring of Li-ion battery," *J. Power Sour.*, vol. 196, pp. 6007–6014, Aug. 2011. doi: 10.1016/j.jpowsour.2011.03.101.

[7] S. M. Rezvanizaniani, Z. Liu, Y. Chen, and J. Lee, "Review and recent advances in battery health monitoring and prognostics technologies for electric vehicle (EV) safety and mobility," *J. Power Sour.*, vol. 256, pp. 110–124, Jun. 2014. doi: 10.1016/j.jpowsour.2011.03.101.

[8] R. Xiong, L. Li, and J. Tian, "Towards a smarter battery management system: A critical review on battery state of health monitoring methods," *J. Power Sources*, vol. 405, no. 5, pp. 18–29, Nov. 2018. doi: 10.1016/j.jpowsour.2018.10.019.

[9] B. Bole, C. S. Kulkarni, and M. Daigle, "Adaptation of an electrochemistry-based li-ion battery model to account for deterioration observed under randomized use," in *Proc. Annu. Conf. Prognostics Health Manage. Soc.*, 2014, pp. 1–9.

[10] X. Hu, S. Li, and H. Peng, "A comparative study of equivalent circuit models for Li-ion batteries," *J. Power Sour.*, vol. 198, pp. 359–367, Jan. 2015. doi: 10.1016/j.jpowsour.2011.10.013.

[11] C. Huang, Z. Wang, Z. Zhao, L. Wang, C. S. Lai, and D. Wang, "Robustness evaluation of extended and unscented Kalman filter for battery state of charge estimation," *IEEE Access*, vol. 6, pp. 27617–27628, 2018. doi: 10.1109/ACCESS.2018.2833858.

[12] L. Zhang, Z. Mu, and C. Sun, "Remaining useful life prediction for lithium-ion batteries based on exponential model and particle filter," *IEEE Access*, vol. 6, pp. 17729–17740, Mar. 2018. doi: 10.1109/ACCESS.2018.2816684.

[13] D. Liu, X. Yin, Y. Song, W. Liu, and Y. Peng, "An on-line state of health estimation of lithium-ion battery using unscented particle filter," *IEEE Access*, vol. 6, pp. 40990–41001, Jul. 2018. doi: 10.1109/ACCESS.2018.2854224.

[14] X. H. Su, S. Wang, M. Pecht, L. L. Zhao, and Z. Ye, "Interacting multiple model particle filter for prognostics of lithium-ion batteries," *Microelectron. Rel.*, vol. 70, pp. 59–69, Mar. 2017. doi: 10.1016/j.microrel.2017.02.003.

[15] J. Wang *et al.*, "Cycle-life model for graphite-$LiFePO_4$cells," *J. Power Sour.*, vol. 196, no. 8, pp. 3942–3948, Apr. 2011. doi: 10.1016/j.jpowsour.2010.11.134.

[16] D. Wang, Q. Miao, and M. Pecht, "Prognostics of lithium-ion batteries based on relevance vectors and a conditional three-parameter capacity degradation model," *J. Power Sour.*, vol. 239, pp. 253–264, Oct. 2013. doi: 10.1016/j.jpowsour.2013.03.129.

[17] D. Wang, F. Yang, K.-L. Tsui, Q. Zhou, and S. J. Bae, "Remaining useful life prediction of lithium-ion batteries based on spherical cubature particle filter," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 6, pp. 1282–1291, Jun. 2016. doi: 10.1109/TIM.2016.2534258.

[18] D. Yang, X. Zhang, R. Pan, Y. Wang, and Z. Chen, "A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve," *J. Power Sour.*, vol. 384, pp. 387–395, Apr. 2018. doi: 10.1016/j.jpowsour.2018.03.015.

[19] Y. Zhang, R. Xiong, H. He, and M. Pecht, "Lithium-ion battery remaining useful life prediction with Box–Cox transformation and Monte Carlo simulation," *IEEE Trans. Ind. Electron*, vol. 66, no. 2, pp. 1585–1597, Feb. 2019. doi: 10.1109/TIE.2018.2808918.

[20] J. Wu, C. Zhang, and Z. Chen, "An online method for lithium-ion battery remaining useful life estimation using importance sampling and neural networks," *Appl. Energy*, vol. 173, pp. 134–140, Jul. 2016. doi: 10.1016/j.apenergy.2016.04.057.

[21] Y. Zhou and M. Huang, "Lithium-ion batteries remaining useful life prediction based on a mixture of empirical mode decomposition and ARIMA model," *Microelectron. Rel.*, vol. 65, pp. 265–273, Oct. 2016. doi: 10.1016/j.microrel.2016.07.151.

[22] X. Li, X. Shu, J. Shen, R. Xiao, W. Yan, and Z. Chen, "An on-board remaining useful life estimation algorithm for lithium-ion batteries of electric vehicles," *Energies*, vol. 10, no. 5, p. 691, May 2017. doi: 10.3390/en10050691.

[23] G. O. Sahinoglu, M. Pajovic, Z. Sahinoglu, Y. Wang, P. V. Orlik, and T. Wada, "Battery state-of-charge estimation based on regular/recurrent gaussian process regression," *IEEE Trans. Ind. Electron*, vol. 65, no. 5, pp. 4311–4321, May 2018. doi: 10.1109/TIE.2017.2764869.

[24] M. Seeger, "Gaussian processes for machine learning," *Int. J. Neural Syst.*, vol. 14, no. 2, pp. 69–106, 2004. doi: 10.1142/S0129065704001899.

[25] J. Hu and J. Wang, "Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression," *Energy*, vol. 93, pp. 1456–1466, Dec. 2015. doi: 10.1016/j.energy.2015.10.041.

[26] D. Liu, J. Pang, J. Zhou, Y. Peng, and M. Pecht, "Prognostics for state of health estimation of lithium-ion batteries based on combination Gaussian process functional regression," *Microelectron. Rel.*, vol. 53, no. 6, pp. 832–839, Jun. 2013. doi: 10.1016/j.microrel.2013.03.010.

[27] R. R. Richardson, M. A. Osborne, and D. A. Howey, "Gaussian process regression for forecasting battery state of health," *J. Power Sour.*, vol. 357, pp. 209–219, Jul. 2017. doi: 10.1016/j.jpowsour.2017.05.004.

[28] Y. He, J. Shen, J. Shen, and Z. Ma, "State of health estimation of lithium-ion batteries: A multiscale Gaussian process regression modeling approach," *AIChE J.*, vol. 61, no. 5, pp. 1589–1600, May 2015. doi: 10.1002/aic.14760.

[29] J. Yu, "State of health prediction of lithium-ion batteries: Multiscale logic regression and Gaussian process regression ensemble," *Rel. Eng. Syst. Saf.*, vol. 174, pp. 82–95, Jun. 2018. doi: 10.1016/j.ress.2018.02.022.

[30] D. Liu, J. Zhou, H. Liao, Y. Peng, and X. Peng, "A health indicator extraction and optimization framework for lithium-ion battery degradation modeling and prognostics," *IEEE Trans. Syst., Man, Cybern. Syst*, vol. 45, no. 6, pp. 915–928, Jun. 2015. doi: 10.1109/TSMC.2015.2389757.

[31] D. Liu, Y. Song, L. Li, H. Liao, and Y. Peng, "On-line life cycle health assessment for lithium-ion battery in electric vehicles," *J. Cleaner Prod.*, vol. 199, pp. 1050–1065, Oct. 2018. doi: 10.1016/j.jclepro.2018.06.182.

[32] Y. Zhou, M. Huang, Y. Chen, and Y. Tao, "A novel health indicator for on-line lithium-ion batteries remaining useful life prediction," *J. Power Sour.*, vol. 321, pp. 1–10, Jul. 2016. doi: 10.1016/j.jpowsour.2016.04.119.

[33] A. Widodo, M.-C. Shim, W. Caesarendra, and B.-S. Yang, "Intelligent prognostics for battery health monitoring based on sample entropy," *Expert Syst. Appl.*, vol. 38, no. 9, pp. 11763–11769, Sep. 2011. doi: 10.1016/j.eswa.2011.03.063.

[34] L. Chen, L. Xu, and Y. Zhou, "Novel approach for lithium-ion battery on-line remaining useful life prediction based on permutation entropy," *Energies*, vol. 11, no. 4, p. 820, Apr. 2018. doi: 10.3390/en11040820.

[35] N. Williard, W. He, M. Osterman, and M. Pecht, "Comparative analysis of features for determining state of health in lithium-ion batteries," *Int. J. Prognostics Health Manage.*, vol. 4, no. 1, pp. 1–7, Jun. 2013.

[36] Y. Xing, E. W. M. Ma, K. L. Tsui, and M. Pecht, "An ensemble model for predicting the remaining useful performance of lithium-ion batteries," *Microelectron. Rel.*, vol. 53, no. 6, pp. 811–820, Jun. 2013. doi: 10.1016/j.microrel.2012.12.003.

[37] M. E. Orchard, L. Tang, B. Saha, K. Goebel, and G. Vachtsevanos, "Risk-sensitive particle-filtering-based prognosis framework for estimation of remaining useful life in energy storage devices," *Stud. Inform. Control*, vol. 19, no. 3, pp. 209–218, Sep. 2010.

[38] T. Qin, S. Zeng, J. Guo, and Z. Skaf, "A rest time-based prognostic framework for state of health estimation of lithium-ion batteries with regeneration phenomena," *Energies*, vol. 9, no. 11, p. 896, Nov. 2016. doi: 10.3390/en9110896.

[39] D. Liu, W. Xie, H. Liao, and Y. Peng, "An integrated probabilistic approach to lithium-ion battery remaining useful life estimation," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 3, pp. 660–670, Mar. 2015. doi: 10.1109/TIM.2014.2348613.

JIAN LIU received the B.S. degree from the Department of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, China, in 2016, where he is currently pursuing the M.S. degree in marine engineering with the State Key Laboratory of Ocean Engineering.

His main research interests include the areas of the battery health state estimation and remaining useful life prediction.

ZIQIANG CHEN (SM'18) received the Ph.D. degree in mechanical engineering from Xi'an Jiaotong University, Xi'an, China, in 1999.

He was a Visiting Scholar with the Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI, USA, from 2011 to 2012. Since 2007, he has been with Shanghai Jiao Tong University, Shanghai, China, where he is currently a Professor with the Department of Naval Architecture, Ocean and Civil Engineering. His research interests include the areas of system identification and state estimation, fault diagnosis, discrete-event systems, automotive and electric vehicle control, battery management systems, composite energy storage system, intelligent industrial detection, robot applications, and intelligent manufacturing.

Dr. Chen received the National Scientific Conference Award, the Ministerial Second Prize Award, and the Excellence Award of Chinese Invention Patent.