A Health Indicator Extraction and Optimization Framework for Lithium-Ion Battery Degradation Modeling and Prognostics

Datong Liu, *Member, IEEE*, Jianbao Zhou, Haitao Liao, *Member, IEEE*, Yu Peng, *Member, IEEE*, and Xiyuan Peng

***Abstract*—Maximum releasable capacity and internal resistance are often used as the health indicators (HIs) of a lithium-ion battery for degradation modeling and estimation of remaining useful life (RUL). However, the maximum releasable capacity is usually difficult to estimate in online applications due to complex operating conditions in the field. Moreover, measuring the internal resistance is too expensive to be implemented on-line. In this paper, an HI extraction and optimization framework requiring only the operating parameters of lithium-ion batteries is proposed for battery degradation modeling and RUL estimation. The framework carries out raw HI extraction, transformation, correlation analysis, and verification and evaluation to achieve HI enhancement. In particular, the Box–Cox transformation is adopted to improve the correlation between the extracted HI and the battery's actual degradation state. To estimate the battery's RUL using the enhanced HI, an optimized relevance vector-machine algorithm is utilized, which can be performed in a flexible and agile way. Experimental studies using two different industrial testing data sets illustrate the high efficiency and adaptability of the proposed framework in lithium-ion battery degradation modeling and RUL estimation.**

***Index Terms*—Box–Cox transformation, correlation analysis, health indicator (HI), lithium-ion battery, prognostics and health management (PHM), remaining useful life (RUL).**

I. Introduction

LITHIUM-ION batteries have been widely used in many fields, such as communications, consumer electronics, electric vehicles, navigation, and aviation, due to their high energy density, high output voltage, low self-discharge rate, long lifetime, and many other advantages [1], [2]. To meet the requirements on safety management, and charging and discharging control of lithium-ion batteries, research on battery capacity fade and estimation of remaining useful life (RUL) has become a hotspot in prognostics and health management (PHM), reliability engineering, and other related fields. Indeed, lithium-ion battery RUL estimation is considered one of the most important issues in electronics PHM [3], [4]. The idea is to use the current and historical data of a battery system to forecast the system's future states under certain operating conditions. Essentially, this useful information can provide an alarm before faults reach critical levels, so that timely maintenance can be performed to prevent malfunction, or even catastrophic failures [5]–[7].

In recent years, extensive research has been conducted on lithium-ion battery degradation modeling and RUL estimation. For example, Goebel *et al.* [3] and Olivares *et al.* [8] achieved battery RUL prediction and the uncertainty representation with particle filter (PF). An empirical degradation model is developed considering the coulombic efficiency factor and capacity regeneration phenomenon. Moreover, the PF method as well as its extended versions has been widely adopted in prognostics [9]–[13]. For example, Olivares *et al.* [8] and Orchard *et al.* [14] proposed an RUL estimation framework that combines PF and an anomaly detection model for battery capacity estimation and cycle life prediction, in which the capacity regeneration phenomenon is considered. Recently, Miao *et al.* [11] presented an unscented PF algorithm to improve the accuracy of lithium-ion battery RUL estimation.

As artificial intelligence and machine-learning techniques rapidly advance, data-driven prognostics for lithium-ion battery have become quite popular. For example, autoregressive and moving average (ARMA) models have been used, which extrapolate the battery capacity from the battery's internal parameters [5]. As an alternative, Liu *et al.* [12] developed an adaptive recurrent neural network model for lithium-ion battery prognostics. The electrolyte resistance and charge transfer resistance are adopted as the inputs of the model, and their future values are predicted to compute the corresponding battery's capacity and RUL. Similar to support vector machine (SVM), relevance vector machine (RVM), which is based on sparse Bayesian learning and kernel learning, has been used to quantify RUL estimation uncertainty [15]–[17]. Alternatively, Liu *et al.* [18] used a Gaussian process regression approach to perform state-of-health (SOH) forecasting for lithium-ion batteries. The approach provides both the

Manuscript received February 20, 2014; revised June 10, 2014; accepted October 13, 2014. Date of publication January 22, 2015; date of current version May 13, 2015. This work was supported in part by the National Natural Science Foundation of China under Grant 61301205, in part by the Research Fund for the Doctoral Program of Higher Education of China under Grant 20112302120027, in part by the Natural Scientific Research Innovation Foundation in Harbin Institute of Technology under Grant HIT.NSRIF.2014017, and in part by the China Scholarship Council. This paper was recommended by Associate Editor G. Biswas.

D. Liu, J. Zhou, Y. Peng, and X. Peng are with the Department of Automatic Test and Control, Harbin Institute of Technology, Harbin 150080, China (e-mail: liudatong@hit.edu.cn).

H. Liao is with the Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMC.2015.2389757

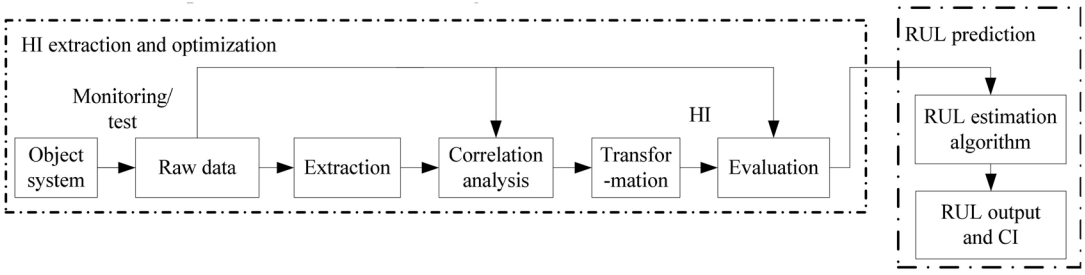

Fig. 1. Framework for HI extraction and transformation and RUL prediction for a general application.

mean and variance of SOH. The mixed Gaussian process functional regression is also introduced to track and predict the global capacity degradation and local capacity regeneration to improve the estimation accuracy.

To improve the performance of battery RUL estimation, developing fusion prognostic methods that complement the capabilities of different approaches has attracted much attention recently. Kozlowski [19] proposed a data-driven RUL prediction approach by combining ARMA model, neural networks, and fuzzy logic. Liu *et al.* [20] explored an ensemble echo state networks to realize satellite lithium battery RUL prediction with high prediction performance. Liu *et al.* [21] developed a fusion prognostic framework to increase the system long-term prediction performance. Xing *et al.* [22] proposed an ensemble model for predicting the RUL of a lithium-ion battery by combining a fused empirical exponential model, a polynomial regression model, and a PF algorithm. In addition, Hu *et al.* [23] combined RVM, SVM, exponential fitting, Bayesian linear regression model, and recurrent neural networks to improve the estimation accuracy and precision.

It is worth pointing out that most of the related research on lithium-ion battery RUL estimation is mainly focused on developing various algorithms to improve estimation accuracy and efficiency. These methods often utilize the capacity [24] or internal resistance as the health indicator (HI) of a battery. However, in on-line applications, it is difficult to perform battery capacity measurement or monitoring. One reason is that these batteries may not be deeply discharged from 100% state-of-charge (SoC) to 0% SoC or be fully charged from 0% SoC to 100% SoC. Moreover, such internal state variables may be inaccessible via regular sensors. There are several types of internal resistance parameters which can also reflect the actual health status of lithium-ion battery. However, directly and precisely measuring these internal parameters, e.g., with electrochemical impedance spectroscopy (EIS) test, is expensive, and is difficult to implement for many online applications [25], [26]. As a result, it is important to extract some performance parameters reflecting the health state of a battery from available monitoring parameters.

To the best of our knowledge, little work has been conducted on the use of operating parameters in battery RUL estimation, which requires more advanced methods for the identification and extraction of degradation features. Widodo *et al.* [27] proposed the use of sample entropy of discharge voltage under a prognostic framework for battery health assessment. This method provides a useful computational tool for assessing the predictability of a time series and can also quantify the regularity of a data sequence. However, this technique is time-consuming and also requires the capacity parameter in evaluating the sample entropy indicator. Williard *et al.* [28] developed two different HIs, the length of time of constant current (CC) and the constant voltage (CV) phases of charging. The final SOH estimation is integrated with several available HIs, in which the two new HIs are evaluated and compared with the battery capacity and resistance.

For lithium-ion battery RUL estimation and other similar applications, HI extraction is indeed critical in characterizing the underlying degradation process. To overcome the aforementioned problems and challenges, we propose a novel method for indirect HI extraction based on the charging and discharging voltage, current, and time for on-line lithium-ion battery degradation analysis. We also develop a framework that can reliably provide accurate on-line battery RUL estimates after HI enhancement. Technically, considering the complicated relationship between the extracted HI and the nonlinearly degrading battery capacity, this paper proposes the use of nonlinear transformation and optimization of the extracted HI to enhance the correlation between the HI and battery capacity. We also present an optimized RVM algorithm to improve the accuracy and stability of RUL estimation and to provide an uncertainty representation for the resulting RUL estimates.

The remainder of this paper is organized as follows. The HI extraction and optimization framework, including the related methodologies, is introduced in Section II. The way of applying the framework to lithium-ion battery RUL estimation is described in detail in Section III. HI evaluation and verification with actual lithium-ion batteries are demonstrated in Section IV. Section V illustrates the efficiency of the proposed framework in lithium-ion battery RUL estimation. Finally, the conclusion is provided in Section VI.

II. Methodology

A. Structure of the Framework

Fig. 1 shows the proposed HI extraction and optimization framework that improves the accuracy of HI in representing the system/component health state. It is worth pointing out that the proposed framework is well suited for on-line applications, which is the main contribution of this paper.

The framework uses the raw testing and monitoring data for HI extraction and optimization. First, the trend of the operating

parameters is analyzed. Then, the HI parameters are extracted from these parameters, and correlation analysis is conducted to verify the representation capability of these HI parameters in degradation quantification. Through the nonlinear Box–Cox transformation, the extracted HI is enhanced to improve the correlation between the HI and the actual battery health status. Finally, the optimized HI is utilized in RUL estimation. In this paper, we use the RVM algorithm for RUL estimation, in which the series reconstruction, RVM modeling training, and iterative prediction are involved.

The detailed procedure for implementing the framework in practice is described as follows.

Step 1: *HI Extraction (Feature Extraction):* With the online monitoring parameters, we extract and format available HI to indicate the health status of a battery. To achieve a more robust HI, segmentation, and fusion technology are applied to improve the stability and reliability of this step.

Step 2: *Correlation Analysis:* Correlation analysis and evaluation are performed to verify the level of correlation between the extracted HI and the battery's health status. To identify the nonlinear transformation parameters, Pearson and Spearman rank correlation analysis methods are implemented.

Step 3: *HI Transformation and Optimization:* The Box–Cox transformation is applied to enhance the HI. In particular, the extracted HI is used as the independent variable, and the degrading battery capacity is treated as the response.

Step 4: *HI Verification and Evaluation:* The similarity measurement and correlation analysis are performed for improving the usability and efficiency of the HI in indicating the actual health state.

Step 5: *RUL Estimation:* A data-driven state prediction algorithm with uncertainty representation and management capability is utilized.

B. Raw HI Extraction

In most industrial applications, an HI is firstly extracted from the available parameters by analyzing the relationship between the parameters and system degradation. In this paper, we denote the system health state in the ith cycle by $U_i, i = 1, 2, \ldots, n$, and the raw extracted HI series corresponding to U_i by $\boldsymbol{X}_i = (x_{i1}, x_{i2}, \ldots, x_{\mathrm{iq}})$.

C. Box–Cox Transformation

Ideally, the raw HI can reflect the degradation process of the system, but its performance may not be fully satisfactory due to the complexity of underlying degradation process. To reduce the possible deviation of the HI from the degradation process, we apply the Box–Cox transformation to enhance the linear relationship between the HI and the system health status.

1) Box–Cox Transformation [29], [30]: Consider a linear regression model

$$\begin{aligned} &y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots, +\beta_q x_{\mathrm{iq}} + \varepsilon_i \\ &\varepsilon_i \sim N\left(0, \sigma^2\right), i = 1, 2, \ldots, n \end{aligned} \tag{1}$$

where q is the number of independent variables, n is the sample size, $\beta_0, \beta_1, \ldots, \beta_q$ are the coefficients, and $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ are random errors that are independent and identically distributed and follow the normal distribution with zero mean and variance σ^2.

In practice, to provide an adequate statistical fitting to collected data using the regression model, the Box–Cox transformation [29] can be implemented. Mathematically, the Box–Cox transformation can be expressed as

$$x_t^{(\lambda)} = h(x, \lambda) = \begin{cases} \frac{x_t^{\lambda} - 1}{z}, & \lambda \neq 0 \\ \ln x_t, & \lambda = 0 \end{cases} \tag{2}$$

and the corresponding inverse transformation is

$$x_t = g\left(x_t^{(\lambda)}, \lambda\right) = \begin{cases} \left(1 + \lambda x_t^{(\lambda)}\right)^{1/\lambda}, & \lambda \neq 0 \\ \exp\left(x_t^{(\lambda)}\right), & \lambda = 0 \end{cases} \tag{3}$$

where λ is the transformation parameter to be determined. The Box–Cox transformation is valid when $x > 0$. For a case with $x < 0$, a shifted power transformation can be used [30].

In this paper, the Box–Cox transformation is applied to the exogenous variables [31] as follows:

$$y_i = \beta_0 + \beta_1 x_{i1}^{(\lambda)} + \beta_2 x_{i2}^{(\lambda)} + \cdots + \beta_q x_{\mathrm{iq}}^{(\lambda)} + \varepsilon_i. \tag{4}$$

2) Parameter Identification for Box–Cox Transformation: Two main approaches can be used to determine λ: the maximum likelihood (ML) [29], [30] method and Bayesian method [31], [32]. In this paper, the ML method is considered.

Taking (4) as an example and assuming that $\varepsilon_i \sim N(0, \sigma^2)$, we have

$$y_i \sim N\left(\boldsymbol{X}_i^{(\lambda)} \boldsymbol{\beta}, \sigma^2\right) \tag{5}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_q)^T$ and $\boldsymbol{X}_i^{(\lambda)} = (1, x_{i1}^{(\lambda)}, \ldots, x_{\mathrm{iq}}^{(\lambda)})$. So, the joint probability density function of $\boldsymbol{Y} = (y_1, y_2, \ldots, y_n)^T$ is

$$f(\boldsymbol{Y}) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{\left(\boldsymbol{Y} - \boldsymbol{X}^{(\lambda)}\boldsymbol{\beta}\right)^T \left(\boldsymbol{Y} - \boldsymbol{X}^{(\lambda)}\boldsymbol{\beta}\right)}{2\sigma^2}\right\} \tag{6}$$

where $\boldsymbol{X}^{(\lambda)} = (\boldsymbol{J}, \boldsymbol{X}_1^{(\lambda)}, \boldsymbol{X}_2^{(\lambda)}, \ldots, \boldsymbol{X}_q^{(\lambda)})$, in which $\boldsymbol{J} = (1, \ldots, 1)^T$ is an n-dimensional vector and $\boldsymbol{X}_j^{(\lambda)} = (x_{j1}^{(\lambda)}, x_{j2}^{(\lambda)}, \ldots, x_{jn}^{(\lambda)})^T, j = 1, \ldots, q$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_q)^T$.

From (6), the log-likelihood function can be expressed as

$$\begin{aligned} L\left(\boldsymbol{\beta}, \sigma^2, \lambda | \boldsymbol{Y}, \boldsymbol{X}\right) = &-\frac{n}{2} ln 2\pi - \frac{n}{2} ln\sigma^2 \\ &- \frac{1}{2\sigma^2}\left(\boldsymbol{Y} - \boldsymbol{X}^{(\lambda)}\boldsymbol{\beta}\right)^T \left(\boldsymbol{Y} - \boldsymbol{X}^{(\lambda)}\boldsymbol{\beta}\right). \end{aligned} \tag{7}$$

Taking the partial derivatives of $L(\boldsymbol{\beta}, \sigma^2, \lambda | \boldsymbol{Y}, \boldsymbol{X})$ with respect to $\boldsymbol{\beta}$ and σ^2 and setting each of the resulting equations to zero yields

$$\hat{\boldsymbol{\beta}}(\lambda) = \left[\left(\boldsymbol{X}^{(\lambda)}\right)^T \left(\boldsymbol{X}^{(\lambda)}\right)\right]^{-1} \left(\boldsymbol{X}^{(\lambda)}\right)^T \boldsymbol{Y} \tag{8}$$

$$\hat{\sigma}^2(\lambda) = \frac{\left(Y - X^{(\lambda)}\beta\right)^T \left(Y - X^{(\lambda)}\beta\right)}{n}. \tag{9}$$

Substituting (9) into (7) yields the profile log-likelihood function as

$$L^*(\boldsymbol{\beta}, \lambda) = C - \frac{n}{2} ln \left\{ \frac{\left(\boldsymbol{Y} - \boldsymbol{X}^{(\lambda)}\boldsymbol{\beta}\right)^T \left(\boldsymbol{Y} - \boldsymbol{X}^{(\lambda)}\boldsymbol{\beta}\right)}{n} \right\}. \quad (10)$$

Let C_1 be the constant value and invariant to λ, then, we have

$$L^*(\boldsymbol{\beta}, \lambda) = C_1 - \frac{n}{2} ln \left\{ \left(\boldsymbol{Y} - \boldsymbol{X}^{(\lambda)}\boldsymbol{\beta}\right)^T \left(\boldsymbol{Y} - \boldsymbol{X}^{(\lambda)}\boldsymbol{\beta}\right) \right\}. \quad (11)$$

To maximize (11) with respect to $(\boldsymbol{\beta}, \lambda)$, it is equivalent to minimizing the value of

$$S(\lambda) = \left(\boldsymbol{e}^*\right)^T \left(\boldsymbol{e}^*\right) \quad (12)$$

where $\boldsymbol{e}^* = (\boldsymbol{Y} - \boldsymbol{X}^{(\lambda)}\boldsymbol{\beta})$. This can be implemented using the Newton's method to first find the value of $\hat{\lambda}$ that minimizes $S(\lambda)$ followed by the determination of the corresponding $\hat{\boldsymbol{\beta}}$ according to (8).

D. Correlation Analysis

To verify the effectiveness of the Box–Cox transformation in enhancing the extracted HI, correlation analysis can be performed.

1) Pearson Correlation Analysis: The Pearson correlation analysis [33] can be applied to quantitatively verify the linear relationship between the extracted HI and actual health status. The Pearson correlation coefficient is computed as

$$r = \frac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2} \sqrt{\sum_{i=1}^{n} \left(Y_i - \bar{Y}\right)^2}} \quad (13)$$

which is between -1 and $+1$. When the correlation equals ± 1, the perfect linear relationship exists between variables X and Y, and when it equals 0, no correlation exists.

2) Spearman Rank Correlation Analysis: Alternatively, Spearman rank correlation analysis [33] can be implemented, which is a nonparametric way of evaluating the strict monotonic relationship between two variables. Denote the Spearman rank correlation coefficient as r_s. The variables X and Y are ranked in a descending order and then the Pearson correlation coefficient is calculated based on the ranks. The value of the Spearman rank correlation coefficients r_s is also between -1 and $+1$.

E. RUL Estimation Algorithm

In the framework, we apply the RVM algorithm in RUL estimation [15], [16], [34]. Given data set $\{z_n, t_n\}_{n=1}^N$, $z_n \in R^q$, $t_n \in R$, the output of the RVM model is

$$z(\boldsymbol{z}, \omega) = \sum_{i=1}^{N} \omega_i K(\boldsymbol{z}, z_i) + \omega_0 \quad (14)$$

where $K(\boldsymbol{z}, z_i)$ is a kernel function, ω_i is the weight of the model, and N is the sample size.

We follow the standard probabilistic formulation and assume that the targets are samples from the model contaminated by additive noise:

$$T_n = z(\boldsymbol{z}_n, \boldsymbol{\omega}) + \xi_n \quad (15)$$

where ξ_n are independent errors following $\xi_n \sim N(0, \sigma^2)$. As a result, the distribution of T_n is $p(t_n/z) \sim N(z(\boldsymbol{z}_n), \sigma^2)$, which is the normal distribution with mean $z(\boldsymbol{z}_n)$ and variance σ^2. Assuming t_n are independent, the likelihood of the complete data set can be written as

$$p\left(\boldsymbol{t}|\boldsymbol{\omega}, \sigma^2\right) = \left(2\pi\sigma^2\right)^{-N/2} \exp\left\{ -\frac{\|\boldsymbol{t} - \boldsymbol{\Phi}\boldsymbol{\omega}\|^2}{2\sigma^2} \right\} \quad (16)$$

where $\boldsymbol{t} = (t_1, \ldots, t_N)^T$, $\boldsymbol{\omega} = (\omega_0, \ldots, \omega_N)^T$, and $\boldsymbol{\Phi}$ is an $N \times (N+1)$ design matrix with $\boldsymbol{\Phi} = [\phi(\boldsymbol{z}_1), \phi(\boldsymbol{z}_2), \ldots, \phi(\boldsymbol{z}_n)]^T$ and $\phi(\boldsymbol{z}_n) = [1, K(\boldsymbol{z}_n, \boldsymbol{z}_1), \ldots, K(\boldsymbol{z}_n, \boldsymbol{z}_n)]$.

Because the number of parameters in the model is the same as the size of training data, the ML estimates of $\boldsymbol{\omega}$ and σ^2 from (16) will lead to severe over-fitting. To avoid this problem, we adopt a Bayesian perspective and superimpose constraints on the parameters by explicitly defining a prior zero-mean Gaussian distribution over $\boldsymbol{\omega}$

$$p(\boldsymbol{\omega}/\boldsymbol{\alpha}) = \prod_{i=0}^{N} \frac{\alpha_i}{\sqrt{2\pi}} \exp\left(\frac{\omega_i^2 \alpha_i}{2}\right) \quad (17)$$

with $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \ldots, \alpha_N\}$ being a vector of $N+1$ hyperparameters.

The posterior distribution over the weights is thus given by

$$p\left(\boldsymbol{\omega}/\boldsymbol{t}, \boldsymbol{\alpha}, \sigma^2\right) = (2\pi)^{-\frac{N+1}{2}} \left|\sum\right|^{-1/2} \exp\left\{ -\frac{(\boldsymbol{\omega} - \boldsymbol{\mu})^T \sum^{-1} (\boldsymbol{\omega} - \boldsymbol{\mu})}{2} \right\} \quad (18)$$

where the posterior covariance matrix and mean vector are

$$\sum = \left(\sigma^{-2}\boldsymbol{\Phi}^T\boldsymbol{\Phi} + \boldsymbol{A}\right)^{-1} \quad (19)$$

$$\boldsymbol{\mu} = \sigma^{-2} \sum \boldsymbol{\Phi}^T \boldsymbol{t} \quad (20)$$

where $\boldsymbol{A} = \text{diag}(\alpha_0, \alpha_1, \ldots, \alpha_N)$.

Because the posterior distributions of many weights are sharply (indeed infinitely) peaked around zero, the sparsity is achieved. These vectors associated with remaining nonzero weights relevance vector are called relevance vectors.

In (18), the distribution of the output can be obtained by

$$p\left(\boldsymbol{t}|\boldsymbol{\alpha}, \boldsymbol{\sigma}^2\right) = \int p\left(\boldsymbol{t}|\boldsymbol{\omega}, \boldsymbol{\sigma}^2\right) \cdot p(\boldsymbol{\omega}|\boldsymbol{\alpha}) d\boldsymbol{\omega}. \quad (21)$$

So, the distribution of the hyperparameters is

$$p\left(\boldsymbol{t}|\boldsymbol{\alpha}, \boldsymbol{\sigma}^2\right) = (2\pi)^{-N/2} \left|\boldsymbol{\sigma}^2\boldsymbol{I} + \boldsymbol{\Phi}\boldsymbol{A}^{-1}\boldsymbol{\Phi}^T\right|^{-1/2} \times \exp\left\{ -\frac{1}{2}\boldsymbol{t}^T \left(\boldsymbol{\sigma}^2\boldsymbol{I} + \boldsymbol{\Phi}\boldsymbol{A}^{-1}\boldsymbol{\Phi}^T\right)^{-1} \boldsymbol{t} \right\}. \quad (22)$$

It is worth pointing out that the values of α and σ^2 that maximize (22) cannot be obtained analytically. Here, we summarize the formulas used in iterative calculation steps.

For $\boldsymbol{\alpha}$, after taking the derivative of (22), equating it to zero and following the approach in [16], we have:

$$\alpha_i^{\text{new}} = \frac{\gamma_i}{\mu_i^2} \quad (23)$$

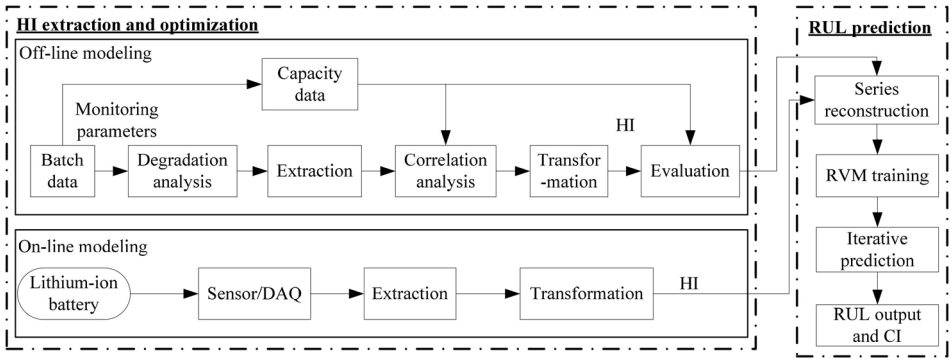

Fig. 2. Framework for HI extraction and transformation and RUL prediction for lithium-ion batteries.

where μ_i is the ith posterior mean weight from (20), and γ_i is given by

$$\gamma_i = 1 - \alpha_i \sum_{ii} \tag{24}$$

where $\sum_{ii}$ is the ith diagonal element of the posterior weight covariance from (19) computed with the current values of α and σ^2. To update the variance σ^2 of the noise, differentiation leads to the reestimate

$$\left(\sigma^2\right)^{\text{new}} = \frac{\|\boldsymbol{t} - \boldsymbol{\Phi}\boldsymbol{\mu}\|}{N - \sum_i \gamma_i} \tag{25}$$

where N is the number of data samples.

Given a new test point z_*, predictions are made for the corresponding target t_*. Specifically, the distribution of t_* is $p(t_*|t) \sim N(\mu_*, \sigma_*^2)$ with mean μ_* and variance σ_*^2, for which

$$\mu_* = \boldsymbol{\mu}^T \phi(z_*) \tag{26}$$

$$\sigma_*^2 = \sigma_{MP}^2 + \phi\,(z_*)^T \sum \phi\,(z_*) \tag{27}$$

where σ_*^2 is attributed to the uncertainty in predicting the weights and the estimated variance σ_{MP}^2 of the noise obtained when iterations end in (24) and (25).

The optimized dynamic gray RVM (DGRVM) [35] algorithm is adopted to RUL estimation. Due to the lack of long-term prediction capability of the RVM algorithm, here, we combine the gray model (GM) and RVM to improve the RUL estimation accuracy. The DGRVM algorithm uses the predicted value as the inputs of the RVM algorithm to exert the strong trend estimation ability of the GM algorithm. The DGRVM algorithm fully considers the degradation features, and adopts the combined prediction model and dynamic training algorithm to solve problems with small data samples and low accuracy of multistep prediction. For more detailed information about DGRVM model, readers are referred to [30].

III. Application to Lithium-Ion Battery RUL Estimation

In this section, we apply the proposed framework (see Fig. 2) to extract and optimize the HI obtained from the monitoring parameters for RUL estimation of lithium-ion batteries.

Part I: HI Extraction and Optimization: During off-line modeling, the degradation trend of the monitoring parameters, including the charging voltage, charging current and discharging voltage, discharging current and temperature, etc., is analyzed first. afterward, HIs are extracted and their effectiveness in quantifying battery degradation is verified by correlation analysis. If necessary, the Box–Cox transformation is applied to enhance the extracted HI so that the correlation between of the HI and the battery capacity is improved. As a result, an indirect optimized HI can be obtained without relying on the battery capacity. For on-line applications, the procedure for the monitoring of parameters is implemented based on the prior knowledge obtained from off-line modeling.

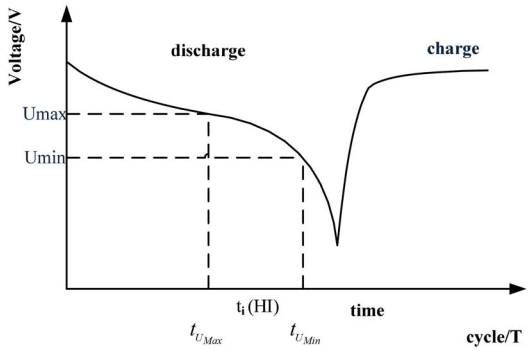

Fig. 3. HI extraction with DVD_ETI.

Part II: Battery RUL Estimation: The extracted and optimized HI is adopted along with the RVM algorithm in RUL estimation for lithium-ion battery.

Fig. 3 shows the flow chart for RUL estimation using the extracted and optimized HI. In this procedure, the correlation analysis and HI transformation techniques are the same as the ones introduced in Section II. In the next, we will mainly discuss the specific HI extraction process and other feature conversion technique related to RUL estimation.

It is worth pointing out that in real-world applications the operating profiles may be different for different batteries. For certain applications, the degradation modes can also be classified into several types. Thus, some off-line experimental data and training parameters can be used as a starting point for on-line applications. In fact, more prior knowledge obtained from off-line experiments can help us optimize and improve the on-line performance of the proposed method.

A. Raw HI Extraction With Monitoring Parameters

Degradation modeling and RUL estimation with monitoring parameters (i.e., charging and discharging voltage) is one of the most important industrial requirements. Considering lithium-ion batteries of new cell phones/laptops, the operating time will be maximal after a full charge at the very beginning. However, the operating time (discharging time period) after subsequent full charges becomes shorter and shorter. This is because the maximum charging capacity fades with the cyclic charge/discharge. Meanwhile, there is a certain relationship between the discharging time period and the capacity of a lithium-ion battery.

The phenomenon and related analysis indicate that the discharging voltage difference of equal time interval (DVD_ETI) can be used as an HI to measure the capacity degradation in each charging and discharging cycle. The HI extraction for a charging and discharging cycle is shown in Fig. 3 [20].

Specifically, The DVD_ETI parameter is defined as the DVD in a certain equal-length time interval for discharging. The degradation indicated by the DVD_ETI parameter is similar to that of battery capacity [20]. In particular, the DVD in the equal-length time interval in the ith cycle is

$$V_{i\_\mathrm{DVD\_ETI}} = U_{t_{\min}} - U_{t_{\max}}, \qquad i = 1, 2, \ldots, n \tag{28}$$

and the raw HI series can be expressed as

$$V_{\mathrm{DVD\_ETI}} = \left\{V_{1\_\mathrm{DVD\_ETI}}, V_{2\_\mathrm{DVD\_ETI}}, \ldots, V_{n\_\mathrm{DVD\_ETI}}\right\}. \tag{29}$$

We develop a three-step raw HI extraction procedure as follows.

Step 1) Extract the monitoring voltage, current, and cycle index in each charging/discharging cycle under the constant-voltage and restricted-current mode.

Step 2) Define the ETI ($t_{\max}$ and $t_{\min}$) and extract the health indicating time series. Here, $t_{\max}/t_{\min}$ is the maximum/minimum time value used as the starting/ending signal to count the amplitude of DVD. Thus, the discharging voltage interval corresponding to ETI between $t_{\max}$ and $t_{\min}$ can be obtained using (28).

Step 3) Convert the discharging voltage interval corresponding to the $t_{\max} - t_{\min,}$ to obtain the DVD_ETI series in each cycle as shown in (29).

B. Box–Cox Transformation, Normalizations, and Correlation Analysis

After the raw HI extraction, to further improve the linearity of the extracted HI and actual health status (i.e., battery capacity), the Box–Cox transformation is applied as introduced in Section II-C.

With the identified parameter λ for the Box–Cox transformation, the linear model for the capacity and DVD_ETI (raw HI) is

$$C = \beta_0 + \beta_1 V^{\lambda}_{\mathrm{DVD\_ETI}}. \tag{30}$$

Denote $U_{\mathrm{BCT\_DVD\_ETI}} = V^{\lambda}_{\mathrm{DVD\_ETI}}$ (transformed and optimized HI), and (30) becomes

$$C = \beta_0 + \beta_1 U_{\mathrm{BCT\_DVD\_ETI}}. \tag{31}$$

Usually, normalization is implemented to lithium-ion battery capacity before being adopted as the HI for RUL estimation, that is

$$\mathrm{HI}_c(i) = \frac{C_i - C_{\min}}{C_{\max} - C_{\min}}, \qquad i = 1 \cdots n \tag{32}$$

where i represents the corresponding cycle number of the battery. After substituting (31) into (32), we have

$$\begin{aligned}\mathrm{HI}_{\mathrm{BCT\_DVD\_ETI}}(i) &= \frac{C_i - C_{\min}}{C_{\max} - C_{\min}} \\ &= \begin{cases} \frac{(\beta_0+\beta_1 U_i)-(\beta_0+\beta_1 U_{\min})}{(\beta_0+\beta_1 U_{\max})-(\beta_0+\beta_1 U_{\min})}, & r_s > 0 \\ \frac{(\beta_0+\beta_1 U_{\max})-(\beta_0+\beta_1 U_i)}{(\beta_0+\beta_1 U_{\max})-(\beta_0+\beta_1 U_{\min})}, & r_s < 0. \end{cases}\end{aligned} \tag{33}$$

Simplifying (33) yields

$$\mathrm{HI}_{\mathrm{BCT\_DVD\_ETI}}(i) = \begin{cases} \frac{U_i - U_{\min}}{U_{\max} - U_{\min}}, & r_s > 0 \\ \frac{U_{\max} - U_i}{U_{\max} - U_{\min}}, & r_s < 0. \end{cases} \tag{34}$$

With (34), we can obtain the normalized $\mathrm{HI}_{\mathrm{BCT\_DVD\_ETI}}$ of $U_{\mathrm{BCT\_DVD\_ETI}}$ for RUL estimation.

C. RUL Estimation and Cycle Life Threshold Conversion

In practice, we usually define a certain level (i.e., 20% or 30%) fade of the rated capacity as the failure threshold. For the extracted and optimized HI of DVD_ETI, the flexibility of defining the corresponding threshold for this HI parameter is also quite important. To determine the threshold, the inverse Box–Cox transformation can be performed. Let the threshold for the capacity be C_p, the corresponding threshold for the raw HI be $V_{\mathrm{DVD\_ETI}\_p}$, and the one for the transformed HI be $U_{\mathrm{BCT\_DVD\_ETI}\_p}$, respectively. We can calculate the $V_{\mathrm{DVD\_ETI}\_p}$ and $U_{\mathrm{BCT\_DVD\_ETI}\_p}$ as follows.

From (30) and (31), we have

$$U_{\mathrm{BCT\_DVD\_ET}} = \frac{C - \beta_0}{\beta_1} \tag{35}$$

$$V_{\mathrm{DVD\_ETI}} = \begin{cases} \left(\frac{(C-\beta_0)}{\beta_1}\right)^{1/\lambda}, & \lambda \neq 0 \\ \exp\left(\frac{(C-\beta_0)}{\beta_1}\right), & \lambda = 0 \end{cases} \tag{36}$$

where parameters β_0 and β_1 can be identified as in Section II-C. By substituting C_p into (35) and (36), we have

$$U_{\mathrm{BCT\_DVD\_ETI}\_p} = \frac{C_p - \beta_0}{\beta_1} \tag{37}$$

$$V_{\mathrm{DVD\_ETI}\_p} = V_{\mathrm{DVD\_ETI}} = \begin{cases} \left(\frac{(C_p-\beta_0)}{\beta_1}\right)^{1/\lambda}, & \lambda \neq 0 \\ \exp\left(\frac{(C_p-\beta_0)}{\beta_1}\right), & \lambda = 0. \end{cases} \tag{38}$$

Then, the corresponding threshold for normalized transformed HI can be obtained by substituting (37) into (34) as

$$\mathrm{HI}_{\mathrm{BCT\_DVD\_ETI}\_p} = \begin{cases} \frac{U_{\mathrm{BCT\_DVD\_ETI}\_p} - U_{\min}}{U_{\max} - U_{\min}}, & r_s > 0 \\ \frac{U_{\max} - U_{\mathrm{BCT\_DVD\_ETI}\_p}}{U_{\max} - U_{\min}}, & r_s < 0. \end{cases} \tag{39}$$

TABLE I
TIME INTERVAL SETTING (UNIT: SECONDS)

t_1=500 (V1)		t_2=1500 (V2)		t_3=2300 (V3)	
min	max	min	Max	min	max
0	500	0	1500	0	2300

With the extracted and optimized HI, as well as the threshold for the novel HI, the DGRVM can be implemented in RUL estimation for lithium-ion battery.

IV. DEMONSTRATION OF LITHIUM-ION BATTERY HI EVALUATION

A. Two Data Sets of Lithium-Ion Batteries

Two lithium-ion battery data sets are studied in this paper to verify the proposed framework. One is obtained from the data repository of the NASA Ames Prognostics Center of Excellence. Another is from the Center for Advanced Life Cycle Engineering (CALCE) at University of Maryland.

The first data set was collected from a battery prognostics test bed at NASA 18 650 comprising commercially available lithium-ion rechargeable batteries. The test bed contains a power supply, programmable dc electronic load, voltmeter, thermocouple sensor, environmental chamber, EIS, PCI extensions for Instrumentation chassis based on data acquisition, and experiment control conditions [36]. The lithium-ion batteries were run through three different operational profiles (charge, discharge, and impedance) at room temperature. Charging was carried out in a CC mode at 1.5 A until the battery voltage reached 4.2 V and then continued in a CV mode until the charge current dropped to 20 mA. Discharge was carried out at a CC level of 2 A until the battery voltage fell to 2.7, 2.5, 2.2, and 2.5 V for batteries nos. 5–7 and 18, respectively. Repeated charge and discharge cycles resulted in accelerated aging of the batteries. The experiments were stopped when the batteries reached the end-of-life with a 30% fade in rated capacity (from 2 to 1.4 Ahr).

The second data set was obtained to discover battery capacity degradation. The cyclic test was implemented with the Arbin BT2000 battery testing system under room temperature. The batteries with 1.1 Ah rated capacity were tested in the experiment with the discharging current 0.55 A (it indicates that the discharging speeds is 0.5 C) [1], [22], [24], [37].

B. Correlation Analysis and Evaluation

We first analyze the correlation of the raw HI and transformed HI with battery capacity to show the necessity of applying Box–Cox transformation. Here, we only apply NASA battery no. B5 to briefly describe the experimental results and related analysis.

1) Qualitative Analysis: To consider the influences of the time intervals on the extracted HI, we extract three types of DVD_ETI series (V1, V2, and V3) corresponding to different time intervals. The time intervals setting are shown in Table I, and the extracted DVD_ETI series collected under different time intervals are shown in Fig. 4.

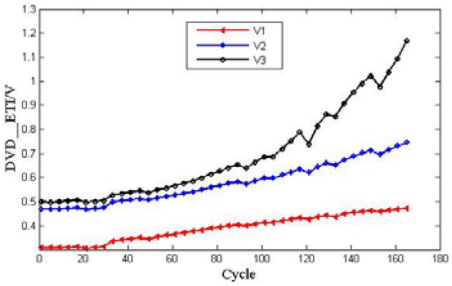

Fig. 4. DVD_ETI with different time intervals.

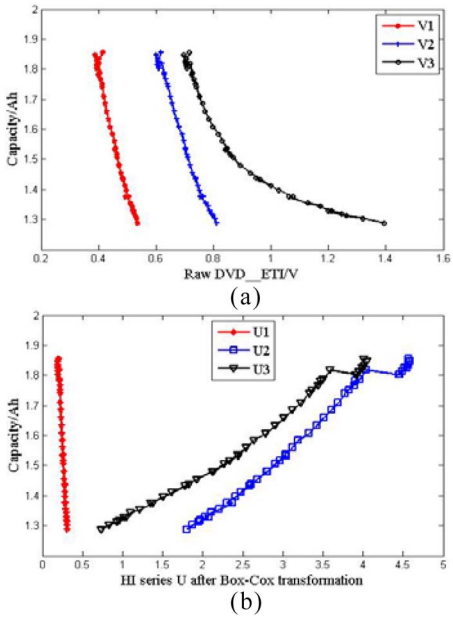

Fig. 5. Correlation relationship of the DVD_ETI series and capacity for NASA battery no. B5. (a) Scatter plot of raw DVD_ETI series and capacity. (b) Scatter plot of transformed HI $U_{\text{BCT\_DVD\_ETI}}$ and capacity.

Scatter plots are first used to present the correlation of the three groups of DVD_ETI series with capacity series before and after the Box–Cox transformation. Fig. 5 shows the raw HI series $V_{\text{DVD\_ETI}}$ and transformed HI $U_{\text{BCT\_DVD\_ETI}}$. We can find that extinguishing monotonic relationship exists between the capacity and raw DVD_ETI series. However, their linear relationships are not stable (the linearity between $V1_{\text{DVD\_ETI}}$, $V2_{\text{DVD\_ETI}}$ and capacity is superior to the linearity between $V3_{\text{DVD\_ETI}}$ and capacity). On the other hand, the monotonic linearity relationship between the transformed parameters $U1_{\text{BCT\_DVD\_ETI}}$, $U2_{\text{BCT\_DVD\_ETI}}$, $U3_{\text{BCT\_DVD\_ETI}}$, and capacity are improved, and the correlation appears to be higher.

2) Quantitative Analysis: To quantitatively evaluate the improvement in correlation by performing the Box–Cox transformation, the Pearson correlation and the Spearman rank correlation of multiple DVD_ETI series and capacity series are calculated. Table II shows the results for NASA battery no. B5.

From Table II, we can see that the correlation coefficient is close to -1, which indicates a strong negative correlation between the capacity and $V1_{\text{DVD\_ETI}}$, $V2_{\text{DVD\_ETI}}$,

TABLE II
RESULTS OF CORRELATION ANALYSIS (NASA BATTERY NO. B5)

After Box-Cox transformation				Before Box-Cox transformation		
factors	λ	Pearson correlation (r)	Spearman rank Correlation (r_s)	factors	Pearson correlation (r)	Spearman rank Correlation (r_s)
$U1$	1	-0.992	**-0.989**	$V1_{DVD\_ETI}$	-0.992	**-0.989**
$U2$	-2	0.990	**0.996**	$V2_{DVD\_ETI}$	-0.980	**-0.996**
$U3$	-2	0.994	**0.996**	$V3_{DVD\_ETI}$	-0.930	**-0.996**

TABLE III
RELATIONSHIP OF PEARSON CORRELATION COEFFICIENT AND PARAMETER λ

Transformation parameter		Pearson correlation coefficients		
λ	Model	$U1$	$U2$	$U3$
-5	$U=V^{-5}$	0.976	0.976	0.960
-4	$U= V^{-4}$	0.978	0.978	0.972
-3	$U = V^{-3}$	0.981	0.981	0.985
-2	$U = V^{-2}$	0.988	**0.990**	**0.996**
-1	$U = V^{-1}$	0.967	0.978	0.977
0	$U =\ln V$	0.976	0.978	0.956
0.5	$U = V^{0.5}$	0.981	0.981	0.944
1	$U = V^{1}$	**0.992**	0.980	0.930
2	$U = V^{2}$	0.967	0.970	0.896
3	$U = V^{3}$	0.941	0.946	0.858
4	$U = V^{4}$	0.913	0.938	0.817
5	$U = V^{5}$	0.904	0.920	0.775

and $V3_{DVD\_ETI}$. Moreover, the correlation coefficients $r1$, $r2$, and $r3$ are different for the same battery. With the Box–Cox transformation, the correlation coefficients r_s1, r_s2, and r_s3 are much closer to 1 or −1, and the absolute values of these coefficients are almost equal, which indicates that the Box-Cox transformation improves the stability and robustness of the extracted HIs. This will benefit RUL estimation with these optimized HI.

In Table II, the values of transformation parameter λ are different for different DVD_ETI series. The absolute values of the Spearman rank correlation coefficients r_s remain the same after the Box–Cox transformation. For the Pearson correlation coefficients r, the one for $U1$ keeps unchanged because of the high linear relationship between $V1_{DVD\_ETI}$ and capacity. As a result, the corresponding transformation parameter is $\lambda = 1$, which means the data is not changed.

With the Pearson correlation analysis, the extracted DVD_ETI, and battery capacity is approximately linear, but, with different extraction methods, the correlation degrees are different. The Spearman correlation analysis shows that the relationship of DVD_ETI and capacity is not strictly monotonic, and the monotonic features are different for the extracted parameters that are greatly influenced by parameters setting (see Fig. 6). Considering these two factors, we discuss two aspects in this paper; one is to improve the linearity using the Box–Cox transformation for variables, and the other aspect is to use a fusion method, to some extent, to improve the accuracy of the capacity degradation modeling.

In Table III, with the parameters identification approach addressed in Section II-C, the highest Pearson correlation coefficients are 0.992, 0.990, and 0.996 when parameter λ is set to be 1, and −2, and −2, respectively (corresponding to three group of DVD_ETI series). The results show that the

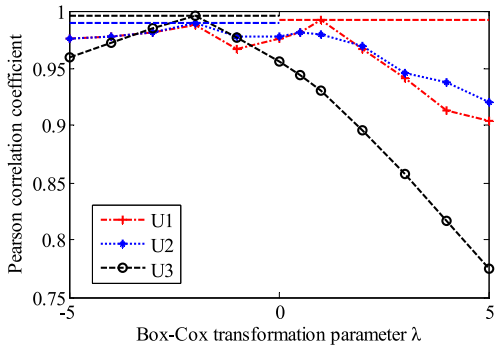

Fig. 6. Changes of the Pearson correlation coefficients with parameter λ.

TABLE IV
TIME INTERVAL SETTINGS FOR NASA BATTERIES

Index	t_1		t_2		t_3	
	min	*max*	*min*	*max*	*min*	*Max*
B5	0	500	0	1500	0	2300
B6	0	400	0	1200	0	2000
B7	0	500	0	1500	0	2000
B18	0	500	0	1500	0	2400

transformed feature parameter and the capacity have a strong linear relationship, which illustrate the effectiveness of the proposed method.

C. HI Performance Evaluation

1) Evaluation of NASA Battery: Using (28), we extract feature DVD_ETI of four lithium-ion batteries. The experimental settings for NASA battery nos. B5–B7 and B18 are shown in Table IV.

Here, the root mean squared error (RMSE) and the fitness degree ($0 < R^2 < 1$) are used in evaluating the performance of the extracted HI

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(h-\hat{h}\right)^2}{n}} \tag{40}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(h-\hat{h}\right)^2}{\sum_{i=1}^{n}\left(h-\bar{h}\right)^2} \tag{41}$$

where h is the actual HI (as the battery capacity), $\hat{h}$ is the estimated HI with transformation, and $\bar{h}$ is the actual mean value of h, and n is the sample size.

The detailed experimental results are shown in Fig. 7 and Table V. In Fig. 7, "HIx" represents the normalized HI corresponding to the normalized capacity, and

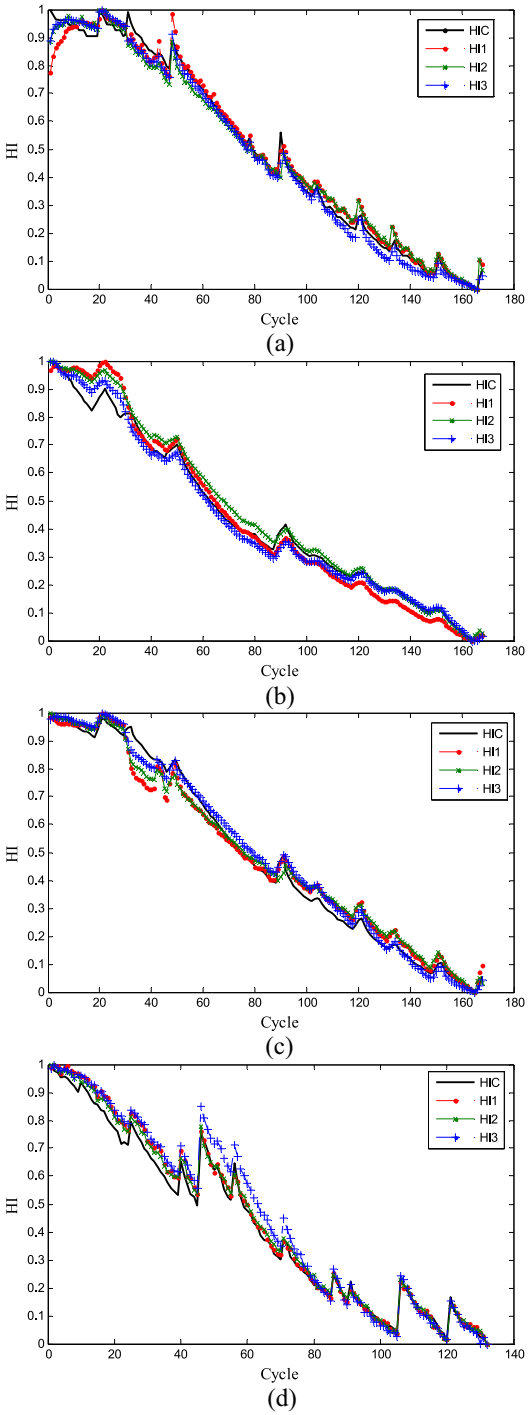

Fig. 7. HI for various lithium-ion batteries. HI for lithium-ion battery with nos. (a) B5, (b) B6, (c) B7, and (d) B18.

HI1, HI2, and HI3 mean the extracted HIs with the proposed method using three different time intervals. As shown in the figure, the three HI factors are very close to the actual capacity HI.

As shown in Table V, for the four lithium-ion batteries under three different time interval settings, 12 groups of extracted HI factors are obtained and compared with the corresponding capacity HIs. We can see that the overall R^2 values are between 0.9670 and 0.9931, which indicate adequate statistical fit.

TABLE V
HI EVALUATED RESULTS OF FOUR NASA LITHIUM-ION BATTERIES

Index	λ	HI	R^2	*RMSE*
B5	1	HI1	0.9883	0.0361
	-2	HI2	0.9881	0.0364
	-2	HI3	0.9931	0.0278
B6	-1	HI1	0.9748	0.0457
	1	HI2	0.9806	0.0401
	-2	HI3	0.9910	0.0273
B7	1	HI1	0.9816	0.0447
	1	HI2	0.9840	0.0416
	-3	HI3	0.9925	0.0286
B18	-1	HI1	0.9884	0.0323
	-1	HI2	0.9912	0.0281
	-3	HI3	0.9670	0.0545

TABLE VI
TIME INTERVAL SETTINGS FOR CALCE BATTERIES

Index	t_1		t_2		t_3	
	min	*max*	*min*	*max*	*min*	*Max*
CS2-35	0	750	0	1500	0	2250
CS2-36	0	630	0	1260	0	1890
CS2-37	0	870	0	1740	0	2610

TABLE VII
HI EVALUATED RESULTS OF THREE CALCE LITHIUM-ION BATTERIES

Index	λ	HI	R^2	*RMSE*
CS2-35	-3	HI1	0.9698	0.0357
	-3	HI2	0.9780	0.0304
	-5	HI3	0.9858	0.0245
CS2-36	-2	HI1	0.9798	0.0335
	-2	HI2	0.9743	0.0377
	-4	HI3	0.9950	0.0166
CS2-37	-4	HI1	0.9675	0.0384
	-3	HI2	0.9701	0.0369
	-5	HI3	0.9626	0.0412

Moreover, the RMSE values are between 0.0278 and 0.0545. Both criteria indicate that the proposed method can provide an appropriate HI and improve the similarity between the HI and capacity. Especially, the R^2 values are all close to 1. Because the capacity can represent the actual battery degradation, the interested degradation process can be effectively modeled with this extracted HI. For each battery, the different HIs corresponding to different time intervals are in the same order of magnitude with less difference, indicating that the proposed method is robust against different settings for sample time interval.

2) CALCE Battery Evaluation: To further illustrate the effectiveness of the proposed method, we study the other experimental data set provided by the CALCE. Because of its similarity with the NASA battery data set, we will only briefly discuss the related results, and the correlation analysis and the analysis on the influence of parameters will not be repeated. The parameter settings are presented in Table VI, and the experimental results are shown in Table VII.

From Table VII, we can see that, comparing to the capacity HI, the overall fitting degrees R^2 are between 0.9626 and 0.9950. The RMSE are between 0.0166 and

TABLE VIII
FAILURE THRESHOLD AND ACTUAL LIFE CYCLE FOR NASA BATTERIES

Index	Failure threshold	Actual cycle life (cycles)
B5	1.38Ah	128
B6	1.38Ah	112
B7	1.42Ah	159
B18	1.38Ah	100

0.0412, which are lower than the values of NASA batteries, due to the slower degradation cycle life and lower failure threshold. Again, the proposed method can offer satisfied HI that is close to the actual capacity HI.

V. DEMONSTRATION OF RUL ESTIMATION

After the evaluation and verification of the extracted and optimized HI, we perform RUL estimation based on these two lithium-ion battery data sets.

The numerical experiments are conducted as follows.

1) Compare the RUL estimates obtained from the raw HI ($V_{\text{DVD\_ETI}}$) and the transformed HI, to test the coherence of the proposed HI framework.
2) Compare the RUL estimates from the transformed HI and battery capacity, to measure the equivalence of the proposed HI framework in estimating battery RUL.

A. NASA Battery RUL Estimation

For the four NASA lithium-ion batteries, the actual thresholds in terms of the capacity as well as the actual cycle life are shown in Table VIII.

In the experiments, we carry out RUL estimation with different starting points to fully verify the effectiveness of the proposed method. According to the size of the lithium-ion battery data sets, the RUL estimates are achieved from the starting points of 60th cycles and 80th cycles, respectively.

To evaluate the performance of RUL estimation, we use both the absolute error (AE) and the improved percentage of estimation accuracy (η_{AE}) defined by

$$\text{AE} = \left|R - \hat{R}\right| \tag{42}$$

$$\eta_{\text{AE}} = \frac{\text{AE}_2 - \text{AE}_1}{R} \tag{43}$$

where R is the actual RUL value, and $\hat{R}$ is the estimated RUL value, and AE_1 and AE_2 are the AEs of RUL estimates with the two approaches. If $\eta_{\text{AE}} > 0$, it means the second approach has a lower accuracy compared to the first approach; otherwise, the first approach is more accurate.

1) RUL Estimation With Raw DVD_ETI (V_{DVD_ETI}) and the Transformed HI ($U_{BCT\_DVD\_ETI}$): The experimental settings are as follows. The kernel function for the RVM algorithm is chosen as the Gaussian function, for which the kernel parameter is set to be 3. Small adjustment will be done for individual battery data set. The variance of noise is $\sigma^2 = var(y) \times 0.1$, where y is the target data for the training data set. The maximum number of iterative steps and the convergence condition are that iter $= 1000$ and $\tau_{\max} = 1e5$. In addition, the hyperparameter is set to be $\{\alpha_i = 1\}_{i=1}^{N}$ where N is the length of training data set. The comparison of the experimental results for NASA battery no. B5 is shown in Fig. 8.

TABLE IX
RUL ESTIMATION FOR NASA BATTERIES

Index	HI	Threshold	Actual RUL value	Predicted RUL value	95% Confidence intervals	AE	η_{AE}
B5	HI1	0.1827	68	75	[71, 81]	7	30.89%
	$V1_{DVD\_ETI}$	0.5018		96	[95, 97]	28	
	HI2	0.1911		75	[66, 80]	7	1.47%
	$V2_{DVD\_ETI}$	0.7603		60	[57, 66]	8	
	HI3	0.1297		75	[64, 81]	7	19.12%
	$V3_{DVD\_ETI}$	1.0658		88	[87, 89]	20	
B6	HI1	0.2823	52	60	[49, 68]	8	30.77%
	$V1_{DVD\_ETI}$	0.2668		32	[25, 33]	24	
	HI2	0.2861		56	[42, 64]	4	55.77%
	$V2_{DVD\_ETI}$	0.5215		29	[20, 32]	33	
	HI3	0.2349		54	[45, 62]	2	38.46%
	$V3_{DVD\_ETI}$	0.6941		30	[36, 38]	22	
B7	HI1	0.0388	99	105	[87, 119]	6	47.47%
	$V1_{DVD\_ETI}$	0.2801		46	[50, 59]	53	
	HI2	0.0480		105	[97, 117]	6	45.45%
	$V2_{DVD\_ETI}$	0.5356		48	[51, 54]	51	
	HI3	0.0228		74	[68, 81]	25	32.32%
	$V3_{DVD\_ETI}$	1.0031		156	[155, 157]	57	
B18	HI1	0.0775	40	49	[46, 54]	9	137.50%
	$V1_{DVD\_ETI}$	0.2798		104	[103, 105]	64	
	HI2	0.0835		46	[39, 59]	6	55.00%
	$V2_{DVD\_ETI}$	0.5467		68	[67, 69]	28	
	HI3	0.0625		50	[30, 56]	10	62.50%
	$V3_{DVD\_ETI}$	0.9766		75	[72, 77]	35	

In Fig. 8, the x-axis is the charge and discharge cycles, the y-axis is the HI (normalized transformed HI and raw HI). The figure shows the predicted HI values, and the upper limits and lower limits of the 95% confidence interval. By comparing Fig. 8(a)–(f), respectively, one can see that the predicted transformed HI is very close to the actual value near the failure threshold. Comparatively, the predicted value of the raw DVD_ETI series is not satisfactory.

Here, we will not show pictorial results but present more quantitative experimental results. Table IX gives the predicted RUL values along with the 95% confidence intervals, and the values of AE, and η_{AE} (here, we give one group of results with the prediction starting point of 60th cycle). By examining the results, we can see that the RUL estimate obtained from the optimized HI is quite close to the actual RUL value, and the AE is less than 10 in most cases. It is worth pointing out that only one group of predicted values shows decreased estimation accuracy (the AE is 25 for the HI3 of battery no. B7). This is caused by the use of the same parameters of RVM for all the batteries, which may not be suitable for battery no. B7. At the same time, its actual RUL is 99 cycles, so this result is still acceptable.

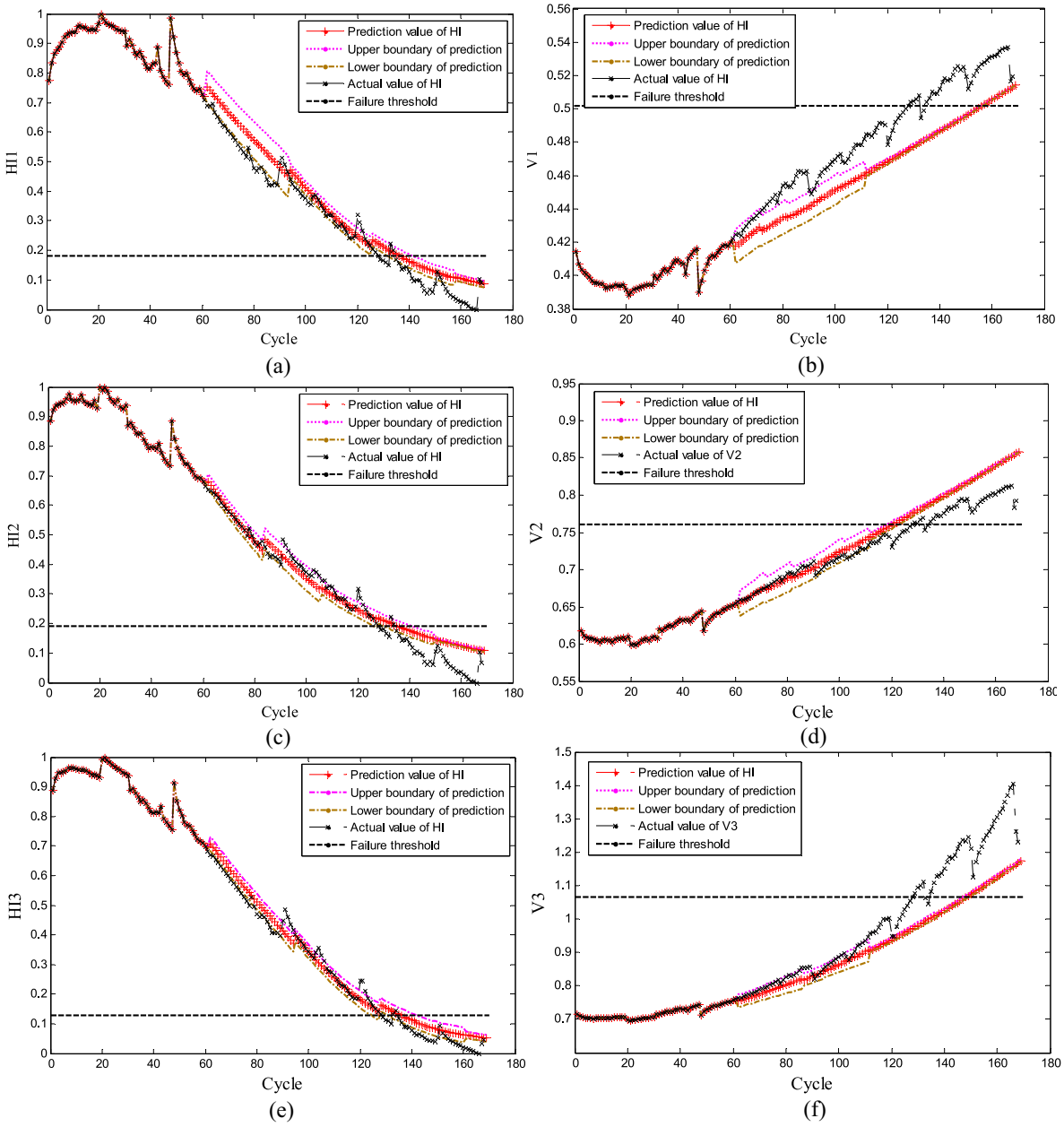

Fig. 8. RUL estimation for lithium-ion battery with no. B5 (starting at 60th cycle). RUL estimation with (a) HI1, (b) V1, (c) HI2, (d) V2, (e) HI3, and (f) V3.

Meanwhile, for the four battery samples using two predicted start points and under the same prediction conditions, RUL predicted value with the transformed and optimized HI is superior to the raw DVD_ETI series, in which the estimation accuracy is improved. The improvement of the estimation accuracy is between 19.1% and 137.5%, except that 1.47% for the RUL estimation of HI2. This further proves the effectiveness and reasonableness for the Box-Cox transformation and optimization on battery DVD_ETI series.

2) RUL Estimation With Normalized Transformation HI ($U_{BCT\_DVD\_ETI}$) and Capacity: The similar experiment is carried out to evaluate the equivalence of the new HI and the capacity for RUL estimation. The setting and flow of the experiment are the same as the experiments in Section V-A2. The experimental results are shown in Table X (here, we give one group of results with the prediction starting point of 80th cycle).

From Table X, for the capacity HI and the optimized HI, the predicted RUL value is very close to the actual RUL value, where the AE is less than 10 in most cases. The estimation accuracy is not so good for battery no. B7, which again is caused by the same reason described above.

Compared to the estimation results with capacity, the estimation accuracy is increased for five groups of results and decreased for the other five groups of results, and two groups of results are equal among the 12 groups of predicted results. The difference between the RUL estimates obtained with the HI and capacity may be caused by the local noise, but the results are quite close. In summary, the new HI proposed in

TABLE X
COMPARISON OF RUL ESTIMATION WITH THE CAPACITY AND DIFFERENT HIS (NASA)

Index	HI	Actual RUL value	Predicted RUL value	95% Confidence intervals	AE	η_{AE}
B5	HI1	48	46	[28, 57]	2	-2.09%
	HI2		50	[43, 58]	2	-2.09%
	HI3		45	[38, 53]	3	-4.17%
	capacity		49	[41, 54]	1	/
B6	HI1	32	35	[17, 52]	3	6.25%
	HI2		37	[19, 47]	5	0.00%
	HI3		31	[23, 49]	1	12.50%
	capacity		27	[15, 44]	5	/
B7	HI1	79	72	[61, 87]	7	46.84%
	HI2		73	[66, 84]	6	48.10%
	HI3		71	[64, 80]	8	45.57%
	capacity		35	[30, 38]	44	/
B18	HI1	20	20	[9, 40]	0	0.00%
	HI2		19	[9, 39]	1	-5.00%
	HI3		27	[21, 35]	7	-35.00%
	capacity		20	[10, 29]	0	/

TABLE XI
FAILURE THRESHOLD AND ACTUAL CYCLE LIFE FOR CALCE BATTERIES

Index	Failure threshold	Actual cycle life (cycles)
CS2-35	0.88Ah	606
CS2-36	0.88Ah	522
CS2-37	0.88Ah	595

this paper can be used to achieve RUL estimation as good as the one obtained using battery capacity.

B. CALCE Battery RUL Estimation

We also used another group of lithium-ion batteries from the CALCE to further illustrate the effectiveness of the proposed method. The threshold and the actual cycle life for the selected three batteries are shown in Table XI. The same experimental settings and procedures are utilized as the NASA batteries.

1) RUL Estimation With Raw DVD_ETI (V_{DVD_ETI}) and the Transformed HI ($U_{BCT\_DVD\_ETI}$): According to the length of CALCE battery data set, the predicted starting points for the battery CS2-35 and CS2-37 are the 234th cycle, and 193rd cycle for battery CS-36. The experimental results are shown as Table XII.

From Table XII, we can see that for the three batteries, the use of transformed HI results in more accurate RUL estimates compared to those obtained using the raw DVD_ETI series. The improvement of estimation accuracy is between 7.60% and 31.58%.

2) RUL Estimation With Normalized Transformation HI ($U_{BCT\_DVD\_ETI}$) and Capacity: The experimental results are shown in Table XIII (here, we give the results for the batteries with different prediction starting points: 343th cycle, 298th cycle, and 343th cycle). From Table XIII, we can see that, for the three batteries, in most cases, the AE of the RUL estimation are between 2 and 30. Compared with the method using the raw DVD_ETI series as the HI, the new HI can realize increasing for the RUL estimation for most of battery samples. Only one RUL result has a lower accuracy, for which the decrease is no more than 7.14%. The same results can be obtained for the NASA batteries.

TABLE XII
RUL ESTIMATION FOR CALCE BATTERIES

Index	HI	Actual RUL value	Predicted RUL value	95% Confidence intervals	AE	η_{AE}
CS2-35	HI1	372	364	[349, 376]	8	18.28%
	$V1_{DVD\_ETI}$		448	[445, 451]	76	
	HI2		391	[385, 397]	19	30.65%
	$V2_{DVD\_ETI}$		505	[502, 506]	133	
	HI3		352	[343, 361]	20	7.80%
	$V3_{DVD\_ETI}$		421	[418, 424]	49	
CS2-36	HI1	329	369	[354, 375]	8	31.00%
	$V1_{DVD\_ETI}$		471	[468, 474]	24	
	HI2		311	[306, 320]	4	7.60%
	$V2_{DVD\_ETI}$		372	[366, 375]	33	
	HI3		342	[333, 348]	2	9.42%
	$V3_{DVD\_ETI}$		285	[282, 288]	22	
CS2-37	HI1	361	384	[346, 397]	23	24.38%
	$V1_{DVD\_ETI}$		472	[466, 478]	111	
	HI2		355	[343, 364]	6	15.79%
	$V2_{DVD\_ETI}$		298	[289, 304]	63	
	HI3		376	[355, 382]	15	31.58%
	$V3_{DVD\_ETI}$		490	[487, 496]	129	

TABLE XIII
COMPARISON OF RUL ESTIMATION WITH CAPACITY AND HI (CALCE)

Index	HI	Predicted RUL value	95% Confidence intervals	AE	η_{AE}
CS-35	HI1	243	[210 282]	20	2.28%
	HI2	261	[219 303]	2	9.13%
	HI3	258	[213 297]	5	7.98%
	capacity	237	[207 267]	26	/
CS-36	HI1	228	[171 255]	4	8.72%
	HI2	239	[216 264]	15	6.70%
	HI3	228	[207 255]	4	8.72%
	capacity	254	[237 282]	30	/
CS-37	HI1	249	[234 267]	3	1.19%
	HI2	276	[237 288]	24	-7.14%
	HI3	273	[228 285]	21	-5.95%
	capacity	258	[252 285]	6	/

VI. CONCLUSION

In this paper, we propose a novel HI extraction and transformation framework for lithium-ion battery prognostics. With the Box–Cox transformation, the extracted HI with battery

DVD is optimized to enhance the performance in RUL estimation. Moreover, the RVM algorithm is utilized to realize indirect RUL estimation with the extracted and optimized HI. Satisfactory RUL estimates can be obtained with high prediction accuracy and stability. Experimental results show that the correlation of the extracted HI and battery capacity is increased after the Box–Cox transformation, and the proposed HI framework is quite effective and efficient in estimating the RUL of lithium-ion batteries in on-line applications.

It is worth pointing out that the DVD_ETI used in this paper is only suitable for cases where the current does not change too much in each discharge cycle. For a satellite system in which the payloads power consumption is relatively stable in-orbit, the DVD_ETI can be used as the HI. On the other hand, if the operating current changes significantly, we can use the discharging power difference in equal-length time interval. Moreover, since the charging current is relatively stable in most industrial applications, the charging voltage difference can also be considered as an alternative.

Certainly, other robust HIs should be studied for more complicated operating conditions. This is the main challenge for the RUL estimation of lithium-ion batteries working under time-varying conditions. Peterson *et al.* [38] developed a realistic operating schedule for the EVs considering different running modes. Millner [39] and Ning and Popov [40] focused on the aging models, considering the parameters impact involving charging rate, depth of discharge, end of charge voltage, etc. In addition, the degradation modes and cycle life features of batteries may be significantly different for different applications. To overcome the related big challenges, joint studies on dynamic operating conditions and accelerated aging profiles will provide attractive tools for data-driven RUL estimation of lithium-ion batteries.

In our future research, dynamic and adaptive transformation of HI will be studied. In addition, kernel-based and wavelet-based transformation, as well as correlation analysis in the transformed domain will be investigated to help identify the parameters of transformation. Especially, the influence of dynamic loads to the DVD_ETI parameters will be further discussed. As an application, lithium-ion battery degradation modeling and HI evaluation will also be considered under different operating environments.

REFERENCES

[1] W. He, N. Williard, M. Osterman, and M. Pecht, "Prognostics of lithium-ion batteries based on Dempster–Shafer theory and the Bayesian Monte Carlo method," *J. Power Sources*, vol. 196, no. 23, pp. 10314–10321, Dec. 2011.

[2] 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.

[3] K. Goebel, B. Saha, A. Saxena, J. R. Celaya, and J. P. Christophersen, "Prognostics in battery health management," *IEEE Instrum. Meas. Mag.*, vol. 11, no. 4, pp. 33–40, Aug. 2008.

[4] Y. Peng, D. Liu, and X. Peng, "A review: Prognostics and health management," *J. Electron. Meas. Instrum.*, vol. 24, no. 1, pp. 1–9, Mar. 2010.

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

[6] P. Wang, B. D. Youn, and C. Hu, "A generic probabilistic framework for structural health prognostics and uncertainty management," *Mech. Syst. Signal Process.*, vol. 28, pp. 622–637, Apr. 2012.

[7] K. T. Huynh, I. T. Castro, A. Barros, and B. Christophe, "On the use of mean residual life as a condition index for condition-based maintenance decision-making," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 44, no. 7, pp. 877–893, Jul. 2014.

[8] B. E. Olivares, M. A. C. Munoz, M. E. Orchard, and 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.*, vol. 62, no. 2, pp. 364–376, Feb. 2013.

[9] M. Dalal, J. Ma, and D. He, "Lithium-ion battery life prognostic health management system using particle filtering framework," *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.*, vol. 225, no. 1, pp. 81–90, Mar. 2011.

[10] J. Liu, W. Wang, and F. Ma, "A regularized auxiliary particle filtering approach for system state estimation and battery life prediction," *Smart Mater. Struct.*, vol. 20, no. 7, pp. 1–9, Jul. 2011.

[11] Q. Miao, L. Xie, H. Cui, W. Liang, and M. Pecht, "Remaining useful life prediction of lithium-ion battery with unscented particle filter technique," *Microelectron. Reliab.*, vol. 53, no. 6, pp. 805–810, Jun. 2013.

[12] J. Liu, A. Saxena, K. Goebel, B. Saha, and W. Wang, "An adaptive recurrent neural network for remaining useful life prediction of lithium-ion batteries," in *Proc. Annu. Conf. Prognostics Health Manage. Soc.*, Portland, OR, USA, 2010, pp. 1–9.

[13] B. Zhang *et al.*, "A probabilistic fault detection approach?: Application to bearing fault detection," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 2011–2018, May 2011.

[14] M. E. Orchard, P. Hevia-Koch, B. Zhang, and L. Tang, "Risk measures for particle-filtering-based state-of-charge prognosis in lithium-ion batteries," *IEEE Trans. Ind. Electron.*, vol. 60, no. 11, pp. 5260–5269, Nov. 2013.

[15] J. Zhou, D. Liu, Y. Peng, and X. Peng, "Combined sparse Bayesian learning strategy for remaining useful life forecasting of lithium-ion battery," in *Proc. 2nd Int. Conf. Instrum. Meas. Comput. Commun. Control (IMCCC)*, Harbin, China, 2012, pp. 457–461.

[16] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, Sep. 2001.

[17] B. Saha, K. Goebel, S. Poll, and J. Christophersen, "An integrated approach to battery health monitoring using Bayesian regression and state estimation," in *Proc. IEEE Autotestcon*, Baltimore, MD, USA, 2007, pp. 646–653.

[18] 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. Reliab.*, vol. 53, no. 6, pp. 832–839, Jun. 2013.

[19] J. Kozlowski, "Electrochemical cell prognostics using online impedance measurements and model-based data fusion techniques," in *Proc. IEEE Aerosp. Conf.*, vol. 7. Big Sky, MT, USA, 2003, pp. 3257–3270.

[20] D. Liu, H. Wang, Y. Peng, W. Xie, and H. Liao, "Satellite lithium-ion battery remaining cycle life prediction with novel indirect health indicator extraction," *Energies*, vol. 6, no. 8, pp. 3654–3668, 2013.

[21] J. Liu, W. Wang, F. Ma, Y. B. Yang, and C. S. Yang, "A data-model-fusion prognostic framework for dynamic system state forecasting," *Eng. Appl. Artif. Intell.*, vol. 25, no. 4, pp. 814–823, Jun. 2012.

[22] 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. Reliab.*, vol. 53, no. 6, pp. 811–820, Jun. 2013.

[23] C. Hu, B. D. Youn, P. Wang, and J. T. Yoon, "Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life," *Reliab. Eng. Syst. Safety*, vol. 103, pp. 120–135, Jul. 2012.

[24] 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 Sources*, vol. 239, pp. 253–264, Oct. 2013.

[25] J. Gomez, R. Nelson, E. E. Kalu, M. H. Weatherspoon, and J. P. Zheng, "Equivalent circuit model parameters of a high-power Li-ion battery: Thermal and state of charge effects," *J. Power Sources*, vol. 196, no. 10, pp. 4826–4831, May 2011.

[26] S. Lee, J. Kim, J. Lee, and B. H. Cho, "State-of-charge and capacity estimation of lithium-ion battery using a new open-circuit voltage versus state-of-charge," *J. Power Sources*, vol. 185, no. 2, pp. 1367–1373, Dec. 2008.

[27] 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.

[28] 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, 2013.
[29] G. E. P. Box and D. R. Cox, "An analysis of transformations," *J. R. Stat. Soc. B*, vol. 26, no. 2, pp. 211–252, 1964.
[30] A. R. M. Sakia, "The Box–Cox transformation technique: A review," *J. R. Stat. Soc. D,* vol. 41, no. 2, pp. 169–178, 1992.
[31] D. Zhong, B. Liu, S. Zhang, and W. Zhang, "Research for the method of parameter estimation to Box–Cox transformation models," *J. Syst. Eng.*, vol. 8, no. 2, pp. 4–46, 1993.
[32] J. W. Osborne, "Improving your data transformations: Applying the Box–Cox transformation," *Pract. Assessment, Res. Eval.*, vol. 15, no. 12, pp. 1–9, 2010.
[33] W. J. Conover, *Practical Nonparametric Statistics*, 3rd ed. New York, NY, USA: Wiley, 1999.
[34] H. Li, D. Pan, and C. L. P. Chen, "Intelligent prognostics for battery health monitoring using the mean entropy and relevance vector machine," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 44, no. 7, pp. 851–862, Jul. 2014.
[35] J. Zhou, Y. Ma, Y. Peng, and X. Peng, "Remaining useful life estimation with dynamic grey relevance vector machine for lithium-ion battery," *Int. J. Adv. Comput. Technol.*, vol. 5, no. 6, pp. 460–469, Mar. 2013.
[36] B. Saha and K. Goebel, *Battery Data Set, NASA Ames Prognostics Data Repository*, NASA Ames, Moffett Field, CA, USA, 2007. [Online]. Available: http://ti.arc.nasa.gov/project/prognostic-data-repository
[37] W. He, N. Williard, M. Osterman, and M. Pecht, "Prognostics of lithium-ion batteries using extended Kalman filtering," in *Proc. Int. Microelectron. Assembly Packag. Soc. (IMAPS) Adv. Technol. Workshop High Rel. Microelectron. Military Appl.*, Fountain Hills, AZ, USA, 2011, pp. 1–4.
[38] S. B. Peterson, J. Apt, and J. F. Whitacre, "Lithium-ion battery cell degradation resulting from realistic vehicle and vehicle-to-grid utilization," *J. Power Sources*, vol. 195, no. 8, pp. 2385–2392, Apr. 2010.
[39] A. Millner, "Modeling lithium ion battery degradation in electric vehicles," in *Proc. IEEE Conf. Innovat. Technol. Efficient Rel. Elect. Supply*, Waltham, MA, USA, 2010, pp. 349–356.
[40] G. Ning and B. N. Popov, "Cycle life modeling of lithium-ion batteries," *J. Electrochem. Soc.*, vol. 151, no. 10, pp. A1584–A1591, 2004.

Datong Liu (M'11) received the B.Sc. degree in automatic test and control with a minor in computer science and technology, and the M.Sc. and Ph.D. degrees in instrumentation science and technology, all from the Harbin Institute of Technology (HIT), Harbin, China, in 2003, 2005, and 2010, respectively.

From 2013 to 2014, he was a Visiting Scholar with the University of Arizona, Tucson, AZ, USA. He is currently an Associate Professor with the Department of Automatic Test and Control, School of Electrical Engineering and Automation, HIT. His current research interests include automatic testing, machine learning, and data mining for test data processing of complex system, data-driven prognostics, lithium-ion battery prognostics, and system health management for aerospace.

Jianbao Zhou received the Ph.D. degree in instrumentation science and technology from the Harbin Institute of Technology (HIT), Harbin, China, in 2013.

He is currently a Postdoctoral Fellow with the School of Electrical Engineering and Automation, HIT. His current research interests include lithium-ion battery prognostics and prognostics for embedded platform.

Haitao Liao (M'12) received the Ph.D. degree in industrial and systems engineering from Rutgers University, New Brunswick, NJ, USA.

He is currently an Associate Professor with the Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ, USA, where he is also the Director of the Reliability and Intelligent Systems Engineering Laboratory. His current research interests include accelerated testing, probabilistic risk assessment, maintenance models and optimization, service part logistics, and prognostics.

Yu Peng (M'10) received the B.S. degree in measurement technology and instrumentation, and the M.Sc. and Ph.D. degrees in instrumentation science and technology from the Harbin Institute of Technology (HIT), Harbin, China, in 1996, 1998, and 2004, respectively.

He is currently a Full Professor with the Department of Automatic Test and Control, School of Electrical Engineering and Automation, HIT, where he is also the Vice Dean of the School of Electrical Engineering and Automation. His current research interests include automatic test technologies, virtual instruments, system health management, and reconfigurable computing.

Xiyuan Peng received the Ph.D. degree in instrumentation science and technology from the Harbin Institute of Technology (HIT), Harbin, China, in 1992.

He is currently a Full Professor with the Department of Automatic Test and Control, School of Electrical Engineering and Automation, HIT, where he is also the Dean. His current research interests include automatic test technologies, advanced diagnostics, and prognostics.