Microelectronics Reliability 53 (2013) 832–839

Contents lists available at SciVerse ScienceDirect

Microelectronics Reliability

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

Prognostics for state of health estimation of lithium-ion batteries based on combination Gaussian process functional regression

Datong Liu [a,*], Jingyue Pang [a], Jianbao Zhou [a], Yu Peng [a], Michael Pecht [b]

[a] Department of Automatic Test and Control, Harbin Institute of Technology, Harbin 150080, China
[b] CALCE, University of Maryland, College Park, MD 20742, USA

ARTICLE INFO

Article history:
Received 15 June 2012
Received in revised form 12 March 2013
Accepted 23 March 2013
Available online 17 April 2013

ABSTRACT

State of health (SOH) estimation plays a significant role in battery prognostics. It is used as a qualitative measure of the capability of a lithium-ion battery to store and deliver energy in a system. At present, many algorithms have been applied to perform prognostics for SOH estimation, especially data-driven prognostics algorithms supporting uncertainty representation and management. To describe the uncertainty in evaluation and prediction, we used the Gaussian Process Regression (GPR), a data-driven approach, to perform SOH prediction with mean and variance values as the uncertainty representation of SOH. Then, in order to realize multiple-step-ahead prognostics, we utilized an improved GPR method—combination Gaussian Process Functional Regression (GPFR)—to capture the actual trend of SOH, including global capacity degradation and local regeneration. Experimental results confirm that the proposed method can be effectively applied to lithium-ion battery monitoring and prognostics by quantitative comparison with the other GPR and GPFR models.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Lithium-ion batteries are core components in a wide variety of systems. Therefore, the reliability of lithium-ion batteries is a subject of great interest to the electronics industry [1]. Conventionally, battery reliability monitoring during usage has two different aspects: state of charge (SOC) and state of health (SOH) [2]. Compared to the study of SOC estimation methods which attracts more attentions of current research work, research on SOH estimation and prediction is still in its initial stage. With the increasing demand for lithium-ion batteries, and SOH estimation plays an important part in battery prognostics as qualitative measure for the battery to store and deliver energy in the system. The prognostics of the SOH can indicate the performance degradation and prevent possible accidents, so more research needs to be conducted to develop prognostics algorithms for SOH estimation.

An extended Kalman filter (EKF) has been applied for real-time prediction of SOC and SOH of automotive batteries [3]. Neuro-fuzzy and decision theoretic methods have been utilized to fuse feature vectors derived from battery health sensor data to estimate SOC, SOH, and state of life (SOL) [4–9]. Other approaches, such as combinations of regressions [10], neural networks [11], fuzzy logic [12], and distributed active learning, have been used to predict remaining useful life (RUL) of batteries [13]. However, a common disadvantage of these prediction methods is the lack of uncertainty expression and management for prognostics. In industrial applications, uncertainty in models and data can lead to poor reliability prediction. To address this problem, prognostic algorithms with uncertainty representation have been developed [14]. Relevance vector machines (RVMs) have been applied to implement regression models for SOH prognosis [5]. Particle filters (PFs) have also been applied to predict the RUL of lithium-ion batteries based on impedance spectroscopy data. For example, researchers have used a Rao–Blackwellized PF (RBPF) to reduce the uncertainty in prediction frameworks [15]. However, the implementation of impedance measurement not only requires expensive equipment but also is more time-consuming. In addition, PF is a model-based method that requires physical or electrochemical knowledge to model degradation trends in batteries. As a result, a prediction approach without complicated system models should be used for lithium-ion battery RUL prediction to achieve low computing complexity as well as uncertainty representation.

In this study, we used a Gaussian process regression (GPR) model for battery prognostics. A GPR model is particularly useful because of its flexibility and ability to provide uncertainty representation and facilitate accurate prediction even in the absence of physical models [16]. Recently, GPR algorithms have been applied in the domain of battery prognostics. Researchers have performed regression for internal parameters of batteries over time based on GPR models and have transferred the predicted values to the capacity domain to indicate capacity decay with time [17]. The results are acceptable, but the extrapolation performance deteriorates rapidly when test data are "distant" from the training data

* Corresponding author. Tel.: +86 45186413533; fax: +86 45186402953.
E-mail address: liudatong@hit.edu.cn (D. Liu).

0026-2714/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.microrel.2013.03.010

[18]. Thus, when we make k-steps-head prognostics and k is large, the prediction based on the GPR model cannot obtain satisfactory results. In addition, under certain operating conditions, there is the regeneration phenomenon (also called the self-recharge phenomenon in this paper) where a battery shows a sudden (and temporary) incremental increase in the capacity available for the next cycle [19]. The regeneration phenomenon influences the prediction of RUL. Moreover, SOH prognostics based on the GPR model, without taking the regeneration phenomenon into account, leads to bad point and variance prediction. In order to solve these problems, in this study we performed a combination Linear Gaussian Process Functional Regression (LGPFR) to perform multi-step-ahead prognostics for the state of health of lithium-ion batteries. Experimental results from the processing of a NASA battery data set proved the effectiveness of the combination LGPFR compared with the basic GPR model and the other Gaussian Process Functional Regression (GPFR) models.

This paper is organized as follows. First, in Section 2, the motivation for this study and an overview of related studies are provided. The prediction methods of GPR, GPFR, and improved combination GPFR are presented in Section 3. The prognostic prediction for SOH of lithium-ion batteries are described in detail in Section 4. Finally, conclusions and future work are discussed in Section 5.

2. Motivation and related studies

2.1. The SOH of lithium-ion batteries

Battery operation is a dynamic process, and its performance is strongly influenced by ambient environmental and load conditions [14]. Many methods have been developed to measure the health of lithium-ion batteries. There are some parameters that describe the health of lithium-ion batteries, such as SOH, SOC, SOL, and so on. SOH estimation is a critical part of battery prognostics. It can be used to show the performance degradation of batteries and decrease the possibility of accidents. There are two main methods for calculating the SOH of batteries [20]. First, battery impedance may be used to indicate battery SOH. Its definition can be given by:

$$\mathrm{SOH} = \frac{R_i}{R_o} \times 100\% \tag{1}$$

where R_i is the ith impedance measurement that is varied with the charge and discharge cycles, and R_0 is the initial impedance. Second, battery capacity, C, can also be introduced to determine the battery SOH, which is the method applied in this paper to measure the SOH. This method can be expressed as:

$$\mathrm{SOH} = \frac{C_i}{C_0} \times 100\% \tag{2}$$

where C_i is the ith capacitance value degenerated with cycles and C_0 is the initial capacity. There have been several studies which estimated battery SOH using battery impedance or battery capacity [21,22].

2.2. The regeneration phenomenon in battery degradation

In this study, we used capacity to indicate the SOH of lithium-ion batteries that can be calculated through the current integral over time from a fully charged state to a fully discharged state. The capacity of a lithium-ion battery shows the degradation trend during usage because of the side reactions that occur between the electrodes and electrolyte of the battery [2]. However, when the battery rests during charge/discharge profiles, the reaction products have the opportunity to dissipate. Thus, the electrochemical performance will relatively recover compared to the former cycle in the degradation process of the lithium-ion batteries. Obviously, this can increase the available capacity for the next cycle [23]. This phenomenon, which emerges during the use of lithium-ion batteries, is called regeneration or self-recharge. The regeneration phenomenon in lithium-ion batteries is introduced and analyzed in [24]. This phenomenon can alter the trend of the SOH prediction curve, thus affecting the performance of prognostic models based on Bayesian algorithms [19]. In order to determine the impact on degradation prediction and improve the prediction performance taking the regeneration phenomenon into account, researchers have estimated the SOH and predicted the RUL of lithium-ion batteries based on a particle-filtering prognostic framework. These methods are used to detect and isolate the influence of the regeneration phenomena within the life cycle model, thus improving the initial conditions that are used to generate long-term predictions [19].

2.3. Gaussian process regression

The GPR model has been applied in many areas due to its advantages of being flexible, probabilistic, and nonparametric [25–27]. It can model the behavior of any system through the combination of the appropriate Gaussian process and realize prognostics combined with prior knowledge based on a Bayesian framework [28]. The basic GPR model has a zero mean function and a standard covariance function, described as a diagonal squared exponential covariance function. In order to solve different problems, transformations have been made to the basic GPR. For example, the Approximate Gaussian Process Regression (AGPR) was developed based on a finite dimensional approximation to solve the scaling problems in a typical Gaussian process [29]. Online Sparse Matrix Gaussian Process Regression (OSMGPR) has been used for online learning [30]. Gaussian Process Dynamic Model (GPDM) was proposed for nonlinear time series analysis and prediction and was found to generalize well for small datasets [31] can improve the multi-step-ahead prognostic performance by considering the mean structure and covariance structure simultaneously [18].

To perform prognostics for the SOH of lithium-ion batteries, we focused on k-step-ahead prognostics in this study. We wanted to perform prognostics with off-line data, and so we did not update the model with dynamic on-line data. Based on the application problems of k-step-ahead and offline prognostics, because of the un-matched performance of the GPR model in multi-step-ahead prognostics [18], the GPFR model is applied to achieve prognostics of lithium-ion batteries. In order to capture the self-recharge phenomenon, a combination GPFR was used to perform prognostics for SOH. Section 4 presents the results of our experiment to compare the three types of models involving basic GPR, GPFR and combination GPFR models.

3. Methods

3.1. Guassian process regression

The Gaussian process (GP) is defined as a collection of a finite number of random variables $\{f(x_i)|x_i \in x\}$ indexed by a set x. The stochastic process is specified by giving the probability distribution for every finite subset of variables $f(x_i)$ in a consistent manner [16]. In a regression and prediction case, x will be the input space with a 1-dimension time series, which is the number of inputs. In our prognostics study, x is the number of charge/discharge cycles. A Gaussian process $f(x)$ can be fully described by its mean $m(x)$ and co-variance function $k(x_i, x_j)$, which are defined as follows [16]:

$$m(x) = E(f(x)) \tag{3}$$

$$k(x_i, x_j) = E[(f(x_i) - m(x_i)) \cdot (f(x_j) - m(x_j))] \tag{4}$$

This can be described as $f(x) \sim \mathrm{GP}[m(x), k(x_i, x_j)]$. In the basic GPR model, the mean function is zero and the covariance function is the squared exponential covariance function. The index set x is the set of input points; these input points do not need time vectors. However, this set of points should be a time vector in prognostics. The covariance function $k(x_i, x_j)$ can be interpreted as the measurement of the distance between the input points x_i and x_j. In the system models, the covariance function is usually composed of two main parts, $k(x_i, x_j) = k_f(x_i, x_j) + k_n(x_i, x_j)$, where k_f represents the functional part which would be used to describe the unknown system model, and k_n indicates the noise part to represent the noise of the model. The possible selections for k_f are discussed in [28].

Here we list the covariance functions that we applied in battery health prognostics [28,32].

The squared exponential covariance function:

$$k_f = \sigma_{f_1}^2 \exp\left(-\frac{(x_i - x_j)^2}{2l_1^2}\right) \tag{5}$$

The periodic covariance function:

$$K_f = \sigma_{f_2}^2 \exp\left(-\frac{2}{l_2^2}\sin^2\left(\frac{\omega}{2\pi}(x_i - x_j)\right)\right) \tag{6}$$

The constant covariance function:

$$k_n(x_i, x_j) = \sigma_n^2 \tag{7}$$

The constant covariance function is used as the noise, which is generally considered to be white Gaussian noise. There are some free parameters in the covariance functions, that is, $\Theta = [\sigma_{f_1}, \sigma_{f_2}, l_1, l_2, w, \sigma_n]^T$. Here $\sigma_{f_1}^2$ and $\sigma_{f_2}^2$ are the signal variance representing the vertical scale of variations of a typical function, σ_n^2 is the noise variance, l_1 and l_2 reflect the length-scale, and w is the angular frequency [32]. Generally, the hyper-parameters are needed to be optimized with the maximization of the log-likelihood function given by [28,32]:

$$\begin{aligned} L &= \log p(y|x, \Theta) \\ &= -\frac{1}{2}\log\left(\det(K_f + \sigma_n^2 I)\right) - \frac{1}{2}y^T\left[K_f + \sigma_n^2 I\right]^{-1}y - \frac{N}{2}\log 2\pi \end{aligned} \tag{8}$$

Here I is the unit matrix of n dimensions and n is the number of training data sets. Given a set of training points $\{(x, y)|i = 1, \ldots, n\}$, we can derive the posterior distribution over functions by imposing a restriction on prior joint distribution. When a Gaussian process is applied to a general regression problem, we describe the target y given the effect of noise, which is $y = f(x) + \varepsilon$. Where ε is the white Gaussian noise, and $\varepsilon \sim N(0, \sigma_n^2)$. If $f(x)$ is Gaussian process, the joint distribution of limited observations from y is also a Gaussian process. Once a posterior distribution is derived, it can be used to estimate predictive values for the test data points [33]. The following equations describe the predictive distribution for GPR [16], where x represents the training data points and x' stands for the test inputs.

Prior:

$$\begin{pmatrix} y \\ f' \end{pmatrix} \sim \left(0, \begin{pmatrix} k_f(x,x) + \sigma_n^2 & k_f(x,x') \\ k_f(x,x')^T & k_f(x',x') \end{pmatrix}\right) \tag{9}$$

Posterior:

$$\overline{f'}|x, y, x' \sim N(f', \mathrm{cov}(f')) \tag{10}$$

where $\overline{f'} = E[f'|x, y, x'] = k_f(x, x')[k_f(x, x) + \sigma_n^2 I]^{-1}y$

$$\mathrm{cov}(f') = k_f(x', x') - [k_f(x', x) + \sigma_n^2 I]^{-1}k_f(x, x')$$

3.2. Gaussian process functional regression model

GPR can obtain satisfied regression performance for the interpolation test data. However, there will be rapid performance deterioration or model mismatching for its extrapolation when the test data are distant from the training data. Consequently, multi-step-ahead forecasting is a typical application. When k is small, a GPR model can achieve a satisfactory prediction result [34]. Related analysis and discussion with experimental and evaluated results can be referred in [34]. Otherwise, when k is large, the prediction performance decreases rapidly, as described in [18]. In [18], the researchers applied the GPFR model to improve the performance of multiple-step-ahead forecasting with GPR. The reason is that the GPFR model considers the mean function and covariance function simultaneously compared to the basic GPR. Moreover, the mean prediction based on GPR is described as $\overline{f'}$ in Eq. (10) which only depends on $k_f(x, x')[k_f(x, x) + \sigma_n^2 I]^{-1}y$; but the mean prediction of GPFR $\overline{f'}$ in Eq. (12) is also determined by $m(x)$. $m(x)$ is the optimized mean function achieved by training the input data points, and it can track the trend of the training data so as to make an impact on the prediction output. As a result, the GPFR model can improve the multi-step-ahead prediction. Due to the multi-step-ahead prediction in prognostics for SOH of lithium-ion batteries, we adopted the GPFR model in this study. We apply the GPFR model to realize the mean function and covariance function training simultaneously. In order to improve the flexibility of the model, we select the mean function vaguely by the introduction of enough parameters, such as the linear mean, which is shown as $m(x) = ax + b$, and the quadratic polynomial is described as $m(x) = ax^2 + bx + c$. Then we can derive the corresponding prior distribution and posterior distribution in a manner similar to the case of GPR model and described as follows:

Prior:

$$\begin{pmatrix} y \\ f' \end{pmatrix} \sim \left(\begin{bmatrix} u \\ u^* \end{bmatrix}, \begin{pmatrix} k_f(x,x) + \sigma_n^2 & k_f(x,x') \\ k_f(x,x')^T & k_f(x',x') \end{pmatrix}\right) \tag{11}$$

Posterior:

$$\overline{f'}|x, y, x' \sim N(f', \mathrm{cov}(f')) \tag{12}$$

where $\overline{f'} = E[f'|x, y, x'] = m(x) + k_f(x, x')[k_f(x, x) + \sigma_n^2 I]^{-1}(y - m)$

$$\mathrm{cov}(f') = k_f(x', x') - [k_f(x', x) + \sigma_n^2 I]^{-1}k_f(x, x')$$

A larger amount of data can enable the GPR model to make better prognostic predictions. Here, we suppose that the training data and test data have been confirmed. We use the equation $u = m(x_i)$, $i = 1, 2, \ldots, n$ for the training means [28] and for the test means u^*. For the covariance, we use $k_f(x, x) + \sigma_n^2$ for the training set covariance, $k_f(x, x')$ for training-test set covariance, and $k_f(x', x')$ for test set covariance. In order to simplify the training procedure, we optimize the optimal hyper-parameters in the mean structure and covariance structure by using the maximization of the log-likelihood function, which is also described in [18]. The prediction framework is shown in Fig. 1.

From Fig. 1 it can be seen that the most important step in prognostics based on GPFR is the selection of the mean function and the covariance function. Generally, the Gaussian process requires some prior knowledge about the forms of the covariance function and mean function; the forms should vary with the prediction. We can select different covariance and mean functions for training. Also we can optimize the parameters using a conjugate gradient method. For example, we can maximize the marginal likelihood

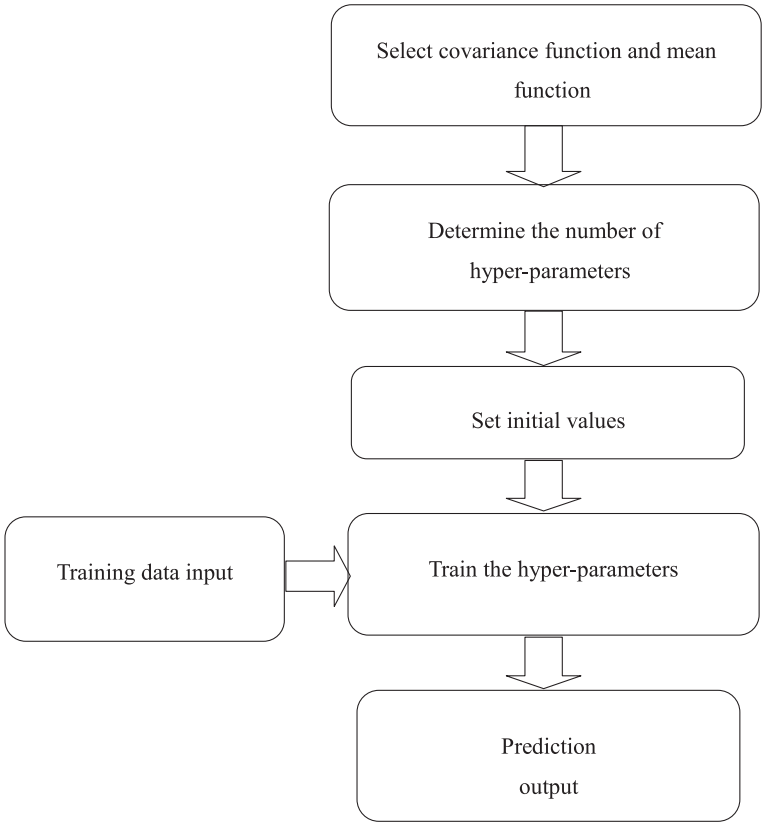

Fig. 1. The prediction framework based on the GPFR model.

of the observed data. In this study, we made prognostic predictions based on different GPFR models compared with the basic GPR model in order to find an effective model to track the SOH trend.

3.3. Combination Gaussian process functional regression

The covariance function of GPFR is generally single. As there are several trends in our training data, a single covariance function cannot describe different trends well. For lithium-ion batteries, with the charge/discharge cycles, the available capacity shows two kinds of changes including the degradation trend and local regeneration. Moreover, the degradation trend and local regeneration present opposite features. When we make prognostics based on GPFR model, both of the changes must be taken into account. Accordingly, the GPFR model, with only one type of covariance function, cannot fulfill the complicated multi-step-ahead predictions. First, the squared exponential covariance function is commonly used to describe the trend of training data points. Second, the regeneration phenomenon can be regarded as the local vibrations in the normal degradation trend. A periodic covariance function is generally used to model a function within a specific period [28]. Although the local regeneration is not strictly periodic, we can realize parameter optimization by training the input data points to estimate the approximate periodic characteristics. So we apply a periodic covariance function to describe the approximate periodic phenomenon of local regeneration.

Therefore, we combined the period covariance function with a standard covariance function to obtain a combination GPFR model to achieve high prediction performance. Both LGPFR (GPFR with linear mean function) and QGPFR (GPFR with quadratic polynomial mean function) can obtain better prediction results at point estimation. Here we present the combination GPFR model:

$$m(x) = ax + b$$
$$k_f(x,x) = \sigma_{f_1}^2 \exp\left(-\frac{(x_i-x_j)^2}{2l_1^2}\right) + \sigma_{f_2}^2 \exp\left(-\frac{2}{l_2^2}\sin^2\left(\frac{\omega}{2\pi}(x_i-x_j)\right)\right) \tag{13}$$

$$m(x) = ax^2 + bx + c$$
$$k_f(x,x) = \sigma_{f_2}^2 \exp\left(-\frac{(x_i-x_j)^2}{2l_2^2}\right) + \sigma_{f_2}^2 \exp\left(-\frac{2}{l_2^2}\sin^2\left(\frac{\omega}{2\pi}(x_i-x_j)\right)\right) \tag{14}$$

4. Experiments and analysis

4.1. Raw data from lithium-ion battery

The lithium-ion battery data that we used to make prognostic predictions based on the different models was obtained from the data repository of the NASA Ames Prognostics Center of Excellence (PCoE). This data set had been sampled from a battery prognostics test bed at NASA comprising commercially available Li-ion 18650 rechargeable batteries along with data obtained from a power supply, programmable DC electronic load, voltmeter, thermocouple sensor, environmental chamber, electro-chemical impedance spectrometry (EIS), PXI chassis based on DAQ, and experiment control conditions [35,36].

The lithium-ion batteries were run through three different operational profiles (charge, discharge, and impedance) at room temperature. Charging was carried out in a constant current (CC) mode at 1.5A until the battery voltage reached 4.2 V and then continued in a constant voltage (CV) mode until the charge current dropped to 20 mA. Discharge was carried out at a constant current (CC) level of 2A until the battery voltage fell to 2.7 V, 2.5 V, 2.2 V,

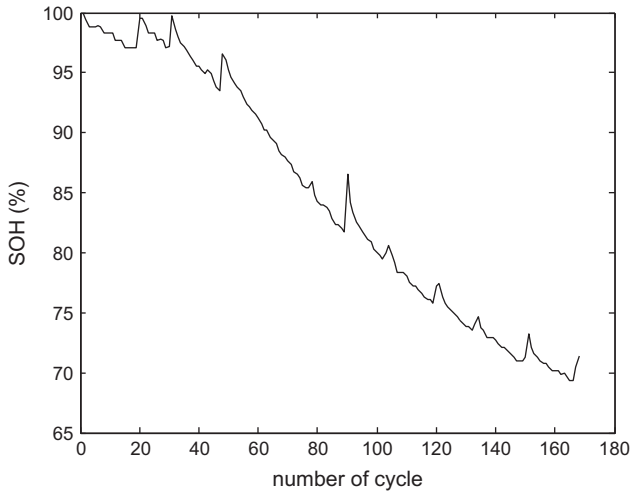

Fig. 2. The SOH of Battery No. 5.

and 2.5 V for batteries No. 5, No. 6, No. 7, and No. 18, respectively. Impedance measurement was conducted by an electrochemical impedance spectroscopy (EIS) frequency sweep from 0.1 Hz to 5 kHz. Repeated charge and discharge cycles resulted in accelerated aging of the batteries. The experiments were stopped when the batteries reached the end-of-life (EOL) criterion, which was a 30% fade in rated capacity (from 2 A h to 1.4 A h) [20]. The data from battery No. 5 was selected to implemented in experiments, including training and testing. The experiment was performed in an accelerated aging profile. In real industrial applications, the number of degradation data samples may be much greater than this aging dataset. The results of the test of SOH conducted for battery No. 5 is shown in Fig. 2.

Fig. 2 contains aging information about battery SOH values from cycle 1 to cycle 168. In order to verify the effectiveness of the GPR model for SOH prognostics, more data samples are required as the training data. Therefore, to reduce the influence of a small amount of data samples and consider the end-of-life (EOL) criteria, the number of training data samples we selected was 100 (that is the data from cycle 1 to cycle 100 are selected as the training set). The prediction input is the data set $\{c(i), i\}$, $i = 1, 2, \ldots, 100$, where i is the number of cycles ranging from 1 to 100, and $c(i)$ is the corresponding capacity at the ith charge/discharge cycle. Fig. 2 shows that the SOH is higher at cycle 90 than in previous cycles due to significant regeneration during the resting period. The regeneration phenomenon also emerges in other cycles which greatly influences the prognostics for RUL estimation. Therefore, the optimum model must be selected to respond to the degradation trend and regeneration phenomenon in lithium-ion batteries.

4.2. Experiments with GPR and GPFR models

The two parts representing the properties of the Gaussian process are the mean function and covariance function, as discussed in Section 3. In the basic GPR model, the mean function is assumed to be zero and the covariance function is the squared exponential covariance function. The basic GPR model can be described as:

$$\begin{aligned} m(x) &= 0 \\ k_f &= \sigma_f^2 \exp\left(-\frac{(x_i - x_j)^2}{2l^2}\right) \end{aligned} \tag{15}$$

The hyper-parameters $\Theta = [l, \sigma_f]^T$ are optimized with the maximization of the log-likelihood function. As described in Section 4.1, 100 data points are used to train the Gaussian process model and achieve prognostics prediction for 68 steps of available capacity. In the following experiments, we compare and evaluate the proposed models.

The prognostic procedures based on the GPR/GPFR model are shown as follows:

(a) Select the training data sets $\{c(i), i\}_{i=1}^{trainN}$, where i is the number of charge/discharge cycles, $c(i)$ is the corresponding available capacity at the ith charge/discharge cycle. As introduced in Section 4.1, *trainN* equals 100 and i ranges from 1 to100.
(b) Initial hype-parameters as random values, which are included in the mean function and covariance function. Here we set the initial hyper-parameters as $\Theta = [l, \sigma_f]^T = [1, 1]^T$ in the GPR model; $\Theta = [a, b, \sigma_{f1}, l_1, \sigma_{f2}, l_2, \omega]^T = [0, 0, 0.1, 1, 0.2, 1, 5]^T$ in the combination LGPFR; and $\Theta = [a, b, c, \sigma_{f1}, l_1, \sigma_{f2}, l_2, \omega]^T = [0, 0, 0, 0.1, 1, 0.2, 1, 5]^T$ in the combination QGPFR.
(c) Optimize the hyper-parameters with the maximization of the log-likelihood function.
(d) Set the prediction steps m.
(e) Use the training data $\{c(i), i\}_{i=1}^{trainN}$ and the test inputs $i' = trainN + 1, \ldots, trainN + m$to be the input of the GPR/GPFR with the optimal hype-parameters. Then the prediction output will be derived including the mean and variance values for every setting step.
(f) Convert the prediction capacity to SOH with Eq. (2).

The prediction results based on a basic GPR model are shown in Fig. 3, and the prognostics results also include 95% confidence bounds for uncertainty representation.

As shown in Fig. 3, the grey region represents the confidence intervals. The upper and lower boundaries are far from the actual SOH points. Moreover, the distance between the boundary and the actual SOH data points increases significantly with the continued prediction process. The large confidence intervals indicate the high uncertainty of the prediction result. The mean prediction output marked with the red line is further away from the actual SOH for the subsequent cycles. We can find that both the point prediction result as well as the uncertainty representation are not satisfied. Experiments have shown the limitations of multi-step-ahead prognostics for lithium-ion batteries based on the basic GPR model. In this case, we performed experiments with a linear mean function, called the LGPFR model, and a quadratic polynomial mean function, called QGPFR, to improve the flexibility and accuracy of the Gaussian process model. In these two types of models, the

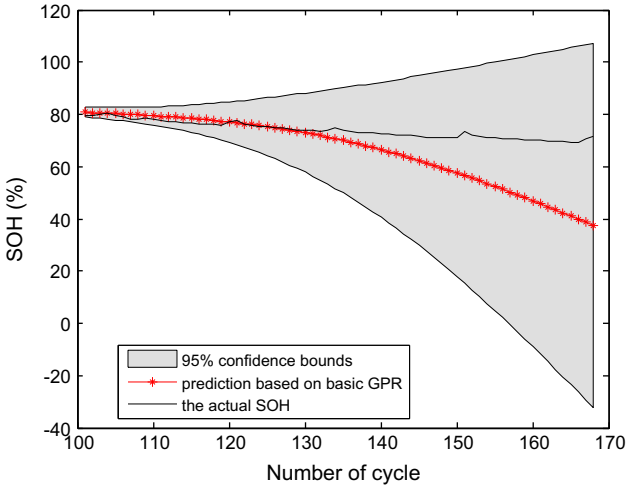

Fig. 3. Battery health prognostics based on basic GPR model with 95% confidence bounds.

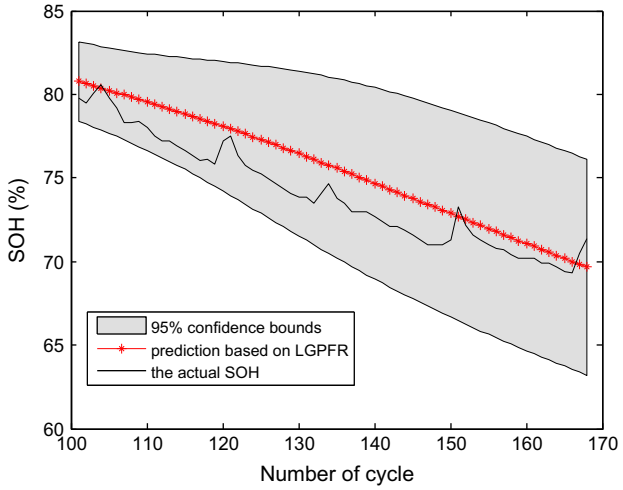

Fig. 4. Battery health prognostics based on LGPFR with 95% confidence bounds.

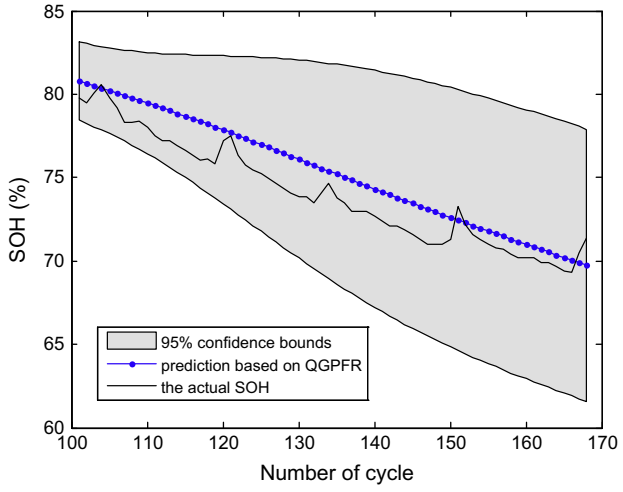

Fig. 5. Battery health prognostics based on QGPFR with 95% confidence bounds.

linear mean function is $m(x) = ax + b$, and the quadratic polynomial mean function is $m(x) = ax^2 + bx + c$. The prediction results with the QGPFR and LGPFR are shown in Figs. 4 and 5.

From Figs. 3–5, we can see that all the prediction results reflect the degradation trend of SOH. Obviously, the predictions based on the LGPFR model and QGPFR model are much better than the basic GPR model. In order to evaluate the two models, we compared the predictions, as shown in Fig. 6, in which only mean prediction values are involved.

As seen in Fig. 6, prediction based on the QGPFR model produced better results than the LGPFR model. The quantitative comparison and evaluation results are shown in Table 1. By analyzing the prediction results, we found that the difference between the QGPFR model and LGPFR model was small. From the above experiments we can see that the prediction results cannot capture the self-recharge phenomenon, even if these two models provide more precise RUL prediction results than the basic GPR model. In the next subsection, we evaluate and verify the proposed combined GPFR model and compare it with the models above.

4.3. Experiments with combination GPFR models

The prediction results shown in Section 4.2 are unable to capture the regeneration phenomenon during the lithium-ion battery

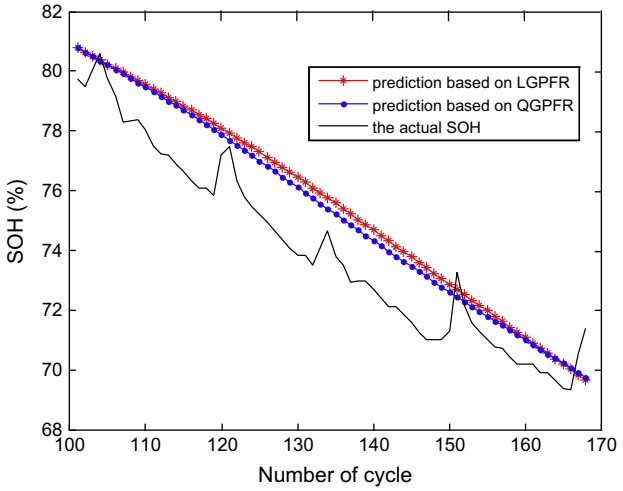

Fig. 6. Battery health prognostics based on LGPFR and QGPFR.

Table 1
Comparison of five GPR models for battery No. 5.

Error criteria	Basic GPR	LGPFR	QGPFR	Combination LGPFR	Combination QGPFR
MAPE	0.121	0.230	0.019	0.016	0.021
RMSE	13.03	1.71	1.50	1.36	1.80

degradation process. Therefore, we performed prognostics for the SOH of a lithium-ion battery based on the combination GPFR, including the combination LGPFR and the combination QGPFR, as introduced in Section 3.3.

The lithium-ion battery SOH prediction results included 95% confidence bounds, as shown in Figs. 7 and 8.

From Figs. 7 and 8, the prediction results with the combination LGPFR model and combination QGPFR model are much closer to the actual SOH curve. Especially for the local regeneration points, the combined models showed satisfactory prediction and approximation ability. In other words, the combined GPFR strategy with the period covariance function and the standard covariance function can respond to the self-recharge phenomenon of battery degradation. Compared to the basic GPR and GPFR models, much better estimation and prediction of SOH can be obtained. The

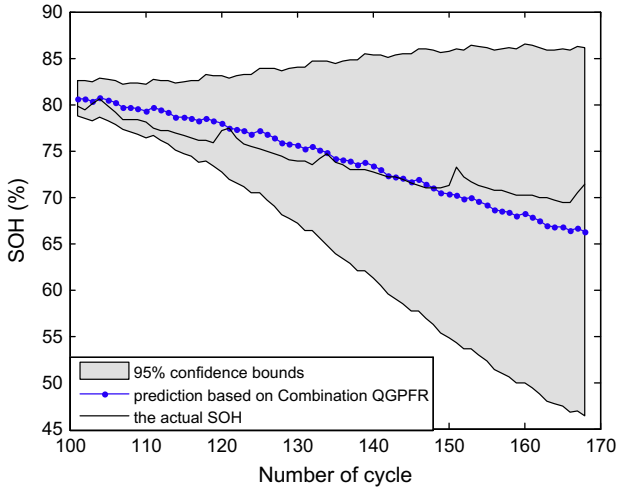

Fig. 7. Battery health prognostics based on combination QGPFR with 95% confidence bounds.

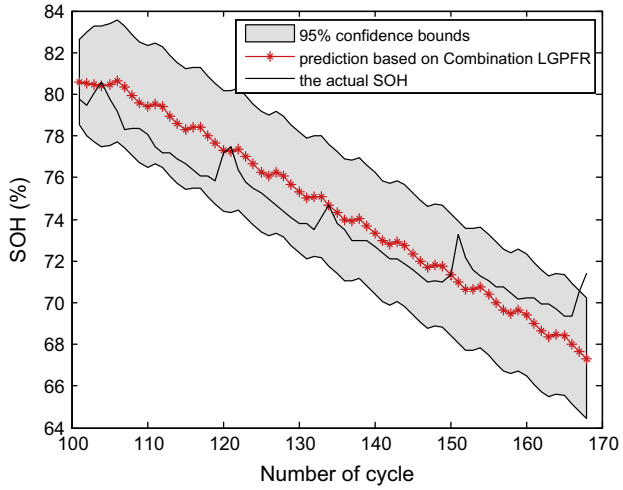

Fig. 8. Battery health prognostics based on combination LGPFR model with 95% confidence bounds.

combination LGPFR model shows similar confidence intervals for the whole testing data sample, even if the test data is distant from the training data set. However, prediction results from the QGPFR model shown in Fig. 7 indicate that the confidence bounds of the prediction value would gradually increase with the prognostic process. The reason is that the combination QGPFR model is more complex than the combination LGPFR, which leads to the conservative prediction result with increasing uncertainty intervals. In order to make the prognostics better, we adopted the effective model while considering the mean and variance prediction simultaneously. Fig. 9 shows details about the mean prediction of the two models.

From Fig. 9, it is shown that the prognostics results only reflect the self-recharge phenomenon in the form of fixed cycles, which may be different from the actual regeneration cycles. The reason is that the selection of the covariance function in this paper is smooth and steady. In addition, we make prognostics with off-line data, and the model is not updated with real-time updating data samples. Furthermore, it is meaningful that the uncertainty intervals of prognostics results based on the combination LGPFR can approximate the actual SOH from the whole prediction interval. In other words, the intervals remain stable for multi-step-ahead prognostics.

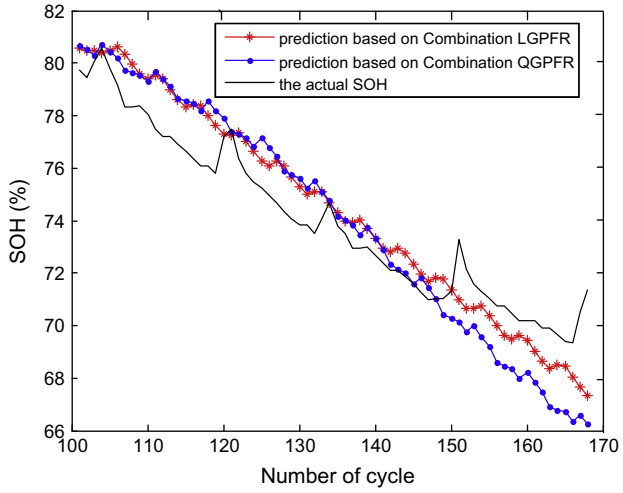

Fig. 9. Battery health prognostics compared with combination LGPFR and Combination QGPFR.

4.4. Comparison and discussion

In Sections 4.2 and 4.3, we can see that the basic GPR model cannot make better prognostic prediction for multi-step-ahead prediction. The application of LGPFR and QGPFR for prognostics improves the case. However, the prediction trend cannot respond to the regeneration phenomenon. There is a regeneration phenomenon in the training data that is not reflected in the prediction curve. The experiments based on a combination LGPFR and a combination QGPFR show the performance improvement on tracking the degradation trend and capturing the self-recharge phenomenon. In order to verify the availability and stability of the combination LGPFR model, we conducted prognostics for SOH on batteries No. 6 and No. 7, which had different charge/discharge profiles, as described in Section 4.1. The experiment was performed based on five different models (basic GPR, LGPFR, QGPFR, combination LGPFR, and combination QGPFR) to compare their results. We also used 100 data points $\{c(i), i\}_{i=1}^{100}$ to train our model and made prognostics for 68 steps of available capacity. Here, two criteria, including the mean absolute percentage error (MAPE) and root mean square error (RMSE), were introduced to evaluate the prediction performance.

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{16}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{17}$$

The quantitative analysis results of the experiment on batteries No. 5, No. 6, and No. 7 based on the five different models are shown in Tables 1–3, respectively.

As shown in Table 1, the GPFR model (including LGPFR and QGPFR) obtain much better prediction performance than the basic GPR model. For example, the prediction RMSE on battery No. 5 based on the basic GPR was 13.03, while the prediction RMSE based on the LGPFR and QGPFR was only 1.71 and 1.5 respectively. This means that the prediction precision with the GPFR improved greatly compared to the basic GPR. Moreover, the prognostic MAPE and RMSE based on the combination LGPFR were the smallest compared to the other four models; the prediction MAPE and RMSE were 0.016 and 1.36, respectively. From Tables 2 and 3, we can get the same results as Table 1. Although the QGPFR can derive a better point prediction than the combination LGPFR from Table 2, the prediction output on battery No. 6 only shows a degradation trend similar to the linear downward trend. In other words, the prediction on battery No. 6 with the QGPFR achieved results similar to the prediction output on battery No. 5 with the QGPFR, which

Table 2
Comparison of five GPR model for battery No. 6.

Error criteria	Basic GPR	LGPFR	QGPFR	Combination LGPFR	Combination QGPFR
MAPE	0.270	0.103	0.077	0.102	0.290
RMSE	22.51	6.90	5.12	6.86	20.44

Table 3
Comparison of five GPR model for battery No. 7.

Error criteria	Basic GPR	LGPFR	QGPFR	Combination LGPFR	Combination QGPFR
MAPE	0.192	0.019	0.054	0.017	0.026
RMSE	20.70	1.59	5.52	1.73	2.69

also did not match the actual self-recharge. Moreover, the confidence intervals become large in the final prediction steps. In brief, the prognostics based on LGPFR and QGPFR can improve the performance of a multi-step-ahead model compared to the basic GPR model. Furthermore, the combination LGPFR constructed with prior knowledge is more suitable to lithium-ion battery SOH prognostics. In conclusion, if we only focus on mean prediction without uncertainty, the combination LGPFR model and combination QGPFR model can also produce good prognostic results. While the uncertainty intervals are considered and concentrated, the combination LGPFR is a better choice for lithium-ion battery RUL prediction.

5. Conclusions

In this paper, we presented a battery SOH estimation approach based on different models of GPR algorithms. In order to improve the poor performance of long-term prediction in multi-step-ahead prognostics based on basic GPR, we applied the GPFR algorithms to perform prognostics for battery SOH. Moreover, considering the effect of the regeneration phenomenon on RUL prognostics and the improvement of RUL prediction precision, we proposed a new combination of covariance functions and mean functions for GPFR with the prior knowledge, called the combination LGPFR, to perform prognostic predictions of battery SOH. Our experiments confirmed that the proposed approach achieved a balance between point estimation and uncertainty representation for lithium-ion batteries with the self-recharge phenomenon.

In the future, the dynamic training strategy and on-line modeling approach and non-stationary model should be considered for real-time processing. Moreover, lithium-ion battery degradation may vary under different operating conditions for various industrial applications, and the effort should be focused on improving the adaptability of the proposed method.

Acknowledgements

This research work is supported by Research Fund for the Doctoral Program of Higher Education of China (20112302120027), and the Program for New Century Excellent Talents (NCET-10-0062), and the Twelfth Government Advanced Research Fund (51317040302).

References

[1] Jingliang Z, Jay L. A review on prognostics and health monitoring of Li-ion battery. J Power Sources 2011;196(15):6007–14.
[2] Wei H, Nick W, Michael O, Michael P. Prognostics of Lithium-ion Batteries using Extended Kalman Filtering. In: IMAPS advanced technology workshop on high reliability microelectronics for military applications; 2011. p. 1–4.
[3] Bhangu BS, Bentley P, Stone DA, Bingham CM. Nonlinear observers for predicting state-of-charge and state-of-health of lead-acid batteries for hybrid-electric vehicles. IEEE Trans Vehicular Technol 2011;54(3):783–94.
[4] Kozlowski JD. Electrochemical cell prognostics using online impedance measurements and model-based data fusion techniques. In: Proceedings of IEEE aerospace conference; 2003. p. 3257–70.
[5] Bhaskar S, Kai G, Jon C. Comparison of prognostic algorithms for estimating remaining useful life of batteries. Trans Inst Measure Control 2009;31:293–308.
[6] Nick W, Wei H, Michael P. Model based Battery Management System for Condition based Maintenance. In: MFPT 2012 proceedings: the prognostics and health management solutions conference; 2012. p. 1–13.
[7] Chaochao C, Michael P. Prognostics of lithium-ion batteries using model-based and data-driven methods. In: 2012 Prognostics & system health management conference. (PHM-2012 Beijing); 2012. p. 1–6.
[8] Wei H, Williard N, Chaochao C, Michael P. State of charge estimation for electric vehicle batteries under an adaptive filtering framework. In: 2012 Prognostics & System Health Management Conference. (PHM-2012 Beijing); 2012. p. 1–5.
[9] Heng-Juan C, Qiang M, Wei L, Zhonglai W, Michael P. Application of unscented particle filter in remaining useful life prediction of lithium-ion batteries. In: 2012 Prognostics & system health management conference (PHM-2012 Beijing); 2012. p. 1–6.
[10] Pattipati B, Pattipati K, Christopherson JP, Namburu SM, Prokhorov DV, Liu Q. Automotive battery management systems. In: 2008 IEEE Autotestcon; 2008. p. 581–6.
[11] Kejun Q, Chengke Z, Yue Y, Allan M. Temperature effect on electric vehicle battery cycle life in vehicle-to-grid applications. In: 2010 China international conference on electricity distribution (CICED); 2010. p. 1–6.
[12] Alvin JS, Craig F, Pritpal S, Terrill A, David ER. Determination of state-of-charge and state-of-health of batteries by fuzzy logic methodology. J Power Sources 1999;80(1–2):293–300.
[13] Huimin C, Li XR. Distributed active learning with application to battery health management. In: Proceedings of the 14th international conference on information fusion (FUSION); 2011. p. 1–7.
[14] Saha B, Goebel K, Poll S, Christophersen J. An integrated approach to battery health monitoring using Bayesian regression and state estimation. In: 2007 IEEE Autotestcon; 2007. p. 646–53.
[15] Saha B, Goebel K, Poll S, Christophersen J. Prognostics methods for battery health monitoring using a Bayesian framework. Instrument Measure, IEEE Trans 2009;58:291–6.
[16] Williams CKI, Rasmussen CE. Gaussian processes for regression. In: Advances in neural information processing systems 8. MIT Press; 1996.
[17] Goebel K, Saha B, Saxena A, Celaya J, Christophersen J. Prognostics in battery health management. IEEE Instrum Meas Mag 2008;4:33–40.
[18] Shi JQ, Wang B. Gaussian process functional regression modeling for batch data. Biometrics 2007;63:714–23.
[19] Olivares BE, Cerda Munoz MA, Orchard ME, Silva JF. Particle-filtering-based prognosis framework for energy storage devices with a statistical characterization of state-of-health regeneration phenomena. IEEE T Instrum Meas 2013;62:364–76.
[20] Le D, Tang XD. Lithium-ion battery state of health estimation using ah-v characterization. In: Annual conference of the prognostics and health management society; 2011. p. 367–73.
[21] Dai HF, Wei XZ, Sun ZC. A new SOH prediction concept for the power lithium-ion battery used on HEVs. In: Vehicle power and propulsion conference; 2009. p. 1649–53.
[22] Kim IS. A technique for estimating the state of health of lithium batteries through a dual-sliding-mode observer. IEEE Trans Power Electron 2010;25:1013–22.
[23] Widodo A, Shim MC, Caesarendra W, Yang BS. Intelligent prognostics for battery health monitoring based on sample entropy. Expert Syst Appl 2011;38:11763–9.
[24] Orchard M, Tang L, Saha B, Goebel K, Vachtsevanos G. Risk-sensitive particle-filtering-based prognosis framework for estimation of remaining useful life in energy storage devices. Studies Inform Control 2010;19:209–18.
[25] Kocijan J, Tanko V. Prognosis of gear health using gaussian process model. In: 2011 IEEE international conference on computer as a tool (EUROCON); 2011. p. 1–4.
[26] Yunong Z, Leithead WE, Leith DJ. Time-series Gaussian process regression based on toeplitz computation of O(N2) operations and O(N)-level storage. In: Decision and control, 2005 and 2005 European control conference; 2005. p. 3711–16.
[27] Tao C, Julian M, Elaine M. Gaussian process regression for multivariate spectroscopic calibration. Chem Intell Lab Syst 2007;87:59–71.
[28] Rasmussen CE, Williams CKI. Gaussian processes for machine learning. Cambridge (MA): MIT Press; 2006.
[29] Anton S, Volker T. Transductive and Inductive methods for approximate gaussian process regression. In: Advances in nerual information processing systems 15. MIT Press; 2003.
[30] Ananth R, Ming-Hsuan Y. Online sparse matrix gaussian process regression and vision applications. Lect Notes Comput Sci 2008;5302:468–82.
[31] Jack MW, David JF, Aaron H. Gaussian process dynamical models. Adv Neural Inform Process Syst (NIPS) 2005;18:1441–8.
[32] MacKay D. Introduction to Gaussian processes. Tech-nical report. UK: Cambridge University; 1997. <Seehttp://wol.ra.phy.cam.ac.uk/mackay/README.html>.
[33] Saha S, Saha B, Saxena A, Goebel K. Distributed prognostic health management with gaussian process regression, In: IEEE Aerospace conference; 2010. p. 1–8.
[34] Girard A, Murray-Smith R. Gaussian processes: prediction at a noisy input and application to iterative multiple-step ahead fore-casting of time-series. Switching and learning in feedback systems. New York: Springer-Verlag; 2005.
[35] Saha B, Goebel K. Uncertainty management for diagnostics and prognostics of batteries using Bayesian techniques. In: Proceedings of the IEEE aerospace conference; 2008. p. 1–8.
[36] Saha B, Goebel K. Battery data set, NASA ames prognostics data repository. NASA Ames, Moffett Field, CA; 2007. <http://ti.arc.nasa.gov/project/prognostic-data-repository>.