

Research Article

A Hybrid Prognostic Approach for Remaining Useful Life Prediction of Lithium-Ion Batteries

Wen-An Yang,¹ Maohua Xiao,² Wei Zhou,³ Yu Guo,¹ and Wenhe Liao¹

¹College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

²College of Engineering, Nanjing Agricultural University, Nanjing 210031, China

³Nanjing Surveying and Mapping Instrument Factory, Nanjing 210003, China

Correspondence should be addressed to Wen-An Yang; dreamflow@nuaa.edu.cn

Received 3 July 2015; Revised 30 October 2015; Accepted 1 November 2015

Academic Editor: Chuan Li

Copyright © 2016 Wen-An Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Lithium-ion battery is a core component of many systems such as satellite, spacecraft, and electric vehicles and its failure can lead to reduced capability, downtime, and even catastrophic breakdowns. Remaining useful life (RUL) prediction of lithium-ion batteries before the future failure event is extremely crucial for proactive maintenance/safety actions. This study proposes a hybrid prognostic approach that can predict the RUL of degraded lithium-ion batteries using physical laws and data-driven modeling simultaneously. In this hybrid prognostic approach, the relevant vectors obtained with the selective kernel ensemble-based relevance vector machine (RVM) learning algorithm are fitted to the physical degradation model, which is then extrapolated to failure threshold for estimating the RUL of the lithium-ion battery of interest. The experimental results indicated that the proposed hybrid prognostic approach can accurately predict the RUL of degraded lithium-ion batteries. Empirical comparisons show that the proposed hybrid prognostic approach using the selective kernel ensemble-based RVM learning algorithm performs better than the hybrid prognostic approaches using the popular learning algorithms of feedforward artificial neural networks (ANNs) like the conventional backpropagation (BP) algorithm and support vector machines (SVMs). In addition, an investigation is also conducted to identify the effects of RVM learning algorithm on the proposed hybrid prognostic approach.

1. Introduction

Lithium-ion batteries are significant energy solution for many systems (e.g., satellite, spacecraft, and electric vehicles) due to their high energy density, high galvanic potential, lightness of weight, and long lifetime compared to lead-acid, nickel-cadmium, and nickel-metal-hydride cells [1]. Their failure can lead to reduced capability, downtime, and even catastrophic breakdowns. For example, in November 2006, The National Aeronautics and Space Administration's Mars Global Surveyor stopped working after the radiator for its batteries was positioned towards the sun causing an increase in the temperature of the batteries, which resulted in lost charge capacity [2]. Battery health management would greatly enhance the reliability of such systems. Thus, this raises the challenging issue of remaining useful life (RUL) prediction in relation to lithium-ion batteries.

In the past few years, much research effort has been devoted to developing approaches to lithium-ion battery degradation modeling and RUL prediction. In general, these approaches can be classified into categories of model-based and/or data-driven methodologies. The model-based methodologies attempt to constitute physical models of the lithium-ion battery for RUL prediction. Recently, various Bayesian filtering models such as Kalman filter [3], extended Kalman filter [4–6], particle filter [7–9], and unscented particle filter [10] have been extensively used to construct exhaustive models of deteriorating lithium-ion batteries. However, uncertainty due to assumptions and simplifications in the models may impose severe limitations upon their applicability in practical applications. In order to overcome the aforementioned problems that can occur with the model-based methodologies, intensive research has been conducted into the utilization of various data-driven methodologies,

for example, autoregressive moving average (ARMA) models [11], artificial neural networks (ANNs) [12], and support vector machines (SVMs) [13], to model lithium-ion battery degradation and to predict the RUL of lithium-ion batteries. Data-driven techniques utilize monitored operational data related to lithium-ion battery health. Compared with the model-based methodologies, the data-driven methodologies may be more appropriate when the understanding of first principles of system operation is not comprehensive or when the system is so complex such that developing an accurate model is prohibitively expensive but sufficient data are available for constructing a map of the performance degradation space. Furthermore, rapid development has recently been achieved in automatic data collection and processing of real-time field data, which hugely facilitate the continuous monitoring of the state of health of operating lithium-ion batteries and the lean management of the related large amount of reference data. The most natural data-driven methodology for RUL prediction is to fit a curve of the available data of the lithium-ion battery degradation evolution using regression models and then to extrapolate the curve to the criteria indicating failure. In practice, however, the lithium-ion battery degradation history available may be short and incomplete and even differ significantly because of different operating conditions, so that a common extrapolation may lead to large errors and unreliable results. The same problem arises when employing ARMA models, although the method can handle the situation in which more run-to-failure data are unavailable or insufficient. With respect to ANNs, they have the advantages of superior learning, noise suppression, and parallel computation abilities. However, despite their advantages, ANNs also have some disadvantages: (1) design and training often lead to a complex and time-consuming task, in which architecture and many training parameters must be tuned; (2) minimization of the training errors can result in poor generalization performance; and (3) performance can be degraded when working with low-sized datasets. With respect to SVMs, they are powerful in solving problems with small samples, nonlinearities, and local minimum. However, despite their advantages, SVMs also have some disadvantages: (1) by assuming an explicit loss function (usually, the ϵ -insensitive loss function), one assumes a fixed distribution of the residuals; (2) the soft margin parameter C must be tuned usually through cross-validation methods, which result in time-consuming tasks; (3) the kernel function used in SVM must satisfy Mercer's theorem to be valid; and (4) sparsity is not always achieved and a high number of support vectors are thus obtained.

More recently, some researchers have attempted to combine model-based and data-driven methods for RUL prediction of lithium-ion batteries in order to leverage the strength from both data-driven methodology and model-based methodology and have obtained promising results [14]. Most of the combination of model-based and data-driven methods in literature has focused on the utilization of relevance vector machines (RVMs) in place of ANNs or SVMs as the prognostic technique. RVM, a general Bayesian probabilistic framework of SVM, can efficiently

alleviate some of these shortcomings of SVMs [15]. Saha et al. employed a RVM to find the most representative relevant vectors to fit the capacity degradation data of lithium-ion batteries [16]. Maio et al. combined a RVM and an exponential function to predict the RUL of bearings [17]. Zio and Maio employed a RVM to find the most representative relevant vectors to fit a crack growth model for predicting RUL [18]. Wang et al. employed a RVM to find the most representative relevant vectors to fit the three-parameter capacity degradation model to predict the RUL of lithium-ion batteries [19]. A review of the related literature also indicates that similar idea has already been investigated in the area of applying SVM to RUL prediction. Benkedjouh et al. [20] employed a SVM to find the most representative support vectors to fit a power model for RUL prediction of the cutting tool. Also based on a similar idea, Benkedjouh et al. employed a SVM to find the most representative support vectors to fit an exponential regression for bearing performance degradation assessment and RUL estimation [21]. The ability to extract the relevant vectors is very useful for making good predictions, as the relevant vectors can be used to find the representative training vectors containing the cycles of the relevant vectors and the predictive values at the cycles of the relevant vectors. A review of the related literature [16–21] also indicates that, for the hybrid prognostic approaches that are based on RVM learning algorithm, their RUL prediction performances are very sensitive to kernels choice and kernel parameters setting. A kernel (or kernel parameter setting) that works well for one situation might not be the appropriate choice for the other. However, no systematic methodology as yet has been established for determining the optimal kernel type and kernel parameters for the RVM learning algorithm. Most of the previous work in the area of applying RVM to RUL prediction determined single kernel and kernel parameters by trial and error and did not deal with automatic kernel choice and kernel parameters optimization.

According to the literature review given above, the aim of this study is to develop a hybrid prognostic approach of physical laws and data-driven modeling that integrates selective kernel ensemble-based RVM (a data-driven methodology) and exponential regression (a model-based methodology) for on-line RUL prediction of lithium-ion batteries. The choice of kernel (and kernel parameters) of RVM is evolutionarily determined via coevolutionary swarm intelligence, without the need of any human intervention. A sum of two exponential functions' model is fitted to these relevant vectors to predict the RUL of degraded lithium-ion batteries. The experimental results indicate that the proposed hybrid prognostic approach can accurately predict the RUL of degraded lithium-ion batteries. Empirical comparisons show that the proposed hybrid prognostic approach using the selective kernel ensemble-based RVM learning algorithm performs better than the hybrid prognostic approaches using popular learning algorithms of feedforward artificial neural networks (ANNs) like the conventional backpropagation (BP) algorithm and support vector machines (SVMs). The proposed hybrid prognostic approach using the selective kernel ensemble-based RVM learning algorithm outperforms

the hybrid prognostic approaches using the single kernel-based RVM learning algorithm and the Ensemble All-based RVM learning algorithm.

The rest of this study is organized as follows. Section 2 gives a review of the RVM basic framework. Section 3 presents a selective kernel ensemble-based RVM learning algorithm. Section 4 describes a hybrid prognostic approach for RUL prediction of lithium-ion batteries. Section 5 conducts an investigation to identify the effects of RVM learning on the hybrid prognostic approach. Section 6 provides an empirical comparison of the proposed hybrid prognostic approach with other existing approaches. Section 7 presents a concluding summary and suggests some directions for future research.

2. Review of Relevance Vector Machine

RVM is a Bayesian form representing a generalized linear model of identical functional form of SVM. Unlike SVM, RVM can provide probabilistic interpretation of its outputs [15]. As a supervised learning, RVM starts with a dataset of input-target pairs $\{\mathbf{x}_n, t_n\}_{n=1}^N$. The aim is to learn a model of the dependency of the targets on the inputs to make accurate prediction of t for previously unseen values of \mathbf{x} . Typically, the predictions are based on a function $y(\mathbf{x})$ defined over the input space, and learning is the process of inferring (perhaps the parameters of) this function. In the context of SVM, this function takes the following form:

$$y(\mathbf{x}) = \sum_{n=1}^N w_n K(\mathbf{x}, \mathbf{x}_n) + w_0, \quad (1)$$

where $\{w_n\}$ are the model “weights,” w_0 is bias, and $K(\mathbf{x}, \mathbf{x}_n)$ is a kernel function.

By considering only the scalar valued output we follow the standard probabilistic formulation and add additive noise with output samples for better data overfitting, which is described as follows:

$$t_n = y(\mathbf{x}_n) + \varepsilon_n, \quad (2)$$

where ε_n are independent samples from some noise process which is further assumed to be zero-mean Gaussian noise with variance σ^2 .

The likelihood of the complete dataset can be written as

$$p(\mathbf{t} | \mathbf{w}, \sigma^2) = (2\pi\sigma^2)^{-(N/2)} \exp\left\{-\frac{1}{2\sigma^2} \|\mathbf{t} - \Phi\mathbf{w}\|^2\right\}, \quad (3)$$

where $\mathbf{t} = [t_1, t_2, \dots, t_N]^T$, $\mathbf{w} = [w_0, w_1, w_2, \dots, w_N]^T$, and Φ is the $N \times (N + 1)$ “design” matrix with $\Phi = [\boldsymbol{\varphi}(\mathbf{x}_1), \boldsymbol{\varphi}(\mathbf{x}_2), \dots, \boldsymbol{\varphi}(\mathbf{x}_N)]^T$, wherein $\boldsymbol{\varphi}(\mathbf{x}_n) = [1, K(\mathbf{x}_n, \mathbf{x}_1), K(\mathbf{x}_n, \mathbf{x}_2), \dots, K(\mathbf{x}_n, \mathbf{x}_N)]^T$.

Maximizing likelihood prediction of \mathbf{w} and σ^2 in (3) often leads to overfitting. Therefore, a preference for smoother functions is encoded by choosing a zero-mean Gaussian prior distribution over \mathbf{w} :

$$p(\mathbf{w} | \boldsymbol{\alpha}) = \prod_{i=0}^N N(w_i | 0, \alpha_i^{-1}), \quad (4)$$

where $\boldsymbol{\alpha}$ is a vector of $(N + 1)$ hyperparameters.

Using Bayes’ rule, the posterior over all unknowns can be computed; that is,

$$p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2 | \mathbf{t}) = \frac{p(\mathbf{t} | \mathbf{w}, \boldsymbol{\alpha}, \sigma^2) p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2)}{p(\mathbf{t})}. \quad (5)$$

However, we cannot compute the solution of the posterior $p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2 | \mathbf{t})$ in (5) directly. But we can decompose the posterior as $p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2 | \mathbf{t}) = p(\mathbf{w} | \mathbf{t}, \boldsymbol{\alpha}, \sigma^2) p(\boldsymbol{\alpha}, \sigma^2 | \mathbf{t})$, where

$$\begin{aligned} p(\mathbf{w} | \mathbf{t}, \boldsymbol{\alpha}, \sigma^2) &= \frac{p(\mathbf{t} | \mathbf{w}, \sigma^2) p(\mathbf{w} | \boldsymbol{\alpha})}{p(\mathbf{t} | \boldsymbol{\alpha}, \sigma^2)} \\ &= (2\pi)^{-(N+1)/2} |\boldsymbol{\Sigma}|^{-1/2} \\ &\quad \cdot \exp\left\{-\frac{1}{2} (\mathbf{w} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{w} - \boldsymbol{\mu})\right\}, \end{aligned} \quad (6)$$

where the posterior covariance and mean are expressed as follows:

$$\begin{aligned} \boldsymbol{\Sigma} &= (\boldsymbol{\Phi}^T \mathbf{B} \boldsymbol{\Phi} + \mathbf{A})^{-1} \\ \boldsymbol{\mu} &= \boldsymbol{\Sigma} \boldsymbol{\Phi}^T \mathbf{B} \mathbf{t} \end{aligned} \quad (7)$$

with $\mathbf{A} = \text{diag}(\alpha_0, \alpha_1, \dots, \alpha_N)$ and $\mathbf{B} = \sigma^{-2} \mathbf{I}_N$. Thus, RVM method becomes the search for the best hyperparameters posterior mode. Predictions for new data are then made according to integration of the weights to obtain the marginal likelihood for the hyperparameters:

$$\begin{aligned} p(\mathbf{t} | \boldsymbol{\alpha}, \sigma^2) &= \int p(\mathbf{t} | \mathbf{w}, \sigma^2) p(\mathbf{w} | \boldsymbol{\alpha}) d\mathbf{w} = (2\pi)^{-N/2} \\ &\quad \cdot |\mathbf{B}^{-1} + \boldsymbol{\Phi} \mathbf{A}^{-1} \boldsymbol{\Phi}^T|^{-1/2} \\ &\quad \cdot \exp\left\{-\frac{1}{2} \mathbf{t}^T (\mathbf{B}^{-1} + \boldsymbol{\Phi} \mathbf{A}^{-1} \boldsymbol{\Phi}^T)^{-1} \mathbf{t}\right\}. \end{aligned} \quad (8)$$

The hyperparameters $\boldsymbol{\alpha}$ and σ^2 which maximize (8) are obtained by using an alternate reproduction approach [15], because values of $\boldsymbol{\alpha}$ and σ^2 cannot be directly calculated in closed form. Suppose that the values of $\boldsymbol{\alpha}_{MP}$ and σ_{MP}^2 that can maximize (8) are obtained. Then we can compute the predictive distribution for a new input \mathbf{x}_* by using (6):

$$\begin{aligned} p(t_* | \mathbf{t}, \boldsymbol{\alpha}_{MP}, \sigma_{MP}^2) \\ = \int p(t_* | \mathbf{w}, \sigma_{MP}^2) p(\mathbf{w} | \mathbf{t}, \boldsymbol{\alpha}_{MP}, \sigma_{MP}^2) d\mathbf{w}. \end{aligned} \quad (9)$$

Since both terms in the integral are Gaussian, one can easily compute the probability as follows:

$$p(t_* | \mathbf{t}, \boldsymbol{\alpha}_{MP}, \sigma_{MP}^2) = N(t_* | y_*, \sigma_*^2), \quad (10)$$

where the mean and variance of the predicted value are, respectively,

$$y_* = \boldsymbol{\mu}^T \boldsymbol{\varphi}(\mathbf{x}_*), \quad (11)$$

$$\sigma_*^2 = \sigma_{MP}^2 + \boldsymbol{\varphi}(\mathbf{x}_*)^T \boldsymbol{\Sigma} \boldsymbol{\varphi}(\mathbf{x}_*). \quad (12)$$

The variance of the predicted value (i.e., (12)) is the sum of the variance associated with noise in the training data and uncertainty associated with prediction of weights.

3. Selective Kernel Ensemble-Based Relevance Vector Machine

As mentioned in Section 1, kernel types and kernel parameters have significant influences on the generalization capability of the RVM learning. Generally, commonly used basic kernels for RVM learning include Gaussian kernel (i.e., (13)), Exponential kernel (i.e., (14)), Laplacian kernel (i.e., (15)), Polynomial kernel (i.e., (16)), Sigmoid kernel (i.e., (17)), Cauchy kernel (i.e., (18)), and Multiquadric kernel (i.e., (19)):

$$K_{\text{Gau}}(\mathbf{x}, \mathbf{x}_i) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_i\|^2}{2\sigma_{\text{Gau}}^2}\right), \quad (13)$$

$$K_{\text{Exp}}(\mathbf{x}, \mathbf{x}_i) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_i\|}{2\sigma_{\text{Exp}}^2}\right), \quad (14)$$

$$K_{\text{Lap}}(\mathbf{x}, \mathbf{x}_i) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_i\|}{\sigma_{\text{Lap}}}\right), \quad (15)$$

$$K_{\text{Poly}}(\mathbf{x}, \mathbf{x}_i) = (\mathbf{x}^T \mathbf{x}_i + 1)^d, \quad (16)$$

$$K_{\text{Sig}}(\mathbf{x}, \mathbf{x}_i) = \tanh(\xi \mathbf{x}^T \mathbf{x}_i + \eta), \quad (17)$$

$$K_{\text{Cau}}(\mathbf{x}, \mathbf{x}_i) = \frac{1}{1 + \|\mathbf{x} - \mathbf{x}_i\|^2 / \sigma_{\text{Cau}}^2}, \quad (18)$$

$$K_{\text{Mul}}(\mathbf{x}, \mathbf{x}_i) = \sqrt{\|\mathbf{x} - \mathbf{x}_i\|^2 + \tau^2}, \quad (19)$$

where σ_{Gau} , σ_{Exp} , σ_{Lap} , d , ξ , η , σ_{Cau} , and τ are kernel parameters that need to be finely tuned. It is impossible to fully determine which one is the best kernel for all problems, because the choice of a kernel depends on the problem at hand. For example, Gaussian kernel is a local kernel and Polynomial kernel is a global kernel [22]. In the case of local kernel, only the data that are close or in the proximity of each other have an influence on the kernel values [22]. In the case of global kernel, samples that are far away from each other still have an influence on the kernel value [22]. With respect to Gaussian kernel and Polynomial kernel only, the mixture of these two basic kernels has been demonstrated to substantively improve the generalization performance of the SVM [23, 24]. However, for many existing basic kernels mentioned above, this combination of basic kernels can also be different for different problems. In one extreme case where all of the individual basic kernels are completely identical, the size of the combination can be reduced without sacrificing the generalization performance of the RVM. In addition, in some scenarios, eliminating some unacceptable basic kernels and meanwhile selecting several acceptable ones to construct a kernel ensemble may be better than combining all of those basic kernels. In this study, each kernel applied to RVM learning algorithm is a selective kernel ensemble of these basic kernels. It should be noted that although

the multikernel idea has been successfully used in several machine learning models [25–28] that assume a weighted linear sum of basic kernel and estimate the kernel weights during training, to the best of the authors' knowledge, it is the first time that a multikernel version of RVM with adaptive kernel selections, adaptive kernel combinations, and adaptive kernel parameters optimization is proposed. The selective kernel ensemble can be expressed as follows:

$$K(\mathbf{x}, \mathbf{x}_i) = \sum_{h=1}^H \lambda_h \rho_h K_h(\mathbf{x}, \mathbf{x}_i) \quad (20)$$

$$\sum_{h=1}^H \lambda_h = 1$$

$$0 \leq \lambda_h \leq 1$$

$$\rho_h = \begin{cases} 1, & \text{being selected} \\ 0, & \text{otherwise,} \end{cases}$$

where H is the number of basic kernels under consideration and equals 7 in this study, $K_h(\mathbf{x}, \mathbf{x}_i)$ denotes the h th basic kernel, λ_h stands for the weight assigned to $K_h(\mathbf{x}, \mathbf{x}_i)$, and ρ_h represents the selection label assigned to $K_h(\mathbf{x}, \mathbf{x}_i)$.

3.1. Selection of Candidate Basic Kernels. Instead of combining all of candidate basic kernels, selective kernel ensemble tries to select an optimal subset of individual basic kernels to constitute a selective convex combination. However, selecting an optimal subset from candidate basic kernels is not an easy task since the space of possible subsets is very large ($2^H - 1$) for a basic kernel population of size H . It is very difficult if not impractical to use exhaustive search to find an optimal subset if and especially when H is a large number. In this study, discrete particle swarm optimization (DPSO) [29] algorithm is used for obtaining an optimal subset from candidate basic kernels. Each dimension of a particle in DPSO is encoded by binary bit, where each element of "1" (i.e., $\rho_h = 1$) denotes an individual basic kernel appearing in the selective kernel ensemble while "0" (i.e., $\rho_h = 0$) denotes its absence, $h = 1, 2, \dots, 7$. The optimal subset of individual basic kernels can be obtained according to the best evolved selective label vector $\Theta = (\rho_1, \rho_2, \dots, \rho_7)$ that can achieve the maximum fitness value. Thus, such a DPSO bit representation gets rid of the tedious trial-and-error search for an optimal subset of basic kernels.

3.2. Determination of Kernel Parameters and Additional Weights. Although utilization of selective kernel ensemble can relieve the influence of kernel types on the generalization capability of RVM, it involves 7 additional weight coefficients ($\lambda_1, \lambda_2, \dots, \lambda_7$). In addition, more component basic kernels mean more kernel parameters. It is not easy to determine the optimal values of all these design parameters, including kernel parameters (σ_{Gau} , σ_{Exp} , σ_{Lap} , d , ξ , η , σ_{Cau} , and τ) and convex combination coefficients ($\lambda_1, \lambda_2, \dots, \lambda_7$) that can allow the RVM to achieve the

maximum performance. In this circumstance, manual trial-and-error method is absolutely tedious and unacceptable. Moreover, manual trial-and-error method does not necessarily guarantee a good decision, because these parameters usually interact with each other nonlinearly. In this study, these 7 additional weight coefficients ($\lambda_1, \lambda_2, \dots, \lambda_7$) together with kernel parameters ($\sigma_{\text{Gau}}, \sigma_{\text{Exp}}, \sigma_{\text{Lap}}, d, \xi, \eta, \sigma_{\text{Cau}}$, and τ) constitute a general real-value parameter vector $\Psi = \{\lambda_1, \lambda_2, \dots, \lambda_7, \sigma_{\text{Gau}}, \sigma_{\text{Exp}}, \sigma_{\text{Lap}}, d, \xi, \eta, \sigma_{\text{Cau}}, \tau\}$, which will be represented in the population of continuous particle swarm optimization (CPSO) [30]. Thus, such a CPSO real-value representation gets rid of the tedious trial-and-error search for optimal kernel parameters and additional weights.

3.3. Coevolution of DPSO and CPSO. As mentioned in Sections 3.1 and 3.2, one swarm population DPSO with population size s_{DPSO} and the other swarm population CPSO with population size s_{CPSO} are involved in equipping the RVM with adaptive kernel selections, adaptive kernel combinations, and adaptive kernel parameters optimization. From a pure DPSO perspective, this suffices for the design of the RVM with the best kernel selection, but without taking kernel parameters and weights in kernel combination into account; that is, only good kernel selection obtained with DPSO may not necessarily mean good RVM performance. Similarly, only good kernel parameters and weights in kernel combination may not necessarily evoke maximum RVM performance. Therefore, the evolution of kernel selections by DPSO and the evolution of kernel combinations and kernel parameters by CPSO should be taken into consideration simultaneously. Inspired by the coevolution of swarms, a coevolutionary PSO scheme is proposed in this section. In the proposed coevolutionary PSO scheme, the DPSO and the CPSO interact with each other through the fitness evaluation. Within each iteration, the DPSO is run for a certain number (g_DPSO) of generations; then the CPSO is run for a certain number (g_CPSO) of generations; this process is repeated until either an acceptable solution has been obtained or the maximum number (max_i_PSO) of iterations has been reached. The global best in the population of DPSO is the final solution for the selection label vector, and the global best in the population of CPSO is the final solution for the general parameter vector with regard to kernel parameters and additional weight coefficients. The procedure of coevolution of DPSO and CPSO is outlined in the following pseudocode.

Step 1. Initialize randomly one swarm population DPSO with population size s_{DPSO} .

Step 2. Initialize randomly the other swarm population CPSO with population size s_{CPSO} .

Step 3. Run the DPSO for g_DPSO generations.

Step 4. Reevaluate the personal best values for the CPSO if it is not the first cycle.

Step 5. Run the CPSO for g_CPSO generations.

Step 6. Reevaluate the personal best values for the DPSO.

Step 7. Go back to Step 3. Repeat this procedure until a termination criterion is reached.

In the above coevolutionary PSO scheme, when one PSO is running, the other PSO serves as its ecological environment; that is, for each PSO its ecological environment has varied from iteration to iteration. Therefore, the personal best obtained in the previous iteration has to be reevaluated in accordance with the new ecological environment before playing its coevolving role. It is also worth noting that, in each generation of the coevolution, the real weights are normalized so that the selected individual basic kernels are combined using a weighted average. Hence, this study uses a quite simple normalization scheme as follows:

$$\tilde{\lambda}_h = \frac{\lambda_h \rho_h}{\sum_{h=1}^7 \lambda_h \rho_h}. \quad (21)$$

4. Hybrid Prognostic Approach for RUL Prediction

As a lithium-ion battery ages, its maximum capacity begins to deteriorate over time. If the maximum capacity falls below 80% of its initial rated capacity, the battery is considered to be unable to provide reliable power supplies and needs to be replaced. In the current academia/industry practices, reliability of a lithium-ion battery for providing reliable power supplies is ensured via the prediction of the remaining maximum capacity. In this study, a hybrid prognostic approach that integrates selective kernel ensemble-based RVM learning algorithm and exponential regression is proposed for RUL prediction of lithium-ion batteries. Figure 1 shows an overall flowchart of the proposed hybrid prognostic approach.

4.1. Capacity Degradation Data Collection. To develop the degradation model, four lithium-ion batteries under test went through the full charge and discharge procedure. These four lithium-ion batteries in the following text are referred to as A1, A2, A3, and A4, respectively. Noting that, these batteries have a graphite anode and a lithium cobalt oxide cathode which were verified using electron dispersive spectroscopy (EDS). The rated capacity of the tested lithium-ion battery was 0.9 Ah. Multiple charge-discharge tests were performed with an Arbin BT2000 battery testing system under ambient temperature (around 25°C). The discharge current was 0.45 A. Cut-off voltage was 2.5 V. The failure threshold of the lithium-ion batteries was 0.72 Ah. The discharge capacity was recorded after each full charge-discharge cycle. Herein, N successive capacity degradation measurements are denoted as $\mathbf{Q} = (q_1, q_2, \dots, q_N)^T$ and their corresponding cycles are $\mathbf{I} = (1, 2, \dots, N)^T$. Therefore, the battery capacity degradation condition can be monitored through the measurements of the input-target pairs $\{l, q_l\}$, where $l = 1, 2, \dots, N$. Inspection of the battery capacity degradation state is made at the predefined inspection cycles of (C_1, C_2, \dots, C_J) , respectively. At each predefined inspection cycle C_i , along the developing lithium-ion battery degradation-to-failure trajectory, the last input-target pair of $\{l, q_l\}$ ($l = C_i$) is recorded and appended to the vector of the input-target pairs $\{l, q_l\}$ ($l = 1, 2, \dots, C_i - 1$)

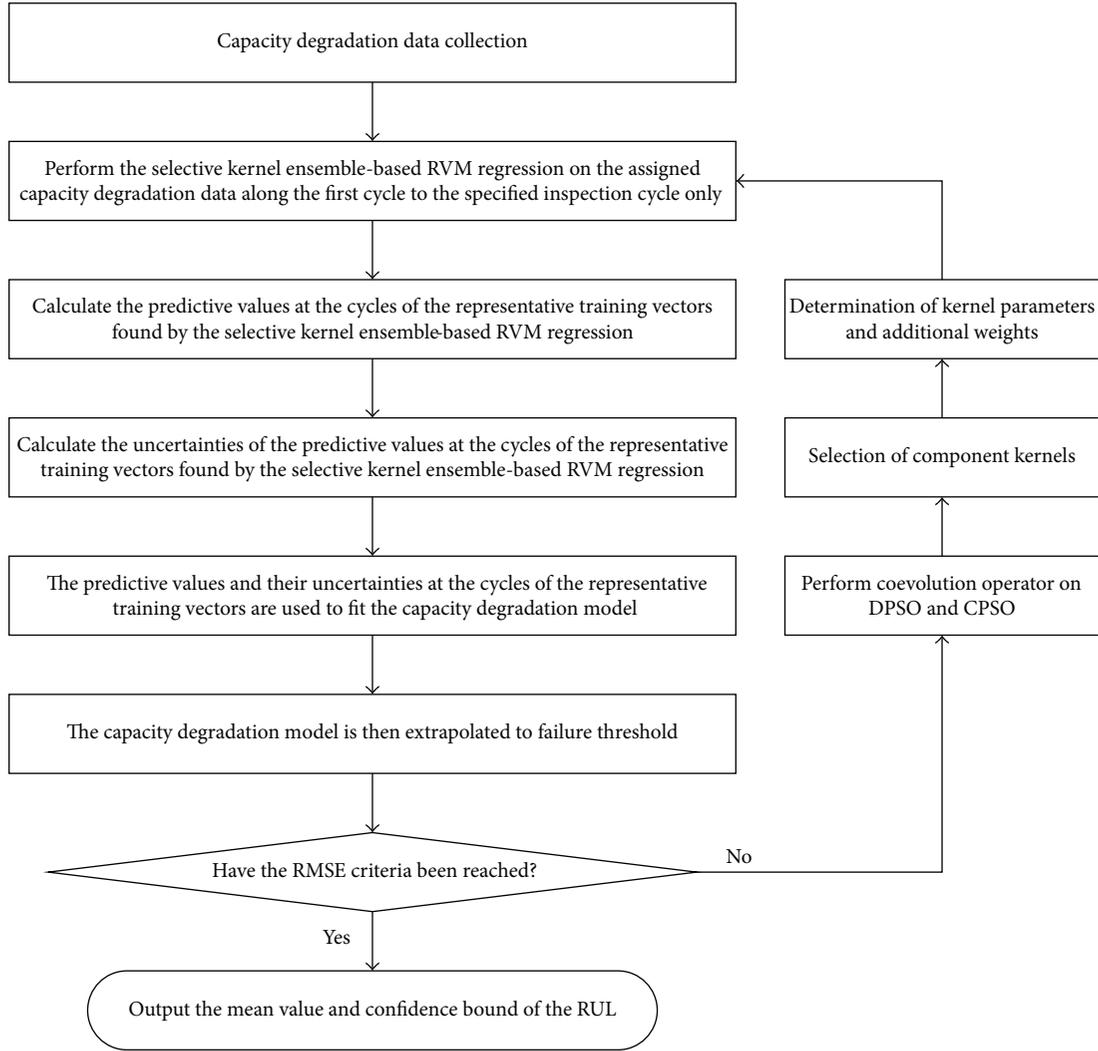


FIGURE 1: Flowchart of the proposed hybrid prognostic approach for RUL prediction of lithium-ion batteries.

collected at the previous $(C_i - 1)$ inspections, so that the capacity degradation data used for degradation model development at the inspection cycle C_i is collected.

4.2. Degradation Model Formulation. At each inspection cycle C_i , the selective kernel ensemble-based RVM learning algorithm is performed on the N available input-target pairs of data $\{l, q_l\}$, where $l = 1, 2, \dots, N$, and thus the $M \leq N$ most representative input-target pairs of data $\{l_m^*, q_{lm}^*\}$ identified by the RVM regression, that is, the relevant vectors whose corresponding basis functions are associated with the remaining nonzero weights, are collected in a sparse dataset, where $m = 1, 2, \dots, M$ and y_{lm}^* is the capacity estimate provided by the RVM in correspondence with l_m^* . For convenience of expression, the sparse dataset is referred to as $\{\bar{l}_M^*, \bar{q}_{lM}^*\}$. Then, fitting to the sparse dataset $\{\bar{l}_M^*, \bar{q}_{lM}^*\}$ is performed to identify the unknown parameters of the model adopted. Finally, the fitted model is extrapolated up to the predefined failure threshold of 0.72 Ah to predict the RUL at inspection cycle C_i , $RUL(C_i)$.

An important issue in developing a capacity degradation model is determining the fitted model, which influences substantially the prognostics performance of the proposed hybrid prognostic approach. The appropriate fitted model depends on the battery under consideration. Goebel et al. [31] used a sum of two exponential functions to model the increase of internal impedance due to solid-electrolyte interface thickening with time. As battery capacity degradation is closely related to the internal impedance increase, potential models for capacity degradation can also be exponential models. Following up Goebel et al. [31] work, He et al. [2] have experimentally demonstrated that the sum of two exponential functions can well describe the capacity degradation trend of many different batteries:

$$q_l = a \cdot \exp(b \cdot l) + c \cdot \exp(d \cdot l), \quad (22)$$

where q_l is the capacity of the battery at the cycle l ; a and b are the parameters associated with the internal impedance; and c and d are the parameters associated with the aging rate. To demonstrate the suitability of the model in (22)

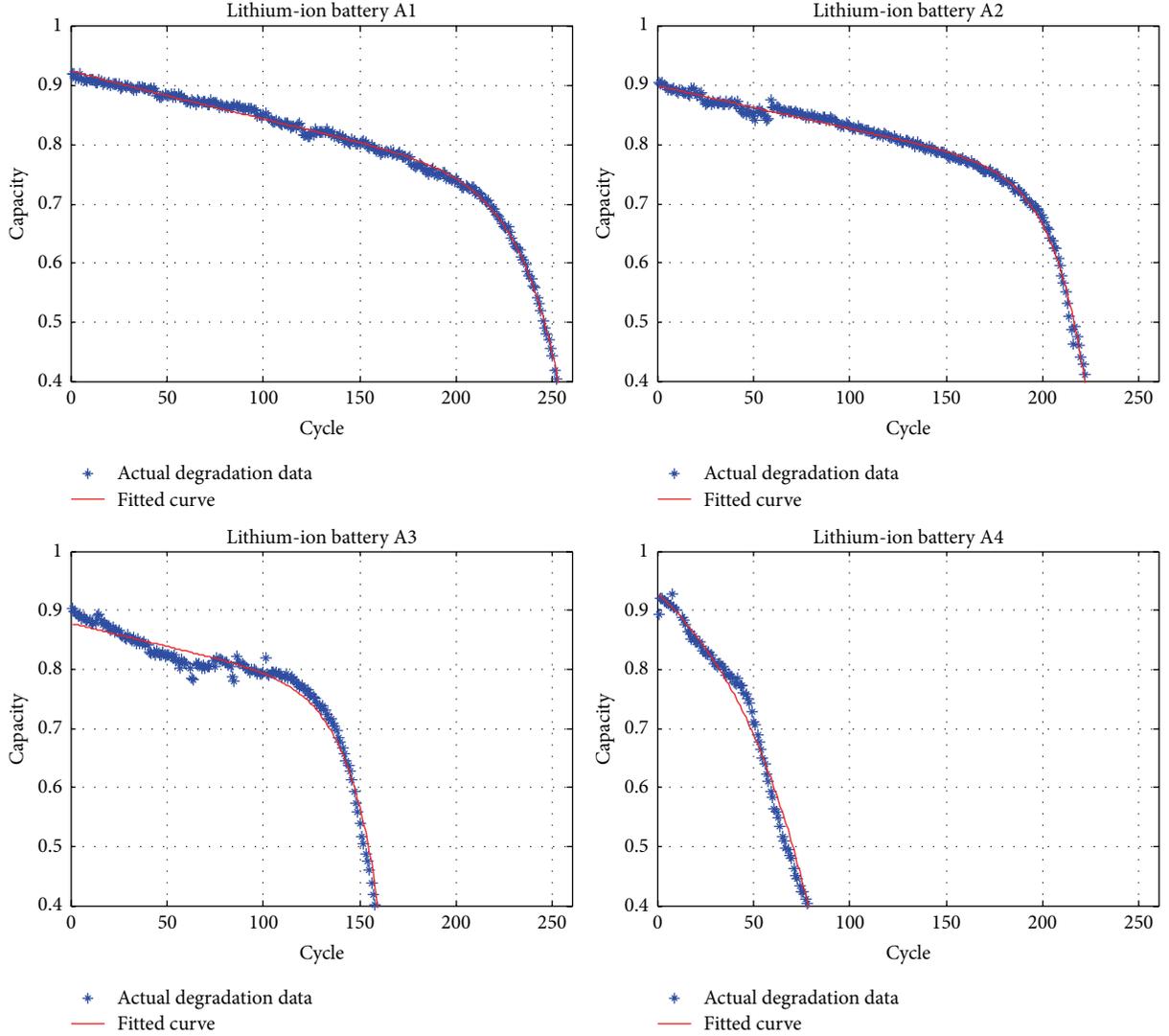


FIGURE 2: Degradation data and fitting curves of the four lithium-ion batteries A1, A2, A3, and A4.

in depicting battery degradation being addressed, Figure 2 shows the curve fitting result (solid line) of (22) to the capacity data of these four batteries, which indicates that the sum of two exponential functions closely agrees with the values of the measured capacity degradation data. Hence, in this study, the sum of two exponential functions was used to fit the degradation curves of the lithium-ion batteries on the basis of the sparse dataset $\{\bar{l}_M^*, \bar{q}_{lM}^*\}$.

4.3. RUL Prediction. The RUL of lithium-ion batteries can be obtained by extrapolating the fitted model to a predefined failure threshold. The predicted RUL at the inspection cycle C_i is then derived by projecting the M state estimates, namely, $\mu^T \varphi(l_1^*), \mu^T \varphi(l_2^*), \dots, \mu^T \varphi(l_{M-1}^*)$ and $\mu^T \varphi(l_M^*)$, into the future until the future cycle C_θ at which the predictive future capacity degradation value hits the predefined failure threshold. Thus, the predicted RUL at the inspection cycle C_i

can be expressed as the differences between the inspection cycle C_i and the future cycle C_θ and can be calculated as

$$\text{RUL}(C_i) = C_\theta - C_i. \quad (23)$$

4.4. Experiment and Results. In order to demonstrate the performance of the proposed hybrid prognostic approach for on-line prediction of lithium-ion battery RUL, four lithium-ion batteries A1, A2, A3, and A4 are employed in this experiment. These four illustrative batteries are in exactly the same experimental environments. For detailed information on these four lithium-ion batteries, please refer to Section 4.1. The battery capacity data used in this study is provided by the Center for Advanced Life Cycle Engineering, University of Maryland [32].

4.4.1. Relative Parameter Settings. This study uses the root mean squared error (RMSE) as a measure of accuracy

index to evaluate the performance of the proposed hybrid prognostic approach. The closer the value of the RMSE is to 0, the better the performance of the fitted model (i.e., (22)) is. Let $\text{RMSE}_{\Theta\Psi}$ denote the RMSE of the fitted model enabled by the sparse dataset $\{\bar{l}_M^*, \bar{q}_{LM}^*\}$ which was found by the selective kernel ensemble-based RVM with the selection vector Θ and the general parameter vector Ψ . Thus, $f(\Theta, \Psi) = 1/\text{RMSE}_{\Theta\Psi}$ is taken as the fitness function of the proposed hybrid prognostic approach.

In order to apply the developed coevolution of DPSO and CPSO for equipping the RVM with adaptive kernel selections, adaptive kernel combinations, and adaptive kernel parameters optimization, after a small number of simple trials relative parameter settings are determined. For the step of selection of candidate basic kernels, parameters of DPSO are set as follows: s_DPSO, 30; acceleration coefficients c_1 and c_2 , 1.0 and 0.5; initial and final inertia weight, 0.9 and 0.2; initial and final inertia velocity, 4 and -4 ; fitness function, $f(\Theta, \Psi) = 1/\text{RMSE}_{\Theta\Psi}$. For the step of determination of kernel parameters and additional weight coefficients, parameters of CPSO are set as follows: s_CPSO, 60; acceleration coefficients c_1 and c_2 , 1.0 and 0.5; initial and final inertia weight, 0.9 and 0.4; initial and final inertia velocity, 4 and -4 ; fitness function, $f(\Theta, \Psi) = 1/\text{RMSE}_{\Theta\Psi}$. For the step of coevolution of DPSO and CPSO, parameters of coevolutionary PSO are set as follows: g_DPSO: 10; g_CPSO: 15; i_max_DPSO: 50.

4.4.2. Experimental Results. To show the robustness of the proposed hybrid prognostic approach, we will perform four independent experiments for each battery. For this purpose, four different inspection cycles corresponding to 60%, 70%, 80%, and 90% data partition rates have been used for on-line prediction of lithium-ion battery RUL. That is, the inspection cycle was set at the cycle steps that separate the whole battery capacity degradation data available into two parts, where the first 60%, 70%, 80%, and 90% are used for RVM learning. Here we assume that the RVM learning with less than 60% battery capacity degradation data is inadequate.

The first case refers to the lithium-ion battery A1. The proposed hybrid prognostic approach is applied to the lithium-ion battery A1 capacity degradation data plotted with dots in Figures 3–6. As aforementioned in Section 4.4.1, four different inspection cycles corresponding to 60%, 70%, 80%, and 90% data partition rates have been used for on-line prediction of lithium-ion battery RUL; that is, predictions of the RUL of the lithium-ion battery A1 are calculated at the inspection cycles of 125, 146, 167, and 188, respectively. Table 1 summarizes the predicted RUL and the actual RUL when the inspection cycles of 125, 146, 167, and 188 were chosen for battery A1. In the second case, the lithium-ion battery A2 is investigated. The proposed hybrid prognostic approach is then conducted on the lithium-ion battery A2 capacity degradation data, which are plotted as the dots in Figures 7–10. Predictions of the RUL of the lithium-ion battery A2 are calculated at the inspection cycles of 113, 132, 151, and 170, respectively. Table 2 summarizes the predicted RUL and the actual RUL when the inspection cycles of 113, 132, 151, and 170

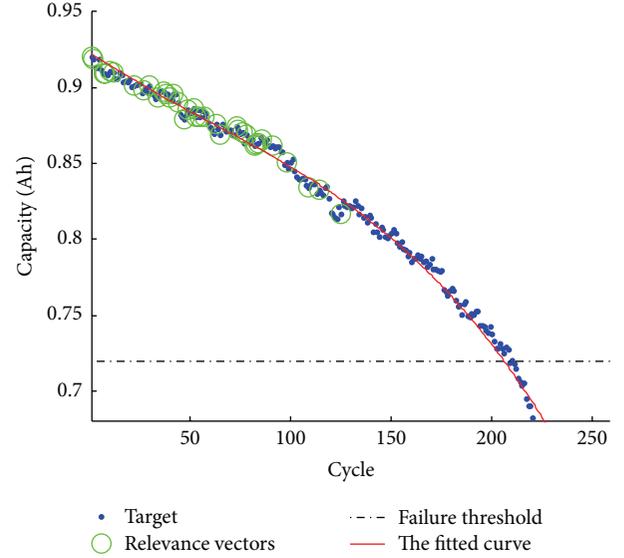


FIGURE 3: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 125 for lithium-ion battery A1.

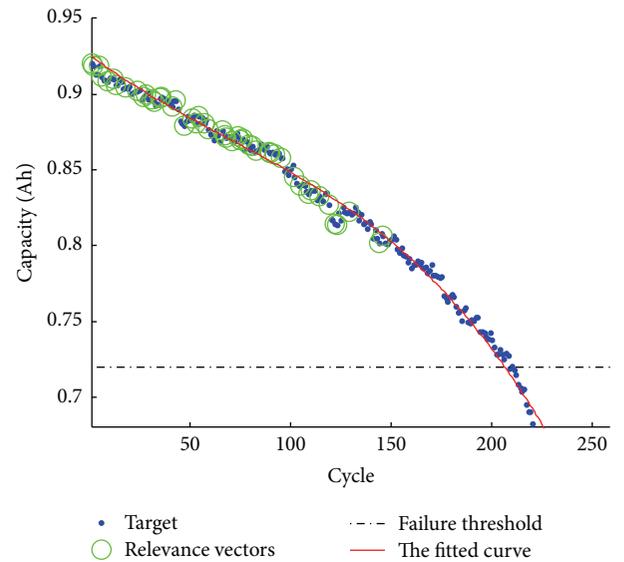


FIGURE 4: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 146 for lithium-ion battery A1.

were chosen for battery A2. In the third case, a lithium-ion battery A3 is investigated. Battery A3 capacity degradation data, plotted with the dots in Figures 11–14, are analyzed by the proposed hybrid prognostic approach. Predictions of the RUL of the lithium-ion battery A3 are calculated at the inspection cycles of 79, 92, 106, and 119, respectively. Table 3 summarizes the predicted RUL and the actual RUL when the inspection cycles of 79, 92, 106, and 119 were chosen for battery A3. In the fourth case, a lithium-ion battery A4 is investigated. Battery A4 capacity degradation data, plotted with the dots in Figures 15–18, are analyzed by the proposed hybrid prognostic approach. Predictions of the RUL of the lithium-ion battery A4 are calculated at the inspection cycles

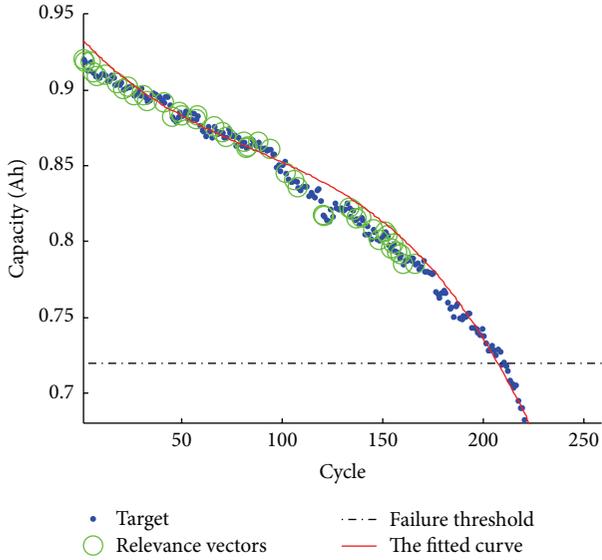


FIGURE 5: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 167 for lithium-ion battery A1.

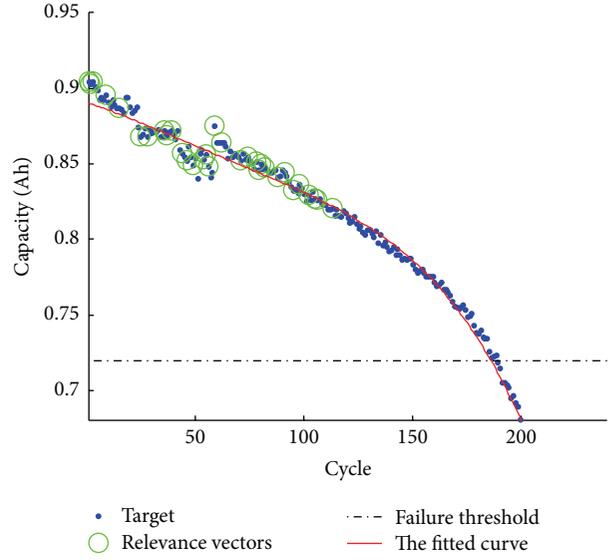


FIGURE 7: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 113 for lithium-ion battery A2.

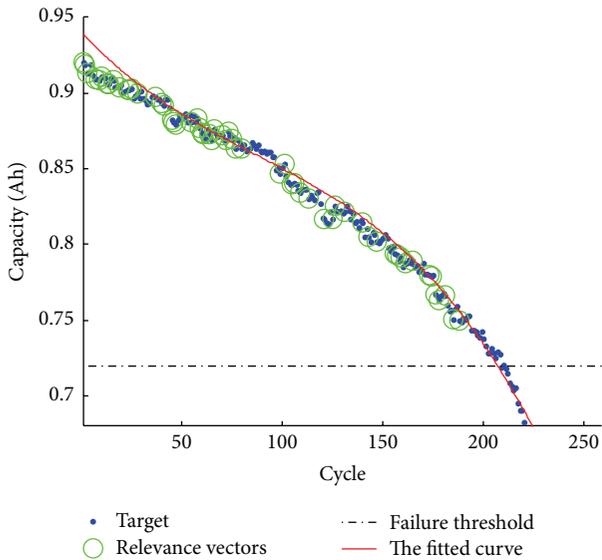


FIGURE 6: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 188 for lithium-ion battery A1.

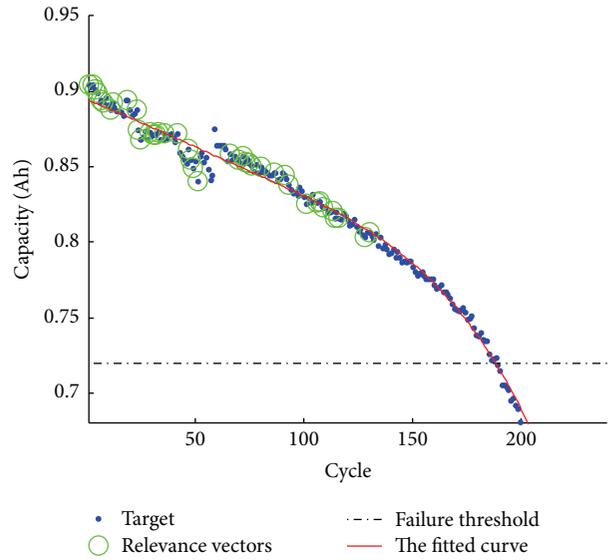


FIGURE 8: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 132 for lithium-ion battery A2.

of 29, 34, 38, and 43, respectively. Table 4 summarizes the predicted RUL and the actual RUL when the inspection cycles of 29, 34, 38, and 43 were chosen for battery A4. Note that the relevant vectors are highlighted by the circles in Figures 3–18. Also note that the parameter values are obtained by fitting the capacity degradation values predicted by the selective kernel ensemble-based RVM at the cycles of the representative training vectors. As seen in Figures 3–18, the proposed hybrid prognostic approach can effectively identify the lithium-ion battery capacity degradation trajectory, except in the third case of inspection number = 79 and the fourth case of inspection number = 29. The results in Tables 1–4 also revealed the good agreement of the predicted RUL and the

actual RUL. It can be concluded from Figures 3–18 and Tables 1–4 that the proposed hybrid prognostic approach may be a promising tool for lithium-ion battery RUL prediction.

5. Effects of RVM Learning on Hybrid Prognostic Approach

In order to investigate the effects of RVM learning on the proposed hybrid prognostic approach, comparison of the proposed hybrid prognostic approach using the selective kernel ensemble-based RVM learning algorithm with the hybrid prognostic approaches using the single kernel-based RVM learning algorithm and the Ensemble All-based RVM

TABLE 1: Values of the predicted RUL and the actual RUL for lithium-ion battery A1.

Inspection cycle (C_i)	Kernel mixed coefficient							Kernel parameters							Predicted RUL (C_i)	Actual RUL (C_i)	
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	σ_{Gau}	σ_{Exp}	σ_{Lap}	d	ξ	η	σ_{cau}			τ
125	0.4	N/A	N/A	0.4	0.2	N/A	N/A	5.56	N/A	N/A	0.98	109.2	8795	N/A	N/A	83	84
146	0.4	N/A	N/A	0.4	0.2	N/A	N/A	5.56	N/A	N/A	0.98	120.3	89.95	N/A	N/A	63	63
167	0.4	N/A	N/A	0.4	0.2	N/A	N/A	5.56	N/A	N/A	0.98	127.7	89.95	N/A	N/A	42	42
188	0.4	N/A	N/A	0.4	0.2	N/A	N/A	12.59	N/A	N/A	0.98	129.7	89.95	N/A	N/A	20	21

TABLE 2: Values of the predicted RUL and the actual RUL for lithium-ion battery A2.

Inspection cycle (C_i)	Kernel mixed coefficients							Kernel parameters							Predicted RUL (C_i)	Actual RUL (C_i)	
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	σ_{Gau}	σ_{Exp}	σ_{Lap}	d	ξ	η	σ_{Cau}			τ
113	0.6	N/A	N/A	0.3	0.1	N/A	N/A	2.85	N/A	N/A	9.75	100.28	81.43	N/A	N/A	74	76
132	0.6	N/A	N/A	0.3	0.1	N/A	N/A	2.98	N/A	N/A	9.78	70.28	81.43	N/A	N/A	58	57
151	0.6	N/A	N/A	0.3	0.1	N/A	N/A	2.98	N/A	N/A	8.79	70.28	71.43	N/A	N/A	39	38
170	0.6	N/A	N/A	0.3	0.1	N/A	N/A	2.98	N/A	N/A	8.79	129.7	89.95	N/A	N/A	20	19

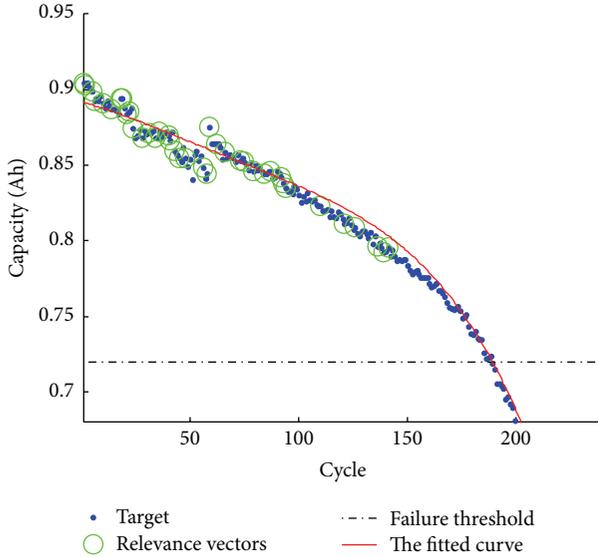


FIGURE 9: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 151 for lithium-ion battery A2.

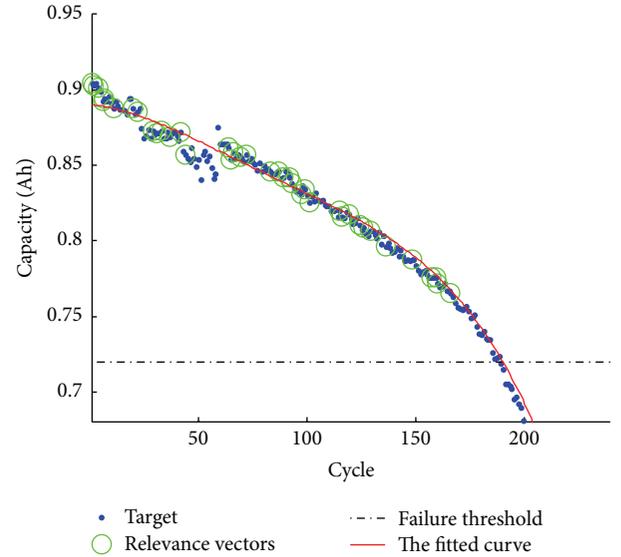


FIGURE 10: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 170 for lithium-ion battery A2.

learning algorithm was made in this section. It should be noted that herein the best kernel-based RVM learning algorithm means that the RVM only adopts the best performing component kernel among all available basic kernels for supervised learning, while the Ensemble All-based RVM learning algorithm means that the RVM adopts all of those available basic kernels for supervised learning. The results are summarized in Table 5 in terms of the predicted RUL and actual RUL. As can be seen, the proposed hybrid prognostic approach using the selective kernel ensemble-based RVM learning algorithm performed better (and in most cases substantially better) than the compared hybrid prognostic approach using the best kernel-based RVM learning algorithm. It can also be concluded from Table 5 that the proposed hybrid prognostic approach outperformed the

compared hybrid prognostic approach using the Ensemble All-based RVM learning algorithm, even though selective kernel ensemble-based RVM learning algorithm only uses a far smaller number of basic kernels. Taking the lithium-ion battery A1, for example, the size of the selective kernel ensemble-based RVM learning algorithm is about only 43% (3.0/7.0), 43% (3.0/7.0), 43% (3.0/7.0), and 43% (3.0/7.0) of the size of the Ensemble All-based RVM learning algorithm for four inspection cycles of 125, 146, 167, and 188, respectively. Significant improvement obtained not only demonstrates better generalization performance of selective kernel ensemble-based RVM learning algorithm but also proves the feasibility and necessity of removing redundant basic kernels in Ensemble All-based RVM learning algorithm. Therefore, the step of eliminating some unacceptable basic kernels and

TABLE 3: Values of the predicted RUL and the actual RUL for lithium-ion battery A3.

Inspection cycle (C_i)	Kernel mixed coefficient							Kernel parameters							Predicted RUL (C_i)	Actual RUL (C_i)	
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	σ_{Gau}	σ_{Exp}	σ_{Lap}	d	ξ	η	σ_{Cau}			τ
79	N/A	0.8	N/A	N/A	0.1	N/A	0.1	N/A	100.4	N/A	N/A	3.457	2.178	N/A	9.866	46	53
92	N/A	0.8	N/A	N/A	0.1	N/A	0.1	N/A	1.33	N/A	N/A	3.9209	3.987	N/A	59.91	38	40
106	0.67	0.165	N/A	N/A	0.165	N/A	N/A	0.1003	1.1895	N/A	N/A	150.2544	108.6072	N/A	N/A	24	26
119	0.67	0.165	N/A	N/A	0.165	N/A	N/A	0.10	0.1081	N/A	N/A	110.254	108.608	N/A	N/A	13	13

TABLE 4: Values of the predicted RUL and the actual RUL for lithium-ion battery A4.

Inspection cycle (C_i)	Kernel mixed coefficient							Kernel parameters							Predicted RUL (C_i)	Actual RUL (C_i)	
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	σ_{Gau}	σ_{Exp}	σ_{Lap}	d	ξ	η	σ_{Cau}			τ
29	0.8	N/A	N/A	0.1	0.1	N/A	N/A	1.79	N/A	N/A	1.29	0.58	2.178	N/A	N/A	16	19
34	0.8	N/A	N/A	0.1	0.1	N/A	N/A	0.42	N/A	N/A	1.68	0.59	0.84	N/A	N/A	14	14
38	0.8	N/A	N/A	0.1	0.1	N/A	N/A	0.45	N/A	N/A	0.48	0.59	0.84	N/A	N/A	10	10
43	0.8	N/A	N/A	0.1	0.1	N/A	N/A	0.21	N/A	N/A	0.58	0.49	0.53	N/A	N/A	5	5

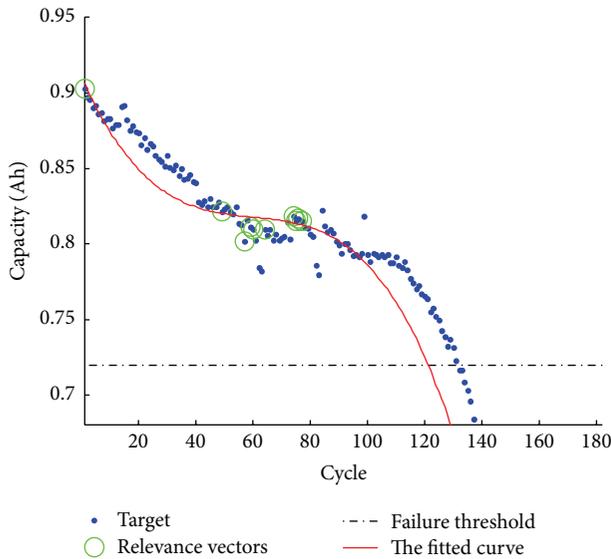


FIGURE 11: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 79 for lithium-ion battery A3.

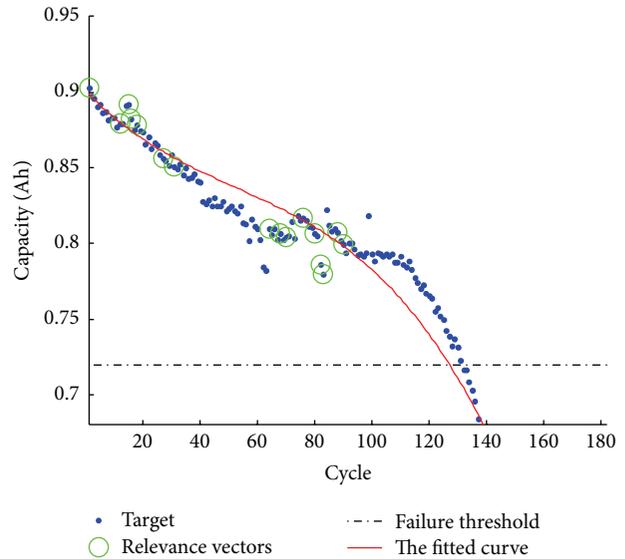


FIGURE 12: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 92 for lithium-ion battery A3.

meanwhile selecting several acceptable ones to construct a basic kernel combination plays a crucial role in enhancing the generalization capability of RVM.

6. Comparison with Existing Approaches

In this section, performances of the proposed hybrid prognostic approach using the selective kernel ensemble-based RVM learning algorithm were compared with those of the hybrid prognostic approaches using the popular algorithms of feedforward ANNs like the conventional BP algorithm and SVMs on four lithium-ion batteries A1, A2, A3, and A4. Although there are many variants of BP algorithm, a faster BP algorithm called Levenberg-Marquardt algorithm is used in ANNs. The activation function used is a simple sigmoidal

function $g(x) = 1/(1 + \exp(-x))$, whereas the kernel function used in SVMs is Gaussian.

Table 6 presents the results in the test set for the hybrid prognostic approaches using the selective kernel ensemble-based RVM learning algorithm, ANN, and SVM. We show the RUL prediction result. Also, we include the number of relevant/support vectors in the learning algorithms, which is related to prognostic model structural complexity (sparsity). As seen in Table 6, the proposed hybrid prognostic approach using the selective kernel ensemble-based RVM learning algorithm can provide more accurate RUL prediction of degraded lithium-ion batteries than the hybrid prognostic approaches using ANN and SVM can. This is mainly because (1) the selective kernel ensemble-based RVM learning algorithm can overcome some shortcomings of the ANN, such

TABLE 5: Comparison of the hybrid prognostic approach's performances using different RVM learning algorithms.

Battery name	Inspection cycle (C_i)	Predicted RUL (C_i)			Actual RUL (C_i)
		SKE	BK	EA	
A1	125	83	88	90	84
	146	63	60	67	63
	167	42	43	46	42
	188	20	23	25	21
A2	113	74	81	78	76
	132	58	61	60	57
	151	39	42	44	38
	170	20	23	24	19
A3	79	46	48	57	53
	92	38	36	35	40
	106	24	23	30	26
	119	13	10	16	13
A4	29	16	14	15	19
	34	14	16	12	14
	38	10	12	10	10
	43	5	4	6	5

SKE: the proposed hybrid prognostic approach that integrates selective kernel ensemble-based RVM and exponential regression; BK: another hybrid prognostic approach that integrates the best kernel-based (i.e., the best performing component kernel among all available basic kernels) RVM with exponential regression; EA: the other hybrid prognostic approach that integrates the Ensemble All-based (i.e., combining all of those available basic kernels) RVM with exponential regression.

TABLE 6: Comparison of the predicted RUL and sparse results for four lithium-ion batteries.

Battery name	Inspection cycle (C_i)	Predicted RUL (C_i)			Actual RUL (C_i)	Relevant/support vectors		
		SKE-RVM	SVM	BP		SKE-RVM	SVM	BP
A1	125	83	89	96	84	42	54	—
	146	63	66	71	63	46	62	—
	167	42	47	51	42	61	70	—
	188	20	25	27	21	60	75	—
A2	113	74	83	90	76	30	51	—
	132	58	61	66	57	40	55	—
	151	39	41	44	38	61	83	—
	170	20	25	27	19	52	86	—
A3	79	46	61	66	53	11	32	—
	92	38	45	48	40	23	38	—
	106	24	22	24	26	4	17	—
	119	13	15	16	13	4	15	—
A4	29	16	14	15	19	5	12	—
	34	14	16	18	14	4	14	—
	38	10	8	13	10	4	15	—
	43	5	6	7	5	6	18	—

SKE-RVM: selective kernel ensemble-based RVM learning algorithm.

as overfitting and local minima; (2) the selective kernel ensemble-based RVM learning algorithm used the selective convex combination and thus the RVM has stronger generalization capability; (3) the evolution of kernel parameters and kernel weights via CPSO can improve the generalization performance of the selective kernel ensemble-based RVM learning algorithm. Moreover, the selective kernel ensemble-based RVM learning algorithm also provides the capability

of yielding a decision function that is much sparser than SVM; that is, the number of relevant vectors is much smaller than that of support vectors, while maintaining the prediction accuracy. This can lead to a significant reduction in prognostic model structural complexity, thereby making it more suitable for on-line real-time RUL prediction. In addition, RVM does not need the tuning of a soft margin parameter (C) as in SVM during the training phase.

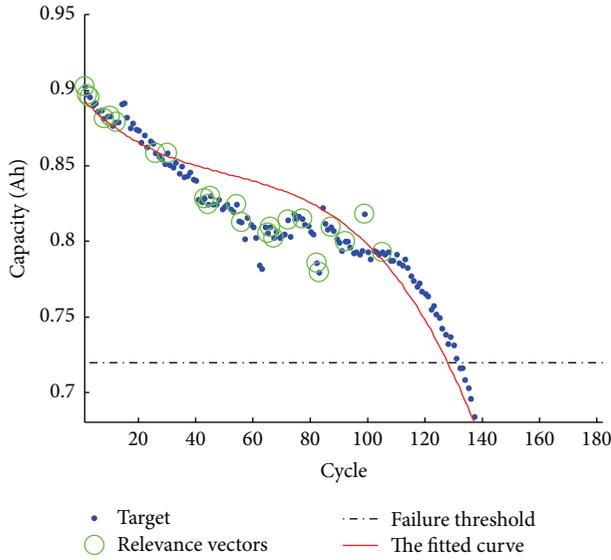


FIGURE 13: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 106 for lithium-ion battery A3.

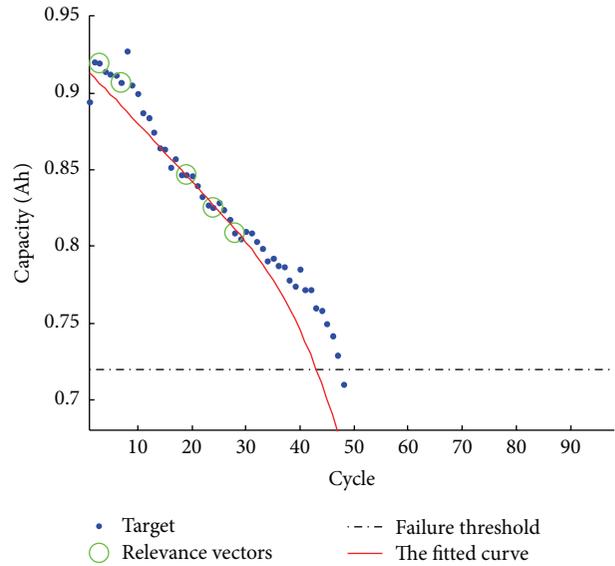


FIGURE 15: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 29 for lithium-ion battery A4.

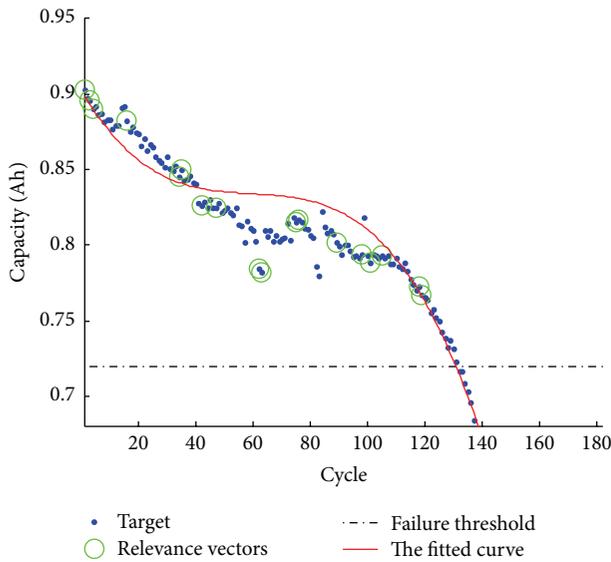


FIGURE 14: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 119 for lithium-ion battery A3.

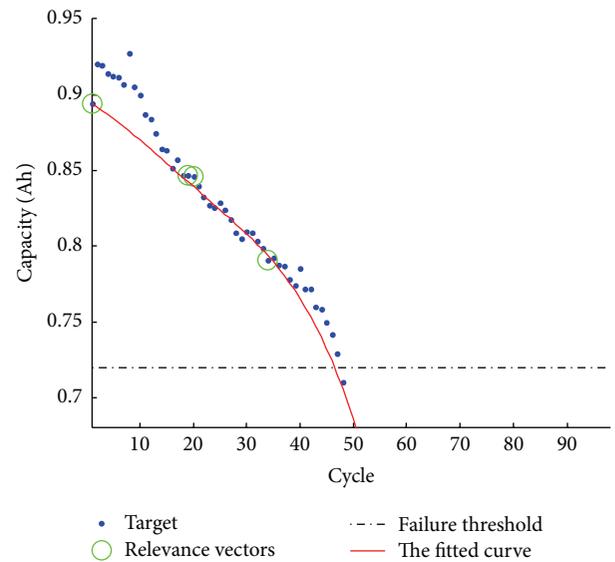


FIGURE 16: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 34 for lithium-ion battery A4.

7. Conclusions

Lithium-ion battery is a core component of many systems and is critical to the systems' functional capabilities. Battery failure could lead to reduced performance, operational impairment, and even catastrophic failure, especially in aerospace systems. Therefore, RUL prediction of degraded lithium-ion batteries is very helpful for preventing long-term breakdown or catastrophic failure. Using selective kernel ensemble-based relevance vector machine and exponential regression together, an automatic, effective but simpler-to-use hybrid prognostic approach was proposed for RUL prediction of degraded lithium-ion batteries. Four lithium-ion batteries,

namely, A1, A2, A3, and A4, were considered in this study. The experimental result indicated that the proposed hybrid prognostic approach can accurately predict the RUL of degraded lithium-ion batteries. Empirical comparisons showed that the proposed hybrid prognostic approach performed better than the hybrid prognostic approaches using the popular algorithms of feedforward ANNs like the conventional BP algorithm and SVMs. This study also demonstrates that the hybrid prognostic approach using the selective kernel ensemble-based RVM learning algorithm outperformed the hybrid prognostic approaches using the single kernel-based RVM learning algorithm and the Ensemble All-based RVM learning algorithm.

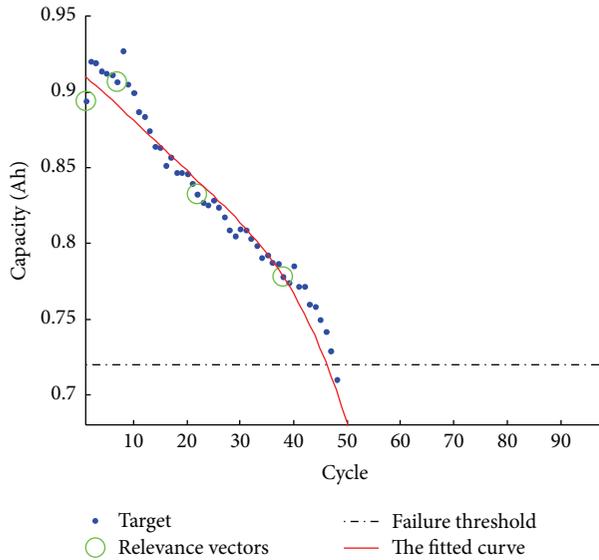


FIGURE 17: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 38 for lithium-ion battery A4.

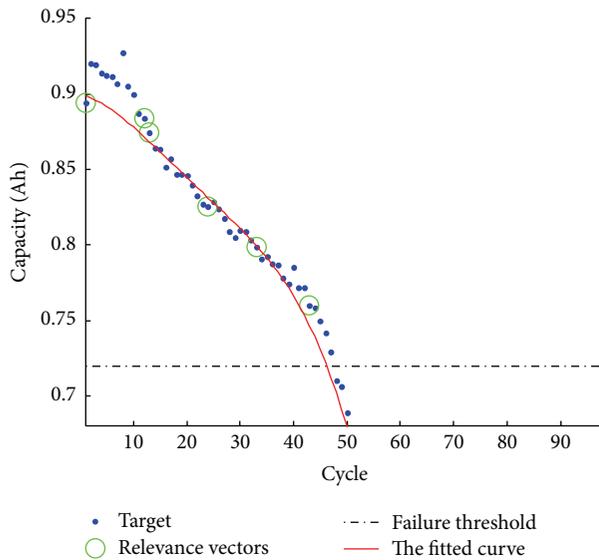


FIGURE 18: Predictive results obtained by the hybrid prognostic approach at inspection cycle of 43 for lithium-ion battery A4.

Three research directions are worth pursuing. First, although this study considers the application of lithium-ion batteries, the proposed hybrid prognostic approach can be modified and extended to other types of battery (e.g., li-polymer battery [33], lead-acid battery [34]). Second, bearings, gearboxes, and oil sand pumps are core components of all kinds of machinery, and maintenance of bearings, gearboxes, and oil sand pumps is essential. The proposed hybrid prognostic approach can also be extended to deal with gear, bearing, and oil sand pump performance degradation assessment and RUL estimation [35–38]. Third, in this study, only seven basic kernels were adopted. Hence, including other types of basic kernels (e.g., Wavelet kernel [39], Cauchy

kernel [40], and Rational Quadratic kernel [41]) in the RVM learning is another further research direction that is also worth pursuing.

Competing Interests

The authors declare that there are no competing interests regarding the publication of this paper.

Acknowledgments

The research is funded partially by the National Science Foundation of China (51405239), National Defense Basic Scientific Research Program of China (A2620132010, A2520110003), Fundamental Research Funds for the Central Universities (1005-YAH15055), Jiangsu Provincial Natural Science Foundation of China (BK20150745, BK20140727), Jiangsu Province Science and Technology Support Program (BE2014134), and Jiangsu Postdoctoral Science Foundation of China (1501024C). The authors would like to express sincere appreciation to Professor Pecht and Center for Advanced Life Cycle Engineering, University of Maryland, for their efforts to make battery dataset available and permission to use dataset.

References

- [1] I.-S. Kim, "A technique for estimating the state of health of lithium batteries through a dual-sliding-mode observer," *IEEE Transactions on Power Electronics*, vol. 25, no. 4, pp. 1013–1022, 2010.
- [2] 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," *Journal of Power Sources*, vol. 196, no. 23, pp. 10314–10321, 2011.
- [3] B.-H. Seo, T. H. Nguyen, D.-C. Lee, K.-B. Lee, and J.-M. Kim, "Condition monitoring of lithium polymer batteries based on a sigma-point Kalman filter," *Journal of Power Electronics*, vol. 12, no. 5, pp. 778–786, 2012.
- [4] B. S. Bhangu, P. Bentley, D. A. Stone, and C. M. Bingham, "Non-linear observers for predicting state-of-charge and state-of-health of lead-acid batteries for hybrid-electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 3, pp. 783–794, 2005.
- [5] W. He, N. Williard, M. Osterman, and M. Pecht, "Prognostics of lithium-ion batteries using extended Kalman filtering," in *Proceedings of the International Microelectronics and Packaging Society (IMAPS) Advanced Technology Workshop on High Reliability Microelectronics for Military Applications*, pp. 17–19, Linthicum Heights, Md, USA, September 2011.
- [6] D. Andre, A. Nuhic, T. Soczka-Guth, and D. U. Sauer, "Comparative study of a structured neural network and an extended Kalman filter for state of health determination of lithium-ion batteries in hybrid electricvehicles," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 3, pp. 951–961, 2013.
- [7] B. Saha and K. Goebel, "Modeling Li-ion battery capacity depletion in a particle filtering framework," in *Proceedings of the Annual Conference of the Prognostics and Health Management Society (PHM '09)*, pp. 1–10, San Diego, Calif, USA, October 2009.
- [8] M. Dalal, J. Ma, and D. He, "Lithium-ion battery life prognostic health management system using particle filtering framework,"

- Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 225, no. 1, pp. 81–90, 2011.
- [9] Y. J. Xing, E. W. M. Ma, K.-L. Tsui, and M. Pecht, “An ensemble model for predicting the remaining useful performance of lithium-ion batteries,” *Microelectronics Reliability*, vol. 53, no. 6, pp. 811–820, 2013.
 - [10] Q. Miao, L. Xie, H. J. Cui, W. Liang, and M. Pecht, “Remaining useful life prediction of lithium-ion battery with unscented particle filter technique,” *Microelectronics Reliability*, vol. 53, no. 6, pp. 805–810, 2013.
 - [11] B. Long, W. Xian, L. Jiang, and Z. Liu, “An improved autoregressive model by particle swarm optimization for prognostics of lithium-ion batteries,” *Microelectronics Reliability*, vol. 53, no. 6, pp. 821–831, 2013.
 - [12] A. Nuhic, T. Terzimehic, T. Soczka-Guth, M. Buchholz, and K. Dietmayer, “Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods,” *Journal of Power Sources*, vol. 239, no. 1, pp. 680–688, 2013.
 - [13] C. Weng, Y. Cui, J. Sun, and H. Peng, “On-board state of health monitoring of lithium-ion batteries using incremental capacity analysis with support vector regression,” *Journal of Power Sources*, vol. 235, pp. 36–44, 2013.
 - [14] C. C. Chen and M. Pecht, “Prognostics of lithium-ion batteries using model-based and data-driven methods,” in *Proceedings of the 3rd Annual IEEE Prognostics and System Health Management Conference (PHM '12)*, pp. 1–6, Beijing, China, May 2012.
 - [15] M. E. Tipping, “Sparse Bayesian learning and the relevance vector machine,” *Journal of Machine Learning Research*, vol. 1, no. 3, pp. 211–244, 2001.
 - [16] B. Saha, K. Goebel, S. Poll, and J. Christophersen, “Prognostics methods for battery health monitoring using a Bayesian framework,” *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 2, pp. 291–296, 2009.
 - [17] F. Di Maio, K.-L. Tsui, and E. Zio, “Combining relevance vector machines and exponential regression for bearing residual life estimation,” *Mechanical Systems and Signal Processing*, vol. 31, pp. 405–427, 2012.
 - [18] E. Zio and F. Di Maio, “Fatigue crack growth estimation by relevance vector machine,” *Expert Systems with Applications*, vol. 39, no. 12, pp. 10681–10692, 2012.
 - [19] D. Wang, Q. Miao, and M. Pecht, “Prognostics of lithium-ion batteries based on relevance vectors and a conditional three-parameter capacity degradation model,” *Journal of Power Sources*, vol. 239, pp. 253–264, 2013.
 - [20] T. Benkedjouh, K. Medjaher, N. Zerhouni, and S. Rechak, “Remaining useful life prediction based on nonlinear feature reduction and support vector regression,” *Engineering Applications of Artificial Intelligence*, vol. 26, no. 7, pp. 1751–1760, 2013.
 - [21] T. Benkedjouh, K. Medjaher, N. Zerhouni, and S. Rechak, “Health assessment and life prediction of cutting tools based on support vector regression,” *Journal of Intelligent Manufacturing*, vol. 26, no. 2, pp. 213–223, 2015.
 - [22] A. J. Smola, *Learning with kernels [Ph.D. thesis]*, TU Berlin, 1998.
 - [23] G. F. Smits and E. M. Jordaán, “Improved SVM regression using mixtures of kernels,” in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '02)*, vol. 3, pp. 2785–2790, IEEE, Honolulu, Hawaii, USA, May 2002.
 - [24] A. T. Quang, Q.-L. Zhang, and X. Li, “Evolving support vector machine parameters,” in *Proceedings of the International Conference on Machine Learning and Cybernetics*, pp. 548–551, Beijing, China, November 2002.
 - [25] M. Girolami and S. Rogers, “Hierarchic bayesian models for kernel learning,” in *Proceedings of the 22nd International Conference on Machine Learning (ICML '05)*, pp. 241–248, August 2005.
 - [26] M. Gönen and E. Alpaydın, “Multiple kernel learning algorithms,” *Journal of Machine Learning Research*, vol. 12, pp. 2211–2268, 2011.
 - [27] S. R. Gunn and J. S. Kandola, “Structural modelling with sparse kernels,” *Machine Learning*, vol. 48, no. 1–3, pp. 137–163, 2002.
 - [28] M. Hu, Y. Chen, and J. T.-Y. Kwok, “Building sparse multiple-kernel SVM classifiers,” *IEEE Transactions on Neural Networks*, vol. 20, no. 5, pp. 827–839, 2009.
 - [29] J. Kennedy and R. C. Eberhart, “A discrete binary version of the particle swarm algorithm,” in *Proceedings of the IEEE International Conference on Computational Cybernetics and Simulation*, pp. 4104–4108, IEEE Press, Orlando, Fla, USA, October 1997.
 - [30] J. Kennedy and R. C. Eberhart, “Particle swarm optimization,” in *Proceedings of the IEEE International Conference on Neural Networks*, vol. 4, pp. 1942–1948, IEEE, Perth, Australia, November–December 1995.
 - [31] K. Goebel, B. Saha, A. Saxena, J. R. Celaya, and J. P. Christophersen, “Prognostics in battery health management,” *IEEE Instrumentation and Measurement Magazine*, vol. 11, no. 4, pp. 33–40, 2008.
 - [32] M. Pecht, *CALCE Battery Data*, University of Maryland, College Park, Md, USA, 2014.
 - [33] B. Saha, E. Koshimoto, C. C. Quach et al., “Predicting battery life for electric UAVs,” in *Proceedings of the AIAA Infotech at Aerospace Conference and Exhibit*, St. Louis, Mo, USA, March 2011.
 - [34] E. Frisk, M. Krysander, and E. Larsson, “Data-driven lead-acid battery prognostics using random survival forests,” in *Proceedings of the Annual Conference of the Prognostics and Health Management Society*, Fort Worth, Tex, USA, September 2014.
 - [35] D. Wang and P. W. Tse, “Prognostics of slurry pumps based on a moving-average wear degradation index and a general sequential Monte Carlo method,” *Mechanical Systems and Signal Processing*, vol. 56, pp. 213–229, 2014.
 - [36] D. Wang, Q. Miao, Q. H. Zhou, and G. W. Zhou, “An intelligent prognostic system for gear performance degradation assessment and remaining useful life estimation,” *Journal of Vibration and Acoustics*, vol. 137, no. 2, Article ID 021004, 2015.
 - [37] D. Wang and C. Shen, “An equivalent cyclic energy indicator for bearing performance degradation assessment,” *Journal of Vibration and Control*, 2014.
 - [38] D. Wang, P. W. Tse, W. Guo, and Q. Miao, “Support vector data description for fusion of multiple health indicators for enhancing gearbox fault diagnosis and prognosis,” *Measurement Science and Technology*, vol. 22, no. 2, Article ID 025102, 2011.
 - [39] L. Zhang, W. Zhou, and L. Jiao, “Wavelet support vector machine,” *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, vol. 34, no. 1, pp. 34–39, 2004.
 - [40] J. Basak, “A least square kernel machine with box constraints,” in *Proceeding of the IEEE 19th International Conference on Pattern Recognition (ICPR '08)*, pp. 1–4, IEEE, Tampa, Fla, USA, December 2008.
 - [41] D. K. Duvenaud, *Automatic model construction with Gaussian processes [Ph.D. thesis]*, University of Cambridge, 2014.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

