IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 21, 2017, accepted June 7, 2017, date of publication June 16, 2017, date of current version July 17, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2716353

Capacity Prognostics of Lithium-Ion Batteries using EMD Denoising and Multiple Kernel RVM

CHAOLONG ZHANG[1,2], YIGANG HE[2], LIFENG YUAN[2], AND SHENG XIANG[2]

[1]School of Physics and Electronic Engineering, Anqing Normal University, Anqing 246011, China
[2]School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China

Corresponding author: Chaolong Zhang (zhangchaolong@126.com) and Yigang He (18655136887@163.com)

This work was supported by the National Natural Science Foundation of China under Grant 51577046 and Grant 51607004, in part by the State Key Program of National Natural Science Foundation of China under Grant 51637004, in part by the National Key Research and Development Plan ''Important Scientific Instruments and Equipment Development'' under Grant 2016YFF0102200, in part by the Anhui Provincial Natural Science Foundation under Grant 1608085QF157, in part by the Key Projects of Anhui Province University Outstanding Youth Talent Support Program under Grant gxyqZD2016207, and in part by the Anhui Provincial College Student's Creative Lab Construction Plan under Grant 2016ckjh112.

ABSTRACT Lithium-ion batteries are crucial to many types of electric equipments. Hence, lithium-ion battery capacity prognostic is significantly important, and it is yet very hard for the measured battery data that are regularly polluted by miscellaneous noises. In this paper, a battery capacity prognostic approach using the empirical mode decomposition (EMD) denoising method and multiple kernel relevance vector machine (MKRVM) approach is presented. The EMD denoising method is employed to process the measured capacity data to produce noise-free capacity data. The battery capacity prediction model using MKRVM is constructed based on the denoised capacity data. The MKRVM's kernel keeps diversity by using multiple heterogeneous kernel learning method. Meanwhile, sparse weights of basic kernel functions are yielded by using particle swarm optimization (PSO) algorithm. The measured battery capacity data are used to demonstrate the effect of EMD denoising method, and battery capacity prediction experiments reveal that the proposed MKRVM approach can predict the battery's future capacity precisely.

INDEX TERMS Lithium-ion battery, capacity prognostic, EMD denoising, MKRVM, PSO.

I. INTRODUCTION

Lithium-ion batteries are the crucial components in many types of electric equipments including mobile phones, electric vehicles, portable computers, and other portable electric equipments for their long cycle life, light weight, high cell voltage, high power density and no memory effect. However, irreversible physical and chemical diversifications in lithium-ion batteries occur with the battery's aging and usage. Therefore, the state of health (SOH) of the battery decreases by degrees until the battery failure eventually, which usually makes the portable electric equipments unusable. As a result, many works focus on prognostics and health management (PHM) of lithium-ion batteries in recent years [1]–[28].

Lithium-ion battery's PHM includes predicting lithium-ion battery's future capacity and estimating its remaining useful life (RUL). Model-based method and data-driven method are normally used battery capacity prognostic approaches. Model-based method adopts mathematical representation or failure physics model in order to reflect the understanding of battery capacity degradation and battery failure. Kalman filtering [7]–[9], physics-based model [10], non-linear model [11], [12], sparse Bayesian predictive model [13] and particle filtering [14]–[19] are currently used model-based battery capacity prognostic methods. However, it is hard to derive a generally accepted and precisely analytical model to track the degradation of battery capacity. Data-driven method employs machine learning algorithms to predict the battery's future capacity and deduce the battery RUL. The method avoids constructing mathematical or physical models in battery capacity prognostic works [20]–[28]. Artificial neural network is a classical data-driven method in the battery capacity prognostic works [20], [21], but the method has inadequacies of low computation speed, intangible network structure and inferior generalization performance. Support vector machine (SVM) [29] and relevance vector machine (RVM) [30] are commonly used data-driven methods in the recent battery capacity prognostic works [22]–[28]. Compared to the SVM, RVM is a Bayesian sparse kernel technique with usage of much fewer kernel function computations and higher performance.

The kernel is one of the essential ingredients of the machine learning algorithms such as SVM, RVM and minimax probability machine, etc. [22]–[35], where it is responsible for mapping the inner product of data points to a high-dimensional feature space. Meanwhile, the single kernel learning method is commonly used in the RVM prognostic works [25]–[28], [33]. However, the method may discard the useful information in training data and thus result in suboptimal generalization performance. For the purpose of generating better classification and prediction performance in the machine learning works, learning the optimal kernel has attracted the increasing attentions [36]–[42]. Multiple kernel learning method considers the kernel as a convex combination of basis kernel functions, and the method has shown better generalization performance than the single kernel learning method in the relevant works [39]–[42]. With respect to the multiple kernel learning method applied in the RVM, a multiple homogeneous kernel learning approach to generate the optimal kernel for the RVM is presented in work [42]. Nevertheless, the method doesn't consider the heterogeneous kernel functions which keep diversity of the kernel in the kernel learning procedure.

The measured battery capacity data are regularly polluted by miscellaneous noises for the reason of instrument errors, interference factors in the measurement, uncertain load and other unknown reasons in batteries. Capacity prognostic based on the data polluted by noises usually causes inaccurate results. Therefore, removing the noises from the measured capacity data and extracting the noise-free data are obviously important. Empirical mode decomposition (EMD) is a method for analyzing nonlinear and non-stationary signal by adaptively decomposing the signal into a series of zero-mean oscillation modes called intrinsic mode functions (IMFs) [43], and method's main advantage is that its basis functions are extracted from the signals in contrast to wavelet methods whose basis functions are fixed [26], [28], [43]–[45].

On the basis of the above discussions, a battery capacity prognostic approach is presented in the paper. An EMD denoising method is employed to process the measured data for the purpose of removing the noises and extracting the noise-free battery capacity data. A MKRVM is utilized to construct a prediction model based on the denoised capacity data. The proposed MKRVM's kernel is a weighted combination of basic heterogeneous kernel functions, and its sparse weights are generated by using the particle swarm optimization (PSO) algorithm. Battery capacity data measured in the data repository of NASA Ames Prognostics Center of Excellence [46] are used to demonstrate the effect of the denoising method, and verify the predict performance of the proposed MKRVM.

This paper is organized as the following: Section 2 illustrates the EMD denoising method to yield noise-free battery capacity data. Section 3 presents the MKRVM, and its sparse weights generation by using the PSO algorithm. Section 4 introduces the experiment data and procedure, and then gives the experiment results and discussions. Section 5 draws the conclusions eventually.

II. EMD DENOISING

The measured capacity data of batteries are regularly polluted by miscellaneous noises. Experiment with data polluted by noises cannot make accurate capacity prediction. Therefore, it is necessary to process the capacity data for the purpose of extracting the noise-free data. EMD denoising method is employed to address the concern in the work.

The gradual decreased battery capacity is a commonly used battery SOH indicator charactering its health degradation. Assuming that the measured capacity data $m(c)$ are composed of

$$m(c) = x(c) + \sigma(c) \tag{1}$$

where $x(c)$ are noise-free battery capacity data; σ (c) are noises; c denotes the cycle, which is a time index.

EMD denoising method includes two processes: signal $m(c)$ sifting and choosing relevant mode IMFs. $m(c)$ are decomposed into IMFs and a residual $r_N(c)$ by using the sifting process

$$m(c) = \sum_{i=1}^{N} \mathrm{IMF}_i(c) + r_N(c) \tag{2}$$

where N is the number of the IMFs.

For the most important structures of the signal $m(c)$ are comprised by relevant mode IMFs and residual. Correlation coefficient r_i of $m(c)$ and $\mathrm{IMF}_i(c)$ is employed to identify the relevant mode $\mathrm{IMF}_k(c)$ [45]

$$r_i = \sum_{i=1}^{N} m(c)\mathrm{IMF}_i(c) \Big/ \sqrt{\sum_{i=1}^{N} m^2(c) \sum_{i=1}^{N} \mathrm{IMF}_i^2(c)} \tag{3}$$

where $k > i$, and $k \in [1, N]$. r_k is firstly greater than a correlation coefficient threshold T. The relevant mode IMF(c)s and $r_N(c)$ constitute the noise-free battery capacity data $x(c)$

$$x(c) = \sum_{i=k}^{N} \mathrm{IMF}_i(c) + r_N(c) \tag{4}$$

III. MKRVM AND ITS SPARSE WEIGHTS GENERATION

Generally, the capacity sample data are divided into training data and testing data. The training data are used to construct a prediction model by prognostic approach, and the testing data are employed to evaluate the model's performance. Assume $\{x_i\}_{i=1}^N$ and $\{c_i\}_{i=1}^N$ are input time index data and capacity data of the training data, respectively, where N is the size of the training data. In the work, the MKRVM is employed to set up a model of the dependency of the $\{c_i\}_{i=1}^N$ on the $\{x_i\}_{i=1}^N$ to make accurate prediction of c_* for new input x_*.

A. MKRVM APPROACH

RVM is a Bayesian treatment and it is a special case of a sparse kernel model [30].

The target c_i is produced by

$$c_i = y(x_i; \mathbf{w}) + \varepsilon_i \tag{5}$$

where ε_i denote noises; $\mathbf{w} = (w_1, w_2, \ldots, w_N)^{\mathrm{T}}$ is a weight vector.

Assume c_i is independent, and the complete dataset's likelihood is

$$p(\mathbf{c}|\mathbf{w}, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\{-\frac{1}{2\sigma^2}||\mathbf{c} - \boldsymbol{\varphi}\mathbf{w}||^2\} \quad (6)$$

where $\mathbf{w} = (w_0, w_1, \ldots, w_N)^{\mathrm{T}}$; $\mathbf{c} = (c_1, c_2, \ldots, c_N)^{\mathrm{T}}$; $\boldsymbol{\varphi} = [\varphi\ (x_1), \varphi\ (x_2), \ldots, \varphi\ (x_N)]^{\mathrm{T}}$ with $\varphi\ (x_i) = [1, K(x_i, x_1), K(x_i, x_2), \ldots, K(x_i, x_N)]^{\mathrm{T}}$; $K(x_i, x)$ is kernel.

Maximum likelihood estimations of $\mathbf{w}$ and σ^2 in Eq. (6) usually bring about over-fitting. For the purpose of constraining the parameters, an explicit zero-mean Gaussian prior probability distribution is defined as

$$p(\mathbf{w}|\boldsymbol{\alpha}) = \prod_{i=0}^{N} N(w_i|0, \alpha_i^{-1}) \quad (7)$$

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

The posterior probability of the unbeknown parameters is defined as

$$p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2|\mathbf{c}) = \frac{p(\mathbf{c}|\mathbf{w}, \boldsymbol{\alpha}, \sigma^2)p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2)}{\int p(\mathbf{c}|\mathbf{w}, \boldsymbol{\alpha}, \sigma^2)p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2)d\mathbf{w}d\boldsymbol{\alpha}d\sigma^2} \quad (8)$$

However, $\int p(\mathbf{c}|\mathbf{w}, \boldsymbol{\alpha}, \sigma^2)p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2)d\mathbf{w}d\boldsymbol{\alpha}d\sigma^2$ is difficult to be calculated. Hence, $p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2|\mathbf{c})$ is instead as

$$p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2|\mathbf{c}) = p(\mathbf{w}|\boldsymbol{\alpha}, \sigma^2, \mathbf{c})p(\boldsymbol{\alpha}, \sigma^2|\mathbf{c}) \quad (9)$$

The weights' posterior distribution is generated as

$$\begin{aligned} &p(\mathbf{w}|\mathbf{c}, \boldsymbol{\alpha}, \sigma^2) \\ &= \frac{p(\mathbf{c}|\mathbf{w}, \sigma^2)p(\mathbf{w}|\boldsymbol{\alpha})}{p(\mathbf{c}|\boldsymbol{\alpha}, \sigma^2)} \\ &= (2\pi)^{-(N+1)/2}|\boldsymbol{\Sigma}|^{-1/2} \exp\{-\frac{1}{2}(\mathbf{w} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{w} - \boldsymbol{\mu})\} \end{aligned} \quad (10)$$

where $\boldsymbol{\mu} = \sigma^{-2}\boldsymbol{\Sigma}\varphi^{\mathrm{T}}\mathbf{c}$ with $\boldsymbol{\Sigma} = (\sigma^{-2}\boldsymbol{\varphi}^{\mathrm{T}}\boldsymbol{\varphi} + A)^{-1}$ and $A = diag(\alpha_0, \alpha_1, \ldots, \alpha_N)$.

Because of the unified hyperpriors, $p(\mathbf{c}|\boldsymbol{\alpha}, \sigma^2)$is defined as

$$\begin{aligned} p(\mathbf{c}|\boldsymbol{\alpha}, \sigma^2) &= \int p(\mathbf{c}|\mathbf{w}, \sigma^2)p(\mathbf{w}|\boldsymbol{\alpha})d\mathbf{w} \\ &= (2\pi)^{-N/2}|\sigma^2\mathbf{I} + \boldsymbol{\varphi}A^{-1}\boldsymbol{\varphi}^{\mathrm{T}}|^{-1/2} \\ &\quad \times \exp\{-\frac{1}{2}\mathbf{c}^{\mathrm{T}}(\sigma^2\mathbf{I} + \boldsymbol{\varphi}A^{-1}\boldsymbol{\varphi}^{\mathrm{T}})^{-1}\mathbf{c}\} \end{aligned} \quad (11)$$

The maximum a posterior (MP) estimation of the weights is obtained by the posterior mean, which relies on the value of $\boldsymbol{\alpha}$ and σ^2. The estimations of $\boldsymbol{\alpha}_{\mathrm{MP}}$ and σ^2_{MP} are generated by maximizing the marginal likelihood. The iterative formulas of $\boldsymbol{\alpha}_{\mathrm{MP}}$ and σ^2_{MP} are

$$\alpha_i^{new} = \frac{1 - \alpha_i N_{ii}}{\mu_i^2} \quad (12)$$

$$(\sigma^2)^{new} = \frac{||\mathbf{c} - \boldsymbol{\varphi}w||^2}{N - \boldsymbol{\Sigma}_i(1 - \alpha_i N_{ii})} \quad (13)$$

where N_{ii} is the posterior weight covariance's i-th diagonal element.

With respect to a new input x_*, its corresponding output c_*'s probability distribution is

$$p(c_*|\mathbf{c}, \boldsymbol{\alpha}_{\mathrm{MP}}, \sigma^2_{\mathrm{MP}}) = \int p(c_*|\mathbf{w}, \sigma^2_{\mathrm{MP}})p(\mathbf{w}|\mathbf{c}, \boldsymbol{\alpha}_{\mathrm{MP}}, \sigma^2_{\mathrm{MP}})d\mathbf{w} \quad (14)$$

The distribution is a Gaussian form for both integrated terms are Gaussian

$$p(c_*|\mathbf{c}, \boldsymbol{\alpha}_{\mathrm{MP}}, \sigma^2_{\mathrm{MP}}) = N(c_*|y_*, \sigma_*^2) \quad (15)$$

where $y_* = \mu^{\mathrm{T}}\varphi\ (x_*)$ and $\sigma_*^2 = \sigma^2_{\mathrm{MP}} + \varphi\ (x_*)^{\mathrm{T}}\boldsymbol{\Sigma}\varphi(x_*)$.

$K(x_i, x)$ in the $\varphi\ (x_i)$ is applied to map the inner product of data points to a high-dimensional feature space, and the kernel is significantly important to the prediction performance of the RVM. Gaussian kernel function is a commonly used kernel function in the RVM for its powerful nonlinear processing capability, and the function is defined as

$$K_1(x_i, x) = \exp(-\frac{\|x_i - x\|^2}{2\gamma^2}) \quad (16)$$

where γ is the width factor.

Polynomial kernel function is a global kernel function, and it is a beneficial complement to the Gaussian kernel function in many machine learning works [39]–[41]. The function is defined as

$$K_2(x_i, x) = [(x_i^{\mathrm{T}} \cdot x) + 1]^d \quad (17)$$

where d is the degree.

Proper kernel is needed in the RVM for the purpose of making accurate prediction. Single kernel learning method has been used to generate kernel in the RVM prognostic work [25]–[28], [33]. However, the method may discard useful information in the training data and thus result in suboptimal generalization performance. Compared to the single kernel learning method, multiple kernel learning is a novel and effective method to consist the optimal kernel for the reason of a convex combination of finitely many basic kernel functions can always generate the optimal kernel. The MKRVM's kernel is produced by using multiple heterogeneous kernel learning method, and the MKRVM is used to predict the battery's future capacity in the work. The multiple heterogeneous kernel learning method is defined as

$$K(x_i, x) = \sum_{m=1}^{M} d_m K_m(x_i, x) \quad (18)$$

where d_m is the weight of the m-th kernel function with $d_m \geq 0$ and $\sum_{m=1}^{M} d_m = 1$. $K(x_i, x)$ in the work is composed of the Polynomial kernel function of degree 1 to 3 and Gaussian kernel function with 10 different width factors.

Since a sparse weights solution is helpful in precision and fitting quality [39]–[42], the weight is cleared when the corresponding basic kernel function contributes less to the prediction performance. PSO algorithm has shown superior

efficiency in the parameter estimation [47]–[49]. This algorithm can seek the kernel function's weights by using a fitness function and easily generate an effectively sparse weights solution. Therefore, the sparse weights of the MKRVM are produced by using the PSO algorithm in the work.

B. PSO ALGORITHM

PSO algorithm is an effective evolutionary method presented in 1995 [50]. Referencing birds' flocking, particles generated in the algorithm seek the optimal solution for the target.

The particle k's optimal location can be selected as

$$P_k = \begin{cases} P_k, & f(P_k) < f(X_k(t+1)) \\ X_k(t+1), & f(P_k) \geq f(X_k(t+1)) \end{cases} \tag{19}$$

where $k = 1, 2, \cdots, s$, and s is the size of swarm; $X_k \in [X_{min}, X_{max}]$, and it is the particle k's location; t is the current iteration number; $f(x)$ is a minimized objective fitness function.

The global optimal location of the swarm is elected as

$$P_g(t) = \arg\min \{f(P_k)\} \tag{20}$$

The location and speed evolution equations are

$$V_k(t+1) = wV_k(t) + c_1 r_1(P_k - X_k(t)) + c_2 r_2(P_g - X_k(t)) \tag{21}$$

$$X_k(t+1) = X_k(t) + V_k(t+1) \tag{22}$$

where V_k is the particle k's speed; c_1 and c_2 are accelerating factors; w is inertia weight; r_1 and r_2 are random numbers between 0 and 1.

C. PROCEDURE OF WEIGHTS SEEKING

The weights are mapped to the multi-dimensional locations of every particle and the optimization target is to seek a best particle with the optimal location.

The seeking flowchart is shown in Fig. 1 and the seeking steps are as follows

1) Initialize the PSO algorithm's particles.
2) Calculate each particle's fitness value. Mean square error (MSE) is adopted as the fitness function, and it reflects the difference between the predicted data and initial data

$$\text{MSE} = \frac{\sum_{r=1}^{R} [z^*(r) - z(r)]^2}{R} \tag{23}$$

where $r = 1, 2, \ldots R$, and R is the size of data; $z^*(r)$ are the predicted data and $z(r)$ are the initial data.
3) Produce each particle's optimal location and the global optimal location of the swarm in accordance with Eq. (19) and Eq. (20).
4) Update the speed and location of each particle according to Eq. (21) and Eq. (22).
5) Repeat the step 2) to 4) until the maximum iteration.

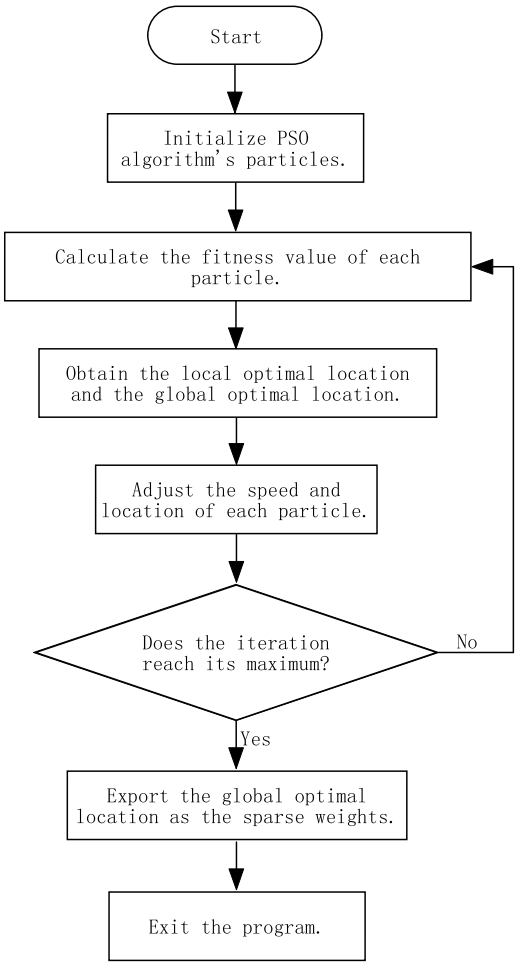

FIGURE 1. Flowchart of sparse weights generation by using PSO algorithm.

6) Acquire the optimal multi-dimensional locations and considered them as the sparse weights of the kernel functions.
7) Exit the program.

When a basic kernel function contributes less to the prediction performance in the weights seeking procedure, its weight would be cleared by PSO algorithm. The remained weights are sparse weights generated by PSO for the MKRVM.

IV. PROGNOSTICS EXPERIMENT

A. EXPERIMENT DATA AND PROCEDURE

The lithium-ion battery capacity data are measured in the data repository of NASA Ames Prognostics Center of Excellence [46], and they are used to demonstrate the proposed capacity prognostic approach in the work. In the NASA data collected procedure, lithium-ion batteries are testing at 25 °C under three different operational profiles: charge, discharge and impedance. Charging is using a 1.5 A constant current until the battery's voltage reaches 4.2 V, and then the voltage is maintained until the current drops to 20mA. Discharging is performing at a 2 A constant current until the voltage drops to 2.7 V, 2.2 V and 2.5 V, which are corresponding to

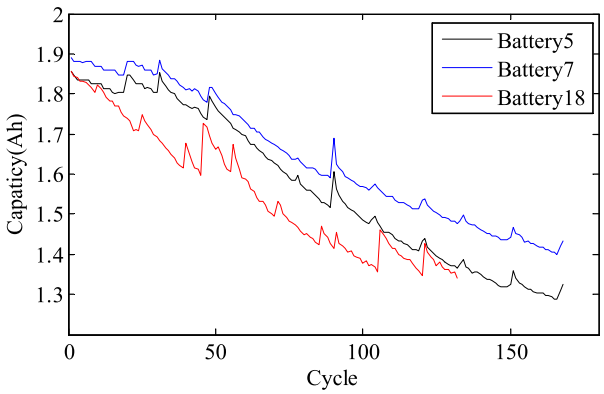

FIGURE 2. Measured capacity data of batteries.

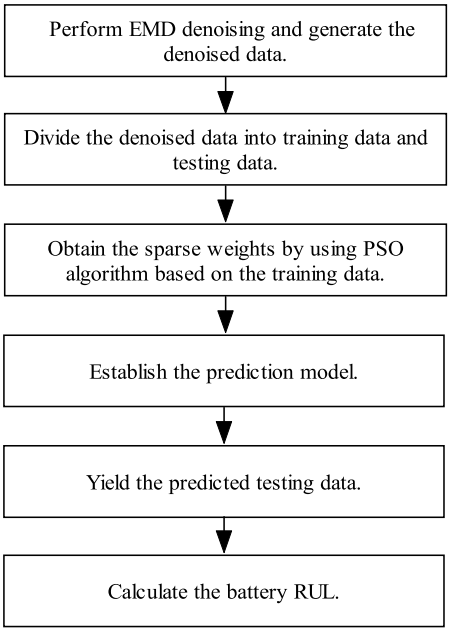

FIGURE 3. Flowchart of experiment steps.

battery 5, battery 7 and battery 18, respectively. Impedance measurement refers to an electrochemical impedance spectroscopy frequency sweep ranging from 0.1 Hz to 5 kHz. The measured battery 5, battery 7 and battery 18 capacity data are shown in Fig. 2. A cycle refers to once charging and discharging operation. It can be observed that every capacity trajectory descends with charging and discharging, and at some cycle ascends rapidly, shortly and irregularly due to instrument errors, interference factors in the measurement, uncertain load and other unknown reasons in batteries. The lengths of capacity data of battery 5, battery 7 and battery 18 are 168, 168 and 132 cycles, respectively.

In the work, RUL refers to the remaining useful cycle that the battery performance guarantees, and it is the end of life (EOL) cycle minus the current cycle. The detailed experiment steps of battery capacity prognostic are shown in Fig. 3 and described as follows

1) Perform EMD denoising method on basis of the measured data, and the denoised data are generated.

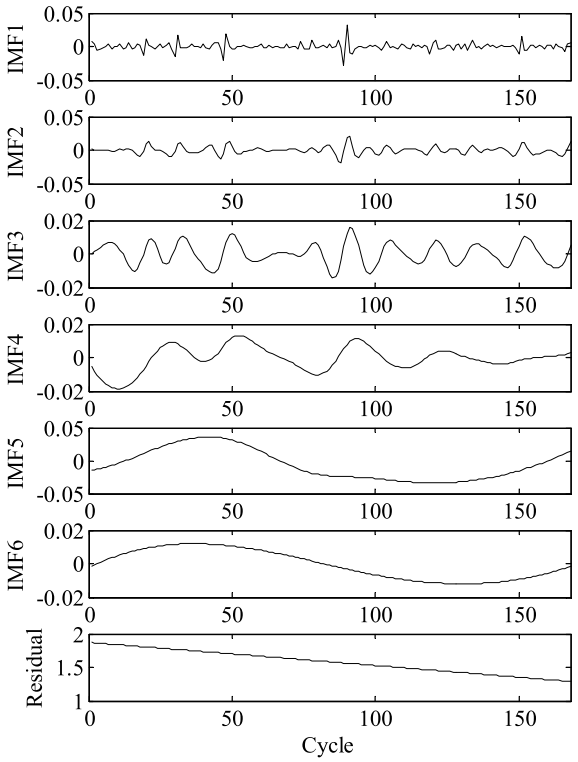

FIGURE 4. EMD results of the measured battery 5's capacity.

TABLE 1. Correlation coefficients between decompose results and the measured capacity data of the batteries

	Battery 5	Battery 7	Battery 18
IMF1	0.0404	0.0289	0.0618
IMF2	0.0819	0.0581	0.1359
IMF3	0.1112	0.1369	0.0778
IMF4	-0.0795	-0.0512	0.1218
IMF5	0.6637	0.2329	0.2392
IMF6	0.8835	0.8155	0.2892
Residual	0.9877	0.9893	0.9745

2) Divide the denoised data into training data and testing data.
3) By using PSO algorithm, sparse weights of MKRVM are yielded on basis of the training data.
4) Construct a capacity prediction model by using MKRVM which employs the generated sparse weights.
5) Yield the predicted testing data by the model.
6) Calculate the estimated RUL of the battery.

B. EXPERIMENT RESULTS AND ANALYSIS

EMD denosing is employed to denoise the measured capacity data. The measured capacity data of battery 5, battery 7 and battery 18 are decomposed into 6 IMFs and a residual. Fig. 4, Fig. 5 and Fig. 6 demonstrate the EMD results of battery 5, battery 7 and battery 18, respectively. The correlation coefficients between decompose results and the measured capacity data of the batteries are calculated according to Eq. (3), and shown in Table 1. The correlation coefficient threshold is

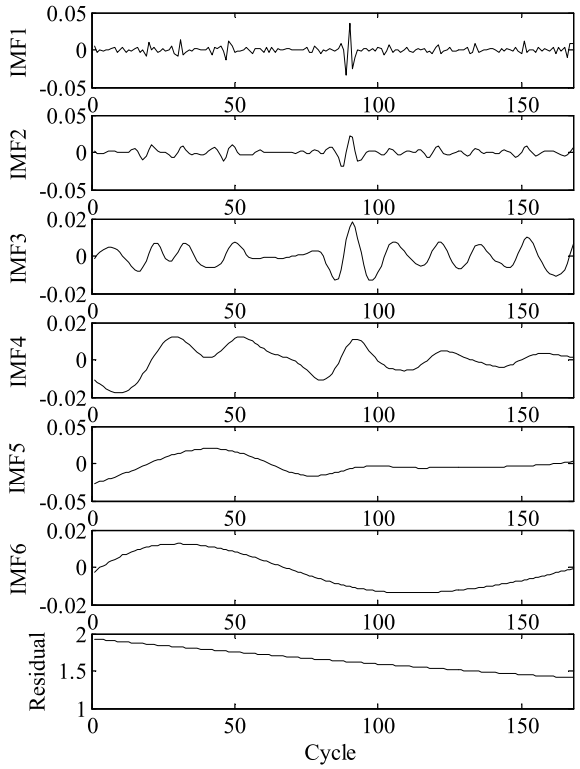

FIGURE 5. EMD results of the measured battery 7's capacity.

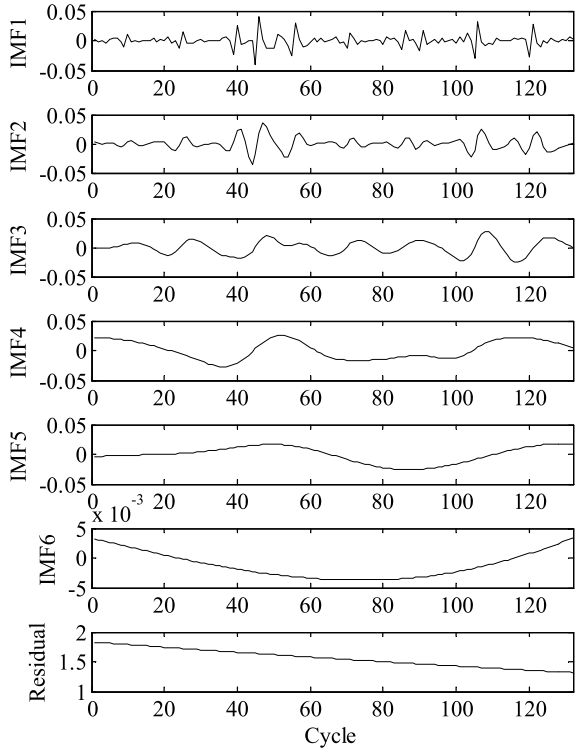

FIGURE 6. EMD results of the measured battery 18's capacity.

set to 0.2. Then, the denoised capacity data are generated in according with Eq. (4) based on the decompose results whose correlation coefficients are higher than 0.2.

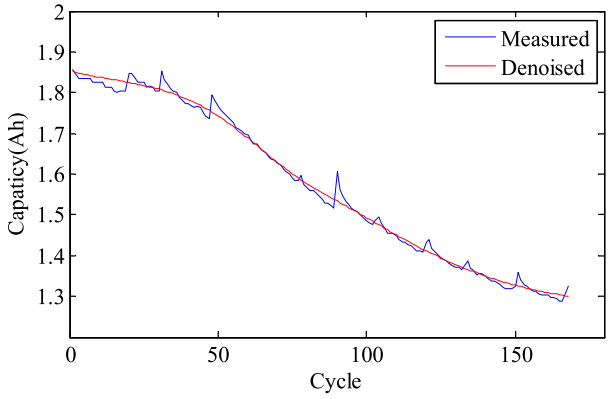

FIGURE 7. The denoised capacity data of battery 5.

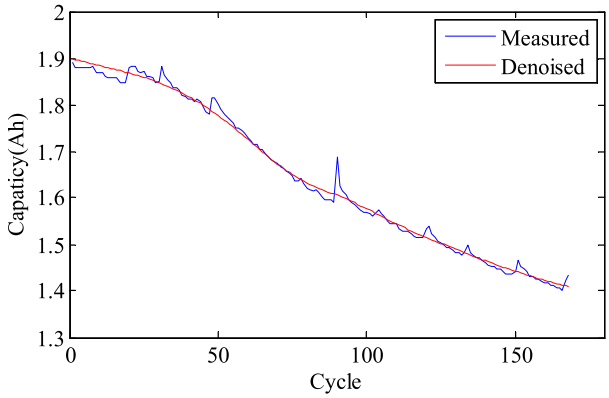

FIGURE 8. The denoised capacity data of battery 7.

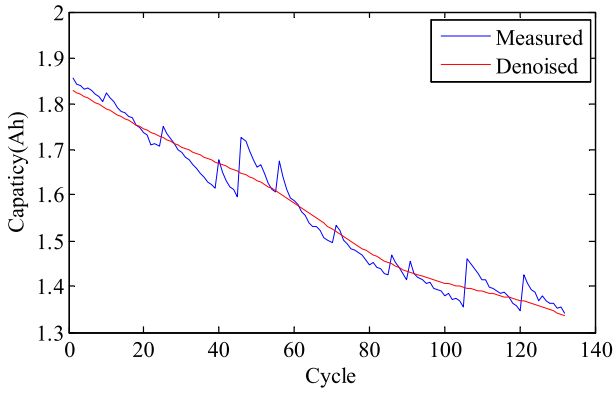

FIGURE 9. The denoised capacity data of battery 18.

The generated denoised capacity data of battery 5, battery 7 and battery 18 are shown in Fig. 7, Fig. 8 and Fig. 9. In the figures, the trajectories of the denoised data descend by degrees when the batteries are charging and discharging in the experiments. The denoised capacity data are all close to the measured capacity data obviously, which reflects that denoised capacity data can character the downward trend of the denoised capacity data in the testing.

In the experiments, each nominal capacity of lithium-ion battery is 2 Ah and the EOL threshold is 1.38 Ah. The lowest capacity of battery 7 is 1.4005 Ah, which is higher than

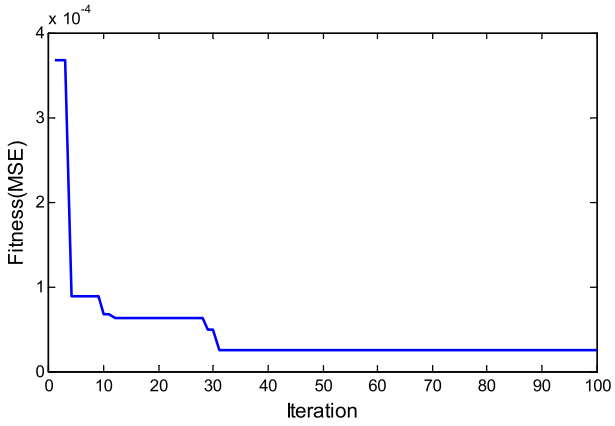

FIGURE 10. The sparse weights generate procedure of battery 5 case.

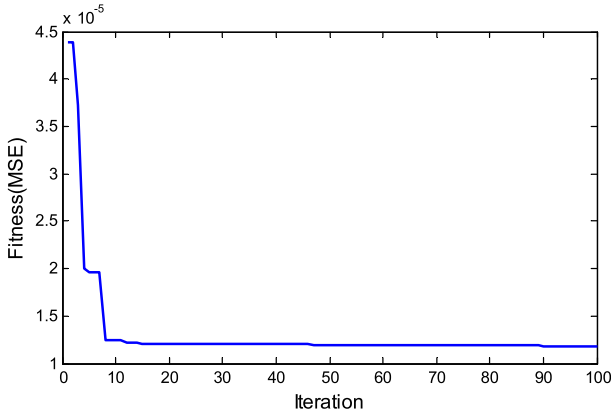

FIGURE 11. The sparse weights generate procedure of battery 18 case.

the EOL threshold. Therefore, the measured capacity data of battery 7 are only used to show the effect of the EMD denoising method, and not applied to predict the battery's future capacity. The battery capacity prediction experiments include a battery 5 capacity prediction case and a battery 18 capacity prediction case. The lengths of the training data in the two battery cases are set to 80 and 70, respectively.

The PSO algorithm's swarm size is set to 10; the maximum iteration is set to 100; the locations of particles are limited between 0 and 1; c_1 and c_2 are both set to 2; w is linearly decreased from 0.95 to 0.4 with the evolution. On basis of the experiences of former works [25]–[28], [33], [42], the degrees of the polynomial kernel function and the width factors of the Gaussian kernel function are set to 1, 2, 3, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, which are corresponding to weight 1, weight 2, weight 3, weight 4, weight 5, weight 6, weight 7, weight 8, weight 9, weight 10, weight 11, weight 12, weight 13, respectively.

Fig. 10 and Fig. 11 show the generate procedures of sparse weights by using the PSO algorithm in the battery 5 and battery 18 capacity prediction cases. As displayed in the figures, the PSO algorithm can seek the sparse weights for the capacity prediction cases. The generated sparse weights of two cases are shown in Table 2.

TABLE 2. Sparse weights generated of batteries.

Weight	Battery 5	Battery 18
Weight 1	0.0176	0
Weight 2	0	0.1029
Weight 3	0.1093	0
Weight 4	0	0.0087
Weight 5	0.0127	0
Weight 6	0.0277	0
Weight 7	0	0.1492
Weight 8	0	0.1791
Weight 9	0.1909	0
Weight 10	0.1155	0.2580
Weight 11	0.2694	0
Weight 12	0.0103	0.1372
Weight 13	0.2466	0.1649

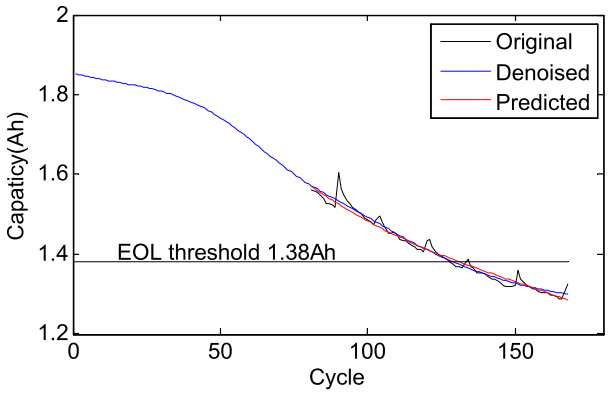

FIGURE 12. Prediction results of battery 5 case.

Adopting the generated sparse weights, the MKRVM is used to predict the battery's future capacity. The estimated RUL, absolute error (AE), relative accuracy (RA), $\alpha - \lambda$ *accuracy* [51] and MSE are employed as measure metrics. The metrics are defined as

$$\mathrm{AE} = \mathrm{RUL_{es}} - \mathrm{RUL} \tag{24}$$

$$\mathrm{RA} = 1 - \frac{|\mathrm{RUL_{es}} - \mathrm{RUL}|}{\mathrm{RUL}} \tag{25}$$

$$\alpha - \lambda\ accuracy = \begin{cases} Yes & if\ \mathrm{RUL_{es}} \in [\mathrm{B_1}, \mathrm{B_2}] \\ No & if\ others \end{cases} \tag{26}$$

where $\mathrm{RUL_{es}}$ refers to the estimated RUL; $\mathrm{B_1}$ and $\mathrm{B_2}$ are confidence intervals which are equal to $\mathrm{RUL} * (1 - \alpha)$ and $\mathrm{RUL} * (1+\alpha)$, respectively, and α is a bound set to 0.1.

The RUL of the denoised data of battery 5 is 49, and the RUL of the denoised data of battery 18 is 45. The predictions of the two cases are displayed in Fig. 12 and Fig. 13. Estimated RULs, AEs, RAs, MSE1s, MSE2s of two cases are shown in Table 3, where MSE1 refers to the MSE of the predicted testing data and the denoised testing data, and MSE2 denotes the MSE of the predicted testing data and the measured testing data. The figures obviously show that the MKRVM can predict the comparatively accurate future capacity of two batteries. Meanwhile this can also be proved by the MSE1s and MSE2s in Table 3, which are very low

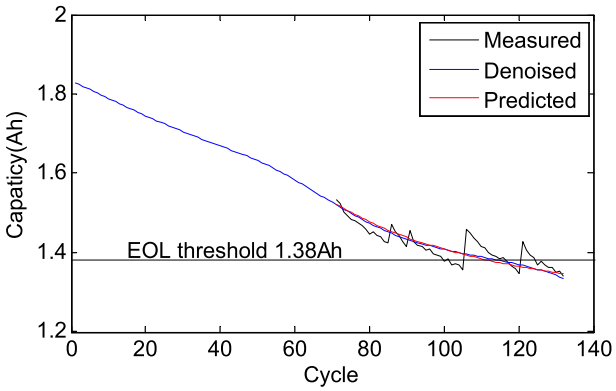

FIGURE 13. Prediction results of battery 18 case.

TABLE 3. Estimated RULs, AEs, RAs and MSEs of two cases.

Measure metric	Battery 5	Battery 18
Estimated RUL	51	42
AE	2	-3
RA	95.9%	93.3%
MSE1	4.4972e-05	1.6437e-05
MSE2	2.4354e-04	6.6203e-04

TABLE 4. Compare results of two methods.

Method	Measure metric	Battery 5	Battery 18
RVM	Estimated RUL	52	41
	AE	3	-4
	MSE1	8.1706e-05	4.1624e-05
	MSE2	2.6501e-04	7.2523e-04
SVM	Estimated RUL	53	51
	AE	4	6
	MSE1	1.4115e-04	2.8595e-04
	MSE2	3.7265e-04	8.6736e-04

values reflecting that the predicted testing data are close to the denoised testing data and measured testing data in the two cases. RAs are both beyond 90% in the two cases indicating high prediction precisions produced by the MKRVM. Meanwhile, two $\alpha - \lambda$ *accuracy*s are both *Yes* in the two cases for the estimated RULs are both within the confidence intervals.

In order to prove the predict performance of the proposed MKRVM approach, the proposed MKRVM approach is compared with RVM optimized by PSO algorithm approach [33] and SVM improved by PSO algorithm approach [52]. The measured data and denoised data of the battery 5 and 18 are used as experiment data. The measure metrics include estimated RULs, AE, MSE1 and MSE2. The results are recorded in Table 4. From the table it can be seen that the RVM and SVM are also effective in battery capacity prediction. Meanwhile, MKRVM provides smaller MSE1 and MSE2 than the RVM and SVM which reflects that the data predicted by the MKRVM are more closely to the measured data and denoised data. Meanwhile, the MKRVM produces lower AE than the RVM and SVM which indicates that the MKRVM can yield more precise prediction than the RVM and SVM. Therefore, the proposed MKRVM approach observably precedes the RVM and SVM approaches in the battery capacity prediction.

V. CONCLUSION

An EMD denoising method has been presented to process the measured battery capacity data for the purpose of generating the noise-free capacity data. Meanwhile, the method has been demonstrated by the measured battery 5, battery 7 and battery 18 capacity data.

A MKRVM approach has been proposed to predict the battery's future capacity in the work. Meanwhile, its sparse weights are produced by using PSO algorithm. Battery 5 and battery 18 capacity prediction cases' results have validated that the proposed MKRVM approach can predict the battery's future capacity precisely. Furthermore, a compare experiment has demonstrated that the proposed MKRVM approach can yield higher prediction precision than RVM and SVM.

There are two future works associated with our work. First, it would be interesting to apply the EMD denosing and MKRVM methods to other prognostic applications. Second, it is promising to develop another advanced kernel function as basis kernel function, and apply it in our MKRVM method.

REFERENCES

[1] X. Hu, C. M. Martinez, and Y. Yang, "Charging, power management, and battery degradation mitigation in plug-in hybrid electric vehicles: A unified cost-optimal approach," *Mech. Syst. Signal Process.*, vol. 87, pp. 4–16, Mar. 2017.

[2] C. Zou, C. Manzie, and D. Nešić, "A framework for simplification of PDE-based lithium-ion battery models," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 5, pp. 1594–1609, Sep. 2016.

[3] C. Zou, C. Manzie, D. Nesic, and A. G. Kallapur, "Multi-time-scale observer design for state-of-charge and state-of-health of a lithium-ion battery," *J. Power Sour.*, vol. 335, pp. 121–130, Dec. 2016.

[4] C. Zou, A. G. Kallapur, C. Manzie, and D. Nešić, "PDE battery model simplification for SOC and SOH estimator design," in *Proc. 54th IEEE Conf. Decision Control (CDC)*, Osaka, Japan, Sep. 2015, pp. 1328–1333.

[5] Z. Wang, J. Hong, P. Liu, and L. Zhang, "Voltage fault diagnosis and prognosis of battery systems based on entropy and Z-score for electric vehicles," *Appl. Energy*, vol. 196, pp. 289–302, Jun. 2017.

[6] L. Zheng, L. Zhang, J. Zhu, G. Wang, and J. Jiang, "Co-estimation of state-of-charge, capacity and resistance for lithium-ion batteries based on a high-fidelity electrochemical model," *Appl. Energy*, vol. 180, pp. 424–434, Oct. 2016.

[7] S. Yuan, H. Wu, and C. Yin, "State of charge estimation using the extended Kalman filter for battery management systems based on the ARX battery model," *Energies*, vol. 6, no. 1, pp. 444–470, Jan. 2013.

[8] Z. Chen, Y. Fu, and C. C. Mi, "State of charge estimation of lithium-ion batteries in electric drive vehicles using extended Kalman filtering," *IEEE Trans. Veh. Technol.*, vol. 62, no. 3, pp. 1020–1030, Mar. 2013.

[9] X. Zheng and H. Fang, "An integrated unscented Kalman filter and relevance vector regression approach for lithium-ion battery remaining useful life and short-term capacity prediction," *Reliab. Eng. Syst. Safety*, vol. 144, pp. 74–82, Dec. 2015.

[10] M. Safari, M. Morcrette, A. Teyssot, and C. Delacourt, "Multimodal physics-based aging model for life prediction of Li-Ion batteries," *J. Electrochem. Soc.*, vol. 156, no. 3, pp. A145–A153, Jan. 2009.

[11] S. Tang, C. Yu, X. Wang, X. Guo, and X. Si, "Remaining useful life prediction of lithium-ion batteries based on the Wiener process with measurement error," *Energies*, vol. 7, no. 2, pp. 520–547, Feb. 2014.

[12] W. Xian, B. Long, M. Li, and H. Wang, "Prognostics of lithium-ion batteries based on the Verhulst model, particle swarm optimization and particle filter," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 1, pp. 2–17, Jan. 2014.

[13] X. Hu, J. Jiang, D. Cao, and B. Egardt, "Battery health prognosis for electric vehicles using sample entropy and sparse Bayesian predictive modeling," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2645–2656, Apr. 2016.

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

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

[16] L. Chao, Q. Lai, T. Ge, H. Yu, L. Wang, and N. Ma, "A lead-acid battery's remaining useful life prediction by using electrochemical model in the particle filtering framework," *Energy*, vol. 120, no. 1, pp. 975–984, Feb. 2017.

[17] Z. Liu, G. Sun, S. Bu, J. Han, X. Tang, and M. Pecht, "Particle learning framework for estimating the remaining useful life of lithium-ion batteries," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 2, pp. 280–293, Feb. 2017.

[18] W. Yan, B. Zhang, W. Dou, D. Liu, and Y. Peng, "Low-cost adaptive Lebesgue sampling particle filtering approach for real-time Li-Ion battery diagnosis and prognosis," *IEEE Trans. Autom. Sci. Eng.*, to be published.

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

[20] W. X. Shen, C. C. Chan, E. W. C. Lo, and K. T. Chau, "A new battery available capacity indicator for electric vehicles using neural network," *Energy Convers. Manage.*, vol. 43, no. 6, pp. 817–826, Apr. 2002.

[21] J. Wu, C. Zhang, and Z. Chen, "An online method for lithium-ion battery remaining useful life estimation using importance sampling and neural networks," *Appl. Energy*, vol. 173, no. 1, pp. 134–140, Jul. 2016.

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

[23] M. A. Patil *et al.*, "A novel multistage support vector machine based approach for Li ion battery remaining useful life estimation," *Appl. Energy*, vol. 159, no. 1, pp. 285–297, Dec. 2015.

[24] H. Dong, X. Jin, Y. Lou, and C. Wang, "Lithium-ion battery state of health monitoring and remaining useful life prediction based on support vector regression-particle filter," *J. Power Sour.*, vol. 271, no. 20, pp. 114–123, Dec. 2014.

[25] H. Li, D. Pan, and C. L. P. Chen, "Intelligent prognostics for battery health monitoring using the mean entropy and relevance vector machine," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 44, no. 7, pp. 851–862, Jul. 2014.

[26] C. Zhang, Y. He, L. Yuan, S. Xiang, and J. Wang, "Prognostics of lithium-ion batteries based on wavelet denoising and DE-RVM," *Comput. Intell. Neurosci.*, vol. 2015, no. 1, Aug. 2015, Art. no. 14. [Online]. Available: https://www.hindawi.com/journals/cin/2015/918305/cta/

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

[28] D. Liu, J. Zhou, D. Pan, Y. Peng, and X. Peng, "Lithium-ion battery remaining useful life estimation with an optimized relevance vector machine algorithm with incremental learning," *Measurement*, vol. 63, pp. 143–151, Mar. 2015.

[29] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[30] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, no. 3, pp. 211–244, Jun. 2001.

[31] B. Gu and V. S. Sheng, "A robust regularization path algorithm for ν-support vector classification," *IEEE Trans. Neural. Netw. Learn. Syst.*, vol. 28, no. 5, pp. 1241–1248, Feb. 2016.

[32] B. Gu, V. S. Sheng, Z. Wang, D. Ho, S. Osman, and S. Li, "Incremental learning for ν-support vector regression," *Neural Netw.*, vol. 67, pp. 140–150, Jul. 2015.

[33] C. Zhang, Y. He, L. Yuan, and F. Deng, "A novel approach for analog circuit fault prognostics based on improved RVM," *J. Electron. Test.*, vol. 30, no. 3, pp. 343–356, Jun. 2014.

[34] B. Gu, V. S. Sheng, K. Y. Tay, W. Romano, and S. Li, "Incremental support vector learning for ordinal regression," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 7, pp. 1403–1416, Jul. 2015.

[35] B. Gu, X. Sun, and V. S. Sheng, "Structural minimax probability machine," *IEEE Trans. Neural. Netw. Learn. Syst.*, vol. 28, no. 7, pp. 1646–1656, Jul. 2017.

[36] J. Liu, J. Chen, S. Chen, and J. Ye, "Learning the optimal neighborhood kernel for classification," in *Proc. Int. Joint Conf. Artif. Intell.*, 2009, pp. 1144–1149.

[37] I. W. Tsang and J. T.-Y. Kwok, "Efficient hyperkernel learning using second-order cone programming," *IEEE Trans. Neural Netw.*, vol. 17, no. 1, pp. 48–58, Jan. 2006.

[38] Y. Chen, M. R. Gupta, and B. Recht, "Learning kernels from indefinite similarities," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, New York, NY, USA, 2009, pp. 145–152.

[39] M. Hu, Y. Chen, and J. T. Y. Kwok, "Building sparse multiple-kernel SVM classifiers," *IEEE Trans. Neural Netw.*, vol. 20, no. 5, pp. 827–839, May 2009.

[40] H. Yang, Z. Xu, J. Ye, I. King, and M. R. Lyu, "Efficient sparse generalized multiple kernel learning," *IEEE Trans. Neural Netw.*, vol. 22, no. 3, pp. 433–446, Mar. 2011.

[41] C. Cortes, M. Mohri, and A. Rostamizadeh, "L2 regularization for learning kernels," in *Proc. 25th Conf. Uncertain. Artif. Intell.*, Montreal, QC, Canada, 2009, pp. 109–116.

[42] K. Blekas and A. Likas, "Sparse regression mixture modeling with the multi-kernel relevance vector machine," *Knowl. Inf. Syst.*, vol. 39, no. 2, pp. 241–264, May 2014.

[43] P. Flandrin, G. Rilling, and P. Goncalves, "Empirical mode decomposition as a filter bank," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 112–114, Feb. 2004.

[44] M. Hassan, S. Boudaoud, J. Terrien, B. Karlsson, and C. Marque, "Combination of canonical correlation analysis and empirical mode decomposition applied to denoising the labor electrohysterogram," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 9, pp. 2441–2447, Sep. 2011.

[45] G. Yang, Y. Liu, Y. Wang, and Z. Zhu, "EMD interval thresholding denoising based on similarity measure to select relevant modes," *Signal Process.*, vol. 109, pp. 95–109, Apr. 2015.

[46] B. Saha and K. Goebel. (2007). *Battery Data Set. NASA Ames Prognostics Data Repository*. [Online]. Available: http://ti.arc.nasa.gov/project/prognostic-data-repository

[47] S. Lin, K. Ying, S. Chen, and Z. Lee, "Particle swarm optimization for parameter determination and feature selection of support vector machines," *Expert Syst. Appl.*, vol. 35, no. 4, pp. 1817–1824, Nov. 2008.

[48] X. Guo, J. Yang, C. Wu, and Y. Liang, "A novel LS-SVMs hyper-parameter selection based on particle swarm optimization," *Neurocomputing*, vol. 71, no. 16, pp. 3211–3215, Oct. 2008.

[49] C. Gao, J. Huang, Y. Sun, and S. L. Diao, "Particle swarm optimization based RVM classifier for non-linear circuit fault diagnosis," *J. Central South Univ.*, vol. 19, no. 2, pp. 459–464, Jan. 2012.

[50] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proc. 6th Int. Symp. Micro Mach. Human Sci.*, Nagoya, Japan, 1995, pp. 39–43.

[51] A. Saxena, J. Celaya, B. Saha, S. Saha, and K. Goebel, "Metrics for offline evaluation of prognostic performance," *Int. J. Prognostics Health Manage.*, vol. 1, no. 1, pp. 1–20, Jan. 2010.

[52] C. Sudheer, R. Maheswaran, B. K. Panigrahi, and S. Mathur, "A hybrid SVM-PSO model for forecasting monthly streamflow," *Neural Comput. Appl.*, vol. 24, no. 6, pp. 1381–1389, May 2014.

CHAOLONG ZHANG received the M.Sc. degree from the Nanjing University of Information Science and Technology in 2008. He is currently pursuing the Ph.D. degree with the School of Electrical Engineering and Automation, Hefei University of Technology. He is also an Associate Professor with the School of Physics and Electronic Engineering, Anqing Normal University. His current research interests include fault diagnostics and prognostics of analog and mixed-signal circuits, and battery capacity prognostic.

YIGANG HE received the Ph.D. degree from Xi'an Jiaotong University in 1996. He is currently a Professor, a Doctoral Supervisor, and the Head of the School of Electrical Engineering and Automation, Hefei University of Technology. He is one of the winners of the National Distinguished Young Scientists Foundation. His research interests are in the areas of circuit theory and its applications, testing and fault diagnosis of analog and mixed-signal circuits, smart grid, radio frequency identification technology, and intelligent signal processing.

LIFENG YUAN received the Ph.D. degree in electrical engineering from Hunan University, China, in 2003 and 2011, respectively. She is currently a Professor with the School of Electrical Engineering and Automation, Hefei University of Technology, China. Her research interests are in the area of test theory for circuit and system.

SHENG XIANG received the M.Sc. degree from HeFei University of Technology in 2014. He is currently pursuing the Ph.D. degree with the HeFei University of Technology. His research interests are applications, testing, and fault diagnosis of power electronic system.