IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 20, 2020, accepted November 8, 2020, date of publication November 16, 2020,
date of current version November 27, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3038182

State of Health Diagnosis and Remaining Useful Life Prediction for Lithium-ion Battery Based on Data Model Fusion Method

XIANGBO CUI, (Graduate Student Member, IEEE), AND TETE HU
State Key Laboratory of High Performance Complex Manufacturing, School of Mechanical and Electrical Engineering, Central South University, Hunan 410083, China

Corresponding author: Xiangbo Cui (xiangb.c@csu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51675539.

ABSTRACT Accurate state-of-health (SOH) diagnosis and remaining useful life (RUL) prediction of lithium-ion batteries (LIBs) play an extremely important role in ensuring safe and reliable operation of electric and hybrid vehicles. However, due to the complex electrochemical properties, it is difficult to achieve the goal of accurate diagnosis and prediction. Here, we propose a novel data-model fusion method to perform accurate SOH estimation and RUL prediction for LIBs, which considers nonlinear dynamics of not only discharging process but also charging process. A long short-term memory (LSTM) network is first employed to model battery SOH dynamics. A neural network (NN) model is then developed to describe battery capacity degradation mechanism according to the prior knowledge extracted from the charging process. Finally, an unscented Kalman filter (UKF) algorithm is incorporated with the LSTM network and NN model to filter out the noises and further reduce the estimation errors. Different from the traditional model fusion approaches, this proposed method uses full information from all sensors, and with no need for any physical model. Experiments and verification demonstrate both the effectiveness of this proposed method and its superior modeling performance as compared with several commonly used methods.

INDEX TERMS State-of-health diagnosis, remaining useful life prediction, long short-term memory, model fusion, unscented Kalman filter.

I. INTRODUCTION

In recent years, with the rapid development of energy storage technology, LIBs have been widely used in the field of engineering applications, e.g., electric vehicles (EVs), hybrid EVs and so forth, due to the advantages of high energy density, high power density, and long lifetime [1], [2]. One of the most important issues in the application of LIBs is to meet the safety-critical and energy-efficient requirements, in which effective SOH diagnosis and RUL prediction are considered as a key enabler, because the electrical properties, stability and safety alterations often change with battery SOH and RUL [3], [4]. SOH is defined as a percentage of internal resistance or capacity, which is utilized to describe the aging level of battery in each charge-discharge cycle. And RUL is defined as the number of remaining useful charge-discharge cycles at a specific cycle, which is calculated by the k-step-ahead projection of SOH [5], [6]. Therefore, the estimation accuracy of SOH directly affects the results of RUL prediction.

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

Accurately estimating SOH and predicting RUL are of great significance to provide the performance variance and ensure the reliable operation of battery during its whole storage life time. However, the complex internal electrochemical properties and unpredictability of physical measurements lead to the considerably complicated degradation behaviors of LIBs. Thus, it is considered as a great challenge to perform an accurate SOH estimation and RUL prediction.

Several lines of past work have contributed to estimate SOH and predict RUL of a battery. Some intelligent algorithms are applied to estimate SOH by using process parameters such as voltage, temperature and current measurements. Generally, these processes and algorithms can be classified into the following two categories: 1) Mode-based method; 2) Data-driven method.

A. MODE-BASED METHOD

Mode-based methods are conducted to capture the degradation dynamics based on the models that are used to describe the dynamic properties of battery, such as electrochemical models [7], equivalent circuit models [8], or empirical models

[9]. These estimation methods are often implemented by using the advanced filtering techniques, such as particle filter (PF) [10], extended Kalman filtering (EKF) [11], [12], slide mode observer [13], and Lyapunov-based adaptive observer [14], [15]. These methods could obtain good robustness and are also easy to conduct for online estimation, but they all require precise mathematical models. This is very difficult to satisfy in actual industrial processes due to lack of sufficient physical insights of aging dynamics.

B. DATA-DRIVEN METHOD

Benefitting from the massive historical data and no request of an explicit physical model, the data-driven methods have been widely used in the field of battery SOH and RUL dynamics in recent years. These modeling methods are usually conducted by using a large number of offline data to train and establish the nonlinear approximate models between the input and the output features. The commonly used methods include EKF method [16], Wiener model [17], NN model [18], [19], Gaussian regression model [20], support vector machine (SVM) [21], fuzzy logic model [22], and deep learning method [23], [24]. Although these methods have numerous successful applications in estimating SOH and predicting RUL, they do not consider the dynamic natures of battery aging mechanism, especially the aging dynamics between the two adjacent cycles. This may lead to poor performance in practical applications due to the presence of noise and uncertainty. Besides, the deep knowledge extracted from the charging process is also not considered, which can be applied to optimize the traditional data-driven methods.

Therefore, in order to further improve the accuracy of diagnosis and prediction, some combined model-based and data-driven methods are proposed to overcome the limitations of the aforementioned methods. These methods, include Brownian motion model and PF method [25], [26], exponential model and PF method [27], NN and PF method [28], and SVM and PF method [29], are employed to track the degradation behavior of battery capacity. In these methods, the PF algorithm is conducted based on the physical model; but, it's very difficult to obtain the explicit physical model in practical applications due to the complex nonlinear dynamic feature of LIBs. In addition, these methods have poor performance of long-term prediction because they are not suitable for big data processing. Thus, these methods have some limitations.

Inspired by model fusion methods, in this paper, we proposed a novel data-model fusion method to estimate SOH and predict RUL for LIBs. A LSTM network was first trained to model the complex SOH dynamics according to the discharge voltage, temperature, and current measurements. In addition, it is well known that the NN modeling method represents the nonlinear behavior of a time series well; given this, it was used here to establish the state space model for describing battery capacity degradation mechanism. Then, a SOH estimator combining the aforementioned two models and UKF algorithm was proposed to filter out the noises and reduce the estimation errors. Unlike the traditional model fusion methods, this modeling method fully considers the data information of all sensors, and does not require any physical models. Besides, the deep transcendental knowledge extracted from the charging process is considered in the modeling process. In addition, this proposed method makes full use of the advantage of both machine learning and model-based filtering technique. Thus, it can effectively improve the estimation performance of SOH and RUL of LIBs.

This paper is organized as follows: Review of LSTM network is described in Section II, modeling and methods in Section III, experiments and verification in Section IV, summary and conclusion in Section V.

II. REVIEW OF LSTM NETWORK

As a variant of recursive neural network (RNN), the LSTM network has been widely used in the field of state estimation of LIBs, because it has been proven to have strong robustness to dynamic load, hysteresis, aging, and parameter uncertainty. By using an input gate, a forget gate, and an output gate, the LSTM unit can decide what to remember and what to forget, and is thus capable of dealing with long-term dependencies. In LSTM framework, the output can be represented by the composite function as follows:

$$f_t = \sigma(W_{hf} \cdot h_{t-1} + W_{xf} \cdot x_t + b_f) \tag{1}$$

$$i_t = \sigma(W_{hi} \cdot h_{t-1} + W_{xi} \cdot x_t + b_i) \tag{2}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \tag{3}$$

with

$$\tilde{c}_t = \tanh(W_{hc} \cdot h_{t-1} + W_{xc} \cdot x_t + b_c) \tag{4}$$

$$o_t = \sigma(W_{ho} \cdot h_{t-1} + W_{xo} \cdot x_t + b_o) \tag{5}$$

$$h_t = o_t \cdot \tanh(c_t) \tag{6}$$

where the initial hidden state C_{t_0} is set to a zero matrix; $\sigma(\cdot)$ is the sigmoid function; $tanh(\cdot)$ is the hyperbolic tangent function; f_t, i_t, o_t, c_t are the forget, input, output gates, and memory cell, respectively; h_t is the output of the hidden layer at time t. Every gate can be seen as a single layer fully connected neural network model with a sigmoid function. Its input is a vector, which consists of the time series output of the previous time and the input of the current time. Its output is a real vector between 0 and 1, which is able to control the memory and forgetting degree of previous information and current information. Each gate has its set of network weight matrix and bias that are denoted by W and b. The subscripts of W represent the transformation occurring between the two respective components, such as W_{hf} represents the weight connecting the time series output of the previous time and the forget gate output, W_{xf} represents the weight connecting the input of the current time and the forget gate output, b_f represents the bias of forget gate.

III. MODELING AND METHODS

The main purpose of this study is developed a novel data-model fusion method to provide a prognostic framework for SOH and RUL of LIBs, as indicated in Fig. 1. Herein,

a LSTM network is first developed to represent battery SOH dynamics, a NN-based state-space model is then constructed to describe the aging mechanism of battery. Finally, an UKF algorithm is incorporated with the two models to estimate SOH and predict RUL.

A. LSTM-BASED SOH MODELING

1) DEFINITION OF SOH AND RUL

SOH is a health index to describe the aging status of battery in each charge-discharge cycle, which is usually represented by parameters such as capacity, internal resistance or number of cycles [30]. In this work, the capacity ratio is selected to define battery SOH; and SOH in the kth cycle can be described by the following equation:

$$SOH_k = \frac{C_k}{C_0} \times 100\% \tag{7}$$

where SOH_k and C_k represent the SOH and capacity in the kth cycle, respectively; C_0 is the nominal capacity.

RUL represents the number of cycles that is available before the battery fails. The capacity value of battery failure is called end of life (EOL) threshold, which is often regarded as the capacity reaches 70–80% of the initial value. The definition of EOL is given as follows:

$$EOL = C_0 \cdot \eta \tag{8}$$

where η is a coefficient $\eta \in [0.7, 0.8]$.

The actual and predicted RUL are separately given as:

$$RUL = T_{EOL} - T_t \tag{9}$$

$$\widehat{RUL} = \hat{T}_{EOL} - T_t \tag{10}$$

where T_{EOL} and $\hat{T}_{EOL}$ are the cycles at the actual and predicted EOL, respectively; T_t is the current charge-discharge cycles of battery.

2) LSTM-BASED SOH MODELING

According to Eq. (7), estimating SOH is equivalent to estimate battery's capacity. The capacity training model can be established by using the following equation:

$$C_k = f([x^h_{k,t}, x^h_{k,t+1}, \cdots, x^h_{k,t+w-1}]) \tag{11}$$

where C_k is the actual capacity in the kth cycle; $x_{k,t}$ is a historical input vector at sampling time t from the kth cycle; w is the sliding window (SW) length; $f(\cdot)$ is a LSTM model established by using Eqs. (1)-(6). By using the collected dataset, this LSTM model is easily constructed using the back propagation through time (BPTT) training algorithm according to the modeling error $E_L(t)$. In this paper, the SW length is set as $w \in [1, m]$, and m is the length of $x_{k,t}$. It is important to note that $x_{k,t}$ can be either any of measured voltage, current, temperature and sampling time or any combination of these parameters.

When $i \geq 1$, the on-line prediction of capacity based on LSTM model at cycle number $k+i$ can be obtained as follows:

$$\hat{C}_{k+i} = \hat{f}([x^o_{k+i,t}, x^o_{k+i,t+1}, \cdots, x^o_{k+i,t+w-1}]) \tag{12}$$

where $\hat{C}_{k+i}$ represents the predicted capacity in the $k+i$th cycle; $x^o_{k+i,t}$ is the online observed vector at sampling time t from $k+i$th cycle; $\hat{f}(\cdot)$ is the trained LSTM model. Note that the data structure of $x^o_{k+i,t}$ is the same as that of $x_{k,t}$. When the capacity of battery is obtained, the SOH can then be easily calculated by using model (7).

B. NN-BASED CAPACITY DEGRAATION MODELING

Charging of battery was often carried out in a constant current (CC) and constant voltage (CV) mode. Generally, in the CV stage, the battery current will change in different aging times, which can be considered as exponentially related to the number of cycles. In addition, the charged capacity also changes with the aging times during the CV stage. Hence, an exponential function can be used to escribe the charged current dynamics during the CV stage, which is given as follows [31]:

$$I(t) = \alpha \cdot \exp(-\beta \cdot t) + \gamma \tag{13}$$

where I is the charging current; t is the sampling time; α, β and γ are the model parameters, which need to be identified using the sampling data.

Here, two polynomial function models are used to describe the relationship among the model parameter β, charged capacity Q_{cv}, and the cycle number k, respectively, which are given as follows:

$$\beta_k = a_1 \cdot k^3 + a_2 \cdot k^2 + a_3 \cdot k + a_4 \tag{14}$$

$$Q_{cv,k} = b_1 \cdot k^3 + b_2 \cdot k^2 + b_3 \cdot k + b_4 \tag{15}$$

where β is obtained by model (13); Q_{cv} is the charged capacity during the CV stage, which can be obtained by using ampere hour integration method; $a_1, a_2, a_3, a_4, b_1, b_2, b_3$ and b_4 are the model parameters that need to be identified; k is the cycle number.

Here, the battery capacity C_k is used as the state variable to describe the degradation mechanism of battery. The model parameter β_k in Eq. (14), charged capacity $Q_{cv,k}$ during the CV stage in Eq. (15), and C_{k-1} are taken as the input variables. Thus, the state-space equation can be expressed by using the following equation:

$$C_k = h(C_{k-1},\ \beta_k, Q_{cv,k}) + r_k \tag{16}$$

where C_k is the capacity in the kth cycle; $h(\cdot)$ is an unknown nonlinear function; r_k denotes the modeling error. It's usually very difficult to find a definite analytic expression for model (16) because of the complex electrochemical reaction process.

Hence, a RBF neural network model is used here to approximate model (16) to describe the aging dynamics between the two adjacent cycles because it can achieve high prediction accuracy without requiring prior knowledge. The output of kth cycle is represented as follows:

$$y_k = \sum_{i=1}^{h} \omega_{k,i} \exp(-\frac{1}{2\sigma_k^2} \left\| v_k - c_{k,i} \right\|) \tag{17}$$

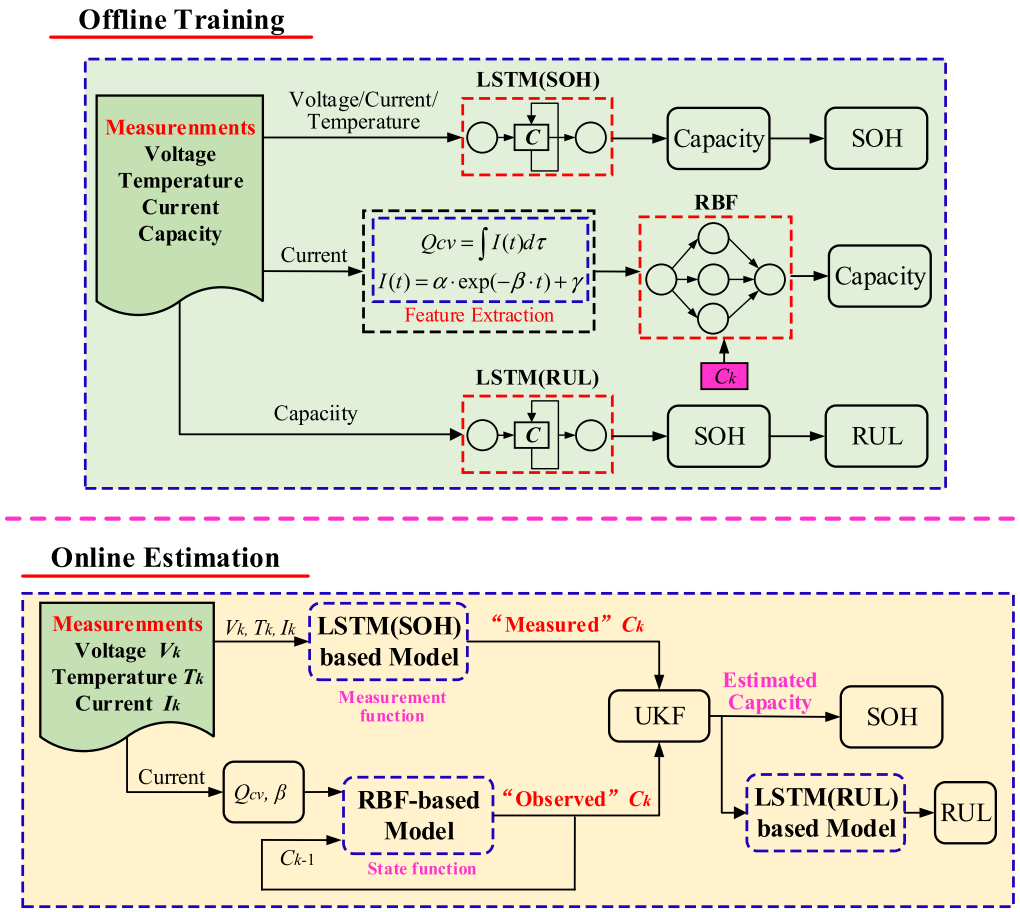

FIGURE 1. Framework of SOH estimation and RUL prediction.

Here, $\omega_{k,i}$ is the weight of the kth cycle connecting hidden layer and output layer; $c_{k,i}$ is the center; $v_k = [C_{k-1}, \beta_k, Q_{cv,k}]$ is the input vector; σ_k^2 is the width; h is the number of hidden neurons. By using the identified data from each cycle, this NN model is easily constructed using the traditional gradient descent algorithm according to the modeling error of $E_R(t)$.

For $n \geq 1$, the on-line prediction of capacity based on NN-based model at cycle number $k+n$ can be updated as follows:

$$\hat{C}_{k+n} = \hat{h}(\hat{C}_{k+n-1},\ \hat{\beta}_{k+n}, \hat{Q}_{cv,k+n}) \tag{18}$$

where $\hat{C}_{k+n}$ is the predicted capacity at cycle number $k+n$; the predicted capacity $\hat{C}_{k+n-1}$, $\hat{\beta}_{k+n}$, and $\hat{Q}_{cv,k+n}$ are used as the online inputs for RBF network; $\hat{h}(\cdot)$ is the trained RBF model.

C. UKF-BASED MODEL FUSION METHOD

UKF is a model-based filtering algorithm on account of discrete system, which specializes in dealing with uncertainty and nonlinearity. This algorithm has a better ability to strip noise out of a stream of data, and can achieve better estimated performance for the nonlinear systems and is easier to implement compared with the EKF method. In UKF approach, UT transformation is used to deal with the nonlinear transfer of mean value and covariance, thus it is no need to compute the Jacobi matrix and perform any form of approximation to the nonlinear system function [32]. Therefore, it can greatly reduce the computational complexity. Hence, an UKF algorithm was used to incorporate with the LSTM network and NN model to filter out the noises and further reduce the estimation errors to obtain an optimized estimation of battery SOH.

1) UKF ALGORITHM

Consider a nonlinear system $y = f(x)$, where x is an n-dimensional state vector with mean vector $\tilde{x}$ and covariance matrix P_x, the posterior statistics of y is determined by $2n+1$ sigma points χ_i with corresponding weights W_i according to:

$$\begin{aligned}
\chi_0 &= \bar{x} \\
\chi_i &= \bar{x} + (n+\lambda)(\sqrt{P_x})_i \\
\chi_{i+n} &= \bar{x} - (n+\lambda)(\sqrt{P_x})_{i-n} \\
W_0^m &= \lambda/(n+\lambda) \\
W_0^c &= \lambda/(n+\lambda) + (1-\alpha^2+\beta) \\
W_i^m &= W_i^c = 1/\{2(n+\lambda)\}
\end{aligned} \tag{19}$$

where $\lambda = \alpha(n+k) - n$ is the scaling factor, and adjusting its value can improve the accuracy; $(\cdot)_i$ represents the ith column of the matrix; α determines the distribution of sigma points around the mean vector, which is usually set as $\alpha \in [0, 1]$; and k is usually set as $k = 3 - n$; β is the state distribution parameter.

In this paper, α and β are empirically chosen as 0.01 and 2.The sigma points are projected through the nonlinear system process $y_i = f(\chi_i)$. The mean and covariance of y are obtained as follows:

$$\bar{y}_i = \sum_{i=0}^{2n} W_i y_i \tag{20}$$

$$P_y = \sum_{i=0}^{2n} W_i^c (y_i - y)(y_i - y)^T \tag{21}$$

Generally, UKF has a predictive updating structure. For each time step k, firstly, the state mean vector and covariance matrix are predicted based on the state vector in step k-1; then, in the update step, the predicted mean vector and covariance matrix are updated based on the new measured values obtained in the current time step.

2) SOH ESTIMATION BASED ON THE PROPOSED METHOD

LIB is a strongly nonlinear system due to its complex electrochemical reaction process, which leads to different properties under different conditions; this thus brings great difficulties to the observation. Considering the advantages of data-driven methods in modeling process, a data-model fusion method is proposed here to estimate battery SOH and predict RUL. In particular, this proposed model fusion method is performed only using the information of all sensors. In the process of online estimation SOH, model (11) is chosen as the measurement function, and the measurement vector is the outputs of LSTM network. Model (16) is used as the state transition function, and the state vector C is the output of model (18). The UKF algorithm is used to filter out the measurement noise and reduce some other uncertainties, and update the estimation vector of SOH. The state-space model is described as follows:

State function:

$$\begin{aligned} C_k &= h(C_{k-1},\ \beta_k, Q_{cv,k}) \\ &= \sum_{i=1}^{h} \omega_{k,i} \exp(-\frac{1}{2\sigma_k^2} \left\| v_k - c_{k,i} \right\|) + q \end{aligned} \tag{22}$$

Measurement function:

$$\mathrm{LSTM}_k = f([x_{k,t}^h, x_{k,t+1}^h, \cdots, x_{k,t+w-1}^h]) = C_k + p \tag{23}$$

where $q \sim N(0, Q)$ and $p \sim N(0, R)$ are the Gaussian state noise and measurement noise, respectively. This data-model fusion method fully considers the characteristic parameters of charging and discharging process such as current, voltage, temperature and capacity. At the same time, the UKF algorithm is applied to reduce the measurement noise and uncertainty. Therefore, this proposed method can effectively improve the estimation accuracy of SOH.

3) RUL PREDICTION BASED ON THE PROPOSED METHOD

When RUL prediction is carried out, the prediction model is established using the history data of battery; and then the model is applied to conduct multi-step prediction. Assuming that the length of SW is w, the number of real values in the previous sliding window is k; and then the predicted SOH value at t step are given as follows:

$$M(\cdot) = g([s_t^h, s_{t-1}^h, \cdots, s_{t-w+1}^h]) \tag{24}$$

$$\hat{s}_{t+1} = M([\hat{s}_t, \cdots, \hat{s}_{t-w+k+1}, s_{t-w+k}^o, \cdots, s_{t-w+1}^o]) \tag{25}$$

where $g(\cdot)$ denotes an unknown nonlinear function; $M(\cdot)$ denotes the trained model; s_t^h denotes the history data of SOH at t prediction step; $\hat{s}_{t+1}$ is the online prediction SOH at $t+1$ step. Note that the $t+1$ prediction step represents the $t+1$th cycle, not the sampling time in a cycle. s_{t-w+k}^o is the observed value of SOH, which is updated by models (22) and (23). Here, the LSTM network is selected to train the model (24). Based on this, the SOH value at T_t-th charging-discharging cycle can be predicted. If the predicted SOH reaches the EOL threshold, the battery RUL can be obtained by model (10).

It should be noted that the proposed method in this paper is trained on dataset obtained from 18650 LIBs, and the same architecture can be trained on a completely different battery. However, the architecture and network parameters, such as the learning rate, will not change. In practical application, if it is a completely different type of battery, it needs to retrain the network framework. If the battery type is the same, then the training results in this paper can be directly applied.

IV. EXPERIMENTS AND VERIFICATION

In this section, we extracted the features that effectively reflect the aging natures of LIBs, and the collected data was conducted to model and verify the effectiveness of the proposed modeling architecture.

A. DATA DESCRIPTION

The 18650 LIBs used in this paper are from the NASA AmesPrognostics Center of Excellence with the rated capacity of 2Ah. And the chemistry of li-ion cell is indicated in Table 1 [33]. The experimental data were collected from the battery prognostics test bed, including power supply, DC electronic load, electrochemical impedance spectroscopy (EIS), voltmeter, thermocouple sensor, thermal chamber, peripheral component interconnect (PCI) extensions for instrumentation chassis based on data acquisition, and experimental control conditions. The LIBs used here were run through three different operational profiles (charge, discharge, and impedance) at different ambient temperatures. Charging process was carried out in a constant current (CC) mode at 1.5A until the battery voltage reached 4.2V, then continued in a constant voltage (CV) mode until the charge current dropped to 20mA. After

TABLE 1. Li-ion cell chemistry.

Positive Electrode	8wt% PVDF binder
	4wt% SFG-6 graphite
	4wt% carbon black
	84wt% $LiNi_{0.8}Co_{0.15}Al_{0.05}O_2$
Negative Electrode	8wt% PVDF binder
	92wt% MAG-10 graphite
Electrolyte	1.2 M $LiPF_6$ in EC: EMC (3:7 wt%)
Separator	25μm thick PE (Celgard)

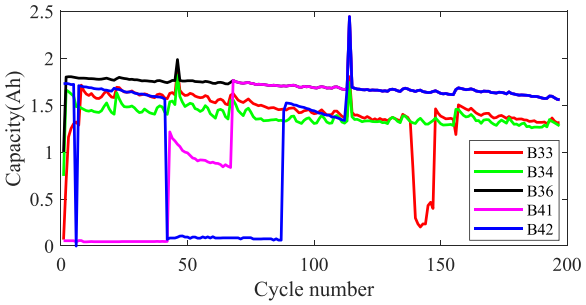

FIGURE 2. Discharged capacity versus cycle number.

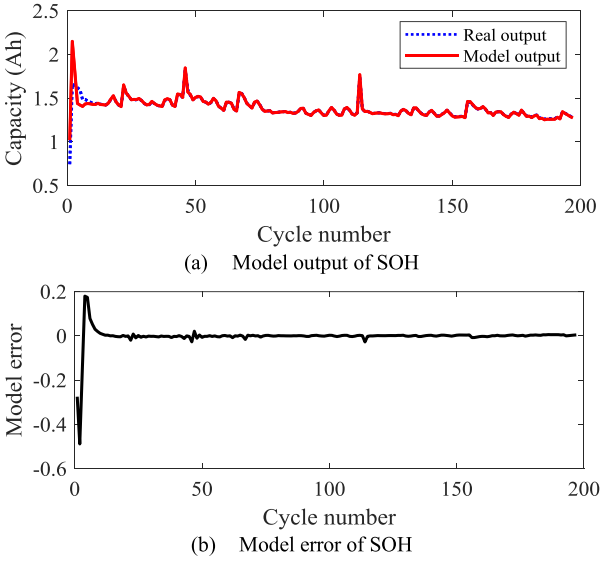

FIGURE 3. SOH modeling process using LSTM.

that, different exciting current were used to discharge until the battery voltages reached 2.7, 2.5, 2.2 and 2V. Note that all the charging models were the same, but the cutoff voltage for different batteries were different.

In order to verify the effectiveness of the proposed method, the batteries (B33, B34, B36, B41, and B42) run at two different ambient temperatures are selected here. Where batteries B33, B34 and B36 were run at temperature of 24°, batteries B41 and B42 were run at 4°C. The features of discharge process, including current, temperature, voltage, and capacity, were extracted. During the experiments, the number of cycles of batteries B41and B42 were shorter because they discharged at a lower ambient temperature. Thus, in order to increase the number of cycles and fully verify the effectiveness and correctness of the proposed architecture, the data sets of batteries B41and B42 were combined with B36 datasets. In the experiments, a total of 197 cycles were collected, and each cycle contained a charging-discharge process. The degeneration capacity curves for all the selected batteries were indicated in Fig. 2.

As shown in Fig. 2, it is obvious that the capacity of all the selected LIBs degrade with the cycle number increases, but the degradation rates of all batteries are different. This is because the discharge capacities are related to several parameters, such as charge current and voltage, discharge voltage, current and temperature, etc. In addition, it can also be seen the battery capacities are not always declining, but suddenly rising in a relatively smooth stage at some times. This is because the ion concentration inside the battery tends to balance due to the diffusion effect. This phenomenon is called battery self-charging effect, which is mostly obvious in type B33 and B34. Besides, the maximum observed capacity of B33 is lower than 1.4Ah (EOL) at 143th cycle, which implies the serious performance degradation.

B. OFFLINE TRAINING FOR THE PROPOSED METHOD

1) TRAINING FOR SOH ESTIMATING MODEL

In this section, the discharge datasets of B34, including the temperature, voltage, and current measurements, were used to train SOH estimating model, and the rest battery datasets were used for online prediction. During the training process, the w is set as 20; the data sets of the first 80 cycles were used for modeling, and the remaining data were used for validation. For the LSTM network, according to the practical experience and prior process knowledge, the initial learning rate and break error were set as 1e-2 and 1e-6, respectively. And the max iterations was set as 300. The model output and absolute error over the whole training data and validation data are shown in Fig. 3.

Two evaluation criteria, including root mean square error (RMSE) and mean absolute percentage error (MAPE), are used here to evaluate the estimating accuracy of SOH, which are given as follows:

$$RMSE = \sqrt{\sum_{k=1}^{m}(C(k) - \hat{C}(k))^2/m} \tag{26}$$

$$MAPE = \frac{1}{m}\sum_{k=1}^{m}\left|\frac{C(k) - \hat{C}(k)}{\hat{C}(k)}\right| \tag{27}$$

where m is the total cycle number, $C(k)$ and $\hat{C}(k)$ represent the real and predicted capacity, respectively.

2) TRAINING FOR CAPACITY DEGRADATION MODEL

Here, the charge datasets of B34 were used for training the capacity degradation model, and the rest battery datasets were used for online prediction. Firstly, the current data of CV

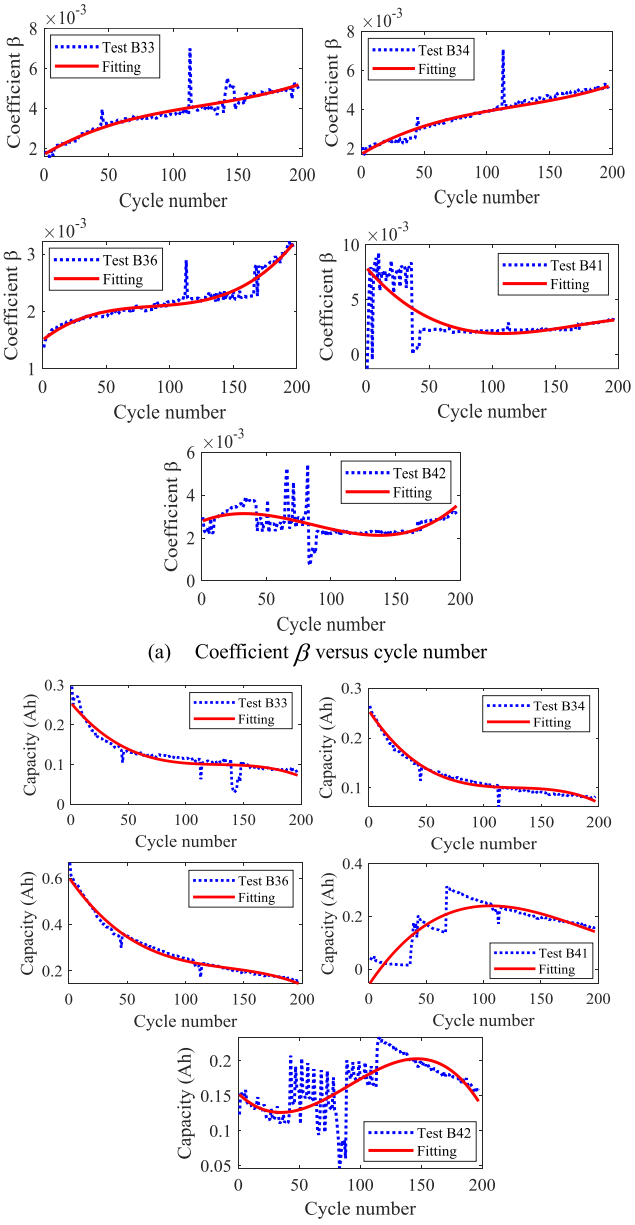

FIGURE 4. Coefficient β and charged capacity during the CV stage.

stage was used to identify the model parameters β and Q_{cv}. Then, we fitted the nonlinear relationship between β and Q_{cv} with respect to the cycle number according to models (14) and (15). The fitting results of these two parameters were shown in Fig. 4. Then, a RBF network was used to train the capacity degradation model. In the training process, the data sets of the first 80 cycles were used for modeling, while the remaining data sets were used for validation. According to the practical experience and prior process knowledge, the initial learning rate and break error were set as 1e-2 and 1e-4, respectively. And the max iterations was set as 200. The model output and absolute error over the whole training data and validation data are shown in Fig. 5.

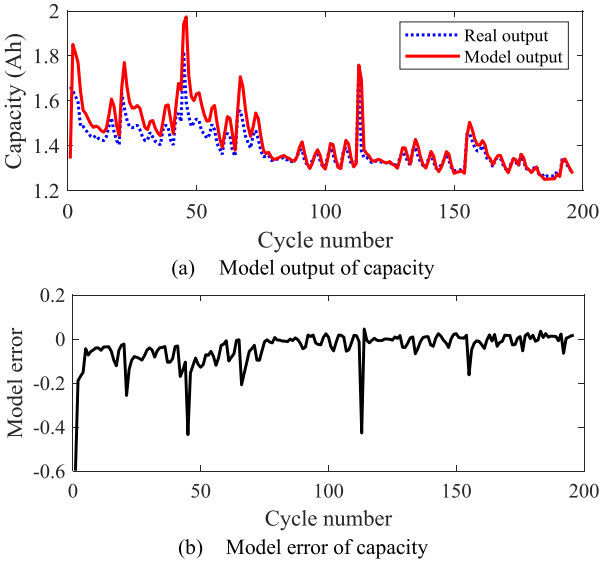

FIGURE 5. Capacity modeling process using RBF.

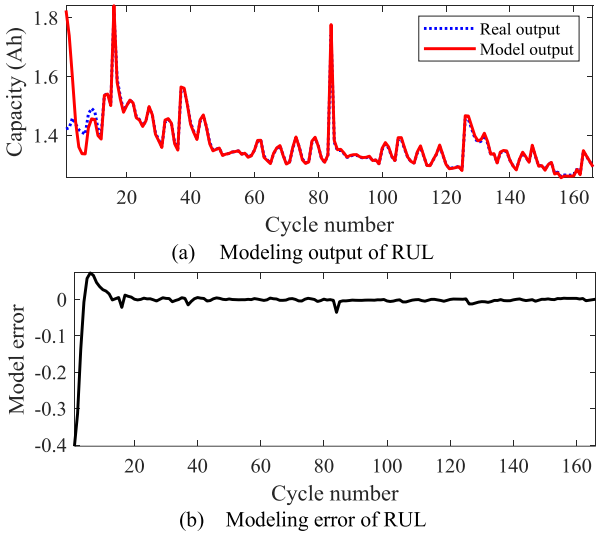

FIGURE 6. RUL modeling process using LSTM.

3) TRAINING FOR RUL PREDICTION MODEL

There is no need to perform the inputs selection for RUL prediction model due to its capacity-to-capacity structure. Its training process was the same as the SOH modeling process. The w is set as 30. The data sets of the first 80 cycles were used for modeling, and the remaining data sets were used for validation. The initial learning rate and break error were set as 1e-2 and 1e-6, respectively, And the max iterations was set as 300. The model output and absolute error over the whole training data and validation data are shown in Fig. 6.

C. ONLINE VERIFICATION BASED ON THE PROPOSED METHOD

1) ONLINE ESTIMATING FOR SOH

By using models (12) and (18), the datasets from B33, B36, B41, and B42 were used to evaluate the performance of

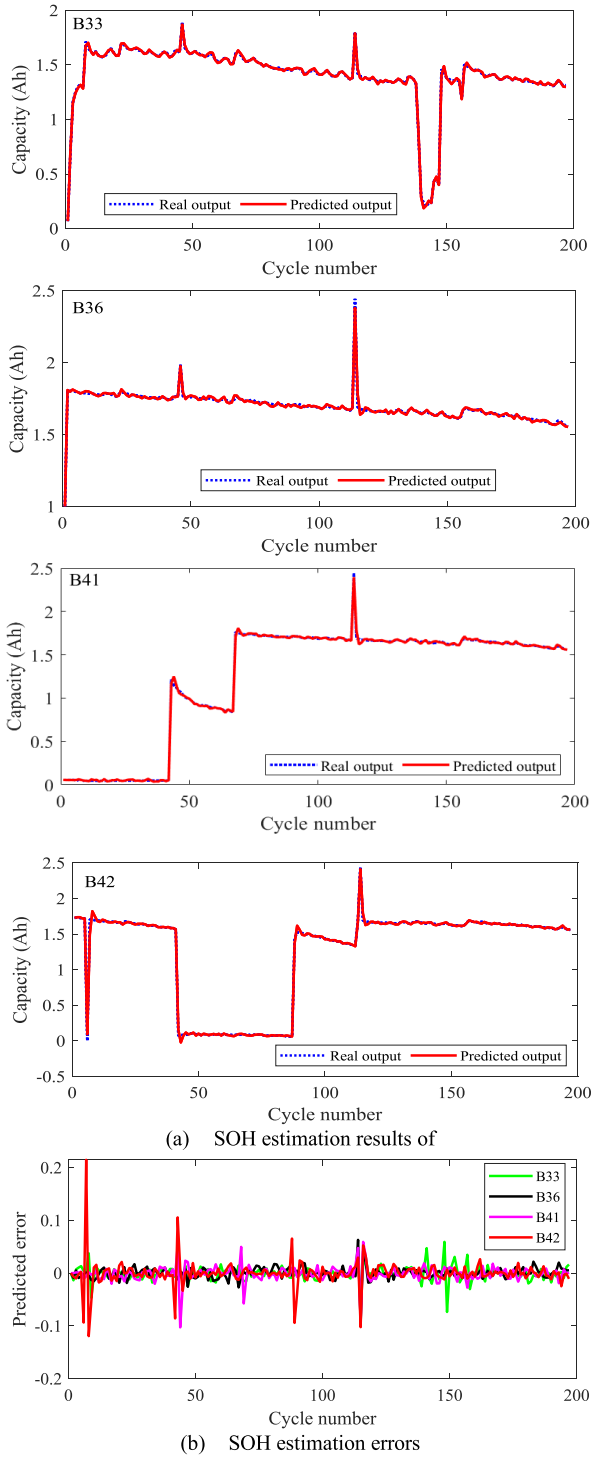

FIGURE 7. SOH prediction results and estimation errors.

the proposed method. SOH estimation results and estimation errors over the whole cycle numbers for the four cells are indicated in Fig. 7.

As shown in Fig. 7, it is obvious that the proposed method was capable of tracking battery SOH dynamics for both high and low ambient temperatures with a good approximation

TABLE 2. Comparison of SOH estimation accuracy on RMSE.

Method	B33	B36	B41	B42
NN [18]	9.716e-4	7.125e-4	0.0036	0.0031
LSTM [23]	9.333e-4	6.858e-4	0.0024	0.0022
Proposed method	1.777e-4	1.521e-4	2.523e-4	6.899e-4

TABLE 3. Comparison of SOH estimation accuracy on MAPE.

Method	B33	B36	B41	B42
NN [18]	0.5732	0.0184	0.5744	3.6281
LSTM [23]	0.0537	0.0157	0.4431	3.3119
Proposed method	0.0079	0.0046	0.0324	0.0541

error. This means that the proposed modeling method can be used to estimate SOH under a wide temperature range.

The predictive capability of this proposed method was also compared with conventional NN method [18], and LSTM method [23]. All these methods were carried out by using the same experimental data, model parameters, and data structure. For comparison, the error indexes of different modeling methods, such as RMSE and MAPE, are calculated; where the smaller the error indexes, the better the prediction performance. The error indexes RMSE and MAPE for the test data of different methods are shown in Tables 2 and 3, respectively. It is clear that the proposed method performed substantially better when compared with other existing methods due to its smaller error indexes. This is because the proposed modeling method fully considers the data information of all sensors, and does not require any physical model. Besides, it considers the dynamic natures of battery aging mechanism, especially the aging dynamics between the two adjacent cycles. In addition, it is also because that the proposed method is carried out by taking advantage of both machine learning and model-based filtering technique, which suppresses the system measurement noises and further reduces the estimation errors. Thus, it can obtain a better estimation ability than the traditional modeling method, such as NN method and LSTM method.

2) ONLINE PREDICTING FOR RUL

Then, we will discuss the battery RUL prediction results. In this case studies, the B34 and B36 are more suitable to select as the research object according to Fig. 2. Here, the EOL values of B34 and B36 were set to 70% and 81% of the nominal capacity, respectively. The cycle numbers of 40 and 60 were selected as the starting point (SP) to predict battery RUL, respectively. The predicted capacities of these two cells are shown in Fig. 8. Absolute error (AE) was used to evaluate the RUL prediction results of different methods as follows:

$$AE = RUL - \widehat{RUL} \tag{28}$$

where RUL and $\widehat{\text{RUL}}$ are the actual and predicted RUL, respectively.

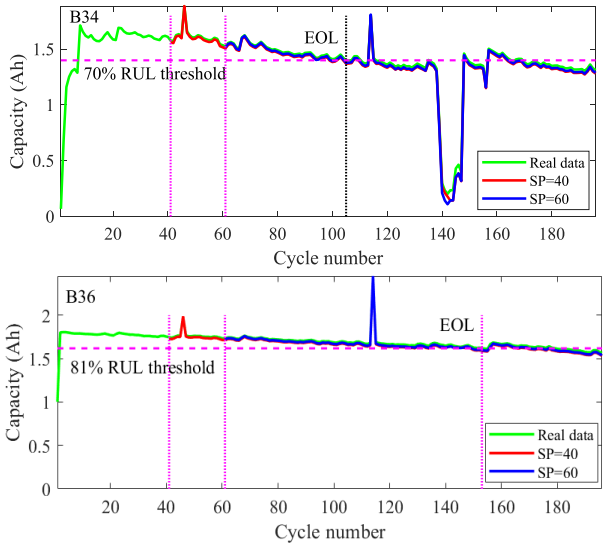

FIGURE 8. RUL prediction results of B34 and B36.

TABLE 4. Comparison of RUL prediction results.

Method	B34		B36	
	SP=40	SP=60	SP=40	SP=60
NN [18]	54	36	103	84
LSTM [23]	56	37	106	86
Proposed method	59	40	108	89
Real RUL	65	45	113	93

As shown in Fig. 8, it is clear that the proposed method effectively established battery SOH dynamics for different SP. Further, this proposed modeling method has good prediction ability for battery RUL.

In addition, we also conducted a comparison between the proposed method and the commonly used RUL prediction algorithms for different SP. The error index AE of different SP for the test data was calculated, which are shown in Table 4 and Table 5, respectively. From these two Tables, it is obvious that the proposed method performed a good predictive ability for battery RUL as indicated by its smaller error index. This is because the proposed method uses full data information of all sensors and does not require any battery models. In addition, this proposed approach fully takes into account the advantage of both machine learning and model-based filtering technique, which further reduces the estimation errors. Thus, it can obtain a better estimation ability than the traditional modeling method.

In summary, both of our experimental schemes indicated that the proposed modeling method performed very well in estimating SOH and predicting RUL of LIBs. The superiority of this proposed method compared with other commonly used algorithms can be further classified from the following several aspects: First, it uses full data information from all sensors, and does not require any physical models, which reduces the mode error caused by battery modeling; in addition, it full accounts for the dynamic natures of battery aging

TABLE 5. Comparison of RUL prediction errors.

Method	B34		B36	
	SP=40	SP=60	SP=40	SP=60
NN [18]	11	9	10	9
LSTM [23]	9	8	7	7
Proposed method	6	5	5	4

mechanism, especially the aging dynamics between the two adjacent cycles. Moreover, this proposed modeling method take full advantage of both machine learning and model-based filtering technique, which suppresses the system measurement noises and further reduces the estimation errors.

V. CONCLUSION

In this paper, we propose a novel data-model fusion method to estimate SOH and predict RUL for LIBs by using machine learning and model-based filtering technique. This proposed modeling method fully took into account the dynamic natures of battery aging mechanism, especially the aging dynamics between the two adjacent cycles. Moreover, the proposed modeling method had a model free and data-driven features, which reduced the model errors and uncertain interference in battery modeling. In addition, an UKF algorithm was incorporated with the LSTM and NN to filter out the noises in the network output, and further improve the accuracy of SOH and RUL. The effectiveness of this proposed approach was verified by using four cell cycle experiments, and its superior estimating ability was demonstrated by comparing with other traditional methods. In the future, this method will be extended to SOC and SOH co-estimation considering battery aging properties.

REFERENCES

[1] Y. Xie, X.-J. He, X.-S. Hu, W. Li, Y.-J. Zhang, B. Liu, and Y.-T. Sun, "An improved resistance-based thermal model for a pouch lithium-ion battery considering heat generation of posts," *Appl. Thermal Eng.*, vol. 164, Jan. 2020, Art. no. 114455.
[2] Z. Song, H. Wang, J. Hou, H. F. Hofmann, and J. Sun, "Combined state and parameter estimation of lithium-ion battery with active current injection," *IEEE Trans. Power Electron.*, vol. 35, no. 4, pp. 4439–4447, Apr. 2020.
[3] J. Wu, Y. Wang, X. Zhang, and Z. Chen, "A novel state of health estimation method of Li-ion battery using group method of data handling," *J. Power Sources*, vol. 327, pp. 457–464, Sep. 2016.
[4] X. Hu, F. Feng, K. Liu, L. Zhang, J. Xie, and B. Liu, "State estimation for advanced battery management: Key challenges and future trends," *Renew. Sustain. Energy Rev.*, vol. 114, Oct. 2019, Art. no. 109334.
[5] A. El Mejdoubi, A. Oukaour, H. Chaoui, H. Gualous, J. Sabor, and Y. Slamani, "State-of-charge and state-of-health lithium-ion batteries' diagnosis according to surface temperature variation," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2391–2402, Apr. 2016.
[6] 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, Jan. 2014.
[7] A. V. Virkar, "A model for degradation of electrochemical devices based on linear non-equilibrium thermodynamics and its application to lithium ion batteries," *J. Power Sources*, vol. 196, no. 14, pp. 5970–5984, Jul. 2011.
[8] Z. Wei, J. Zhao, D. Ji, and K. J. Tseng, "A multi-timescale estimator for battery state of charge and capacity dual estimation based on an online identified model," *Appl. Energy*, vol. 204, pp. 1264–1274, Oct. 2017.
[9] Y. Yuan, H.-T. Zhang, Y. Wu, T. Zhu, and H. Ding, "Bayesian learning-based model-predictive vibration control for thin-walled workpiece machining processes," *IEEE/ASME Trans. Mechatronics*, vol. 22, no. 1, pp. 509–520, Feb. 2017.

[10] S. Schwunk, N. Armbruster, S. Straub, J. Kehl, and M. Vetter, "Particle filter for state of charge and state of health estimation for lithium-iron phosphate batteries," *J. Power Sources*, vol. 239, no. 1, pp. 705–710, 2013.
[11] S. Sepasi, R. Ghorbani, and B. Y. Liaw, "Inline state of health estimation of lithium-ion batteries using state of charge calculation," *J. Power Sources*, vol. 299, pp. 246–254, Dec. 2015.
[12] W. He, N. Williard, M. Osterman, and M. Pecht, "Prognostics of lithium-ion batteries based on Dempster–Shafer theory and the Bayesian Monte Carlo method," *J. Power Sources*, vol. 196, no. 23, pp. 10314–10321, Dec. 2011.
[13] I.-S. Kim, "A technique for estimating the state of health of lithium batteries through a dual-sliding-mode observer," *IEEE Trans. Power Electron.*, vol. 25, no. 4, pp. 1013–1022, Apr. 2010.
[14] J. Wei, G. Dong, and Z. Chen, "Lyapunov-based state of charge diagnosis and health prognosis for lithium-ion batteries," *J. Power Sources*, vol. 397, pp. 352–360, Sep. 2018.
[15] H. Chaoui, N. Golbon, I. Hmouz, R. Souissi, and S. Tahar, "Lyapunov-based adaptive state of charge and state of health estimation for lithium-ion batteries," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1610–1618, Mar. 2015.
[16] D. Wang and K.-L. Tsui, "Brownian motion with adaptive drift for remaining useful life prediction: Revisited," *Mech. Syst. Signal Process.*, vol. 99, pp. 691–701, Jan. 2018.
[17] Q. Zhai and Z.-S. Ye, "RUL prediction of deteriorating products using an adaptive Wiener process model," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 2911–2921, Dec. 2017.
[18] J. Wu, C. Zhang, and Z. Chen, "An online method for lithium-ion battery remaining useful life estimation using importance sampling and neural networks," *Appl. Energy*, vol. 173, pp. 134–140, Jul. 2016.
[19] G. Ma, Y. Zhang, C. Cheng, B. Zhou, P. Hu, and Y. Yuan, "Remaining useful life prediction of lithium-ion batteries based on false nearest neighbors and a hybrid neural network," *Appl. Energy*, vol. 253, Nov. 2019, Art. no. 113626.
[20] D. Yang, X. Zhang, R. Pan, Y. Wang, and Z. Chen, "A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve," *J. Power Sources*, vol. 384, pp. 387–395, Apr. 2018.
[21] V. Klass, M. Behm, and G. Lindbergh, "A support vector machine-based state-of-health estimation method for lithium-ion batteries under electric vehicle operation," *J. Power Sources*, vol. 270, pp. 262–272, Dec. 2014.
[22] A. J. Salkind, C. Fennie, P. Singh, T. Atwater, and D. E. Reisner, "Determination of state of charge and state of health of batteries by fuzzy logic methodology," *J. Power Sources*, vol. 80, nos. 1–2, pp. 293–300, 1999.
[23] P. Li, Z. Zhang, Q. Xiong, B. Ding, J. Hou, D. Luo, Y. Rong, and S. Li, "State-of-health estimation and remaining useful life prediction for the lithium-ion battery based on a variant long short term memory neural network," *J. Power Sources*, vol. 459, May 2020, Art. no. 228069.
[24] Y. D. Tan and G. K. Zhao, "A novel state-of-health prediction method for lithium-ion batteries based on transfer learning with long short-term memory network," *IEEE Trans. Ind. Electron.*, vol. 67, no. 10, pp. 8723–8731, 2020.
[25] G. Dong, Z. Chen, J. Wei, and Q. Ling, "Battery health prognosis using Brownian motion modeling and particle filtering," *IEEE Trans. Ind. Electron.*, vol. 65, no. 11, pp. 8646–8655, Nov. 2018.
[26] G. Dong, F. Yang, Z. Wei, J. Wei, and K.-L. Tsui, "Data-driven battery health prognosis using adaptive Brownian motion model," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4736–4746, Jul. 2020.
[27] L. Zhang, Z. Mu, and C. Sun, "Remaining useful life prediction for lithium-ion batteries based on exponential model and particle filter," *IEEE Access*, vol. 6, pp. 17729–17740, 2018.
[28] C. Zhang, Y. Zhu, G. Dong, and J. Wei, "Data-driven lithium-ion battery states estimation using neural networks and particle filtering," *Int. J. Energy Res.*, vol. 43, no. 14, pp. 8230–8241, Aug. 2019.
[29] J. Wei, G. Dong, and Z. Chen, "Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5634–5643, Jul. 2018.
[30] D. Liu, J. Pang, J. Zhou, Y. Peng, and M. Pecht, "Prognostics for state of health estimation of lithium-ion batteries based on combination Gaussian process functional regression," *Microelectron. Rel.*, vol. 53, no. 6, pp. 832–839, Jun. 2013.
[31] A. Eddahech, O. Briat, and J.-M. Vinassa, "Determination of lithium-ion battery state-of-health based on constant-voltage charge phase," *J. Power Sources*, vol. 258, pp. 218–227, Jul. 2014.
[32] F. Yang, S. Zhang, W. Li, and Q. Miao, "State-of-charge estimation of lithium-ion batteries using LSTM and UKF," *Energy*, vol. 201, Jun. 2020, Art. no. 117664.
[33] K. Goebel, B. Saha, A. Saxena, J. R. Celaya, and J. P. Christophersen, "Prognostics in battery health management," *IEEE Instrum. Meas. Mag.*, vol. 11, no. 4, pp. 33–40, Aug. 2008.

XIANGBO CUI (Graduate Student Member, IEEE) received the B.E. degree from the College of Mechanical Engineering, Xi'an University of Science and Technology, Xi'an, China, in 2013, and the M.E. degree from the School of Mechanical and Electrical Engineering, Xiamen University of Technology, Xiamen, China, in 2017. He is currently pursuing the Ph.D. degree in mechanical engineering with the School of Mechanical and Electrical Engineering, Central South University, Changsha, China.

His current research interests include machine learning, battery management systems, and intelligent control.

TETE HU received the B.E. degree in mechanical engineering from the Department of Mechanical Engineering, Hefei University, Hefei, China, in 2013, and the M.E. degree in mechanical engineering from the School of Mechanical Engineering and Electronic Information, China University of Geosciences, Wuhan, China, in 2017. He is currently pursuing the Ph.D. degree in mechanical engineering with the School of Mechanical and Electrical Engineering, Central South University, Changsha, China.

His current research interests include machine learning, design and control of smart material-based actuators, and their use in the field of soft and bio-inspired robotics.