Model migration neural network for predicting battery aging trajectories

Xiaopeng Tang, Kailong Liu, *Member*, *IEEE*, Xin Wang, Furong Gao, James Macro, and W. Dhammika Widanage, *Member*, *IEEE*

***Abstract*—Accurate prediction of batteries' future degradation is a key solution to relief users' anxiety on battery lifespan and electric vehicle's driving range. Technical challenges arise from the highly nonlinear dynamics of battery aging. In this paper, a feed-forward migration neural network is proposed to predict the batteries' aging trajectories. Specifically, a base model that describes the capacity decay over time is first established from the existed battery aging dataset. This base model is then transformed by an input-output slope-and-bias-correction (SBC) method structure to capture the degradation of target cell. To enhance the model's nonlinear transfer capability, the SBC-model is further integrated into a four-layer neural network, and easily trained via the gradient correlation algorithm. The proposed migration neural network is experimentally verified with four different commercial batteries. The predicted RMSEs are all lower than 2.5% when using only the first 30% of aging trajectories for neural network training. In addition, illustrative results demonstrate that a small size feed-forward neural network (down to 1-5-5-1) is sufficient for battery aging trajectory prediction.**

***Index Terms*—Electric Vehicle, Lithium-ion Battery Management, Model Migration, Neural Network, Aging Trajectory Prediction**

STATEMENT OF ORIGINALITY

This work is an original work that has not been previously submitted, and is not under consideration for publication elsewhere.

I. INTRODUCTION

Lithium-ion (Li-ion) batteries have gained significant market share as the main power source for electrical vehicles (EVs) [1] or hybrid electrical vehicles (HEVs) [2], owing to their high energy density and high efficiency [3]. However, battery health is a key element to affect battery performance regarding the applications such as charging management [4,5], energy management [6], thermal management [7,8] and equalization management [9,10]. To ensure the reliability and safety of battery operation, effective health prognostic solutions are required. In response to this challenge, numerous approaches on the battery state-of-health (SoH) estimations have been designed to provide the current information of battery health [11,12]. However, SoH is not enough for further energy scheduling and management due to the fact that EV users require to know the remaining service life the batteries still have [13,14]. It is vital therefore to develop effective approaches for battery aging trajectory predictions, helping to guarantee that batteries are able to operate within the reliable conditions and reduce the users' anxieties on battery duration [15,16].

To achieve effective predictions of battery aging trajectories, various approaches have been proposed in the literature. These approaches can generally be divided into three categories, first-principle model based methods, empirical model based methods and data-driven model based methods.

For first-principle model based methods, a suitable electrochemical model that describes the knowledge of battery's dynamic and aging, along with several observers, are utilized to predict the battery degradation trajectories [17]. Although battery electrochemical dynamics can be well described by these methods, some problems still exist as: 1) there exists a large amount of parameters for this type of model, which are difficult to be identified accurately [18]. 2) these electrochemical models generally contain partial differential equations, leading to large computational and memory consumption [19]. Therefore, the popularity of first-principle-based methods in EV applications is lower than that of the empirical or data-driven based methods [20].

Rather than using complex electrochemical models, empirical model based methods portray the battery capacity degradation over time with mathematic functions such as single exponential function [21], dual exponential function [22], power function [23], hybrid linear function [24], and polynomial function [25]. Then different data-fitting techniques such as particle filters [22], Kalman filters [26] or some offline optimization algorithms [27,28] are employed to identify the models' parameters. Due to their relative simple structure, these methods are easy and straightforward to be implemented in real applications. However, the parameters in these simplified models usually lack physical meanings. It would be difficult to add reasonable constraints during the identification process to prevent potential over-fitting [29]. As a result, this kind of methods would be sensitive to the noise and easy to diverge in the prediction phase [30]. In

Manuscript received 17-Nov-2019; revised 28-Jan-2020; and accepted 02-Mar-2020. This work was supported by the National Natural Science Foundation of China project (61433005), EU funded project under grant No. 685716, Hong Kong Research Grant Council (16207717), as well as Guangdong Scientific and Technological Project (Grant 2017B010120002). (*Corresponding authors: Kailong Liu and Furong Gao.*)

X. Tang, X. Wang and F. Gao are with the Department of Chemical and Biological Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR, and F. Gao is also with Guangzhou HKUST Fok Ying Tung Research Institute, Guangzhou 511458, China (xtangai@connect.ust.hk; wangx@connect.ust.hk; kefgao@ust.hk).

K. Liu, James Macro, and W. Dhammika Widanage are with the WMG, The University of Warwick, Coventry, CV4 7AL, United Kingdom (kliu02@qub.ac.uk, kailong.liu@warwick.ac.uk; James.Marco@warwick.ac.uk; Dhammika.Widanalage@warwick.ac.uk).

addition, the prediction performance is highly dependent on the mathematical form of the empirical function, the selection of which generally requires trial-and-error efforts [20].

Data-driven approaches, which generate the future predictions based on the collected data, have also been widely adopted. Through using advanced machine learning techniques such as support vector regression [26], Gaussian process regression [31,32] and neural network (NN) [33]–[35], these methods do not assume any battery degradation mechanism a-priori, and turn out to be suitable for general battery types. With the development of vehicle data centers and cloud platforms [36], data-driven methods would gain a foreseeable popularity. However, there still exists some open technical issues that should be considered when applying such methods for the battery aging trajectory prediction. First, the knee-point effect [37]. It is known that the degradation rate of a battery is nonlinear. For some battery types, there exists a point (known as the knee point) that the degradation rates before and after are significantly different. In this case, the knee point information has to be included in the training dataset to make the predictions more accurate. Noting that battery degradation takes several years, an overly long experimental time is required to accumulate such data. Second, the computational issue. As the battery decay can be seen as a time-series process, therefore a state-of-the-art technique is to use NNs that contain internal feed-backs such as long short-term memory (LSTM) recurrent neural network (RNN) [33] to model battery time-series aging dynamics. In general, this NN is accurate, but computationally complex due to its gates structure. Third, the network initialization. Various initialization methods would lead to different network training results [38]. In this case, even if the same dataset is used to train NN, various aging trajectories could be generated, resulting in the increase of users' anxiety on battery lifespan [23]. However, as a pure data-driven method, how to select a suitable initialization solution for battery aging trajectory prediction is still an open issue [20].

To improve the performance of existing data-driven methods, the model migration technique from the field of injection molding processes [39] has been introduced to the research field of battery management by Tang *et al.* in 2019 [40]. In the reported method, a base model developed from the existing dataset is transformed through the slope-and-bias-correction (SBC) method to model the aging process of a target battery. In the SBC-model migration approach, the base model that covers the entire aging information can be fully utilized. Therefore, as long as the base model matches the degradation curve of the target cell, this method is able to achieve satisfactory performance through only training based on the first 30% of aging trajectory. However, in general cases, the degradation process of a target cell would be different from that of the base model. Since SBC can only provide linear transforms of the input and output from base model, there is no guarantee that the nonlinear relation between the base and target processes can be reliably compensated.

Based upon the above discussions and driven by the purpose of predicting nonlinear battery aging trajectories with limited training data, a feed-forward migration neural network that merges the NN and the SBC-model migration is proposed. Several key features make this work distinguishable from the relevant literature. First, by integrating the SBC-migration model into the NN, the knee point does not need to be covered by the online accumulated training dataset. Second, the nonlinear fitting performance is significantly improved with the integration of SBC-migration technique and NN structure. Third, a small-size feed-forward NN (down to 1-5-5-1) trained by the light-weight gradient correction algorithm is sufficient for aging trajectory prediction. Further, the proposed migration NN is not sensitive to the tunable parameters such as learning rate or initial weight metrics.

The rest of this article is organized as follows. Section 2 gives a detailed description of the datasets. Then the elaborations of fundamentals behind the conventional empirical model-based prediction method, the slope-and-bias-correction-based model migration method, the proposed model migration neural network method, and the benchmarking algorithms are presented in Section 3. Section 4 gives the detailed experimental verification, and finally, Section 5 summaries this study.

II. Description of the Datasets

To verify the effectiveness of the proposed method, four battery aging datasets are utilized, while each dataset is composed of aging data from two batteries.

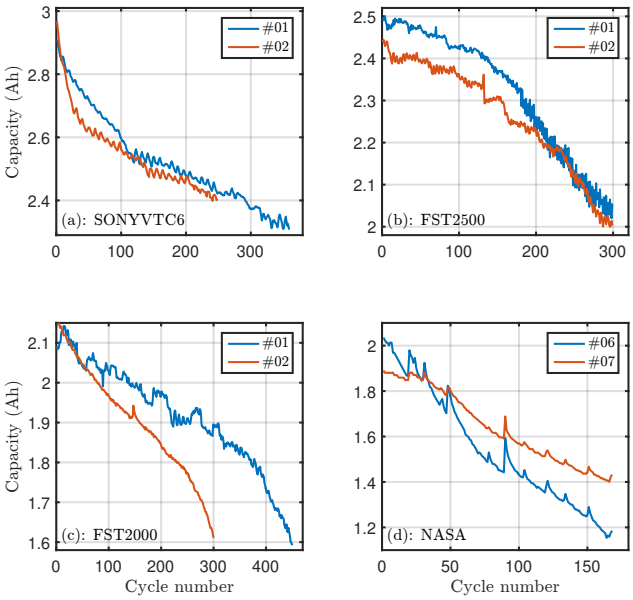

Fig. 1. Datasets for verification. (a): SONYVTC6 batteries; (b): FST2500 batteries; (c): FST2000 batteries; and (d): NASA #06 and #07 batteries.

The first dataset was collected from SONYVTC6 batteries, with a rated capacity of 3 Ah. For the first battery in this dataset, a cyclic aging profile of constant-current-constant-voltage (CCCV) charging [41,42] and constant-current (CC) discharging is carried out. The current rate of this cyclic profile was 1C in the CC phase, and the cutoff conditions were set as 4.2V, 2.75V and 0.05C for the upper voltage limit, lower voltage limit, and current limit, respectively. For the second SONY battery, the aging profile was the same as the first

one except that the upper voltage limit became 4.4V, in other words, the battery was overcharged. The aging experiment was carried out in Guangzhou HKUST Fok Ying Tung Research Institute, and a UPower testing system described in [43] was utilized to implement this experiment. The lifespan of the SONYVTC6 batteries under nominal operating condition is about 360 cycles.

The second dataset was collected from FST2500 batteries, with a rated capacity of 2.5Ah. The batteries were cycled based on a CCCV charging-CC discharging profile with the cut-off conditions of 4.2V upper voltage, 2.75V lower voltage, and 0.05C current limit, respectively. For the first battery, its current rate was 0.4C in the CC phase, while for the second cell, the corresponding current rate became 0.2C. The aging experiments were also carried out using the UPower battery testing system, and a lifespan of 300 cycles could be expected for this battery type.

The third dataset was collected from FST2000 batteries, with a rated capacity of 2 Ah. Two cells were cycled using a CCCV charging-CC discharging profile with a 1C current rate. However, their cut-off conditions were different. For the first one, the cutoff conditions were 4.2V, 2.75V and 0.05C. However, for the second one, the cutoff conditions became 4.2V, 1.9V and 0.05C, which means that the second battery is over-discharged. This experiment was also carried out in our research institute. The cyclic test of the first battery was carried out using a SUNWAY battery testing system as described in [44], while the second battery was cycled using the UPower battery tester. Under nominal conditions, the expected lifespan of FST2000 batteries is about 450 cycles.

In addition, cell #06 and cell #07 in the NASA dataset [45] were selected. The NASA battery type was $LiNi_{0.8}Co_{0.15}Al_{0.05}O_2$, with a rated capacity of 2 Ah. The CCCV charging-CC discharging profile profiles were also used to test their aging performances. Specifically, a charging rate of 0.75C and a discharging rate of 1C were selected. The cut-off voltage for charging is 4.2V, and the cut-off current rate is 0.01C. For cell #06, the cut-off discharging voltage is 2.5V, but only 2.2V for cell #07. It should be noted that this dataset was tested 12 years ago, and the lifespan of the battery is relatively low ($\leq$ 170 cycles). However, NASA dataset is one of the earliest open-access battery aging dataset, and has been widely used for algorithm verification until very recently [46]. Therefore, this dataset could enrich our experimental results.

For all these datasets, the referenced battery capacity is calculated by integrating the discharging current over each cycle, and the aging trajectories of these selected batteries are shown in Fig. 1. It should be pointed out that the experiments are carried out in room temperature without precise control, which brings additional noise to the measured capacity. Due to the joint effect of the uncontrolled temperature, internal chemical materials and battery manufacturing limitations, the aging trajectories of these cells are distinct, containing the cases of convex, concave, and close-to-linear capacity fade trajectories.

Remark 1: The main purpose of this paper is to predict the battery's future degradation, and we do not limit the selection of aging indicator. In our study, followed the same way as many related works [46,47], the battery full discharging capacity has been selected as health indicator for evaluating battery aging state. Some other criteria from partial charging/discharging data such as impedance [48] and incremental-capacity-based features [49] can be also utilized to reflect battery SoH. After we adopt mature extraction techniques from [43,50] to obtain these criteria, the related SoH aging trajectory could be also conveniently predicted through using our proposed model migration NN.

III. Technique

This section describes the proposed model migration neural network. For the purpose of comparison and motivating other algorithms, the conventional empirical model-based prediction technique is introduced first. Then, the standard input-output slope-and-bias correction (SBC) based model migration method is discussed. After that, the SBC-based model migration is extended to a model-migration neural network, followed by a gradient-correction (GC)-based training algorithm. Finally, two commonly used benchmarking algorithms, namely, the empirical prediction and the NN-based prediction, are briefly introduced.

A. Conventional empirical model-based prediction

The decrease in actual capacity is widely utilized to describe the degree of battery degradation. However, the capacity of the commercial batteries vary from case to case. Therefore, normalization is required to compare the capacity of different batteries. Following a widely used engineering definition for the battery SoH [51], it is defined here as

$$SoH(k) = \frac{C_n(k)}{C_n(1)} \tag{1}$$

where $C_n(k)$ represents the actual battery capacity at time k, which is usually sampled at each operating cycle, and $C_n(1)$ stands for the battery capacity that is measured at the first cycle.

With this definition of battery SoH, the relation between battery degradation and cycle number k could be formulated as a single-input-single-output (SISO) function as follows,

$$SoH(k) = y_k = f(k) \tag{2}$$

After defining $f(\cdot)$ based on the partial aging data, it can be extended to predict the aging trajectory of the entire lifespan. The detailed procedure (EPredict) is described in Algorithm 1.

Algorithm 1 Empirical model-based prediction algorithm

1: **procedure** EPredict($SoH_{1:L}, l+L$)
2: Select a mathematical form for $f(\cdot)$;
3: Set the input of $f(\cdot)$ as $1:L$;
4: Set the target output of $f(\cdot)$ as $y_{1:L} = SoH_{1:L}$;
5: Offline determine $\hat{f}(\cdot)$;
6: $\hat{y}_{L+l} = \hat{f}(L+l)$;
7: **end procedure**

In Algorithm 1, several model forms such as single exponential function [21], dual exponential function [22], power function [23], hybrid linear function [24], and polynomial function [25] are generally used for $f(\cdot)$, there also exists lots of algorithms to identify these models' parameters. Among these methods, MATLAB nonlinear curve fitting algorithm has been packed into the standard toolbox and widely used in different areas. Therefore, it is selected for the offline parameter identification in this study.

B. Slope-and-bias-correction-based model migration

Algorithm 1 is effective under the conditions where sufficient aging data has been accumulated [40], but its prediction accuracy is usually low with reduced training data due to the effect of the knee point and the measurement noise.

In order to handle the above issues, an SBC-based model can be utilized to model battery degradation as follows,

$$y_k = F(k) = x_1 \cdot f(x_2 \cdot k + x_3) + x_4 \tag{3}$$

where $\boldsymbol{x} = [x_1, x_2, x_3, x_4]$ are the migration factors under identification, while $f(\cdot)$ is regarded as a *base model*, which should be determined in advance.

A two-stage strategy can be applied for determining the detailed form of $F(\cdot)$. First, the base model $f(\cdot)$ should be determined offline using existing datasets that cover most of the battery lifespan (e.g., dataset collected from some accelerated aging experiments). Then, the online measured partial aging trajectory from time 1 to time L is utilized to identify $\boldsymbol{x}$. Therefore, the overall procedure for aging trajectory prediction using SBC-based model migration can be summarized in Algorithm 2 (SBCPREDICT).

Algorithm 2 SBC-migrated model-based prediction algorithm

1: **procedure** SBCPREDICT($SoH_{1:L}, l+L$)
2: Select a mathematical form for $f(\cdot)$;
3: Select an existing dataset for $f(\cdot)$;
4: Offline determine $f(\cdot)$;
5: Set the input of $F(\cdot)$ as $1:L$;
6: Set the target output of $F(\cdot)$ as $y_{1:L} = SoH_{1:L}$;
7: Identify the $\hat{\boldsymbol{x}}$ in $\hat{F}(\cdot)$;
8: $\hat{y}_{L+l} = \hat{F}(L+l) = \hat{x}_1 \cdot f(\hat{x}_2 \cdot (L+l) + \hat{x}_3) + \hat{x}_4$;
9: **end procedure**

Noting that the base model $f(\cdot)$ has the information covering most of the battery SoH range, and that $\boldsymbol{x}$ is identified through partial aging data accumulated from time 1 to L, the prediction of SBC-based migration model can still use the full SoH range battery degradation information. This is quite different to the case of just using an empirical model, which uses only partial information to generate the predictions.

There are three issues that need to be further clarified. First, obtaining the training data for base model could be quicker if accelerated aging experiments are carried out. For instance, finding the knee point of the red curve shown in Fig. 1(a) requires less than 20% of the total experimental time. Second, the mathematical form of $f(\cdot)$ will not affect the logic of the model migration. However, the quality of $f(\cdot)$ will affect the accuracy of migrated model. For instance, $F(\cdot)$ in (3) will remain linear if the base model $f(\cdot)$ is linear. Third, different algorithms could be applied to identify $\boldsymbol{x}$ (line 7 of Algorithm 2). For instance, gradient correction [51] could provide a computing friendly solution, while Bayes Monte Carlo [22] methods would have a better accuracy.

C. Model migration network

The good fitting performance of the SBC-based model migration relies on a strong assumption that the base model and the target aging process maintains some similarities. This assumption, however, might not be suitable for all testing cases. To handle this issue, the SBC-based model migration technique is integrated into a feed-forward neural network in our study, with the expectation to improve the ability of handling the nonlinearty behaviour in the battery aging process. This integration includes two steps:

First, (3) is formally converted into a neural network with the result as shown in Fig. 2. Specifically, the generated network has one input (the cycle number k), and one output (the SoH). Following (3), biases have been added to both the input and output. The weighted summation of the cycle number and the bias is treated as the input of $f(\cdot)$, and the weighted summation of the output of $f(\cdot)$ and the second bias is treated as the input of $h(\cdot)$. When $h(x) = x$ is selected, the network shown in Fig. 2 is equivalent to (3) mathematically.

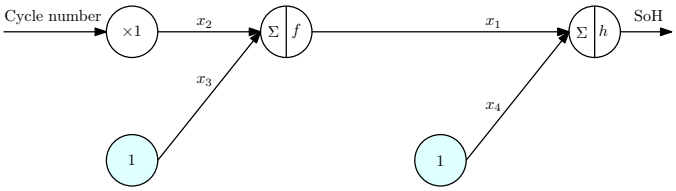

Fig. 2. Illustration of the SBC-based model migration using a neural network structure.

Next, the network shown in Fig. 2 is further enhanced. The number of hidden-neurons could be improved to N. In addition, to avoid simply using linear combinations of the SBC model output, another hidden layer with K neurons is added. The output bias is also shifted to the last hidden layer. Then, a migration NN shown in Fig. 3 can be obtained, where $g(\cdot)$ is the activation function of the additional hidden layer, which is usually set as a nonlinear SISO function.

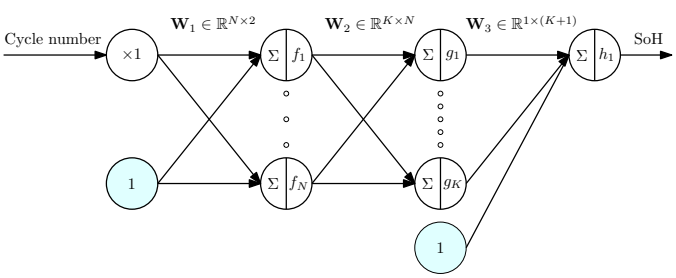

Fig. 3. Framework of the proposed migration NN.

It should be noted that for battery aging trajectory prediction tasks, the training data does not cover the range of the prediction. Therefore, an activation function with saturation property (such as sigmoid) is not suitable here [52]. In response, a simple and effective leakly rectified linear unit ($LReLU$) [53] is selected as:

$$g(x) = LReLU(x) = \max\{0, x\} + \min\{0, x\} \cdot \lambda \quad (4)$$

where λ is commonly set as a small positive value, e.g., 0.05. The $LReLU$ function is light-weight as there is no need to calculate the exponential response compared with the conventional sigmoid activation function.

Mathematically, each $f_i(\cdot)$ for $i \in [1, N]$ and $g_j(\cdot)$ for $j \in [1, K]$ in Fig. 3 can be different. However, from the view of NN design, a general multi-layer feed-forward NN with each neuron sharing the same nonlinear activation function is already sufficient to approximate a convex function with a desired accuracy [54]. Therefore, there is no necessity to improve the complexity of the network by selecting different internal activation functions.

The process of using the proposed migration network for aging prediction is almost the same as Algorithm 2: after determining the base model $f(\cdot)$ offline, $\mathbf{W}_i$ for $i \in \{1, 2, 3\}$ is trained using the data collected from time 1 to L. By changing the input, the SoH at any cycle can then be predicted accordingly. The detailed algorithm (NNPREDICT) is listed in Algorithm 3 as follows,.

Algorithm 3 Migration network-based prediction algorithm

procedure NNPREDICT($SoH_{1:L}, l + L$)
  Select a mathematical form for $f(\cdot)$;
  Select an existing dataset for identifying $f(\cdot)$;
  Offline determine $f(\cdot)$;
  Set the network size as $\{1 - N - K - 1\}$;
  Add bias to the input of the 2nd and 4th layer;
  Set the input of network as $1 : L$;
  Set the target output of the network as $y_{1:L} = SoH_{1:L}$;
  $[\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3]$ = NETTRAIN($SoH_{1:L}$)
  $\hat{y}_{L+l}$ =NETSIM($\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3, L + l$);
end procedure

In Algorithm 3, two functions are called, namely NETSIM and NETTRAIN. NETSIM is utilized to generate the output of the migration network with the provided $\mathbf{W}_i$ and the desired cycle number k. This algorithm is implemented by calculating the network output in a layer-by-layer manner, and its detailed implementation can be found in Algorithm 4.

On the other hand, NETTRAIN is used to train the $\mathbf{W}_i$ with the provided input and the corresponding desired output (the input here is the cycle number 1 to L, which can be omitted, as the desired output is already labelled as $SoH_{1:L}$). There exists many efficient algorithms to train the network. However, to show that the training process of the proposed algorithm could be implemented in a computationally efficient way, an example of using the standard gradient-correction [55] is provided.The detailed process of this algorithm is shown in Algorithm 5.

Algorithm 4 Migration network simulation algorithm

function $[SoH]$ = NETSIM($\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3, k$)
  $y_0 = [k]$;
  $y_{0b} = [k; 1]$; // Add bias
  $v_1 = \mathbf{W}_1 \cdot y_{0b}$;
  $y_1 = f(v_1)$;
  $v_2 = \mathbf{W}_2 \cdot y_1$;
  $y_2 = g(v_2)$;
  $y_{2b} = [y_2; 1]$; // Add bias
  $v_3 = \mathbf{W}_3 \cdot y_{2b}$;
  $y_3 = h(v_3)$;
  $SoH = y_3$;
  return SoH;
end function

Algorithm 5 Migration network training algorithm

1: **function** $[\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3]$ = NETTRAIN($SoH_{1:L}$)
2: Initialize: $N, K, \mathbf{W}_i$ and η_i for $i \in \{1, 2, 3\}$;
3: Initialize: $\text{Ep}_{\max} = 10000$;
4: **for** cnt = 1: $\text{Ep}_{\max}$ **do** // Each cycle
5: **for** $k = 1 : L$ **do** // Each data point
6: $y_0 = [k]$;
7: $y_{0b} = [k; 1]$; // Add bias
8: $v_1 = \mathbf{W}_1 \cdot y_{0b}$;
9: $y_1 = f(v_1)$
10: $v_2 = \mathbf{W}_2 \cdot y_1$;
11: $y_2 = g(v_2)$;
12: $y_{2b} = [y_2; 1]$; // Add bias
13: $v_3 = \mathbf{W}_3 \cdot y_{2b}$;
14: $y_3 = h(v_3)$;
15: $e_3 = SoH_k - y_3$;
16: $\delta_3 = h'(v_3) \cdot e_3$;
17: $e_{2b} = \mathbf{W}_3^T \cdot \delta_3$;
18: $e_2 = e_{2b}(1 : \text{end} - 1)$ // Remove bias
19: $\delta_2 = g'(v_2) \cdot e_2$;
20: $e_1 = \mathbf{W}_2^T \cdot \delta_2$;
21: $\delta_1 = f'(v_1) \cdot e_1$;
22: $\mathbf{W}_3 = \mathbf{W}_3 + \eta_3 \cdot \delta_3 \cdot y_{2b}$;
23: $\mathbf{W}_2 = \mathbf{W}_2 + \eta_2 \cdot \delta_2 \cdot y_1$;
24: $\mathbf{W}_1 = \mathbf{W}_1 + \eta_1 \cdot \delta_1 \cdot y_{0b}$;
25: **end for**
26: $\hat{y}_{1:L}$ = NETSIM($\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3, [1 : L]$);
27: **if** $\text{rms}(y_{1:L} - \hat{y}_{1:L}) \leq \Delta_{Th}$ **then**
28: **break**;
29: **end if**
30: **end for**
31: **return** $[\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3]$;
32: **end function**

It can be seen from Algorithm 5 that the equations from line 6 to line 14 are similar as those in Algorithm 4, which actually represent the feed-forward propagation. The equations from line 15 to line 24 are the standard procedure of gradient-correction-based back propagation. Following line 3, the train-

ing process will be stopped when the overall algorithm has been run for 10,000 times. The 10,000 here is empirically selected, which is sufficiently large for general training cases. However, over-fitting is very likely to happen if the root-mean-square-error (RMSE) of the training is too small. For a light-weight gradient-based algorithm, a simple and efficient way to prevent over-fitting is to stop training process before RMSE reaches its minimum (also known as early stopping technique, see [56] and [57] for details). As described in line 27 of Algorithm 5, the training will also stop when the RMSE of training is smaller or equal to a pre-determined threshold Δ_{Th}. The value of Δ_{Th} should be slightly larger than the fitting RMSE of the base model, but these two values should maintain the same level of magnitude.

D. Benchmarking algorithms

In this subsection, two benchmarking algorithms are designed to verify the superiorities of the proposed method.

Benchmark 1 applies a NN with the structure of Fig. 3 to do the prediction. In this structure, $f_i(\cdot)$ is selected as $LReLU(\cdot)$. In this way, the proposed model migration network becomes a conventional 4-layer feed-forward NN. The training process can be implemented through following Algorithm 5.

Benchmark 2 predicts the aging trajectory by using Algorithm 1, where $f(\cdot)$ is selected to be the dual exponential function as follows [22]

$$f(k) = a \cdot e^{b \cdot k} + c \cdot e^{d \cdot k} \tag{5}$$

where a, b, c, d are the parameters of this dual exponential function.

IV. Experimental results

The experimental verification is carried out in this section. First, the parameter configurations are described in detail to guarantee the repeatability of the proposed method. Then, the prediction accuracy is evaluated under the cases of using 70% and 30% aging trajectories for training, respectively. In addition, discussions of modeling functions are carried out, followed by the numerical analysis of the size of the NN.

A. Parameter configurations

According to the logic of the model migration, a base model is required before predicting the aging trajectories of the target cells. In our experiment, the blue curves shown in Fig. 1 are selected as the target cells, while the batteries with the red curves are used to fit the base model. Without loss of generality, a piece-wise cubic interpolation function is fitted through the MATLAB toolbox, and the threshold for training RMSE $\Delta_{Th} = 0.95\%$ is selected based on the fitting accuracy of the four base models. $N = K = 5$ is selected as the size of the network, and the learning rate $\eta_i = 0.01$ is selected for $i = \{1, 2, 3\}$.

Compared to conventional NN, the weight metrics in the migration NN have clear meanings. According to the training process of equation (3), $\boldsymbol{x} = [1, 1, 0, 0]$ can be selected as the initial value, so that without any training information, $F(\cdot)$ should be the same as the $f(\cdot)$. Following this idea, it is reasonable to use the following method to initialize the weight matrices $\mathbf{W}_i$ as:

$$\mathbf{W}_1 = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{bmatrix}_{N \times 2} + 0.05 \cdot randn(N, 2) \tag{6a}$$

$$\mathbf{W}_2 = \begin{bmatrix} 1/N & \cdots & 1/N \\ \vdots & \ddots & \vdots \\ 1/N & \cdots & 1/N \end{bmatrix}_{K \times N} + 0.05 \cdot randn(K, N) \tag{6b}$$

$$\mathbf{W}_3 = \begin{bmatrix} \frac{1}{K} & \cdots & \frac{1}{K} & 0 \end{bmatrix}_{1 \times (K+1)} + 0.05 \cdot randn(1, K+1) \tag{6c}$$

where $randn(a, b)$ returns an a-by-b random matrix of normally distributed random numbers. When ignoring the random part in (6), the input of $g(\cdot)$ will be greater than 0, in which case $g(x) = x$ holds. In other words, the average of the N base models is utilized to predict the aging trajectory when no additional training data is available. This initialization agrees with the basic idea of the model migration.

To verify that the proposed method is not sensitive to the tuning parameters such as the learning rate, the above-stated parameter configuration is applied to the four different aging datasets, in which the batteries' lifespan vary from 170 cycles to 450 cycles. The algorithm performance is verified in the next subsections.

B. Prediction with 70% data for training

To test the performance of both the proposed and the benchmarking algorithms, we start with the case that most of the aging data is available. To be specific, the first 70% of the aging trajectories are used to train the corresponding model. Then the remaining 30% of the aging trajectory is predicted. For instance, when handling the SONYVTC6 battery, the first 252 cycles are used to train models, and the well-trained models are utilized to forecast the battery degradation of the remaining 108 cycles. The results are shown in Fig. 4 and Table I.

All three algorithms (proposed, benchmark 1 and benchmark 2) are able to achieve a prediction RMSE within 2.5% in at least one testing scenario, while the overall performance of our proposed method is the best one among these three algorithms. This means that both the proposed and the benchmarking algorithms could effectively predict the battery aging trajectory under the conditions that sufficient data is provided. However, the benchmark 1 algorithm can only provide an accurate prediction when the training data covers the knee point of the aging trajectory, as shown in Fig 4-(a), (b) and (d). When it comes to the Fig 4-(c), the conventional NN is not able to accurately predict the changes in battery degradation rate, and the RMSE of this prediction becomes 4.04%. For the second benchmarking algorithm, it does not work well when the measurement noise is large. Even though 70% data is provided for training, the results of scenario (c) and (d) still diverge.

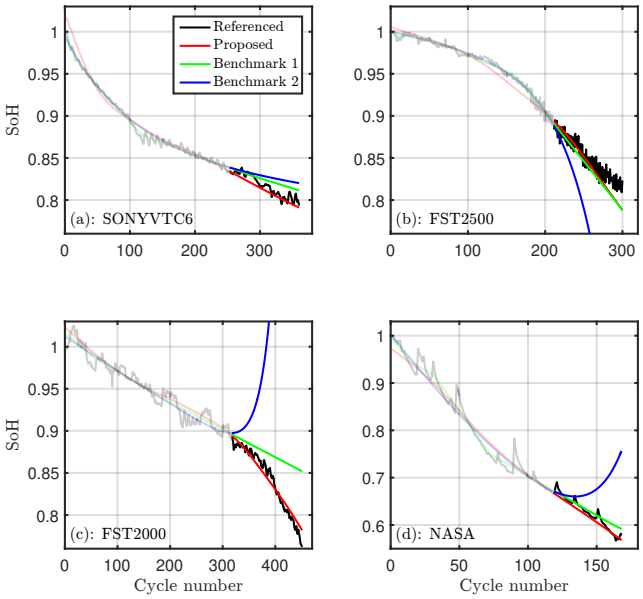

Fig. 4. Predicted aging trajectory when using 70% data for training.

TABLE I
RMSE (%) OF THE PREDICTIONS WHEN USING 70% DATA FOR TRAINING.

	SONYVTC6	FST2500	FST2000	NASA
Proposed	0.65	1.43	0.83	1.06
Benchmark 1	0.98	1.53	4.04	1.18
Benchmark 2	1.43	14.7	71.3	7.90

C. Prediction with 30% data for training

It is more interesting to investigate the algorithms' performance when the training data is limited. In this subsection, only the first 30% aging trajectory is utilized for training process. The corresponding prediction results are shown in Fig. 5 and Table II.

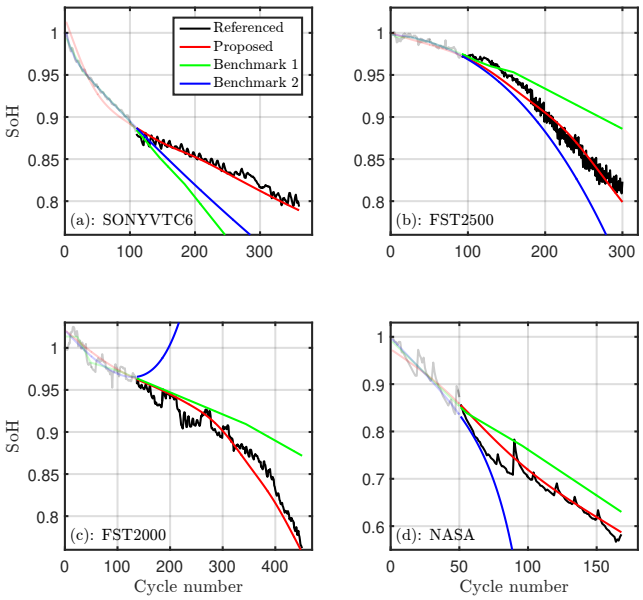

Fig. 5. Predicted aging trajectory when using 30% data for training.

TABLE II
RMSE (%) OF THE PREDICTIONS WHEN USING 30% DATA FOR TRAINING.

	SONYVTC6	FST2500	FST2000	NASA
Proposed	0.67	1.02	1.45	2.30
Benchmark 1	8.43	4.03	4.04	5.28
Benchmark 2	5.28	3.91	>100	>100

In the training phase, it can be seen that the output of all three algorithms can follow the referenced aging trajectories with high accuracy. However, the benchmarking algorithms deviate significantly from the referenced trajectory in the prediction phase, leading to unrealistic prediction error (>100%) at the end of battery life.

In scenario (a), the training data covers only the first 108 cycles. However, there exists a knee point around 125$^{\text{th}}$ cycle. Without this information, the RMSE of the two benchmarking algorithms both exceed 5%, as these algorithms would only propagate the aging trajectory based on the historical aging trend. In comparison, the proposed method can predict the change in degradation speed with the help of the base model. As a result, its prediction RMSE is still better than 1%.

In scenario (b), the battery degradation rate of the first 90 cycles is significantly slower than that of the last 90 cycles. Again, the proposed method can accurately predict this phenomenon and generate a reasonable prediction with 1.02% RMSE. However, it becomes different for the benchmarking algorithms. The first benchmarking algorithm tends to use a linear method to propagate the existing aging trajectory, while the second one extends the aging trajectory with an exponential method. As a result, the predicted SoH of first benchmarking algorithm is higher than the referenced value, while that of the second benchmark is lower. The RMSE of the benchmarking algorithms are 4.03% and 3.91%, respectively.

The dataset of the target cell is highly noisy in scenario (c). In this case, the second benchmark diverges. The first benchmarking algorithm can present a relatively non-diverging result under noise-polluted condition thanks to the piece-wise linear structure in the $LReLU$ function, but similar to the case of using 70% data for training, it cannot predict the change in degradation rate. Even though the degradation rate of the base model and the target cell are quite different (as shown in Fig. 1(c)), the RMSE of the proposed prediction is still better than 1.5%.

In scenario (d), the proposed method still exhibits the highest accuracy, and benchmark 1 turns out to be better than benchmark 2. In fact, the second benchmark provides a totally wrong trend of the degradation curve due to the large noise in the first 30% trajectory. The NN, on the other hand, shows better robustness with the help of the over-fitting prevention technique.

By comparing the Fig. 5-(a), (b) and (d), it can be found that if the aging trajectory of the target cell presents a linear trend, the accuracy of the second benchmark will become better, as the activation function is a piece-wise linear function. However, if the degradation trend contains some exponential curves, the second benchmark would be better. This means that the different selection of modeling functions $f(\cdot)$ in the NN or empirical model can lead to different results. This phenomenon

would be further discussed in the next subsection.

D. Further discussions

To further understand the properties of the proposed model migration NN, the effects of modeling functions and the the size of NN are discussed below.

1) Modeling functions: In the empirical model-based prediction, we have multiple choices for the modeling function $f(\cdot)$. Similarly, the activation functions in a NN are also a variable. In this subsection, the performance of different selections are compared quantitatively.

When using the empirical model-based prediction, the performance of the linear functions (lr), single-exponential function (se), dual-exponential function (de), 2-order polynomial function (2p) and power functions (pr) are tested. Their mathematical expressions are listed as follows:

$$f_{lr}(x) = a \cdot x + b \tag{7a}$$
$$f_{se}(x) = a + b \cdot e^{c \cdot x} \tag{7b}$$
$$f_{de}(x) = a \cdot e^{b \cdot x} + c \cdot e^{d \cdot x} \tag{7c}$$
$$f_{2p}(x) = a + b \cdot x + c \cdot x^2 \tag{7d}$$
$$f_{pr}(x) = a + b \cdot x^c \tag{7e}$$

In addition, when using the NN based prediction, the performance of the different transfer functions are tested. To be specific, we change the transfer functions of the first hidden layer from the SBC-model to the corresponding activation functions, and keep the other parts of the network unmodified. The candidate functions include $LReLU$, $logsig$, $radbas$, $tansig$ and $satlin$, and their definitions are listed below:

$$LReLU(x) = \max\{0, x\} + \lambda \cdot \min\{0, x\} \tag{8a}$$
$$logsig(x) = \frac{1}{1 + e^{-x}} \tag{8b}$$
$$radbas(x) = e^{-x^2} \tag{8c}$$
$$tansig(x) = \frac{2}{1 + e^{-2x}} - 1 \tag{8d}$$
$$satlin(x) = \min\{\max\{0, x\}, 1\} \tag{8e}$$

In this test, the dataset of FST2500 batteries is selected. 50% data is used for model training, and the remaining 50% aging trajectories are utilized for prediction purpose. The results of empirical fittings are shown in Fig. 6-(a). It should be known that initialization[1] of the NN could significantly affect the prediction performance, three groups of NN-based prediction results are shown in Fig. 6-(b), (c) and (d), respectively.

It can be seen that the fitting (training) results of the conventional neural network can well track the global referenced aging trajectory. However, detailed predictions are not accuracy and consistent. The prediction accuracy can improve under some situations, such as the result of $tansig$ in Fig. 6-(b), but these results are almost not repeatable as they highly rely on the random initialization. On the contrary, the empirical fitting is able to provide similar predicting results under multiple times testing, but all candidates could not provide accurate predictions when the training data size is limited. Although the random initialization in (6) leads to slight differences in the results of our proposed method, the overall output is still satisfactory. The reason is mainly due to that the mathematical expectations of $\mathbf{W}_i$ already have clear meanings when being initialized: we tend to use the base model for prediction when no new data is available. Such an initialization could provide a reasonable starting point for the NN training. From the perspective of NN design, it is also interesting to note that through properly selecting the activation functions, a general feed-forward NN turns out to be capable of accurately predicting the lifespan of a battery.

[1]Implemented by MATLAB neural network tool box: `train(net,X,T)`.

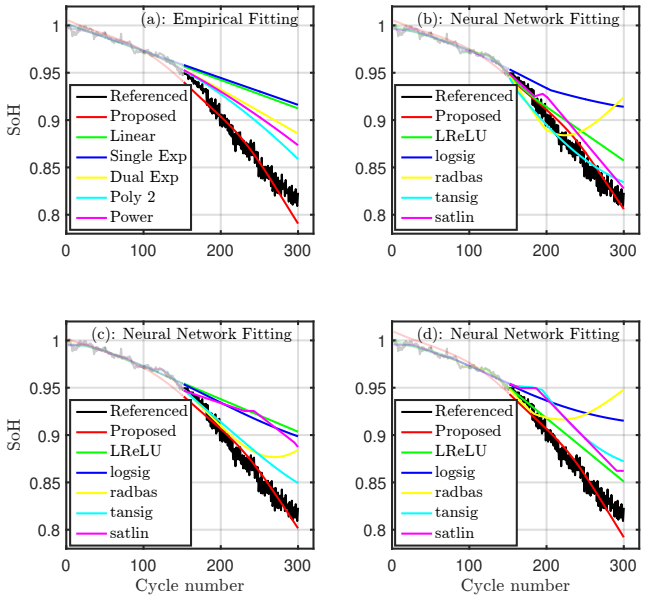

Fig. 6. Predicted aging trajectory when using different modeling functions or activation functions.

2) Network Size: A key motivation of developing model-migration NN is that the conventional SBC-based migration might not be able to compensate all nonlinearties between the base model and the target process, especially when their similarity is low. To verify the improvement of integrating an SBC model migration into a feed-forward NN, the proposed algorithm is tested with different network sizes. Noting that when $N = K = 1$ holds (in Fig. 3), the proposed NN actually becomes the conventional SBC-based model migration. In this test, the base model is established based on the NASA cell #07, while the target cell is the FST2500 cell #01. We set $N = K = Sz$ in this test, and the performance for $Sz \in [1, 5]$ is evaluated. 85% aging trajectory is utilized for model training, and the result is shown in Fig. 7.

Compared with the case of Fig. 1-(b), it can be seen from Fig. 7-(a) that the similarity between the base model and the target aging process becomes low. When $Sz = 1$, the performances of both training and prediction phase are not satisfactory. Obvious under-fitting problems can be observed between cycle 100 and 200. The last predicted 15% aging trajectory is also far away from its true trend. Here, the limitation is mainly caused by the mismatches between the base model and the target degradation process. However,

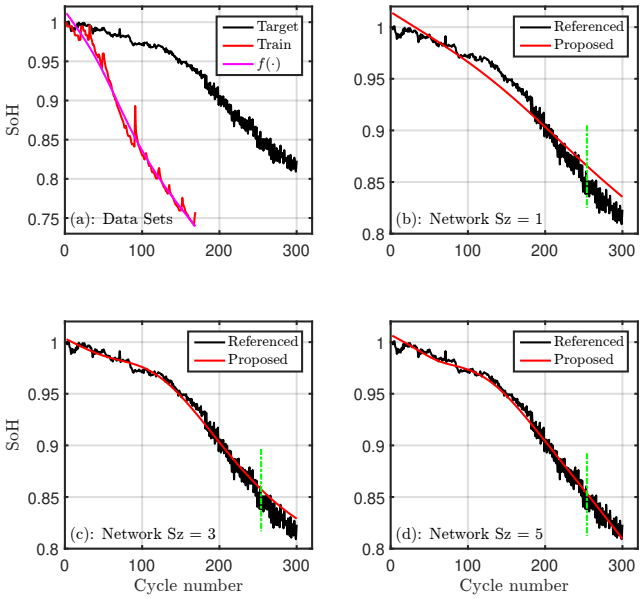

Fig. 7. Predicted aging trajectory with different network size.

when we set $Sz = 3$, the algorithm's performance in the training phase becomes satisfactory. Although the predicted aging trajectory is slightly higher than the true curve, it is much better than the case of $Sz = 1$. As expected, when the network Sz increases to 5, the prediction of the future capacity degradation also becomes more accurate. By comparing the results of Fig. 7-(b) and (d), the model migration NN does present better performance to consider nonlinear behaviors than the conventional SBC-based NN.

3) Minor changes in training data proportion: With a limited training data size, the influence of the noise is, in general, not negligible. Therefore, it is also important to check the network's performance under minor variations of parameters. To implement this task, the first dataset is utilized and the proposed migrated NN is trained using the first 21%∼30% of the aging data, with an interval of 1%. The corresponding results are provided in Fig. 8, and the RMSE of the predictions are listed in Table III.

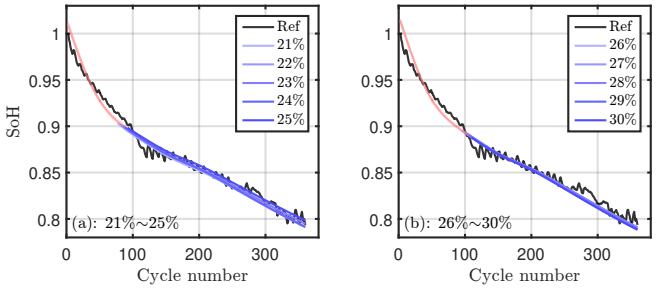

Fig. 8. Aging trajectory of SONYVTC6 battery predicted using different size of training data. (a): with 21%∼25%; (b): with 26%∼30%.

It can be seen that the prediction results would vary with different percentages of training data. However, this change is relatively small (0.5%-0.71%), and the RMSEs of all predictions are smaller than 1%. This result implies that the propose NN is robust over the minor changes of the training data percentage.

4) Comparison with the existing results: To further illustrate the superiority of the proposed algorithm, we compare our results with the existing publications. Noting that different equipment, batteries, and testing conditions may be used in different articles, it is, in fact, difficult to conduct a fair comparison. In response, the results of cell #06 in the widely used open-access NASA aging data set is selected to provide a fair comparing. Refs [58]–[60] are several highly related works from reputable journals within 2 years, and it can be seen from Table IV that the proposed method can achieve the best prediction performance under the case of minimum training data (30%). The complexity of our proposed algorithm is also lowest.

V. Conclusion

This paper develops a feed-forward migration neural network to predict the aging trajectories of Li-ion batteries. Through combining the benefits of both NN technique and model migration concept for battery aging prediction, several interesting conclusions can be observed: 1) As long as the activation functions are properly selected based on the accelerated aging curve, the light-weighted feed-forward model migration NN (1-5-5-1) would be sufficient for general aging trajectory prediction tasks for a battery system. 2) Knee point is not necessary to be covered when predicting the aging trajectory with the proposed model migration NN. 3) When only using 30% of the aging trajectory for training, a prediction RMSE within 2.5% can be achieved under highly noisy conditions. 4) The prediction result is stable even if the NN is randomly initialized. 5) Without any requirement of battery electrochemical information, the proposed model migration NN can be conveniently utilized to other battery types for effective predictions of their aging trajectories.

In the future, the proposed method will be enhanced through two aspects. First, improve the accelerated aging test so that the experimental time could be reduced. Then the obtained base model could better mimic the target degradation process. Second, develop better network structure to preserve the flexibility of model migration and reduce the number of parameters that require to be identified. In addition, more tests, especially those with extreme conditions, could be carried out to further verify the effectiveness of the proposed method.

Acknowledgment

The first author would like to thank the continuing support from the Guangzhou HKUST Fok Ying Tung Research Institute during the Hong Kong's unrest and the outbreak of the Covid-19. The institute provides him with not only a safe environment to continue his PhD study, but also a chance to keep contributing to Hong Kong from the perspective of engineering research even in this difficult time.

TABLE III
RMSE OF THE PREDICTIONS WITH 21%∼30% TRAINING DATA.

Training data	21%	22%	23%	24%	25%	26%	27%	28%	29%	30%
RMSE(%)	0.60	0.59	0.50	0.57	0.56	0.62	0.59	0.59	0.67	0.71

TABLE IV
COMPARISON OF THE RESULTS THAT OBTAINED FROM CELL #06 IN NASA DATA SET.

Method	Training data	RMSE	Complexity
Proposed 1-5-5-1 NN	30%	2.30%	Low
Proposed 1-5-5-1 NN	70%	1.06%	Low
AUKF-GA-SVR [58]	43%	2.55%	High
AUKF-GA-SVR [58]	64%	2.42%	High
PA-LSTM [59]	46%	4.99%	High
PA-LSTM [59]	83%	2.93%	High
LR-GPR [60]	36%	2.92%	High

REFERENCES

[1] L. Lu, X. Han, J. Li, J. Hua, and M. Ouyang, "A review on the key issues for lithium-ion battery management in electric vehicles," *Journal of Power Sources*, vol. 226, pp. 272 – 288, 2013.

[2] Y. Wang, G. Gao, X. Li, and Z. Chen, "A fractional-order model-based state estimation approach for lithium-ion battery and ultra-capacitor hybrid power source system considering load trajectory," *Journal of Power Sources*, vol. 449, p. 227543, 2020.

[3] Z. Wei, J. Zhao, C. Zou, T. M. Lim, and K. J. Tseng, "Comparative study of methods for integrated model identification and state of charge estimation of lithium-ion battery," *Journal of Power Sources*, vol. 402, pp. 189 – 197, 2018.

[4] C. Zou, C. Manzie, and D. Nešić, "Model predictive control for lithium-ion battery optimal charging," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 2, pp. 947–957, 2018.

[5] K. Liu, C. Zou, K. Li, and T. Wik, "Charging pattern optimization for lithium-ion batteries with an electrothermal-aging model," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 12, pp. 5463–5474, 2018.

[6] Q. Ouyang, Z. Wang, K. Liu, G. Xu, and Y. Li, "Optimal charging control for lithium-ion battery packs: A distributed average tracking approach," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2019.

[7] Y. Shang, K. Liu, N. Cui, N. Wang, K. Li, and C. Zhang, "A compact resonant switched-capacitor heater for lithium-ion battery self-heating at low temperatures," *IEEE Transactions on Power Electronics*, pp. 1–1, 2019.

[8] K. Liu, K. Li, and C. Zhang, "Constrained generalized predictive control of battery charging process based on a coupled thermoelectric model," *Journal of Power Sources*, vol. 347, pp. 145–158, 2017.

[9] F. Feng, X. Hu, J. Liu, X. Lin, and B. Liu, "A review of equalization strategies for series battery packs: variables, objectives, and algorithms," *Renewable and Sustainable Energy Reviews*, vol. 116, p. 109464, 2019.

[10] X. Tang, C. Zou, T. Wik, K. Yao, Y. Xia, Y. Wang, D. Yang, and F. Gao, "Run-to-run control for active balancing of lithium iron phosphate battery packs," *IEEE Transactions on Power Electronics*, pp. 1–1, 2019.

[11] X. Hu, F. Feng, K. Liu, L. Zhang, J. Xie, and B. Liu, "State estimation for advanced battery management: Key challenges and future trends," *Renewable and Sustainable Energy Reviews*, vol. 114, p. 109334, 2019.

[12] A. Guha and A. Patra, "State of health estimation of lithium-ion batteries using capacity fade and internal resistance growth models," *IEEE Transactions on Transportation Electrification*, vol. 4, no. 1, pp. 135–146, March 2018.

[13] Z. Yang, D. Patil, and B. Fahimi, "Online estimation of capacity fade and power fade of lithium-ion batteries based on input–output response technique," *IEEE Transactions on Transportation Electrification*, vol. 4, no. 1, pp. 147–156, 2017.

[14] K. Liu, K. Li, Q. Peng, and C. Zhang, "A brief review on key technologies in the battery management system of electric vehicles," *Frontiers of mechanical engineering*, vol. 14, no. 1, pp. 47–64, 2019.

[15] M. Jafari, A. Gauchia, S. Zhao, K. Zhang, and L. Gauchia, "Electric vehicle battery cycle aging evaluation in real-world daily driving and vehicle-to-grid services," *IEEE Transactions on Transportation Electrification*, vol. 4, no. 1, pp. 122–134, 2017.

[16] Y. Liu, J. Li, Z. Chen, D. Qin, and Y. Zhang, "Research on a multi-objective hierarchical prediction energy management strategy for range extended fuel cell vehicles," *Journal of Power Sources*, vol. 429, pp. 55 – 66, 2019.

[17] Y. Zhang, R. Xiong, H. He, X. Qu, and M. Pecht, "State of charge-dependent aging mechanisms in graphite/li(nicoal)o2 cells: Capacity loss modeling and remaining useful life prediction," *Applied Energy*, vol. 255, p. 113818, 2019.

[18] C. Lyu, Y. Song, J. Zheng, W. Luo, G. Hinds, J. Li, and L. Wang, "In situ monitoring of lithium-ion battery degradation using an electrochemical model," *Applied Energy*, vol. 250, pp. 685 – 696, 2019.

[19] R. Gu, P. Malysz, H. Yang, and A. Emadi, "On the suitability of electrochemical-based modeling for lithium-ion batteries," *IEEE Transactions on Transportation Electrification*, vol. 2, no. 4, pp. 417–431, 2016.

[20] Y. Li, K. Liu, A. M. Foley, A. Zlke, M. Berecibar, E. Nanini-Maury, J. V. Mierlo, and H. E. Hoster, "Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review," *Renewable and Sustainable Energy Reviews*, vol. 113, p. 109254, 2019.

[21] 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. 17 729–17 740, 2018.

[22] W. He, N. Williard, M. Osterman, and M. Pecht, "Prognostics of lithium-ion batteries based on dempstershafer theory and the bayesian monte carlo method," *Journal of Power Sources*, vol. 196, no. 23, pp. 10 314 – 10 321, 2011.

[23] X. Tang, K. Liu, X. Wang, B. Liu, F. Gao, and W. D. Widanage, "Real-time aging trajectory prediction using a base model-oriented gradient-correction particle filter for lithium-ion batteries," *Journal of Power Sources*, vol. 440, p. 227118, 2019.

[24] C. Hu, H. Ye, G. Jain, and C. Schmidt, "Remaining useful life assessment of lithium-ion batteries in implantable medical devices," *Journal of Power Sources*, vol. 375, pp. 118 – 130, 2018.

[25] Y. Sun, X. Hao, M. Pecht, and Y. Zhou, "Remaining useful life prediction for lithium-ion batteries based on an integrated health indicator," *Microelectronics Reliability*, vol. 88-90, pp. 1189 – 1194, 2018, 29th European Symposium on Reliability of Electron Devices, Failure Physics and Analysis (ESREF 2018).

[26] Z. Xue, Y. Zhang, C. Cheng, and G. Ma, "Remaining useful life prediction of lithium-ion batteries with adaptive unscented kalman filter and optimized support vector regression," *Neurocomputing*, 2019.

[27] K. Liu, X. Hu, Z. Yang, Y. Xie, and S. Feng, "Lithium-ion battery charging management considering economic costs of electrical energy loss and battery degradation," *Energy Conversion and Management*, vol. 195, pp. 167–179, 2019.

[28] P. G. Nieto, E. Garca-Gonzalo, F. S. Lasheras, and F. de Cos Juez, "Hybrid psosvm-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability," *Reliability Engineering & System Safety*, vol. 138, pp. 219 – 231, 2015.

[29] K. Liu, Y. Li, X. Hu, M. Lucu, and D. Widanalage, "Gaussian process regression with automatic relevance determination kernel for calendar aging prediction of lithium-ion batteries," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2019.

[30] Y. Zhang, R. Xiong, H. He, and M. G. Pecht, "Lithium-ion battery remaining useful life prediction with boxcox transformation and monte carlo simulation," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 2, pp. 1585–1597, Feb 2019.

[31] K. Liu, X. Hu, Z. Wei, Y. Li, and Y. Jiang, "Modified gaussian process regression models for cyclic capacity prediction of lithium-ion batteries," *IEEE Transactions on Transportation Electrification*, pp. 1–1, 2019.

[32] R. R. Richardson, M. A. Osborne, and D. A. Howey, "Gaussian process regression for forecasting battery state of health," *Journal of Power Sources*, vol. 357, pp. 209–219, 2017.

[33] Y. Zhang, R. Xiong, H. He, and M. G. Pecht, "Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 5695–5705, July 2018.

[34] B. Yang, R. Liu, and E. Zio, "Remaining useful life prediction based on a double-convolutional neural network architecture," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 12, pp. 9521–9530, Dec 2019.
[35] K. Liu, Y. Shang, Q. Ouyang, and D. Widanalage, "A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery," *IEEE Transactions on Industrial Electronics*, pp. 1–1, 2020.
[36] Y. Zhao, P. Liu, Z. Wang, L. Zhang, and J. Hong, "Fault and defect diagnosis of battery for electric vehicles based on big data analysis methods," *Applied Energy*, vol. 207, pp. 354 – 362, 2017.
[37] C. Zhang, Y. Wang, Y. Gao, F. Wang, B. Mu, and W. Zhang, "Accelerated fading recognition for lithium-ion batteries with nickel-cobalt-manganese cathode using quantile regression method," *Applied Energy*, vol. 256, p. 113841, 2019.
[38] K. Liu, K. Li, Q. Peng, Y. Guo, and L. Zhang, "Data-driven hybrid internal temperature estimation approach for battery thermal management," *Complexity*, vol. 2018, 2018.
[39] J. Lu, K. Yao, and F. Gao, "Process similarity and developing new process models through migration," *AIChE journal*, vol. 55, no. 9, pp. 2318–2328, 2009.
[40] X. Tang, C. Zou, K. Yao, J. Lu, Y. Xia, and F. Gao, "Aging trajectory prediction for lithium-ion batteries via model migration and bayesian monte carlo method," *Applied Energy*, vol. 254, p. 113591, 2019.
[41] K. Liu, K. Li, Z. Yang, C. Zhang, and J. Deng, "An advanced lithium-ion battery optimal charging strategy based on a coupled thermoelectric model," *Electrochimica Acta*, vol. 225, pp. 330–344, 2017.
[42] K. Liu, K. Li, H. Ma, J. Zhang, and Q. Peng, "Multi-objective optimization of charging patterns for lithium-ion battery management," *Energy Conversion and Management*, vol. 159, pp. 151–162, 2018.
[43] X. Tang, C. Zou, K. Yao, G. Chen, B. Liu, Z. He, and F. Gao, "A fast estimation algorithm for lithium-ion battery state of health," *Journal of Power Sources*, vol. 396, pp. 453 – 458, 2018.
[44] X. Tang, K. Yao, B. Liu, W. Hu, and F. Gao, "Long-term battery voltage, power, and surface temperature prediction using a model-based extreme learning machine," *Energies*, vol. 11, no. 1, 2018.
[45] B. Saha and K. Goebel, "Battery data set," *NASA AMES prognostics data repository*, vol. PP, no. 99, pp. 1–1, 2007.
[46] P. Tagade, K. S. Hariharan, S. Ramachandran, A. Khandelwal, A. Naha, S. M. Kolake, and S. H. Han, "Deep gaussian process regression for lithium-ion battery health prognosis and degradation mode diagnosis," *Journal of Power Sources*, vol. 445, p. 227281, 2020.
[47] X. Tang, Y. Wang, K. Yao, Z. He, and F. Gao, "Model migration based battery power capability evaluation considering uncertainties of temperature and aging," *Journal of Power Sources*, vol. 440, p. 227141, 2019.
[48] X. Zhou, J. Huang, Z. Pan, and M. Ouyang, "Impedance characterization of lithium-ion batteries aging under high-temperature cycling: Importance of electrolyte-phase diffusion," *Journal of Power Sources*, vol. 426, pp. 216–222, 2019.
[49] X. Li, Z. Wang, and J. Yan, "Prognostic health condition for lithium battery using the partial incremental capacity and gaussian process regression," *Journal of Power Sources*, vol. 421, pp. 56–67, 2019.
[50] T. Murariu and C. Morari, "Time-dependent analysis of the state-of-health for lead-acid batteries: An eis study," *Journal of Energy Storage*, vol. 21, pp. 87–93, 2019.
[51] X. Tang, Y. Wang, C. Zou, K. Yao, Y. Xia, and F. Gao, "A novel framework for lithium-ion battery modeling considering uncertainties of temperature and aging," *Energy Conversion and Management*, vol. 180, pp. 162 – 170, 2019.
[52] D. Stursa and P. Dolezel, "Comparison of relu and linear saturated activation functions in neural network for universal approximation," in *2019 22nd International Conference on Process Control (PC19)*, June 2019, pp. 146–151.
[53] Y. Liu, X. Wang, L. Wang, and D. Liu, "A modified leaky relu scheme (mlrs) for topology optimization with multiple materials," *Applied Mathematics and Computation*, vol. 352, pp. 188 – 204, 2019.
[54] E. J. Hartman, J. D. Keeler, and J. M. Kowalski, "Layered neural networks with gaussian hidden units as universal approximations," *Neural Computation*, vol. 2, no. 2, pp. 210–215, June 1990.
[55] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.
[56] R. Gencay and Min Qi, "Pricing and hedging derivative securities with neural networks: Bayesian regularization, early stopping, and bagging," *IEEE Transactions on Neural Networks*, vol. 12, no. 4, pp. 726–734, July 2001.
[57] Y. Shao, G. N. Taff, and S. J. Walsh, "Comparison of early stopping criteria for neural-network-based subpixel classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 1, pp. 113–117, Jan 2011.
[58] Z. Xue, Y. Zhang, C. Cheng, and G. Ma, "Remaining useful life prediction of lithium-ion batteries with adaptive unscented kalman filter and optimized support vector regression," *Neurocomputing*, vol. 376, pp. 95 – 102, 2020.
[59] J. Qu, F. Liu, Y. Ma, and J. Fan, "A neural-network-based method for rul prediction and soh monitoring of lithium-ion battery," *IEEE Access*, vol. 7, pp. 87 178–87 191, 2019.
[60] J. Yu, "State of health prediction of lithium-ion batteries: Multiscale logic regression and gaussian process regression ensemble," *Reliability Engineering & System Safety*, vol. 174, pp. 82–95, 2018.

Xiaopeng Tang received the B.S. in automation from the University of Science and Technology of China, He Fei, China, in 2015.

He is currently working towards the Ph.D. degree in chemical and biological engineering with the Hong Kong University of Science and Technology, Hong Kong, China. His research interests include modeling and control of energy storage systems.

Mr. Tang has been a recipient of the National scholarship in 2015, and Hong Kong PhD Fellowship Scheme in 2017. He is also serving as the reviewer for many leading journals in the research field, such as IEEE TRANSACTIONS ON TRANSPORTATION ELECTRIFICATION, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE TRANSACTIONS ON POWER ELECTRONICS, APPLIED ENERGY, and ENERGY CONVERSION AND MANAGEMENT.

Kailong Liu (M18) is a Senior Research Fellow in the Warwick Manufacturing Group, University of Warwick, United Kingdom. He received the B.Eng. degree in electrical engineering and the M.Sc. degree in control theory and control engineering from Shanghai University, China, and the Ph.D. degree in electrical engineering from Queens University Belfast, United Kingdom, in 2011, 2014, and 2018, respectively.

He was a Visiting Student Researcher at the Tsinghua University, China, in 2016. His research interests include modeling, optimization and control with applications to electrical/hybrid vehicles, energy storage, battery manufacture and management. Dr. Liu was the student chair of IEEE QUB student branch and a recipient of awards such as EPSRC Scholarship, Santander International Scholarship, and QUB ESM International Scholarship.

Xin Wang received the B.S. degree in chemical engineering and technology from Beijing Institute of Technology, Beijing, China, in 2003, and the M.Phil. degree in Nano Science and Technology from The Hong Kong University of Science and Technology, Hong Kong, China, in 2011.

She is currently working towards the Ph.D. degree in chemical and biological Engineering with the Hong Kong University of Science and Technology. Her research interests include data-mining and energy system modeling.

Furong Gao received the B.Eng. degree in automation from the East China Institute of Petroleum, Beijing, China, in 1985, and the M.Eng. and Ph.D. degrees in chemical engineering from McGill University, Montreal, QC, Canada, in 1989 and 1993, respectively.

He was a Senior Research Engineer with Moldflow International, Melbourne, Australia, from 1993 to 1995. He is currently a Chair Professor with the Department of Chemical and Biological Engineering, Hong Kong University of Science and Technology, Hong Kong, China. His research interests include process monitoring and fault diagnosis, batch process control, polymer processing control, and optimization. He received a number of best paper awards, and is on editorial boards of a number of journals of his area.

James Marco received the Eng.D. degree from the University of Warwick, Coventry, U.K., in 2000.

In his early career, James worked for several years within the automotive industry leading engineering research teams for Ford (North America and Europe), Jaguar Cars, Land Rover and Daimlerchrysler (Germany).

James Marco is currently a professor of systems modeling and control with the University of Warwick. He is also a Chartered Engineer and a Fellow of the Institution of Engineering and Technology (IET). His research interests include systems engineering, real-time control, energy storage modeling, design optimization, and the design of energy management control systems.

W. Dhammika Widanage is an Assistant Professor in modelling and energy storage at WMG, Warwick University and was the recipient of thee WMG Early Career Researcher of the Year award in 2016. His research is in system identification theory, applied across several applications including batteries.

He leads the battery modelling research in the department and recently secured funding to lead the modelling activity as a PI (for WMG) on the Faraday Multiscale Modelling project, as a Co-I on the EPSRC Prosperity Partnership with Jaguar Land Rover and a PI and Co-I on four Innovate UK projects (PI and three Co-I). He is a member of the European Materials Modelling Community Interoperability and Repository Advisory Group (EMMC-IRAG) and supported the compilation of the Materials modelling-Terminology, Classification and Metadata document which is now publicly available to increase the use of material modelling with industrial end-users.