(article starts on next page)

Combining ICA Feature-Based Offline and Online Machine Learning to Estimate State of Health of Lithium-Ion Batteries

Chengqi She, Yang Li, Torsten Wik, and Changfu Zou

Abstract— This article proposes an adaptive state of health (SOH) estimation method for lithium-ion batteries using machine learning. Practical problems with cell inconsistency and online implementability are specifically solved using a proposed individualized estimation scheme blending offline model migration with online ensemble learning. First, based on the data of pseudo-open-circuit voltage measured over the battery lifespan, a systematic comparison of different incremental capacity features is conducted to identify a suitable SOH indicator. Next, a pool of candidate models, composed of slope-bias correction (SBC) and radial basis function neural networks (RBFNNs), are trained offline. For online operation, the prediction errors due to cell inconsistency in the target new cell are then mitigated by a proposed modified random forest regression (mRFR) based ensemble learning process with high adaptability. The results show that compared to prevailing methods, the proposed SBC-RBFNN-mRFR-based scheme can achieve considerably improved SOH estimation accuracy (15%) with only a small amount of early-age data and online measurements are needed for practical operation.

I. INTRODUCTION

Lithium-ion (Li-ion) batteries have been reckoned as the backbone of electric vehicles (EVs) and key components of modern grid systems due to their salient merits of high energy and power densities, low self-discharge rate, and ever-declining costs in recent years [1]. However, the energy storage capacity and power capability of Li-ion batteries can gradually reduce caused by various aging mechanisms, leading to limited service life and degraded system performance over time [2]. To ensure the safe, reliable, and efficient use of Li-ion batteries, the indicator of battery health, namely the state of health (SOH), must be precisely monitored and predicted, which forms a fundamental functionality of Li-ion battery management systems (BMSs) [3].

The SOH is a number defined by comparing an aged battery parameter to its pristine value at the beginning of life, commonly defined using battery capacity or internal resistance. Many battery SOH estimation algorithms have been developed in the literature. In the last decade, the data-driven methods have received rapidly growing research attention, as their model-free and easy-to-implement natures are in favor of real-world applications [4]. Among many data-driven techniques, incremental capacity analysis (ICA) is one of the most extensively investigated methods. In ICA,

the incremental capacity (IC) curves are generated based on the capacity–voltage relationship obtained from constant-current charging/discharging. The movement of the IC curve is closely related to the phase transitions and phase equilibria during the lithiation and de-lithiation processes inside the Li-ion cells [5]. Thus, features of interest (FOIs) can be reasonably extracted from the IC curves to reveal the hidden relationships between the direct measurements to SOH [6].

FOIs are usually constructed by investigating the positions and amplitudes of the points of interest (POIs) on the IC curve. For instance, Li *et al.* compared the estimation accuracy of the IC features generated from different POIs [7]. The main POIs that have been identified as qualified for battery SOH estimation are the peak and valley points, especially in high-voltage regions [8], and it usually requires the battery to cycle under a full charging/discharging protocol in the lifelong operation. Unfortunately, such a cycling process is rarely experienced in real-world applications, and in many cases, these POIs are inappropriate to be used for full life-cycle prediction. For example, the peak points on the IC curves may decrease and vanish towards the battery's end of life. As a result, the effectiveness of the SOH estimation algorithms based on these features would be significantly reduced for aged batteries. To address this problem, Li *et al.* focused on the datasets with a selected state of charge (SOC) range during actual charging processes and a model was proposed to adopt multipoint features as the model inputs [9]. Indeed, by selecting several points distributed in a region with a drastically changed IC curve, the risk of losing a single point feature as battery ages can be lowered. Nevertheless, both features extracted from a single point and multiple points are prone to measurement noises. In contrast, using area features of the IC curves can effectively reduce the sensitivity to these noises and mitigate the influence of applied filtering algorithms, thereby achieving high accuracy for battery SOH estimation [10].

Most of the algorithms based on IC features are investigated and validated at the battery cell level. A common problem found in the above-mentioned SOH estimation algorithms is that the effects of cell inconsistency are overlooked. In fact, cell inconsistency is inevitable due to manufacturing tolerance: even the same type of batteries from the same manufacturing batch will exhibit different characteristics. The inconsistency can be magnified with the cell being used due to unbalanced aging trajectories caused by thermal, electrical, and mechanical nonuniformity in the battery packs. Hence, the suitability for extrapolating the predictive model derived by fitting the data from one cell or one pack to other

Chengqi She, Yang Li, Torsten Wik, and Changfu Zou are with the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg 41296, Sweden. (yangli@ieee.org; tw@chalmers.se; changfu.zou@chalmers.se)

Chengqi She is also with the National Engineering Laboratory for Electric Vehicles, Beijing Institute of Technology, Beijing, 100081, China. (shechengqi@bit.edu.cn)

individuals is not guaranteed.

In this paper, suitable IC features are first selected and extracted to provide the information for training a pool of predictive models fused with slope-bias correction (SBC) and radial basis function neural networks (RBFNNs). The peak value in the high-voltage region of IC curves is used as the input of the proposed SBC-RBFNN models, where highly nonlinear relationships of the corrected parameters in the SBC method are identified by RBFNNs. To handle the cell inconsistency between the candidate offline models and the new target cell, a modified random forest regression (mRFR) based online ensemble learning is developed for the first time to realize individualized estimation by incorporating the adaptively generated weighting into the offline SBC-RBFNN models. The applicability and the effectiveness of our proposed method are validated through comparative studies with several state-of-the-art SOH estimation techniques using the data collected from laboratory tests.

II. Dataset Description and Data Preprocessing

The Oxford Battery Degradation Dataset is used in this work. The dataset contains measurements of battery aging data from eight commercial Kokam pouch cells of 740-mAh nominal capacity, with graphite-based negative electrode and lithium cobalt oxide/lithium nickel cobalt oxide positive electrode [11]. The cells were tested in a thermal chamber at 40 °C and repeatedly exposed to a 1C or C/25 constant-current-constant-voltage charging profile, followed by a driving cycle discharging profile obtained from the Urban Artemis profile. Characterization measurements with a sampling frequency of 1 Hz were taken every 100 driving cycles. More detailed descriptions of this dataset can be found in [11] and [12].

The voltage data measured under very low current rates in the laboratory dataset usually contain the most pertinent knowledge of the OCV, and we denote them as the pseudo-open-circuit voltage (pseudo-OCV) data, e.g., the C/25 constant current in the characterization processes in the Oxford Battery Degradation Dataset. Although close to the true OCV data, the pseudo-OCV data are disturbed by measurement errors and affected by the inconsistent variation of the internal resistance. Directly using the pseudo-OCV data for feature extraction can thus lead to inaccurate SOH estimation results, and it is beneficial to approximate the true OCV curves before conducting feature extraction.

Usually, when the current rate applied to the battery is very low, the overpotentials due to polarization and hysteresis are negligible, and the impact of temperature variation on the equilibrium voltage due to self-heating can also be ignored [13]. In this condition, the pseudo-OCV, or the terminal voltage during low-rate charging and discharging, can be simply expressed as

$$V_{ch} = OCV_{ch} + I_{ch}R \tag{1}$$

$$V_{dc} = OCV_{dc} - I_{dc}R \tag{2}$$

where the symbols V, I, and R represent the terminal voltage (pseudo-OCV), current magnitude, and internal resistance,

respectively, and the subscripts ch and dc denote the charging and discharging processes, respectively. Since in this dataset the batteries were fully charged and fully discharged in each characterization, the SOC and the SOH can be defined and calculated by

$$SOC_t = Q_t/Q_n \tag{3}$$

$$SOH_n = Q_n/Q_0 \tag{4}$$

where Q_t represents the Coulomb-counting capacity at time t of the nth characterization, Q_n denotes the charging capacity of the nth characterization, and Q_0 is the cell capacity at the beginning of life.

Considering the consecutive charging/discharging process with the same current rate and ambient temperature, at the same SOC level $SOC_{ch} = SOC_{dc}$, we have $OCV_{ch} = OCV_{dc}$. Applying this condition to (1) and (2) yields the expression of the internal resistance

$$R = (V_{ch} - V_{dc})/(I_{ch} + I_{dc}) = (V_{ch} - V_{dc})/(2I) \tag{5}$$

where $I_{ch} = I_{dc} = I$. According to (1), the true OCV can be approximated by the following corrected OCV, i.e.,

$$OCV_{corrected} = V_{ch} - I_{ch}R. \tag{6}$$

III. Development of SOH Estimation Algorithm

A. Development of Offline Models Based on SBC-RBFNN

The individual candidate models to be developed should be able to cover the entire space of the possible cell variation and simple to train. Since model migration can properly balance the individuality and similarity in generating these models, it is well-suited for our present investigation. Several model migration methods have been proposed in the literature to reliably deal with the predictive performance of similar processes and to save experimental resources, amongst which the SBC is considered to be one of the most effective approaches yet simple to implement [14]. In the SBC, the inconsistency between the base model $f(\cdot)$ and the new model is completely parameterized by the input slope $W^{[1]}$ and bias $b^{[1]}$ as well as the output slope $W^{[2]}$ and bias $b^{[2]}$. The general equation $g(x)$ for the SBC model is

$$g(x) = W^{[2]}f(W^{[1]}x + b^{[1]}) + b^{[2]} \tag{7}$$

In our proposed three-layer SBC-RBFNN framework, Layers 1 and 2 are described by the SBC model (7), and an RBFNN model is connected to the output of the SBC structure for its excellent capability to capture the underlying nonlinear relationship. In an RBFNN, the activation function can map the input feature to a high dimensional space and transform the nonlinear relationship into a linear one.

The numbers of nodes of the three layers in the SBC-RBFNN are denoted by l, m, and n, respectively. The IC values of peak point C are the inputs of the base model (Layer 1), corrected by the SBC function (Layer 2). The outputs of Layer 2 can be written as a vector $g(x) = z = [z_1, z_2, \cdots, z_m]^\top \in \mathbb{R}^m$, which acts as the input vector of the RBF layer (Layer 3). As the model output, the SOH values

are predicted by the RBFNN model, in which the Gaussian function is used as the transfer function, i.e.,

$$\kappa(z, z_{cj}) = \exp\left((- \parallel z - z_{cj} \parallel^2)/2\sigma_j^2\right) \quad (8)$$

where $\kappa(\cdot, \cdot)$ represents the kernel function, $j \in \{1, 2, \cdots, n\}$ is the index of the node in the RBF layer, and z_{cj} and σ_j are the RBF center vector and the standard deviation of the Gaussian function of the jth node, respectively. The RBFNN function can thus be written as

$$y = \sum_{j=1}^{n} w_j^{[3]} \cdot \kappa(z, z_{cj}) + b^{[3]} \quad (9)$$

where we denote $W^{[3]} = [w_1^{[3]}, w_2^{[3]}, \cdots, w_n^{[3]}]^\top$ and $b^{[3]}$ as the weights and bias of the RBF layer, respectively.

The offline SBC-RBFNN model structure is obtained by combining (7)–(9), denoted by $y = h(x)$. To obtain the parameters for one individual model, including the weights $W^{[1]}$, $W^{[2]}$, and $W^{[3]}$, the biases $b^{[1]}$, $b^{[2]}$, and $b^{[3]}$, as well as the center vector $[z_{c1}, z_{c2}, \cdots, z_{cn}]$, the datasets are randomly split into the base model group, training group, and testing group, and they are used for base model generation, SBC-RBFNN model training, and model verification, respectively. The root-mean-square error (RMSE) between the model and the expected output, calculated by

$$RMSE = \sqrt{\frac{1}{M} \sum_{k=1}^{M} (h(x_k) - y_k)^2} \quad (10)$$

where k and M denote the index and the total number of data samples, respectively. Once the predictive model is established, the remaining laboratory datasets of battery cells are used to verify and test the model accuracy.

B. Online Model Adaptation Using mRFR

In the previous subsection, the N SBC-RBFNN models offline trained from historical data are expected to provide sufficient information to cover the most representative aging behaviors of all individual cells, and they form a pool of candidate models for individualized estimation. In order to estimate the SOH for a new target cell, of which only limited early-age data and online measurements are available, it is essential to develop an algorithm to exploit these N offline models for online use. A general and simple strategy is to find a proper blending scheme by weighted averaging, i.e.,

$$\bar{h}_k = \sum_{i=1}^{N} w_i h_i(x_k) \quad (11)$$

where x_k and $\bar{h}_k$ represents the blended model input and model output at time instant k, respectively, $h_i(\cdot)$ is the ith offline SBC-RBFNN model established based on the method described in Section III-A, and our objective is to find proper weights $\{w_1, w_2, \cdots, w_N\}$ for online operation.

A modified random forest regression (mRFR) based online ensemble learning algorithm is proposed to adaptively generate and adjust the weights w_i of each offline model, with the structure illustrated in Fig. 1. A similar "bootstrap" technique
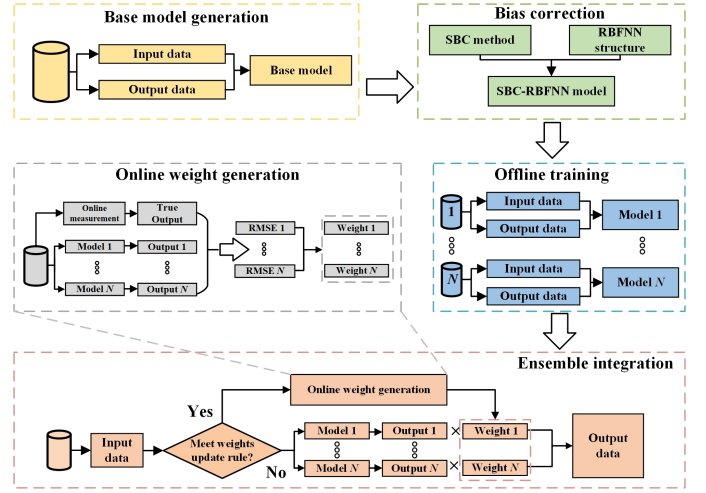

Fig. 1. Flowchart of the online ensemble learning process.

in the conventional RFR [15] is adopted to randomly sample from the dataset with replacement, resulting in N different and uncorrelated decision trees, namely the offline models. The way to plant trees is different from the conventional RFR and has been described in Section III-A. Once the N offline SBC-RBFNN models with corresponding model parameters have been determined, based on the real-time measurements that can reflect the user's real-world behaviors, the weights w_i are next generated and updated online according to the following steps:

First, for each cell or vehicle i, its RMSE over a selected time horizon $\mathcal{P}$ in one online adaptive event is calculated:

$$RMSE_i = \sqrt{\frac{1}{\mathcal{P}} \sum_{p=1}^{\mathcal{P}} (h_i(x_p) - h'_p)^2} \quad (12)$$

where $h_i(x_p)$ and h'_p are the estimated output from the ith model and the true output data (i.e., true SOH), respectively, both obtained at time step p. To calibrate the true SOH in terms of the battery capacity, a full discharge and charge process is needed. This can be readily triggered in experimental studies, but rarely occurs for real-world EV battery systems, which is also the reason to estimate SOH. A practical scenario is to generate SOH measurements when EVs do regular maintenance in the service center, which can be every six months or 10,000 km, for instance. In that case, the online adaption is then activated in a much slower timescale than the offline models.

Next, we consider the importance of each model inversely proportional to its calculated RMSE, and the weighting factor for the ith model is given by
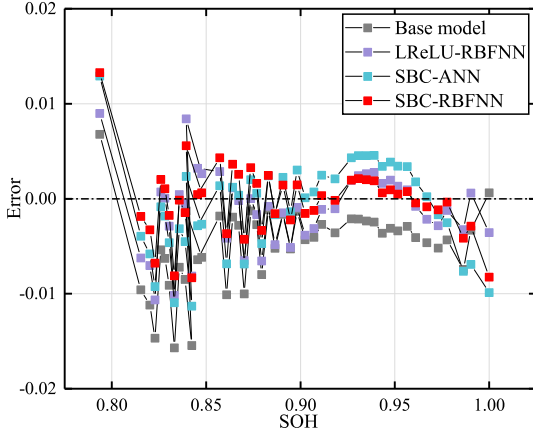
$$r_i = \left(\sum_{i=1}^{N} RMSE_i\right) / RMSE_i \quad (13)$$

Finally, normalizing r_i yields the weight for the ith offline model, i.e.,

$$w_i = r_i / \sum_{i=1}^{N} r_i \quad (14)$$
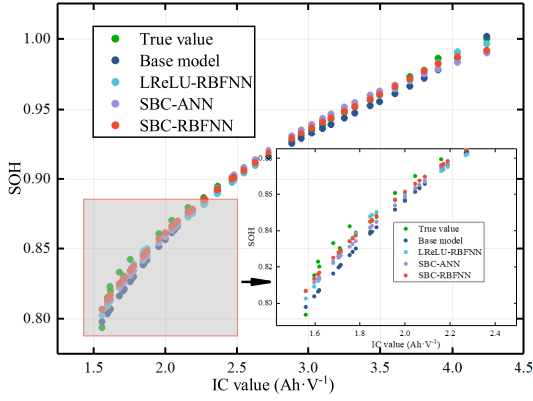
(a)

(b)

Fig. 2. (a) Predictive result of No. 6 cell in Case 1. (b) Predictive error of No. 6 cell in Case 1.

Fig. 3. Comparison of average REPs between different offline models based on the laboratory dataset.

with which the estimated SOH can be obtained based on (11). When the sampling time interval of p is larger than that of k, the weights w_i will be held constant until the next SOH measurement comes at $p + 1$.

IV. RESULTS AND DISCUSSION

A. Offline Model Verification

For the proposed offline SBC-RBFNN model, first, a cubic polynomial is found to be a good candidate to describe the base model in this study by trial and error. Backpropagation is chosen for the training of the SBC-RBFNN, where the gradient descent method is used to solve the fitting problem and the learning rate is set as 0.01 based on extensive training and testing. The maximum number of iterations for training is set to 1000 and the stop criterion is that the RMSE for the training set drops below 0.0001. The numbers of nodes in the three hidden layers are set to $l = 5$, $m = 5$, and $n = 25$, respectively, determined by trial-and-error.

Three benchmark algorithms are designed to verify the stability and effectiveness of the proposed offline model and they are described as follows. Benchmark 1 directly uses the base model generated from one cell or vehicle dataset to estimate the battery SOH of others. Such a method is widely used in previous research works such as [7]. Benchmark 2
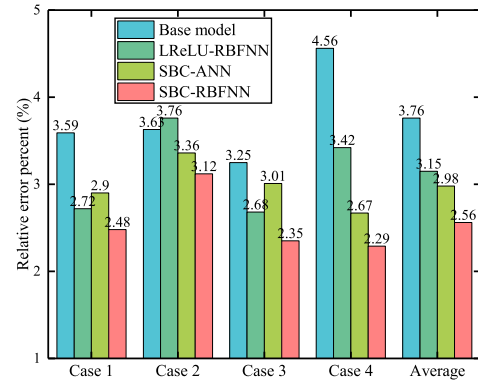
is a Leaky Rectified Linear Unit (LReLU) based RBFNN model. Due to its capability to deal with the negative part of datasets, the LReLU is superior to the traditional ReLU function widely used in ANN [16]. The structure and the configuration of the LReLU-RBFNN model are similar to that of the SBC-RBFNN model. The only difference is that the activation function in Layer 1 is replaced by the LReLU function. The major difference between the proposed SBC-RBFNN and the LReLU-RBFNN is that the latter lacks prior knowledge about the research subject while such information is available for the proposed SBC-RBFNN by model migration. Benchmark 3 is an SBC-ANN model proposed in [14], which is used to verify the superiority of the RBFNN. The training method, learning rate, and nodes in Layer 1 and Layer 2 of the SBC-ANN are both set the same as the SBC-RBFNN for a fair comparison.

The laboratory datasets are used to examine the performance of the proposed model following the methodology in Section III-A. Table I presents the verification results for all four dataset combinations, where the predictive RMSEs are calculated by (10). In Table I, it is clearly shown that the predictive accuracy can be significantly improved by using model migration: The SBC-RBFNN model can better rebuild the nonlinear relationships with a smaller predictive error compared to the SBC-ANN model. Furthermore, although trained by more datasets, the LReLU-RBFNN model still generates more significant predictive errors than the proposed SBC-RBFNN model. In fact, the LReLU-RBFNN even performs worse than the SBC-ANN model in terms of the average RMSE. This result exhibits the importance of the experience and information buried in the base model. The graphical results in Fig. 2 also show the superiority of the SBC-RBFNN model, especially in the region below 90% SOH, which indicates that the proposed method is more reliable than the two prevailing methods under comparison.

The relative error percentage (REP) is next used to signify the influences of estimated errors on the overall process of battery aging. The REP for the ith model, E_{ri}, is defined by

$$E_{ri} = E_i / S_i \times 100\% \qquad (15)$$

where E_i and S_i represent the predictive RMSE and the

TABLE I

RMSEs of Predicted SOH Between Different Offline Models based on the Laboratory Datasets.

Base Model	Training Cells	Model Type	Test Cells					Average
	Case 1		No.1	No.4	No.5	No.6	No.7	
No.2	⌐	Base Model	0.00498	0.00583	0.00620	0.00650	0.01215	0.00713
⌐	No.2, No.3 and No.8	LReLU-RBFNN	0.00583	0.00458	0.00379	0.00399	0.00921	0.00548
No.2	No.3 and No.8	SBC-ANN	0.00807	0.00527	0.00387	0.00474	0.00759	0.00591
		SBC-RBFNN	0.00716	0.00421	0.00293	0.00374	0.00727	**0.00506**
	Case 2		No.1	No.2	No.5	No.7	No.8	
No.4	⌐	Base Model	0.00820	0.00985	0.00421	0.00906	0.00734	0.00773
⌐	No.3, No.4 and No.6	LReLU-RBFNN	0.00938	0.01017	0.00372	0.00896	0.00797	0.00804
No.4	No.3 and No.6	SBC-ANN	0.00611	0.00870	0.00380	0.00897	0.00801	0.00712
		SBC-RBFNN	0.00613	0.00785	0.00322	0.00889	0.00697	**0.00661**
	Case 3		No.3	No.4	No.5	No.6	No.8	
No.1	⌐	Base Model	0.00493	0.00576	0.00633	0.00670	0.00972	0.00669
⌐	No.1, No.2 and No.7	LReLU-RBFNN	0.00692	0.00606	0.00293	0.00396	0.00828	0.00563
No.1	No.2 and No.7	SBC-ANN	0.00690	0.00649	0.00539	0.00574	0.00640	0.00618
		SBC-RBFNN	0.00559	0.00471	0.00305	0.00368	0.00762	**0.00493**
	Case 4		No.2	No.3	No.4	No.5	No.6	
No.7	⌐	Base model	0.01458	0.01098	0.00897	0.00631	0.00712	0.00959
⌐	No.7, No.1 and No.8	LReLU-RBFNN	0.01453	0.00935	0.00558	0.00321	0.00412	0.00736
No.7	No.1 and No.8	SBC-ANN	0.00834	0.00508	0.00527	0.00422	0.00498	0.00558
		SBC-RBFNN	0.00797	0.00470	0.00422	0.00330	0.00393	**0.00482**

range of the SOH for the ith cell, respectively. Comparisons of the calculated REPs are illustrated in Fig. 3 in terms of the average relative errors over the battery lifetime. It can be seen from Fig. 3 that by using the SBC-RBFNN method, the predictive accuracy is 32% higher than the direct estimated results using the base model, while the latter is the most commonly used method in the literature. In addition, by incorporating the model migration, we observe an improvement of about 19% and 14% in the predictive accuracy, compared to using the LReLU-RBFNN and the ANN-based methods, respectively.

B. Online Model Verification

The performance of the mRFR-based online adaptive method for individualized battery SOH estimation will be examined in this subsection. The core step of the proposed method is the weight generation (12)–(14), where the ensemble weights w_i are determined by the predictive errors of each offline model. By using the laboratory dataset, two offline models are trained. Datasets of three cells are extracted to produce the base model for the benchmark method, and the offline SBC-RBFNN models for ensemble learning to realize a fair comparison. Two weight update rules are examined. In the first rule, the predictive RMSEs of the first five data are chosen to generate w_i. In the second rule, two kinds of updating intervals are chosen. The estimated results using the offline SBC-RBFNN will be used as the benchmark in this subsection. The benchmark method is configured according to the suggestion given in [14], where the first 30% and 50% of the data are used to train the individual prediction model. The predictive accuracy is tested using the remaining data in the dataset.

The numerical results are compared in Table II based on the laboratory datasets. Considering that the data sizes for each of the eight cells are different, the training data percentage of the online ensemble learning method for each case is provided in the table. It shows that in order to improve the predictive accuracy in terms of the average RMSE, the size of the training dataset has to be increased for the offline SBC-RBFNN model. In contrast, increasing the data size for training is not necessary for our proposed ensemble learning method: Only a small amount of data are needed to achieve about 90% improvement in the estimation accuracy compared to using the offline SBC-RBFNN model trained by the first 30% of data. Compared to the offline model trained with the first 50% of data, the percentage of the improvement on the performance is still high up to 80%. It is worth noting that the estimation accuracy using the two weights generation methods is very close, and both methods are very effective. In this case, with the high quality of the laboratory dataset, only a few measurements at the early operating stage are needed to determine the individual battery aging pattern.

V. Conclusions

An online adaptive ensemble learning scheme based on a combination of offline model training and online weight generation has been proposed to deal with the divergence problem in the battery SOH estimation caused by inherent inconsistency between individual research objects. The effectiveness of the IC features for battery SOH estimation was discussed based on high-quality laboratory datasets. A pretrained model based on a combination of the SBC method with an RBFNN structure was developed. Four different dataset settings were chosen to test the stability of the

TABLE II

RMSEs of Predicted SOH Between Different Online Adaptive Methods Based on the Laboratory Datasets.

Methods	Base Model	Offline Model	Online Adaptive	Test Cells					Average
			Case 1	No.1	No.4	No.5	No.6	No.7	
SBC-RBFNN	No.2, No.3 and No.8	——	First 30% of the data	0.02608	0.03323	0.04294	0.06725	0.04321	0.04255
			First 50% of the data	0.02084	0.01701	0.01905	0.03417	0.02843	0.02390
mRFR	No.2	No.3 and No.8	First 5 data (9%)	0.00548	0.00451	0.00292	0.00393	0.00685	**0.00474**
			Every 10 cycles (10%)	0.00596	0.00422	0.00281	0.00383	0.00744	0.00485
			Every 5 cycles (20%)	0.00582	0.00406	0.00282	0.00371	0.00736	0.00475
			Case 2	No.1	No.2	No.5	No.7	No.8	
SBC-RBFNN	No.3, No.4 and No.6	——	First 30% of the data	0.06054	0.04565	0.04184	0.06345	0.07366	0.05703
			First 50% of the data	0.03307	0.03211	0.02489	0.02871	0.04746	0.03325
mRFR	No.4	No.3 and No.6	First 5 data (7%)	0.00412	0.00455	0.00358	0.00814	0.00601	**0.00528**
			Every 10 cycles (11%)	0.00564	0.00544	0.00352	0.00815	0.00574	0.00570
			Every 5 cycles (22%)	0.00446	0.00530	0.00341	0.00802	0.00567	0.00537
			Case 3	No.3	No.4	No.5	No.6	No.8	
SBC-RBFNN	No.1, No.2 and No.7	——	First 30% of the data	0.04937	0.04338	0.04704	0.03540	0.04041	0.04312
			First 50% of the data	0.02983	0.02306	0.01173	0.02419	0.03261	0.02428
mRFR	No.1	No.2 and No.7	First 5 data (9%)	0.00474	0.00539	0.00311	0.00420	0.00564	0.00462
			Every 10 cycles (10%)	0.00473	0.00380	0.00305	0.00401	0.00489	0.00409
			Every 5 cycles (20%)	0.00446	0.00350	0.00292	0.00351	0.00425	**0.00373**
			Case 4	No.2	No.3	No.4	No.5	No.6	
SBC-RBFNN	No.2, No.3 and No.8	——	First 30% of the data	0.04932	0.05399	0.03572	0.04574	0.05480	0.04792
			First 50% of the data	0.02878	0.02459	0.02619	0.01488	0.01808	0.02250
mRFR	No.7	No.1 and No.8	First 5 data (9%)	0.00368	0.00364	0.00411	0.00525	0.00511	0.00436
			Every 10 cycles (10%)	0.00403	0.00403	0.00448	0.00509	0.00491	0.00451
			Every 5 cycles (20%)	0.00374	0.00363	0.00428	0.00500	0.00484	**0.00430**

SBC-RBFNN model among several existing offline training methods. The results based on the laboratory datasets showed that the proposed SBC-RBFNN model can significantly improve the predictive accuracy. An online adaptive scheme was next constructed by synthesizing the offline models and an ensemble integration process using a proposed modified random forest regression method. The weights for each offline model were generated by online measuring a few data from fixed cycle or mileage interval, which markedly reduced the requirement of datasets compared with the online adaptive methods in previous research.

REFERENCES

[1] A. Manthiram, "A reflection on lithium-ion battery cathode chemistry," *Nature Commun.*, vol. 11, no. 1, p. 1550, 2020.
[2] Y. Li, M. Vilathgamuwa, S. S. Choi, B. Xiong, J. Tang, Y. Su, and Y. Wang, "Design of minimum cost degradation-conscious lithium-ion battery energy storage system to achieve renewable power dispatchability," *Appl. Energy*, vol. 260, p. 114282, 2020.
[3] J. Tian, R. Xiong, W. Shen, and F. Sun, "Electrode ageing estimation and open circuit voltage reconstruction for lithium ion batteries," *Energy Storage Mater.*, vol. 37, pp. 283–295, 2021.
[4] C. She, Z. Wang, F. Sun, P. Liu, and L. Zhang, "Battery aging assessment for real-world electric buses based on incremental capacity analysis and radial basis function neural network," *IEEE Trans. Ind. Informat.*, vol. 16, no. 5, pp. 3345–3354, May 2020.
[5] M. Dubarry, B. Y. Liaw, M.-S. Chen, S.-S. Chyan, K.-C. Han, W.-T. Sie, and S.-H. Wu, "Identifying battery aging mechanisms in large format Li ion cells," *J. Power Sources*, vol. 196, no. 7, pp. 3420–3425, Apr. 2011.
[6] M. Berecibar, I. Gandiaga, I. Villarreal, N. Omar, J. V. Mierlo, and P. V. den Bossche, "Critical review of state of health estimation methods of Li-ion batteries for real applications," *Renew. Sustain. Energy Rev.*, vol. 56, pp. 572–587, Apr. 2016.
[7] Y. Li, M. Abdel-Monem, R. Gopalakrishnan, M. Berecibar, E. Nanini-Maury, N. Omar, P. van den Bossche, and J. V. Mierlo, "A quick on-line state of health estimation method for Li-ion battery with incremental capacity curves processed by Gaussian filter," *J. Power Sources*, vol. 373, pp. 40–53, Jan. 2018.
[8] Y. Gao, J. Jiang, C. Zhang, W. Zhang, Z. Ma, and Y. Jiang, "Lithium-ion battery aging mechanisms and life model under different charging stresses," *J. Power Sources*, vol. 356, pp. 103–114, Jul. 2017.
[9] X. Li, Z. Wang, and J. Yan, "Prognostic health condition for lithium battery using the partial incremental capacity and Gaussian process regression," *J. Power Sources*, vol. 421, pp. 56–67, May 2019.
[10] 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," *J. Power Sources*, vol. 396, pp. 453–458, Aug. 2018.
[11] C. Birkl, "Diagnosis and prognosis of degradation in lithium-ion batteries," Ph.D. dissertation, University of Oxford, 2017.
[12] R. R. Richardson, C. R. Birkl, M. A. Osborne, and D. A. Howey, "Gaussian process regression for in situ capacity estimation of lithium-ion batteries," *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 127–138, 2018.
[13] A. Fly and R. Chen, "Rate dependency of incremental capacity analysis (dQ/dV) as a diagnostic tool for lithium-ion batteries," *J. Energy Storage*, vol. 29, p. 101329, Jun. 2020.
[14] X. Tang, K. Liu, X. Wang, F. Gao, J. Macro, and W. D. Widanage, "Model migration neural network for predicting battery aging trajectories," *IEEE Trans. Transport. Electrific.*, vol. 6, no. 2, pp. 363–374, Jun. 2020.
[15] Y. Li, C. Zou, M. Berecibar, E. Nanini-Maury, J. C.-W. Chan, P. van den Bossche, J. Van Mierlo, and N. Omar, "Random forest regression for online capacity estimation of lithium-ion batteries," *Appl. Energy*, vol. 232, pp. 197–210, 2018.
[16] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical evaluation of rectified activations in convolutional network," *ArXiv Preprint*, 2015.