# Capacity Prediction and Validation of Lithium-Ion Batteries Based on Long Short-Term Memory Recurrent Neural Network

**ZHENG CHEN**[1,2], **(Senior Member, IEEE), QIAO XUE**[1], **YITAO WU**[1], **SHIQUAN SHEN**[1], **YUANJIAN ZHANG**[3], **(Member, IEEE), AND JIANGWEI SHEN**[1]

[1]Faculty of Transportation Engineering, Kunming University of Science and Technology, Kunming 650500, China
[2]School of Engineering and Materials Science, Queen Mary University of London, London E1 4NS, U.K.
[3]School of Mechanical and Aerospace Engineering, Queen's University Belfast, Belfast BT9 5AG, U.K.

Corresponding author: Jiangwei Shen (shenjiangwei6@kust.edu.cn)

**ABSTRACT** Capacity prediction of lithium-ion batteries represents an important function of battery management systems. Conventional machine learning-based methods for capacity prediction are inefficient to learn long-term dependencies during capacity degradations. This paper investigates the deep learning method for lithium-ion battery's capacity prediction based on long short-term memory recurrent neural network, which is employed to capture the latent long-term dependence of degraded capacity. The neural network is adaptively optimized by the Adam optimization algorithm, and the dropout technique is exploited to prevent overfitting. Based on the offline cycling aging data of batteries, the capacity prediction performance is validated and evaluated. The experimental results demonstrate that the proposed algorithm can accurately track the nonlinear degradation trend of capacity within the whole lifespan with a maximum error of only 2.84%.

**INDEX TERMS** Lithium-ion batteries, capacity prediction, aging factors, long short-term memory (LSTM).

## NOMENCLATURE

### A. ABBREVIATIONS

| | |
|---|---|
| EVs | electric vehicles |
| EOL | end of life |
| BMS | battery management system |
| SOC | state of charge |
| SOH | state of health |
| GPR | Gaussian process regression |
| ECM | equivalent circuit model |
| OCV | open circuit voltage |
| KF | Kalman filter |
| PF | particle filter |
| RC | resistance-capacitance |
| RLS | recursive least squares |
| IC | incremental capacity |
| RUL | remaining useful life |
| SVM | support vector machine |

| | |
|---|---|
| HFs | healthy features |
| RNN | recurrent neural networks |
| LSTM | long short-term memory |
| LOOCV | leave-one-out cross validation |
| IR | internal resistance |
| SEI | solid electrolyte interphase |
| CEI | cathode electrolyte interface |
| DIC | discharge incremental capacity |
| GRA | grey relational analysis |
| MAE | maximum absolute error |
| MSE | mean square error |
| RMSE | root mean square error |

### B. SYMBOLS

| | |
|---|---|
| $F_1$ | the internal resistance of battery |
| $F_2$ | the average temperature of battery in a charging and discharging cycle |
| $X_i$ | the aging factors dataset in the training set |

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

| | |
|---|---|
| $X_i^*$ | the aging factors dataset in the test set |
| $\hat{y}_i$ | the predicted value of capacity |
| $y_i^*$ | the observed value of capacity in the test set |
| $h_{t-1}$ | the last output of cell state in the LSTM-RNN architecture |
| $x_t$ | the current cell input |
| $f_t$ | the outputs of forget gate |
| $\sigma$ | the sigmoid function |
| $W_f$ | the weight matrix of forget gate |
| $b_f$ | the bias of forget gate |
| $i_t$ | the sigmoid layer of input gate |
| $a_t$ | the output of tanh layer for the input gate |
| $W_i$ | the weight matrix of sigmoid layer of input gate |
| $W_c$ | the weight matrix of tanh layer of input gate |
| $b_i$ | the bias of sigmoid layer of input gate |
| $b_c$ | the bias of tanh layer of input gate |
| $c_t$ | the current cell state |
| $h_t$ | the output of LSTM-RNN at current moment $t$ |
| $W_o$ | the weight matrix of the output layer |
| $b_o$ | the bias of output layer |
| $o_t$ | the sigmoid layer of output gate |
| $J(\theta)$ | the objective function with $\theta$ |
| $g_t$ | the gradient with parameter $\theta_{t-1}$ |
| $m_t$ | the exponential moving average of gradient |
| $v_t$ | the exponential moving average of squared gradient |
| $\beta_1$ | the exponential decay factors accounting for weight distribution |
| $\beta_2$ | the exponential decay factors that controls the influence incurred by squared gradient |
| $\hat{m}_t$ | the modified values of $m_t$ |
| $\hat{v}_t$ | the modified values of $v_t$ |
| $\alpha$ | the learning rate of Adam algorithm |
| $\varepsilon$ | a smooth coefficient |
| $p$ | the probability of dropout for the neuron |
| $X_{Ci}$ | the complete features dataset of $i$th cell |
| $R^2$ | goodness-of-fit |
| $n$ | the total sample number |
| $\bar{y}_i$ | the sampling average value |

## I. INTRODUCTION

Lithium-ion batteries have been widely deployed in electric vehicles (EVs) and energy storage systems of power grids due to their high energy/power density, no memory effect and long lifespan [1]. However, with the cyclic charging and discharging operations, battery's capacity degradation and electrical performance deterioration can influence vehicle operation performance and safety. In particular, when the capacity decreases below 80% of its initial value, lithium-ion batteries turn to be unstable and degrade faster than before, implying that they reach end of life (EOL) [2], and the continued operation of batteries may lead to irreversible damage. As such, accurate diagnosis for battery health condition

becomes an indispensable task [3]. In practice, a serviceable battery management system (BMS) is essential to ensure operating efficacy and battery safety [4]. One main function of BMS is to conduct inner status estimation of batteries, such as state of charge (SOC) and state of health (SOH) [5]. Accurate capacity information can supply an important foundation for SOC and SOH estimation, and also provide valuable indexes to end-users and battery manufactures [6].

To now, extensive research has been conducted to improve capacity prediction accuracy of batteries. The conventional capacity estimation methods can be categorized into two types: model-based methods and data-driven methods [7]. Reference [8] proposes a two-stage scheme for battery capacity estimation according to the variation of thermal dynamics. In the first stage, the estimation for battery core temperature and heat generation is implemented, and then a joint estimation for both SOC and capacity are exerted in the following stage. Reference [9] builds a capacity fading model based on the sample entropy, which is employed to calculate the battery surface temperature in the charging process. Through considering the influence of heat generation on capacity attenuation, the particle filter (PF) is exploited to estimate the battery remaining capacity. Reference [10] presents a data-driven diagnostic technique for capacity estimation based on Gaussian process regression (GPR), in which the voltage measurement over a short period of galvanostatic phase is considered as the model input. In [11], considering the consistency among cells connected in series, the variation characteristics of voltage are extracted from two different cycles by conducting the dynamic time warping algorithm. Based on the extracted feature, a three-step capacity estimation method with the theoretical foundation of shape invariance of the charging voltage is proposed to calculate the capacity difference between two adjacent cells. In these mentioned prediction methods, the estimated capacity value is severed as an intermediary for other state estimation. Considering the coupling relationship, joint estimation of capacity with other battery states, such as co-estimation of capacity and SOC, are usually carried out sequentially.

For the co-estimation of capacity and SOC, some joint algorithms become attracted due to their satisfactory precision and robustness [12]. Based on effective electrical models such as equivalent circuit model (ECM), a number of advanced filter algorithms (such as Kalman filter (KF) and PF) can then be adopted to conduct the joint estimation of battery status and model parameters [13]. In [14], a second-order resistance-capacitance (RC) ECM is established, and the square root cubature KF is employed to estimate the SOC. Meanwhile, the capacity, as one of the key parameters of model, is identified by the genetic algorithm (GA). Reference [15] proposes a multiscale dual H-infinity filter to estimate the SOC and capacity of battery in real time with different timescales for reaction to slow varying battery capacity and fast varying battery state. To address the different variation rates of model parameters, [16] presents a joint algorithm integrated by KF and the recursive least

square (RLS) method to estimate SOC and capacity, in which the model parameters are adaptively updated by a vector-type RLS. For the sake of enhancing the estimation precision. Reference [17] constructs a serially connected battery pack model based on a second-order RC ECM of cell. Then, a multiscale extended KF algorithm is employed to accurately estimate SOC, model parameter and capacity of single cell in battery packs.

In addition to SOC, accurate estimation of capacity, as mentioned above, is also of importance for health diagnosis of battery. As an indicator of assessing the battery degradation status, SOH is usually indexed by the ratio of current maximum useful capacity over the rated value [18]. In [19], a fusion method incorporating partial incremental capacity (IC) analysis and a dual GPR model is proposed to estimate the SOH of lithium-ion batteries. To improve the SOH estimation accuracy and reliability, [20] extracts four feature vectors representing the degradation status of battery from the charging voltage curves. Consequently, the SOH prediction is attained via the well-tuned GPR model with the extracted features as the inputs. By incorporating the critical features derived from battery operation data set, [21] proposes a real-time estimator for remaining useful life (RUL) prediction based on the SVM model. In [22], the SVM model is constructed with a radial basis function kernel, and the feature variables are extracted from partial charging voltage curves to construct the training dataset. In addition, the kernel parameters of SVM model are optimized by the grid search method. Reference [23] leverages the conjugate gradient method and multi-island GA to optimize the hyper parameters of GPR model, and the characteristic parameters of constant-current charging process are extracted as the healthy features by the IC analysis method. On this basis, the SOH estimation is attained by combing the extracted features and the optimized GPR model. In short, all of these data-driven methods need healthy features to establish a mapping relationship between SOH and feature variables. In other words, a reliable SOH estimation strongly requires proper feature extraction to perform qualified SOH diagnosis [24]. However, lithium-ion battery degradation is consecutive and generally involves hundreds to thousands of cycles, and the later degradation evolution is highly related with the former degradation information throughout these cycle operations. Moreover, the healthy features extracted from the charging and discharging profiles also show a specific variation trend with the aging. These variables can be regarded as a time series signal, of which the current values may exhibit long-term dependencies with historical values. Nevertheless, the conventional data-driven methods, such as SVM and GPR, are inefficient to learn the long-term dependencies, thus it remains challenging to maintain high estimation accuracy for long-term capacity prediction [25].

Presently, deep learning network has received widely attention and has been progressively applied in the language modeling [26] and image recognition [27]. As a kind of deep learning network, long short-term memory recurrent neural network (LSTM-RNN) is employed to solve the problem with long-term dependences. LSTM-RNN can reserve the key information from the degradation data via effective learning of long-term dependence based on the specific gate [28]. Given long-term characteristic of battery degradation, LSTM-RNN may be a suitable solution to learn the long-term degradation trend of capacity variation. Reference [29] exploits the LSTM-RNN to learn the long-term dependence of degraded capacity of supercapacitor, and the experimental results show that the LSTM-RNN can predict the RUL of the supercapacitor on the rest testing data with 2.61% root mean square error (RMSE). Reference [30] employs the LSTM-RNN to predict RUL of lithium-ion batteries. The elastic mean squared back-propagation algorithm and Monte Carlo simulation are respectively applied to adaptively optimize the network and generate a probabilistic RUL prediction. However, the LSTM-RNN in [29], [30] is trained based on the historical capacity degradation data, and then one- and multi-step forward RUL prediction is performed. In addition, [29] reveals that the capacity degradation trajectory of lithium-ion batteries is approximate to the linear degradation, thus the decline rate of capacity under the whole cycle life is similar. Nonetheless, the degradation rate of battery is significantly different in the beginning and ending stage of whole lifespan. Therefore, the prediction accuracy of capacity based on only partial degradation data for model training needs to be further analyzed. Furthermore, when incomplete offline data is available, whether the LSTM-RNN can also accurately predict the battery remaining capacity in the whole lifespan still needs to be investigated and validated.

Motivated by this, the capacity prediction of lithium-ion batteries based on the LSTM-RNN is carefully conducted. Firstly, the LSTM-RNN is optimized based on the Adam optimization algorithm, and the dropout technology is employed to prevent the network from overfitting. Then, the optimized LSTM-RNN is exploited to achieve the capacity prediction of lithium-ion batteries. Whereupon, this paper conducts the validation and comparison of capacity prediction effectiveness of LSTM-RNN for the lithium-ion batteries from the following four aspects. (1) The influence of aging factors on the performance of capacity prediction for lithium-ion batteries based on LSTM-RNN is discussed. (2) The capacity prediction results based on the LSTM-RNN are compared with the prediction results of SVM, GPR and Elman NN. (3) Through one battery's whole cycle life data for model training, others battery's data are validated to examine the prediction performance of the built LSTM-RNN model. (4) The leave-one-out cross validation (LOOCV) method is applied to evaluate the performance of the LSTM-RNN with the aging factors as model inputs.

The remainder of this paper is structured as follows. The battery life cycle test is introduced, and the experimental data are analyzed in Section II. Section III illustrates the detailed capacity prediction process of lithium-ion batteries based on LSTM-RNN. The validation and comparison of prediction
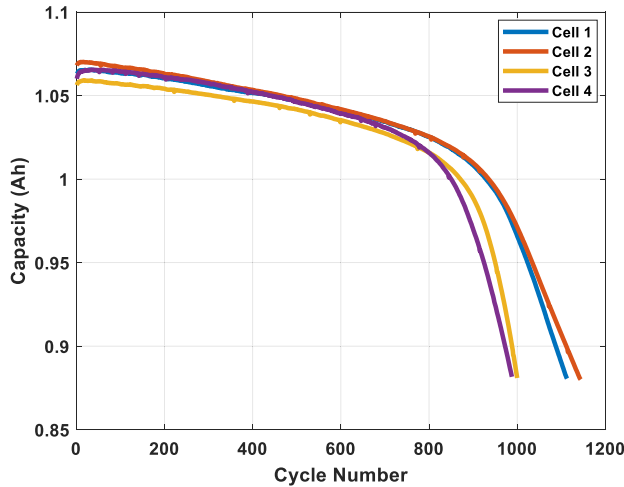
results are elaborated in Section IV, and Section V concludes the study.

## II. BATTERY AGING EXPERIMENTAL AND DATA ANALYSIS

In this section, the cycle life experimental and the degradation data acquired from a huge cycling data repository are introduced. Based on the cycling data, the aging factors are extracted to represent the battery capacity variation. After that, the framework and process of capacity prediction are illustrated in a schematic diagram.

### A. BATTERY AGING EXPERIMENTAL AND CAPACITY DEGRADATION DATA

In this study, the cyclic aging data of lithium-ion batteries are obtained from an open source experimental data repository [31], which collects the cyclic life tests of a variety of commercial LFP/graphite batteries (nominal capacity of 1.1 Ah and rated voltage of 3.3 V). The upper and lower cut-off voltages of the battery are 3.6 V and 2.0 V, respectively. The charging policy follows a form of C1(Q1)-C2 mode, where C1 and C2 denote the first and second constant current stage, and Q1 denotes the SOC at which the current changes. The second current step ends at 80% SOC, after that the cell is charged with 1C (C denotes the rate capacity value, i.e., 1.1) constant current (CC)-constant voltage (CV) mode, and the cells are discharged with 4C current. During the experiment, the surface temperature and internal resistance are measured and recorded. Note that the internal resistance measurement is conducted during charging at 80% SOC by imposing 10 pulses of ±3.6C current with the duration of 33 ms [31]. Seven cells' data (labeled as Cells 1 to 7) are selected from this dataset to investigate the performance and effectiveness of the LSTM-RNN model for capacity prediction.

The curves of degradation capacity are shown in Fig. 1, which highlight that the degradation trajectories of four cells remain almost the same, indicating that the degradation
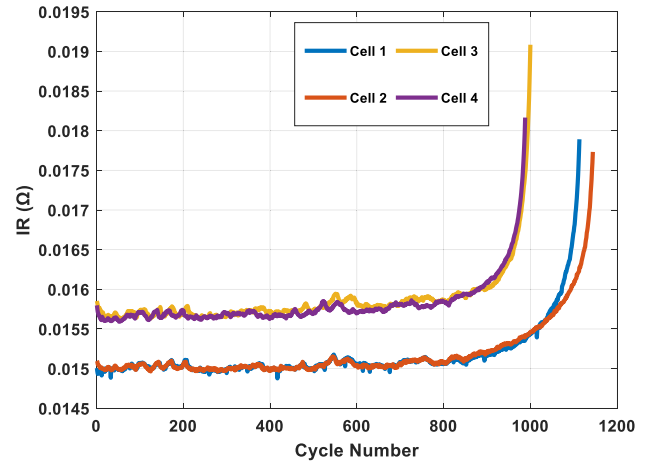
mechanism is nearly consistent for the same type of lithium-ion batteries. The cycle life experiments for all batteries are terminated when the batteries reached 80% of nominal capacity, i.e., 0.88 Ah. It can also be found that the degradation slope is relatively flat before 90% SOH. However, when the SOH drops less than 90%, the capacity degradation shows an exponential decline trend with faster dropping speed. Besides, the electric characteristics will gradually deteriorate during the aging process, the thermal characteristics of batteries will also vary with aging [32]. Next, the aging factors will be extracted from electric and thermal characteristics variation of the battery.

### B. AGING FACTORS EXTRACTION AND ANALYSIS

From the perspective of electric characteristics, one main change during degradation is that the internal resistance (IR) will gradually increase. During the battery aging process, the formation and thickening of the SEI film, the cathode electrolyte interface (CEI) formation and the internal structure disordering can lead to the increase of IR. However, the aforementioned issues cannot be measured directly; and by contrast, the measured IR, as a representative variable, will vary in a nonlinear manner relating to the capacity degradation [32]. As can be seen from Fig. 2, significant variation of IR does not obviously appear in the cycle of [1, 800] but with an exponential variation trend after cycle 800. The IR increase represents the capacity degradation to some extent with the form of inverse proportional function. In other words, the more obviously IR increases, the faster the capacity declines. Therefore, the battery IR, denoted by $F_1$, can be selected as an aging factor.

Considering the battery's thermal characteristics, the surface temperature at each moment is recorded during the experiment. Due to the IR increase and active materials loss of contact caused by the current collector corrosion, binder decomposition and electrolyte loss, the generation of ohmic heat and the heat distribution inside battery differ greatly under the same charge/discharge C-rate when the battery
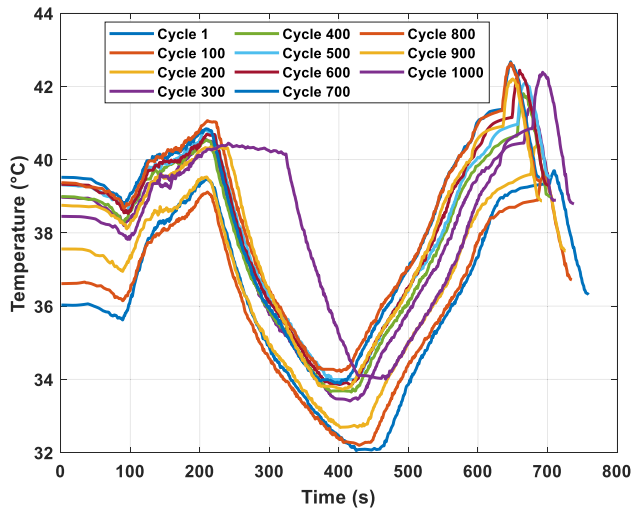
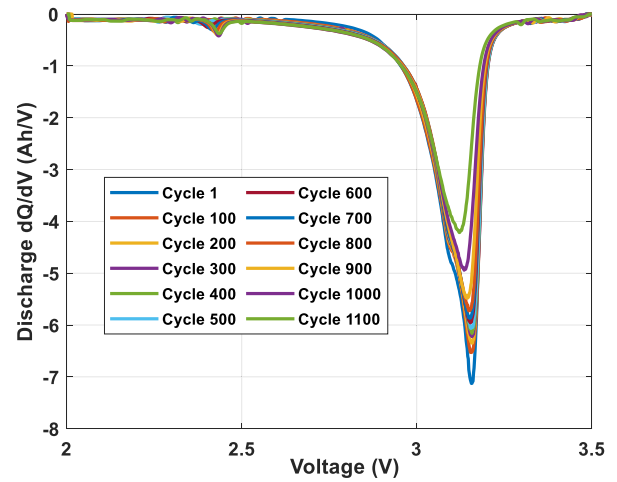**FIGURE 3.** The temperature variation curves of Cell 1 at different cycles.



**FIGURE 5.** The variation curves of discharge incremental capacity at different cycle numbers for Cells 1.
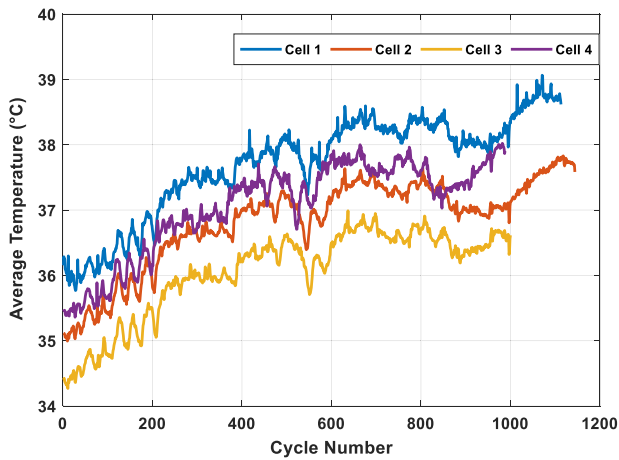


**FIGURE 4.** The evolution trend of average temperature with cycle numbers for Cells 1-4.
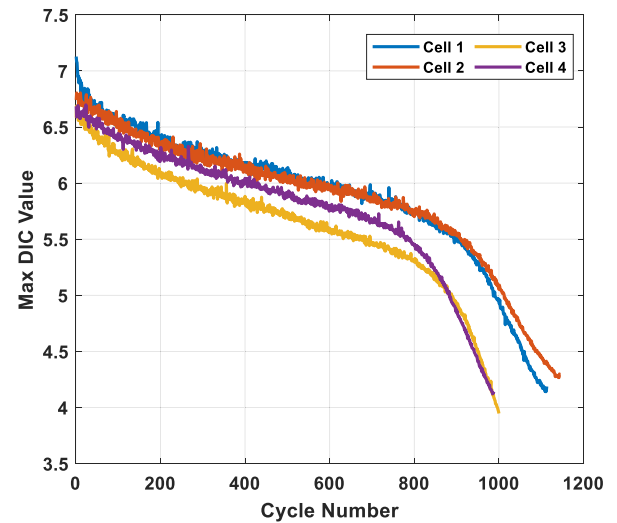


**FIGURE 6.** The evolution trend of max DIC value with cycle numbers for Cells 1-4.

ages [33]. Fig. 3 shows the temperature variation curves of Cell 1 at different cycles. It can be seen that the temperature shows an augmented trend with the increase of cycle number. Intuitively, the average temperature of each cycle is calculated to analyze the thermal characteristics of battery. The variation of average temperature with different cycle times is shown in Fig. 4. It is clearly observed that the average temperature increases progressively with the cycle experiment. Based on the variation relationship of capacity and average temperature with the cycle number, it can be concluded that the average temperature can also represent the capacity degradation. Consequently, the average temperature of each cycle can be selected as another aging factor, denoted by $F_2$. Except for the IR and temperature, the aging factors can also be extracted from the charge/discharge voltage profiles. Since the experimental battery in this study is discharged with constant current, the incremental capacity analysis during the discharging process is conducted. The discharging incremental capacity (DIC) curves at different cycles for Cell 1 are

shown in Fig. 5. In addition, Fig. 6 shows the variation curves of peak absolute value with cycle numbers for Cells 1 to 4. As can be seen, the absolute value of peak decreases gradually with the increase of cycle number, implying that the absolute value of DIC peak point can effectively characterize the battery degradation. Hence, the absolute value of DIC peak can also be considered as one aging factor, called F3. Next step, the implied relationships between aging factors and capacity will be analyzed.

### C. CORRELATION ANALYSIS OF AGING FACTORS BASED ON GRA

As discussed previously, the IR, average temperature and the absolute value of DIC peak, denoted as $F_1$, $F_2$ and $F_3$, are selected as the aging factors to characterize the capacity degradation. To further analyze the relationship between aging factors and capacity, we took cell 1 as an example, and the variation relationships between the aging factors and
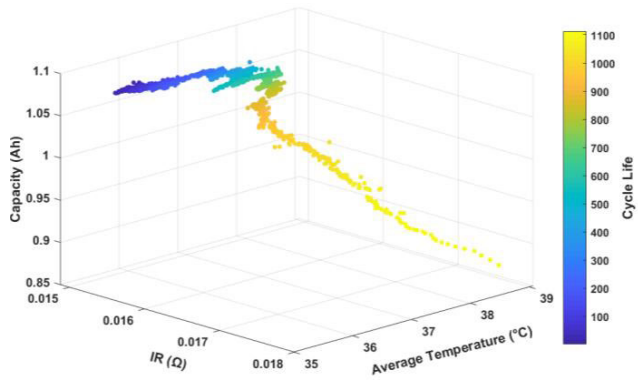
**FIGURE 7.** The variation relationship between capacity and $F_1$ and $F_2$ with cycle numbers of Cell 1.
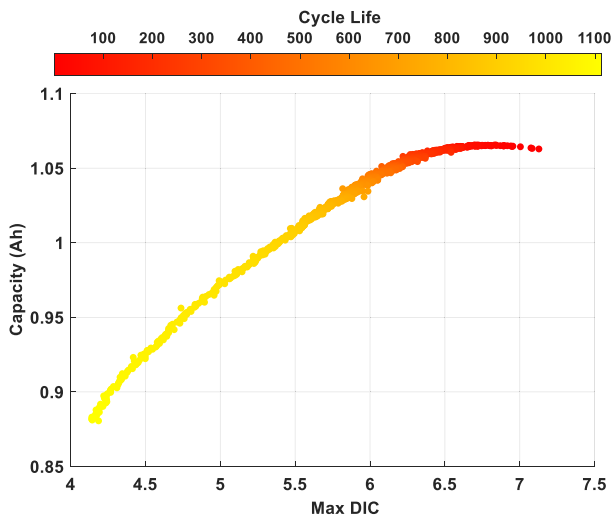


**FIGURE 8.** The variation relationship between capacity and $F_3$ with cycle numbers of Cell 1.

capacity with respect to cycle life are shown in Fig. 7 and Fig. 8, where the color scale represents the cycle life. As can be found, $F_1$ and $F_2$ increase and $F_3$ decreases with the capacity degradation. Additionally, in the early and middle phases of cycle life (1 to 800 cycles), the capacity degrades with a slow speed, and $F_1$ remains almost unchanged; and in contrast, $F_2$ increases obviously and $F_3$ gradually decreases with the increase of cycle numbers. Comparatively, in the ending phase of cycle life (800 to 1100), the capacity degradation and the increase of $F_1$ are faster, and the increase rate of $F_2$ becomes slower and more stabilized; however, $F_3$ still decreases obviously. It can be concluded that the change of $F_1$ is not obvious, whereas the variation of $F_2$ is relative larger in the early cycle life. In the later cycle life stage, the changes of $F_1$ and $F_2$ are opposite to that of the early stage. Furthermore, there exists obvious change in $F_3$ throughout the whole cycle life. To sum up, a kind of mapping relationship between the aging factors and capacity really exists in different cycle life phases. In this study, the correlation between the aging factors and battery capacity is further evaluated by grey relational analysis (GRA). As a crucial method based on the

**TABLE 1.** Grey relational grades between features and capacity.

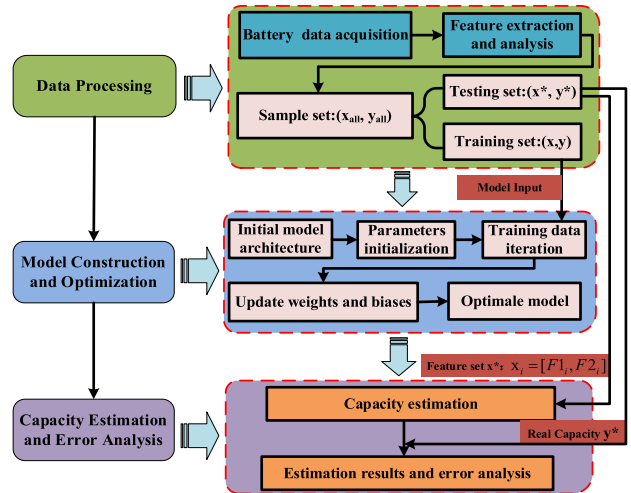| Aging Factors | Battery Number | | | |
|---|---|---|---|---|
| | Cell 1 | Cell 2 | Cell 3 | Cell 4 |
| $F_1$ | 0.5883 | 0.5902 | 0.5691 | 0.5761 |
| $F_2$ | 0.7646 | 0.7819 | 0.7818 | 0.7708 |
| $F_3$ | 0.5769 | 0. 6022 | 0.5793 | 0.5873 |



**FIGURE 9.** The framework and flowchart of capacity prediction.

grey system theory, GRA evaluates the correlation among the elements according to the similarity and dissimilarity of their variation trend. The quantitative analysis based on the GRA is to obtain the correlations between reference and comparative sequences, as detailed in [34]. Through GRA, the correlation grades between aging factors and capacity of each cell are acquired, as shown in Table 1. Particularly, the correlation grade of $F_2$ is greater than 0.75 for all the cells, which means the selection of aging factors is feasible for capacity estimation.

### D. THE FRAMEWORK AND FLOWCHART FOR CAPACITY PREDICTION

Fig. 9 shows the framework and flowchart of capacity prediction based on the LSTM-RNN model. As can be seen, the whole prediction process contains the experimental data processing, the model construction and the capacity prediction modules. In the data processing module, the aging factors data set $X_i = [F_{1i}, F_{2i}, F_{3i}]$, where the subscript $i$ represents the cycle number, is structured based on the extracted characteristic features from the aging experimental data. The sample set is divided into the training set and the test set. In the model construction and optimization module, the architecture and network layers of LSTM-RNN is firstly designed, and the model parameters are initialized. Then, the aging factors data set $X_i$ and the corresponding capacity value $y_i$ in the training set are considered as the LSTM-RNN model's input and output, respectively. The optimal model parameters are searched via test and cross validation. In the capacity prediction and
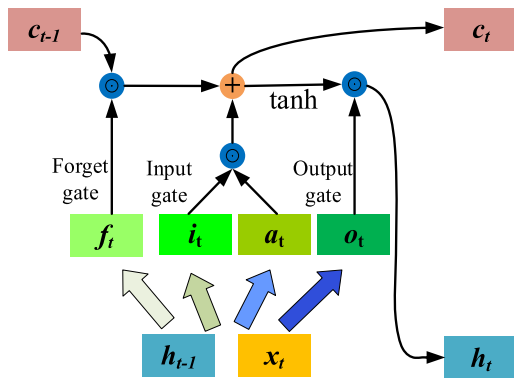
**FIGURE 10.** The network architecture of LSTM.

error analysis module, similarly, the aging factors data set $X_i^*$ in the test set is inputted into the well-tuned model, and then the output $\hat{y}_i$ is collected as the prediction value of battery capacity. By calculating evaluation criteria and comparing the predicted value $\hat{y}_i$ with the observed value $y_i^*$, the prediction effectiveness of LSTM-RNN model is assessed.

## III. METHODOLOGIES

This section elaborates the mechanism and derivation of related model and algorithms applied for the capacity prediction, including the LSTM-RNN, the Adam optimization algorithm and the dropout technique. In addition, the evaluation criteria and LOOCV method are addressed to evaluate the performance of LSTM-RNN based capacity estimation algorithm.

### A. THE ARCHITECTURE OF LSTM-RNN

LSTM-RNN is a kind of specialized RNN for solving vanishing gradient problems and gradient explosion problems with long-term dependency [35]. Compared with the simple RNN, the LSTM-RNN adds a state $c$ in the hidden layer to keep the long-term state, and this newly added state $c$ is called the cell state [29]. The structure of LSTM-RNN is shown in Fig. 10. Note that the subscript $t$ of each vector represents the moment state, which denotes the generality of LSTM-RNN applications. For the capacity prediction of lithium-ion batteries, the moment state means the cycle number. At moment $t$, there are three inputs for the LSTM-RNN: the input variable $x_t$ of the current time network, the output value $h_{t-1}$ and the cell state $c_{t-1}$ in the previous step. Meanwhile, the LSTM-RNN has two outputs: the output value $h_t$ and cell state $c_t$ at current moment $t$.

Similar with classic RNNs, LSTM-RNN is composed of the input layer, hidden layer and output layer. However, the hidden layer in LSTM-RNN is with a specialized memory mechanism, instead of a general neuron. The internal state of LSTM-RNN at moment $t$ is called $c_t$, which is critical to the network and locates at the heart of each neuron that is linearly activated. The internal state can be regarded as a carrier, to which the information has been added or from which has been removed. This information processing can be carefully

regulated by the so-called gate [30]. The gate is a distinctive feature of LSTM-RNN, which actually denotes the fully connected layers. There are three gates, namely, the input gate, forget gate and output gate, in a LSTM-RNN architecture. Any read or modification operation can be achieved through controlling of these three gates. Additionally, the information selection of gate is mainly conducted by the sigmoid function, tanh function or matrix multiplication [24].

It can be seen from Fig. 10 that the first step of applying the LSTM-RNN is to decide what information should be discarded by the forget gate, which reads $h_{t-1}$ and $x_t$, and outputs a value $f_t$ between 0 and 1, where the upper bound 1 indicates that the information should be totally kept; and by contrast, the lower bound 0 means that it should be thoroughly discarded. The next step is to determine what information should be stored in the memory gate. One part of the input gate $i_t$, called the sigmoid layer, decides what information should be updated, and another part, called the tanh layer, creates the candidate vector $a_t$, which is added to the current cell state. Finally, by means of the updated cell state $c_t$ and the value $o_t$ of output gate, the output of LSTM-RNN can be calculated. Based on the previous moment output $h_{t-1}$ and the input of current moment $x_t$, the state values of three gates and the candidate vector $a_t$ can be formulated as:

$$\begin{cases} f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ a_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \\ o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \end{cases} \quad (1)$$

where $h_{t-1}$ is the last output of cell state, $x_t$ is the current cell input, $\sigma$ represents the sigmoid function, $W_f$ is the weight matrix of forget gate, and $b_f$ is the bias of forget gate; $W_i$ and $W_c$ denote the weight matrix of sigmoid layer and tanh layer of input gate, respectively; $b_i$ and $b_c$ represent the bias of sigmoid layer and tanh layer of input gate; $W_o$ and $b_o$ denote the weight matrix and bias of the output layer. When the state values of each gate are determined, the current cell state $c_t$ and the output of LSTM-RNN can be calculated, as:

$$\begin{cases} c_t = f_t \times c_{t-1} + i_t \times a_t \\ h_t = o_t \times \tanh(c_t) \end{cases} \quad (2)$$

Based on the above discussion, LSTM-RNN can reach the purpose of learning the long-term dependences of capacity degradation and performing one- or multi-step forward prediction. Next, the training algorithm will be detailed to search the optimal weight matrices and biases for capacity prediction.

### B. OPTIMIZATION TRAINING FOR LSTM-RNN

In this study, the Adam optimization algorithm is employed to optimize the parameters of LSTM-RNN. The Adam algorithm is a first-order gradient optimization method that mainly accounts for optimizing the gradient of stochastic objective function based on adaptive estimates of lower-order moments. Compared with traditional random gradient

descent algorithms, it advances higher computational efficiency, lower RAM occupation, less turning labor and better dominance in solving large-scale parameter optimization. Reference [36] experimentally validates that the Adam algorithm is more efficient in solving deep learning problems, compared with the RMSprop [37] and AdaGrad algorithm [38]. The parameter updating process of Adam algorithm is detailed as follows. Firstly, at step $t$, the gradient of optimization objective is calculated, as:

$$g_t = \nabla_\theta J(\theta_{t-1}) \tag{3}$$

where $J(\theta)$ represents the objective function with $\theta$, $g_t$ denotes the gradient with $\theta_{t-1}$. At step $t$, the exponential moving average value of both gradient and squared gradient $m_t$ and $v_t$, are respectively calculated, as:

$$\begin{cases} m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \\ v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \end{cases} \tag{4}$$

where $\beta_1$ and $\beta_2$ denote the exponential decay factors for weight distribution and influence incurred by squared gradient. In general, the initial value of $m_0$ and $v_0$ is set to zero, $m_t$ and $v_t$ are adjusted to zero in the initial stage of training process. Thus, a modification will be applied to reduce the training error, as:

$$\begin{cases} \hat{m}_t = m_t / (1 - \beta_1) \\ \hat{v}_t = v_t / (1 - \beta_2) \end{cases} \tag{5}$$

where $\hat{m}_t$ and $\hat{v}_t$ denote the modified values of $m_t$ and $v_t$. The parameters are updated as:

$$\theta_t = \theta_{t-1} - \alpha \times \hat{m}_t / (\sqrt{\hat{v}_t} + \varepsilon) \tag{6}$$

where $\alpha$ denotes the learning rate, and $\varepsilon$ expresses the smooth coefficient for avoiding the denominator from zero. The remaining parameters of Adam algorithm are set to $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\alpha = 0.001$, and $\varepsilon = 10^{-8}$.

## C. DROPOUT TECHNOLOGY TO PREVENT LSTM-RNN FROM OVERFITTING

Overfitting refers to the model's ability of fitting the training data set well but showing inferior fitting effect in the test data set [29]. To address this issue, the dropout technique is employed to prevent the LSTM-RNN from overfitting [39]. Generally, the error back-propagation method is applied to iteratively adjust the parameters for each Mini-Batch in the RNN training process. The key idea of dropout technique is that it removes the neurons from the layers of RNN during the training process to prevent the model from overfitting. The neurons along with all its connections are temporarily discarded from the network, as shown in Fig. 11. It is essentially a random process during which one stochastic neuron is selected to remove. Therefore, each neuron will be retained with a fixed probability $p$, which is set to 0.4 in this paper. It can be seen from Fig. 11 (b) that the NN model after applying the dropout technique is equivalent to sampling a condensed network from it. The condensed network consists
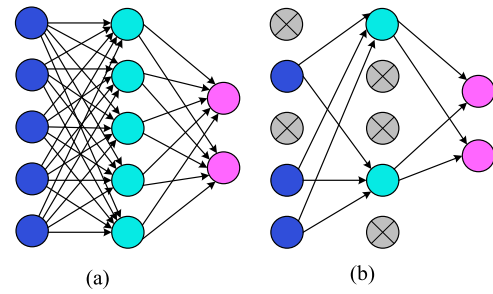


**FIGURE 11.** The schematic diagram of dropout neural network model. (a) standard neural network with 1 hidden layer; (b) neural network after applying Dropout technology.



**FIGURE 12.** The schematic of leave-one-out cross validation process.

of all remaining neurons and their connections after removing the discarded neurons. Hence, training a neural network with dropout can be regarded as training many condensed networks with extensive weight sharing, where each condensed network is trained rarely [30]. By this manner, the network becomes less sensitive to the specific weights of neurons, which in turn results in that the network is with the better generalization capability.

## D. LEAVE-ONE-OUT CROSS VALIDATION

In this paper, the LOOCV method is employed to verify the performance of the LSTM-RNN for capacity prediction [40]. The schematic diagram of LOOCV applied in this study is illustrated in Fig. 12. The complete feature data set contains four subsets $X_{C1}$, $X_{C2}$, $X_{C3}$ and $X_{C4}$ which are combined with four cells' aging factors extracted from the experimental data. Each subset is composed of three features vectors $\{F_1, F_2, F_3\}$, namely IR, average temperature and absolute value of DIC peak. As illustrated in Fig. 1, the degradation curves of four cells are similar, indicating the degradation mechanism is coincident for one type batteries. Therefore, it is feasible to train model with one cell's data and test with others cells' data for validating the prediction effectiveness of the proposed LSTM-RNN model. In this work, we suppose one cell's data as the test dataset and compile other three cells' data together as the training dataset, as shown in Fig. 12. Since the validation datasets are not imported in the training

process, the trained model can provide an approximately unbiased estimation [13]. The training and test process is repeated four times, and thus each battery cell is used as the test dataset, and we can conclude that it is equivalent to perform a 4-fold cross validation for the LSTM-RNN model. After each iteration, the prediction error and evaluation criteria are calculated to assess the model performance. Next, the evaluation criteria applied in this study are introduced.

### E. THE PERFORMANCE EVALUATION CRITERIA

To assess the prediction performance, the maximum absolute error (MAE), mean square error (MSE), RMSE and goodness-of-fit $R^2$ are considered as the evaluation criteria. MAE, MSE and RMSE evaluate the average prediction performance, of which the smaller value implies better prediction precision. By contrast, $R^2$, varying within [0, 1], evaluates the correctness of trained model, and the higher value (closer to 1) of $R^2$ indicates more similar prediction result, compared to the real attribution. These four criterions are formulated as:

$$\begin{cases} MAE = \max |y_i - \hat{y}_i| \\ MSE = \dfrac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2 \\ RMSE = \sqrt{\dfrac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2} \\ R^2 = 1 - \sum_{i=1}^{m} (y_i - \hat{y}_i)^2 / \sum_{i=1}^{m} (y_i - \bar{y}_i)^2 \end{cases} \quad (7)$$

where *n* represents the total sample number; $y_i$ and $\hat{y}_i$ are the real value and predicted value of target variable for the *i*th sample, respectively; and $\bar{y}_i$ represents the average value.

In the next step, a series of validations are conducted, followed by the detailed comparison and discussions.

## IV. RESULTS AND DISCUSSION

In this study, four cells' data are employed to validate the effectiveness of the proposed LSTM-RNN model for capacity prediction. The capacity prediction results under different conditions are discussed, including the influence of aging factors for model inputs, the comparisons of LSTM-RNN with traditional SVM, GPR and Elman NN, as well as the prediction results in terms of different cells' data for training.

### A. INFLUENCE WITH AGING FETURES AS MODEL INPUT ON THE PREDICTION PERFORMANCE

To analyze the influence on the LSTM-RNN model caused by the aging factors as model inputs, the historical capacity degradation data and extracted aging factors are respectively employed as the inputs for model training. When considering the historical capacity data as model inputs during the training process, the model output is the observed capacity value of the next cycle corresponding to the current input cycle. In contrast, the observed capacity value of current cycle is regarded as model output while taking the aging factors as
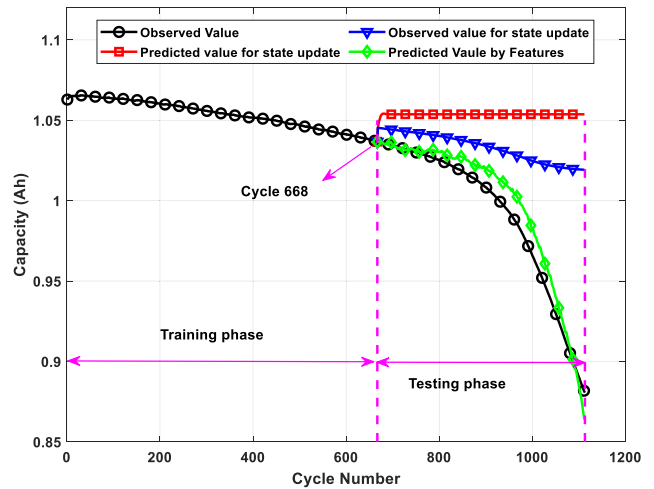


**FIGURE 13.** The capacity prediction results of Cell 1 with different model input and conditions for state updating of network.

model input. Therefore, the data length of prediction results has one cycle difference with different model input. To make the prediction results of different features as model input are consistent, the prediction with historical capacity data as model input starts from the last cycle of training set. In addition, when the LSTM-RNN model executes one- or multi-step forward prediction with historical capacity data as model input, it will obtain different predicted values with disparate variables for state update of the network. For comparison, the observed value and the predicted value of current cycle are respectively exploited to update the network state for next cycle's prediction. The prediction results of taking observed value and predicted value to update state are synchronously compared with the prediction results with the aging factors as model input.

Taking Cell 1 as an example, 60% of cycle life data is employed for model training, and the rest 40% is utilized for test. The predicted results and corresponding errors are shown in Figs. 13 and 14. We can find that when the predicted value is employed to update the network state, all the predicted results remain almost the same, indicating that the model cannot identify the degradation pattern in this case. When the observed value is employed to update the network state, the predicted results show a slight degradation trend in the global view but distinctly deviate from the observed capacity degradation trajectory. When the aging factors are taken as the model input, the capacity degradation trend can be well tracked by the LSTM-RNN model, and the maximum prediction error is less than 2%, as show in Fig. 14. Note that the battery degradation can be divided into two stages according to the capacity decline rate in this study. One stage is a linear degradation with a slower decline rate, e.g., the cycle range [1, 800], and the other stage is an exponential degradation with a faster decline rate, such as cycles 800 to 1100. The experimental results show that when the degradation rate of capacity is distinctly different in the early and later of cycle life period, the LSTM-RNN cannot identify the battery
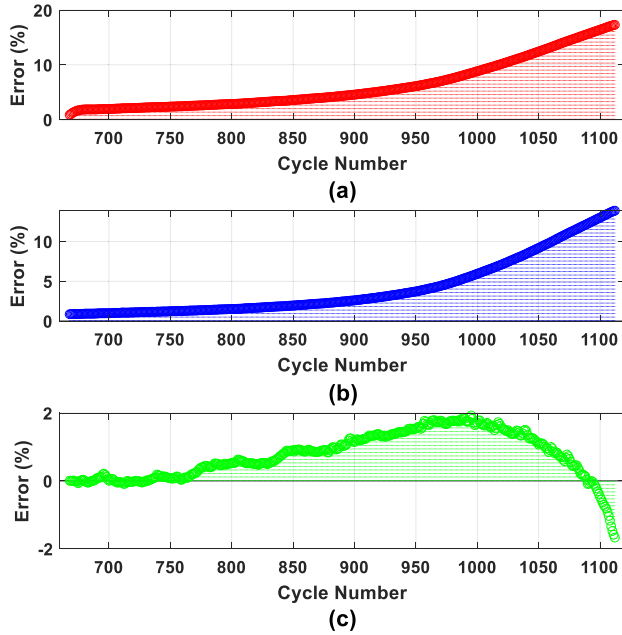
**FIGURE 14.** The capacity prediction errors of Cell 1. (a) The predicted errors with predicted value for state update; (b) the predicted errors with observed value for state update; (c) the predicted errors with aging factors as model input.
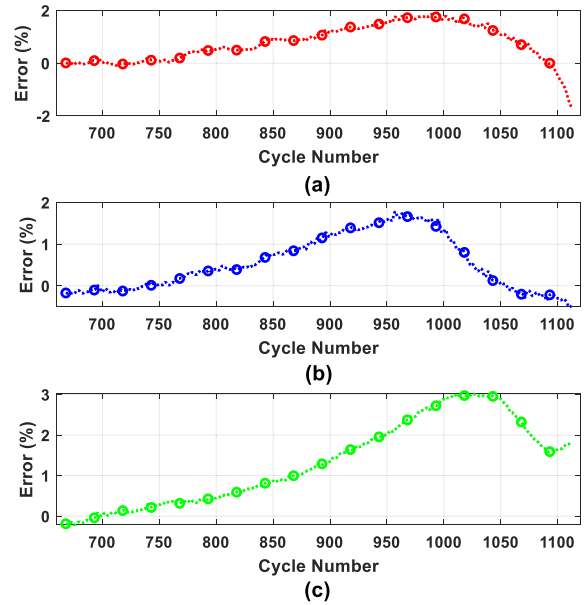


**FIGURE 15.** The capacity prediction of Cell 1 with different aging factors as model input.



**FIGURE 16.** The capacity prediction errors of Cell 1. (a) The predicted errors with $F_1$ and $F_2$ as model input; (b) The predicted errors with $F_1$, $F_2$ and $F_3$ as model input; (c) The predicted errors with only $F_3$ as model input.
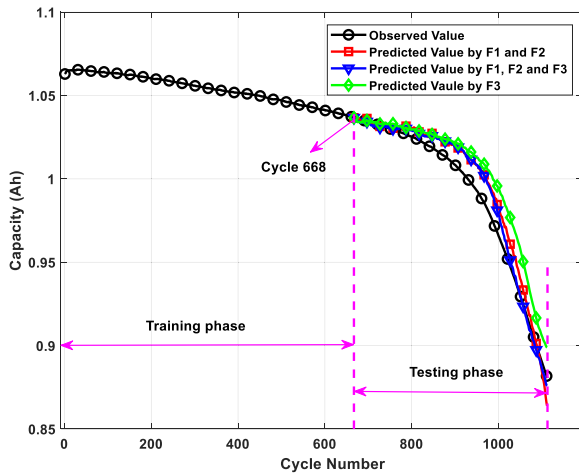
degradation pattern with the historical capacity data as the model input. However, as long as some effective aging factors such as the IR and temperature of battery can be extracted as the input for model training, the LSTM-RNN can predict the remaining capacity of battery with preferable accuracy and strong robustness.

To further analyze the influence of aging factors on the capacity prediction, the absolute value of DIC peak is extracted as one aging factor, namely, $F_3$. The prediction results of Cell 1 with different aging factors as the model input are shown in Figs. 15 and 16, respectively. As can be seen, all the prediction values with different aging factors as the model inputs can well track the capacity degradation

trajectory. As can be obviously seen from Fig. 16 (b), when $F_3$ is added as the model input, the maximum prediction error is 1.78%, which does not decrease much, compared with the maximum error with only $F_1$ and $F_2$ as the model inputs. Moreover, when the model input is only $F_3$, the maximum prediction error reach 3.02%, as shown in Fig. 16 (c). Note that the extraction of $F_3$ requires differential and interpolation calculation, significantly increasing the computation burden and the algorithm's complexity. Furthermore, constant current charging/discharging operations are difficult to encounter in practical applications. Compared with $F_1$ and $F_2$ that can be directly measured, the extraction of $F_3$ is more complex. To sum up, the subsequent discussion of capacity prediction in this study is based on only $F_1$ and $F_2$ as the model inputs hereinafter.

## B. COMPARISON OF PREDICTION RESULTS WITH DIFFERENT METHODS

To further evaluate the performance of LSTM-RNN model, the single GPR, SVM and Elman NN algorithms are respectively applied for the capacity prediction of Cell 2. For the sake of fair comparison, 60% of the cycle data (1 to 686 cycles) are utilized to train the model, and the remaining 40% data (687 to 1144 cycles) are employed to verify the precision. The aging factors are taken as the model inputs, and the predicted results and errors are shown in Figs. 17-18 and Table 2. It can be seen from Fig. 17 that the LSTM-RNN can precisely track the degradation trajectory of capacity in the whole test dataset and can achieve the preferable prediction accuracy. Although the other three methods can roughly reflect the variation trend of capacity, the prediction errors are far more than that of LSTM-RNN. From
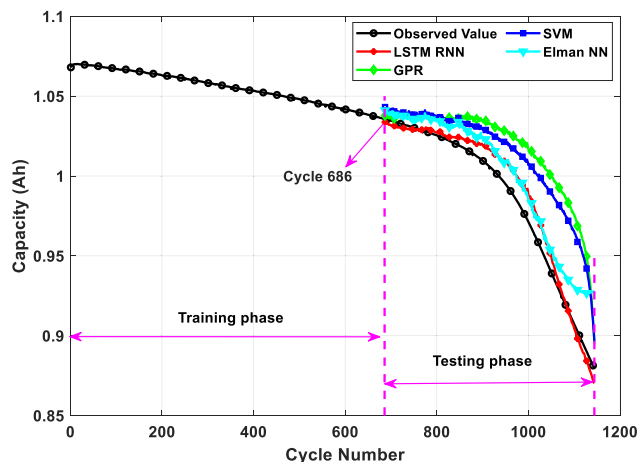
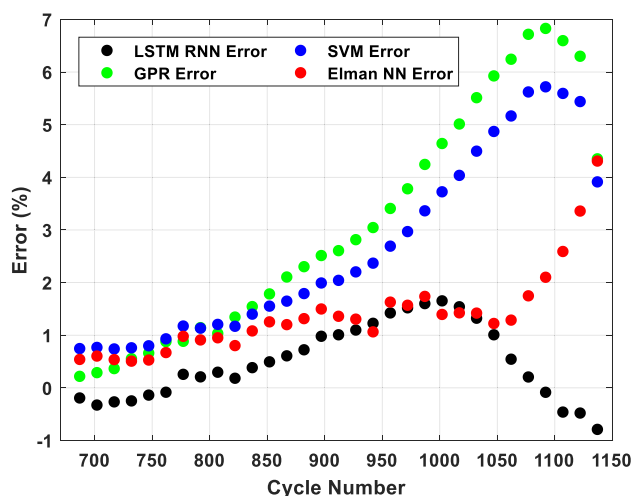**FIGURE 17.** The prediction results for Cell 2 with different methods.



**FIGURE 18.** The prediction errors for Cell 2 with different methods.

**TABLE 2.** Capacity prediction errors for cell 2 with different methods.

| Methods | Evaluation Criteria | | | |
|---------|---------------------|---|---|---|
| | MAE (%) | MSE ($10^{-5}$) | RMSE (%) | Average Time Cost (s) |
| LSTM-RNN | 1.96 | 7.81 | 0.88 | 56.46 |
| GPR | 6.85 | 140 | 3.81 | 153.56 |
| SVM | 5.86 | 100 | 3.17 | 193.55 |
| Elman NN | 4.81 | 25.8 | 1.61 | 32.93 |

**TABLE 3.** Capacity prediction errors for cells 5 to 7.

| Battery Number | Error Criterion | | | |
|----------------|-----------------|---|---|---|
| | *MAE* (%) | *MSE* ($10^{-4}$) | *RMSE* (%) | $R^2$ |
| Cell 5 | 2.04 | 1.29 | 1.14 | 0.9704 |
| Cell 6 | 1.44 | 5.24 | 0.72 | 0.9797 |
| Cell 7 | 2.54 | 2.22 | 1.49 | 0.9241 |

indicate that the GPR, SVM and Elman NN are not qualified for the time series prediction with large sample data and long-term dependence. To sum up, the proposed LSTM-RNN algorithm not only exhibits higher prediction accuracy and faster operation, but also shows more robustness in predicting capacity degradation with long-term dependence.

### C. CAPACITY PREDICTION WITH SINGLE BATTERY DATA

To further validate the performance of LSTM-RNN for capacity prediction, the experimental data of another three cells, i.e., Cells 5, 6, and 7, are analyzed. Similarly, 60% cycle data are employed to train the model for each single battery; in other words, the prediction of Cells 5 to 7 starts at cycle 617, 554 and 564, respectively. Note that the aging factors, which are exploited for the LSTM-RNN model input, are only the IR and average temperature. The prediction results and errors of Cells 5 to 7 are shown in Figs. 19 to 21 and listed in TABLE 3. As can be seen, the maximum prediction error of these three batteries is 2.54%, which is acceptable for capacity prediction. It can also be seen from Figs. 19 to 21 that the prediction error gradually increases with the increment of capacity degradation but declines quickly at the EOL of battery. The prediction error reveals that the LSTM-RNN can better predict the battery capacity in the whole cycle lifespan with the aging factors as the model inputs. In addition, the MSE and RMSE for these three cells are $1.29 \times 10^{-4}$, $5.24 \times 10^{-5}$, $2.22 \times 10^{-4}$, and 1.14%, 0.72%, 1.49%, respectively; manifesting that the proposed LSTM-RNN model leads to preferable prediction performance. The $R^2$ of Cells 5 to 7 are respectively 0.9704, 0.9797 and 0.9241, which illustrate the prediction value is holistically consistent with the real capacity. To sum up, by using the aging factors as the model inputs, the LSTM-RNN can predict the battery capacity with preferable accuracy.

### D. CAPACITY PREDICTION WITH MULTIPLE BATTERY DATA

To analyze the capability of degradation mechanism identification for the same type battery based on the LSTM-RNN

Table 2, we can find that the MAE of GPR, SVM and Elman NN are respectively 6.85%, 5.86% and 4.81%, in contrast to that of LSTM-RNN with 1.95%. Meanwhile, the MSE and RMSE of GPR, SVM and Elman NN are one order magnitude more than that of LSTM-RNN. From the perspective of prediction errors and accuracy, the LSTM-RNN algorithm outperforms GPR, SVM and Elman NN algorithms.

During the model training and optimization process, the consumption time for each iteration is recorded. The average consumption time of each method in the model training is calculated, as show in Table 2. As can be seen, the time cost of Elman NN is shortest, which is 32.93 s, followed by the LSTM-RNN, which lasts 56.46 s. Owing to the calculation of kernel functions and optimization of complex hyperparameters, the SVM and GPR respectively cost 193.55 s and 153.56 s for model training, which are much longer than that of LSTM-RNN. It is worth noting that the capacity degradation rate gradually increases with the cycling experiment, and the prediction errors of GPR, SVM and Elman NN also gradually increase, as show in Fig. 18. The results
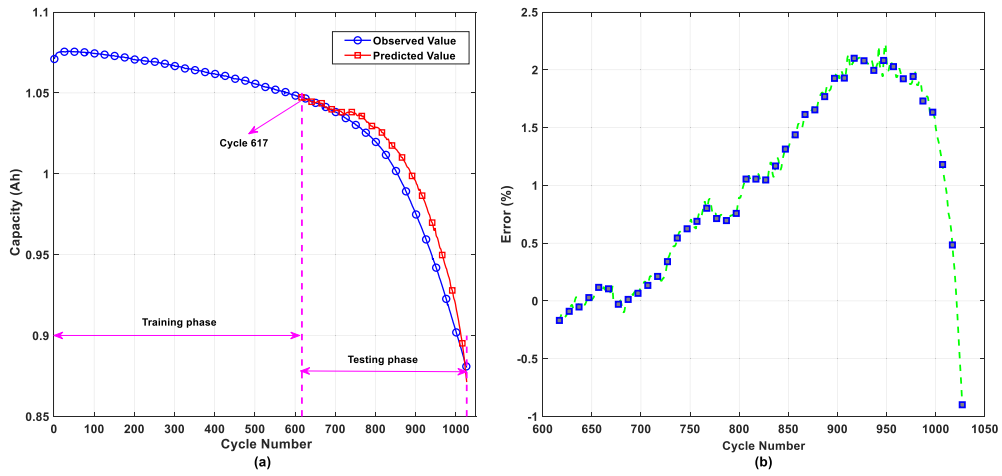
**FIGURE 19.** The capacity prediction of Cell 5. (a) The capacity prediction results; (b) The prediction error.
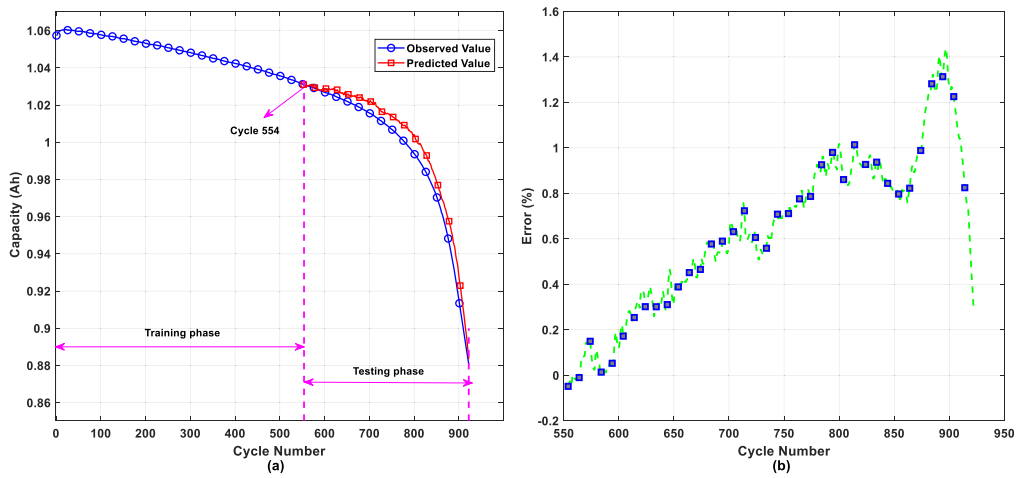


**FIGURE 20.** The capacity prediction of Cell 6. (a) The capacity prediction results; (b) The prediction error.
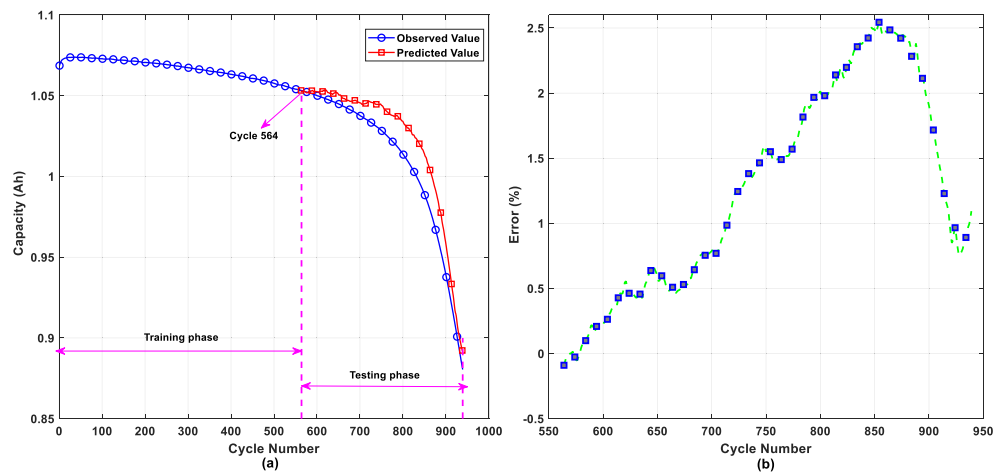


**FIGURE 21.** The capacity prediction of Cell 7. (a) The capacity prediction results; (b) The prediction error.

model, we employed the whole cycle life data of Cell 1 as the training data and the other cell's data for test. Fig. 22 and

Table 4 sketch the prediction results and corresponding errors. As shown in Fig. 22, the capacity degradation trajectories of
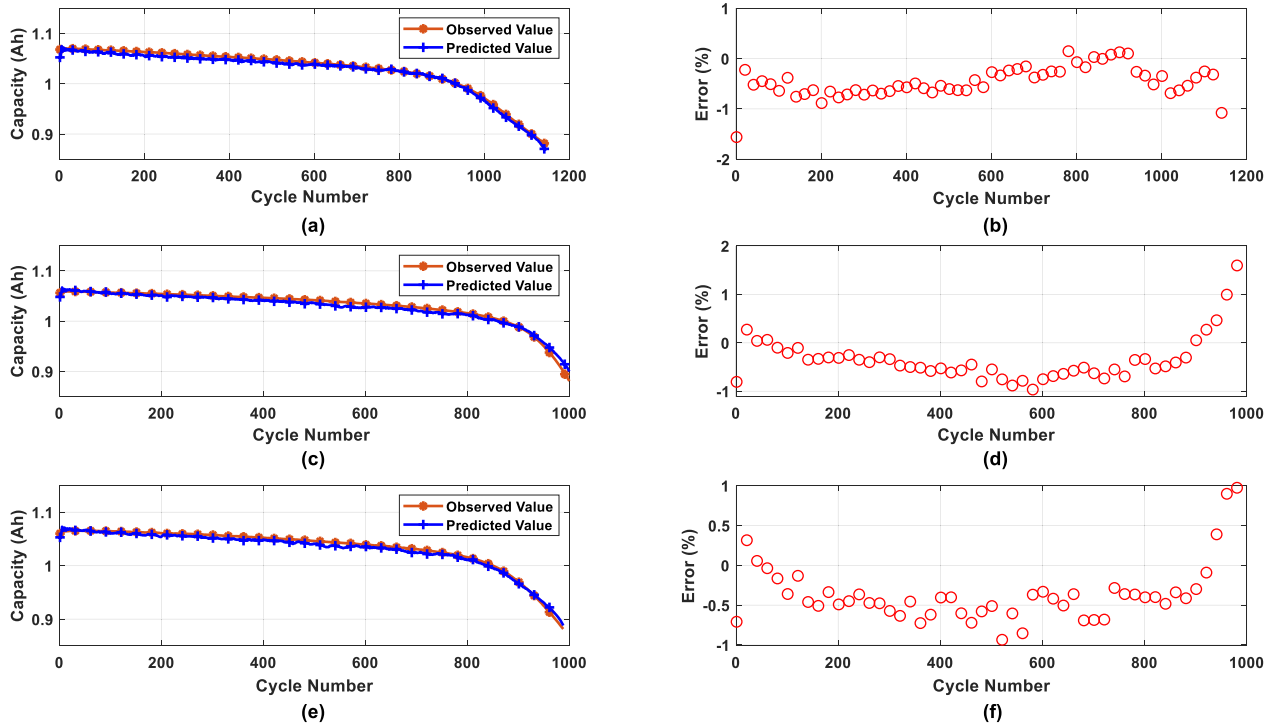
**FIGURE 22.** The capacity prediction results and error of Cells 2 to 4 with data of cell 1 for training. (a)-(b) the prediction results and error of Cell 2; (c)-(d) the prediction results and error of Cell 3; (e)-(f) the prediction results and error of Cell 4.

**TABLE 4.** Capacity prediction errors for cells 2 to 4.

| Training | Testing | Error Criterion | | | |
|---|---|---|---|---|---|
| | | MAE (%) | MSE ($10^{-5}$) | RMSE (%) | $R^2$ |
| Cell 1 | Cell 2 | 1.56 | 2.64 | 0.51 | 0.9872 |
| Cell 1 | Cell 4 | 2.00 | 3.61 | 0.60 | 0.9690 |
| Cell 1 | Cell 5 | 1.01 | 2.53 | 0.50 | 0.9835 |

Cells 2 to 4 are well tracked by the LSTM-RNN model. The MAE is 1.56%, 2.00% and 1.01%, respectively. From Table 4, we can find that the MSE and RMSE of Cell 3 estimation are the largest, i.e., $3.61 \times 10^{-5}$ and 0.60%; whereas the $R^2$ is least, which is 0.9690. It can be obviously seen from Fig. 2 that the IR value of Cell 3 is the most among the four cells when reaching its EOL, and its value is about 0.001 ohm larger than that of the other three cells. It can also be found from Fig. 4 that the average temperature of Cell 3 is the least at every cycle. The slight difference in the aging factors results in larger capacity prediction error of Cell 3 than those of Cells 2 and 4. Nevertheless, the prediction errors of Cells 2 and 4 are mostly less than 1%, except some individual points where the error is relative larger. The MSE and RMSE of Cells 2 and 4 are less than $2.70 \times 10^{-5}$ and 0.51%, which can be regarded as a preferable accuracy for capacity prediction. Moreover, the $R^2$ of Cells 2 and 4 are 0.9872 and 0.9835, demonstrating that the predicted values are highly consistent with the observed values. In summary, the experimental results manifest that even only the complete cycle life data of one cell are employed to train the model, the LSTM-RNN can still accurately predict the capacity of other batteries with the same type.

## E. THE VALIDATION OF CAPACITY PREDICTION BASED ON LOOCV

According to the LOOCV principle shown in Fig. 12, the data of four cells are randomly combined into one group, and thus they are divided into four data group, each of which contains a training dataset and a test dataset. In this study, the training dataset is assembled from three cells' data, and the remaining cell's data is utilized for test. This validation process is repeated for four times until each cell is employed for test in turn. Therefore, a 4-fold cross validation is performed for the LSTM-RNN model.

The validation results of capacity prediction are shown in Fig. 23, and the corresponding errors are illustrated in the Fig. 24 and Table 5. As can be seen from Fig. 24, the maximum prediction errors of Cells 1, 2 and 4 are lower than 2%, whereas that of cell 3 reaches 2.84%. The MSE, RMSE and $R^2$ of Cell 3 are $2.84 \times 10.5$, 0.50% and 0.9787, respectively. It is worth noting that the prediction error of Cell 3 is the largest among those of the four cells. This prediction results are in line with the previous conclusion that the slight difference in the aging factors can lead to larger prediction error, as drawn in Section 4.4. Compared with the prediction results of Cell 3 with only the data of Cell 1 for model training, when the data of Cells 1, 2 and 4 are employed for training, the prediction accuracy is not significantly improved, as shown in Fig. 24 (c) and Table 5. It can be therefore concluded that increasing the amount of training data cannot distinctly improve the prediction accuracy. Furthermore, the experimental results indicate that when the prediction model is fixed, the capacity prediction accuracy is
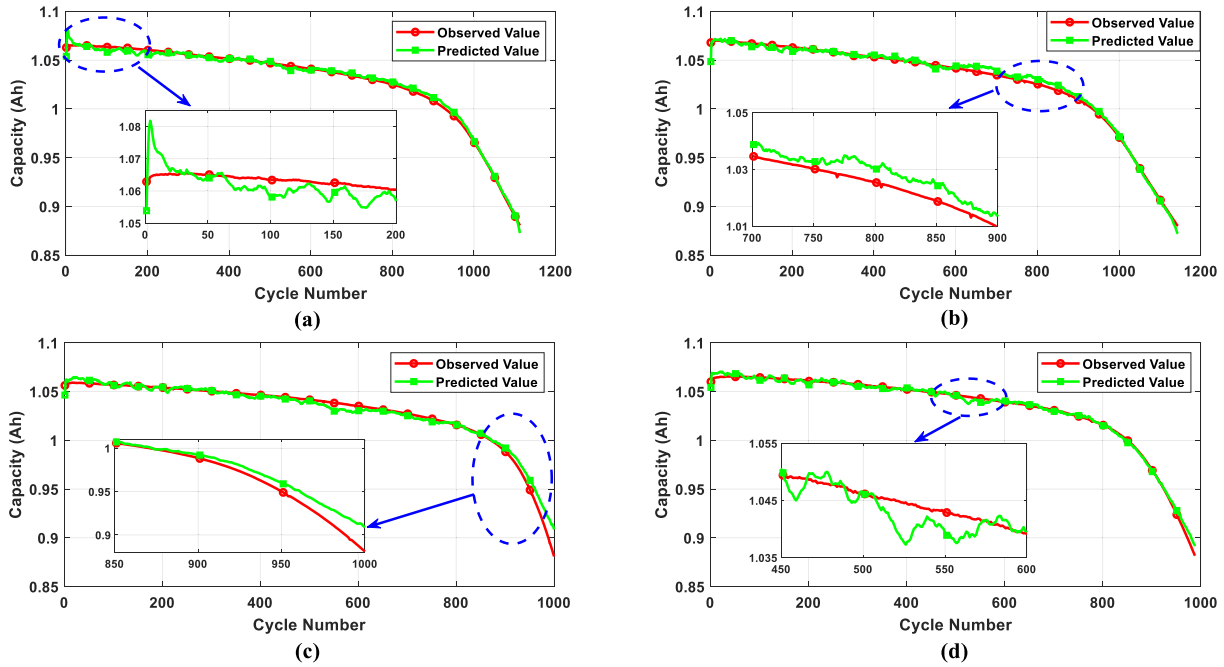
**FIGURE 23.** The capacity prediction results of Cells 1 to 4 based on LOOCV. (a) the prediction results of cell 1 with data of Cells 2, 3 and 4 for training; (b) the prediction results of Cell 2 with data of Cells 1, 3 and 4 for training; (c) the prediction results of Cell 3 with data of Cells 1, 2 and 4 for training; (d) the prediction results of Cell 4 with data of Cells 1, 2 and 3 for training.
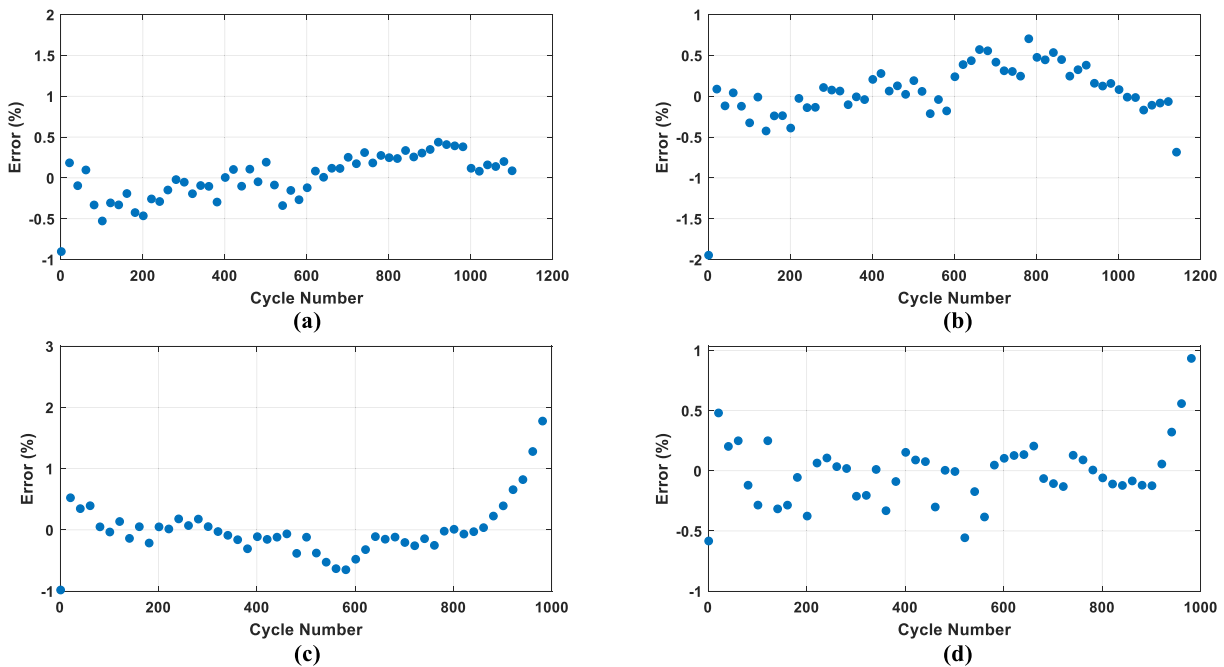


**FIGURE 24.** The capacity prediction error of cells 1 to 4 based on LOOCV. (a)-(d) the prediction errors of cells 1 to 4.

not much related to the amount of training data but depends on the effectiveness of the extracted aging factors. It can be seen from Fig. 23 that the LSTM-RNN model can accurately predict the global trend of capacity degradation, whereas the predicted values fluctuate in the vicinity of the observed values. In addition, the $R^2$ of Cells 1, 2, and 4 are 0.9957, 0.9955 and 0.9960, which are quite close to 1, highlighting

that the predicted values are very similar to the observed values. It can be noted that the capacity prediction for Cells 1 to 4 is attained based on different datasets for model training, and the maximum prediction error is less than 3%, highlighting that the proposed model is stable and reliable. To sum up, the experimental results manifest that when the effective aging factors are extracted for model training, the

**TABLE 5.** Capacity prediction errors for cells 2 to 4.

| Battery | Error Criterion | | | |
|---|---|---|---|---|
| Number | MAE (%) | MSE ($10^{-5}$) | RMSE (%) | $R^2$ |
| Cell 1 | 1.74 | 0.75 | 0.27 | 0.9957 |
| Cell 2 | 1.95 | 0.92 | 0.30 | 0.9955 |
| Cell 3 | 2.84 | 2.48 | 0.50 | 0.9787 |
| Cell 4 | 1.03 | 0.61 | 0.25 | 0.9960 |

LSTM-RNN model can precisely learn the degradation pattern of battery and predict the battery capacity with preferable accuracy.

## V. CONCLUSION

The key challenge of capacity prediction for lithium-ion batteries based on data-driven methods lies in effective extraction of key aging factors and accurate modeling of the long-term dependences of capacity degradation. In this paper, the LSTM-RNN algorithm is employed to construct the data driven-based capacity prediction for lithium-ion batteries. To improve the prediction performance of LSTM-RNN model, the Adam optimization algorithm is leveraged to find the optimal model parameters, and the dropout technique is exploited to prevent the network from overfitting. The reliability and robustness of LSTM-RNN for capacity prediction is validated based on the leave-one-out cross validation. The experimental results validate the LSTM-RNN model can well track the nonlinear capacity degradation trajectory. Meanwhile, even when only one battery data is employed for model training, the capacity prediction error of other cells is still less than 2%. Moreover, two conclusions can be drawn based on the leave-one-out cross validation. Firstly, when different training and test dataset are employed, the LSTM-RNN model can accurately predict the battery capacity with a maximum error of 2.84%, manifesting that the proposed method has preferable prediction accuracy and strong robustness. Secondly, when the model can learn the capacity degradation pattern in the whole lifespan of battery, increasing the amount of training data does not distinctly reduce the prediction error. The prediction accuracy mainly depends on the reliability and validity of the extracted aging factors. This work highlights the feasibility of applying the LSTM-RNN to predict capacity of lithium-ion batteries.

## REFERENCES

[1] Y. Shang, N. Cui, B. Duan, and C. Zhang, "Analysis and optimization of star-structured switched-capacitor equalizers for series-connected battery strings," *IEEE Trans. Power Electron.*, vol. 33, no. 11, pp. 9631–9646, Nov. 2018.
[2] Y. Zhang, R. Xiong, H. He, and M. Pecht, "Validation and verification of a hybrid method for remaining useful life prediction of lithium-ion batteries," *J. Cleaner Prod.*, vol. 212, pp. 240–249, Mar. 2019.
[3] X. Xu, C. Yu, S. Tang, X. Sun, X. Si, and L. Wu, "State-of-health estimation for lithium-ion batteries based on Wiener process with modeling the relaxation effect," *IEEE Access*, vol. 7, pp. 105186–105201, 2019.
[4] X. Shu, G. Li, J. Shen, W. Yan, Z. Chen, and Y. Liu, "An adaptive fusion estimation algorithm for state of charge of lithium-ion batteries considering wide operating temperature and degradation," *J. Power Sources*, vol. 462, Jun. 2020, Art. no. 228132.
[5] H. Chaoui and H. Gualous, "Adaptive state of charge estimation of lithium-ion batteries with parameter and thermal uncertainties," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 2, pp. 752–759, Jun. 2016.
[6] Z. Chen, Q. Xue, R. Xiao, Y. Liu, and J. Shen, "State of health estimation for lithium-ion batteries based on fusion of autoregressive moving average model and elman neural network," *IEEE Access*, vol. 7, pp. 102662–102678, 2019.
[7] D. Zhou, H. Yin, W. Xie, P. Fu, and W. Lu, "Research on online capacity estimation of power battery based on EKF-GPR model," *J. Chem.*, vol. 2019, pp. 1–9, Jan. 2019.
[8] D. Zhang, S. Dey, H. E. Perez, and S. J. Moura, "Real-time capacity estimation of lithium-ion batteries utilizing thermal dynamics," *IEEE Trans. Control Syst. Technol.*, vol. 28, no. 3, pp. 992–1000, May 2020.
[9] J. Li, C. Lyu, L. Wang, L. Zhang, and C. Li, "Remaining capacity estimation of li-ion batteries based on temperature sample entropy and particle filter," *J. Power Sources*, vol. 268, pp. 895–903, Dec. 2014.
[10] 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, Jan. 2019.
[11] Y. Liu, C. Zhang, J. Jiang, Y. Jiang, L. Zhang, and W. Zhang, "Capacity estimation of serial lithium-ion battery pack using dynamic time warping algorithm," *IEEE Access*, vol. 7, pp. 174687–174698, 2019.
[12] Q. Yu, R. Xiong, R. Yang, and M. G. Pecht, "Online capacity estimation for lithium-ion batteries through joint estimation method," *Appl. Energy*, vol. 255, Dec. 2019, Art. no. 113817.
[13] 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, Dec. 2018.
[14] Z. Chen, J. P. Xiao, X. Shu, S. Q. Shen, J. W. Shen, and Y. G. Liu, "Model-based adaptive joint estimation of the state of charge and capacity for Lithium–Ion batteries in their entire lifespan," *Energies*, vol. 13, no. 6, p. 15, Mar. 2020.
[15] C. Chen, R. Xiong, and W. Shen, "A lithium-ion battery-in-the-loop approach to test and validate multiscale dual h infinity filters for state-of-charge and capacity estimation," *IEEE Trans. Power Electron.*, vol. 33, no. 1, pp. 332–342, Jan. 2018.
[16] Z. Wei, J. Zhao, D. Ji, and K. J. Tseng, "A multi-timescale estimator for battery state of charge and capacity dual estimation based on an online identified model," *Appl. Energy*, vol. 204, pp. 1264–1274, Oct. 2017.
[17] C. Yang, X. Wang, Q. Fang, H. Dai, Y. Cao, and X. Wei, "An online SOC and capacity estimation method for aged lithium-ion battery pack considering cell inconsistency," *J. Energy Storage*, vol. 29, Jun. 2020, Art. no. 101250.
[18] Y. Deng, H. Ying, H. Zhu, K. Wei, J. Chen, F. Zhang, and G. Liao, "Feature parameter extraction and intelligent estimation of the state-of-health of lithium-ion batteries," *Energy*, vol. 176, pp. 91–102, Jun. 2019.
[19] 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.
[20] D. Yang, X. Zhang, R. Pan, Y. Wang, and Z. Chen, "A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve," *J. Power Sources*, vol. 384, pp. 387–395, Apr. 2018.
[21] M. A. Patil, P. Tagade, K. S. Hariharan, S. M. Kolake, T. Song, T. Yeo, and S. Doo, "A novel multistage support vector machine based approach for li ion battery remaining useful life estimation," *Appl. Energy*, vol. 159, pp. 285–297, Dec. 2015.
[22] Z. Chen, M. Sun, X. Shu, R. Xiao, and J. Shen, "Online state of health estimation for lithium-ion batteries based on support vector machine," *Appl. Sci.*, vol. 8, no. 6, p. 925, Jun. 2018.
[23] Z. Wang, J. Ma, and L. Zhang, "State-of-health estimation for lithium-ion batteries based on the multi-island genetic algorithm and the Gaussian process regression," *IEEE Access*, vol. 5, pp. 21286–21295, 2017.
[24] Y. Wu, Q. Xue, J. Shen, Z. Lei, Z. Chen, and Y. Liu, "State of health estimation for lithium-ion batteries based on healthy features and long short-term memory," *IEEE Access*, vol. 8, pp. 28533–28547, 2020.
[25] C. S. Wang, N. Y. Lu, S. L. Wang, Y. H. Cheng, and B. Jiang, "Dynamic long short-term memory neural-network-based indirect remaining-useful-life prognosis for satellite lithium-ion battery," *Appl. Sci.-Basel*, vol. 8, no. 11, Nov. 2018, Art. no. 2078.

[26] S. Yoon, H. Yun, Y. Kim, G.-T. Park, and K. Jung, "Efficient transfer learning schemes for personalized language modeling using recurrent neural network," in *Proc. Workshops 31st AAAI Conf. Artif. Intell.*, 2017, pp. 457–463.

[27] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, "Show and tell: A neural image caption generator," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3156–3164.

[28] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," 2015, *arXiv:1506.00019*. [Online]. Available: http://arxiv.org/abs/1506.00019

[29] Y. Zhou, Y. Huang, J. Pang, and K. Wang, "Remaining useful life prediction for supercapacitor based on long short-term memory neural network," *J. Power Sources*, vol. 440, Nov. 2019, Art. no. 227149.

[30] 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 Trans. Veh. Technol.*, vol. 67, no. 7, pp. 5695–5705, Jul. 2018.

[31] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggedakis, M. Z. Bazant, S. J. Harris, W. C. Chueh, and R. D. Braatz, "Data-driven prediction of battery cycle life before capacity degradation," *Nature Energy*, vol. 4, no. 5, pp. 383–391, May 2019.

[32] X. Han, L. Lu, Y. Zheng, X. Feng, Z. Li, J. Li, and M. Ouyang, "A review on the key issues of the lithium ion battery degradation among the whole life cycle," *eTransportation*, vol. 1, Aug. 2019, Art. no. 100005.

[33] R. Zhao, J. Liu, and J. Gu, "The effects of electrode thickness on the electrochemical and thermal characteristics of lithium ion battery," *Appl. Energy*, vol. 139, pp. 220–229, Feb. 2015.

[34] X. Li, Z. Wang, L. Zhang, C. Zou, and D. D. Dorrell, "State-of-health estimation for li-ion batteries by combing the incremental capacity analysis method with grey relational analysis," *J. Power Sources*, vols. 410–411, pp. 106–114, Jan. 2019.

[35] T. Ergen and S. Serdar Kozat, "Efficient online learning algorithms based on LSTM neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 8, pp. 3772–3783, Aug. 2018.

[36] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980

[37] M. C. Mukkamala and M. Hein, "Variants of rmsprop and adagrad with logarithmic regret bounds," in *Proc. 34th Int. Conf. Mach. Learn.*, 2017, pp. 2545–2553.

[38] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," *J. Mach. Learn. Res.*, vol. 12, no. 7, pp. 257–269, 2010.

[39] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, Jun. 2014.

[40] C. Qi, J. Diao, and L. Qiu, "On estimating model in feature selection with cross-validation," *IEEE Access*, vol. 7, pp. 33454–33463, 2019.

**QIAO XUE** received the B.S. degrees in traffic engineering from the Kunming University of Science and Technology, Kunming, China, in 2018, where he is currently pursuing the M.S. degrees in transportation engineering.

His research interests include the state of health monitoring for battery and battery management system.

**YITAO WU** received the B.S. and M.S. degrees in vehicle engineering from the Kunming University of Science and Technology, Kunming, China, in 2017 and 2020, respectively. He is currently pursuing the Ph.D. degree in vehicle engineering with Chongqing University, Chongqing, China.

His research interests include the energy management system of electric vehicle and optimal control.

**SHIQUAN SHEN** received the Ph.D. degree in power machinery and engineering from Tianjin University, Tianjin, China, in 2020. He was a joint Ph.D. student with the Argonne National Laboratory, Lemont, IL, USA, from 2018 to 2019.

He is currently a Lecturer with the Faculty of Transportation Engineering, Kunming University of Science and Technology, Kunming, Yunnan, China. His research interests include energy management of hybrid electric vehicles, model predictive control, and multi-phase flow.

**YUANJIAN ZHANG** (Member, IEEE) received the M.S. degree in automotive engineering from Coventry University, U.K., in 2013, and the Ph.D. degree in automotive engineering from Jilin University, China, in 2018.

In 2018, he joined the University of Surrey, Guildford, U.K., as a Research Fellow in advanced vehicle control. He currently works with Sir William Wright Technology Centre, Queen's University Belfast, U.K. His current research interests include advanced control on electric vehicle powertrains, vehicle-environment-driver cooperative control, vehicle dynamic control, and intelligent control for driving assist system.

**ZHENG CHEN** (Senior Member, IEEE) received the B.S. and M.S. degrees in electrical engineering and the Ph.D. degree in control science engineering from Northwestern Polytechnical University, Xi'an, China, in 2004, 2007, and 2012, respectively.

He was a Postdoctoral Fellow and a Research Scholar with the University of Michigan, Dearborn, MI, USA, from 2008 to 2014. He is currently a Professor with the Faculty of Transportation Engineering, Kunming University of Science and Technology, Kunming, Yunnan, China, and also a Marie-Curie Research Fellow with the School of Engineering and Materials Science, Queen Mary University of London, London, U.K. He has conducted over 30 projects and has published over 100 peer-reviewed journal papers and conference proceedings. His research interests include battery management system, battery status estimation, and energy management of hybrid electric vehicles. He is a Fellow of the Institution of Engineering and Technology. He was a recipient of the Yunnan Oversea High Talent Project, China, and secured second place in the IEEE VTS Motor Vehicles Challenge, in 2017 and 2018, respectively.

**JIANGWEI SHEN** received the B.S. and M.S. degrees in traffic engineering, and power machine and engineering from the Kunming University of Science and Technology, Kunming, China, in 2008 and 2011, respectively. He is currently pursuing the Ph.D. degree in automobile engineering with the Kunming University of Science and Technology, Kunming, China.

He is also an Experimentalist with the Kunming University of Science and Technology. His research interests include the battery management system and energy management of hybrid electric vehicles.

● ● ●