Energy Storage Materials 50 (2022) 139–151

Contents lists available at ScienceDirect

Energy Storage Materials

journal homepage: www.elsevier.com/locate/ensm

Battery degradation prediction against uncertain future conditions with recurrent neural network enabled deep learning

Jiahuan Lu [a], Rui Xiong [a,*], Jinpeng Tian [a], Chenxu Wang [a], Chia-Wei Hsu [c], Nien-Ti Tsou [c], Fengchun Sun [a], Ju Li [b,**]

[a] Department of Vehicle Engineering, School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China
[b] Department of Nuclear Science and Engineering and Department of Materials Science and Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA.
[c] Department of Materials Science and Engineering, National Chiao Tung University, Taiwan.

A B S T R A C T

Accurate degradation trajectory and future life are the key information of a new generation of intelligent battery and electrochemical energy storage systems. It is very challenging to obtain accurate predictions against uncertain application conditions by using only a few known historical data. In this article, we extend the widely studied remaining useful life (RUL) prediction to the prediction of charge and discharge capacity trajectories under both fixed and random future operating conditions. This is achieved by developing a general deep learning framework cored by recurrent neural network (RNN) which integrates future current plan and few early capacity-voltage data as inputs. As a case study, we have experimented with 77 commercial batteries cycled under fixed and random operating conditions. We demonstrate that the median root mean square error (RMSE) of prediction can be within 2.4% for NMC/graphite batteries and 2.3% for LFP/graphite batteries by using 3.8% of the whole life data only. Compared with the existing methods, the proposed framework predicts more accurately and has a very balanced performance for both fixed and random future conditions. This work highlights the promise of actively forecasting the future of batteries based on RNN.

1. Introduction

1.1. Literature review

Lithium-ion batteries (LIB) have been widely applied in a multitude of applications such as electric vehicles (EVs) [1], portable electronics [2], and energy storage stations [3]. The key metric for battery performance is the degradation of battery life caused by many charging and discharging events. In this process, the anode, cathode, electrolyte, and other components of a battery suffer from gradual degradation, leading to capacity and power loss [4,5].

Battery capacity loss is a widely accepted metric of battery life degradation, and it strongly affects the endurance of devices powered by batteries [6], such as the driving range of EVs [7]. Generally, once the battery capacity degrades to a certain threshold, i.e., the so-called end of life (EOL), the battery is no longer considered adequate to meet the requirements of the device and must be replaced [8]. Predicting the degradation of battery life plays a critical role in designing batteries and their management policies, scheduling battery maintenance, as well as screening batteries for pack manufacturing.

Current battery cell life prediction methods include the end-to-end prediction methods and the trajectory prediction methods. The end-to-end methods predict the remaining useful life (RUL) by directly mapping features to battery life. For example, Zhang et al. [9] mapped the collected entire impedance spectra to battery RUL based on Gaussian process regression (GPR). Zhang et al. [10] captured some health indicators highly related to RUL from the time-voltage curve of the batteries and built GPR-based mapping among the RUL and health indicators to realize RUL prediction. Severson et al. [11] extracted several physical features from the operating current, voltage, and temperature data of the first 100 cycles, and established the relationship between these features and the battery life based on regularized regression. The results demonstrated a prediction error of less than 9.1%. Compared with conventional machine learning methods, using deep learning to predict RUL can help establish high-dimensional mapping without feature extraction engineering. Hong et al. [12] utilized a temporal convolutional neural network (CNN) to directly learn the battery raw data (current, voltage, and temperature) and predicted

* Corresponding author at: Department of Vehicle Engineering, School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China.
** Corresponding author at: Department of Nuclear Science and Engineering and Department of Materials Science and Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA.
E-mail addresses: rxiong@bit.edu.cn (R. Xiong), liju@mit.edu (J. Li).

https://doi.org/10.1016/j.ensm.2022.05.007
Received 6 January 2022; Received in revised form 28 April 2022; Accepted 2 May 2022
Available online 4 May 2022

RUL and the corresponding uncertainty based on the raw data collected within 500–10000 s. The main advantage of the end-to-end prediction method is the few requirements of the historical cycle data to implement, benefiting the early prediction of battery life. However, an evident shortcoming is that only a scalar RUL is obtained while the capacity trajectory which comprehensively portrays the aging process is neglected.

The trajectory prediction methods extrapolate the capacity degradation trajectory to the EOL to determine the RUL. This is usually achieved by resorting to a capacity trajectory model. Simple mathematical functions, such as the linear function [13], the polynomial function [14], and the exponential function [15], can be employed as a trajectory model for battery life prediction. Besides, by treating the degradation trajectory of capacity as a time series, some recursive models can also be applied as a trajectory model, such as the autoregressive integrated moving average (ARIMA) model [16,17], GPR model [18], long short-term memory (LSTM) recurrent neural network (RNN) [19], etc. A key benefit of such methods is that they predict the future trajectories, not just a scalar RUL. The capacity trajectory contains more practically useful information. For instance, two batteries may have identical RUL but a different Ah throughput in remaining life [20]. Thus, trajectory prediction methods are more comprehensive to support battery management. However, limited by the parameter identification of the trajectory model, these methods often need a large amount of historical data for training. Studies [19,21] revealed that at least 15–20% of the whole-life cyclic data of a battery is needed for rough prediction of capacity trajectory, which hardly satisfies the demand for early degradation prediction.

More importantly, the existing prediction methods primarily focus on a very specific use scenario where batteries serve under a fixed operating policy of charge and discharge throughout their whole life. For example, two battery degradation datasets from the Center for Advanced Life Cycle Engineering (CALCE) [14,22] and NASA Prognostic Center of Excellence (PCOE) [23] are widely employed to develop life prediction methods, but the batteries in these datasets have almost all experimented under a fixed charge and discharge policy (charged and discharged at constant current) until the EOL. A recent dataset from Stanford-MIT [11,12], which covers 124 commercial 18650-batteries, is relatively diversified, as 72 different charge policies are utilized. However, the operating policy of *each* battery in this dataset is fixed throughout its entire life. This does not fit the practical experience of a typical battery cell in real-life cell phones, electric vehicles or grid-scale energy storage systems, as it may experience multiple user habit regimes. Although NASA PCOE later presents a randomized battery usage dataset including 24 batteries cycled under random walk conditions, they are almost treated as a fixed policy dataset in the current prediction methods [24–26]. Practically, uncertain future operating conditions are quite common, and experimental results have already revealed that this greatly impacts battery capacity degradation trajectory [27].

1.2. Gap analysis and contributions

The major weaknesses of the existing methods can be summarized as follows:

(1) The existing methods are unable to deal with the future varied operating conditions. Most of the prediction methods are developed on the basis of battery cyclic datasets serving under a lifelong fixed operating condition.
(2) The existing methods can hardly predict the capacity degradation trajectory at an early stage. The end-to-end prediction method can achieve early prediction, but it only obtains a scalar RUL, while the trajectory prediction method has conventionally required large numbers of cycle data for accurate predictions.

To resolve these problems, this paper proposes a deep learning-enabled framework to predict the future degradation trajectories of battery charge and discharge capacities under uncertain future service conditions. The main contributions of this paper are:

(1) A large battery degradation dataset covering the fixed and random operating conditions has been experimentally acquired for 77 commercial lithium-ion batteries, all cycled to respective EOL. To our knowledge, this is the first battery degradation dataset that accommodates greatly varying working loads for each cell. We are providing this dataset (2.26 Gigabytes in total) at https://data.mendeley.com/datasets/kw34hhw7xg/2 [28]. Open datasets like these will be essential for the battery community to develop and validate aging prediction methods.
(2) A general RNN-enabled deep learning framework of long-term degradation trajectory prediction that can handle both fixed and varied operating conditions is developed. As an example, the proposed framework cored by gated recurrent unit (GRU) can achieve a median root mean square error (RMSE) of 2.4% for predicting both charge and discharge capacities under the random policies of $LiCoO_2$ $+LiNi_{0.5}Co_{0.2}Mn_{0.3}O_2$ (NMC)/graphite batteries using data of one cycle only. Besides, a median RMSE of 2.3% can be obtained for lithium iron phosphate (LFP)/graphite batteries which survived thousands of cycles using 3.8% of the cycle data only.

1.3. Article organization

The rest of the article is organized as follows. Section 2 introduces the battery experiments and developed dataset. Section 3 describes the methodology of battery degradation prediction in detail. Validation and discussions of the proposed method are elaborated in Section 4. Limitations and outlooks are given in Section 5. Conclusions are drawn in Section 6.

2. Data generation

A type of nominally identical high-energy 18650 lithium-ion batteries manufactured by LISHEN is employed as the experimental subject, which is composed of a positive electrode of $LiCoO_2$ $+LiNi_{0.5}Co_{0.2}Mn_{0.3}O_2$ and a negative electrode of graphite. The rated capacity is 2.4 Ah, the nominal voltage is 3.7 V, and the lower and upper cut-off voltages are 3.0 V and 4.2 V, respectively. Note that even though they are brought as new, they have already undergone formation cycling and there are certainly also intrinsic variabilities due to uncertainties in the material batch and manufacturing process. In our cycling, the cut-off current is 0.048 A. 77 batteries are cycled in our degradation tests. All tests are conducted at 25 ℃, and the batteries operate within the specified lower and upper voltage limits. The experimental data-acquisition platform is detailed in Supplementary Fig. 1 and Supplementary Note 1.

In the first stage, 20 preliminary cycles are applied for simulating the primary battery application to explore the early prediction of battery life. Each preliminary cycle consists of 0.5 C constant-current-constant-voltage (CCCV) charging and 2 C constant current discharging. In the second stage, to investigate the degradation characteristics under different operating policies, the 77 batteries are divided into Group I and Group II for further cyclic degradation experiments. Group I consists of 22 batteries, which are used to perform the cyclic degradation experiments under a specified fixed policy of charge current (1C, 2C, or 3C) and discharge current (1C, 2C, or 3C). Group II consists of 55 batteries, which undergo cyclic degradation experiments under varied operating conditions. Specifically, each operating condition includes a charge policy at a random constant charge current (changes randomly every 5 cycles) that obeys a uniform distribution among three discrete choices of 1C, 2C, and 3C, and a discharge condition at a specified discharge current of 3C.

The diagram of the experimental design of the batteries is described

in Fig. 1. It can be noted the batteries do not show evident degradation at the first stage, while the trajectory significantly varies for different charge policies at the following stage. In this regard, conventional studies [11,12,14,22,23] which assume a fixed operating condition over battery life will not work. The arrangements and results of the cyclic degradation experiments are detailed in Supplementary Note 2.

3. Methodology

3.1. Framework of prediction

The RNN-enabled deep learning framework of battery degradation prediction is described in Fig. 2. It consists of four procedures: the input matrix, the RNN layer (the core layer), the fully connected (FC) layer, and the output layer.

The input matrix contains four components filled with different colors. The first and second components in blue and orange (component I & II) are the current plans designed for users to customize the future operating conditions of the battery, which can be formulated as:

$$\begin{bmatrix} I_1^{\mathrm{c}} & I_2^{\mathrm{c}} & \cdots & I_n^{\mathrm{c}} \\ I_1^{\mathrm{d}} & I_2^{\mathrm{d}} & \cdots & I_n^{\mathrm{d}} \end{bmatrix} \tag{1}$$

where I denotes the expectation of planning operation current at each future cycle, the subscript 1, 2, …, n represents the number of cycles that the battery is planned to work, the superscripts "c" and "d" represent charge and discharge, respectively. The joint input of these two policies is to simulate the degradation caused by charging and discharging. On the other hand, we expect the proposed framework can simultaneously predict both charge and discharge capacity trajectories, not just either one as done in most literature [29]. The third component filled with grey (component III) is the early cycle set, which is a capacity-voltage (Q-V) matrix derived from the charging data of initial cycles. Therefore, it can provide historical degradation data of a given battery to the network. Such data can be easily obtained by cycling batteries a few times, such as tests during battery development or screening after manufacturing, thereby it does not give rise to many testing efforts. The capacity-voltage (Q-V) matrix can be mathematically expressed as:

$$\begin{bmatrix} q_1^{v_1} & q_2^{v_1} & \cdots & q_m^{v_1} \\ q_1^{v_2} & q_2^{v_2} & \cdots & q_m^{v_2} \\ \vdots & \vdots & & \vdots \\ q_1^{v_l} & q_2^{v_l} & \cdots & q_m^{v_l} \end{bmatrix} \tag{2}$$

where superscript (v_1, v_2, …, v_l) represents discretized voltages equally distributed over the voltage range of 3.0–4.2 V, l is the number of the grid points and is set to 120 in this work. The subscript (1, 2, …, m) represents the number of input cycles. q is the cumulative charge capacity, which is calculated by:

$$q_j^{v_k} = \int_{t(v_1)}^{t(v_k)} I_j^{\mathrm{c}} dt, k = 1, 2, ..., l, j = 1, 2, ..., m \tag{3}$$

As n is generally larger than m, the general sequence padding [30] is employed to deal with the fourth component filled with white (component IV) to ensure the consistency of the length of the input series. The padding matrix $\mathbf{P}$ can be formulated as:

$$\mathbf{P} = \left(p_{ij}\right)_{l\times(n-m)}, i = 1, 2, ..., l, j = 1, 2, ..., (n-m) \tag{4}$$

where the padding value p is set to 0 in the present work.

A deep learning network comprising RNN and FC layers in series for prediction is deployed after the input matrix. It has been widely reported that battery capacity degradation is tightly related to the operating history [31]. As a powerful tool for processing time series, the RNN layer is designed as a core of the proposed framework to learn the time dependency of battery degradation, and its output is treated as the input of

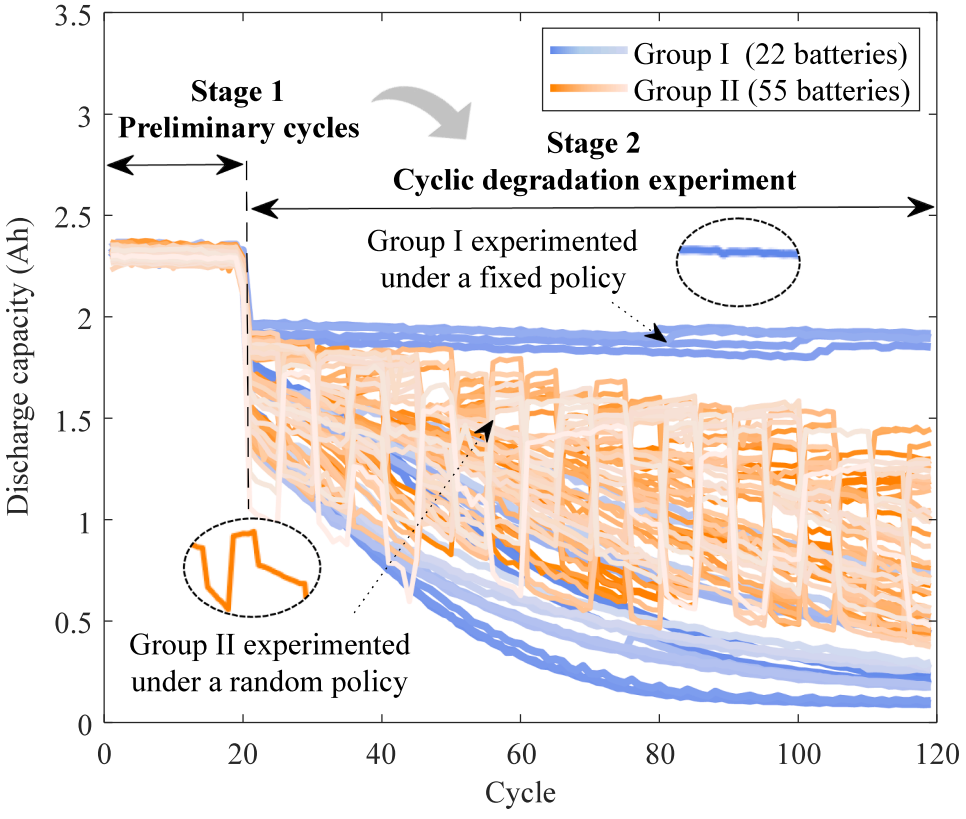

Fig. 1. Diagram of the experimental design of the batteries.

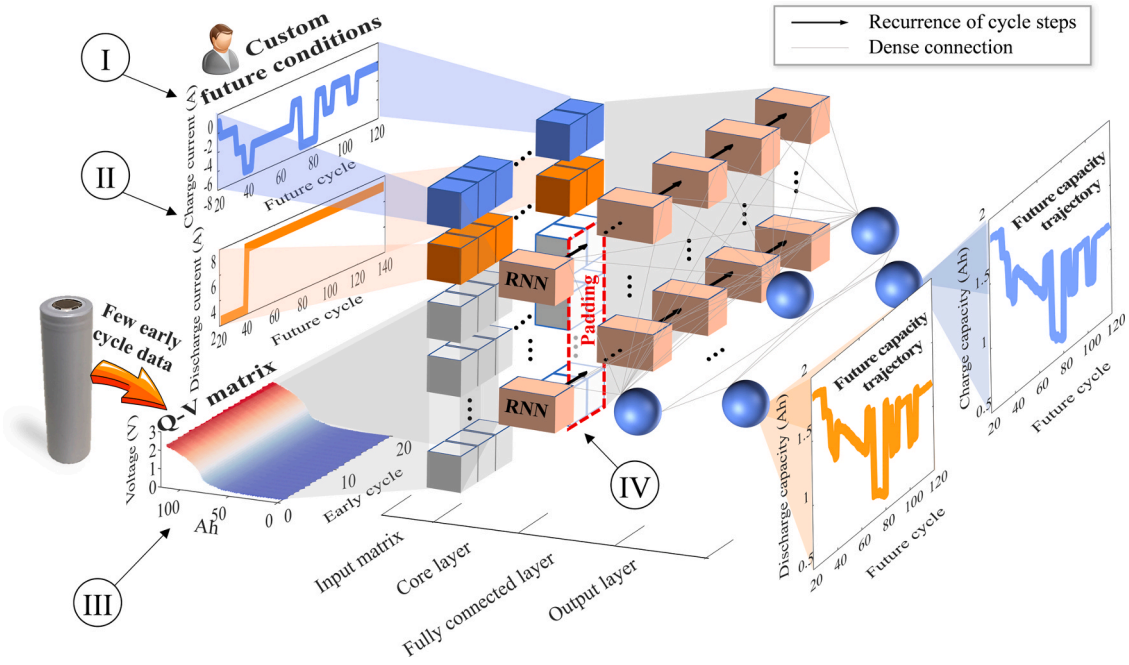

Fig. 2. RNN-enabled deep learning framework of battery life prediction considering varied operating conditions.

the fully connected layer.

The final output layer is composed of two neurons to merge the output of the last deep learning network into the predicted trajectories of both charge and discharge capacity.

As mentioned earlier, the proposed framework enables simultaneous prediction of charge and discharge capacity degradation trajectories. More importantly, the proposed framework can predict future capacity trajectories under custom future operating conditions as it considers not only historical data but also future operating conditions. This offers an opportunity to intelligently plan future operations with consideration of battery degradation [32].

Herein, the proposed framework is evaluated against both fixed and random operating conditions. About 30% of the batteries from Group I and Group II are selected as the training set (23 batteries), respectively. 7 batteries are selected from Group I as the training set to cover every possible combination of the operating policy. 16 batteries are randomly chosen from Group II as the training set. The rest 54 batteries are utilized as the validation set.

3.2. Gated recurrent unit dominated-network

The GRU network, an advanced RNN that is more efficient than the LSTM [33], is employed as an instance to dominate the core of the proposed framework. The GRU cell is the basic unit of a GRU network and is depicted in Fig. S4. Here, $\mathbf{x}_t$ is the matrix of the input series at time step t, $\mathbf{h}_{t-1}$, and $\mathbf{h}_t$ are the matrices of the output series at time step t-1 and t, respectively. r_t is the reset gate for controlling the level of state reset, z_t is the update gate for controlling the level of state update, $\hat{h}_t$ is the candidate state for controlling the level of update added to the hidden state. σ_g and σ_s represent the gate and state activation functions, respectively, which can be written as:

$$\begin{cases} \sigma_g(x) = \dfrac{1}{1+e^{-x}} \\ \sigma_s(x) = \tanh(x) \end{cases} \tag{5}$$

$\boldsymbol{W}$, $\boldsymbol{R}$, and $\boldsymbol{b}$ represent the input weights, recurrent weights, and the bias of the GRU cell, respectively, and they are expressed as

$$\boldsymbol{W} = \begin{bmatrix} W_r \\ W_z \\ W_{\hat{h}} \end{bmatrix}, \boldsymbol{R} = \begin{bmatrix} R_r \\ R_z \\ R_{\hat{h}} \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} b_{Wr} \\ b_{Wz} \\ b_{W\hat{h}} \end{bmatrix} \tag{6}$$

where the subscripts r, z, and $\hat{h}$ represent that the values of input weights, recurrent weights, and the bias belong to the reset gate, the update gate, and the candidate state, respectively. The operation of a GRU cell can be formulated as:

$$\begin{cases} r_t = \sigma_g(W_r \mathbf{x}_t + b_{Wr} + R_r \mathbf{h}_{t-1}) \\ z_t = \sigma_g(W_z \mathbf{x}_t + b_{Wz} + R_z \mathbf{h}_{t-1}) \\ \hat{h}_t = \sigma_s(W_{\hat{h}} \mathbf{x}_t + b_{W\hat{h}} + r_t \odot (R_{\hat{h}} \mathbf{h}_{t-1})) \\ \mathbf{h}_t = (1 - z_t) \odot \hat{h}_t + z_t \odot \mathbf{h}_{t-1} \end{cases} \tag{7}$$

where $\odot$ denotes the element-wise multiplication of two vectors.

Following the GRU network, each FC layer is activated independently at each time step, and the output $\boldsymbol{Y}_t$ of a fully connected layer at time step t can be formulated as:

$$\boldsymbol{Y}_t = \boldsymbol{W}_f \boldsymbol{u}_t + \boldsymbol{b}_f \tag{8}$$

where $\boldsymbol{u}_t$ is the input of the FC layer at time step t. $\boldsymbol{W}_f$ and $\boldsymbol{b}_f$ are the weight matrix and bias vector of the FC layer, respectively.

3.3. Training algorithm

The proposed framework is parameterized based on the experimental data. To this end, the Adam algorithm [34], which has low memory requirements for first-order gradient-based optimization of the stochastic objective function, is adopted. The Adam algorithm iteratively updates the parameters by:

$$\boldsymbol{\theta}_{\kappa+1} = \boldsymbol{\theta}_\kappa - \frac{\alpha \boldsymbol{\eta}_\kappa}{\sqrt{\boldsymbol{v}_\kappa} + \varepsilon} \tag{9}$$

where $\boldsymbol{\theta}_\kappa$ is the parameter vector of the framework at κ-th iteration, α is the learning rate and is set to 0.004, ε is a small value to avoid division by zero and set to 10^{-8}, $\boldsymbol{\eta}_\kappa$ and $\boldsymbol{v}_\kappa$ are updated by:

$$\begin{cases} \boldsymbol{\eta}_\kappa = \beta_1 \boldsymbol{\eta}_{\kappa-1} + (1-\beta_1)\nabla E(\boldsymbol{\theta}_\kappa) \\ \boldsymbol{v}_\kappa = \beta_2 \boldsymbol{v}_{\kappa-1} + (1-\beta_2)[\nabla E(\boldsymbol{\theta}_\kappa)]^2 \\ \widetilde{\boldsymbol{\eta}}_\kappa = \dfrac{\boldsymbol{\eta}_\kappa}{1-\beta_1^\kappa} \\ \widetilde{\boldsymbol{v}}_\kappa = \dfrac{\boldsymbol{v}_\kappa}{1-\beta_2^\kappa} \end{cases} \tag{10}$$

where β_1 and β_2 are decay rates of gradient moving average and squared gradient moving average, defined as 0.9 and 0.999, respectively, $\widetilde{\boldsymbol{\eta}}_\kappa$ and $\widetilde{\boldsymbol{v}}_\kappa$ are bias-corrected values of $\boldsymbol{\eta}_\kappa$ and $\boldsymbol{v}_\kappa$, respectively, $\nabla E(\cdot)$ represents the gradient of the loss function $E(\cdot)$, which is defined as the half mean squared error of the responses of the framework for each time step:

$$E = \frac{1}{2\varsigma}\sum_{i=1}^{\varsigma}\sum_{j=1}^{\zeta}\left(y_{ij} - y'_{ij}\right)^2 \tag{11}$$

where y_{ij} is the i-th target output for response j, y'_{ij} is the i-th predicted output of the framework for response j, ς and ζ are the length of the sequence and the number of the responses, respectively.

In this work, the model is trained based on an NVIDIA Quadro P400 GPU for 200 epochs.

4. Results

4.1. Discussion of different input sizes of early cycle set

The size of the early cycle set determines how much historical data are available to the proposed framework. Therefore, it is a critical parameter impacting the prediction performance. Here, we vary the size of the early cycle set from 0 to 20 cycles to examine the prediction performance. The hyperparameters of the framework are detailed in Supplementary Note 3. The corresponding results are described in Fig. 3. A similar influence of the size of early cycle set can be observed from the

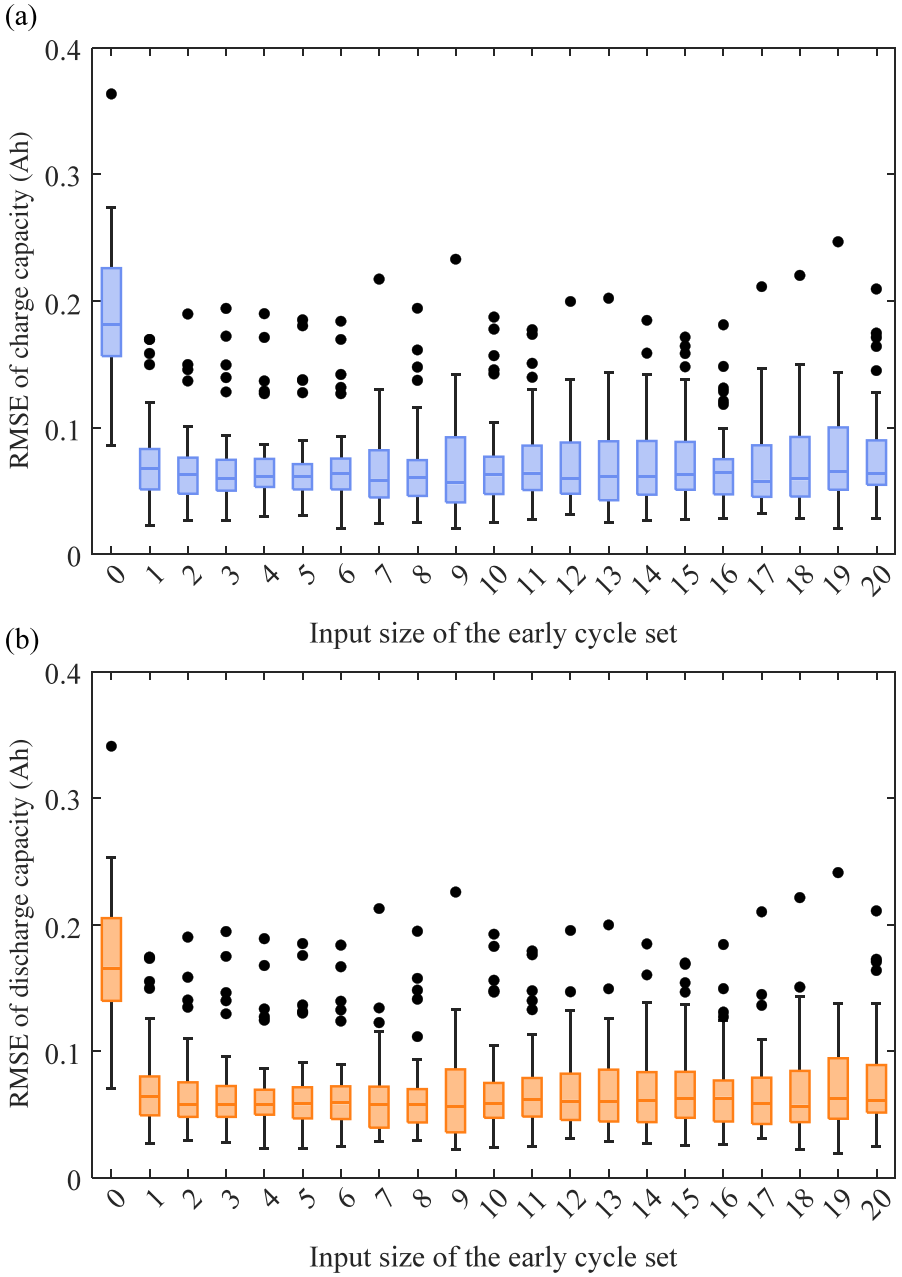

Fig. 3. The influence of the size of the early cycle set on capacity degradation prediction: (a) RMSE of the charge capacity; (b) RMSE of the discharge capacity.

prediction results of charge and discharge capacities. Compared with the case of the zero-size early cycle set (i.e., no historical information is provided), the RMSE of other cases is dramatically low. This indicates that the integration of the current plan and the early cycle set is conducive to degradation prediction. The main reason for this is detailed in Section 4.5. Another important finding from Fig. 3 is that the proposed framework performs well in the case of a small size early cycle set (<9 cycles). In particular, the proposed framework achieves a prediction median RMSE of 2.4% using only the first cycle data. Therefore, the proposed framework can be used for accurate prediction by using little historical data, significantly saving the time to collect battery degradation data.

4.2. Comparison with existing methods

To evaluate the performance of the proposed framework, a comparative study involving the existing mainstream methods is conducted on the validation set. Considering that the end-to-end methods cannot predict the capacity degradation trajectory, we employ four representative trajectory prediction methods as the benchmarks, including two methods based on mathematical functions (particle filter method [15] and Box-Cox transformation method [21]) and the other two based on time series (LSTM RNN method [19] and ARIMA-Empirical mode decomposition method [17]). These methods are abbreviated as the PF, Box-Cox, LSTM RNN, and ARIMA-EMD methods hereinafter, respectively. The key parameters of the four methods are detailed in Supplementary Table 3 and Supplementary Note 4. The proposed framework is set to the same parameters as Supplementary Note 3. Note that the initial model parameters of the PF method and the transformation coefficient of the Box-Cox method are initialized based on the training set. Each existing method is evaluated over a wide range of sizes of the early cycle set (from 1 cycle to 70 cycles), and the most accurate result is selected for the following comparison.

The comparison results of the fixed and random conditions are visualized in Fig. 4. As expected, although existing methods are mostly performing well under fixed working conditions, they cannot ensure reliable predictions under random conditions. In contrast, the proposed framework has a balanced performance for both fixed and random conditions. More importantly, the proposed framework shows a remarkably high precision by using the one-cycle data only.

To further analyze the differences between these methods, the prediction results of two batteries cycled under fixed and random conditions are shown in Fig. 5 and Fig. 6, respectively. It is seen that all the existing methods seem to achieve effective prediction, but none of them can accurately predict the trajectory of the first stage (before the 20th cycle). In particular, owing to disturbance of the first stage, the recursion-dominated LSTM RNN method performs poorly even in the second stage. These results reveal two limitations of the existing methods, i.e., the dependence on the fixed historical policy, and the assumption of an unchanged policy over battery life. These issues preclude the existing methods from performing reliable predictions under time-varying working conditions. By comparison, the proposed framework accurately predicts battery degradation at both the first and second stages (before and after the 20th cycle). This is because the introduction of the current plan enables the framework to notice the transition between the two stages. More importantly, the proposed framework merely requires as less as one early cycle data for accurate prediction.

Fig. 6 reveals that the existing methods, as expected, hardly approximate the capacity degradation trajectory under the random condition. For the PF method and Box-Cox method, the main reason is that their low-freedom mathematical functions cannot account for the fluctuating capacity degradation trajectory cycled under random conditions. Theoretically, the used time series models in the other two methods, particularly in the LSTM RNN method, have high freedom to extrapolate the development of capacity degradation and even future operating conditions. However, as they only take as input historical data, their predictions cannot adaptively accommodate the changes in future current plans. Also, their purely recursive prediction mechanism might lead to error accumulation. In contrast, the proposed framework can map the input of future operating conditions to the capacity trajectory in an end-to-end manner. Thus, it can be observed that the proposed framework achieves an excellent prediction under the random policy by using few early cycle data.

4.3. Performance of different core layers

As a general deep learning-enabled paradigm, the proposed framework is flexible to accept the dominations of different cores. In addition

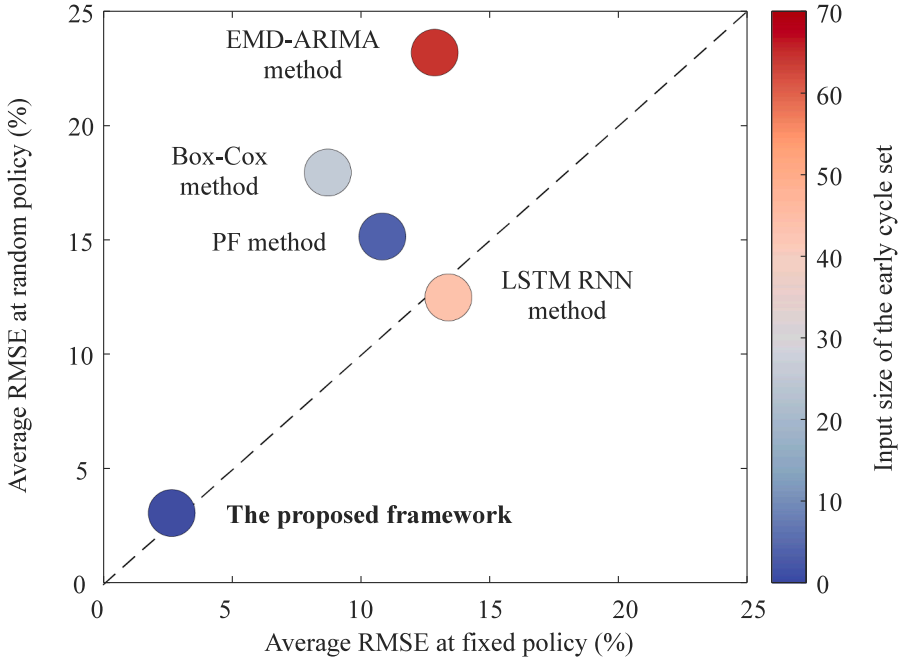

Fig. 4. Comparison of capacity prediction under the fixed and random conditions among the proposed framework and the existing methods.

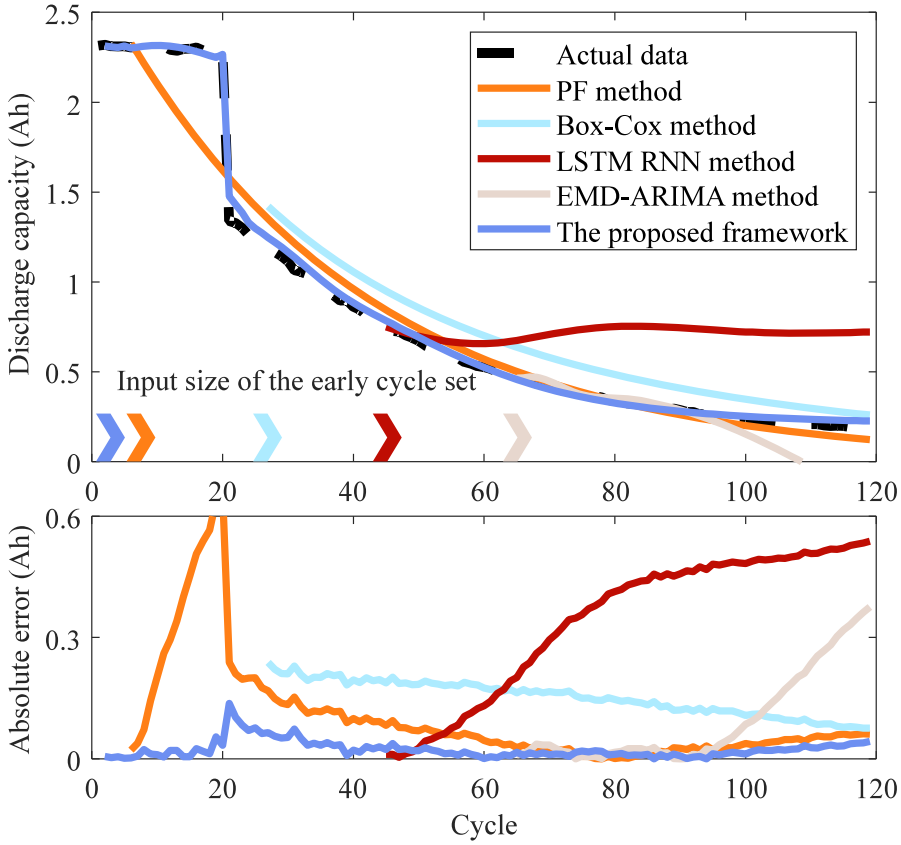

Fig. 5. Comparison of capacity prediction under a fixed condition among the proposed framework and the existing methods.

to the GRU layer, we evaluate the performance of two extra types of layers. One layer composed of FC layers (with the same number of trainable parameters as the proposed network) is introduced, where tanh-layers are designed following the FC layers to activate the nonlinear approximation ability. Another layer using LSTM units [19] instead of the GRU units (with the same hyperparameters as the proposed network) is also employed as a benchmark. The key parameters of these three layers are detailed in Table S4. An early cycle set of the first 20 cycles of the batteries is used to construct the *Q-V* matrix for comparison.

The comparison of the three frameworks dominated by different core layers (termed FC framework, LSTM framework, and GRU framework hereinafter) on capacity degradation prediction is shown in Fig. 7. Compared with the FC framework and LSTM framework, the RMSE of the GRU framework in the great majority of the batteries is relatively low, as shown in Fig. 7(a,b). To clearly exhibit the differences among the three frameworks, Fig. 7(c,d) show the distribution of RMSEs. As expected, the performance of the FC framework is significantly worse than those of the other two frameworks, owing to its intrinsic inability in processing time series [35]. By comparison, both the LSTM framework and GRU framework show excellent performance thanks to the core of the RNN layer. It is generally expected that the LSTM framework could outperform the GRU framework because of its higher degrees of freedom. However, as shown in the results, the median, quartiles, and limits of RMSE of the GRU framework are slightly lower than those of the LSTM framework, although resulting in more outliers. Similar results have also been reported in [36–38] in other fields, which attribute this to the limited data availability [37], the weak contribution of the forget gate [36], etc. Nevertheless, it should be underlined here that there is no theoretical evidence [39] to support which kind of RNN is better, although our work mainly takes the more efficient one, GRU framework, as an instance. We purely give an illustrative example here to show that the proposed framework can be dominated by different RNN cores.

In addition, as our method simultaneously predicts the charge and discharge capacities, the Coulombic efficiency, a crucial indicator of battery degradation [40,41], can be computed as the ratio of discharge capacity to charge capacity to further evaluate the performance, as shown in Fig. 7(e). It is observed that both the LSTM framework and GRU framework perform well, as they did in the prediction of capacity trajectory. Hence, the proposed framework is also able to support the prediction of Coulombic efficiency.

4.4. Performance on LFP/graphite batteries

LFP/graphite batteries have gained in popularity in recent years [42], as they outperform most other NMC/graphite batteries in long-life services. To evaluate the performance of the proposed framework on LFP/graphite batteries, two fast-charging datasets from Stanford-MIT [11,43], covering 124 and 45 commercial LFP/graphite batteries, respectively, are combined herein for modeling and validation, although their fixed policies merely provide unilateral evaluation. The training dataset is divided into two groups: one group (80 batteries) accounting for approximately 50% of batteries are used to iteratively optimize the trainable parameters, and another group (43 batteries) covering about 25% of batteries is used to evaluate the prediction performance of each epoch. The network in the most accurate epoch is chosen as the trained framework. The rest 44 batteries are then employed as the validation dataset to evaluate the prediction performance of the proposed framework. Two batteries are excluded because of the abnormal discharge capacity and service life [11]. The hyperparameters designed on NMC/graphite batteries (detailed in Supplementary Note 3) are shared here to evaluate the friendliness of the proposed framework.

We first examine the performance of capacity degradation prediction of LFP/graphite batteries with different sizes of the early cycle set, as

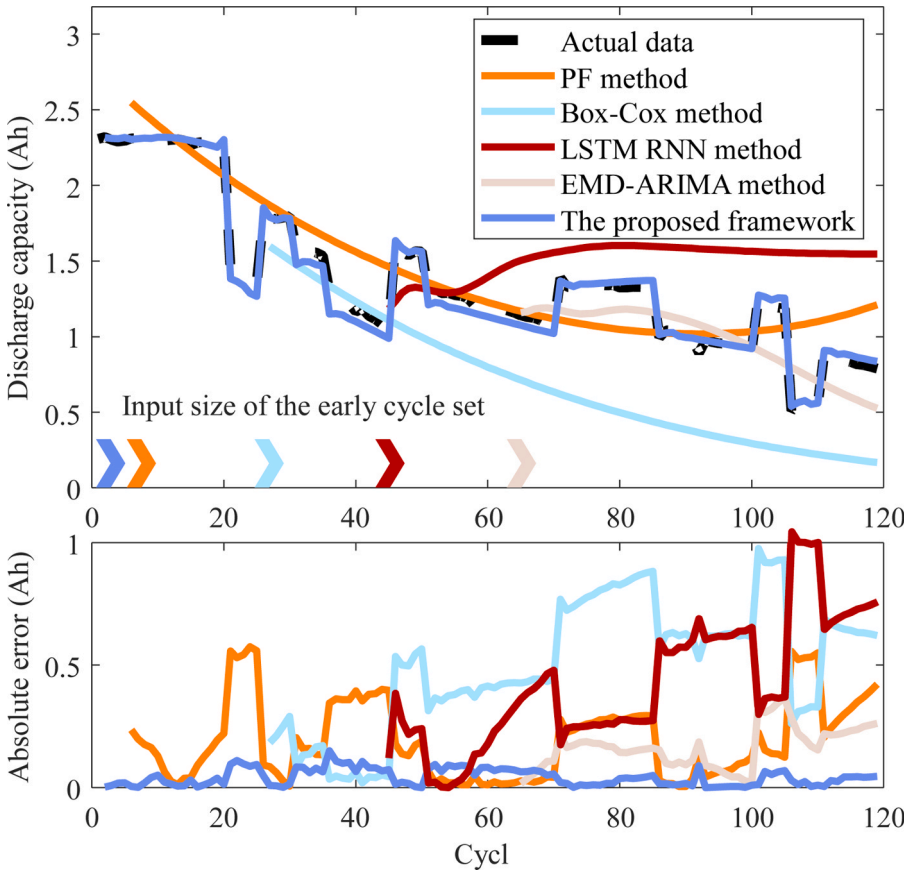

Fig. 6. Comparison of capacity prediction under a random condition among the proposed framework and the existing methods.

shown in Fig. 8(a). It is seen that the RMSE of cases of the other input sizes is relatively low compared with the case of the zero-size early cycle set. This trend is similar to the prediction for NMC/graphite batteries. Taking the nominal capacity as a benchmark, the proposed framework achieves a prediction RMSE of 2.3% (less than 4.5%) using only the first 30 cycle data (3.8% of the whole-life cyclic data). The RUL prediction performance is also investigated as it is concerned by most existing studies [9–11]. The RUL can be easily captured from the predicted capacity trajectories of the proposed framework. Fig. 8(b) shows the relative error of RUL prediction with different input sizes of the early cycle set. The 'full' model from Stanford-MIT is employed as a competitive benchmark herein, and the so-called indexing scheme for $j = 10$ in [11] is selected for evaluation. One handmade feature of the full model, namely the integral of temperature versus time, is removed because of the abnormal temperature collections of some batteries [11]. It is seen that the relative prediction error of both methods is at a low level after the zero-size of the early cycle set. Besides, due to the lack of available early cycles for implementation, the 'full' model is 'on strike' when the input sizes of the early cycle set are less than 20 (shadowed in Fig. 8(b)). The 'full' model provides an error up to 20% even though the early cycles are accumulated to 50 cycles. It is reasonable as the major handmade features of the full model proved to deteriorate when few early cycles are available [11]. The proposed framework, by comparison, achieves a median relative error of 7.2% using only the first 20 cycle data, showing excellent early predictability. The reason for this is that the proposed deep learning-enabled framework can extract dense valuable features from the *Q-V* matrix (including the major handmade feature of the full model). We can also find that, by using the proposed framework, only the first 20 cycle data is required for good RUL prediction, while the first 30 cycle data is needed for capacity trajectory prediction. Such an early 'inflection point' indicates that the proposed framework is also competent for RUL prediction.

Fig. 8(c) further shows the prediction of capacity degradation trajectories of eight batteries with different EOLs, where the long-life batteries (EOL > 1500 cycles) are nearly the representatives of the outliers in Fig. 8(a). In the cases of the normal batteries (EOL < 1500 cycles), accurate prediction is achieved by using only the first 30 cycle data. In the cases of the long-life batteries, it is similar to the normal batteries except that marked underprediction occurs at the tail of capacity trajectories. We speculate that there are two major reasons for this: the first is the difficulty of long-term prediction caused by the instability and higher uncertainty of the close-to-EOL battery; the second is the scarcity of such long-life battery samples for training. To further investigate the development of this underprediction, we then expand the input size from 30 to 50 and 100 to reduce the difficulty of long-term prediction. It is found that, by expanding the input size, this underprediction tends to be relatively alleviated but still stronger than that of the normal batteries. But for these long-life batteries, we believe such a surpassingly advanced prediction is already valuable. In addition, we find a truth that can explain why the inflection point of RUL prediction is ahead of that of capacity trajectory prediction. That is, the terminal of the trajectory (especially for the most long-life battery) is perfectly predicted with different input sizes. This gives us reason to believe that understanding the high-dimensional capacity trajectory is more challenging than predicting the scalar RUL.

Overall, the proposed framework can be applied to predict both capacity trajectory and RUL of LFP/graphite batteries. Compared with the 'full' model from Stanford-MIT, it is mainly powerful in balancing early (<3.8% of early cycle data requirements) and accurate prediction (median error < 2.3%@trajectory and <7.2%@RUL) of both charge and discharge capacity degradation trajectories.

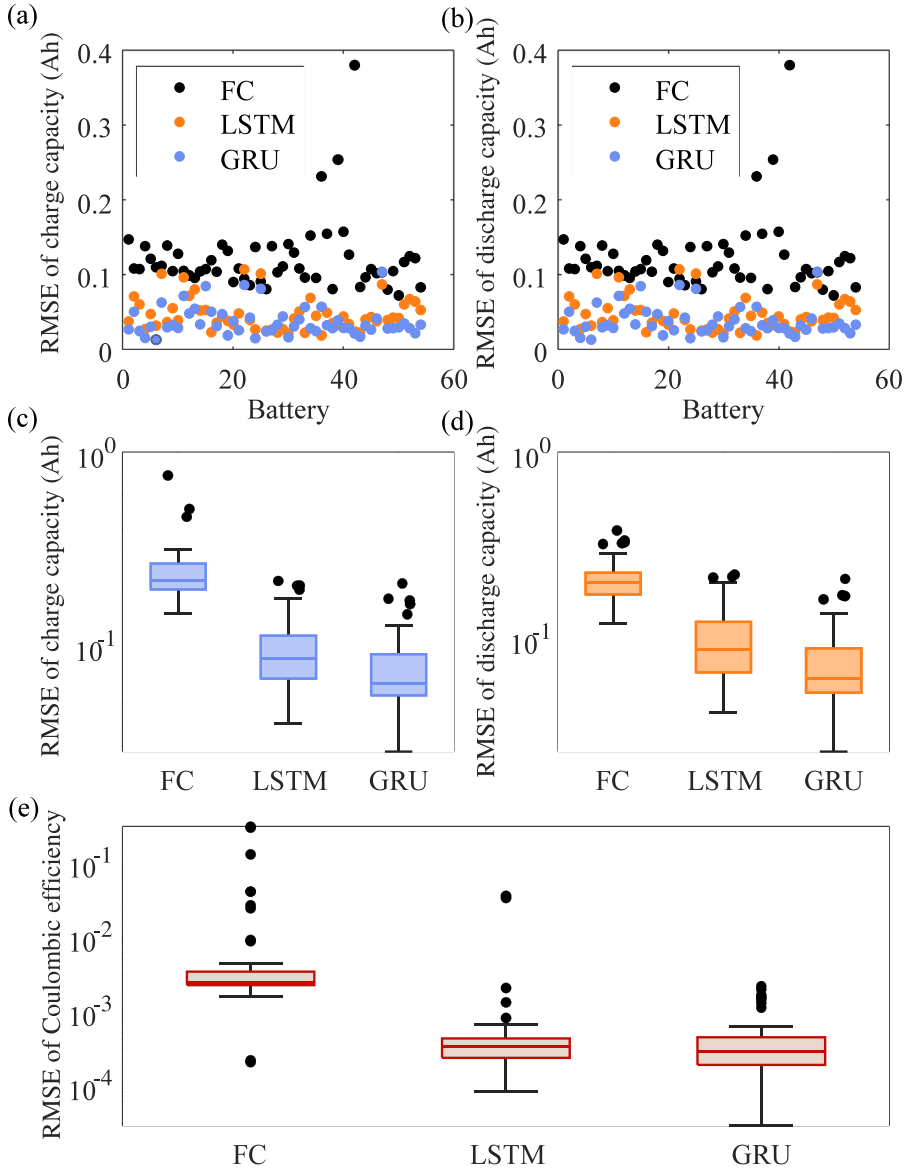

Fig. 7. Comparison of different core layers (FC layer, LSTM layer, and GRU layer) on capacity degradation prediction: (a) RMSE of charge capacity; (b) RMSE of discharge capacity; (c) logarithmic boxplot of the RMSE of charge capacity; (d) logarithmic boxplot of the RMSE of discharge capacity; (e) logarithmic boxplot of the RMSE of Coulombic efficiency.

4.5. Rationalization of predictive performance

To investigate the rationality of the predictive performance, two ancillary frameworks derived from the proposed framework are introduced. The first is to build an ancillary framework without the early cycle set and padding (merely retaining the component I & II in Fig. 2) to decouple the contribution of the current plans. The second is to establish an ancillary framework without the current plans (merely retaining the component III & IV in Fig. 2) to represent the contribution of the early cycle set. For convenience, the first ancillary framework, the second ancillary framework and the proposed framework are hereinafter referred to as the current framework, the 'QV' framework and the 'full' framework, respectively.

The prediction performance of the discharge capacity trajectory of the three frameworks is described in Fig. 9, where the shaded bar denotes that the QV framework is 'on strike' due to the lack of *Q-V* data. We observe that, thanks to the integration of current plans and *Q-V* matrix, the errors of the 'full' framework (as shown in Fig. 9(a)) are lower than those of either of them (as shown in Fig. 9(b,c)), except that at the zero-input size (blaming the participation of full-zero *Q-V* matrix input). It is found that the current framework significantly outperforms the QV framework, which means that the full framework is dominated by the input of future current series. In other words, this proves that the full framework, in this case, is mainly committed to learning the relationship between the future current and capacity degradation trajectory in an end-to-end manner.

The prediction performance of three frameworks on the fixed-policy and random-policy batteries are further separately plotted in Fig. 9(d–i). As the main contributor, the current framework has balanced performance on the two groups of batteries (as shown in Fig. 9(e,h)). As a result, the full framework seems to inherit its advantage as shown in Fig. 9(e,g). By comparison, the QV framework has extremely different performance on the two groups of batteries (as shown in Fig. 9(f,i)). It is logical that the QV framework poorly performs in Fig. 9(f), because the *Q-V* matrix collected under fixed policies hardly benefits the framework to understand how the capacity degrades with the random policies, as the cases of the LSTM RNN method in Fig. 6. In contrast, the prediction of QV framework is also poor but highly scattered on the fixed-policy

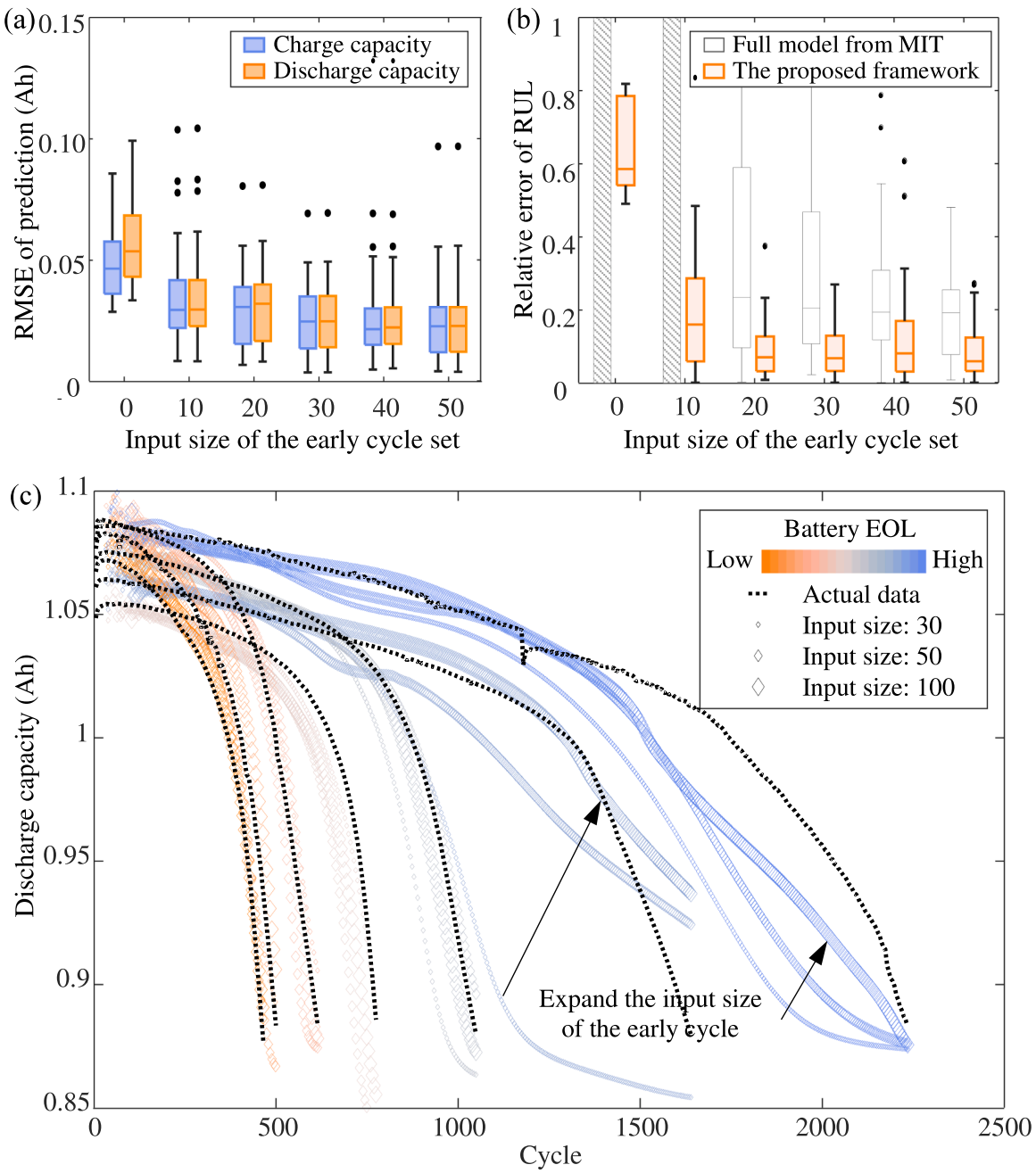

Fig. 8. Result of the proposed framework on LFP/graphite batteries: (a) RMSE of capacity degradation prediction with different input sizes of the early cycle set; (b) relative error of RUL prediction with different input sizes of the early cycle set; (c) prediction of capacity degradation trajectories of eight batteries (differ in EOL) with different input sizes of the early cycle set.

batteries as shown in Fig. 9(i). It is seen that the maximum error is much larger than that of random-policy batteries, while the minimum error is close to that of the current framework. The reason for this extreme is that some fixed-policy batteries are under a similar policy to the *Q-V* matrix collection policy. This indicates that the RNN of the QV framework can recursively deduce the capacity trajectory by learning the degradation rate of the historical *Q-V* matrix, achieving comparable performance to the current framework.

So far, it is of interest to find that although the two ancillary frameworks are structurally identical, their prediction mechanisms may be very different. Hence, we have reason to believe that there are at least two mechanisms included in the full framework: the first is to directly learn the relationship between current conditions and capacity degradation trajectories in an end-to-end manner, like the current framework; the second is to recursively learn the rate of capacity degradation within dense voltage ranges, as the QV framework and the most studies [17,19] do. More importantly, with the help of the powerful high-dimensional approximation and time series prediction capability of the RNN, these two mechanisms can be dynamically weighted by the full framework, rather than focusing on either of them. This is, in summary, the answer to our question regarding the excellent performance of the integration of both the current plan and *Q-V* matrix.

We then examine the prediction performance of the three frameworks on the LFP/graphite batteries, as shown in Fig. 9(j–l). The full framework, as expected, outperforms the two ancillary frameworks, showing the effectiveness of the integration of both the current plan and *Q-V* matrix. It is seen that there is a close contest between the two ancillary frameworks, which implies both two potential mechanisms

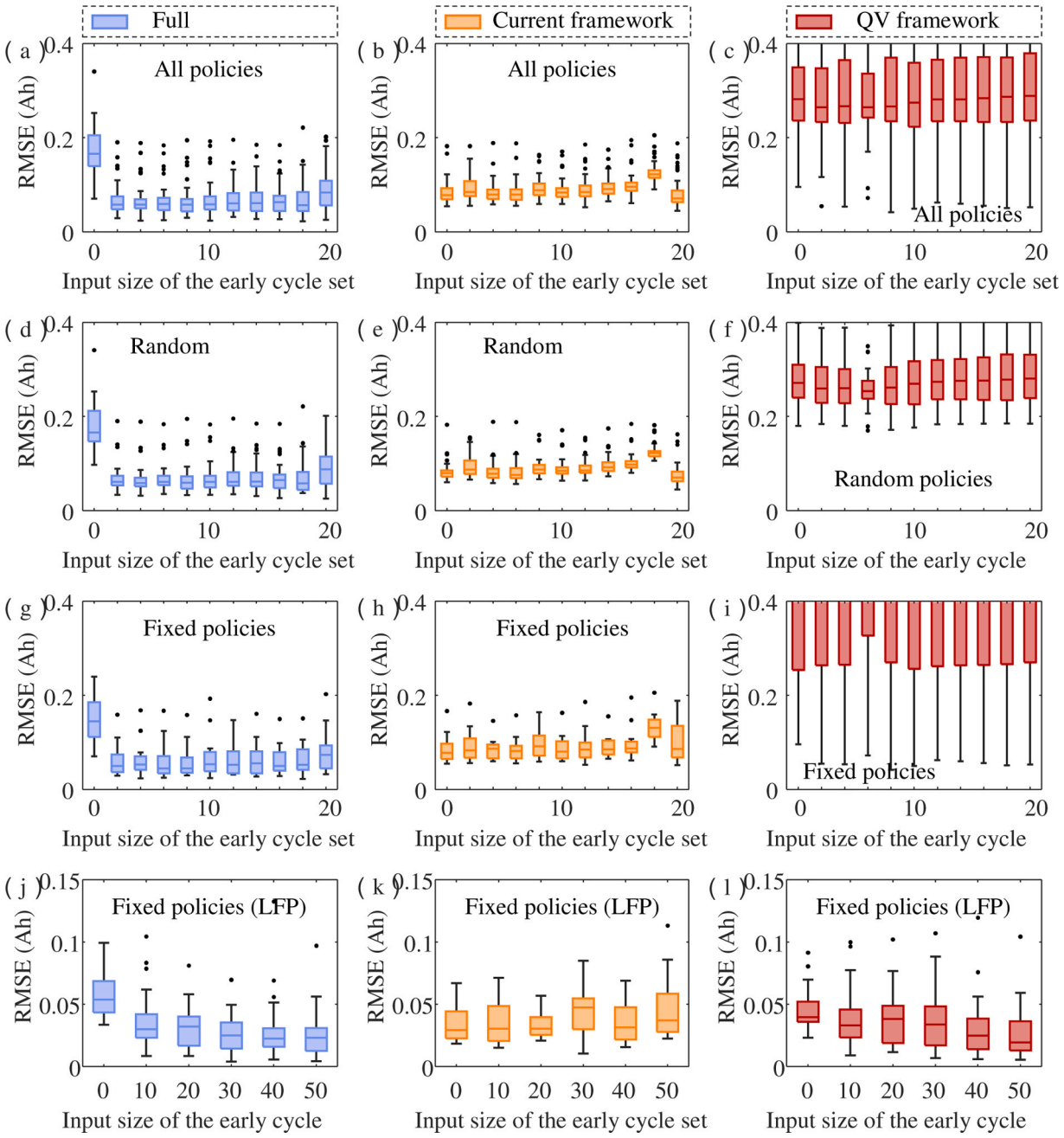

Fig. 9. Performance of discharge capacity trajectory of the full frameworks, current framework and QV framework with different input sizes of the early cycle set: (a) RMSE of all batteries with the full framework; (b) RMSE of all batteries with the current framework; (c) RMSE of all batteries with the QV framework; (d) RMSE of random-policy batteries with the full framework; (e) RMSE of random-policy batteries with the current framework; (f) RMSE of random-policy batteries with the QV framework; (g) RMSE of fixed-policy batteries with the full framework; (h) RMSE of fixed-policy batteries with the current framework; (i) RMSE of fixed-policy batteries with the QV framework; (j) RMSE of fixed-policy LFP/graphite batteries with the full framework; (k) RMSE of fixed-policy LFP/graphite batteries with the current framework; (l) RMSE of fixed-policy LFP/graphite batteries with the QV framework.

contribute to the full framework. Specifically, the full framework tends to be dominated by the current plan at small input size and the *Q-V* matrix at large input size, because the current framework shown in Fig. 9 (k) performs well at small input size while the QV framework shown in Fig. 9(l) begins to excel at large input size. These results are reasonable as the employed two LFP/graphite battery datasets from Stanford-MIT [11] are generated under fixed policies. At the very beginning, the full framework maps the trajectory through the current (i.e., tend to the first mechanism), owing to little *Q-V* evolution information; after a few cycles, since the *Q-V* matrix expands with the expansion of the input size, the full framework recursively deduces the trajectory (i.e., tend to the second mechanism) based on the evolution of *Q-V* matrix within dense voltage ranges. Besides, it can be seen in Fig. 9(l) that the QV framework, due to its recursion, tends to perform better with the expansion of input size, showing a stronger cycle dependence than the first mechanism. However, we do not find the obvious cycle dependence of the full framework in Figs. 3 and 8(a). This is because, in addition to the second mechanism, there is at least the first mechanism to join as one of the contributors to the full framework. Driven by these two mechanisms with different cycle dependence, the full framework tends to be weakly correlated with the cycle.

In total, introducing the two ancillary frameworks help to rationalize

the observation of the full framework. The characteristics embodied in both NMC/graphite and LFP/graphite batteries are consistent.

5. Limitations and outlook

The present study can be improved in the future. First, this study only involves the policies that randomized operating current. As a deep learning-enabled paradigm, the proposed framework does not limit the input feature dimensions. Hence, the proposed framework can be explored to apply to more generalized random policies, such as temperature, depth of discharge, or other factors [44] leading to degradation. Second, the proposed framework should be trained by a few experimental data, so the potential of such a general deep learning framework in transfer learning [45] can be further explored to reduce the required data amount. Finally, the proposed framework can be explored to apply to big data. As a data-driven approach, the proposed framework does not assume specific battery materials or application scenarios. The introduction of big data covering the historical real-world data distribution [46] might lead to more proper and effective learning.

6. Conclusion

In this paper, we firstly present a large battery degradation dataset (includes 77 commercial lithium-ion battery cells) covering the fixed and random operating conditions (published at https://data.mendeley.com/datasets/kw34hhw7xg/2). To our knowledge, this is the first battery degradation dataset that accommodates greatly varying working loads for each cell. We then develop a general deep learning-enabled framework to predict the battery capacity degradation trajectories under both fixed and random future operating conditions by integrating future current plan and few early capacity-voltage data as inputs. Our results demonstrate that the median prediction RMSE is within 2.4% for NMC/graphite batteries and 2.3% for LFP/graphite batteries by using 3.8% of the whole life data only. Compared with the existing methods, the proposed framework predicts more accurately and has a satisfying performance for both fixed and random future conditions. The proposed framework is promising to assist with battery rapid development and customization of battery management software to prolong battery life, and for the intelligent control of many battery cells ("life balancing") in a battery pack.

CRediT authorship contribution statement

Jiahuan Lu: Methodology, Writing – original draft. **Rui Xiong:** Conceptualization, Supervision, Writing – review & editing. **Jinpeng Tian:** Data curation, Writing – review & editing. **Chenxu Wang:** Data curation. **Chia-Wei Hsu:** Writing – review & editing. **Nien-Ti Tsou:** Writing – review & editing. **Fengchun Sun:** Supervision, Writing – review & editing. **Ju Li:** Conceptualization, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The dataset developed in this study is available at https://data.mendeley.com/datasets/kw34hhw7xg/2.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (Grant No. 51922006, 51877009) and the China National Postdoctoral Program for Innovative Talents (Grant No. BX2021035). The battery degradation tests were carried out by the Advanced Energy Storage and Application (AESA) Group at Beijing Institute of Technology.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ensm.2022.05.007.

References

[1] R. Yang, R. Xiong, W. Shen, X. Lin, Extreme learning machine based thermal model for lithium-ion batteries of electric vehicles under external short circuit, Engineering 7 (3) (2021) 395–405, https://doi.org/10.1016/j.eng.2020.08.015.
[2] G. Zubi, R.S. Adhikari, N.E. Sánchez, W. Acuña-Bravo, Lithium-ion battery-packs for solar home systems: layout, cost and implementation perspectives, J. Energy Storage 32 (2020), 101985, https://doi.org/10.1016/j.est.2020.101985.
[3] L. Shen, Q. Cheng, Y. Cheng, L. Wei, Y. Wang, Hierarchical control of DC microgrid for photovoltaic EV charging station based on flywheel and battery energy storage system, Electr. Power Syst. Res. 179 (2020), 106079, https://doi.org/10.1016/j.epsr.2019.106079.
[4] J. Tian, R. Xiong, W. Shen, F. Sun, Electrode ageing estimation and open circuit voltage reconstruction for lithium ion batteries, Energy Storage Mater. 37 (2021) 283–295, https://doi.org/10.1016/j.ensm.2021.02.018.
[5] M.R. Palacín, A. de Guibert, Why do batteries fail? Science 351 (2016), 1253292 https://doi.org/10.1126/science.1253292 (1979)–1253292.
[6] S.L. Wang, W. Tang, C. Fernandez, C.M. Yu, C.Y. Zou, X-Q. Zhang, A novel endurance prediction method of series connected lithium-ion batteries based on the voltage change rate and iterative calculation, J. Clean. Prod. 210 (2019) 43–54, https://doi.org/10.1016/j.jclepro.2018.10.349.
[7] A. Fotouhi, D.J. Auger, K. Propp, S. Longo, M. Wild, A review on electric vehicle battery modelling: From Lithium-ion toward lithium–sulphur, Renew. Sustain. Energy Rev. 56 (2016) 1008–1021, https://doi.org/10.1016/j.rser.2015.12.009.
[8] M.S.H. Lipu, M.A. Hannan, A. Hussain, M.M. Hoque, P.J. Ker, M.H.M. Saad, et al., A review of state of health and remaining useful life estimation methods for lithium-ion battery in electric vehicles: challenges and recommendations, J. Clean. Prod. 205 (2018) 115–133, https://doi.org/10.1016/j.jclepro.2018.09.065.
[9] Y. Zhang, Q. Tang, Y. Zhang, J. Wang, U. Stimming, AA. Lee, Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning, Nat. Commun. 11 (2020) 1706, https://doi.org/10.1038/s41467-020-15235-7.
[10] S. Zhang, X. Guo, X. Zhang, Multi-objective decision analysis for data-driven based estimation of battery states: a case study of remaining useful life estimation, Int. J. Hydrog. Energy 45 (2020) 14156–14173, https://doi.org/10.1016/j.ijhydene.2020.03.100.
[11] K.A. Severson, P.M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, et al., Data-driven prediction of battery cycle life before capacity degradation, Nat. Energy 4 (2019) 383–391, https://doi.org/10.1038/s41560-019-0356-8.
[12] J. Hong, D. Lee, E.R. Jeong, Y. Yi, Towards the swift prediction of the remaining useful life of lithium-ion batteries with end-to-end deep learning, Appl. Energy 278 (2020), 115646, https://doi.org/10.1016/j.apenergy.2020.115646.
[13] R. Xiong, Y. Zhang, J. Wang, H. He, S. Peng, M. Pecht, Lithium-ion battery health prognosis based on a real battery management system used in electric vehicles, IEEE Trans. Veh. Technol. 68 (2019) 4110–4121, https://doi.org/10.1109/TVT.2018.2864688.
[14] Y. Xing, E.W.M. Ma, K.L. Tsui, M. Pecht, An ensemble model for predicting the remaining useful performance of lithium-ion batteries, Microelectron. Reliab. 53 (2013) 811–820, https://doi.org/10.1016/j.microrel.2012.12.003.
[15] W. He, N. Williard, M. Osterman, M. Pecht, Prognostics of lithium-ion batteries based on Dempster-Shafer theory and the Bayesian Monte Carlo method, J Power Sources 196 (2011) 10314–10321, https://doi.org/10.1016/j.jpowsour.2011.08.040.
[16] D. Liu, W. Xie, H. Liao, Y. Peng, An integrated probabilistic approach to lithium-ion battery remaining useful life estimation, IEEE Trans. Instrum. Meas. 64 (2015) 660–670, https://doi.org/10.1109/TIM.2014.2348613.
[17] Zhou Y., Huang M. Lithium-ion batteries remaining useful life prediction based on a mixture of empirical mode decomposition and ARIMA model. Microelectron. Reliab. 2016;65:265–73. 10.1016/j.microrel.2016.07.151.
[18] A.A. Chehade, AA. Hussein, A collaborative gaussian process regression model for transfer learning of capacity trends between Li-ion battery cells, IEEE Trans. Veh. Technol. 69 (2020) 9542–9552, https://doi.org/10.1109/TVT.2020.3000970.
[19] Y. Zhang, R. Xiong, H. He, MG. Pecht, Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries, IEEE Trans. Veh. Technol. 67 (2018) 5695–5705, https://doi.org/10.1109/TVT.2018.2805189.
[20] M. Dubarry, D. Beck, Big data training data for artificial intelligence-based Li-ion diagnosis and prognosis, J Power Sources (2020), https://doi.org/10.1016/j.jpowsour.2020.228806.
[21] Y. Zhang, R. Xiong, H. He, MG. Pecht, Lithium-ion battery remaining useful life prediction with box–cox transformation and Monte Carlo simulation, IEEE Trans. Ind. Electron. 66 (2019) 1585–1597, https://doi.org/10.1109/TIE.2018.2808918.

[22] Z. Liu, G. Sun, S. Bu, J. Han, X. Tang, M. Pecht, Particle learning framework for estimating the remaining useful life of lithium-ion batteries, IEEE Trans. Instrum. Meas. 66 (2017) 280–293, https://doi.org/10.1109/TIM.2016.2622838.
[23] G. Ma, Y. Zhang, C. Cheng, B. Zhou, P. Hu, Y. Yuan, Remaining useful life prediction of lithium-ion batteries based on false nearest neighbors and a hybrid neural network, Appl. Energy 253 (2019), 113626, https://doi.org/10.1016/j.apenergy.2019.113626.
[24] H. Wang, W. Song, E. Zio, A. Kudreyko, Y. Zhang, Remaining useful life prediction for lithium-ion batteries using fractional Brownian motion and fruit-fly optimization algorithm, Meas. J. Int. Meas. Confed. 161 (2020), 107904, https://doi.org/10.1016/j.measurement.2020.107904.
[25] K. Zhang, P. Zhao, C. Sun, Y. Wang, Z. Chen, Remaining useful life prediction of aircraft lithium-ion batteries based on F-distribution particle filter and kernel smoothing algorithm, Chin. J. Aeronaut. 33 (2020) 1517–1531, https://doi.org/10.1016/j.cja.2020.01.007.
[26] B. Gou, Y. Xu, X. Feng, State-of-health estimation and remaining-useful-life prediction for lithium-ion battery using a hybrid data-driven method, IEEE Trans. Veh. Technol. 69 (2020) 10854–10867, https://doi.org/10.1109/TVT.2020.3014932.
[27] D. Wang, F. Yang, Y. Zhao, KL. Tsui, Battery remaining useful life prediction at different discharge rates, Microelectron. Reliab. 78 (2017) 212–219, https://doi.org/10.1016/j.microrel.2017.09.009.
[28] J. Lu, R. Xiong, J. Tian, C. Wang, C.W. Hsu, N.T. Tsou, et al., Battery degradation dataset (fixed current profiles & arbitrary uses profiles), Mendeley Data V1 (2021), https://doi.org/10.17632/kw34hhw7xg.1.
[29] X. Hu, L. Xu, X. Lin, M. Pecht, Battery lifetime prognostics, Joule 4 (2020) 310–346, https://doi.org/10.1016/j.joule.2019.11.018.
[30] A. Lopez-del Rio, M. Martin, A. Perera-Lluna, R. Saidi, Effect of sequence padding on the performance of deep learning models in archaeal protein functional prediction, Sci. Rep. 10 (2020) 14634, https://doi.org/10.1038/s41598-020-71450-8.
[31] T. Raj, A.A. Wang, C.W. Monroe, DA. Howey, Investigation of path-dependent degradation in lithium-ion batteries**, Batter. Supercaps (2020), https://doi.org/10.1002/batt.202000160.
[32] B. Wu, W.D. Widanage, S. Yang, X. Liu, Battery digital twins: perspectives on the fusion of models, data and artificial intelligence for smart battery management systems, Energy AI 1 (2020), 100016, https://doi.org/10.1016/j.egyai.2020.100016.
[33] S. Gao, Y. Huang, S. Zhang, J. Han, G. Wang, M. Zhang, et al., Short-term runoff prediction with GRU and LSTM networks without requiring time step optimization during sample generation, J. Hydrol. 589 (2020), 125188, https://doi.org/10.1016/j.jhydrol.2020.125188.
[34] Z. Chang, Y. Zhang, W. Chen, Electricity price prediction based on hybrid model of adam optimized LSTM neural network and wavelet transform, Energy 187 (2019), 115804, https://doi.org/10.1016/j.energy.2019.07.134.
[35] W. de Mulder, S. Bethard, MF. Moens, A survey on the application of recurrent neural networks to statistical language modeling, Comput. Speech Lang. 30 (2015), https://doi.org/10.1016/j.csl.2014.09.005.
[36] S. Mirzaei, J.L. Kang, KY. Chu, A comparative study on long short-term memory and gated recurrent unit neural networks in fault diagnosis for chemical processes using visualization, J. Taiwan Inst. Chem. Eng. 130 (2022), https://doi.org/10.1016/j.jtice.2021.08.016.
[37] G. Ayzel, M. Heistermann, The effect of calibration data length on the performance of a conceptual hydrological model versus LSTM and GRU: a case study for six basins from the CAMELS dataset, Comput. Geosci. 149 (2021), https://doi.org/10.1016/j.cageo.2021.104708.
[38] K.E. ArunKumar, D.V. Kalaga, C.M.S. Kumar, M. Kawaji, TM. Brenza, Forecasting of COVID-19 using deep layer recurrent neural networks (RNNs) with gated recurrent units (GRUs) and long short-term memory (LSTM) cells, Chaos Solitons Fractals 146 (2021), https://doi.org/10.1016/j.chaos.2021.110861.
[39] J. Chung, C. Gulcehre, K. Cho, Y. Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, doi: 10.48550/arXiv.1412.3555.
[40] Y. Jin, S. Li, A. Kushima, X. Zheng, Y. Sun, J. Xie, et al., Self-healing SEI enables full-cell cycling of a silicon-majority anode with a coulombic efficiency exceeding 99.9%, Energy Environ. Sci. 10 (2017) 580–592, https://doi.org/10.1039/c6ee02685k.
[41] C. Fear, T. Adhikary, R. Carter, A.N. Mistry, C.T. Love, PP. Mukherjee, In operando detection of the onset and mapping of lithium plating regimes during fast charging of lithium-ion batteries, ACS Appl. Mater. Interfaces 12 (2020) 30438–30448, https://doi.org/10.1021/acsami.0c07803.
[42] M. Elliott, L.G. Swan, M. Dubarry, G. Baure, Degradation of electric vehicle lithium-ion batteries in electricity grid services, J. Energy Storage 32 (2020), https://doi.org/10.1016/j.est.2020.101873.
[43] P.M. Attia, A. Grover, N. Jin, K.A. Severson, T.M. Markov, Y.H. Liao, et al., Closed-loop optimization of fast-charging protocols for batteries with machine learning, Nature 578 (2020) 397–402, https://doi.org/10.1038/s41586-020-1994-5.
[44] R.D. Deshpande, K. Uddin, Physics inspired model for estimating 'cycles to failure' as a function of depth of discharge for lithium ion batteries, J. Energy Storage 33 (2021), https://doi.org/10.1016/j.est.2020.101932.
[45] J. Tian, R. Xiong, W. Shen, J. Lu, XG. Yang, Deep neural network battery charging curve prediction using 30 points collected in 10 min, Joule 5 (2021), https://doi.org/10.1016/j.joule.2021.05.012.
[46] M.B. Arias, S. Bae, Electric vehicle charging demand forecasting model based on big data technologies, Appl. Energy 183 (2016), https://doi.org/10.1016/j.apenergy.2016.08.080.

Supplementary Information

Battery Degradation Prediction Against Uncertain Future Conditions with Recurrent Neural Network Enabled Deep Learning

Supplementary Note 1. Battery experimental platform for degradation experiments

An experimental platform is established for the battery degradation test, as shown in Fig. S1. It consists of a LANHE CT2001B battery test system to simulate the battery operating conditions, a GDBELL thermal chamber with three chambers to regulate the ambient temperature, and a host computer for sending control commands and receiving experimental data.

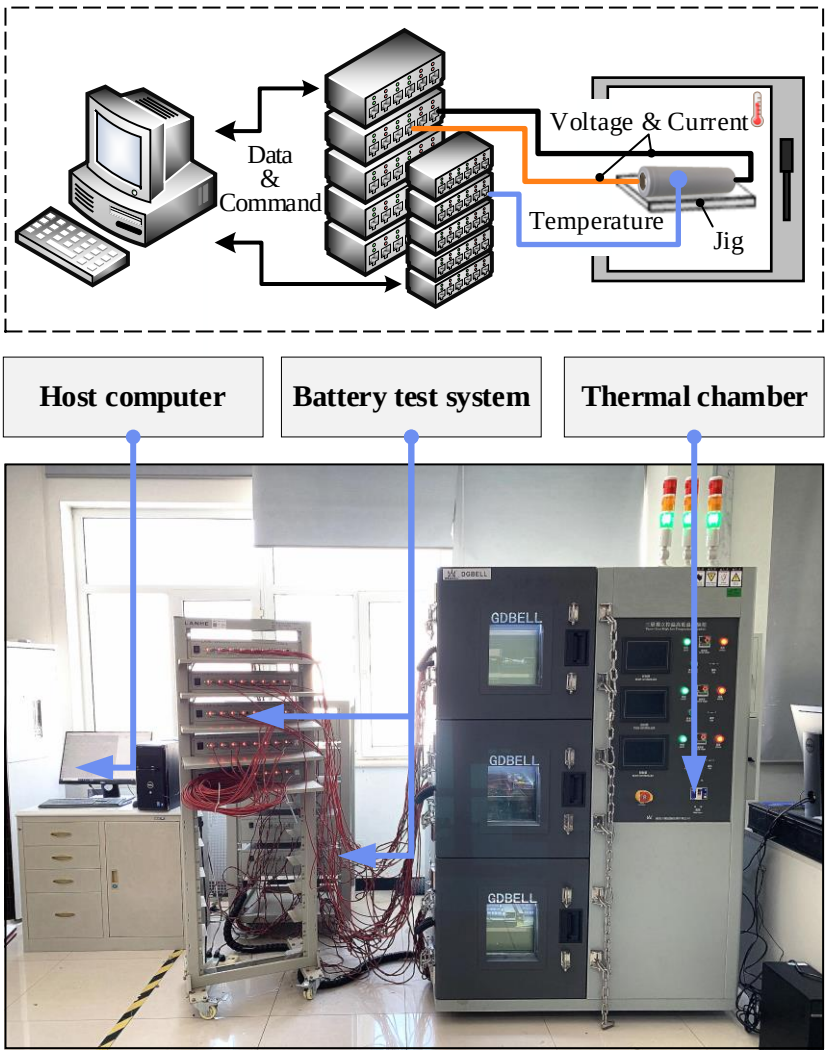

Fig. S1. Battery experimental platform for degradation experiments.

Supplementary Note 2. Arrangements and results of the cyclic degradation experiments.

The workflow diagram and the arrangements of the cyclic degradation experiments are shown in Fig. S2 and Table. S1, respectively. The numbers in the matrix of Table. S1 indicates the number of tested batteries. Notably, thanks to our design, there is no significant order of magnitude gap between the inputs of the proposed framework. Therefore, we do not specifically normalize the input and output data before training.

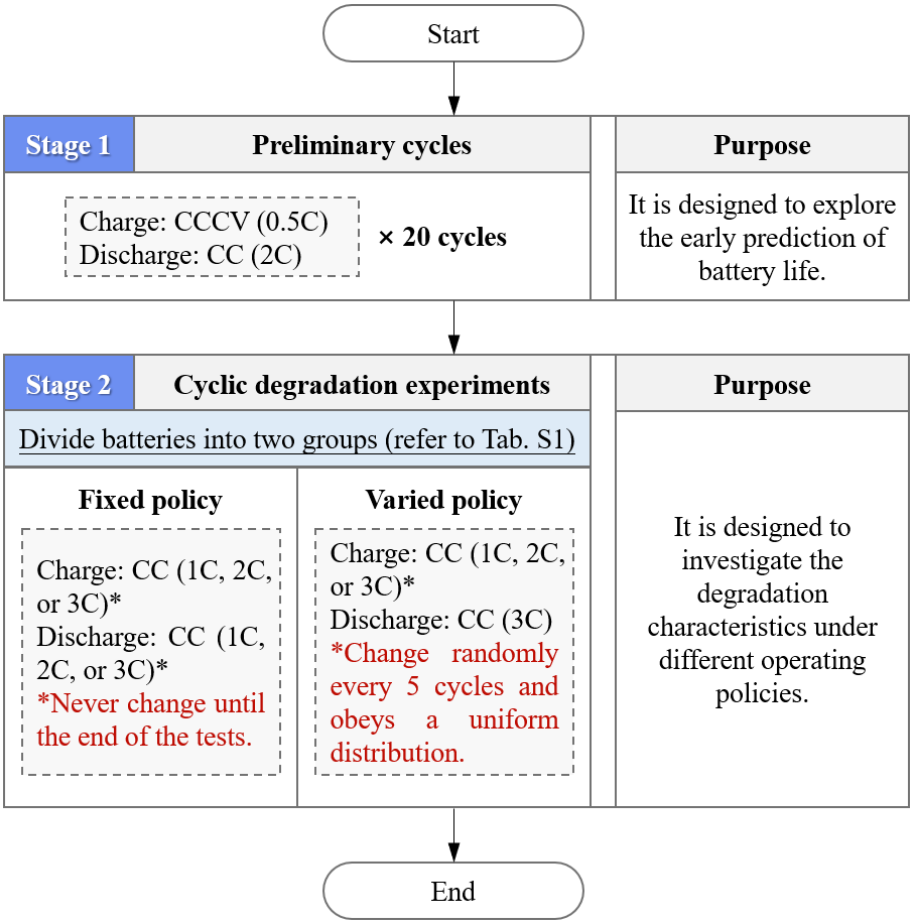

Fig. S2. Workflow diagram of the cyclic degradation experiments.

Table. S1 Test matrix for the cyclic degradation experiments.

Charge / Discharge	1C	2C	3C	Random current (1C~3C)
1C	-	3	3	-
2C	4	3	3	-
3C	-	3	3	55

The capacity degradation trajectory of batteries is plotted in Fig. S3, where the charge and discharge are abbreviated as CHG and DHG, respectively. Fig. S3 (a) shows the degradation of discharge capacity of the batteries from Group I which are charged at different current rates (1C, 2C or 3C) and discharged at a specified current (2C). Due to the one-time switching of different charge current policies, the capacity gap between the lowest and highest charge current at the 120th cycle is up to 1.7 Ah (about 90% capacity reduction). The battery degradation characteristics dominated by the varied charge policy are well disclosed in this figure, which is prevalent in real applications. For instance, the charging rates of EVs may vary for different charging piles [1]. Fig. S3 (b) describes the degradation of discharge capacity of the batteries from Group I which are charged at a specified current (2C) and discharged at different current rates (1C, 2C or 3C), showing a similar phenomenon to Fig. S3 (a). Note that the growing difference among the trajectories is much tinier than that in Fig. S3 (a), which reveals the degradation dominated by the varied charge policy is more intense than that dominated by the varied discharge policy. The Group II is designed to investigate the degradation characteristics dominated by the varied charge policy, as shown in Fig. S3 (c) to Fig. S3 (f) along with the Group I. Fig. S3 (c) and Fig. S3 (e) show the distributions of average discharge and charge capacity during the preliminary tests, clearly depicting the

inconsistency of the experimental batteries in terms of the capacity degradation. Fig. S3 (d) and Fig. S3 (f) are the overview of discharge and charge degradation trajectories of the batteries from Group I and Group II. It is seen that the shape of each trajectory in Group II is extremely fluctuant (up to 1 Ah) compared with that in Group I. Besides, as the capacity degradation of charge and discharge show a similar trend, we focus on discharge capacity in the discussion of battery degradation trajectory.

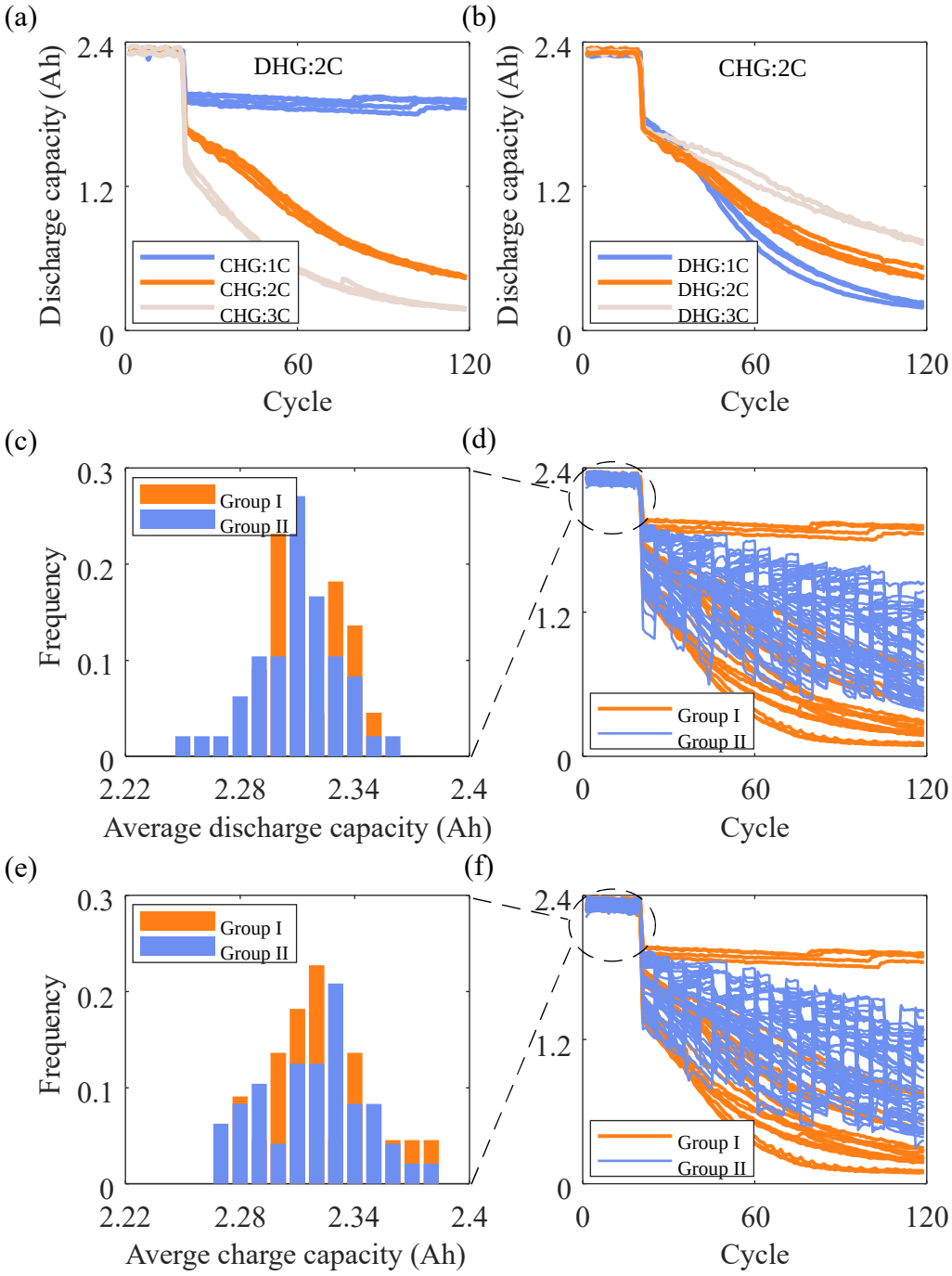

Fig. S3. Capacity degradation of the batteries in the experiments: (a) discharge capacity degradation of batteries charged at 1C, 2C, or 3C, and discharged at 2C; (b) discharge capacity degradation of batteries charged at 2C and discharged at 1C, 2C, or 3C; (c) distribution of average discharge capacity of the first 20 preliminary cycles; (d) Overview of discharge capacity degradation of Group I and Group II; (e) distribution of average charge capacity of the first 20 preliminary cycles; (f) Overview of charge capacity degradation of Group I and Group II.

Supplementary Note 3. Hyperparameters of the proposed framework.

Table. S2 Hyperparameters of the proposed framework

Key parameters of the networks
GRU layer with 256 hidden units ×2;
Fully connected layer with 32 hidden units ×2;
Fully connected layer with 2 hidden units ×1.

Supplementary Note 4. Key parameters of the existing four methods.

Table. S3 Key parameters of the existing four methods

Method	Key parameters
PF method	Model: double exponential model; Number of particles: 200; Covariance of the process noise: 0.00001×diag [1, 1, 1, 1]; Covariance of the observation noise: 0.01.
Box-Cox method	Coefficient of transformation: maximum likelihood estimation.
LSTM RNN method	LSTM layer with 100 hidden units ×1; Dropout layer with dropout factor of 0.2; LSTM layer with 50 hidden units ×1; Dropout layer with dropout factor of 0.2; Fully connected layer with 1 hidden unit ×1.
ARIMA-EMD method	Tolerance of sift relative: 0.2; Model of intrinsic mode functions: ARIMA(3,1,2); Model of residual: ARIMA(3,1,1).

Supplementary Note 5. Key parameters of the involved networks.

Table. S4 Key parameters of the involved networks

Framework	Key parameters of the frameworks
GRU framework (The proposed framework)	GRU layer with 256 hidden units ×2; Fully connected layer with 32 hidden units ×2; Fully connected layer with 2 hidden units ×1.
FC framework	Fully connected layer with 768 hidden units ×1; Tanh layer ×1; Fully connected layer with 768 hidden units ×1; Tanh layer ×1; Fully connected layer with 32 hidden units ×2; Fully connected layer with 2 hidden units ×1.
LSTM framework	LSTM layer with 256 hidden units ×2; Fully connected layer with 32 hidden units ×2; Fully connected layer with 2 hidden units ×1.

Supplementary Note 6. Diagram of a GRU cell.

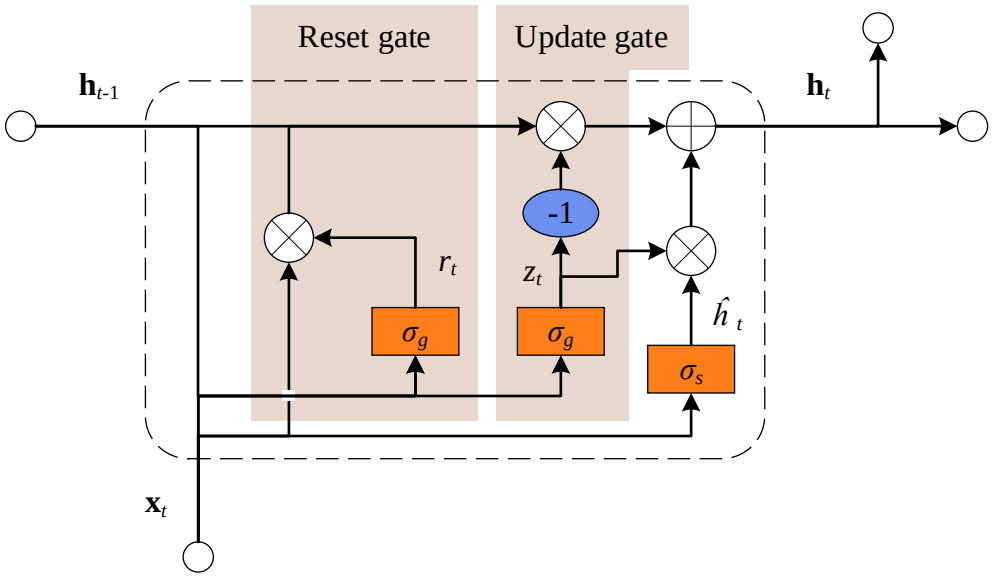

Fig. S4. Diagram of a GRU cell.

Reference

[1] Ma, S.-C. & Fan, Y. A deployment model of EV charging piles and its impact on EV promotion. Energy Policy 146, 111777 (2020).