



Research article

State of health estimation of LIB based on discharge section with multi-model combined

Peng Xu^a, Yuan Huang^{a,*}, Wenwen Ran^a, Shibin Wan^a, Cheng Guo^a, Xin Su^a, Libing Yuan^a, Yuanhong Dan^{b,c}

^a School of Electrical and Electronics Engineering, Chongqing University of Technology, Banan, Chongqing, 400054, China

^b School of Computer Science and Technology, Chongqing University of Technology, Banan, Chongqing, 400054, China

^c Nanjing University of Science and Technology, Xuanwu, Nanjing, China

ARTICLE INFO

Keywords:

Lithium-ion battery

State of health

Bidirectional long short-term memory network

Incremental capacity analysis

Differential thermal voltammetry analysis

Differential temperature analysis

ABSTRACT

Accurate estimation of a battery's state of health (SOH) is essential in battery management systems (BMS). This study considers a complete analysis of combining incremental capacity (IC), differential thermal voltammetry (DTV), and differential temperature (DT) for SOH prediction in cases of discharge. Initially, the IC, DTV, and DT curves were derived from the current, voltage, and temperature datasets, and these curves underwent smoothing through the application of Lowess and Gaussian techniques. Subsequently, discerning healthy features were identified within the domains where the curve exhibited substantial phase transitions. Utilizing Pearson correlation analysis, features exhibiting the utmost correlation with battery capacity degradation were singled out. Finally, the state-of-health (SOH) prediction model was constructed using a bidirectional long short-term memory (BILSTM) neural network. Two datasets were used to validate the model, and the experimental results demonstrated that the SOH prediction had a root mean square error (RMSE) below 1.2% and mean absolute error (MAE) below 1%, which verified the feasibility and accuracy. This approach quantifies the internal electrochemical reactions of a battery using externally measured data, further enabling early SOH predictions.

1. Introduction

With the increasing depletion of oil resources, the fuel car industry is facing challenges, and electric vehicles are becoming a favorable option. Electric vehicles, energy-storage devices, electronic devices, and many other industries cannot function without lithium-ion batteries. Lithium-ion batteries have disadvantages such as conductivity loss (CL), loss of active material (LAM), and loss of lithium inventory (LLI) during use [1]. To prevent the effects of battery deterioration, battery management system (BMS) monitoring assures that electric vehicles can operate safely and stably [2]. The BMS collects and manages not only the charging and discharging of individual cells in the battery package, but also the SOC and state of health (SOH) conditions of the battery [3]. Chemical reactions occur internally during battery cycling, and it is difficult to predict battery aging directly from these internal reactions. The SOH prediction is vital for battery health research and can be easily obtained from external battery characteristics, including the terminal voltage, current, and surface temperature of the cell. At present, research on battery health emphasizes short-term SOH predictions and long-term predictions of remaining useful life (RUL) prediction [4]. The battery SOH attains 70%–80% when it reaches the end of life

* Corresponding author.

E-mail address: huangyuan202102@163.com (Y. Huang).

<https://doi.org/10.1016/j.heliyon.2024.e25808>

Received 16 July 2023; Received in revised form 7 January 2024; Accepted 2 February 2024

Available online 9 February 2024

2405-8440/Â© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

(EOL), which indicates that it should be replaced [5].

Standard SOH prediction methods include modeling and data-driven methods. The modeling methods include electrochemical, empirical, and equivalent-circuit models. In electrochemical models, Zhang and White [6] proposed a method to build a single-particle model to simulate the mechanism changes inside the cell and characterize the solid electrolyte interface (SEI) layer growth. The electrochemical model involves many relevant characteristic parameters, further increasing the complexity of the process of building a perfect electrochemical model. Ramadass et al. [7] proposed to build a semi-empirical model to predict the decline of battery capacity. Although the empirical modeling approach is more efficient than electrochemical modeling, the algorithm accuracy is still affected by the battery model parameters, which makes it difficult to achieve accurate SOH estimation. The equivalent-circuit model must identify and confirm the internal parameters of the battery, which are complicated and not conducive to SOH prediction. Standard SOH prediction using data-driven methods includes statistical-based regression models, support vector regression models, and neural network models. The data-driven method relies on battery capacity-related data to predict the RUL by exploring the operational data, in which the feature extraction strategy can effectively combine electrochemical analysis techniques for SOH or RUL prediction. Lin et al. [8] used the acquired SOH history data as the input, and after determining the model parameters, they built an autoregressive model for current lithium-ion battery SOH estimation. The advantage of this model is that the model parameters are small and easy to determine, and the data volume requirement is not large, which makes it suitable for estimating the historical SOH values of lithium batteries. It is challenging to analyze the implied relationships between multiple input features because of the small amount of model data, further resulting in low accuracy of the estimation results. Guo et al. [9] proposed an integrated support vector regression model that considers the inconsistency of the battery. The advantage of the support vector regression model is that a high accuracy can be maintained even when the size of the sample is small. However, it is difficult to obtain the kernel function for the support vector regression model directly. To scientifically analyze the battery aging data, feature characteristics, and other factors, Zhang et al. [10] proposed Elman network applications that are suitable for small sample situations, and Wu et al. [11] constructed radial basis function neural networks to improve their prediction accuracy. The advantage of the neural network model is that it can solve the diverse, complex, and nonlinear mapping relationships between lithium battery feature vectors and capacity with high accuracy and good adaptability of model estimation.

The use of sensors to measure voltage, current, and temperature data and make health-state predictions without damaging the battery is a popular research topic today. Zhang et al. [12] combined the IC curve and support vector regression model for SOH estimation; this method controls the mean absolute error (MAE) and root mean square error (RMSE) below 2.3%. Li et al. [13] combined the IC curve with a Gaussian regression model to predict the health state of a battery by probability. To suppress the effect of noise, Guo et al. [14] reconstructed the charging voltage curve and used the SVR model for prediction. Zhang et al. [15] divided the IC curve of the charging part into three regions to comprehensively predict the curve. Xu et al. [16] used IC curves as neural network inputs and a CNN for feature value extraction, by combining a CNN and attention mechanisms based on LSTM neural networks. Gismero et al. [17] analyzed the aging of batteries used in the BMW i3 and used the ICA technique for SOH prediction, where the RMSE error was below 2%, thus confirming the effectiveness of the ICA technique in commercial vehicles. Kwon et al. [18] for the first time used a charging voltage curve to extract feature data and combined it with the ICA technique to make predictions using an LSTM neural network. The abovementioned scholars analyzed the IC curves for charging time regions, and for extracting more accurate ICA data, they used low C-rate charging such as C/25 and C/10 for feature extraction. Fly and Chen [19] discussed the trend of ICA peaks while charging C/6 and C/24, and fewer peaks were obtained during 1C charging. In practical applications, fast-charging and high-multiplier charging situations must be considered. The observable peaks decreased during fast charging, and Zhou et al. [20] estimated the SOH of the area generated during fast charging as a new feature input. Ospina Agudelo et al. [21] discussed the case of partial peak area of the IC curve as a health indicator for the 1C case. High C-rate charging can still be used as an indicator of health status, although it leads to a reduction in the peak value. The DTV method, initially introduced by Wu et al. [22], serves as a battery status monitoring technique involving voltage and temperature measurements. Wang et al. [23] further investigated the DTV curve, extracting eigenvalues and integrating them with a Gaussian regression model to predict State of Health (SOH), achieving a maximum relative error below 2%. Yang et al. [24] employed DTV curves to delineate voltage intervals. With the battery charging and discharging cycles, an aging trend was observed in the internal temperature change of the battery. A neural network model incorporating temperature prediction of SOH was validated by a Bayesian optimization model with three sets of data; the dataset MAE was below 5% [25]. The feature parameters extracted from the DT curves were used to predict the SOH in 11 mainstream estimation methods, and all the estimated RMSEs were maintained below 2.5% [26]. Chen et al. [27] obtained the DT curve through the battery's internal temperature-to-time differential, followed by a correlation analysis to find the feature value with the highest correlation to capacity, combined with gated recurrent unitary neural networks to estimate the SOH, and the error was controlled within 2.28%.

To avoid the overreliance of the model on a single feature source, a combination of feature sources was considered for a comprehensive prediction of the SOH. There is a decrease in the IC peak during high C-rate discharge by considering the peak of the IC curve during discharge alone as a feature source. This may lead to inaccurate modeling, and the impact of temperature changes inside the battery could be considered on the basis of capacity changes, as ICA technology is commonly used in charging situations. A mechanistic model based on differential voltage curves, such as IC and DV curves, can be implemented in a BMS that controls three parameters: voltage, current, and temperature [28]. The characteristic data for predicting battery aging consisted of peak and valley data. The characteristic values that correlated with the battery capacity in the range of 80%–100% were selected using Pearson correlation analysis. The battery aging process has time-series characteristics and long-term dependence. Therefore, time-series neural networks with gating mechanisms such as LSTM are suitable for SOH prediction. Compared to the LSTM neural network, the BILSTM neural network can analyze the data from front to back together, which offers a more comprehensive analysis. Bharat et al. [29] discussed BILSTM with LSTM, CNN-LSTM, and CNN-BILSTM neural network models for short term residential load prediction and

3

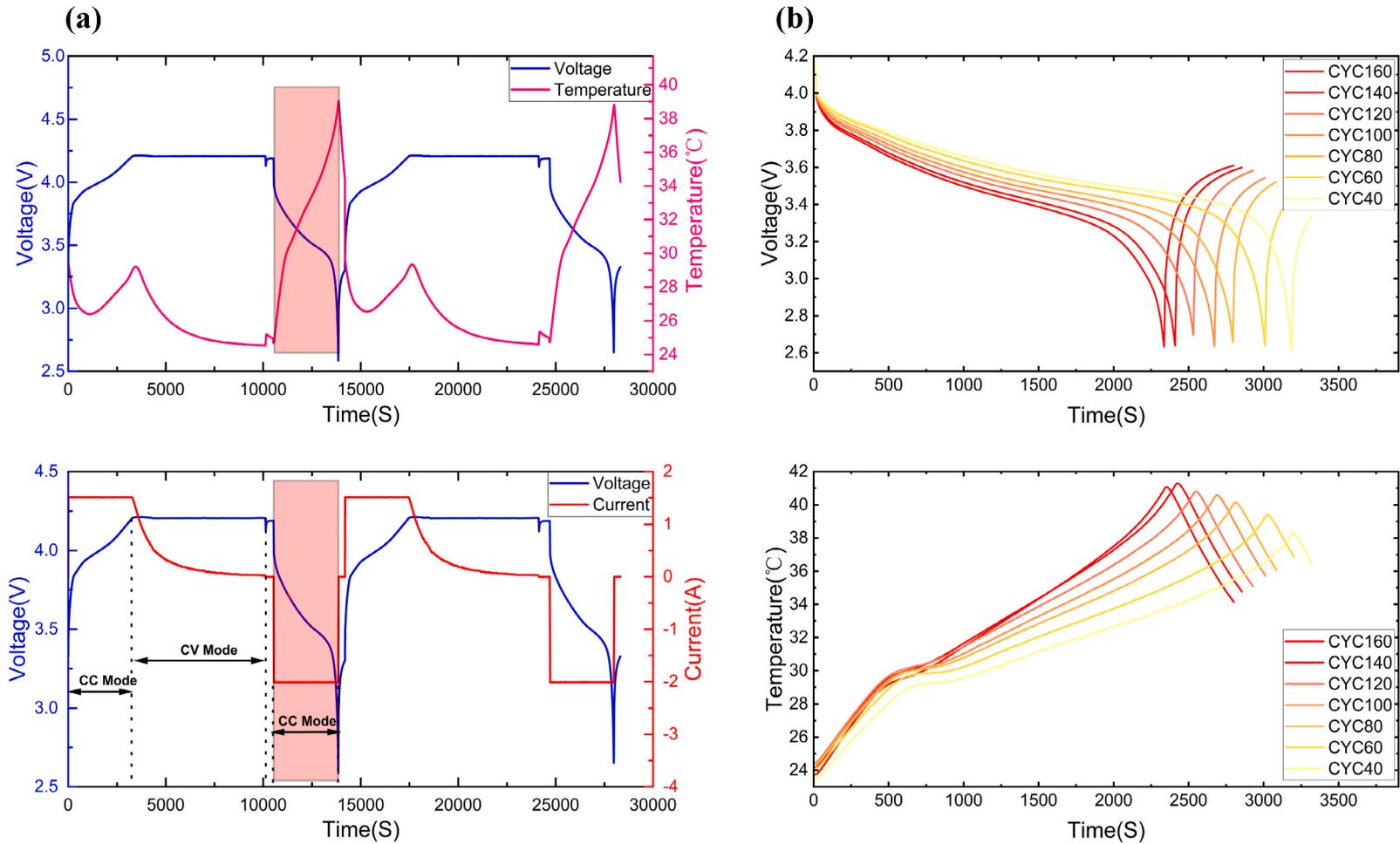


Fig. 1. Variation of voltage, current, and temperature (a) Voltage, current and temperature transformation in two cycles. (b) Voltage and temperature change under different cycles.

obtained the lowest RMSE prediction using the BILSTM model. Zhang et al. [30] compared the prediction time errors of BILSTM and LSTM. The results demonstrated that the BILSTM neural network was more advantageous in extracting time-series data. Panagiotis et al. [31] predicted the battery capacity decline phenomenon under different operating conditions and compared the LSTM, BILSTM, GRU, and BIGRU algorithms; BILSTM had the smallest MAE and RMSE, thus confirming that the BILSTM algorithm is suitable for SOH prediction. The main contributions of this study are as follows:

- (a) The ICA technique is commonly used for low C-rate charging cases. To verify that it is still applicable for high C-rate discharging cases, 1C discharging data are used for feature processing and extraction.
- (b) By combining IC, DTV, and DT technologies, the integrated sensor-measurable quantities captured the internal chemical changes of the battery and predicted the aging of the battery more comprehensively.
- (c) A data-driven approach was used to perform the SOH prediction for both datasets.

This study characterized the discharge part of the data using NASA and Oxford datasets, combined with a data-driven approach to predict the SOH. The remainder of this paper is organized as follows: Section 2 describes the computation of IC, DTV, and DT curves, compares different sampling intervals and filtering methods, and extracts features from peaks and troughs with significant phase changes. Section 3 extracts quantities that are highly correlated with battery capacity as eigenvalues from the Pearson correlation analysis. Section 4 describes the BILSTM network structure and compares the SOH prediction results of the BP, LSTM, and BILSTM neural-network models. Finally, section 5 concludes the study.

1.1. Data analysis

1.1.1. Data collection

The NASA Ames Prognostics Center of Excellence (PCoE) published experimental data on the cycle life of lithium-ion batteries [32]. This dataset collects test data of 18,650 Li-ion batteries for charging, discharging, impedance, and temperature at 24 °C. This study analyzed the data of batteries B05, B06, and B07. These three batteries were charged in the CC-CV mode and discharged in the CC mode. The charging process consisted of charging to 4.2 V at a constant current of 1.5 A and then charging at a constant voltage of 4.2 V until the current was reduced to a 20 mA cut-off current. The discharge process selected 2 A constant current discharge. The battery discharge cutoff voltages were 2.7 V, 2.5 V, and 2.2 V, which collected 168 cycles of charge and discharge data for this dataset.

Many researchers have used predictable data trends during battery cyclic charging to estimate the battery SOH. Voltage and current trends as well as the battery's internal temperature were observed during the discharging process. Fig. 1 (a) depicts the relationship between the voltage and current, voltage, and temperature of the battery during the cycle. The pink plot window indicates the discharge. Fig. 1 (b) depicts the trends in the discharge voltage and temperature of B05 battery at the 40th, 60th, 80th, 100th, 120th, 140th, and 160th cycles. Aging of the battery results in a shorter time to reach the battery cut-off voltage and maximum temperature, which is due to the shuttling of the internal lithium ions between the two poles during the battery cycle, and the CL, LAM, and LLI phenomena that occur during this process [33].

The battery dataset published by the University of Oxford provided an 8-channel battery tester for commercial Kokam pouch batteries with a rated capacity of 740 mAh [34]. The batteries had a cut-off charging voltage of 4.2 V and a cut-off discharging voltage of 2.7 V. The 1C charge/discharge test was carried out at 40 °C. The tests were conducted using an 8-channel battery tester. The sampling interval was 1 Hz. Under the simulation of Artemis urban discharge conditions, battery data were recorded for every 100 cycles. The data descriptions of the two batteries are listed in Table 1.

1.1.2. Feature extraction

The IC curve can effectively quantify the internal degradation. When the battery was discharged in the CC model, a voltage interval decreased as the Li ions moved toward the end of electrodes. Within this interval, the peaks in the IC curve were generated as a result of the increase in power discharge. The DTV method can be used to track the phase changes that occur in the electrode and the resulting entropic heat. The aging process can be reflected by correlating the internal chemical reactions with the external measurements of the battery [35]. Lithium ions migrated from the anode to cathode during battery discharge. The charge imbalance between the cathode and anode was strong such that it generated approximately 4.2 V. The charge imbalance gradually decreased with the gradual migration of lithium ions. During migration, the internal temperature of the battery gradually increased until it reached the battery cut-off voltage. Subsequently, the internal temperature gradually decreased during the battery setting process. The battery degradation process was predicted in connection with the external data, and thus, the IC, DTV, and DT curves were used as important evaluation targets of the battery health state, and equations (1)–(5) to find the three curves.

Table 1
Description of the battery dataset.

Dataset	Charging cut-off voltage [V]	Charging cut-off current [mA]	Charging constant current [C]	Discharge cut-off voltage [V]	Discharge constant current [C]	Temperature [°C]
NASA	4.2	20	0.75	2.7	1	24
Oxford	4.2	–	2	2.7	1	40

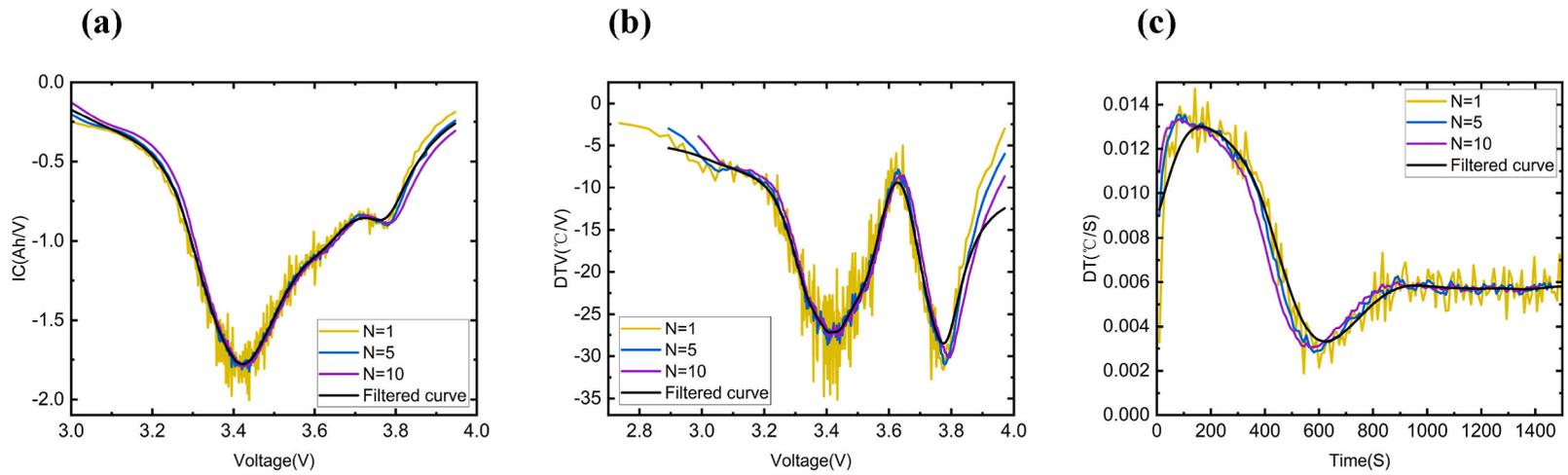


Fig. 2. NASA discharge data curve at different sampling intervals and filtering (a) IC curve at different sampling intervals and filtering. (b) DTV curve at different sampling intervals and filtering. (c) DT curve at different sampling intervals and filtering.

$$Q = \int I dt \tag{1}$$

$$IC = \frac{dQ}{dV} = \frac{I \cdot dt}{dV} = I \cdot \frac{dt}{dV} \tag{2}$$

$$IC = \frac{Q_K - Q_{KN}}{V_K - V_{KN}} = \frac{I_K t_K - I_{KN} t_{KN}}{V_K - V_{KN}} \tag{3}$$

$$DTV = \frac{\Delta T}{\Delta V} = \frac{T_K - T_{KN}}{V_K - V_{KN}} \tag{4}$$

$$DT = \frac{\Delta T}{\Delta t} = \frac{T_K - T_{KN}}{t_K - t_{KN}} \tag{5}$$

Where N is the sampling interval, K is a certain moment, and Q is the battery capacity, which is obtained by integrating the measurable amount of current I as shown in equation (2). The DTV curve can be obtained as mentioned in equation (4). The chemical reaction of the battery varied the internal temperature. The DT curve can be obtained by differentiating the internal temperature variation of the battery. Although the sampling interval $N = 1$ has noise, it can be filtered by increasing the sampling interval. In Fig. 2 (a), although the IC curve gradually smooths out as the sampling interval N increases, the noise is still present. The IC curve was smoothed via Lowess filtering. This filter method fits a polynomial regression curve in the subset by taking a certain percentage of local data, and the curve's trend can be observed. The trend of the curve can be observed in Fig. 3 (a). Both DTV and DT curves were smoothed using Gaussian filtering to obtain smooth curves as depicted in Fig. 2 (b), (c).

For the NASA dataset, as the number of battery cycles increased, there was a leftward shift and an increase in the peak of the IC curve, as depicted in Fig. 3 (a). The peak rise is related to LAM, and the peak shift to the left is related to CL and LLI [36]. Therefore, the IC curve can be used to visualize the internal chemical reactions of the lithium-ion battery during cycling. As depicted in Fig. 4 (a), the

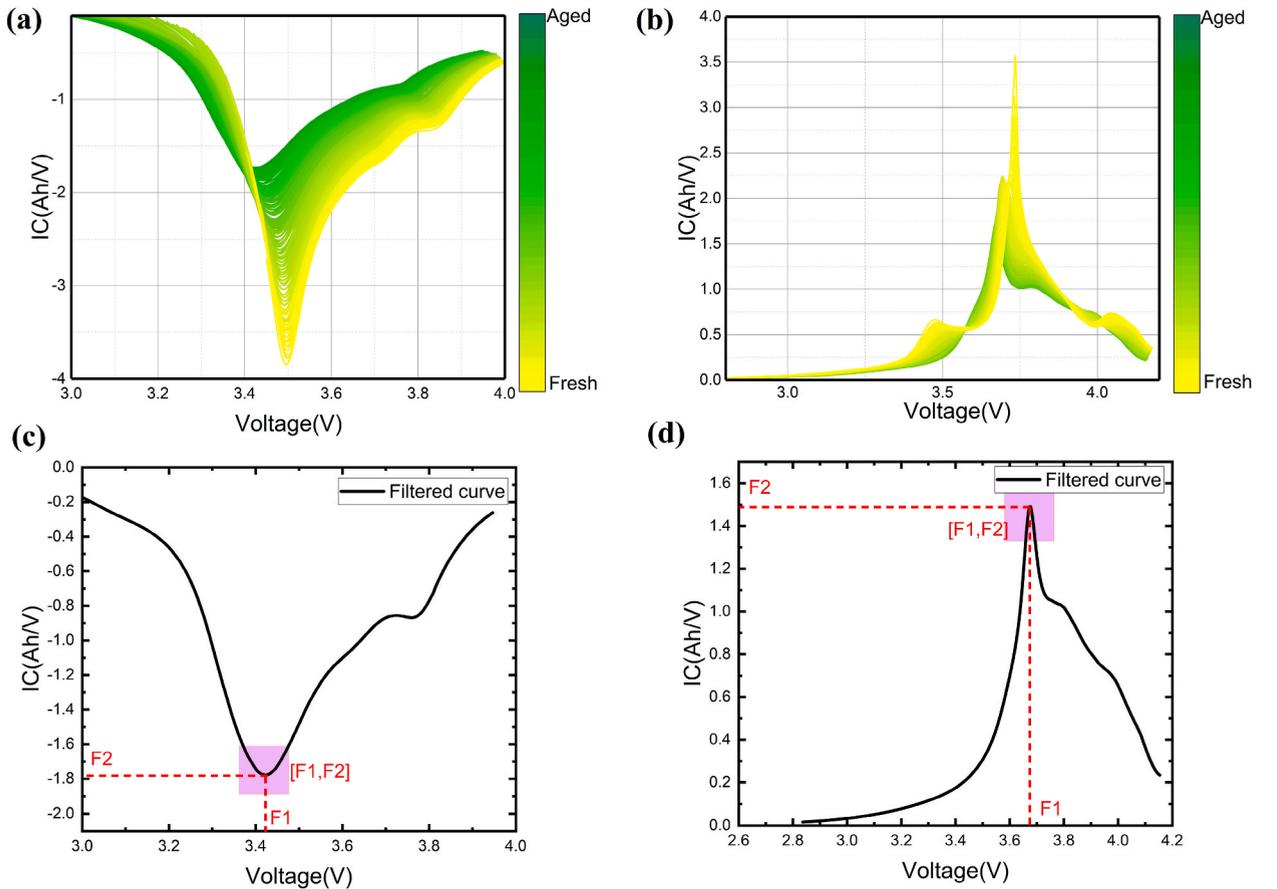


Fig. 3. Discharge IC curve acquisition process (a) Trends in IC curves for NASA datasets. (b) Trends in IC curves for Oxford datasets. (c) Features selection for NASA datasets. (d) Features selection for Oxford datasets.

DTV curve tends to shrink and shift. There was an entropy change inside the battery, and the DTV curve characterized the entropy change in the battery through inflection points. Therefore, the cell SOH can be predicted using these phase change characteristics [37]. The battery phase change process requires energy to drive, which leads to a change in the temperature [38], therefore, the peaks and valleys of the DTV curve were used as characteristic data. The DT curve also exhibited an offset and rise, as depicted in Fig. 5 (a). The peaks and valleys of the three curves are used as important indicators for battery aging prediction. The IC peaks were used as the first group of characteristic data [F1–F2] as depicted in Fig. 3 (c); the peaks and valleys of the DTV curves were used as the second to fourth groups of characteristic data [F3–F8] as depicted in Fig. 4 (c); and the peaks and valleys of the DT curves were used as the fifth to sixth groups of characteristic data [F9–F12], as depicted in Fig. 5 (c).

For the Oxford dataset, the IC, DTV, and DT curve trends are depicted in Fig. 3 (b), 4 (b), and 5 (b), respectively. The peaks and valleys were used as feature values, where the IC peaks were used as the first set of feature data [F1–F2], as depicted in Fig. 3 (d). The characteristic values of the DTV curves were [F3–F8], where the second peak indicated a decreasing trend with increasing number of cycles, that varied significantly in the range of 3.85–4.10 V, as depicted in Fig. 4 (d). The DT curve depicted in Fig. 5 (d) was used as the third feature set, to extract the feature data [F9–F14].

1.2. Pearson correlation coefficient analysis

The features extracted from the IC, DTV, and DT curves were closely related to battery aging. The Pearson correlation analysis method determines the correlation using the ratio of covariance to the standard deviation of two random variables, as shown in equation (6). When the absolute value of the correlation coefficient was between 80% and 100%, there was a strong correlation between these two variables.

$$\rho_{X,Y} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}} = \frac{E((X - EX)(Y - EY))}{\sqrt{D(X)}\sqrt{D(Y)}} \quad (6)$$

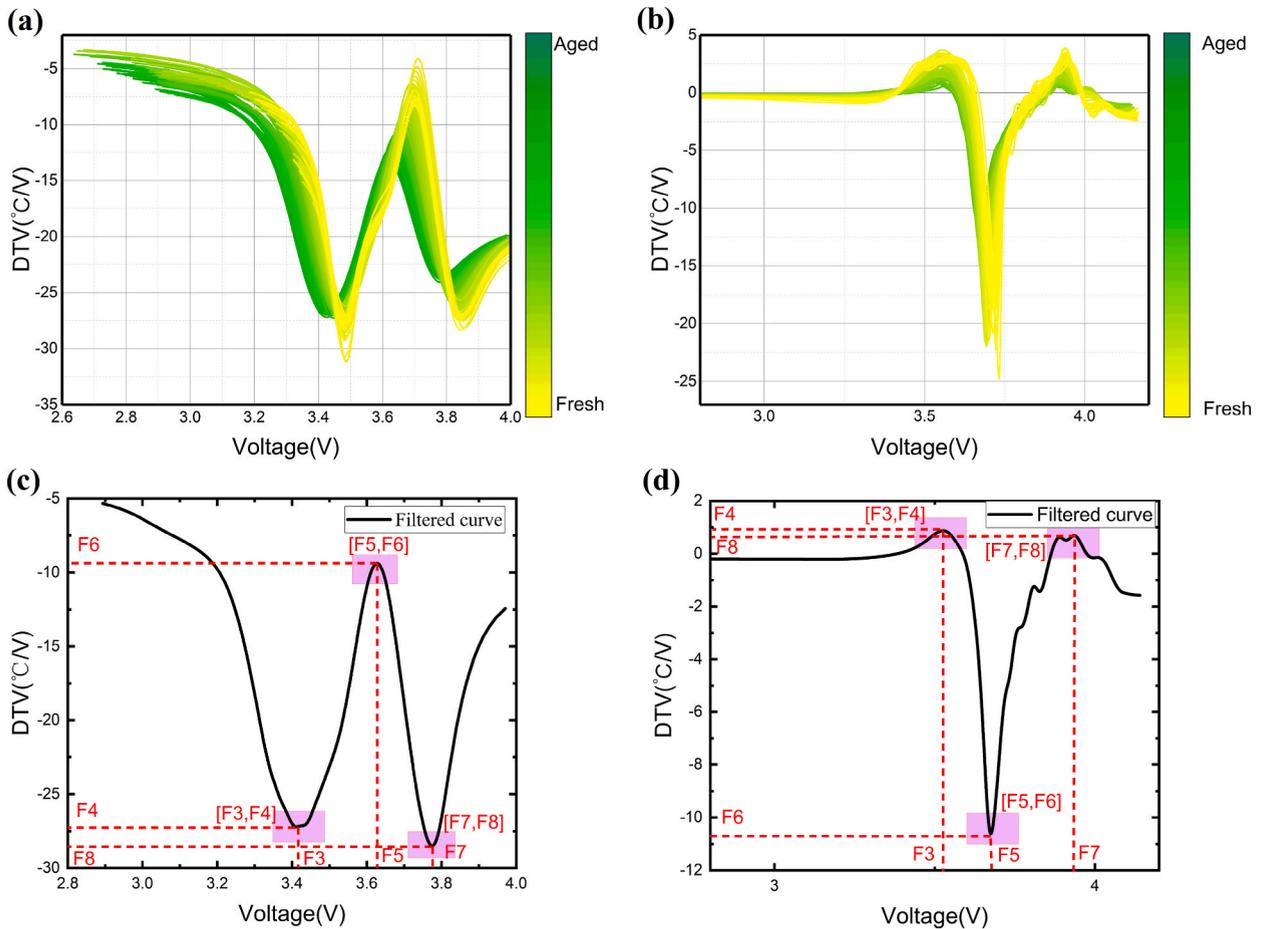


Fig. 4. Discharge DTV curve acquisition process (a) Trends in DTV curves for NASA datasets. (b) Trends in DTV curves for Oxford datasets. (c) Features selection for NASA datasets. (d) Features selection for Oxford datasets.

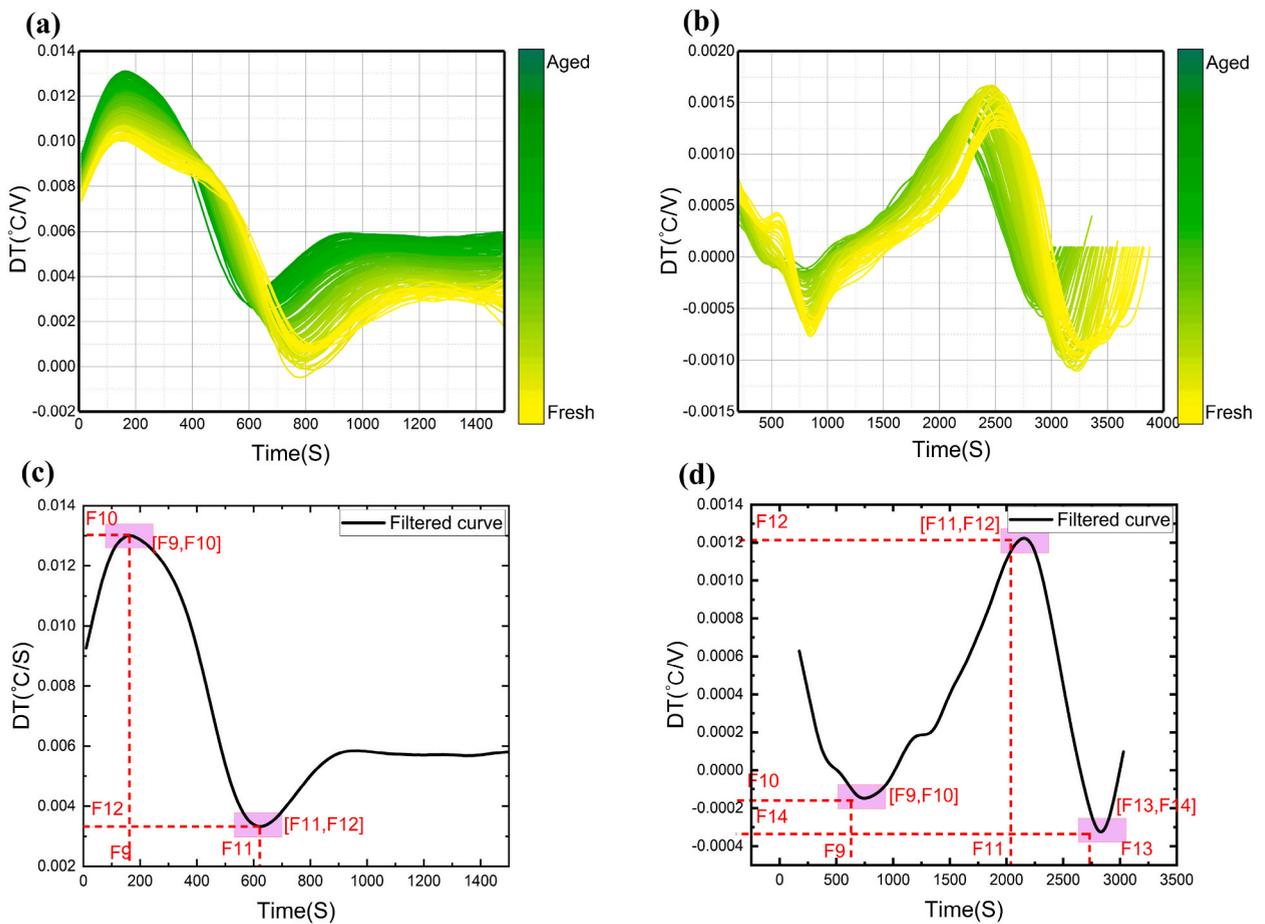


Fig. 5. Discharge DT curve acquisition process (a) Trends in DT curves for NASA datasets. (b) Trends in DT curves for Oxford datasets. (c) Features selection for NASA datasets. (d) Features selection for Oxford datasets.

Where X and Y are any two features, and the correlation is determined by the Pearson ratio value. The correlation analysis for cells B05, B06, and B07 in the NASA dataset and cells Cell1, Cell3, and Cell4 in the Oxford dataset are listed in Table 2.

Fig. 6 depicts the correlation analysis of 12 sets of feature numbers with a battery capacity of for battery B05, where ten sets of feature numbers F1, F2, F3, F5, F6, F7, F8, F10, F11, and F12 were highly correlated with battery capacity. These ten sets of feature data were used as the input of the training set for further SOH prediction.

By considering the NASA dataset B05 as an example, the aging trend of eigenvalue F4 of the first peak and valley of the DTV curve was not evident as depicted in Fig. 7 (a) and Fig. 7 (b), and the trend of F4 is smoother, which is not suitable for its use as an aging indicator. In the DT curve of B05, eigenvalue F9 was concentrated in the range of 150–160 S, and the variation was not prominent, as depicted in Fig. 5 (a), which is poorly correlated with the capacity. In addition, the other characteristic values exhibited significant trends with battery aging and were positively or negatively correlated with capacity degradation. This is favorable for the SOH prediction of batteries. The selection of eigenvalues depended on the trend of the characteristic curve changes, and the quantity with an evident aging trend was selected as the eigenvalue.

Table 2
Battery capacity correlation analysis.

Factor	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14
B05	0.96	-0.99	0.94	-0.71	0.99	0.95	0.98	-0.98	-0.64	-0.98	0.99	-0.96	-	-
B06	0.99	-0.96	0.99	0.18	0.98	0.96	0.96	-0.95	-0.20	-0.99	0.99	-0.95	-	-
B07	0.97	-0.99	0.95	-0.65	0.97	0.98	0.96	-0.93	-0.20	-0.97	0.99	-0.90	-	-
Cell1	0.96	0.94	0.62	0.88	0.98	-0.86	0.88	0.91	0.78	-0.95	0.97	0.33	0.96	-0.81
Cell3	0.95	0.93	0.44	0.88	0.97	-0.8	0.87	0.88	0.77	-0.96	0.97	0.32	0.96	-0.83
Cell4	0.96	0.85	0.68	0.83	0.98	-0.86	0.85	0.90	0.79	-0.92	0.95	0.42	0.99	-0.81

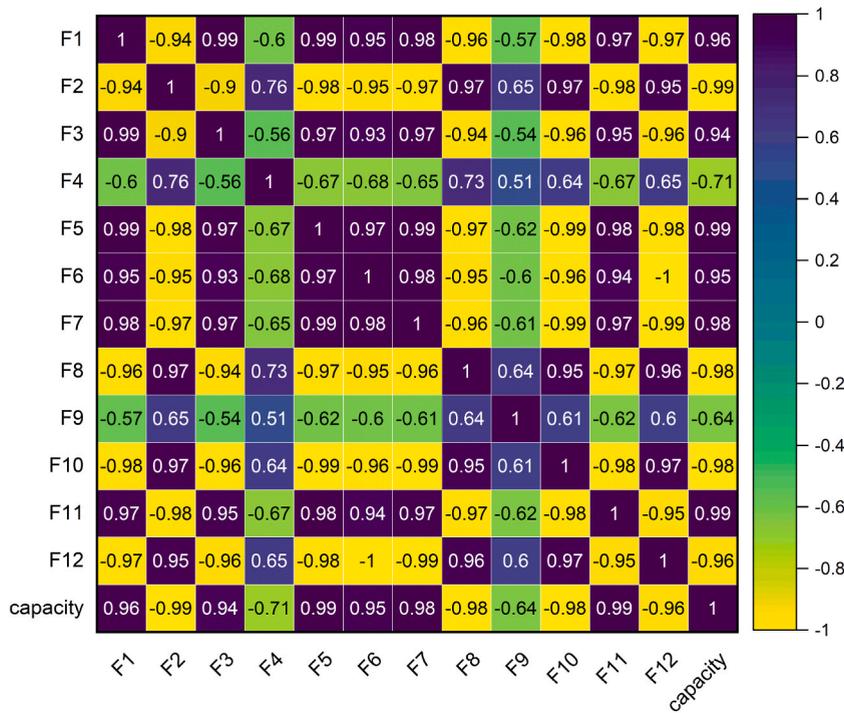


Fig. 6. Pearson correlation coefficient analysis of B05.

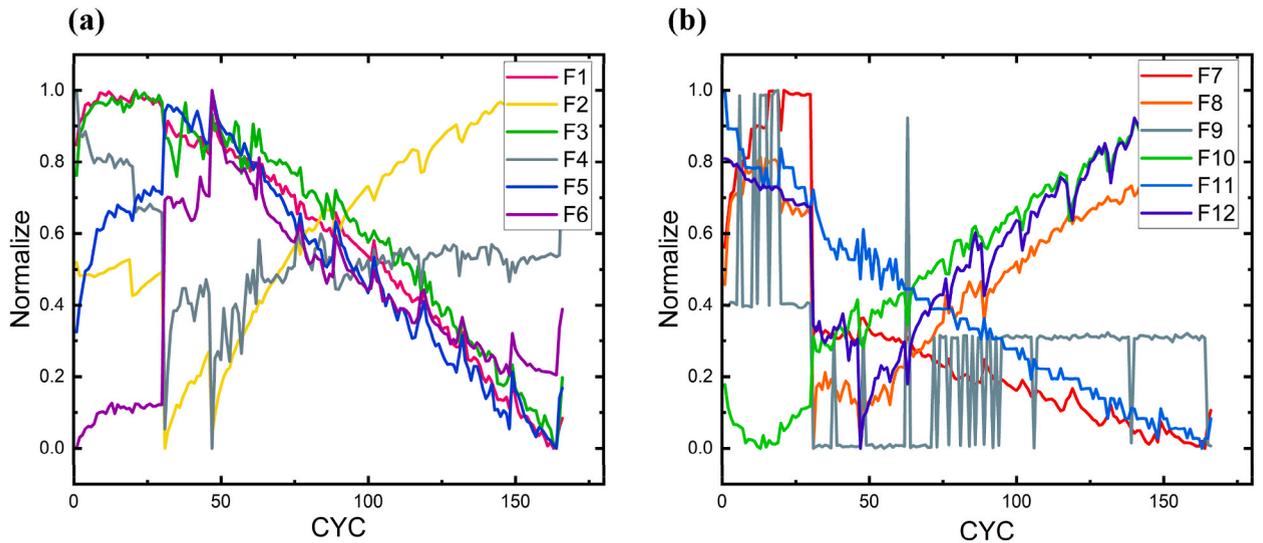


Fig. 7. B05 Changes in features (a) Aging trends in [F1–F6]. (b) Aging trends in [F7–F12].

2. Method

2.1. BILSTM structure

Compared to traditional recurrent neural networks (RNN), LSTM has a gating mechanism. An RNN only has short-term memory capability, which makes it challenging to handle time-series data; moreover, the presence of gradient disappearance and gradient explosion problems occur during backpropagation, which may cause the neural network to fail to learn. The LSTM neural network can effectively solve the gradient disappearance and gradient explosion problems [39]. The LSTM adds a new internal state and gating mechanism based on recurrent networks.

The BiLSTM neural network adds a reverse layer to the LSTM layer, as depicted in Fig. 8. Compared to LSTM, which only analyzes previous information, BiLSTM has a higher prediction capability by combining previous and subsequent information. The prediction results of y were obtained by combining the forward and reverse hidden layers. When the output was at t moment, the forward LSTM layer analyzed the information to moment $t - 1$, whereas the reverse LSTM layer analyzed the information to moment $t + 1$ as shown in the following equations (7)–(9):

$$\vec{h}_t = \sigma(\omega_{hx} x_t + \omega_{hh} h_t + b_{h\rightarrow}) \tag{7}$$

$$\bar{h}_t = \sigma(\omega_{hx} x_t + \omega_{hh} h_t + b_{h\leftarrow}) \tag{8}$$

$$y_t = \omega_{hy} \vec{h}_t + \omega_{hy} \bar{h}_t + b_y \tag{9}$$

Where $x = (x_{t-1}, x_t, x_{t+1})$ denotes the input vector at different moments, $\vec{h} = (\vec{h}_{t-1}, \vec{h}_t, \vec{h}_{t+1})$ is the hidden layer of the forward LSTM layer, $\bar{h} = (\bar{h}_{t-1}, \bar{h}_t, \bar{h}_{t+1})$ is the hidden layer state of the inverse LSTM, ω is the weight of the LSTM layer, and b is the bias of the LSTM layer. The forward and reverse outputs are combined to obtain the output $y = (y_{t-1}, y_t, y_{t+1})$. A BiLSTM neural network combined with the dropout technique was used to effectively prevent overfitting. The structure of SOH prediction is depicted in Fig. 9.

2.2. Hyper-parameter selection

The SOH prediction is strongly dependent on the structure of the neural network, and the selection of hyper-parameters is crucial. The process of hyper-parameter selection for the BiLSTM neural network is listed in Table 2. To evaluate the performance of the model, the RMSE and MAE were used as the evaluation metrics of the algorithm performance. These evaluation indicators are defined by equations (10) and (11).

$$RMSE = \sqrt{\frac{1}{N} \left(\sum_{i=1}^N |\hat{y}_i - y_i| \right)} \tag{10}$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \tag{11}$$

Here, \hat{y}_i represents the SOH prediction results, y_i represents the true SOH value, and N represents the total number of predictions.

Table 3 uses 70% of the data set as the test set and 30% as the validation set. To determine the optimal hyperparameter setting, the hidden layer was first set to one layer compared to two layers, and the error was smaller in one layer. While keeping the other hyperparameters constant, we obtained more minor errors by adding epochs, as shown in the results of the second to sixth tests. The change in the number of hidden layers also affected the error results; thus increasing the dropout rate to 0.001 at 120 neurons, which in turn caused the error to increase. In the ninth test, the neurons, batch size, dropout rate, epoch, and learning rate were set to 145, 15, 0.0001, 800, and 0.001, and the error was minimized such that the RMSE and MAE were 0.47% and 0.36%, respectively. Modifying the dropout rate and batch size led to the deterioration of the model performance, as in the 10th-15th times. The model prediction exhibited poor performance with an increase in the number of hidden layer neurons and layers.

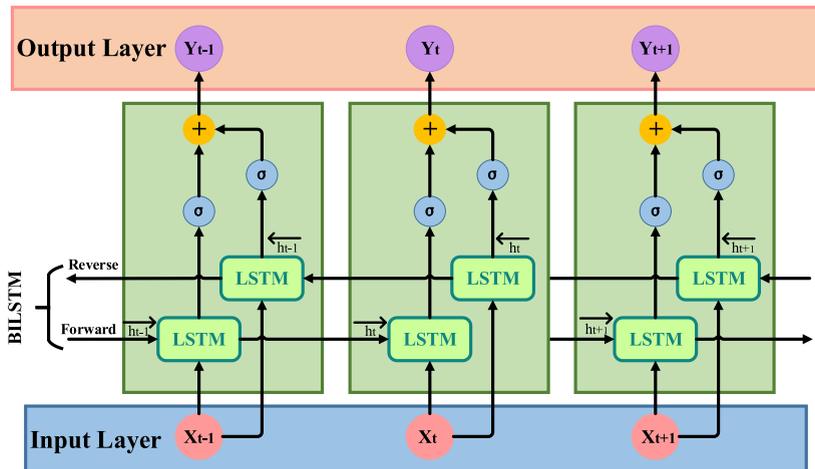


Fig. 8. BiLSTM structure.

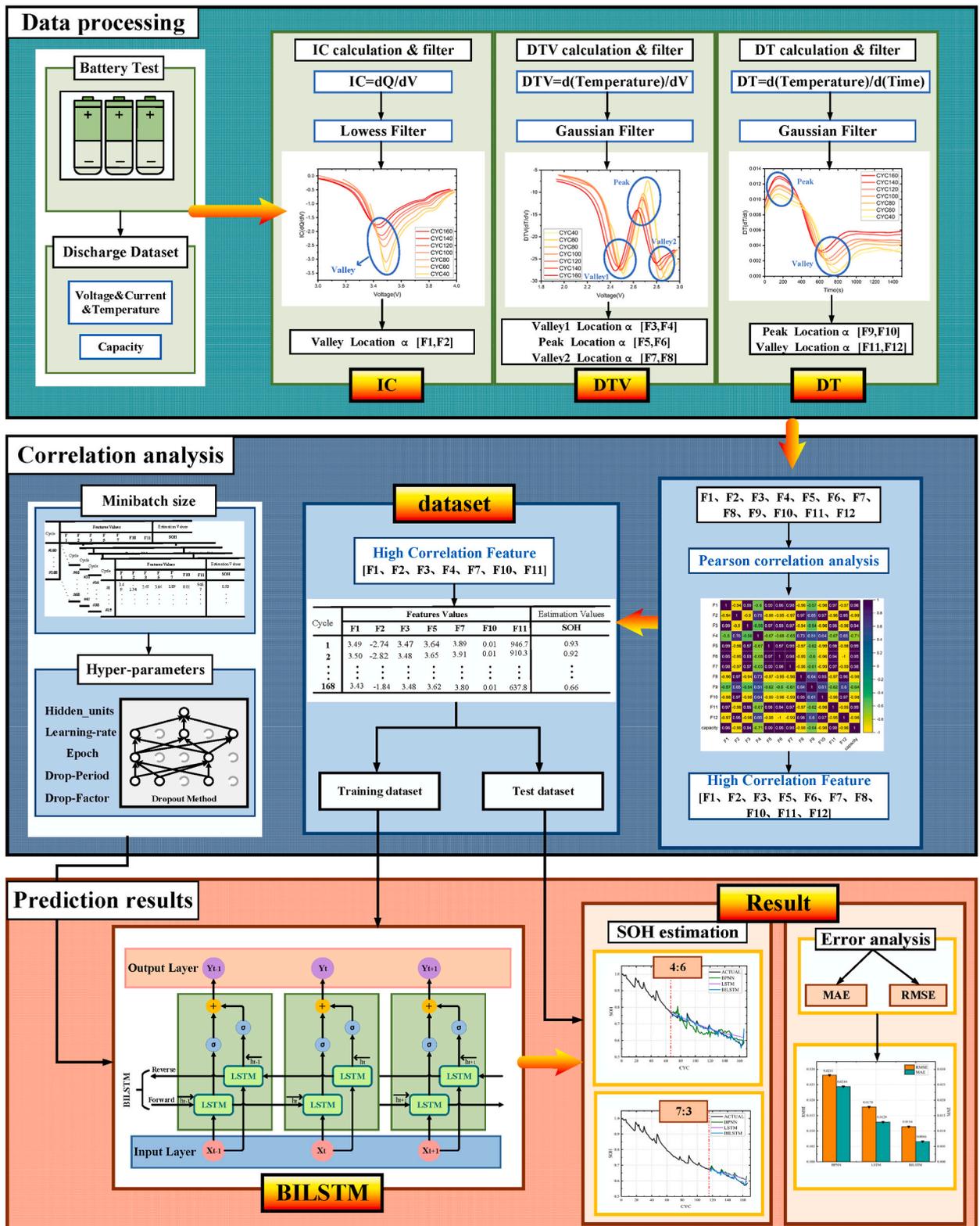


Fig. 9. SOH prediction structure diagram.

Table 3
Model and hyper-parameter selection.

Number	Hyper-parameter Selection						Evaluation Indicators	
	Hidden layer	Neurons	Batch	Dropout rate	Epoch	Learning rate	RMSE (%)	MAE (%)
1	2	110/10	15	0.0001	100	0.001	5.12	4.20
2	1	110	15	0.0001	100	0.001	1.39	0.85
3	1	110	15	0.0001	200	0.001	1.36	1.21
4	1	110	15	0.0001	300	0.001	1.33	1.19
5	1	110	15	0.0001	500	0.001	0.71	0.51
6	1	110	15	0.0001	700	0.001	0.62	0.48
7	1	120	15	0.0001	800	0.001	0.56	0.42
8	1	120	15	0.001	800	0.001	0.83	0.69
9	1	145	15	0.0001	800	0.001	0.47	0.36
10	1	145	15	0.001	800	0.001	0.92	0.79
11	1	145	15	0.005	800	0.001	0.67	0.43
12	1	145	15	0.01	800	0.001	0.75	0.62
13	1	145	20	0.0001	800	0.001	1.65	1.42
14	1	145	40	0.0001	800	0.001	0.82	0.65
15	1	145	60	0.0001	800	0.001	0.91	0.76
16	1	150	15	0.001	1000	0.001	0.58	0.41
17	1	150	15	0.0001	1000	0.0001	7.71	7.63
18	2	110/10	15	0.0001	800	0.001	1.37	0.72
19	3	110/10/10	15	0.0001	800	0.001	0.92	0.79
20	1	100	15	0.0001	800	0.001	0.86	0.59

2.3. Predicted results

The NASA and Oxford datasets were also used for prediction. The NASA dataset 40% was used as the training set, and 60% of the data was predicted, as depicted in Fig. 10 (a), (b), (c); the prediction errors for battery B05, B06, and B07 are shown in Fig. 10 (d), (e), (f); the BILSTM neural network prediction results were better than those of the BP and LSTM neural network predictions shown in Fig. 10 (g), (h), (i). Because of the different capacity regeneration phenomena of each battery dataset, the learning rate, training batch size, maximum number of iterations, and learning rate decay depend on manual adjustments, further resulting in different prediction results for each battery set. The RMSEs of the BILSTM neural network in the SOH prediction results were all less than 1.2%, and the MAE was less than 1%, which proves that the BILSTM neural network model is suitable for battery life prediction. The SOH prediction results were more accurate when 70% of the data were trained and 30% were predicted, as depicted in Fig. 11 (a), (b), (c); the prediction errors are shown in Fig. 11 (d), (e), (f). As the size of the training set increased, the model's performance improved. The RMSE and MAE of the BILSTM prediction results were below 1% shown as Fig. 11 (g), (h), (i), which proves the reliability of the method. To verify the reliability of the method, the Oxford dataset was trained with 40% and 70%, as depicted in Fig. 12 (a), (b), (c) and 13 (a), (b), (c). The BILSTM neural network predicts an RMSE and MAE of less than 1%.

The Oxford dataset exhibits better predictive performance compared to the NASA dataset. Shown as Fig. 12 (d), (e), (f), and Fig. 13 (d), (e), (f). In the case of the Oxford dataset, the phenomenon of capacity regeneration is relatively minimal, making predictions more straightforward. The BP neural network relies on the gradient descent method, which can result in jagged prediction results and subpar predictive performance. On the other hand, the BILSTM neural network considers information from both before and after the data points, leading to higher prediction accuracy compared to the LSTM neural network.

3. Conclusion

This paper proposes a prediction model that combines a data-driven method with IC, DTV, and DT curve analyses. These curves were obtained from the measurable external battery data. The effects of internal battery reactions and battery aging on capacity degradation were comprehensively considered during discharge. This study was based on the NASA and Oxford discharge aging datasets. The IC curves selected the Lowess method for filtering, DT curves selected the Gaussian method for filtering, and based on the DT curve, the DTV curve obtained by differenced voltage was analyzed via Gaussian filtering. Subsequently, the data with prominent features, such as peaks and valleys, were used as health feature variables. Characteristic health variables with a correlation greater than 0.8 were selected, and these characteristics strongly correlated with the battery capacity variation. Based on the LSTM neural network, a BILSTM neural network combined with the dropout technique was selected, and the model performance improved as the size of the training set increased. As listed in Table 4, the MAE was less than 1%, and RMSE was less than 1.2% for different training sets. This indicated that the model was highly reliable. This study considered the effects of the current, voltage, and temperature change processes in the discharge section on battery aging. Future research will help the cloud computing platforms to enable the real-time prediction of the battery's health status.

In this paper, CC discharge data is employed to predict battery capacity degradation. In practical battery usage, the battery undergoes random cycling between various SOCs, which results in inconsistent voltage plateaus during the discharge process and contributes to the irregular trends in characteristic curves. Notably, the NASA random walk dataset employs the pulse discharge method, making it challenging to capture the voltage plateau during the discharge process and, consequently, to extract the IC curve.

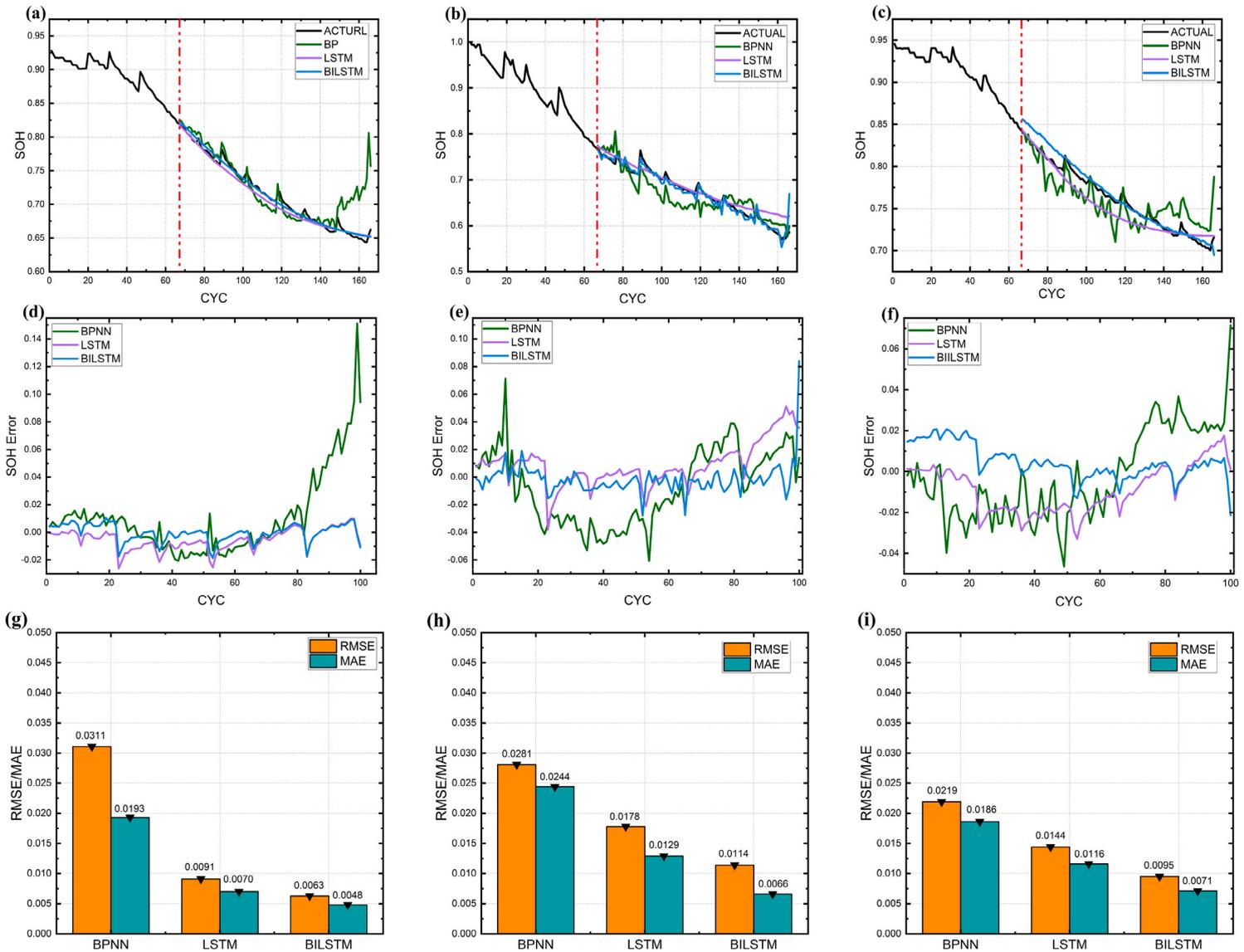


Fig. 10. NASA dataset SOH prediction (training set 40%: prediction set 60%): (a), (d), (g). Bo5 battery's SOH prediction, prediction error, and RMSE and MAE. (b), (e), (h). Bo6 battery's SOH prediction, prediction error, and RMSE and MAE. (c), (f), (i). Bo7 battery's SOH prediction, prediction error, and RMSE and MAE.

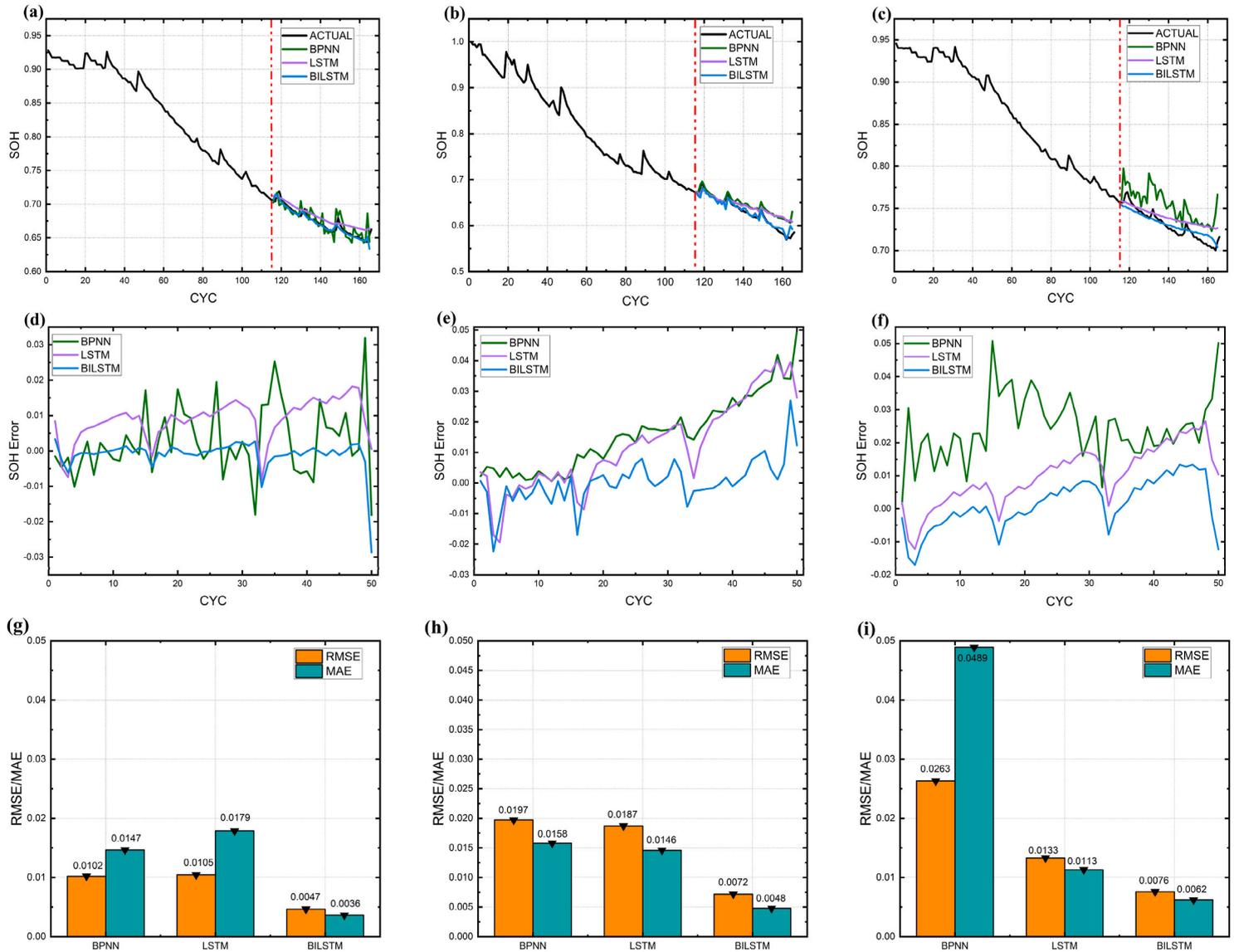


Fig. 11. NASA dataset SOH prediction (training set 70%: prediction set 30%): (a), (d), (g). Bo5 battery's SOH prediction, prediction error, and RMSE and MAE. (b), (e), (h). Bo6 battery's SOH prediction, prediction error, and RMSE and MAE. (c), (f), (i). Bo7 battery's SOH prediction, prediction error, and RMSE and MAE.

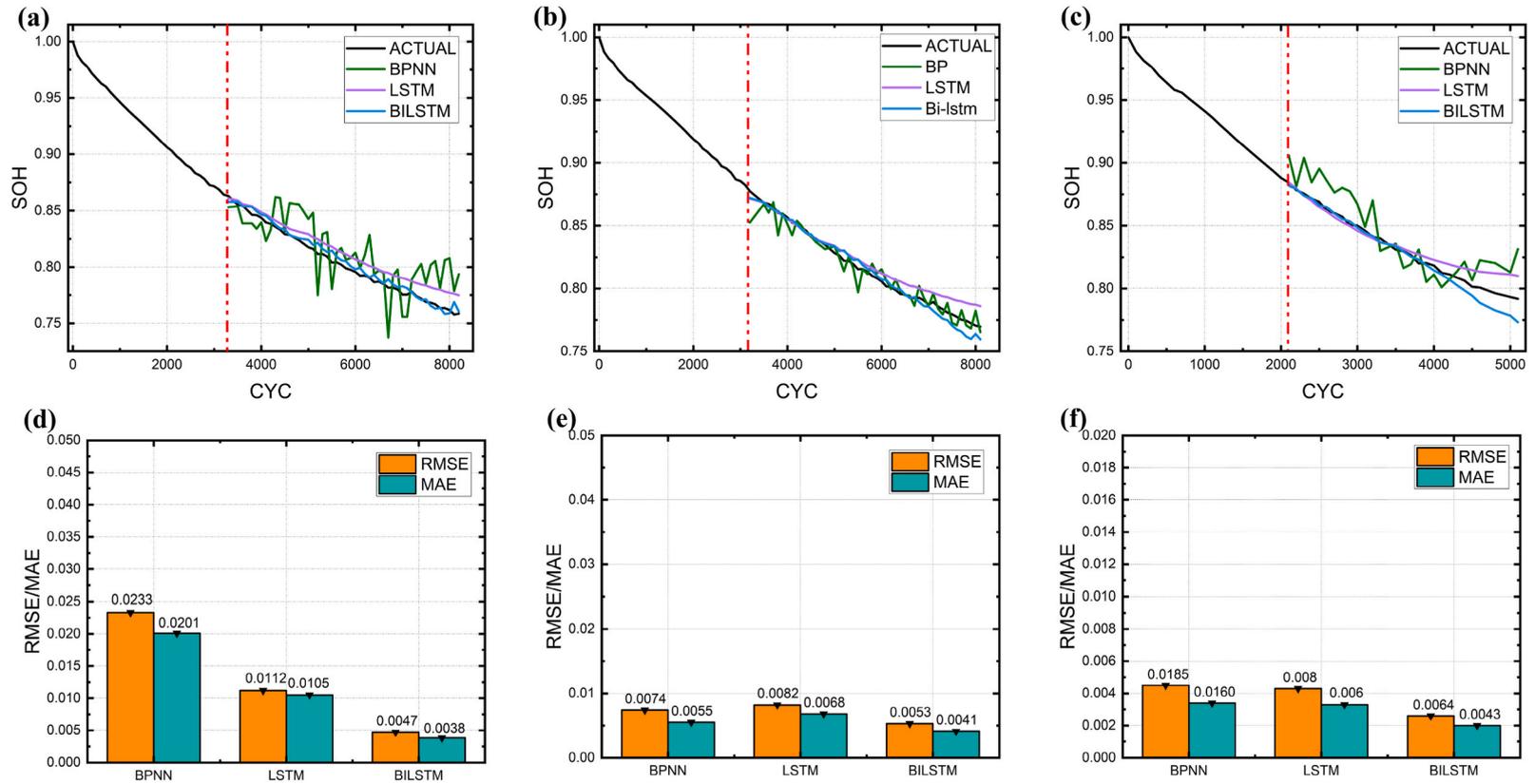


Fig. 12. Oxford dataset SOH prediction (training set 40%: prediction set 60%): (a), (d). Cell1 battery's SOH prediction, and RMSE and MAE. (b), (e). Cell3 battery's SOH prediction, and RMSE and MAE. (c), (f). Cell4 battery's SOH prediction, and RMSE and MAE.

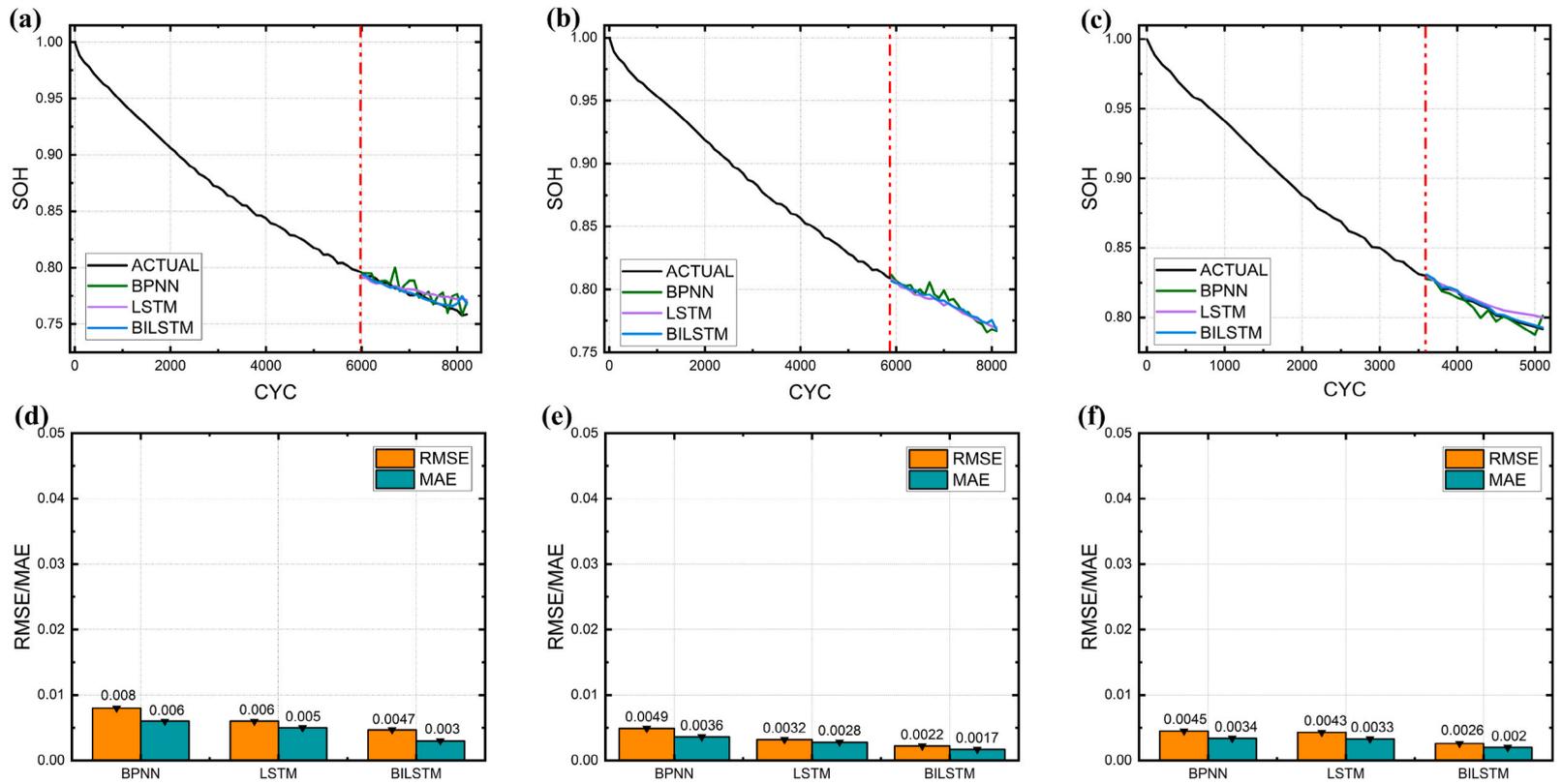


Fig. 13. Oxford dataset SOH prediction (training set 70%; prediction set 30%): (a), (d). Cell1 battery's SOH prediction, and RMSE and MAE. (b), (e). Cell3 battery's SOH prediction, and RMSE and MAE. (c), (f). Cell4 battery's SOH prediction, and RMSE and MAE.

Table 4
Dataset error.

Battery	BPNN				LSTM				BILSTM			
	RMSE (%)		MAE (%)		RMSE (%)		MAE (%)		RMSE (%)		MAE (%)	
	4:6	7:3	4:6	7:3	4:6	7:3	4:6	7:3	4:6	7:3	4:6	7:3
B05	3.11	1.02	1.93	1.47	0.91	1.05	0.70	1.79	0.63	0.47	0.48	0.36
B06	2.81	1.97	2.44	1.58	1.78	1.87	1.29	1.46	1.14	0.72	0.66	0.48
B07	2.19	2.63	1.86	4.89	1.44	1.33	1.16	1.13	0.95	0.76	0.71	0.62
Cell1	2.33	0.80	2.01	0.60	1.12	0.60	1.05	0.50	0.47	0.47	0.38	0.30
Cell3	0.74	0.49	0.55	0.36	0.82	0.32	0.68	0.28	0.53	0.22	0.41	0.17
Cell4	1.85	0.45	1.60	0.34	0.80	0.43	0.60	0.33	0.64	0.26	0.43	0.20

This paper overcomes these challenges by utilizing a combination of IC, DTV, and DT curves for predictive purposes. Using multiple feature curves can help avoid the problem of extracting a single feature being difficult. Even in cases where IC curve extraction proves problematic, the model can still predict SOH through the use of DTV and DT curves. Furthermore, subsequent experiments involving random cycle tests at different SOC levels can be conducted to validate the practicality and effectiveness of the model.

Data availability statement

Data included in article from NASA Ames Prognostics Data Repository <https://www.nasa.gov/intelligent> systems division & Oxford Battery Degradation Dataset 1 - ORA - Oxford University Research Archive.

CRedit authorship contribution statement

Peng Xu: Conceptualization. **Yuan Huang:** Writing – review & editing, Writing – original draft, Visualization, Data curation, Conceptualization. **Wenwen Ran:** Validation, Software. **Shibin Wan:** Formal analysis. **Cheng Guo:** Resources, Project administration. **Xin Su:** Investigation. **Libing Yuan:** Investigation. **Yuanhong Dan:** Software, Project administration, Conceptualization.

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.

Acknowledgments

This work was supported by "High Mobility, Self-Balancing Multipurpose Electric Two-Wheeled Mobility Platform"-Key Research and Development Plan Project of Chongqing Science and Technology Commission, China (2021CCB03). & National Key Laboratory of Transient Physics, Nanjing University of Science and Technology, Nanjing 210094, China (2022-JCJQ-LB-061-07) & Natural Science Foundation of Chongqing, China (No. CSTC2020JCYJ- MSXMX0185).

References

- [1] Carlos Pastor-Fernández, Kotub Uddin, Gael H. Chouchelamane, W. Dhammika Widanage, A comparison between electrochemical impedance spectroscopy and incremental capacity-differential voltage as Li-ion diagnostic techniques to identify and quantify the effects of degradation modes within battery management systems, *J. Power Sources* 360 (2017) 301–318.
- [2] Dongjae Lee, Pyeong-Yeon Lee, Insu Baek, Sanguk Kwon, Deriving the optimized battery model for battery pack and anomaly detection based on the cloud battery management system, *J. Energy Storage* 74 (2023) 109338.
- [3] Mehmet Kurucan, Mete Özbaltan, Zeki Yetgin, Alkan Alkaya, Applications of artificial neural network based battery management systems: a literature review, *Renew. Sustain. Energy Rev.* 192 (2024) 114262.
- [4] Juqiang Feng, Feng Cai, Huachen Li, Kaifeng Huang, A data-driven prediction model for the remaining useful life prediction of lithium-ion batteries, *Process Saf. Environ. Protect.* 180 (2023) 601–615.
- [5] Yi Li, Kailong Liu, M. Aoife, Foley Alana Zülke, Data-driven health estimation and lifetime prediction of lithium-ion batteries: a review, *Renew. Sustain. Energy Rev.* 113 (2019) 109254.
- [6] Qi Zhang, Ralph E. White, Capacity fade analysis of a lithium ion cell, *J. Power Sources* 179 (2) (2008) 793–798.
- [7] P. Ramadass, Bala haran, ralph white, branko N. Popov, mathematical modeling of the capacity fade of Li-ion cells, *J. Power Sources* 123 (2) (2003) 230–240.
- [8] Chun-Pang Lin, Javier Cabrera, Fangfang Yang, Man-Ho Ling, Battery state of health modeling and remaining useful life prediction through time series model, *Appl. Energy* 275 (2020) 115338.
- [9] Yongfang Guo, Kai Huang, Xiangyuan Yu, Yashuang Wang, State-of-health estimation for lithium-ion batteries based on historical dependency of charging data and ensemble SVR, *Electrochim. Acta* 428 (2022) 140940.
- [10] Dengfeng Zhang, Weichen Li, Xiaodong Han, Baochun Lu, Evolving Elman neural networks based state-of-health estimation for satellite lithium-ion batteries, *J. Energy Storage* 59 (2023) 106571.
- [11] Ji Wu, Leichao Fang, Guangzhong Dong, Mingqiang Lin, State of health estimation of lithium-ion battery with improved radial basis function neural network, *Energy* 262 (2023) 125380.
- [12] Yajun Zhang, Yajie Liu, Jia Wang, Tao Zhang, State-of-health estimation for lithium-ion batteries by combining model-based incremental capacity analysis with support vector regression, *Energy* 239 (2022) 121986.

- [13] Xiaoyu Li, Zhenpo Wang, Jinying Yan, Prognostic health condition for lithium battery using the partial incremental capacity and Gaussian process regression, *J. Power Sources* 421 (2019) 56–67.
- [14] YongFang Guo, Kai Huang, XiaoYa Hu, A state-of-health estimation method of lithium-ion batteries based on multi-feature extracted from constant current charging curve, *J. Energy Storage* 36 (2021) 102372.
- [15] Chenghui Zhang, Yongzhe Kang, Bin Duan, Zhongkai Zhou, An adaptive battery capacity estimation method suitable for random charging voltage range in electric vehicles, *IEEE Trans. Ind. Electron.* 69 (9) (2022) 9121–9132.
- [16] Huanwei Xu, Lingfeng Wu, Shizhe Xiong, Wei Li, An improved CNN-LSTM model-based state-of-health estimation approach for lithium-ion batteries, *Energy* 276 (2023) 127585.
- [17] Alejandro Gismero, Kjeld Nørregaard, Bjarne Johnsen, Lasse Stenhøj, Electric vehicle battery state of health estimation using Incremental Capacity Analysis, *J. Energy Storage* 64 (2023) 107110.
- [18] Sanguk Kwon, Dongho Han, Jinhyeong Park, Pyeong-Yeon Lee, Joint state-of-health and remaining-useful-life prediction based on multi-level long short-term memory model prognostic framework considering cell voltage inconsistency reflected health indicators, *J. Energy Storage* 55 (2022) 105731.
- [19] A. Fly, R. Chen, Rate dependency of incremental capacity analysis (dQ/dV) as a diagnostic tool for lithium-ion batteries, *J. Energy Storage* 29 (2020) 101329.
- [20] Ruomei Zhou, Rong Zhu, Cheng-Geng Huang, Weiwen Peng, State of health estimation for fast-charging lithium-ion battery based on incremental capacity analysis, *J. Energy Storage* 51 (2022) 104560.
- [21] Brian Ospina Agudelo, Zamboni Walter, Eric Monmasson, Application domain extension of incremental capacity-based battery SoH indicators, *Energy* 234 (2021) 121224.
- [22] Billy Wu, Vladimir Yufit, Merla Yu, F. Ricardo, Martinez-Botas, Differential thermal voltammetry for tracking of degradation in lithium-ion batteries, *J. Power Sources* 273 (2015) 495–501.
- [23] Zhenpo Wang, Changgui Yuan, Xiaoyu Li, Lithium battery state-of-health estimation via differential thermal voltammetry with Gaussian process regression, *IEEE Trans. Transport. Electrification* 7 (1) (2021) 16–25.
- [24] Jufeng Yang, Yingfeng Cai, Chris Mi, Lithium-ion battery capacity estimation based on battery surface temperature change under constant-current charge scenario, *Energy* 241 (2022) 122879.
- [25] Shuxin Zhang, Zhitao Liu, Hongye Su, State of health estimation for lithium-ion batteries on few-shot learning, *Energy* 268 (2023) 126726.
- [26] Hongqian Zhao, Zheng Chen, Shu Xing, Jiangwei Shen, State of health estimation for lithium-ion batteries based on hybrid attention and deep learning, *Reliab. Eng. Syst. Saf.* 232 (2023) 109066.
- [27] Zheng Chen, Hongqian Zhao, Yuanjian Zhang, Shiquan Shen, State of health estimation for lithium-ion batteries based on temperature prediction and gated recurrent unit neural network, *J. Power Sources* 521 (2022) 230892.
- [28] M. Bercebar, I. Gandiaga, I. Villarreal, N. Omar, Critical review of state of health estimation methods of Li-ion batteries for real applications, *Renew. Sustain. Energy Rev.* 56 (2016) 572–587.
- [29] Bharat Bohara, Raymond I. Fernandez, Vysali Gollapudi, Xingpeng Li, Short-term aggregated residential load forecasting using BiLSTM and CNN-BiLSTM, *International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)* (2022) 37–43.
- [30] Zhen Zhang, Wentao Zhang, Kuo Yang, Shujing Zhang, Remaining useful life prediction of lithium-ion batteries based on attention mechanism and bidirectional long short-term memory network, *Measurement* 204 (2022) 112093.
- [31] Panagiotis Eleftheriadis, Sonia Leva, Emanuele Ogliari, Bayesian hyperparameter optimization of stacked bidirectional long short-term memory neural network for the state of charge estimation, *Sustain. Energy, Grids and Networks* 36 (2023) 101160.
- [32] B. Saha, K. Goebel, Battery Data Set, NASA AMES Prognostics Data Repository, 2007.
- [33] Matthieu Dubarry, Cyril Truchot, Bor Yann Liaw, Synthesize battery degradation modes via a diagnostic and prognostic model, *J. Power Sources* 219 (2012) 204–216.
- [34] C.R. Birkl, *Diagnosis and Prognosis of Degradation in Lithium-Ion Batteries*, University of Oxford, Oxford, 2017.
- [35] Huixing Meng, Mengyao Geng, Te Han, Long short-term memory network with Bayesian optimization for health prognostics of lithium-ion batteries based on partial incremental capacity analysis, *Reliability Engineering & System Safety* 236 (2023) 109288.
- [36] Carlos Pastor-Fernández, Kotub Uddin, Gael H. Chouchelamane, W. Dhammika Widanage, A comparison between electrochemical impedance spectroscopy and incremental capacity-differential voltage as Li-ion diagnostic techniques to identify and quantify the effects of degradation modes within battery management systems, *J. Power Sources* 360 (2017) 301–318.
- [37] Bin Ma, Shichun Yang, Lisheng Zhang, Wentao Wang, Remaining useful life and state of health prediction for lithium batteries based on differential thermal voltammetry and a deep-learning model, *J. Power Sources* 548 (2022) 232030.
- [38] Lisheng Zhang, Wentao Wang, Hanqing Yu, Zheng Zhang, Remaining useful life and state of health prediction for lithium batteries based on differential thermal voltammetry and a deep learning model, *iScience* 25 (12) (2022) 105638.
- [39] Van Houdt Greg, Mosquera Carlos, Nápoles Gonzalo, A review on the long short-term memory model, *Artif. Intell. Rev.* 53 (8) (2020) 5929–5955.