Research Article

Li-Ion Battery State of Health Estimation Based on Short Random Charging Segment and Improved Long Short-Term Memory

Aina Tian,[1] Zhe Chen,[1] Zhuangzhuang Pan,[1] Chen Yang,[2] Yuqin Wang,[1] Kailang Dong,[1] Yang Gao,[3] and Jiuchun Jiang [1]

[1]*Hubei Key Laboratory for High-Efficiency Utilization of Solar Energy and Operation Control of Energy Storage System, Hubei University of Technology, Wuhan 430068, China*
[2]*Central and Southern China Municipal Engineering Design and Research Institute Co, Ltd, Wuhan 430068, China*
[3]*School of Electric Power, South China University of Technology, Guangzhou 510640, China*

Correspondence should be addressed to Jiuchun Jiang; jcjiang@hbut.edu.cn

Lithium-ion batteries have been used in a wide range of applications, including electrochemical energy storage and electrical transportation. In order to ensure safe and stable battery operation, the State of Health (SOH) needs to be accurately estimated. In recent years, model-based and data-driven methods have been widely used for SOH estimation, but due to the uncertainty of battery charging conditions in practice, it is difficult to obtain a fixed local segment. In this paper, the charging curve is first divided into several equal voltage difference segments based on charging segment voltage difference ΔV in order to solve the random charging segment problem. Time interval of equal charge voltage difference of the voltage curve, coefficient of variation and euclidean distance of the charging capacity difference curve are extracted as health features. The improved flow direction algorithmlong short term memory-based SOH assessment method is proposed and verified by the Oxford battery degradation dataset and experimental battery degradation dataset with a maximum error of 0.6%.

1. Introduction

With the advantages of low self-discharge rate, high power density and long life [1], lithium-ion batteries have been widely used in the energy storage systems, electric vehicles, and other fields. In practical applications [2], improper handling of lithium ions can lead to deterioration of battery performance or even abnormal deterioration, which may increase battery safety risks. Accurate battery state of health (SOH) prediction is essential for battery management system (BMS) during battery operation to prevent these problems. SOH is an evaluation index to measure the battery life [3], reflecting the degradation degree and reliability of the battery [4]. With the concern of the battery SOH in automotive applications, this paper focuses on the SOH estimation of batteries.

In recent years, lithium-ion battery SOH estimation methods can be broadly classified into three categories:

experimental analysis method, model-based method, and data-driven method [5, 6]. The experimental analysis method is to evaluate the batteries SOH by directly measuring the parameters of the battery through specific experimental equipment and a standardized operation procedure, which mainly includes the coulomb counting method, open circuit voltage (OCV) method, and electrochemical impedance spectrum (EIS) method [7, 8]. The traditional monitoring method is mainly based on voltage and current, which cannot reflect the internal mechanism, so the accuracy is greatly reduced. Recently, with the development of EIS, it has been possible to estimate the state of health quickly and accurately online [9]. Equivalent circuit model-based methods use a simplified set model of the combination of resistance, inductance, and capacitance to describe the battery behavior and mainly include empirical recession models, equivalent circuit models (ECM), and recession mechanism models [10, 11]. These

methods require detailed cell specifications and demanding computational power to solve complex partial differential equations, which are more complicated in the practical applications.

Data-driven based methods [12] have attracted much attention due to the lack of need to construct a mechanistic model, and they have been widely used for SOH estimation due to a large amount of battery data. A large number of machine learning algorithms have been used for battery SOH estimation, such as support vector machine (SVM) [13], Gaussian process regression (GPR) [14], relevance vector machine (RVM) [15], Bayesian network (BN) [16], extreme learning machine (ELM) [17], long short-term memory (LSTM) neural network [18], and other machine learning methods are applicable to nonlinear systems such as batteries. Zhang et al. [19] review how to use the latest data-driven algorithms to predict the SOH of LIBs, and proposes a general prediction process, including the acquisition of datasets for the charging and discharging process of LIBs, the processing of data and features, and the selection of algorithms. Wen et al. [20] proposed a battery SOH prediction model based on incremental capacity analysis and BP neural network was proposed to predict the SOH of batteries under different ambient temperatures. Considering the problem that BP neural networks tend to fall into local optimum, Gong et al. [21] selected four health indicators during charging and discharging and constructed a battery capacity degradation model based on LSTM neural network, which has a higher estimation accuracy with a maximum error within 3%. Since LSTM is well suited for nonlinear systems such as battery degradation, the IFDA-LSTM is proposed in this paper to solve the battery SOH estimation.

In addition to improving estimation accuracy and algorithmic efficiency, much research has been devoted to the extraction of health features (HFs). For the data-driven method, it is crucial to extract features that are closely related to the battery SOH. In recent years, incremental capacity analysis (ICA) has been considered as an effective method for offline estimation of SOH. Li et al. [22] used incremental capacity (IC) curves to extract health indicators to model battery degradation and applied gray correlation analysis and entropy weighting to assess the relevance and importance of HIs, respectively. Liu et al. [23] used the discharge voltage difference at equal voltage intervals as a HF to estimate the battery SOH. Patil et al. [24] extracted eight features, such as voltage maximum from the battery discharge curve and another 13 features by equation. However, there are still some problems in estimating the battery SOH based on charging data. The degradation features extracted from the above studies are based on the fact that there are whole or partial charging data, but the electric vehicles (EVs) charging starting state of charge (SOC) and charging duration are usually random, so it is difficult to obtain the corresponding features under some charging cycles. Shu et al. [25] proposed a machine learning method for online diagnosis of the SOH of batteries. A prediction model for future voltage profiles is established based on the ELM algorithm with the short-term charging data. This method allows the estimation error of

SOH to be kept within 2%. Qian et al. [26] presented a one-dimensional convolution neural network- (1D CNN-) based method that takes random segments of charging curves as inputs to perform capacity estimation for lithium-ion batteries. When the segmentation data are selected from the last part of charging curves, RMSEs of all of the network models are larger than 7.5%.

SOH estimation based on short random charging segment and improved LSTM is proposed in this paper. The main contributions of this paper are as follows:

(1) The charge capacity difference curve shifts downward and the rise rate in the region during the battery degradation is slowed down. In this paper, the coefficient of variation (CV) and Euclidean distance (ED) of the charging capacity difference curve are extracted as HFs to characterize the SOH of the battery. As the battery performance degrades, the charging voltage curve shifts to the right, and the moment of rising to a certain voltage constantly advances. Time interval of equal charge voltage difference (TIECVD) is extracted as a HF; the ability of the features to characterize the battery SOH over a relatively small range of random charging segments is further enhanced.

(2) The improved flow direction algorithm-LSTM (IFDA-LSTM) based SOH estimation optimization model is constructed, and LSTM neural network combined with IFDA is used to optimize the number of nodes in the implicit layer and the initial learning rate of the LSTM model. The validity of the model is verified by Oxford and experimental datasets.

The rest of the paper is organized as follows: the Oxford degradation and experimental degradation datasets are presented in the second part. The charging capacity difference curve and the voltage-dependent health characteristics are given in the third part. In the fourth part, a SOH estimation model based on IFDA-LSTM is proposed. In the fifth part using simulations of different battery datasets are presented. Conclusions are given in the sixth part.

2. Lithium-Ion Batteries Dataset

2.1. Oxford Dataset. The experimental battery electrodes were made from a mixture of graphite anodes and lithium cobalt oxide (LCO) and nickel cobalt oxide (NCO) cathodes [27], which consisted of eight 0.74 A hr Kokam batteries named #1–#8. Table 1 shows the main parameters of the battery. In the degradation test, the lithium battery was repeatedly charged at a constant current of 1.48 A and discharged to simulate Artemis urban driving conditions, with capacity measured every 100 cycles. The ambient temperature of the batteries was 40°C. Figure 1 shows the Oxford dataset battery capacity.

2.2. Experimental Dataset. The laboratory-built lithium-ion battery degradation experimental rig consists of Galaxy SDJ405F high and low-temperature humidity and heat test chamber, Arbin LBT single-cell battery test equipment,

TABLE 1: Main parameters of Kokam batteries.

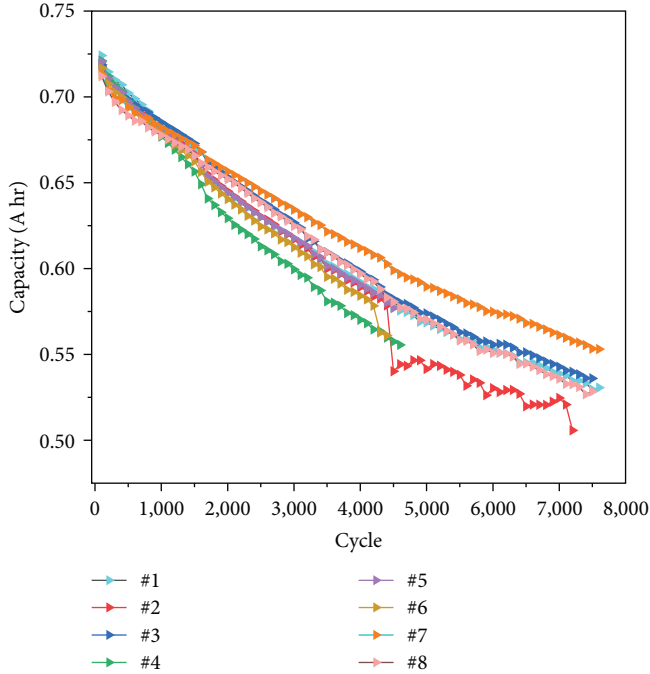| Cathode material	Nominal capacity (A hr)	Nominal voltage (V)	Discharge cutoff voltage (V)	Charge cutoff voltage (V)
LCO and NCO	0.74	3.7	2.7	4.2

FIGURE 1: Capacity of eight cells in the Oxford dataset.

upper computer, test and training host, and other equipment. Its topology is shown in Figure 2.

The test batteries were two Panasonic NCR18650BD ternary lithium batteries, named BD01 and BD02, respectively. The ambient temperature of the batteries was 20°C. Table 2 shows the main parameters of the experimental battery. Two batteries were discharged at 0.61 A with a rated capacity of 3 A hr at an upper and lower limit voltage of 4.2/ 2.75 V. Accelerated battery degradation experiments were conducted with a constant current of 3.2 A and a constant voltage of 4.2 V (current reduced to 0.64A). Considering that the battery is rarely fully discharged in practice, the discharge process was set to a high current of 6.4 A until the voltage cutoff at 2.75 V. In the initial cycle, the average charge capacity of the two cells is calibrated to have SOH equal to 1. The experiment is stopped when the battery capacity degradation to 80%. Figure 3 shows the experimental battery capacity.

3. Healthy Feature Extraction

The discharge curve of the battery changes with the working conditions, which makes it difficult to obtain stable data, while the charging working conditions are often fixed, so it is easier to obtain the effective features. However, the charging interval of EVs is random and the charging starting voltage or SOC is not fixed. Therefore, how to achieve an accurate estimation of battery capacity under short random charging fragment data is the key challenge.

In this paper, the commonly used capacity definition method is used to define the battery SOH, as shown in Equation (1).

$$\text{SOH} = \frac{Q_{\text{now}}}{Q_{\text{rated}}} \times 100\%, \tag{1}$$

where Q_{now} and Q_{rated} indicate the current capacity and rated capacity of the battery, respectively.

This section focuses on analyzing the degradation process of the Oxford dataset, dividing the charging curve into several equal voltage difference segments, extracting TIECVD of the voltage curve, CV and ED of the charging capacity difference curve as HFs within a random segment.

3.1. Charging Fragment Division Method. Due to the uncertainty of the battery charging conditions in practice, it is difficult to obtain a fixed local segment every time, so the charging curve is divided into a number of segments with equal voltage differences.

First, the charging segments are divided. As shown in Figure 4, the charging capacity curve can be divided into multiple equal length segments based on the charging segment voltage difference ΔV. In this case, since the battery is rarely consumed to too low a level before charging [28], there is no need to start the segmentation from the lowest cutoff voltage.

In summary, the charging capacity curve can be divided into N segments; each segment is shifted backward in turn by one step, where the value of 0.01 V is fixed [29], as shown in Equation (2).

$$N = \frac{V_{\text{end}} - V_{\text{start}} - \Delta V}{0.01} + 1, \tag{2}$$

where V_{start} and V_{end} are the starting and ending voltages.

3.2. HFs Related to Charging the Capacity Difference. The charged capacity within each segment $Q = [Q_1, Q_2,\ldots\ldots, Q_n]$ is unknown, while the charging capacity difference $\Delta Q = Q - Q_1$ is known and can be used for quantitative analysis. The degradation evolution of ΔQ during the voltage range (3.6, 4.0 V) of the #1 cell is shown in Figure 5.

It can be seen from Figure 5 that the incremental capacity curve shows a clear pattern of variation with the battery degradation. As shown above, the ΔQ curve shifts down and the rising rate slows down with the battery degradation. The HFs can be extracted in the region.
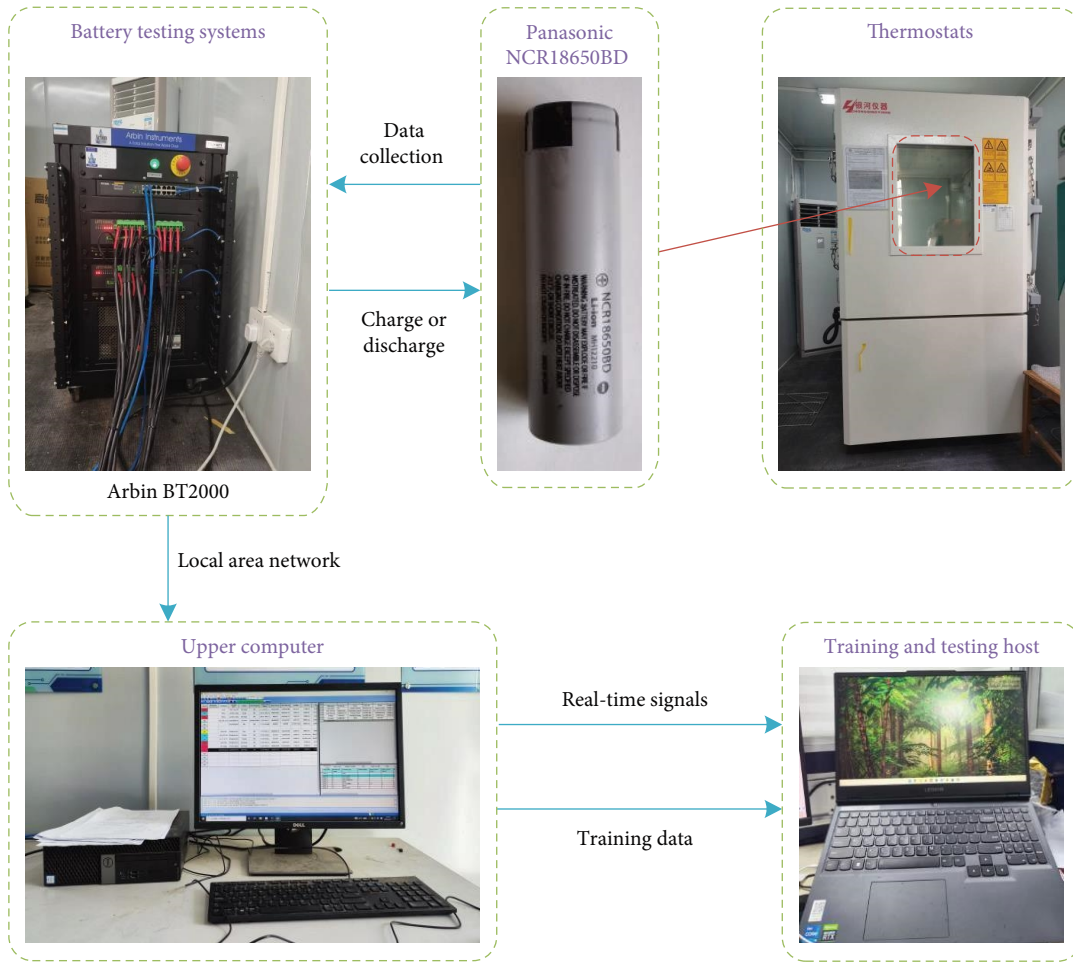
FIGURE 2: Experimental platform topology.

TABLE 2: Main parameters of NCR18650BD battery.

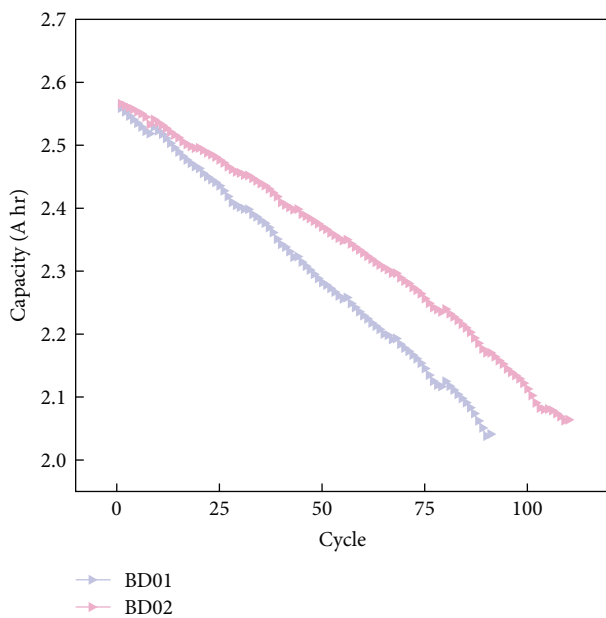| Nominal capacity (A hr)	Nominal voltage (V)	Discharge cutoff voltage (V)	Charge cutoff voltage (V)	Charge temperature
3	3.6	2.75	4.2	20°C

FIGURE 3: Capacity of two experimental cells.
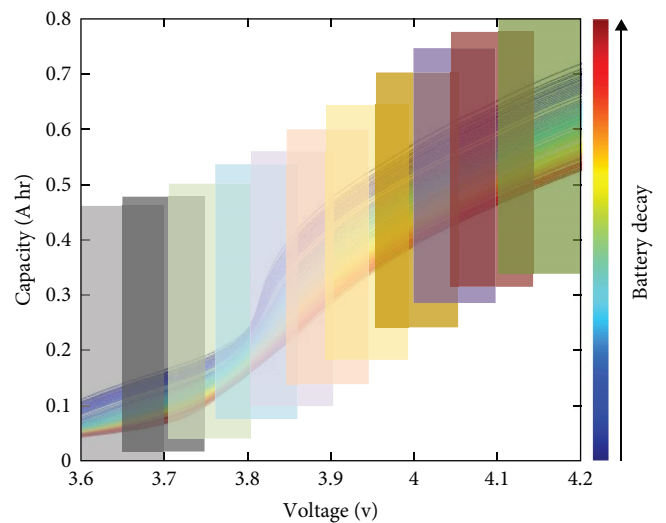
FIGURE 4: Charging capacity curve segmentation diagram.
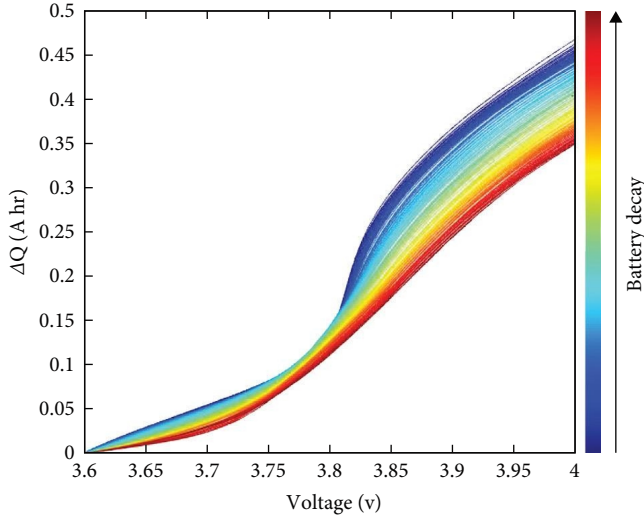
FIGURE 5: ΔQ–V curve for #1 during (3.6 and 4.0 V).

FIGURE 6: Charging voltage curve of #1 battery.

In order to quantify the variation of the charge capacity difference curve, the CV and the ED are calculated in this paper and used as HFs.

The CV is a dimensionless measure to analyze the degree of data dispersion, which is used to compare the degree of variation of the curve. The CV is introduced in the ΔQ–V curve of the charging fragment, as shown in Equation (3).

$$\mathrm{CV} = \frac{\sqrt{\dfrac{\sum\limits_{i}^{M}\left(\Delta Q_i - \overline{\Delta Q}\right)}{M}}}{\overline{\Delta Q}}, \tag{3}$$

where $\overline{\Delta Q}$ is the mean value of ΔQ for each sampling point within the fragment.

The ED is the absolute distance between curves and a relatively simple metric for analyzing the similarity of curves. The ED is introduced into the ΔQ–V curve of the charging segment. The Q–V curves were interpolated to solve for the same number of data elements. The interpolated Q-sequence is of equal length. The ED between two sets of ΔQ sequences $\Delta Q_1 = \{\Delta Q_{11}, \Delta Q_{12}, \ldots, \Delta Q_{1n}\}$ and $\Delta Q_2 = \{\Delta Q_{21}, \Delta Q_{22}, \ldots, \Delta Q_{2n}\}$ is shown in Equation (4).

$$\mathrm{ED} = \sqrt{\sum\limits_{i}^{m}(\Delta Q_{1i} - \Delta Q_{2i})^2}. \tag{4}$$

3.3. HFs Related to Charging Voltage Curve. In the degradation experiment, the charging voltage rose faster as the battery performance degraded. The charging voltage curve of the #1 battery is shown in Figure 6. It is evident that, with the increase of the testing period, the charging voltage curve shifts to the right and the time for it to reach a certain voltage is kept advancing.

Considering the difficulty of obtaining the complete curve, the TIECVD can be extracted as a HF, and the calculation formula is shown in Equation (5).
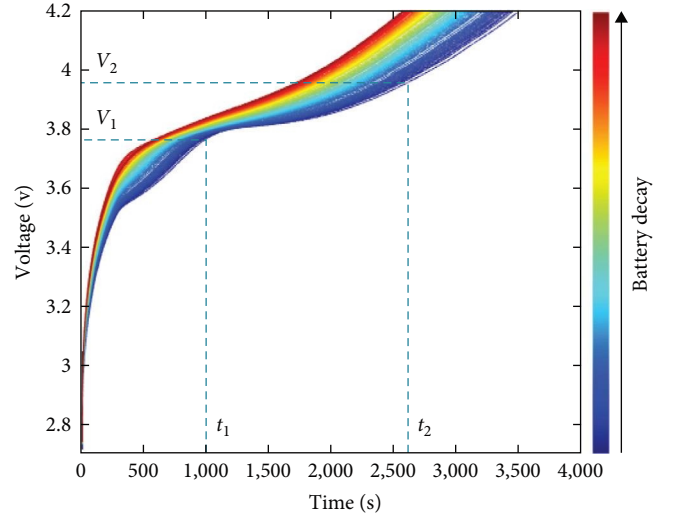
$$\mathrm{TIECVD}(V_1, V_2) = t_1 - t_2 = \Delta t, \tag{5}$$

where V_1 and V_2 are the voltages corresponding to the different times during the charging process.

It can be observed that as the battery ages, the voltage curve shifts upward, indicating a certain relationship between voltage and SOH. Considering that this section requires the simultaneous extraction of both CV and ED at HFs, we choose to set the voltage range for V_1 and V_2 between 3.6 and 4.0 V.

3.4. Correlation Analysis between HFs and SOH. In order to evaluate the correlation between the extracted features and the battery capacity, the Pearson correlation coefficient and Spearman correlation coefficient were introduced, respectively. The ability of the extracted HFs to characterize the degradation of battery capacity was verified by analyzing batteries #1–#8.

The CV and ED were introduced within the ΔQ–V curve of each charging segment, and TIECVD was extracted from the voltage profile. The correlation between the three extracted HFs and the battery capacity was analyzed. As shown in Equation (2), the number of segments N is determined by the segment voltage difference ΔV. When ΔV takes a larger value, the more data within the segment, the more information is described, but at the same time, it increases the difficulty of data acquisition.

In this section, the voltage interval is defined as (3.6 and 4.0 V), the voltage window length $\Delta V = 0.4$ V [30], and every 0.01 V is extracted from ΔV, then each charging interval contains 21 charging segments. Figure 7 shows the three features extracted from the Oxford degradation dataset regarding the capacity difference and voltage, and the features are normalized for the comparison purposes. As can be seen from Figure 7 below, the CV and ED features follow a similar but not identical trend with battery capacity degradation combined with TIECVD. It can be observed that for
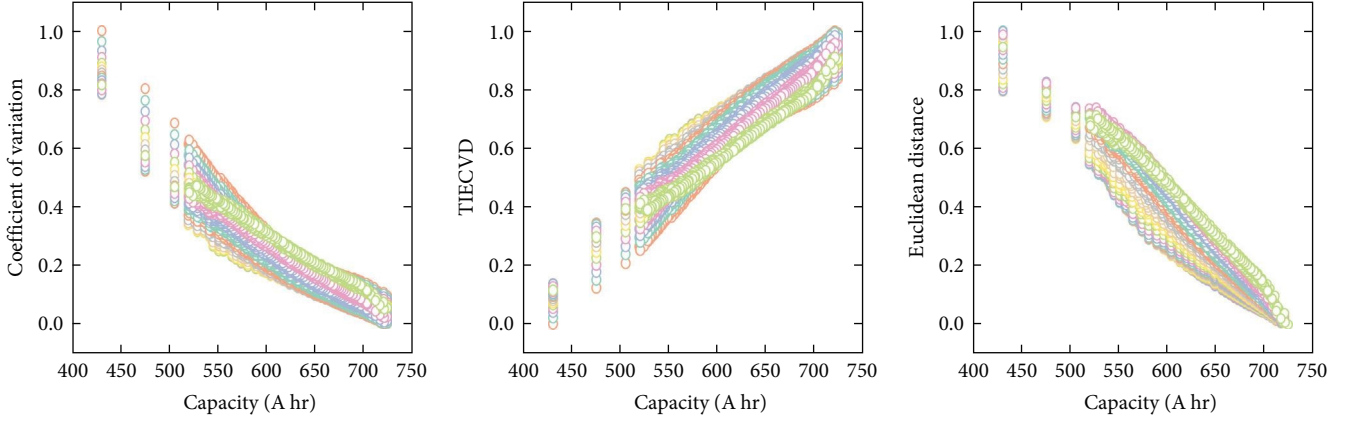
FIGURE 7: Correlation of HFs CV, ED, and TIECVD with cell capacity.

almost all charging segments, the three HFs have a strong correlation with battery capacity.

In order to further determine the correlation between the two, ΔV is taken as $0.40 : -0.01 : 0.20$ V in turn for the analysis here, and the number of segments N is obtained as $21 : 1 : 41$, respectively, where the first segment is the lowest voltage segment. The obtained data were analyzed using Pearson and Spearman correlation coefficients, and the results are shown in Figure 8.

As shown in Figure 8, setting different charging fragment voltage differences ΔV can obtain different numbers of fragments, and the vertical coordinates indicate the number of charging fragments corresponding to different ΔV. The more the color in the graph tends to yellow, the stronger the correlation between the two, and vice versa. The more it tends to blue, the weaker the correlation. CV, ED, and TIECVD extracted within the different fragments have different correlation coefficients with battery capacity. The Pearson and Spearman correlation coefficients are close to 1 within most of the fragments, and the 3 HFs can characterize the capacity degradation of the cell well. However, it can be observed that the correlations start to perform poorly when the number of fragments reaches 24. Collectively, the features are better characterized when the charge is set at a lower voltage difference or with fewer voltage fragments. Considering the practical situation, the setting of charging fragment voltage difference ΔV should not be too small in order to ensure the quality of the extracted HFs.

4. Battery Health Estimation Based on IFDA-LSTM

4.1. LSTM Neural Network. LSTM neural networks can alleviate the problem of gradient disappearance and gradient explosion [31], and show excellent capabilities in terms of robustness and sensitive data. Each recursive module in the LSTM filters information through four hidden layers: forgetting gates, input gates and output gates, and finally a unit state that saves and refreshes long-term memory. The basic structure of the LSTM is depicted in Figure 9.

The forgetting gate is used to control whether or not to forget, and its mathematical expression is Equation (6):

$$f_t = \varphi\big(W_f[h_{t-1}, x_t] + b_f\big), \tag{6}$$

where φ is called the Sigmoid activation function. W_f and b_f are the weighting matrix and bias vector, respectively.

The input gate, which is the second gate, has the expressions Equations (7) and (8):

$$i_t = \varphi(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{7}$$

$$\overline{C_t} = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c). \tag{8}$$

The results of both the previous forgetting gate and the input gate act on the cell state C_t, whose expression is Equation (9).

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \overline{C_t}, \tag{9}$$

where "$\otimes$" means that the two vectors are multiplied by the corresponding elements.

The last stage is the calculation of the new hidden state, whose expressions are Equations (10) and (11).

$$o_t = \varphi(W_0 \cdot [h_{t-1}, x_t] + b_0), \tag{10}$$

$$h_t = o_t \otimes \tanh C_t. \tag{11}$$

However, the selection of the number of nodes in the hidden layer of the LSTM and the optimization of the initial learning rate are both crucial. It is worthwhile to study how to select the parameter values of LSTM and improve the prediction ability of LSTM.

4.2. Improved Flow Direction Algorithm. To improve the FDA [32], the initial solution of the population is generated by a Cubic chaotic map, which makes the position distribution of the water flow more uniform and improves the local search ability of the algorithm. Then the inertia weight strategy is introduced to improve the global optimization and local development ability of the algorithm. Finally, the optimal neighborhood perturbation strategy and greedy strategy
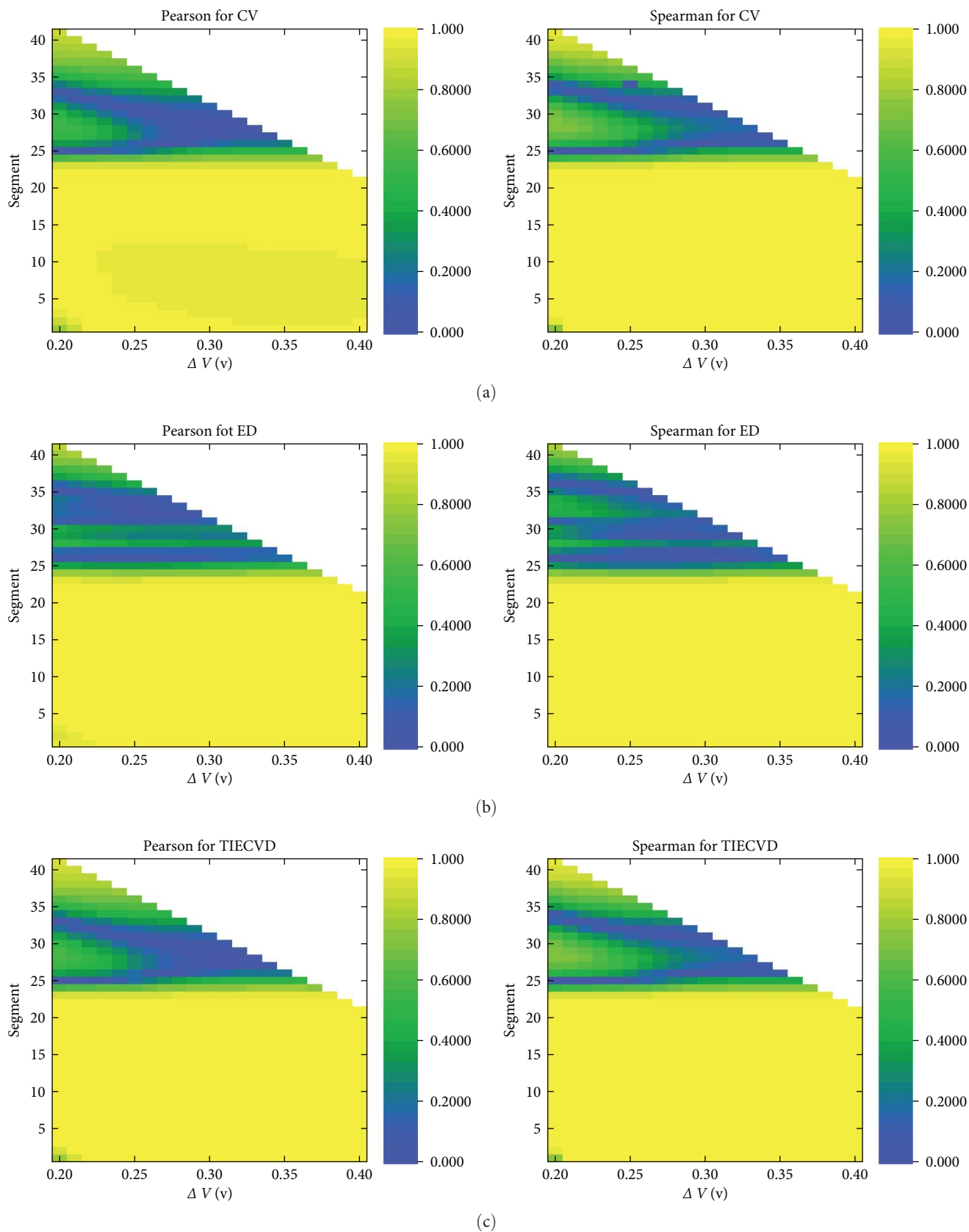
(a)

(b)

(c)

FIGURE 8: Results of correlation coefficient analysis between CV, ED, TIECVD, and cell capacity: (a) correlation coefficient between CV and capacity; (b) correlation coefficient between ED and capacity; (c) correlation coefficient between TIECVD and capacity.
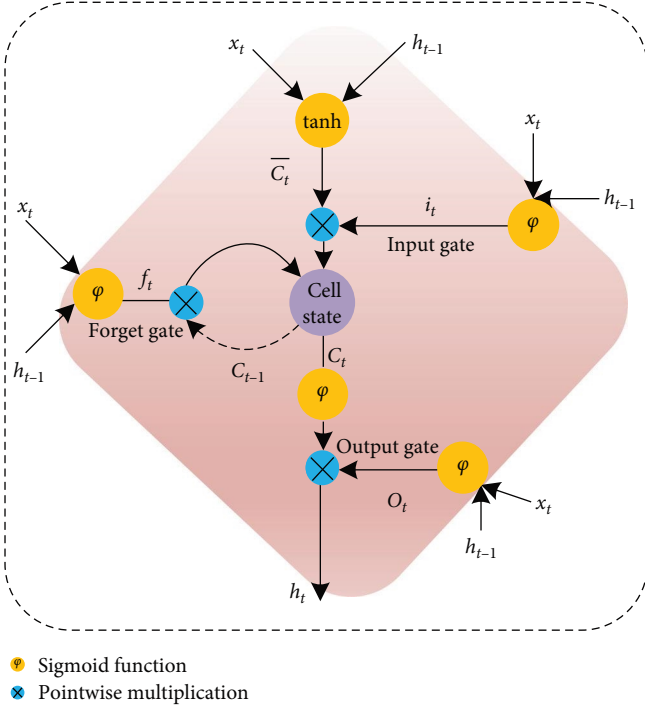
- Sigmoid function
- Pointwise multiplication

FIGURE 9: Structure of LSTM.

are used to update the global optimal solution. The number of nodes in the hidden layer and the initial learning rate of LSTM are determined by the proposed IFDA algorithm.

(A) First, the cubic chaotic mapping shown in Equations (12) and (13) is incorporated into the initialization phase of the FDA algorithm to make the initial location of the water flow more uniformly distributed, thus facilitating a better search in space at the initial stage of the algorithm.

$$\begin{cases} Z(i) = \text{rand}(1, \text{dim}) & i = 1 \\ Z(i) = Z(i-1) \times (1 - Z(i)^2) & i \geq 2 \end{cases}, \quad (12)$$

$$\text{Flow\_}X(i) = l_b + (u_b - l_b) \times Z(i), \quad (13)$$

where the number of variables dim. The upper and lower boundaries u_b and l_b. Flow_$X(i)$ is the location of flow i. The rand is a random value uniformly distributed between 0 and 1. Assume that there exists a β-neighborhood around each flow whose location is generated by the shown in Equation (14):

$$\text{Neighbor\_}X(i) = \text{Flow\_}X(i) + \text{randn}(1, \text{dim}) \times \Delta, \quad (14)$$

where Neighbor_$X(i)$ denotes the ith position of a neighbor, and randn is a normally distributed random value with mean 0 and standard deviation 1. Smaller values Δ result in a search over a smaller range, while larger values allow a search

over a larger range, and to establish a balance between global and local search, Δ is linearly reduced from large values to small values as shown in Equations (15) and (16):

$$\Delta = W \times (\text{rand} \times X_{\text{rand}} - \text{rand} \times \text{Flow\_}X(i) \\ \times \|\text{Best\_}X - \text{Flow\_}X(i)\|), \quad (15)$$

$$X_{\text{rand}} = l_b + \text{rand}(1, \text{dim}) \times (u_b - l_b), \quad (16)$$

where X_{rand} is a random position generated by Equation (17). W is a nonlinear random number weight between 0 and inf, which is calculated as follows:

$$W = \left(\left(1 - \frac{\text{iter}}{\text{Max\_Iter}} \right)^{(2 \times \text{randn})} \right) \times \text{rand}(1, \text{dim}) \\ \times \frac{\text{iter}}{\text{Max\_Iter}} \times \text{rand}(1, \text{dim}). \quad (17)$$

As mentioned above, the flow moves with a velocity of V to the neighbor with the minimum objective function. On the other hand, the velocity of the flow to a neighboring point is directly related to its slope. Accordingly, the flow velocity vector is determined using the shown in Equation (18).

$$V = \text{randn} \times S_f, \quad (18)$$

where S_f denotes the slope vector between the neighbor and the stream's current position. randn generates various solutions to increase the global search. The slope vector of flow i with respect to neighbor j is also determined by the shown in Equation (19).

$$S_f(i, j, d) = \frac{\text{Flow\_fitness}(i) - \text{Neighbor\_fitness}(j)}{\| \text{Flow\_}x(i, d) - \text{Neighbor\_}x(j, d) \|}, \quad (19)$$

where Flow_fitness(i) and Neighbor_fitness(j) denote the target values of flow(i) and neighbor(j), respectively. The parameter d denotes the dimensionality of the problem. The following relation is used to determine the new position, as shown in Equation (20).

$$\text{Flow\_new}X(i) = \text{Flow\_}X(i) + V \\ \times \frac{\text{Flow\_}X(i) - \text{Neighbor\_}X(j)}{\| \text{Flow\_}x(i) - \text{Neighbor\_}x(j) \|}. \quad (20)$$

(B) To further balance the algorithm's global search and local exploration capabilities, a new random inertia weight is added in this paper to define the range of V, as shown in Equations (21) and (22).
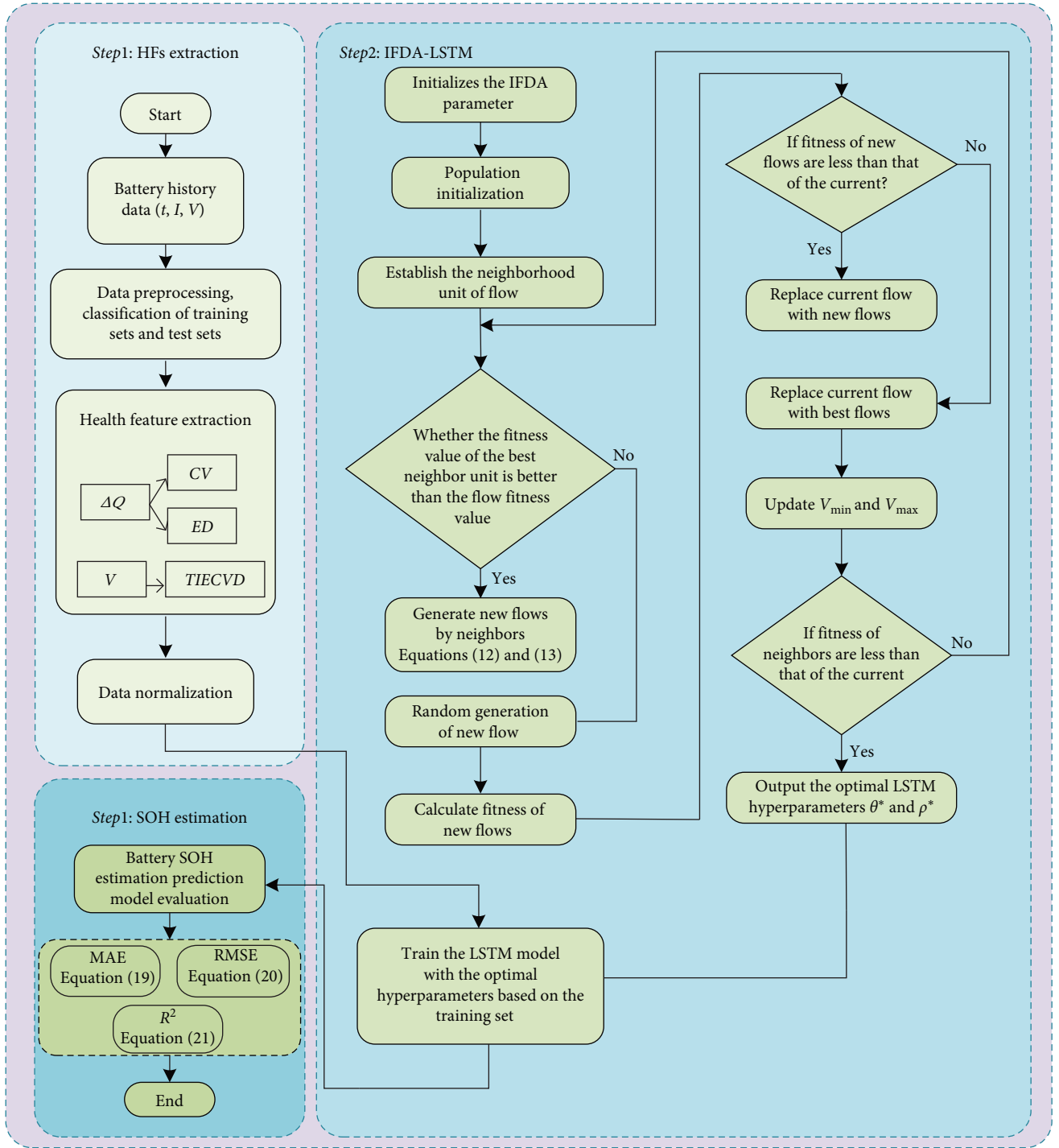
FIGURE 10: Flowchart of SOH estimation based on IFDA-LSTM model.

$$V\_{max} = \begin{cases} p \times (u_b - l_b) & \text{iter} = 1 \\ \left(p - (p - q) \times \left(\dfrac{\text{iter}}{\text{Max\_Iter}}\right)^2\right) \times (u_b - l_b) & \text{iter} > 1 \end{cases},$$

$$(21)$$

$$V\_{max} = \begin{cases} -p \times (u_b - l_b) & \text{iter} = 1 \\ -\left(p - (p - q) \times \left(\dfrac{\text{iter}}{\text{Max\_Iter}}\right)^2\right) \times (u_b - l_b) & \text{iter} > 1 \end{cases},$$

$$(22)$$

TABLE 3: Parameters setups of the models.

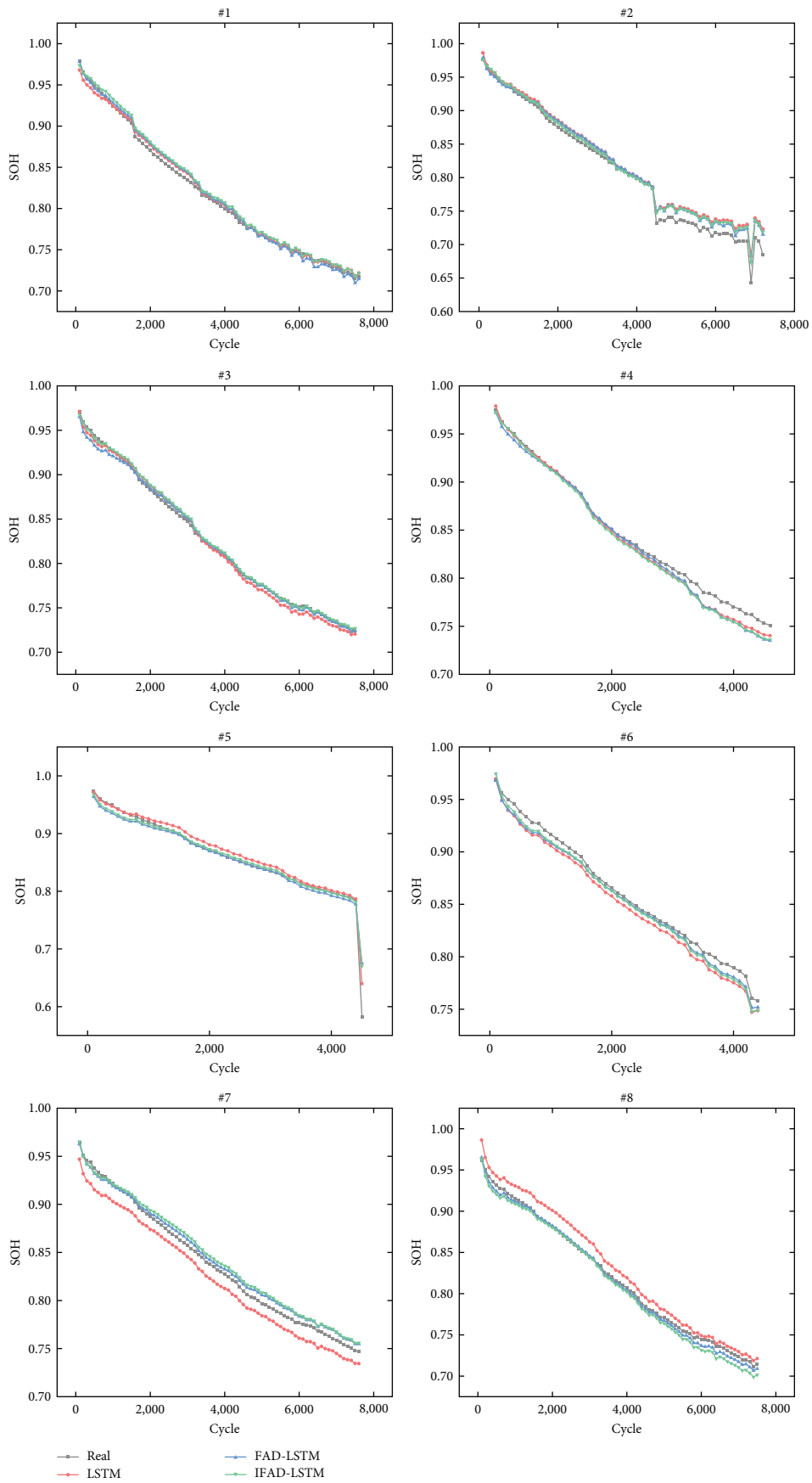| Description	Value
Flow number	20
Number of neighborhoods	1
Max number of iterations	10
Batch size	30
Input layer	CV-ED-TIECVD
Output layer	SOH

FIGURE 11: Battery SOH estimation results based on different methods.

TABLE 4: SOH estimation errors for Kokam cells.

Batteries	MAE (%)			RMSE (%)			R^2		
	LSTM	FDA-LSTM	IFDA-LSTM	LSTM	FDA-LSTM	IFDA-LSTM	LSTM	FDA-LSTM	IFDA-LSTM
#1	1.14	0.35	0.34	1.34	0.60	0.39	0.97	0.98	0.99
#2	1.19	0.93	0.76	1.48	1.14	0.98	0.97	0.98	0.99
#3	1.04	0.34	0.30	1.60	0.42	0.34	0.95	0.97	0.99
#4	1.08	0.62	0.43	1.22	0.79	0.56	0.97	0.98	0.99
#5	0.84	0.53	0.49	1.47	1.30	1.02	0.95	0.96	0.98
#6	1.52	0.75	0.25	1.06	0.79	0.34	0.97	0.98	0.99
#7	1.92	0.57	0.45	1.49	0.62	0.48	0.94	0.97	0.99
#8	1.78	0.85	0.42	1.34	1.01	0.51	0.97	0.98	0.99
Mean	1.31	0.62	0.43	1.38	0.83	0.58	0.96	0.98	0.99

where p and q are the maximum and minimum values of random inertia weights, which are taken as 0.1 and 0.01, respectively, Max_Iter is the maximum number of iterations, and iter is the current number of iterations.

During the flow of water, since the fitness value of any neighboring water flow cannot be smaller than the fitness value of the current water flow, so if the fitness value of the domain water flow is smaller than the fitness value of the current water flow, then it will flow in the same direction as the current water flow. Otherwise, it will flow in the direction of the dominant slope, as shown in Equation (23):

$$
\begin{cases}
\text{Flow\_new}X(i) = \text{Flow\_}X(i) + \text{randn}(1, \dim) \times (\text{Flow\_}X(r) - \text{Flow\_}X(i)) \\
\text{Flow\_new}X(i) = \text{Flow\_}X(i) + \text{randn}(1, \dim) \times (\text{Best\_}X - \text{Flow\_}X(i))
\end{cases}, \quad (23)
$$

(C) During the global position update process, the neighborhood at the optimal position is perturbed so that the convergence rate of the algorithm can be increased and premature maturation of the algorithm can be avoided. The definition is shown in Equation (24):

$$
\begin{cases}
\text{Best\_}X\_\text{new}(i) = \text{Best\_}X(i) + 0.5 \times \text{rand}(1, \dim) & \text{rand} < 0.5 \\
\text{Best\_}X\_\text{new}(i) = \text{Best\_}X(i) & \text{rand} \geq 0.5
\end{cases}, \quad (24)
$$

where Best_X_new is the generated new individual, Best_X is the best individual at the current global update, and rand is a random number with uniform distribution of (0, 1).

Usually, more newly generated neighborhood positions are obtained after optimal neighborhood perturbation, and then a greedy strategy is needed to perform shrinkage ranking and determine whether to keep them. The greedy strategy formula is shown in Equation (25):

$$
\begin{cases}
\text{Best\_}X\_\text{final}(i) = \text{Best\_}X\_\text{new} & f(\text{Best\_}X\_\text{new}) < f(\text{Best\_}X) \\
\text{Best\_}X\_\text{final}(i) = \text{Best\_}X & f(\text{Best\_}X\_\text{new}) \geq f(\text{Best\_}X)
\end{cases}, \quad (25)
$$

where Best_X_final is the optimal individual after the shrinkage ranking of the greedy strategy, and $f(i)$ is the *iter* water flow adaptation value. If the generated water flow adaptation value is smaller than the adaptation value of the original best individual, it replaces the original best individual as the global best individual in the iterations iter; if it is not as good as the original best individual, the global best individual does not change.
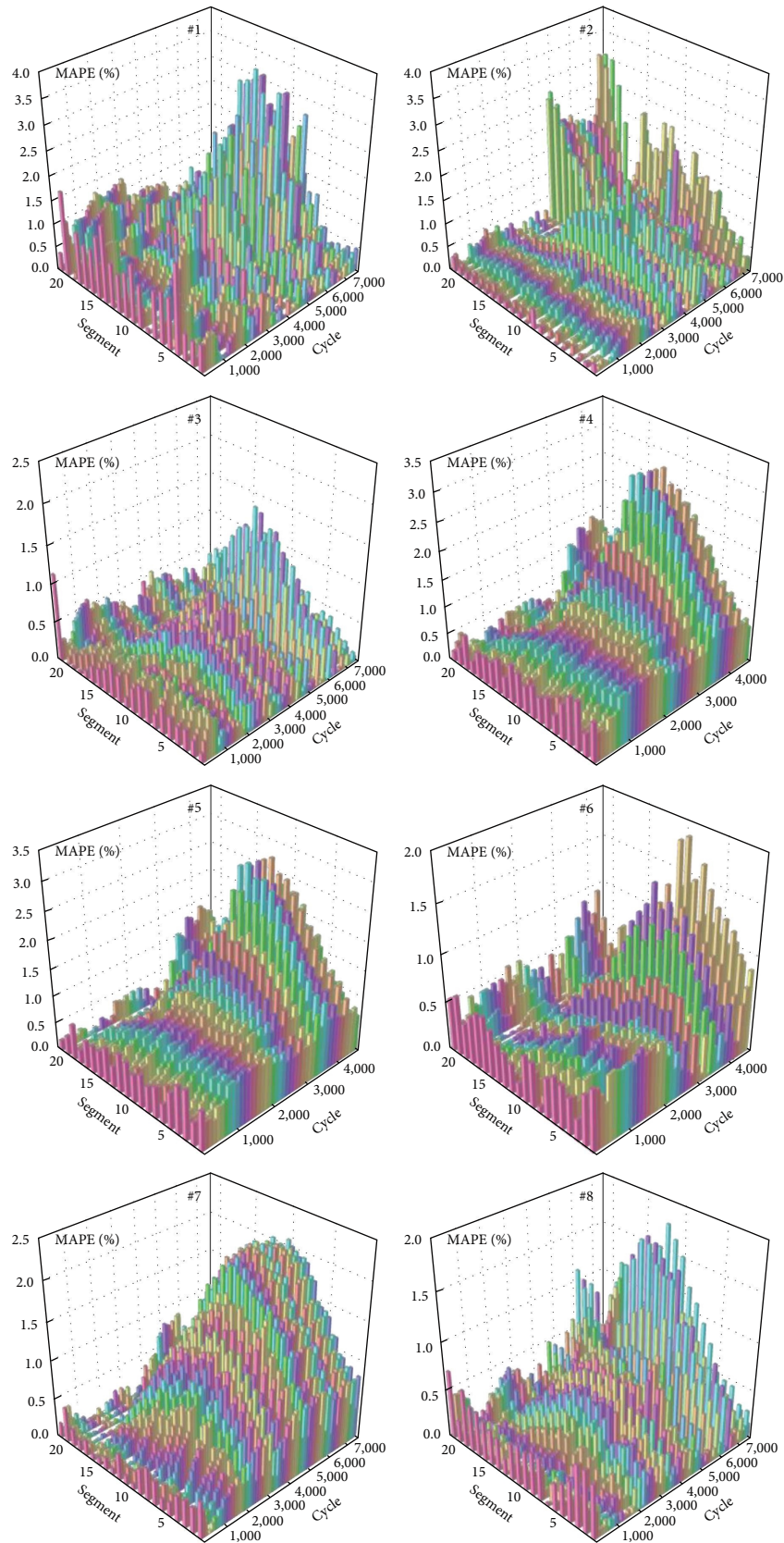
FIGURE 12: Absolute percentage of estimation error within each segment of Kokam cells.
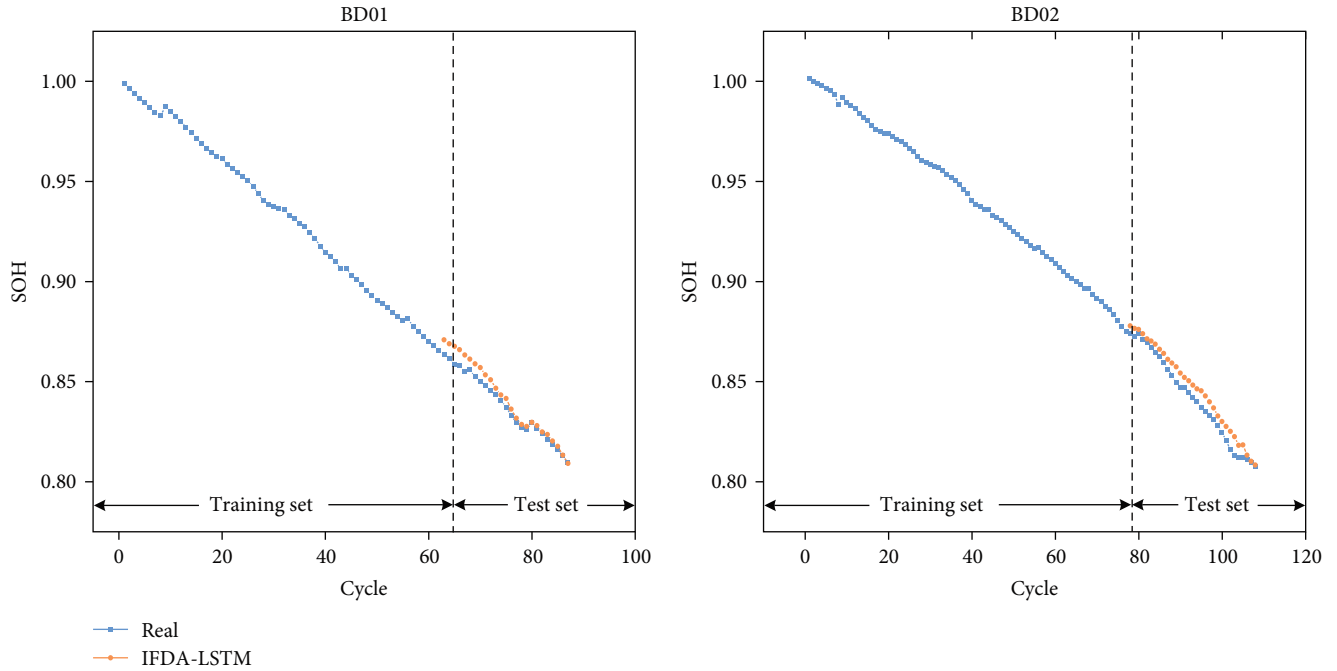
FIGURE 13: SOH estimation results of NCR18650BD battery.

4.3. Batteries SOH Estimation Process. In this paper, the number of nodes in the hidden layer and the initial learning rate of the LSTM are optimized using the IFDA to improve the prediction accuracy of the model. The IFDA can solve the problem of the joint optimization of the parameters in the LSTM neural network model and enhance the estimation accuracy of the network. Compared with the standard FDA, IFDA not only makes the iterative process of model parameters more stable but also significantly faster in finding the optimal global solution. The SOH estimation process of the IFDA-LSTM model is shown in Figure 10. The process is as follows:

Step 1: processing of data acquired in the laboratory;

Step 2: the data were analyzed to extract HFs from the ΔQ curves and voltage curves and the extracted features were normalized;

Step 3: based on the optimized IFDA algorithm in Section 4.2, the hyperparameters of the LSTM neural network are optimized to output the runoff $pso^* = (\theta^*, \rho^*)$ with the best fitness and location. The θ^* and ρ^* are the hyperparameters of the LSTM model;

Step 4: initialize the LSTM parameters and train the model $\text{LSTM}_{\text{best}}$ based on the training set;

Step 5: input test sets data, model output estimation result and error.

Some of the network parameters are shown in Table 3.

The SOH estimation was performed on the test set data using the trained estimation model, and the estimation results were evaluated by error analysis using three metrics: mean absolute error (MAE), root-mean-square error (RMSE),

and correlation coefficient (R^2).

$$\text{MAE} = \frac{1}{M} \sum_{i=1}^{M} \text{SOH}_i' - \text{SOH}_i|, \tag{26}$$

$$\text{RMSE} = \sqrt{\frac{1}{M} \sum_{i=1}^{M} (\text{SOH}_i' - \text{SOH}_i)^2}, \tag{27}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{M} (\text{SOH}_i' - \text{SOH}_i)^2}{\sum_{i=1}^{M} (\overline{\text{SOH}_i} - \text{SOH}_i)^2}, \tag{28}$$

where SOH_i, SOH_i' and $\overline{\text{SOH}_i}$ denote the baseline, estimated, and mean values, respectively, and M denotes the number of test samples. Smaller MAE and RMSE values indicate that the estimation error of the method is smaller. Closer to 1 for $R^2 \in (0, 1)$ represents the better model estimation.

5. Simulation Results Analysis and Discussion

This section focuses on the feasibility of the proposed SOH estimation approach based on the random charge segments, including the SOH estimation of different batteries in different segments and the result analysis.

5.1. Validation of the Estimated Results from the Oxford Dataset. To assess the performance of the proposed IFDA-LSTM approach, two estimation methods are employed for comparison, namely, LSTM and FDA-LSTM. Where IFDA optimizes the number of nodes in the implicit layer and the initial learning rate of the LSTM to improve the prediction accuracy of the model. In this section, the leave-one-out cross-validation (LOOCV) method is used to traverse the eight batteries in the
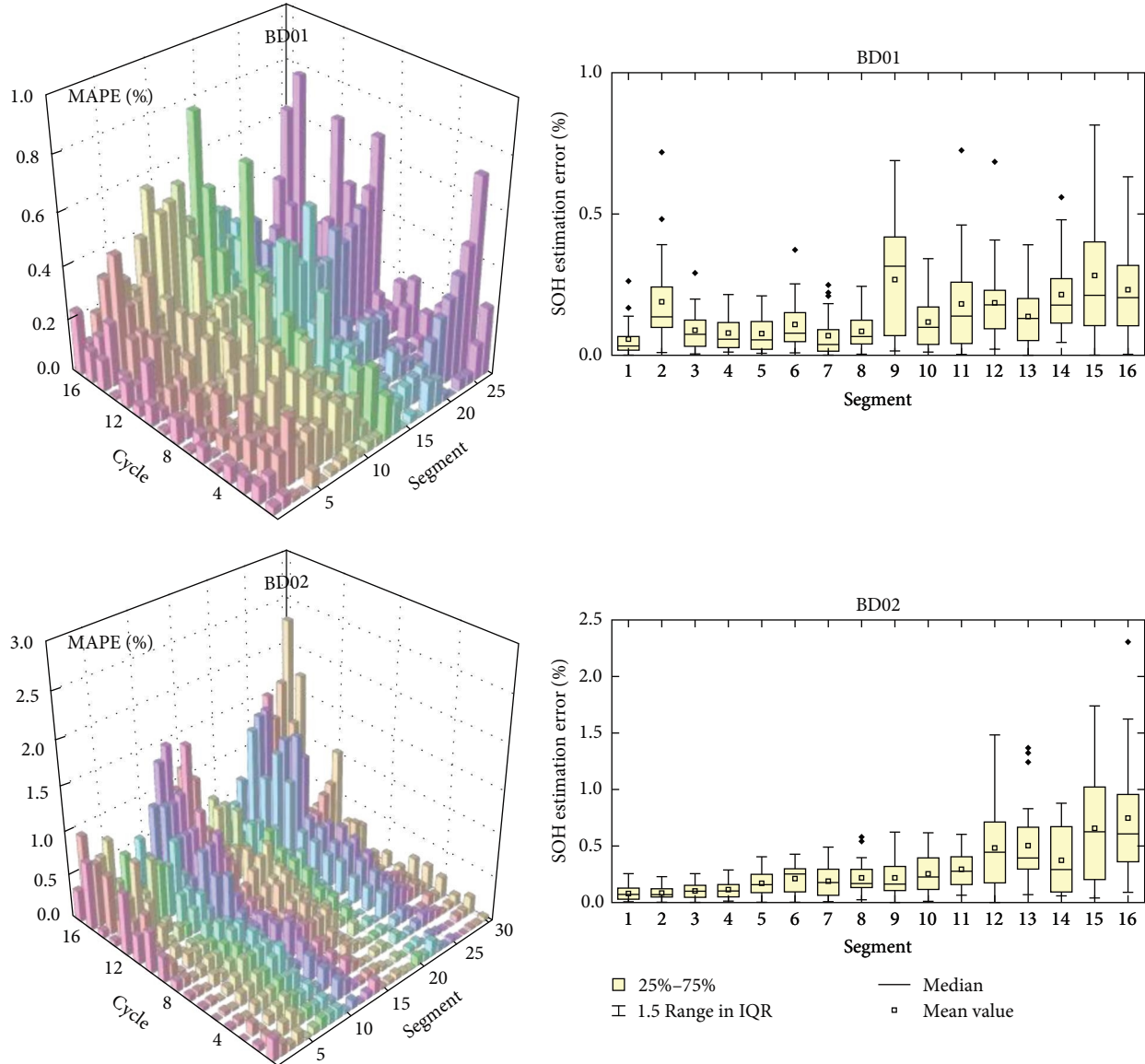
FIGURE 14: Estimation error results within each segment of NCR18650BD battery.

Oxford degradation dataset. The estimated results for batteries in the Oxford dataset #1–#8 are shown in Figure 11.

As shown in Figure 11, the IFDA-LSTM proposed in this paper exhibits high accuracy on all the cells and has higher reliability and performance compared to the several other methods.

Then, the SOH estimation error results for the eight Kokam cells in the Oxford dataset are analyzed. The MAE, RMSE, and R^2 of SOH estimation for cells #1–#8 under LSTM, FDA-LSTM, and IFDA-LSTM are given in Table 4. From Table 4, it can be seen that the estimation errors of the LSTM and FDA-LSTM algorithms are within a certain range, and the estimation results show better results. the MAE, RMSE, and R^2 of IFDA-LSTM are 0.43%, 0.58%, and 0.99%, with the smallest estimation errors, all outperforming several other methods.

To more visually depict the error for each segment, Figure 12 shows a 3D strip plot of the absolute percentage

of the estimated error within the segments using the eight Kokam cells in the Oxford University dataset. The errors are plotted as a function of the number of segments and cycles, where the number of segments per cell is 21.

As shown in Figure 12, the SOH estimation based on random charging segments on the eight cells from #1 to #2.

Performed well overall within each segment. The absolute average estimation errors for each cell were 0.341%, 0.492%, 0.303%, 0.963%, 0.955%, 0.365%, 0.785%, and 0.342%, respectively. Although the estimation errors of the partial cycle count for some segments were relatively large, they were still within acceptable limits, and most of the estimation results were satisfactory.

5.2. Validation of SOH Estimation Based on the Experimental Dataset. To further validate the applicability of the proposed method, SOH estimation is performed on two NCR18650BD batteries, and the first 70% of the historical data are selected

as the training set and the last 30% of the historical data are used as the test set for SOH estimation using the proposed IFDA-LSTM. The SOH estimation results are given in Figure 13. The figure reveals that the proposed IFDA-LSTM has satisfactory accuracy on both of the two NCR18650BD batteries.

The experimentally measured estimated errors within each fragment of the NCR18650BD battery are plotted as 3D bar graphs; as shown in Figure 14, where the number of segments per cell is 16.

As shown in Figure 14, the estimation accuracy of BD01 is particularly impressive in the SOH estimation of the random charging section of the two batteries, with an average estimation error of 0.15% absolute and a maximum estimation error of only 0.81%. The estimation accuracy of BD02 is slightly lower than that of BD01, but it still performs well, with an average estimation error of 0.29% in absolute value and a maximum estimation error of 2.31%.

In summary, the effectiveness of the optimization algorithm proposed in this paper is verified by comparing it with other algorithms. The effectiveness of the proposed SOH estimation method based on random charging segments is demonstrated in the Kokam and NCR18650BD battery.

6. Conclusion

In this paper, a new SOH estimation method based on CV-ED-TIECVD HFs and the IFDA-LSTM estimation model for random charging data of Li-ion batteries are proposed. The method divides the charging curve into several equal voltage difference segments by analyzing the charging process of different degradation batteries, extracts the TIECVD, CV, and ED as HFs, and form a set of CV-ED-TIEVD HFs within the random segments.

The performance of the proposed SOH estimation method based on the random charging segments is validated by the Oxford dataset and experimental dataset, and the results reveal that the proposed method boasts high accuracy and reliability. The results show that the MAE stays within 0.43% and 0.32%, and the RMSE stays within 0.58% and 0.43% for both the Oxford dataset and the experimental dataset. In addition, the proposed method maintains good estimation performance for different datasets. Thus, using CV-ED-TIECVD as HFs makes the features cover more complete information and the results are presented better. The IFDA-LSTM estimation model can accurately estimate the SOH of the battery in short and random charging segment after obtaining HFs charging data.

Data Availability

Datasets were derived from the following public domain resources: https://ora.ox.ac.uk/objects/uuid:03ba4b01-cfed-46d3-9b1a-7d4a7bdf6fac.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Aina Tian contributed in the methodology, writing–- original draft preparation, visualization. Zhe Chen contributed in the data curation, writing–reviewing and editing, and visualization. Zhuangzhuang Pan contributed in the conceptualization and experimental propulsion. Chen Yang contributed in writing guide. Yuqin Wang and Kailang Dong contributed in the software. Yang Gao contributed in the resources. Jiuchun Jiang contributed in the supervision.

Acknowledgments

References

[1] M. S. Whittingham, "Ultimate limits to intercalation reactions for lithium batteries," *Chemical Reviews*, vol. 114, no. 23, pp. 11414–11443, 2014.

[2] A. Tian, C. Yang, Y. Gao et al., "A state of health estimation method of lithium-ion batteries based on DT-IC-V health features extracted from partial charging segment," *International Journal of Green Energy*, vol. 20, no. 9, pp. 997–1011, 2023.

[3] X. Lai, Y. Huang, X. Han, H. Gu, and Y. Zheng, "A novel method for state of energy estimation of lithium-ion batteries using particle filter and extended Kalman filter," *Journal of Energy Storage*, vol. 43, Article ID 103269, 2021.

[4] A. Eddahech, O. Briat, N. Bertrand, J.-Y. Delétage, and J.-M. Vinassa, "Behavior and state-of-health monitoring of li-ion batteries using impedance spectroscopy and recurrent neural networks," *International Journal of Electrical Power & Energy Systems*, vol. 42, no. 1, pp. 487–494, 2012.

[5] M. Berecibar, I. Gandiaga, I. Villarreal, N. Omar, J. Van Mierlo, and P. Van den Bossche, "Critical review of state of health estimation methods of li-ion batteries for real applications," *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 572–587, 2016.

[6] S. Zhang, B. Zhai, X. Guo, K. Wang, N. Peng, and X. Zhang, "Synchronous estimation of state of health and remaining useful lifetime for lithium-ion battery using the incremental capacity and artificial neural networks," *Journal of Energy Storage*, vol. 26, Article ID 100951, 2019.

[7] X. Shu, S. Shen, J. Shen et al., "State of health prediction of lithium-ion batteries based on machine learning: advances and perspectives," *iScience*, vol. 24, no. 11, Article ID 103265, 2021.

[8] L. Ungurean, G. Cârstoiu, M. V. Micea, and V. Groza, "Battery state of health estimation: a structured review of models, methods and commercial devices," *International Journal of Energy Research*, vol. 41, no. 2, pp. 151–181, 2017.

[9] M. Zhang, Y. Liu, D. Li et al., "Electrochemical impedance spectroscopy: a new chapter in the fast and accurate estimation of the state of health for lithium-ion batteries," *Energies*, vol. 16, no. 4, Article ID 1599, 2023.

[10] D. Wang, H. Huang, Z. Tang, Q. Zhang, B. Yang, and B. Zhang, "A lithium-ion battery electrochemical–thermal model for a wide temperature range applications," *Electrochimica Acta*, vol. 362, Article ID 137118, 2020.

[11] J. Wang, P. Liu, J. Hicks-Garner et al., "Cycle-life model for graphite-LiFePO$_4$ cells," *Journal of Power Sources*, vol. 196, no. 8, pp. 3942–3948, 2011.

[12] X. Yu, N. Ma, L. Zheng, L. Wang, and K. Wang, "Developments and applications of artificial intelligence in music education," *Technologies*, vol. 11, no. 2, Article ID 42, 2023.

[13] Y. Zhang, Y. Liu, J. Wang, and T. Zhang, "State-of-health estimation for lithium-ion batteries by combining model-based incremental capacity analysis with support vector regression," *Energy*, vol. 239, Part B, Article ID 121986, 2022.

[14] Z. Wang, C. Yuan, and X. Li, "Lithium battery state-of-health estimation via differential thermal voltammetry with gaussian process regression," *IEEE Transactions on Transportation Electrification*, vol. 7, no. 1, pp. 16–25, 2021.

[15] Z. Chen, N. Shi, Y. Ji, M. Niu, and Y. Wang, "Lithium-ion batteries remaining useful life prediction based on BLS-RVM," *Energy*, vol. 234, Article ID 121269, 2021.

[16] G. Dong, W. Han, and Y. Wang, "Dynamic bayesian network-based lithium-ion battery health prognosis for electric vehicles," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 11, pp. 10949–10958, 2021.

[17] B. Gou, Y. Xu, and X. Feng, "An ensemble learning-based data-driven method for online state-of-health estimation of lithium-ion batteries," *IEEE Transactions on Transportation Electrification*, vol. 7, no. 2, pp. 422–436, 2021.

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

[19] M. Zhang, D. Yang, J. Du et al., "A review of SOH prediction of li-ion batteries based on data-driven algorithms," *Energies*, vol. 16, no. 7, Article ID 3167, 2023.

[20] J. Wen, X. Chen, X. Li, and Y. Li, "SOH prediction of lithium battery based on IC curve feature and BP neural network," *Energy*, vol. 261, Part A, Article ID 125234, 2022.

[21] Y. Gong, X. Zhang, D. Gao et al., "State-of-health estimation of lithium-ion batteries based on improved long short-term memory algorithm," *Journal of Energy Storage*, vol. 53, Article ID 105046, 2022.

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

[23] D. Liu, J. Zhou, H. Liao, Y. Peng, and X. Peng, "A health indicator extraction and optimization framework for lithium-ion battery degradation modeling and prognostics," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 6, pp. 915–928, 2015.

[24] M. A. Patil, P. Tagade, K. S. Hariharan et al., "A novel multistage support vector machine based approach for li ion battery remaining useful life estimation," *Applied Energy*, vol. 159, pp. 285–297, 2015.

[25] X. Shu, G. Li, Y. Zhang, J. Shen, Z. Chen, and Y. Liu, "Online diagnosis of state of health for lithium-ion batteries based on short-term charging profiles," *Journal of Power Sources*, vol. 471, Article ID 228478, 2020.

[26] C. Qian, B. Xu, L. Chang et al., "Convolutional neural network based capacity estimation using random segments of the charging curves for lithium-ion batteries," *Energy*, vol. 227, Article ID 120333, 2021.

[27] D. Yu, D. Ren, K. Dai et al., "Failure mechanism and predictive model of lithium-ion batteries under extremely high transient impact," *Journal of Energy Storage*, vol. 43, Article ID 103191, 2021.

[28] J. Wu, X. Cui, H. Zhang, and M. Lin, "Health prognosis with optimized feature selection for lithium-ion battery in electric vehicle applications," *IEEE Transactions on Power Electronics*, vol. 36, no. 11, pp. 12646–12655, 2021.

[29] Z. Deng, X. Hu, P. Li, X. Lin, and X. Bian, "Data-driven battery state of health estimation based on random partial charging data," *IEEE Transactions on Power Electronics*, vol. 37, no. 5, pp. 5021–5031, 2022.

[30] Z. Deng, X. Hu, Y. Xie et al., "Battery health evaluation using a short random segment of constant current charging," *iScience*, vol. 25, no. 5, Article ID 104260, 2022.

[31] A. H. Mirza, M. Kerpicci, and S. S. Kozat, "Efficient online learning with improved LSTM neural networks," *Digital Signal Processing*, vol. 102, Article ID 102742, 2020.

[32] H. Karami, M. V. Anaraki, S. Farzin, and S. Mirjalili, "Flow direction algorithm (FDA): a novel optimization approach for solving optimization problems," *Computers & Industrial Engineering*, vol. 156, Article ID 107224, 2021.