Int. J. Electrochem. Sci., 17 **(2022)** Article Number: 220725, doi: 10.20964/2022.07.38

International Journal of
ELECTROCHEMICAL
SCIENCE
www.electrochemsci.org

SOH Estimation Method for Lithium-ion Battery Based on Discharge Characteristics

Zhilong Yu[1], *Yekai Zhang*[1], *Lihua Qi*[1], *Ran Li*[2]

[1] College of Automation, Harbin University of Science and Technology, Harbin 150000, Heilongjiang Province, China
[2] Ministry of Education of Automotive Electronics Drive Control and System Integration, Harbin University of Science and Technology, Harbin 150000, Heilongjiang Province, China
[*]E-mail: 2634386312@qq.com

Received: 6 March 2022 / *Accepted:* 19 April 2022 / *Published:* 6 June 2022

In order to accurately estimate the state of health (SOH) of lithium batteries under the condition of constant current charging and discharging, a multi-scale health indicator (HI) selection strategy based on discharge characteristics is proposed in this paper, which can estimate SOH when the battery is not fully charged and discharged. Firstly, the initial voltage drop in the discharge process and the time interval of an equal discharging voltage difference were treated as the indirect HIs of battery capacity estimation. Secondly, the optimal discharge voltage segment was selected using the genetic algorithm based on elitist model (e-GA), the selection criterion of time interval of an equal discharging voltage difference was determined, and its correctness was verified. Finally, a lithium battery degradation model was established by the long and short term memory (LSTM) neural network according to the HI selection strategy as mentioned above, and the experimental results were compared with those as obtained using other selection strategies. As suggested by the experimental results, the aforementioned HI selection strategy can reduce the mean absolute percentage error (MAPE) by 0.41 % on average. In addition, the MAPE of the prediction results fell below 2.22 % in twelve experiments conducted on four batteries at different allocation ratios of training set and test set.

Keywords: Lithium-ion battery; State of health; Elitist model based GA; Health indicator; Long and short term memory neural network

1. INTRODUCTION

Lithium-ion battery is an advanced green energy storage battery, with high specific energy, low self-discharge rate, high safety, long cycle life and other advantages [1]. However, in the long-term operation of any battery, it is inevitable to have attenuation of energy and capacity, and abnormal

attenuation will lead to safety accidents [2]. Therefore, in order to ensure the reliable operation of lithium batteries, there must be a way to determine the health of batteries [3].

The SOH of a battery reflects its service life. It is usually defined as the ratio of the capacity released by the battery from the full charge state to the cut-off voltage at a certain rate to the nominal capacity. It is generally believed that the end-of-life condition is reached when the capacity of the battery declines to 80% of the nominal capacity [4]. Traditionally, the current method, open-circuit voltage method and other methods based on empirical knowledge are combined to estimate SOH of lithium battery, but these methods are time-consuming and have low accuracy, making it difficult to apply them in practice [3]. At present, there are mainly model method and data-driven method to estimate the health status of lithium batteries [5-11]. Among them, the model method is divided into electrochemical mode method and equivalent model method [12-14]. Due to the complex structure and large amounts of calculation, electrochemical model method is rarely applied. Equivalent model method is often combined with Kalman filter [15,16], particle filter and other algorithms [17,18], and the accuracy of its estimation is heavily dependent on the accuracy of model parameters. As a result, the model parameters cannot be updated in real time when battery aging and other factors occur, thus leading to the decline in accuracy of estimation. Data-driven method does not require in-depth understanding as to the internal mechanism of lithium batteries. Instead, it uses machine learning algorithm to build a model based on the historical data of batteries and tries to find the hidden correlation from a large number of data, so that it is widely used with the characteristics of high efficiency and practicality [19].

The data driven modeling based on machine learning has been reported in plenty of literature. Zhou et al [20] used genetic algorithm and Person index to select the optimal charging voltage segment as an indirect health factor to estimate the SOH of batteries, but did not specify the iterative method of genetic algorithm. Person coefficient is a linear correlation coefficient, while for neural network, the mapping between health factors and capacity can be nonlinear. Therefore, the discharge capacity of the optimal voltage segment is not always linearly related to the capacity. Liu et al [21] used the average voltage and temperature under the condition of constant discharge current to map the SOH of the batteries, the selection of its health factors was not targeted, and it was necessary to track the whole discharge process. Li [22] and Chen et al [23] selected the time interval of an equal discharging voltage difference as the input and the actual capacity of the battery as the output to build a degradation model of lithium batteries. The strategy of selecting HI from a single scale is insufficient to predict complex capacity changes, and the generalization ability is weak because the time interval of an equal discharging voltage difference is selected within a fixed voltage segment. Chen et al [11] used the time point of discharge dip as a health indicator, but it was necessary to reduce the prediction error. Jiang et al [24] used ELM network to build a battery degradation model, but the charging and discharging data of lithium batteries were time-series. Besides, the ELM network had poor memory and failed to fully learn the time-series characteristics of the data.

After the advantages and disadvantages of existing methods were analyzed, a data-driven SOH estimation method intended for lithium batteries under the condition of constant current charging and discharging was proposed. Firstly, the initial discharge voltage drop and the time interval of an equal discharging voltage difference were selected as the indirect HIs to map the actual battery capacity at multiple scales. Then, the e-GA algorithm combined with correlation analysis method was used to

optimize the time interval of an equal discharging voltage difference which is one of the HIs, and the HI with the highest correlation with capacity was verified. Finally, a lithium battery degradation model was established by using the above selection strategy of HI to accurately estimate the SOH of batteries.

2. SELECTION OF INDIRECT HI

In order to study the health status of lithium battery, it is necessary to select an appropriate HI to characterize the SOH of battery. HI can be divided into direct HI and indirect HI. Direct HI is the actual capacity of the battery in its current state, but most of it can only be obtained in the laboratory because it is not easy to measure it in real working conditions, which makes it applicable as the basis to measure the accuracy of the prediction. Indirect HI includes terminal voltage, current, temperature, internal resistance and so on. The relationship model between indirect HI and battery capacity was established to predict the actual battery capacity according to partial discharge characteristic.

2.1. Thevenin equivalent model of lithium battery

In the analysis of lithium battery, it is always approximately regarded as the Thevenin equivalent model as shown in Fig.1, where V_{ocv} is the open circuit voltage of the battery, V_d is the terminal voltage of the battery, R is the internal resistance of the power supply, R_1 and C_1 are the internal resistance and capacitance of the battery respectively, and I is the current in the battery.

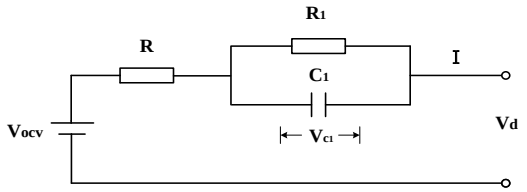

Figure 1. Thevenin equivalent model of lithium-ion battery

According to the equivalent model shown in Fig.1, it is known that before the battery discharge occurs:

$$V_{ocv} = V_d \quad (1)$$

Steady battery constant discharge stage:

$$V_{ocv} - I(R + R_1) = V_d \quad (2)$$

$$V_{ocv} = f(SOC) \quad (3)$$

After discharge:

$$V_{ocv} - V_{c1} \approx V_d \quad (4)$$

Where, V_{c1} is generated by the polarization reaction of the polarization capacitor C_1 and gradually declines to zero with the progress of the polarization reaction.

2.2. Selection of HI

More specific and effective HI can be selected by analyzing the discharge characteristics of lithium batteries. With the charging and discharging of the battery, the internal resistance of the battery will gradually rise. It can be said that there is a strong correlation between the degradation of the battery and the increase of the internal resistance *R*. Therefore, two HIs are selected as follows:

1) The first HI is the initial discharge voltage drop (V_{down}). According to Equations (1) and (2), under constant discharge current, the abrupt change in terminal voltage before and after discharge can indicate the size of *R* of the battery. Therefore, V_{down} is selected as the indirect HI. Take B0005 battery in NASA data set as an example. The terminal voltage curve during discharge was obtained as shown in Fig.2. Then, the terminal voltage value after 35s of discharge was obtained by using cubic spline interpolation to make a difference with the initial terminal voltage, and the V_{down} value was obtained as shown in Fig.3.

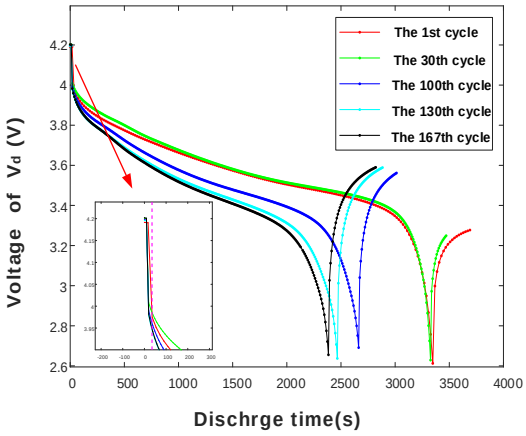

Figure 2. Lithium battery discharge characteristic curve under different numbers of charging and discharging cycles

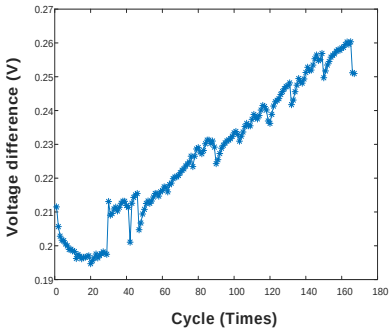

Figure 3. Relation between the terminal voltage drop value and the number of charge and discharge cycles after 35 seconds of discharge

2) The second HI is the time interval to equal discharging voltage difference (T_{AB}). Under the

condition of constant discharge current, the discharge time required for the battery terminal voltage to drop from a high voltage V_A to a low voltage V_B is arranged according to the number of cycles, and the resulting sequence is denoted as T_{AB}. According to Equations (2) and (3), T_{AB} is related to the electric quantity contained in an open-circuit voltage segment of the same span at each stage. When T_{AB} is multiplied by the discharge current, the discharge capacity in the range is C_{AB}. The fixed voltage drop range of 3.8-3.5V is selected to obtain the comparison curve between C_{AB} and the actual capacity of B0005, as shown in Fig.4. It is obvious that T_{AB} is closely related to the actual capacity of the battery. Among them, the actual capacity of the battery is measured using the ampere-hour method from the current curve in the process of discharge to the cut-off voltage in a full charge state.

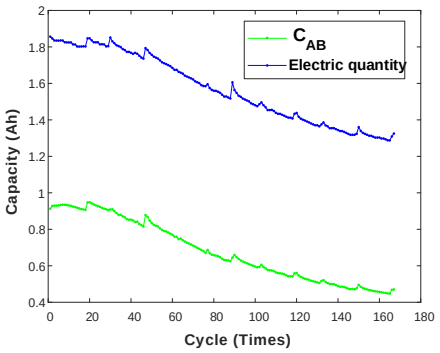

Figure 4. Curve of C_{AB} and battery capacity as the number of charge and discharge cycles increases

To sum up, V_{down} and T_{AB} were treated as indirect HI to map the actual capacity of the battery through an analysis of the discharge characteristics of the batteries.

2.3. Correlation verification

In order to validate the HI selection strategy as mentioned above, Grey relation analysis (GRA) was conducted to verify the correlation between HI and actual capacity. GRA is a method used to measure the degree of correlation among different factors according to the degree of similarity or dissimilarity of developmental trend among these factors. The basic idea is to judge whether the sequence curves are closely related according to the similarity of their geometric shapes. The closer the curves are, the greater the degree of correlation between the corresponding sequences will be; otherwise, the smaller the degree of such correlation will be [22]. The specific process is as follows:

1. The reference sequence y(k) and the comparison sequence x(k) are determined. The actual battery capacity is taken as the reference sequence, and the comparison sequence is the HI.

2. The reference sequence and comparison sequence are standardized to eliminate the differences caused by the dimensional differences between data.

3. The correlation coefficient between the reference sequence and the comparative sequence at each moment can be obtained using the following formula:

$$\xi(k)=\frac{\min\limits_{k}\left|y(k)-x(k)\right|+\rho\max\limits_{k}\left|y(k)-x(k)\right|}{\left|y(k)-x(k)\right|+\rho\max\limits_{k}\left|y(k)-x(k)\right|} \quad (5)$$

Where, ρ is the resolution coefficient, generally ranging from 0 to 1, and it is usually 0.5.

4. The correlation degree r is calculated. Since r is the correlation degree value between the comparison sequence and the reference sequence at each moment, the correlation degree can be obtained by taking the average value,

$$r(k)=\frac{1}{N}\sum_{k=1}^{N}\xi(k) \quad (6)$$

Where, N is the length of the reference sequence.

The correlation of the two selected HIs was calculated respectively, and the results are shown in Tab.1. It can be found out that the two HIs selected have a strong or very strong correlation with the actual capacity, indicating that the selection strategy of HI as mentioned above is reasonable, and the complex trend of degradation of lithium battery capacity is reflected by the multi-scale characteristics.

Table 1. Correlation between the indirect HI and actual capacity of the lithium battery

Compare sequence	Relevancy (r)
V_{down}	0.539 7
T_{AB}（3.8-3.5 V）	0.756 5

3. OPTIMIZE THE SELECTION STRATEGY OF HI

In the selection strategy of HI, V_{down} is selected according to different discharge curves, 10 to 40s voltage drop after discharge can be used, and the difference of grey correlation is not significant. However, the T_{AB} as obtained by selecting different voltage drop segments is different, and the accuracy of prediction results obtained by these segments often varies significantly. Therefore, the generalization ability of the method with fixed voltage drop segments is limited [22,23]. To solve the above problems, e-GA algorithm can be used to select the optimal pressure drop segment and maximize the correlation between T_{AB} and battery capacity sequence.

3.1. Analysis of T_{AB}

By analyzing the curve of V_{ocv} and battery remaining power (SOC), as shown in Fig.5, it can be known that when the selected battery voltage drop interval is V_A to V_B, it becomes V_A+IR to V_B+IR when mapped to open circuit voltage, as $\hat{V}_A$ to $\hat{V}_B$, where R is battery internal resistance. The conclusions can be drawn as follows. On the one hand, after the discharge voltage drop segment is mapped to the open circuit voltage, it will increase as a whole. On the other hand, with the increase of battery internal resistance *R*, the voltage will further increase after mapping.

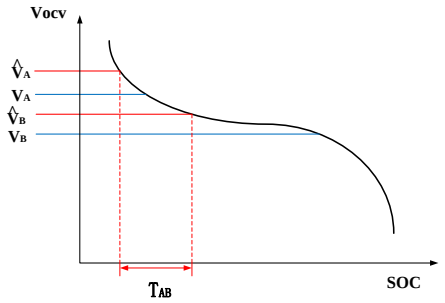

Figure 5. Relation between V_{ocv} and SOC of the lithium battery

To sum up, in the case of specified discharge fragments V_A-V_B, different cycles correspond to different open circuit voltage fragments. Therefore, the position mapped to open circuit voltage gradually shifts upward with the increase of cycles. The discharge time curves under different voltage drops are compared as shown in Fig.6. For the voltage segment of 3.8-3.5V, the voltage moves from the steep part in the front of the plateau stage, making T_{AB} decrease with the increase of cycle times, and the curve shows a downward trend. For the voltage segment of 3.5-3.2V, the voltage moves in the plateau stage and the curve is chaotic. For the voltage segment of 3.0-2.7V, the voltage moves from the steep part behind to the plateau stage, and the curve shows an upward trend. It is generally believed in literature that T_{AB} would decrease with the increase of the number of cycles [22]. However, three trends of discharge time with the number of cycles were obtained through data analysis, which increases the comprehensiveness of the analyzed problem and provides a basis for selecting the range of the optimization interval.

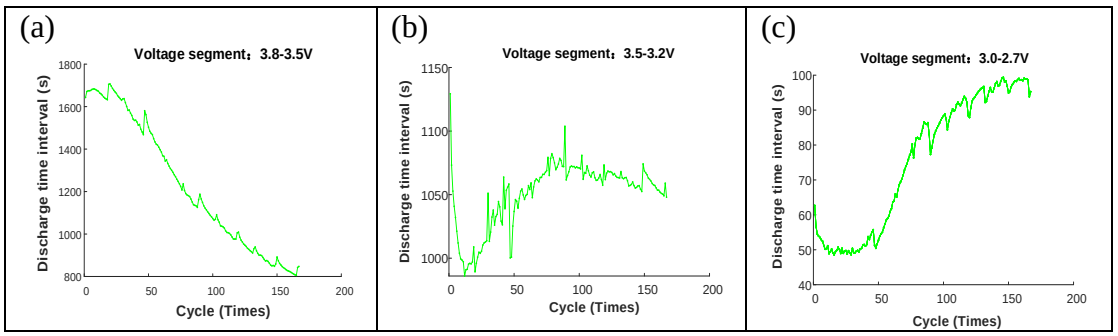

Figure 6. Relation between cycle times and discharge time under different voltage drop segments (a) 3.8-3.5V, (b) 3.5-3.2V, (c) 3.0-2.7V.

3.2. Optimal voltage drop segment selection

The selection of the optimal voltage drop segment requires comprehensive consideration, and 3.8-3.5V is not reasonable. Therefore, in the common voltage range of battery discharge, the correlation degree as obtained by GRA is used as the fitness function of e-GA algorithm to obtain the discharge

voltage drop fragment V_A to V_B corresponding to T_{AB} with the most significant correlation with the actual capacity sequence of the battery.

3.2.1.e-GA algorithm

As an improved algorithm of genetic algorithm (GA), e-GA introduces elite selection as the basic guarantee of convergence into the optimal solution to optimization problem. If the fitness of the best individual of the current population is less than that of the previous generation, the best individual of the previous generation or multiple individuals whose fitness is greater than the fitness of the current best individual will be replicated, and the corresponding number of individuals in the current population with the worst will be randomly replaced. The flow chart of its algorithm is shown in Fig.7.

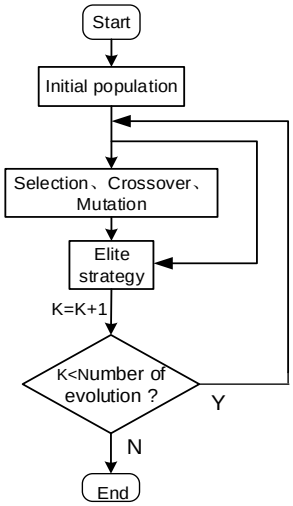

Figure 7. Flow chart of e-GA

3.2.2. Algorithm steps

1) Coding

Firstly, the appropriate voltage drop V_{A-B} and the initial voltage V_A value range are selected. When the value of V_{A-B} is too large, it is the same as the ampere-hour method. If the value of V_{A-B} is too small, the outcome of prediction is poor. Therefore, the voltage drop ranges between 0.1V and 0.2V. According to the analysis of T_{AB}, the value of V_A can be selected in the entire discharge range. However, considering the abrupt voltage drop of initial discharge and the common voltage range, the voltage segment ranging from 3.75V to 3.2V is finally selected as the search range of the optimal voltage drop segment of B0005 battery. Then, V_{A-B} and V_A are encoded in a binary way, to generate a sequence starting from 0 with a certain step size within the value range of V_{A-B} and V_A respectively, and the sequence is encoded as binary. The smaller the step size is, the higher the accuracy is. Thus, the step size is 0.01V and 0.05V, respectively.

Finally, the codes as obtained by V_{A-B} and V_A are combined and verified. Through the method of V_A minus V_{A-B}, the termination voltage V_B is obtained. If the value of V_B exceeds 3.2V, the test is successful.

2）Optimization

According to e-GA algorithm and GRA method as fitness function, the following steps are taken to find the optimal voltage drop fragments V_A and V_B:

1. The first-generation population is generated, and n individuals are randomly coded as the first-generation population. The number of evolutions k is 0.

2. Selection: The fitness of each individual and make roulette is calculated according to the fitness of each individual to select n individuals for n times.

3. Crossover: There are n random matches, and the success rate of each match is the crossover rate. After successful pairing, the binary numbers at the same position of the two individuals are randomly selected for exchange. Then, the two new individuals are tested for coding. If the two new individuals fall short of the requirements, the crossover locations are selected again for exchange until the test is successful.

4. Mutation: A position of an individual is randomly selected and the binary number of the position is inverted. The new individuals obtained are tested for coding, and those not meeting the standards need to be mutated again until the test is successful. Whether variation occurs is related to variation rate. At the same time, for the number of evolution k+1, the population becomes k+1 generation.

5. Elite strategy: The individuals with the highest fitness in the k generation population are replicated to replace the individuals with the lowest fitness in the k+1 generation.

6. The fitness of the optimal individual and the average fitness of the population are recorded, and the number of evolution of the population k is judged. If k falls below the specified number of evolution, it returns to the second step and the cycle is continued until the specified number of evolution is reached.

According to the above steps, the population number n was set to 20, the number of evolution was set to 100, and the crossover rate and mutation rate were set to 0.4 and 0.1, respectively. The relationship between the optimal individual fitness of the population and the evolutionary algebra is shown in Fig.8.

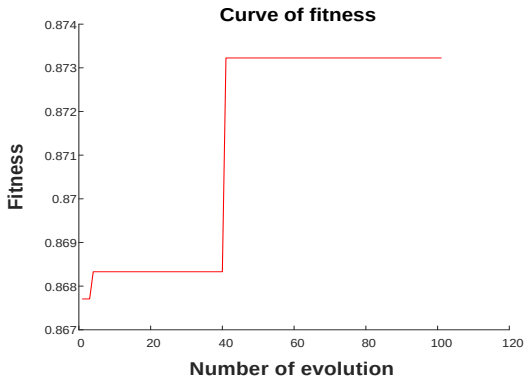

Figure 8. Population fitness changes during the optimization calculation of e-GA

According to Fig.8, the optimal individual fitness as obtained by the e-GA algorithm is 0.873 2, that is, the gray correlation between the optimal T_{AB} and the actual capacity of battery B0005 for 167 cycles, and the corresponding voltage drop range is 3.65V to 3.45V. Compared with 3.8V to 3.5V, the correlation increased by 0.116 7.

3.4. Verify the selection strategy for HI

In order to verify whether the optimized T_{AB} effectively improves the accuracy of prediction, the T_{AB} before and after optimization is selected to map the actual capacity of the battery in combination with V_{down}. The selected training set is NASA B0005 battery. It can be seen from the above section that the voltage drop interval of B0005 battery before and after T_{AB} optimization is 3.8 to 3.5V and 3.65 to 3.45V respectively, V_{down} takes the voltage drop after 35s discharge according to the above paragraph, and the test set is B0007 battery, which is similar to B0005. The prediction results of (back propagation) BP neural network are shown in Fig.9 and Fig.10, respectively. It can be seen from this figure that the optimized selection strategy produces significantly better prediction results.

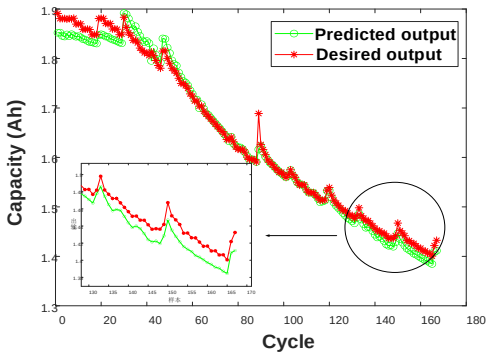

Figure 9. Real remaining capacity and the remaining capacity as calculated before optimizing HI selection strategy

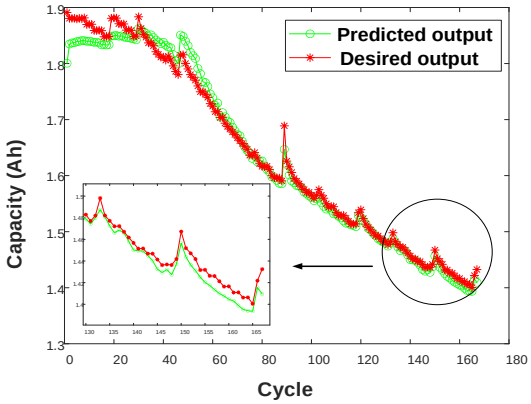

Figure 10. Real remaining capacity and the remaining capacity as calculated after optimizing HI selection strategy

4. DEGRADATION MODEL BASED ON LSTM

4.1. Time series characteristics of LSTM

For the SOH prediction of lithium batteries, due to the capacity regeneration and random interference factors in the process of battery use, there is a close relationship between the data of the last charge and discharge cycle and the data of the previous several cycles. Therefore, it is very important to understand the regularity of time sequence. The LSTM can capture these associations well.

LSTM is an improved algorithm of recursive neural network (RNN). By introducing three "gates" into neurons to control the amount of information retained by the data, the problem of long-term dependence in data is solved. These three gates are the input gate, the forgetting gate and the output gate [23]. The cell structure of LSTM neuron is shown in Fig.11.

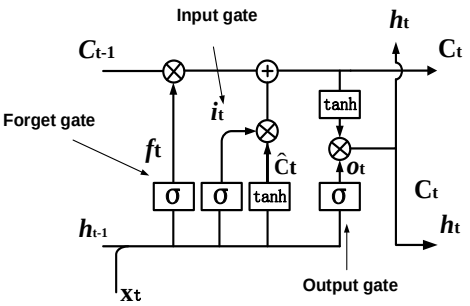

Figure 11. Cell structure diagram of LSTM neuron

Fig.11 shows the sigmoid activation function, the output of which ranges from 0 to 1, where 0 means forgetting lost data and 1 means remembering to save data, and this is where "gates" come in. The hyperbolic tangent (tanh) function, whose output range is -1 to 1, is a common activation function that can restrict data to a certain range. The three gates are generated by three functions respectively, and the specific formula is expressed as follows:

The function of the forgetting gate is to determine how much historical information is discarded by the cellular state,

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, x_t\right] + b_f\right) \quad (7)$$

Where, h_{t-1} is the output state of the last moment, x_t is the input state of the current moment, W_f is the weight matrix of forgetting gate, b_f is the amount of bias of forgetting gate, and f_t is the output of the forgetting gate.

The function of the input gate is to determine how much new information is stored in the cell state at the moment,

$$i(t) = \sigma\left(W_i \cdot \left[h_{t-1}, x_t\right] + b_i\right) \quad (8)$$

$$\hat{C}_t = \tanh\left(W_c \cdot \left[h_{t-1}, x_t\right] + b_c\right) \quad (9)$$

Where, W_i and b_i are the weight matrix and bias of the input gate respectively, W_c and b_c are the weight matrix and bias of cell state respectively, and $\hat{C}_t$ is a transient cellular state that extracts information from the current moment.

The cell state update process is shown in Equation (10). The cell state at the last moment C_{t-1} passes through the forgetting gate f_t, and the transient cell state $\hat{C}_t$ passes through the input gate i_t. The new cell state is obtained by adding the two results.

$$C_t = f_t * C_{t-1} + i_t * \hat{C}_t \quad (10)$$

The function of the output gate is to control how much information about the cell state affects the output value,

$$O_t = \sigma\left(W_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \quad (11)$$

$$h_t = O_t * \tanh\left(C_t\right) \quad (12)$$

Where, W_o and b_o are the weight matrix and bias of the output gate respectively, O_t is output value of output gate, C_t is the cell state at the current time, and h_t is the output of a single LSTM cell.

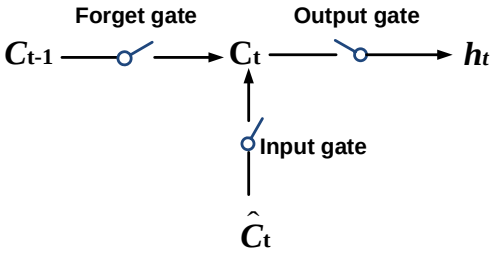

Figure 12. Schematic diagram of gate control device

To sum up, the roles played by the three gates of LSTM can be simply shown in Fig.12. LSTM can selectively store information because of its gating device, which leads to a good processing effect for time series data. Therefore, it is very reasonable to choose LSTM neural network for the SOH prediction of batteries.

4.2. Degradation model of lithium battery

Based on the multi-scale HI selection strategy as proposed above and the LSTM network, which has the ability to remember long-term information, a lithium battery degradation model can be established by collecting historical battery data under the condition of constant discharge current. The model can extract effective information from historical data and predict the next degradation trend of batteries. The degradation model of lithium battery can not only estimate the SOH of the battery effectively, but also provide strong support for the service life prediction of the battery. The structure of a LSTM network is shown in Fig.13.

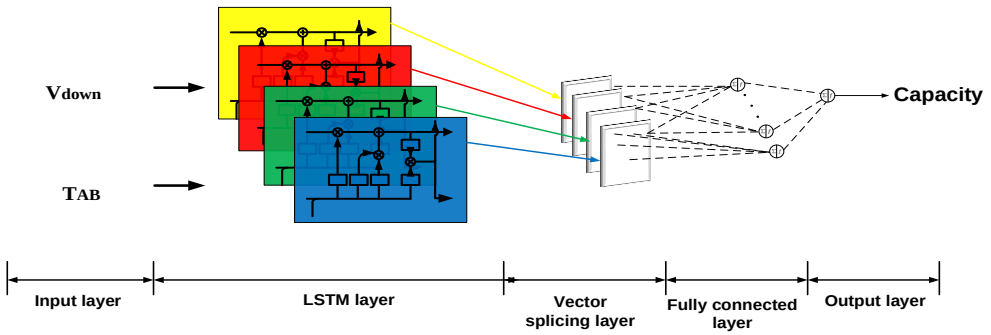

Figure 13. Schematic diagram of LSTM structure

The NASA B0005 battery data set under the condition of constant discharge current was selected as the research object, and the process of establishing the degradation model of lithium battery is shown as follows:

1) Data set proportion selection. Since the number of cycles in the whole life cycle of B0005 is 167, the amount of data is small. Besides, the smaller the proportion of training sets, the greater the significance of degradation model. Therefore, the proportion of selected data sets is 6:4. That is to say, the first 100 are training sets and the last 67 are test sets.

2) The selection of HIs. According to the HI selection strategy as mentioned above, the voltage drop after the initial 35s of the discharge characteristic of B0005 was extracted as HI. Meanwhile, the e-GA algorithm was used to work out the optimal T_{AB} in the first 100 cycles, as shown in Tab.2.

Table 2. Search interval and result of optimal interval of B0005 battery

	Discharge interval	Region of search	Optimal interval	relevancy
B0005	4.2-2.7V	3.75-3.2V	3.6-3.5V	0.8760

Table 3. LSTM hyper-parameter selection

hyper-parameter	value
Number of nodes in hidden layer	200
Number of training	100
Learning rate	The first 30 were 0.005, the rest were 0.001
Activation function of full connection layer	tanh function
Loss function	MSE
Back propagation algorithm	Adam algorithm

3) The standardization of network input. The HIs of the first 100 cycles were standardized and the mean and standard deviations were recorded.

4) LSTM network configuration. The first 100 times of the two HIs were taken as input, and the corresponding actual capacity was taken as output training neural network. Tab.3 shows the hyper-parameter configuration as determined by the number of data and other factors.

5) The output of the prediction results. The HIs of the last 67 cycles were standardized according to the recorded mean value and standard deviation. Then, they were inputted into the trained neural network. The comparison curve between the predicted capacity and the actual capacity obtained is shown in Fig.14.

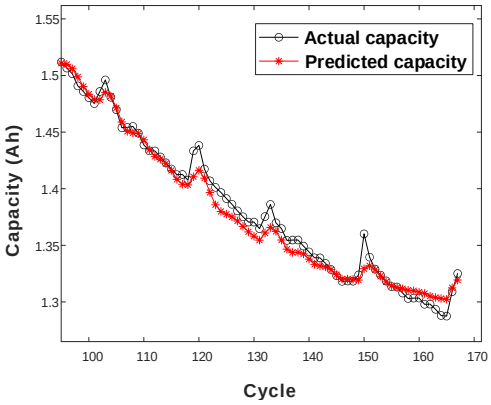

Figure 14. Prediction results of B0005 battery degradation model

According to the prediction results shown in Fig.14, the prediction accuracy of the lithium battery model subsequently constructed is very high, indicating that the battery information extracted by the selection strategy of HI is well absorbed by the LSTM network. Notably, the selected time point for observing the initial discharge voltage drop is not fixed and it can be adjusted according to the battery data and practical conditions.

5. EXPERIMENTAL SIMULATION

5.1. NASA lithium battery data set

NASA lithium battery data set was selected as the experimental object, and it is the commercially available 18650 lithium battery data set published by NASA PCoE Laboratory. The data set contains the charging, discharge and impedance curves of four lithium-ion batteries (B0005, B0006, B0007 and B0018) at room temperature. The four batteries were charged at a constant current of 1.5A until the battery voltage reached 4.2V, and then charging continued at a constant voltage until the charge current dropped to 20 mA. The discharge test was performed with a constant current of 2A until the voltage of

B0005, B0006, B0007 and B0018 dropped to 2.7, 2.5, 2.2 and 2.5 V, respectively [25]. The termination condition of the charging-discharge cycle is a 30% reduction in capacity (2 Ah to 1.4 Ah).

5.2. Comparative experiment on the selection strategies of HI

Based on the NASA lithium battery data set, batteries B0005, B0006, B0007 and B0018 were selected as the experimental subjects to compare the prediction results of four batteries given the same proportional allocation of training set and test set and the selection strategy of HI. The three specific selection strategies of HI are as follows:

1. For a single HI, the discharge time series of same voltage range T_{AB} was taken as the HI, and the range of voltage drop was fixed at 3.8-3.5V.

2. Based on strategy 1, the initial discharge pressure drop V_{down} was added as another HI.

3. Multiple HIs after optimization. The optimal voltage drop segment for T_{AB} was found on the basis of strategy 2, as shown in Tab.4.

Table 4. Search interval and result of optimal interval of four batteries

Batteries	Discharge interval	Region of search	Optimal interval	relevancy
B0005	4.2-2.7V	3.75-3.2V	3.6-3.5V	0.876 0
B0006	4.2-2.5V	3.75-3.0V	3.7-3.54V	0.857 9
B0007	4.2-2.2V	3.75-2.8V	3.75-3.56V	0.905 9
B0018	4.2-2.5V	3.75-3.0V	3.6-3.5V	0.848 0

According to the above-mentioned three selection strategies, LSTM network was used to construct a lithium battery degradation model. For batteries B0005, B0006, and B0007, the first 100 cycles are the training set and the last 67 cycles are the test set. However, the data of B0018 battery is only 130 cycles, so that the first 80 cycles are taken as the training set and the last 50 cycles are treated as the test set. The result of prediction error is shown in Tab.5.

Table 5. Comparison of the estimated errors of remaining capacity of different HI selection strategies

Batteries	MAPE of strategy 1 (%)	MAPE of strategy 2 (%)	MAPE of strategy 3 (%)
B0005	0.85	0.62	0.32
B0006	3.66	2.35	1.33
B0007	0.95	0.64	0.48
B0018	0.60	0.51	0.40

The prediction results of the last 67 or 50 cycles of the four batteries under strategies 1, 2, and 3 are shown in Fig.15, 16, 17, and 18, where the black curve indicates the actual battery capacity and the red denotes the predicted capacity. It can be seen that the prediction error will decrease with the increase

in the number and quality of HIs. The prediction accuracy of the optimized HI selection strategy, namely strategy 3, is higher, and the proposed method of SOH estimation has a certain generalization ability.

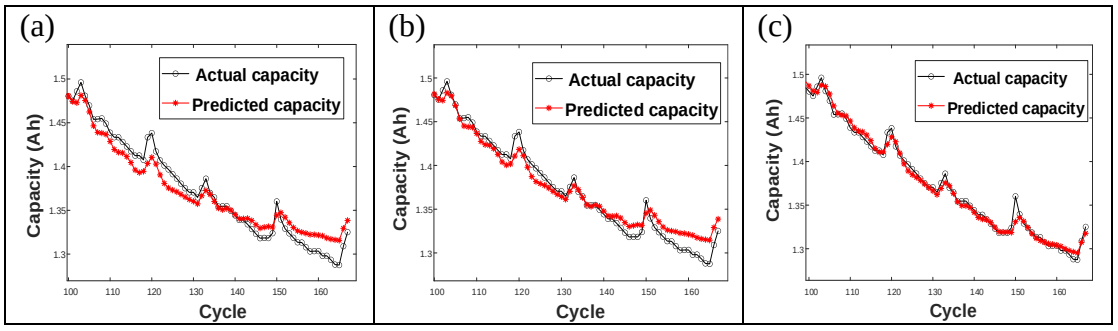

Figure 15. Prediction results of B0005 battery under three HI selection strategies (a) strategy 1, (b) strategy 2, (c) strategy 3

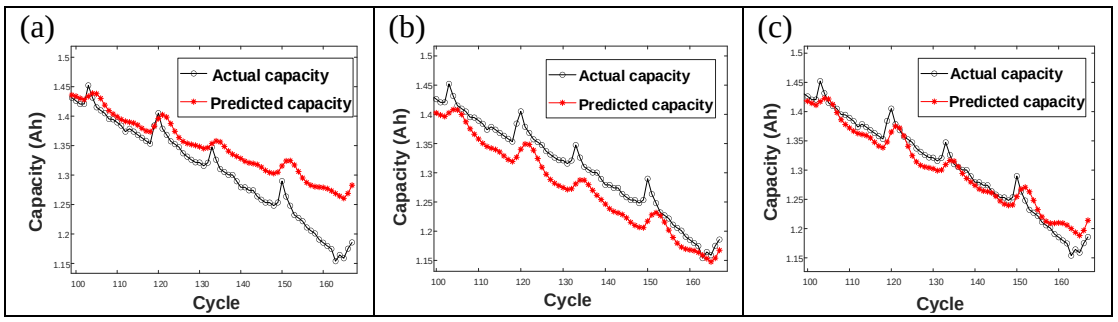

Figure 16. Prediction results of B0006 battery under three HI selection strategies (a) strategy 1, (b) strategy 2, (c) strategy 3

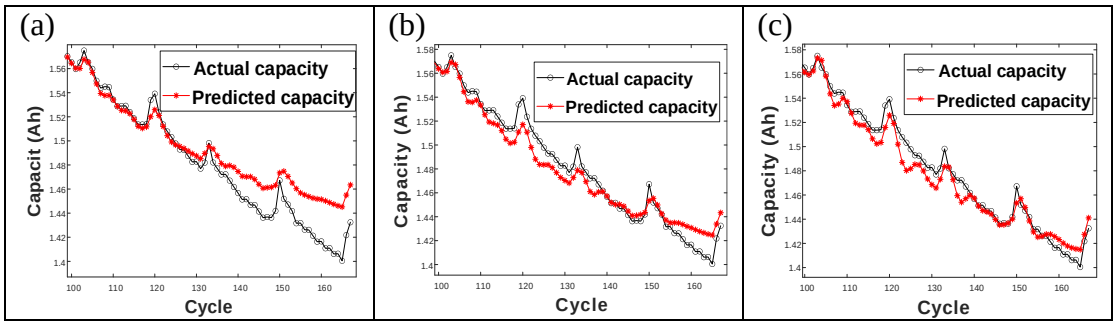

Figure 17. Prediction results of B0007 battery under three HI selection strategies (a) strategy 1, (b) strategy 2, (c) strategy 3

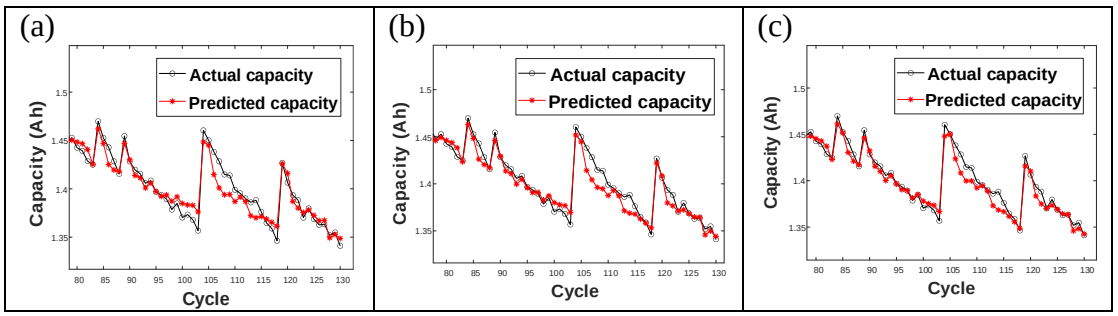

Figure 18. Prediction results of B0018 battery under three HI selection strategies (a) strategy 1, (b) strategy 2, (c) strategy 3

5.3. Influence of data set proportion on the experiment

Since different proportions of training set and test set would have a significant impact on the experimental results, strategy 2 and strategy 3 were selected to construct lithium battery degradation models given different proportions of training set and test set, so as to further explore the performance of the above selection strategies of HI in different proportions of data sets. The specific proportion of data set allocation is as follows:

1) The proportion is about 1:1. The specific proportion of B0005, B0006 and B0007 is 87:80, and it is 70:60 for B0018.

2) The proportion is about 3:2. More specifically, it is 100:67 for B0005, B0006 and B0007, and 80:50 for B0018.

3) The proportion is about 5:2. More specifically, it is 120:47 for B0005, B0006 and B0007, and 100:30 for B0018.

The obtained prediction results are shown in Tab.6. It can be seen that the error of strategy 3 is less significant than that of strategy 2 under different data set proportions, indicating that the excellent selection strategy of HI will have a positive impact on the accuracy of prediction under different conditions.

Table 6. Estimated errors given different proportions of data set allocation

MAPE	Strategy 3			Strategy 2		
(%)	Allocation 1	Allocation 2	Allocation 3	Allocation 1	Allocation 2	Allocation 3
B0005	0.63	0.32	0.23	1.78	0.62	0.41
B0006	2.22	1.33	1.01	3.27	2.35	1.77
B0007	1.32	0.48	0.25	1.36	0.64	0.46
B0018	0.56	0.40	0.48	0.61	0.51	0.44

5.4. Comparison with other models

By comparing the algorithms proposed by Li [22] and Chen [23], Table 7 shows the SOH prediction errors of different neural network models and HI selection strategies.

As shown in Table 7, the MAPE that is based on the LSTM model and combined with an optimized multi-scale HI selection strategy was smaller compared to all the other three models even if the neural network used by the model is better. By comparison, the HI selection strategy proposed in this paper introduced the initial voltage drop as a new HI, and optimized the HI T_{AB} by searching for the optimal equal voltage drop interval, which is effective in improving the prediction performance of battery SOH.

Table 7. Comparison of different SOH prediction models

Battery	Neural network	Number of HI	Equal voltage drop interval	MAPE(%)
B0005	LSTM	2	3.6-3.5V	0.32
B0005	CNN-LSTM [23]	1	3.8-3.5V	0.37
B0005	Elman [22]	1	3.7-3.5V	0.89
B0005	BP [22]	1	3.7-3.5V	2.65

6. CONCLUSION

For the accurate SOH estimation under the condition of constant electric lithium battery, an analysis was first conducted as to the characteristics of constant discharge current of lithium battery, the method of taking the initial discharge voltage drop as the HI was proposed, and the e-GA algorithm was applied to select the optimal time interval of an equal discharge voltage difference as the HI for verification. Then, the proposed single HI strategy, multiple HI strategy and optimized multiple HI strategy were integrated with LSTM neural network to construct a degradation model respectively for the prediction of battery SOH. Finally, these strategies were analyzed and compared using NASA lithium battery experiment data, which reveals that the optimized selection strategy of multiple HI improved the accuracy of prediction and demonstrated a certain generalization ability. This strategy is capable to map battery capacity from multi-scale features, solve the problem that single scale features can reflect the complex trend of capacity degradation, and optimize the HI by using the historical data of each battery, thus improving the generalization ability.

FUNDING
This research was funded by the National Key Research Program of China, Grant number 2016YFC0300104, the National Natural Science Foundation of China (52011530033), the National Natural Science Foundation of China (51877057), and the Harbin Science and Technology Innovation Talent Project (2017RAQXJ069)

References

1. J. Yang, T. Wang, C.Y. Du, F.Q. Min, T.L. Lv, Y.X. Zhang, L.Q. Yan, J.Y. Xie and G.P. Yin, *Energy Storage Sci. Technol.*, 8 (2019) 58.

2. R. Li, W. Li, H. Zhang, Y. Zhou and W. Tian, *Front. Energy Res.*, 9 (2021) 693249.
3. T. Biagetti, E Sciubba, *Energy (Oxford, U. K.)*, 29 (2004) 2553.
4. J. Tian, H.B. Gao, Y.Q. Zhang, Y.T. Wang and D.Z. Hu, *Chin. J. Power Sources.*, 44 (2020) 767.
5. L.D. Couto, J. Schorsch, N. Job, *J. Energy Storage*, 21 (2019) 259.
6. C. Weng, X. Feng, J. Sun and H. Peng, *Appl. Energy*, 180 (2016) 360.
7. Z.C. Liu, X. Yang, W. Yu, B.X. Huang and Y. Wang, *Chin. J. Power Sources*, 43 (2019) 74.
8. Z. Ma, R.X. Yang and Z.P. Wang, *Appl. Energy*, 237 (2019) 836.
9. X.Y. Zhu, W. Li, *ISA Trans.*, 93 (2019) 14.
10. X. Zheng and H. Fang, *Reliab. Eng. Syst. Saf.*, 144 (2015) 74.
11. Y. Chen, M.H. Huang and S.K. Wang, *Autom. Instrum.*, 32 (2017) 69.
12. C.V. Lüders, J. Keil and M. Webersberger, *J. Power Sources*, 414 (2019) 41.
13. X. Lai, W. Gao, Y. Zheng, M.Y. Ou, J. Li, X. Han and L. Zhou, *Electrochim. Acta*, 295 (2019) 1057.
14. Y.F. Zhao, J. Xu, X. Wang and X.S. Mei, *Energy Procedia*, 145 (2018) 357.
15. K.X. Wei and Q.Y. Chen, *Proc. CSEE.*, 34 (2014) 445.
16. D.A. Souza, *J. Mech. Eng. Autom.*, 5 (2015) 17.
17. L.Y. Zhang, L. Zhang, P. Christos and S. Liu, *Wireless Pers. Commun.*, 102 (2018) 2063.
18. D. Wang, F. Yang, K.L. Tsui, Q. Zhou, B.S. Joo, *IEEE Trans. Instrum. Meas.*, 6 (2016) 1282.
19. M.L. Suo, J.F. Li, M. Zhao, M. Zhang, G.J. Shao and Z.B. Wang, *Chin. Battery Ind.*, 24 (2020) 255.
20. Y.F. Zhou, X.X. Sun, L.J. Huang and J.Lian, *J. Harbin Inst. Technol., New Ser.*, 53 (2021) 55.
21. H. Liu, M.X. Hu, Y.H. Zhu, *J. Nanjing Univ. Sci. Technol.*, 42 (2018) 329.
22. L.B. Li, Y.Z. Zhu, Y.J. Tian, Z.T. An and L.L. Wang, Chin. *J. Power Sources*, 43 (2019) 1027.
23. C.Y. Chen, D.W. Chen, *J. Power Sources*, 45 (2021) 589.
24. Y.Y. Jiang, Z. Liu, H. Luo and H. Wang, *J. Electr. Measur. Instr.*, 30 (2016) 179.
25. B. Saha, K. Goebel, S. Poll and J. Christophersen, *IEEE Trans. Instrum. Meas.*, 2 (2009) 291.

© 2022 The Authors. Published by ESG (www.electrochemsci.org). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).