

LIANG, Y., WANG, S., FAN, Y., HAO, X., LIU, D. and FERNANDEZ, C. 2024. State of health prediction of lithium-ion batteries using combined machine learning model based on nonlinear constraint optimization. *Journal of the Electrochemical Society* [online], 171(1), article number 010508. Available from: <https://doi.org/10.1149/1945-7111/ad18e1>

State of health prediction of lithium-ion batteries using combined machine learning model based on nonlinear constraint optimization.

LIANG, Y., WANG, S., FAN, Y., HAO, X., LIU, D. and FERNANDEZ, C.

2024

This is the Accepted Manuscript version of an article accepted for publication in Journal of the Electrochemical Society. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it. The Version of Record is available online at <https://doi.org/10.1149/1945-7111/ad18e1>.

**State of Health Prediction of Lithium-Ion Batteries Using
Combined Machine Learning Model Based on Nonlinear
Constraint Optimization**

Journal:	<i>Journal of The Electrochemical Society</i>
Manuscript ID	JES-110498.R1
Manuscript Type:	Research Paper
Date Submitted by the Author:	03-Nov-2023
Complete List of Authors:	Liang, Yawen; Southwest University of Science and Technology, School of Information Engineering Wang, Shunli; Southwest University of Science and Technology Fan, Yongcun; Southwest University of Science and Technology Hao, Xueyi; Southwest University of Science and Technology, Liu, Donglei; Southwest University of Science and Technology, Fernandez, Carlos; Robert Gordon University
Keywords:	Lithium-ion batteries, State-of-health, Long short-term memory network, Support vector regression, Nonlinear constraint optimization

SCHOLARONE™
Manuscripts

State of Health Prediction of Lithium-Ion Batteries Using Combined Machine Learning Model Based on Nonlinear Constraint Optimization

Yawen Liang,¹ Shunli Wang,^{1,2} Yongcun Fan,² Xueyi Hao,² Donglei Liu,² and Carlos Fernandez³

¹College of Electrical Engineering, Sichuan University, Chengdu, 610065, China

²School of Information Engineering, Southwest University of Science and Technology, Mianyang, 621010, China

³School of Pharmacy and Life Sciences, Robert Gordon University, Aberdeen, AB10-7GJ, United Kingdom

[✉]E-mail: wangshunli@swust.edu.cn

Abstract Accurate State of Health (SOH) estimation of battery systems is critical to vehicle operation safety. However, it's difficult to guarantee the performance of a single model due to the unstable quality of raw data obtained from lithium-ion battery aging and the complexity of operating conditions in actual vehicle operation. Therefore, this paper combines a long short-term memory (LSTM) network with strong temporality, and support vector regression (SVR) with nonlinear mapping and small sample learning. A novel LSTM-SVR combined model with strong input features, less computational burden and multiple advantage combinations is proposed for accurate and robust SOH estimation. The nonlinear constraint optimization is used to assign weights to individual models in terms of minimizing the sum of squared errors of the combined models, which can combine strengths while compensating for weaknesses. Furthermore, voltage, current and temperature change curves during the battery charging were analyzed, and indirect health features (IHF) with a strong correlation with capacity decline were extracted as model inputs using correlation analysis and principal component analysis (PCA). The NASA dataset was used for validation, and the results show that the LSTM-SVR combined model has good SOH estimation performance, with MAE and RMSE all less than 0.75% and 0.97%.

1 Introduction

With the increasing energy and environmental problems, vigorous development of EVs is an effective way to address the resource shortage and ecological issues and a powerful initiative to achieve ecological civilization [1]. The driving range and dynamic performance of EVs depend heavily on the capacity of the BMS [2]. Among them, lithium-ion batteries are gradually becoming one of the main power sources for EVs because of their excellent performance such as high energy density and long life [3]. BMS is developed to ensure the reliability and safety of lithium-ion batteries. SOH helps manage battery life and avoid potential battery failure, which is a crucial function of the BMS [4]. SOH is an indicator of the degradation of a battery compared to a new one. This indicator is not only valuable for the EVs management system to regulate itself to ensure that the overall performance and safety of the EVs are maintained within the desired range but also is essential for ecological sustainability [5]. At present, the recycling of secondary batteries has become a hot topic, providing stronger technical support for future ecological sustainability, and accurate SOH estimation is crucial for the effective utilization of secondary batteries [6]. However, during long-term use, electrodes and electrolytes change. The growth of the SEI of the cell anode, the loss of lithium plating and lithium, and the active material aging and electrode composition can reduce the usable capacity of a cell [7]. Various known/unknown factors interact to accelerate battery aging, which makes it extremely complex and difficult to elucidate battery aging from a theoretical level [8]. In addition, external parameters (e.g., temperature, discharge rate, SOC, etc.) can also contribute to battery aging [9]. Therefore, battery aging is uncontrollable, and accurate and robust SOH estimation for lithium-ion batteries remains an issue.

In recent years, a large body of literature has proposed different classifications of SOH, with the main approaches being experiment prediction approaches and model-based approaches [10]. Among them, experimental prediction techniques include both direct measurement methods and indirect evaluation methods. Direct measurement methods include capacity testing, ohmic internal resistance testing [11], EIS [12], etc. The capacity test is more suitable for laboratory capacity calibration. In practice, accurate detection of battery capacity is difficult because of the non-transitory capacity state of a fully charged battery [13]. Barai et al. [14] applied several battery internal resistance testing techniques to a 20 Ah LiFePO₄/C-6 pouch battery and compared the results. The results show that the calculated resistance is closely related to the time scale of the methods used. However, the relation between cell internal resistance and SOH is influenced by parameters such as discharge depth, SOC and temperature. In addition, the need for a multi-time scale estimator with an appropriate decoupling mechanism to eliminate cross-talk between SOC and SOH is a great challenge for ohmic internal resistance testing [15]. Vyroubal et al. [16] use EIS to obtain the parameters of the battery equivalent circuit. The proposed numerical simulations were compared with the actual

1
2 measurements, and the results obtained by EIS proved to be more satisfactory than the conventional parameter
3 identification methods. Although the EIS method provides vital information on battery life, there are fluctuations in
4 resistive impedance between different batteries. Therefore, EIS is considered an underutilized technology in practical
5 applications [17].
6
7
8
9

10 Indirect evaluation methods include charging curve techniques [18], ICA [19], DVA [20], etc. For example,
11 Wang et al. [21] investigated CV curves to obtain battery health indicators. The battery aging phenomena found in
12 the CV curves were described in detail by a first-order equivalent circuit model, and the CV charging aging factor is
13 a current-dependent parameter that reflects the SOH of batteries. Stroe et al. [22] related the battery capacity decay
14 to the evolution of the four metric points obtained by ICA based on the collected battery calendar aging results. The
15 applicability of the ICA in predicting the capacity decay and SOH of LMO/nmc-based cells was investigated. Zhang
16 et al. [23] proposed a model-free SOH estimation approach incorporating the Coulomb counting method and DVA to
17 achieve fast online SOH calculation in the constant current discharge stage. However, there are two obvious
18 limitations of ICA and DVA. First, the ICA and DVA curves correspond to small current rates of C/25, which are
19 difficult to obtain experimentally and result in high current rates, thus reducing the accuracy and robustness of SOH
20 estimation [24]. Second, ICA and DV curves require more efficient filtering and smoothing techniques to remove
21 noise interference to obtain smoother curves, which further increases the complexity of the analysis [25].
22
23
24
25
26
27
28
29
30
31
32
33

34 While experimental prediction techniques can provide comprehensive battery aging data and accurate SOH
35 estimation, the BMS needs reliable, real-time and accurate SOH estimation [26]. Adaptive filtering approaches and
36 data-driven approaches can accurately estimate battery capacity, resistance and other features, which are closely
37 related to battery capacity decline [27]. Adaptive filtering approaches include electrochemical model-based
38 approaches (e.g., P2D model [28], SPM [29]) and equivalent circuit model-based approaches (Thevenin model [27],
39 FOM [30]). For example, Liu et al. [31] proposed a P2D model-based method for joint SOC and SOH estimation.
40 First, they simplified and reformulated the standard P2D model into a nonlinear state space form. Second, a PF
41 algorithm is used to estimate the average lithium-ion concentration, which is used for SOC estimation, and the
42 average lithium-ion concentration at the predicted cut-off voltage is used for the SOH calibration. However, existing
43 electrochemical model-based SOH estimation approaches provide limited understanding of the electrochemical
44 structure of lithium-ion batteries, which does not fully reflect the actual situation of the lithium-ion batteries, and the
45 precision of the state estimation is questionable [32]. Ling et al. [33] propose a joint SOC and SOH estimation
46 approach for lithium-ion batteries based on DEKF and FOM. The convergence and robustness of this approach are
47 discussed under three operating conditions, and the method is validated using erroneous initial state values and noise
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2 analysis. However, the equivalent circuit model-based approach estimates system parameters and battery state
3 simultaneously, which causes periodic changes in system variables and makes the prediction algorithm unstable [34].
4
5 These results in increased computation time and reduced response speed, which eventually affect the accuracy of
6
7 SOH estimation [35].
8
9

10 In recent years, data-driven approaches have been widely used for SOH estimation on the premise that battery
11 aging is a stable stochastic process [36]. For example, optimization algorithms (e.g., GA [37], BA), EM [38], sample
12 entropy [39], ML techniques (e.g., SVM [40], FNN [41], RNN [42], and RBF [43]). For example, Chang et al. [44]
13 proposed an approach for online SOH estimation under the current discharge based on IC and fusion of GA-WNN.
14 The experiments present that the SOH estimation error is less than 3%. Although the optimization algorithm is highly
15 accurate, it takes a long time to derive the optimal solution matrix, which is hard to apply in real life. EM was obtained
16 by summarizing a large amount of experimental data and the relationship between each relevant parameter (e.g.,
17 temperature, discharge rate, cycle counts, etc.) [45]. Jiang et al. [46] proposed a fusion estimation approach based on
18 an empirical degradation model and data-driven. Among them, four operating characteristics from the battery charging
19 and discharging phase were extracted as model input, and the prediction error of the empirical degradation model
20 with charging and discharging cycles as output. The experimental results present that the average absolute percentage
21 error of the proposed SOH prediction method is about 2%, which has good estimation performance and applicability.
22 Compared with other models, the EM has a simple structure and can be used flexibly in practical applications.
23 However, the EM is heavily influenced by the quantity and quality of experimental data and SOC estimation. In
24 addition, battery aging is subject to the cumulative effects of various known/unknown factors, and it's hard to obtain
25 an accurate SOH estimation considering all parameters [47]. The ML techniques focus on finding the connection
26 between the input excitations and the target response, rather than the battery decline mechanism [48]. Therefore, ML
27 techniques can construct a functional relation between the SOH and the input excitations without exploring the
28 internal mechanism of the battery in depth. Khaleghi et al. [49] establish a NARX model to obtain the dependence
29 between battery HFs and SOH. The experiments show that the maximum root mean square error of NARX for
30 lithium-ion battery health diagnosis is 0.46 for untrained data, which has good accuracy and robustness. Tian et al.
31 [50] propose a migration learning-based SOH prediction method. Among them, a basic model combining LSTM and
32 a fully connected layer was designed for SOH estimation, and a feature expression scoring rule was proposed to
33 assess the relevance of multiple prediction tasks. However, the high quality and quantity of data required by ML
34 techniques, as well as the high cost of computing facilities, remain an issue that needs to be addressed [51]. It is worth
35 noting that the rapidly developing computer technology and the cloud computing technologies allow data to be
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2 assigned to computers for processing to perform fast calculations, which shows great promise for the near future.

3
4 From the above analysis, ML techniques have been a popular choice for SOH estimation of lithium-ion batteries
5
6 in recent years mainly due to the following advantages compared to model-based methods.

7
8 (1) ML techniques do not need to simulate the electrochemical reactions, internal mechanisms and battery
9
10 construction inside the battery, which uses the HFs associated with battery aging as model input to estimate SOH. In
11
12 particular, ML techniques map the nonlinearity of autonomous learning features better than model-based approaches.

13
14 (2) Most existing SOH estimation methods require establish the relationship between SOH and influencing
15
16 factors. However, it is extremely difficult to find a functional relationship that includes multiple or even all
17
18 influencing factors to describe SOH due to many factors that affect the accuracy of SOH. ML models are black-box
19
20 models that autonomously learn the relationship between input features and SOH. In addition, a more comprehensive
21
22 consideration of highly correlated HFs can help achieve more accurate SOH estimation.

23
24 (3) Simple SOH estimation methods are good in terms of computational efficiency and accuracy. However, with
25
26 the strong demand for high-tech technologies for rapid social progress, high accurate and robust SOH estimation
27
28 requires more complex models for implementation. That said, highly accurate and efficient ML methods are an
29
30 obvious choice.

31
32 ML techniques require substantial data related to battery life decline to ensure effective model training and
33
34 accurate SOH estimation and have certain requirements for computing equipment [52]. However, due to the growth
35
36 of computer technology and cloud computing technology, the rapid development of intelligent cloud computing
37
38 technology based on actual battery data obtained from big data platforms, cloud-based machine learning methods
39
40 and other technologies can enable fast calculation and processing of data, as well as effective detection of large
41
42 numbers of lithium-ion batteries in real-time, etc. These show great promise for ML techniques to shine for science
43
44 and technology in the future.

45
46 However, for now, there are still the following difficulties in the SOH estimation of lithium-ion batteries using
47
48 ML techniques in practical applications.

49
50 (1) Most existing ML techniques attempt to build a unified, robust and highly accurate strong ML model. On
51
52 the one hand, strong ML models are usually difficult to build strong estimators in practical applications. Because the
53
54 performance of ML models is very sensitive to the training features, imperfect sensors and poorly correlated training
55
56 samples can affect the performance of the estimator. On the other hand, data processing techniques and computer
57
58 equipment are limitations now. Considering the limitations of industrial cost, data collection and computational
59
60 efficiency, strong estimators are so demanding that they are difficult to implement for practical applications to an

1
2 extent.

3
4 (2) A single estimator is only applicable for fixed-length inputs or specific operating conditions. However,
5 vehicle driving conditions and charging processes are dynamic, and the environmental variables at different trip
6 stages are usually more complex. To achieve accurate SOH estimation for lithium-ion batteries with different length
7 inputs and different operating conditions, a more inclusive and balanced optimization method to increase the SOH
8 estimation accuracy is necessary. This method enables an efficient combination of different excellent algorithms or
9 models to adapt to the dynamic changes in vehicle operation.
10
11

12
13 (3) Input features that have little correlation with the output variables, and too few or too complex input features
14 affect the overall performance of the ML models, resulting in inaccurate SOH estimation. In addition, to achieve
15 efficient SOH online estimation, the computational load requires reduced while maintaining good accuracy. These
16 require that the selected input features are highly relevant and less numerous.
17
18

19
20 To obtain a good SOH estimation effect in practical applications, the two vital directions of feature selection
21 and model optimization must receive more attention [53]. In this paper, a combined ML model with strong input
22 features, small computational load and multi-dominance combination is proposed for accurate and robust SOH
23 estimation of lithium-ion batteries. The proposed SOH estimation method improves the accuracy of SOH estimation
24 under different input feature lengths and different operating conditions by combining ML models with different
25 characteristics. The main contributions are summarized below:
26
27

28
29 (1) The current, voltage and temperature change curves during the battery charging were analyzed and fourteen
30 HFs related to capacity decline were extracted. Next, HFs with high correlation with cell capacity were selected using
31 Pearson and Spearman correlation analysis to characterize the cell capacity decline, and the selected HFs were
32 dimensionally reduced using PCA based on eigenvalue decomposition to obtain IHFs as the input feature of the
33 model.
34

35
36 (2) LSTM network specialized in processing time series data is chosen to store and update vital information of
37 degraded data through effective long-term dependency learning in highly correlated cycles. Since the LSTM network
38 has the disadvantages of being random and not suitable for small sample learning, SVR with nonlinear mapping and
39 small sample learning is used to compensate for these shortcomings. In addition, the computational complexity of
40 SVR is independent of the input dimensionality and thus has high accuracy and good generalization ability. To
41 combine the advantages of both models, a novel LSTM-SVR combined model is proposed for accurate and robust
42 SOH estimation.
43
44

45
46 (3) A nonlinear constraint optimization is used to assign weights to individual models in the LSTM-SVR
47
48
49
50
51
52
53
54
55
56
57
58
59
60

combined model in terms of minimizing the sum of squared errors of the combined model. An effective combination of different models can combine advantages while compensating for disadvantages, thus improving the accuracy and robustness of SOH estimation for different input feature lengths and different operating conditions.

The LSTM-SVR combined model was validated using different batteries, test sets of different proportions and different operating conditions. The results present that the LSTM-SVR combined model has good accuracy and robustness in SOH estimation. In the validation of accuracy and robustness, the MAE and RMSE of the LSTM-SVR combined model are all less than 0.465% and 0.517%, respectively. In the validation of the adaptability for different working conditions, the MAE and RMSE of the LSTM-SVR combined model are all less than 0.406% and 0.815%, respectively.

The remainder of this paper is organized as follows: Chapter 2 presents the selection and optimization of HFs; Chapter 3 presents the methodology of SOH estimation and the development of the combined model; Chapter 4 discusses the experimental results of SOH estimation in detail, as well as summarizes and outlooks the works; the summary is presented in Chapter 5.

2 Health features selection and optimization

2.1 Health Features Extraction

For existing SOH estimation approaches, especially ML techniques, ample and available data is a big problem for SOH estimation due to the time-consuming battery aging experiments [54]. Because battery aging experiments take a long time, ranging from a few months to several years. One solution to this problem is the availability of the public dataset. The discrepancy between an ideal laboratory environment and complex experimental equipment or realistic conditions makes some experiments and procedures challenging to achieve in reality. However, experimental procedures and data can be used for the study of the mechanisms of battery life decline and bring a theoretical basis for ML techniques. Therefore, the NASA battery dataset is selected in this paper. Tab. 1 shows the basic information of the six chosen cells.

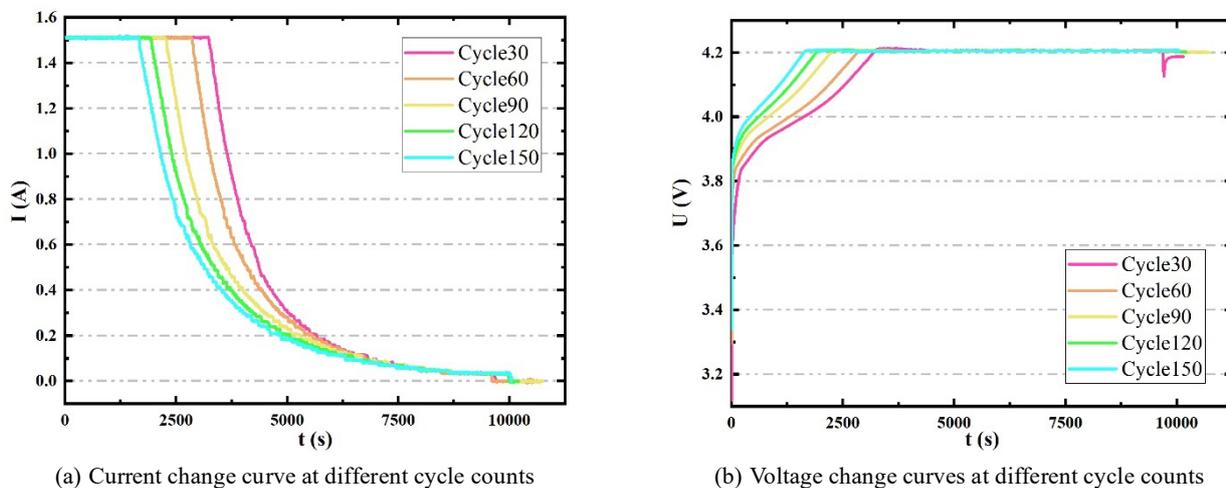
Tab. 1. Parameters of the selected battery

Cell number	Cycle counts	Discharging current /A	Cut-off voltage /V	Ambient temperature /°C	Nominal capacity /Ah	EOL /Ah
B0005	168	2	2.7	24	2	1.40
B0006	168	2	2.5	24	2	1.40
B0030	39	4	2.2	43	2	1.61
B0032	39	4	2.7	43	2	1.65
B0045	71	1	2	4	2	1.4
B0047	71	1	2.5	4	2	1.4

The fundamental reason for this selection is that much of the existing literature selects only B0005 to B0007

and B0018 cells for evaluation method validation. To further evaluate the accuracy and robustness of the proposed SOH estimation method, we selected imperfect states B0030, B0032, B0045 and B0047 cells as auxiliary datasets for model training and testing. Two cells with different cut-off voltages in the same operating condition were chosen to test the generality of the proposed SOH estimation method under different influencing factors.

Some existing ML techniques utilize current, voltage, temperature, and SOC as input features. Although these input features correlate with battery capacity, they are too general, which leads to the neglect of some influences that characterize the aging properties of batteries well. Much literature has demonstrated that capacity decline curves, battery charging and discharging curves, OCV curves and other characterizes that vary widely with battery life can characterize the failure status of lithium-ion batteries [55]. Compared with the input features mentioned above, these curves contain more features that can further reflect the battery aging mechanism, which allows for accurate and robust SOH estimation. The cell charging phase is stable compared to the discharging phase, which facilitates stable analysis and extraction of HFs. The battery charging phase includes CC charging and CV charging. In practice, it is more common to initially charge with CC and then switch to CV, which has a pronounced effect on battery aging. Therefore, in this paper, the voltage, current and temperature change curves during the cell charging phase are analyzed. Taking the B0005 cell as an example, the voltage, current and temperature change curves as well as the capacity decline curve of the cell charging phase are shown in Fig. 1.



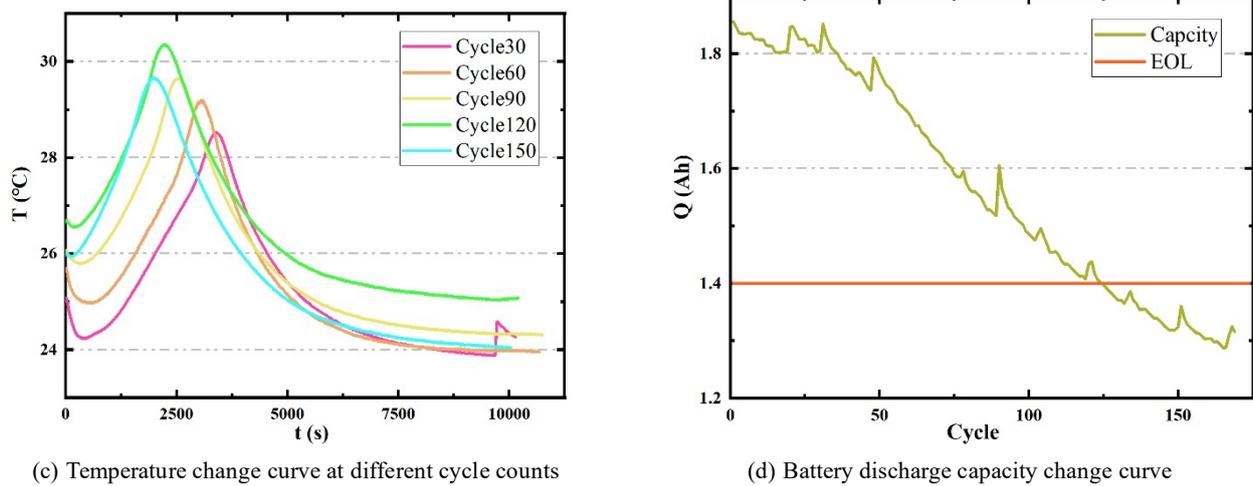
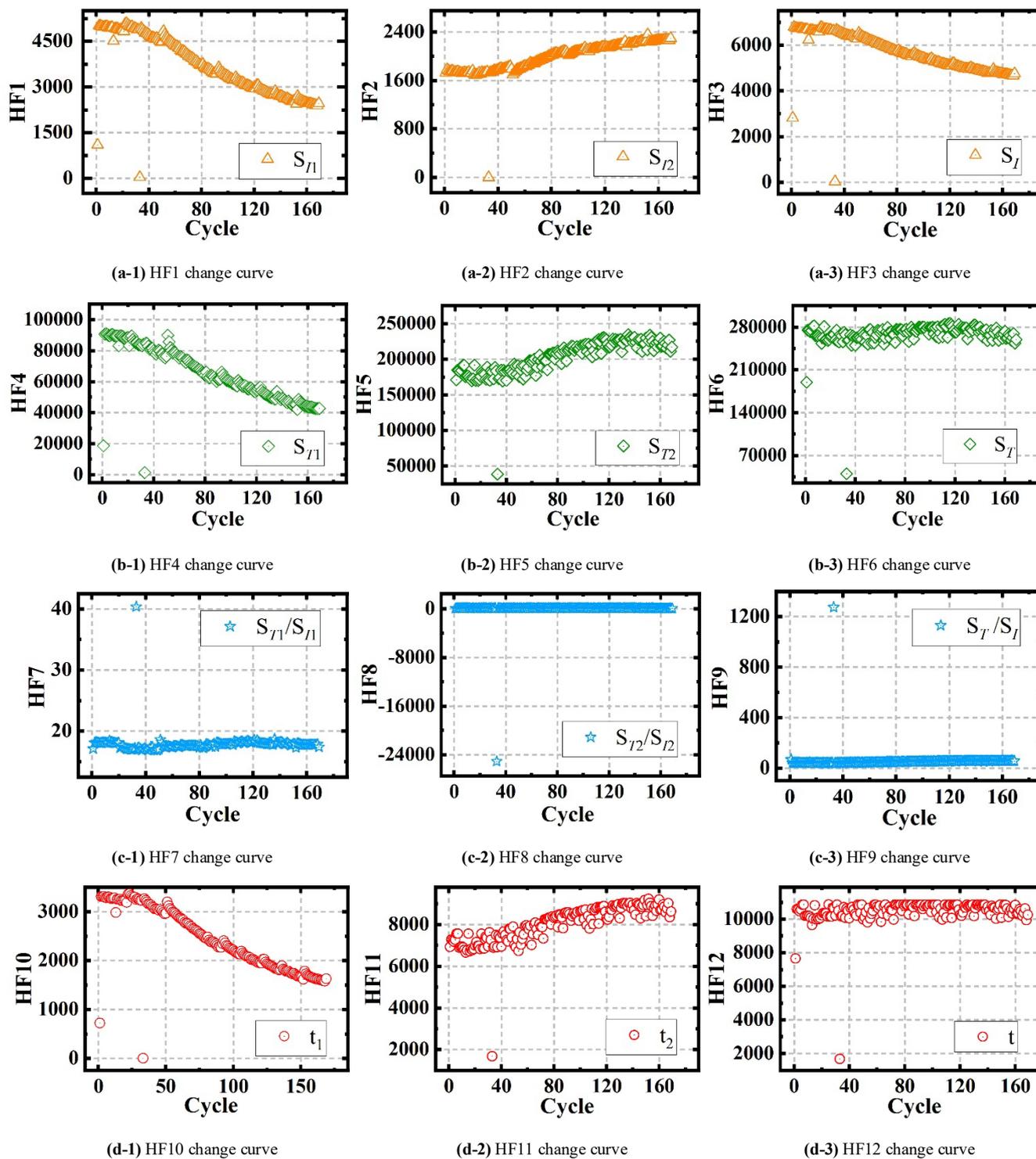


Fig. 1. Lithium-ion battery charging characteristics curve

From Fig. 1 (a) and (b), the current, voltage and temperature change curves during the cell charging phase gradually shift to the left as the cycle counts increase, the CC charging time becomes shorter and the CV charging time becomes longer. In addition, the area between the current change curve and the x-axis during the charging phase increases with the cycle counts. Furthermore, the voltage rise varies with the cycle counts during the CC charging phase. Therefore, the time and the area of the current change curve during the charging phase, as well as the extent of the voltage change between two time points during the CC charging phase can be considered as HFs to characterize the battery capacity decline. From Fig. 1 (c), the battery temperature rises and peaks during the CC charging phase, and then it gradually decreases during the CV charging phase. The area between the temperature change curve and the x-axis increases with the cycle counts during the charging phase, which can be used as HFs to assess the battery aging. Fig. 1 (d) shows the capacity decline curve of the cell. The capacity regeneration phenomena during battery aging pose a challenge for accurate SOH estimation. Therefore, it is critical to take multiple factors affecting battery aging into account and to select appropriate HFs.

Through the above analysis, this paper extracts fourteen HFs that can indirectly reflect the battery aging rules, which are divided into five categories: (1) Area of the current curve during different charging: S_{I1} is the area of the current curve during CC charging, S_{I2} is the area of the current curve during CV charging, S_I is the area of the current curve during CCCV charging; (2) Area of the temperature curve during different charging: S_{T1} is the area of the temperature curve during CC charging, S_{T2} is the area of the temperature curve during CV charging, S_T is the area of the temperature curve during CCCV charging; (3) Ratio of the temperature curve area and current curve area during different charging: S_{T1}/S_{I1} is the ratio of the temperature curve area and current curve area during CC charging, S_{T2}/S_{I2} is the ratio of the temperature curve area and current curve area during CV charging, S_T/S_I is the ratio of temperature curve area and current curve area during CCCV charging; (4) Time of different charging phase and their ratio: t_1 is the CC charging time, t_2 is the CV charging time, t is the CCCV charging time, t_1/t_2

is the ratio of CC charging time to CV charging time; (5) U_{jump} is the extent of the voltage change between two points in time. The relation between the selected HF's and cycle counts is presented in Fig. 2.



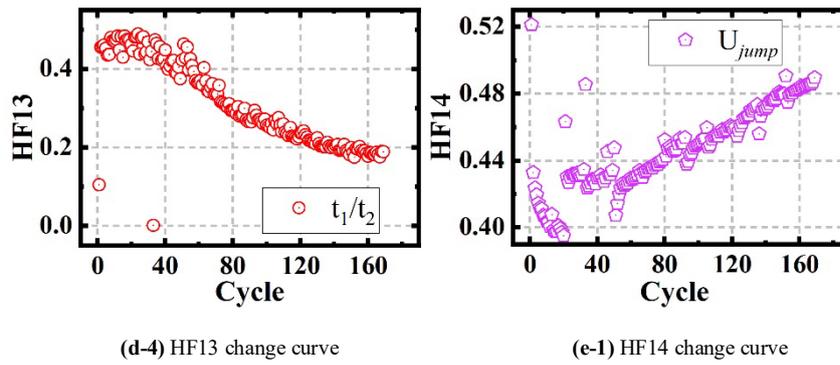


Fig. 2. Relation between selected HFs and cycle counts

From Fig. 2, the selected HFs show regular variations with the increase of the cycle counts, which can characterize battery aging. However, on the one hand, these HFs vary in trend, which makes it difficult to visually determine their correlation with battery capacity decline. On the other hand, these HFs directly used as input features for ML models would cause a large computational load and complex data management, which is not suitable for practical applications. Therefore, the correlation between these HFs and battery capacity requires be further analyzed to select good HFs as the feature input for the machine learning model.

2.2 Health features optimization

Pearson and Spearman correlation coefficients are used to evaluate the correlation between two factors [56]. The larger the absolute value of the two factors, the higher their correlation. It is more reasonable to use the Pearson correlation coefficient when the factors are continuous data, obey a normal distribution, and have a linear relationship. If the above conditions are not fully satisfied, the Spearman correlation coefficient is used to better reflect the correlation between the two factors. In this paper, two correlation coefficients are used to further select more appropriate HFs. The formulas for the Pearson and Spearman correlation coefficients are shown as follows:

$$Pearson = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}} \quad (1)$$

$$Spearman = \frac{\sum_i (x_i - \bar{x})(x_i - \bar{x})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \quad (2)$$

where X and Y represent the total samples of HFs and capacity, respectively; x_i and y_i represent the samples, respectively.

The Pearson and Spearman correlation coefficients of the fourteen HFs with capacity are shown in Tab. 2 and Tab. 3, respectively.

Tab. 2. Pearson correlation coefficient of selected HFs with capacity

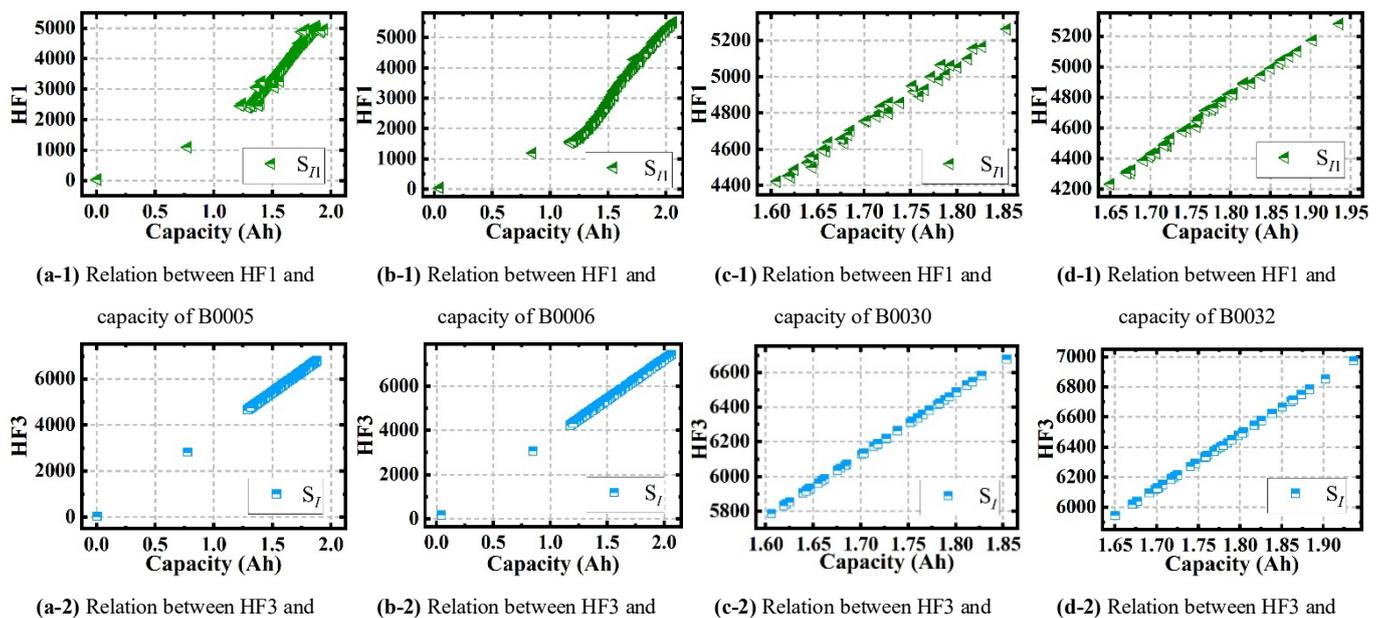
Battery	S_{I1}	S_{I2}	S_I	S_{T1}	S_{T2}	S_T	S_{T1}/S_{I1}	S_{T2}/S_{I2}	S_T/S_I	t_1	t_2	t	t_1/t_2	U_{jump}

B0005	0.9684	-0.3123	1.0000	0.9659	-0.3233	0.4121	-0.5639	0.5099	-0.5670	0.9689	-0.3125	0.4376	0.9316	-0.8531
B0006	0.9684	-0.5305	1.0000	0.9662	-0.5116	0.3384	-0.3960	-0.3195	-0.7876	0.9688	-0.5148	0.3201	0.9448	-0.8881
B0030	0.9950	0.4941	1.0000	0.9238	0.2161	0.8582	-0.1034	-0.2874	-0.8890	0.9947	0.2227	0.8240	0.8840	-0.1503
B0032	0.9983	-0.6310	1.0000	0.9598	0.0340	0.9172	-0.1690	0.4936	-0.9212	0.9974	0.0547	0.8694	0.9355	0.2929
B0045	0.1155	0.9985	1.0000	-0.2721	0.1846	0.1507	-0.5028	-0.9578	-0.9628	0.7983	-0.7906	0.1940	0.7970	-0.8905
B0047	0.6071	0.8152	1.0000	0.5408	-0.2206	0.2417	0.0026	-0.8585	-0.9454	0.6471	-0.7956	-0.5030	0.6682	0.1754

Tab. 3. Spearman correlation coefficient of selected HF's with capacity

Battery	S_{I1}	S_{I2}	S_I	S_{T1}	S_{T2}	S_T	S_{T1}/S_{I1}	S_{T2}/S_{I2}	S_T/S_I	t_1	t_2	t	t_1/t_2	U_{jump}
B0005	0.9993	-0.9160	1.0000	0.9927	-0.8067	-0.1873	-0.2906	0.2979	-0.9532	0.9993	-0.8275	-0.1440	0.9852	-0.9412
B0006	0.9993	-0.8858	1.0000	0.9975	-0.8252	-0.1113	-0.0455	0.5330	-0.9824	0.9993	-0.8528	-0.1795	0.9951	-0.9733
B0030	0.9925	0.4838	1.0000	0.9249	0.1518	0.8830	-0.6445	-0.3350	-0.9148	0.9923	0.1502	0.8178	0.9008	-0.1377
B0032	0.9992	-0.6921	1.0000	0.9589	-0.1099	0.9049	-0.6326	0.5079	-0.9492	0.9976	-0.0771	0.8245	0.9543	0.4972
B0045	-0.3328	0.9992	1.0000	-0.5242	0.1153	0.0831	-0.5561	-0.9829	-0.9831	0.8439	-0.8207	0.0394	0.8439	-0.9488
B0047	0.5454	-0.8079	1.0000	0.4541	-0.3475	0.1687	-0.2622	-0.8481	-0.9297	0.5498	0.7100	-0.4522	0.5801	-0.0762

To ensure that the selected HF's can well characterize the battery capacity decline and that the ML model has less computational load, four to five HF's with high correlation with battery capacity are selected from the two tables as input features of the model. By analyzing Tab. 2 and Tab. 3, for B0005, B0006, B0030 and B0032 cells, $(S_{I1}, S_I, S_{T1}, t_1, t_1/t_2)$ is selected as the input feature of the model; for B0045 and B0047 cells, $(S_{I2}, S_I, S_{T2}/S_{I2}, S_T/S_I)$ is selected as the input features of the model. The relation between the selected HF's and the battery capacity is shown in Fig. 3 and Fig. 4.



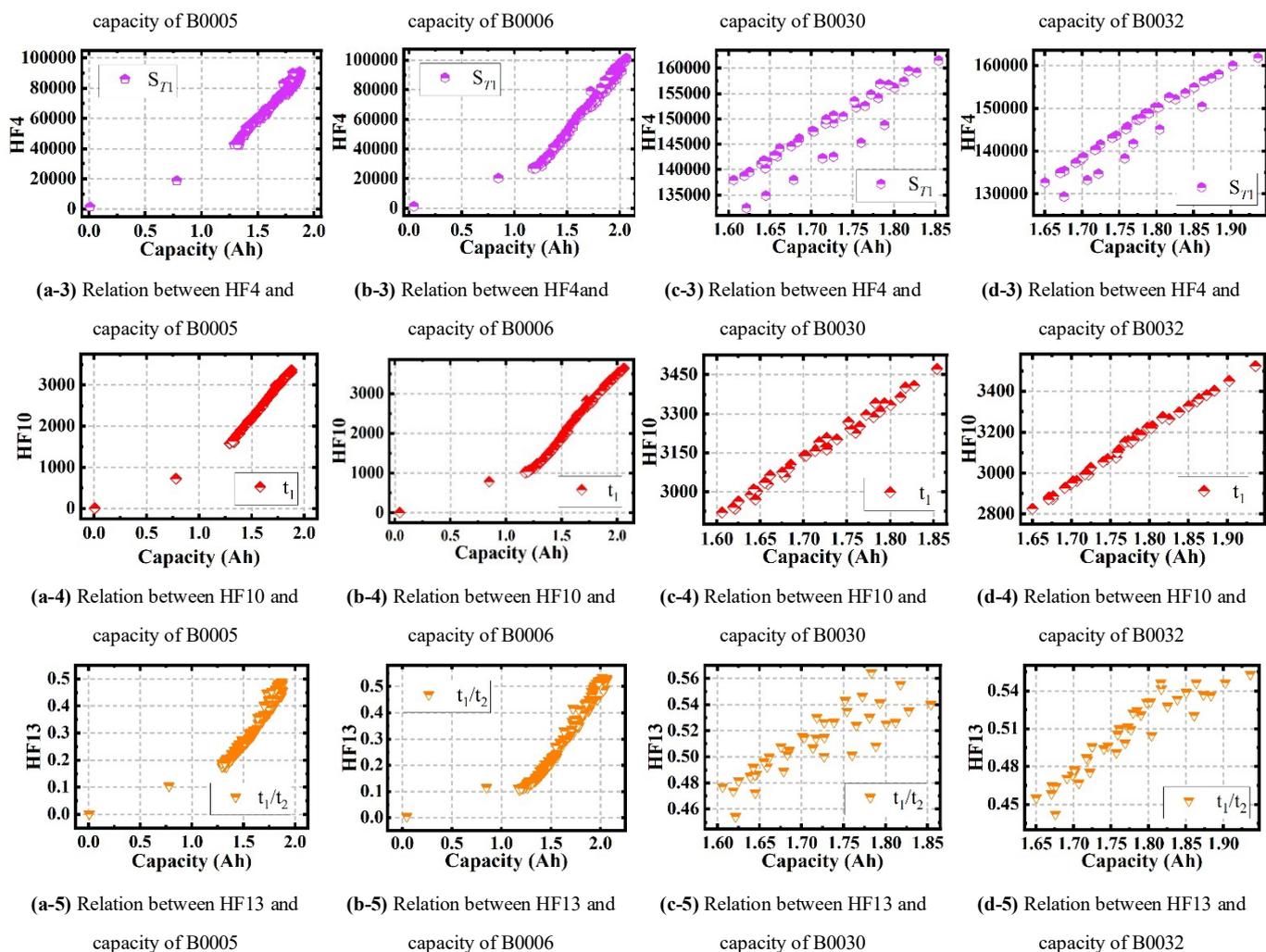


Fig. 3. Relation between the selected HF4 and the capacity of the first four cells

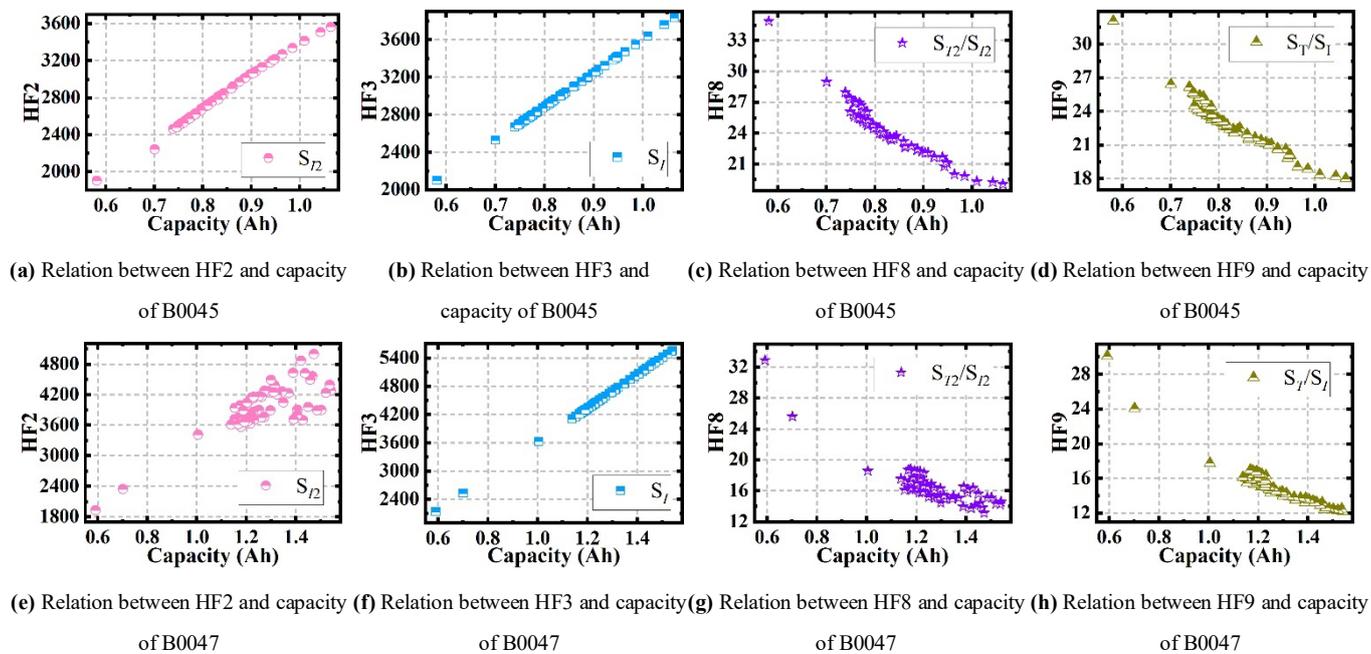


Fig. 4. Relation between the selected HF4 and the capacity of the last two cells

Fig. 3 and Fig. 4 further demonstrate that the fourteen HF4s selected are highly correlated with capacity. However, there is inevitably overlapping information among these HF4s, which can impose a large unnecessary computational

burden on the model prediction. To address this issue, PCA based on eigenvalue decomposition is used to further optimize the selected HFs, which allows overlapping and redundant information in the HFs to be separated to reduce the computational burden of SOH estimation. PCA can maintain the original information of HFs while recombining their duplicate and redundant information into fewer mutually unrelated variables. The specific calculation of PCA is as follows.

First, the $n \times m$ matrix X is composed of n battery samples and m HFs. Centering all sample $x_i, i = 1, 2, \dots, m$ in matrix X , as shown follows:

$$x_i = x_i - \frac{1}{m} \sum_{j=1}^m x_j \quad (3)$$

Thus, the normalized matrix \bar{X} is obtained and its covariance matrix S is calculated as follows:

$$S = \frac{1}{m} \bar{X}^T \bar{X} \quad (4)$$

where S is an $n \times n$ matrix. Then the eigenvalues decomposition is performed, and the eigenvectors u_i and eigenvalues λ_i of S are obtained by the following equation:

$$S u_i = \lambda_i u_i \quad (5)$$

where u_i is the $n \times 1$ vector and λ_i is the scalar.

The eigenvectors are composed of a matrix from top to bottom according to the size of the corresponding eigenvalue, and the first k rows are used to compose matrix U . k is determined by the ratio of the information included in the dimensionality-reduced data to the information in the original data, i.e., the cumulative contribution of the feature values. The matrix \bar{X} is reduced to matrix Z as follows:

$$Z = \bar{X} U \quad (6)$$

where $U = [u_1, u_2, \dots, u_k]$. The contribution of each principal component can be obtained:

$$r_i = \frac{\lambda_i}{\sum_{i=1}^k \lambda_i} \times 100\% \quad (7)$$

HFs after dimensionally reduced by PCA are defined as IHFs, which are used as input features for ML models to estimate SOH.

3 SOH estimation methodology development

3.1 Framework of the proposed methodology

The SOH is defined as the ratio of the current maximum available capacity of a battery to the nominal capacity

of a battery, as shown as follows:

$$SOH = \frac{Q_{max}}{Q_{nominal}} \times 100\% \quad (8)$$

where Q_{max} is the current maximum available capacity of a battery and $Q_{nominal}$ is the nominal capacity of a new battery.

The lithium-ion batteries aging can cover hundreds or even more cycles. The main difficulty in estimating the SOH of lithium-ion batteries is to store and update key information in the aging data over highly correlated cycles through effective long-term dependency learning. To solve this problem, an LSTM network that can capture long-term sequence dependencies better than an RNN network is an ideal choice. Although the LSTM network is highly time-series and can fully exploit the rules of historical information, there is still a vital issue to consider. The LSTM network is random and heavily dependent on large samples. However, in practical applications, large samples increase the data processing and computational burden of the ML models, which is not conducive to accurate SOH online estimation using fewer input features. Although the SVR model has poor timing, it has the advantages of nonlinear mapping and small-sample learning. In addition, the computational complexity of the SVR model is independent of the input dimension, and SVR is more stable compared to the LSTM network with randomness. The SVR model has high accuracy and good generalization capability for SOH estimation.

In practical applications, the environmental variables in different EVs trips are usually complex. To adapt the model to dynamic changes of the actual vehicle operation, different input feature lengths and different operating conditions, as well as to reduce the computational burden while improving the accuracy of the SOH estimation. A novel LSTM-SVR combined model is proposed in this paper, and a nonlinear constraint optimization is used to determine the weights of individual models by minimizing the sum of error squares of the combined model. The design idea of the LSTM-SVR combined model is to combine the advantages of the standard LSTM network and the standard SVR model while compensating for their respective disadvantages. In addition, The LSTM-SVR combined model can improve the accuracy and robustness of SOH estimation under different input feature lengths and different operating conditions.

3.2 LSTM-SVR combined model

The LSTM-SVR combined model is a dynamic combination of the LSTM network and SVR model. A nonlinear constraint optimization is used to minimize the sum of error squares of the combined model to assign weights to individual models. These weights are assigned differently for different working conditions and different input characteristics. The dominant model takes more weight to adapt to obtain accurate and robust SOH estimates. The

construction of the LSTM network, SVR model and LSTM-SVR combined models are described below, respectively.

3.2.1 Introduction of the LSTM network

The framework of the LSTM network is present in Fig. 5. x_k and h_k denote the input features of the HF sequence of lithium-ion batteries at moment k and the output information of SOH estimation, respectively. C_k is the state information of the memory unit at moment k , which stores the state information at the current moment and passes it to the next moment using the state information of the previous moment as well as the current input information. Thus, the historical information of long-series data training is preserved and delivered. σ is a logistic s-type function. The input gate, forget gate and output gate are denoted by i , f and o respectively.

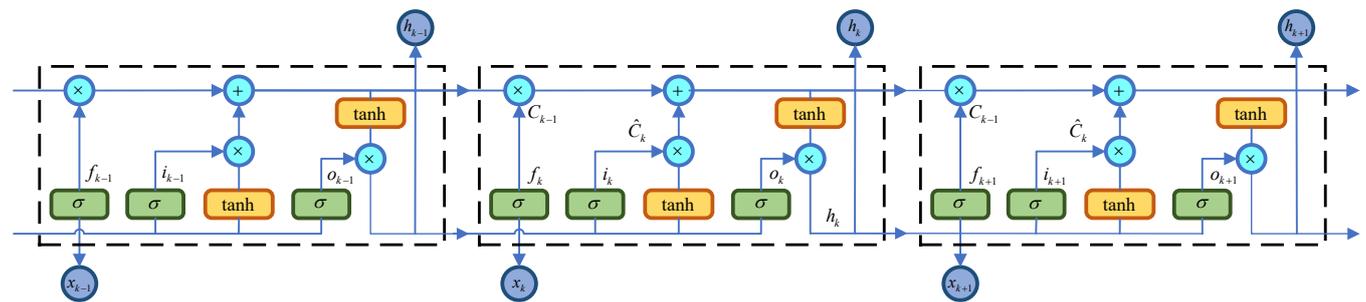


Fig. 5. Structure of the LSTM network

The computation of the state message C_k and the output message h_k relies on three logic gates. The forget gate f_k is used to control the proportion of state information at the previous moment to the current state information. If the forget gate f_k converges to 1, the current LSTM network retains the historical information completely, Conversely, if the forget gate f_k tends to 0, the current LSTM network does not rely on historical information. Input gate i_k and output gate o_k are used to control the update ratio of status information and the input ratio of status information, respectively. The three logic gates are calculated as:

$$i_k = \sigma(W_i[h_{k-1}, x_k] + b_i) \quad (9)$$

$$f_k = \sigma(W_f[h_{k-1}, x_k] + b_f) \quad (10)$$

$$o_k = \sigma(W_o[h_{k-1}, x_k] + b_o) \quad (11)$$

where W_i , W_f and W_o are the weights of the input gate, forget gate and output gate, respectively; b_i , b_f and b_o are the bias of the input gate, forget gate and output gate, respectively. The candidate quantity \hat{C}_k at the current moment is calculated to update the status information C_k .

$$\hat{C}_k = \tanh(W_C[h_{k-1}, x_k] + b_C) \quad (12)$$

$$C_k = f_k C_{k-1} + i_k \hat{C}_k \quad (13)$$

where W_C and b_C are the input weights and biases of C_k , respectively. The output information h_k of the network at the current moment is obtained by the output gate.

$$h_k = o_k \tanh(C_k) \quad (14)$$

In this paper, the LSTM network consists of one output layer, two hidden layers, and one output layer. Three logic gates allow the LSTM network to forget or transfer new information to the memory cell under training. To prevent overfitting of the LSTM network, both hidden layers use the Dropout regularization algorithm, which randomly discards neurons from the model under training. The optimized model can reduce the sensitivity to specific neuron weights and converge to better regularization. In addition, an Adam optimizer was chosen for the LSTM network configuration process to determine the compilation effect. The Adam optimizer is adept at handling non-smooth targets and requires less processing memory. The LSTM network initialization hyperparameters are {Epoch = 300, Dropout ratio = 0.2, Hidden Layer = 2, Initial Learn Rate = 0.0003}. The learning rate determines the adjustment level of the model in each parameter update step. Since a fixed learning rate affects the estimation accuracy of the LSTM network. In this paper, the learning rate schedule based on the exponential method is used for learning rate annealing. The learning rate schedule is calculated as follows.

$$\eta_n = \eta_0 e^{-dn} \quad (15)$$

where η_0 is the initial learning rate, d is the decay hyperparameter, and η_n is the learning rate at the n th step.

3.2.2 Introduction of the SVR framework

The SVR model is designed to address the regression issue for nonlinear separable data. Since the SVR model uses an error tolerance ε , the fitting error is subject to a cost function penalty when it is outside of some defined error range, which facilitates stable SOH estimation. The SVR model framework is presented in Fig. 6.

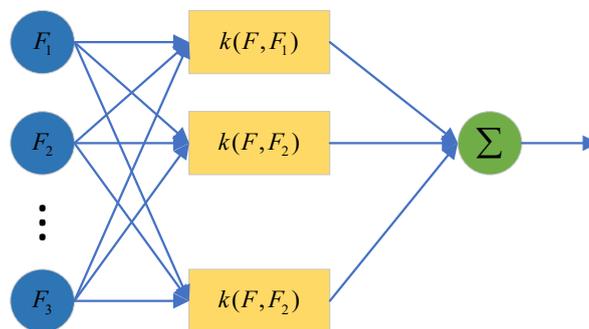


Fig. 6. Structure of the SVR model

From Fig. 6, the first layer is the input features, and the second layer is SVs, which take advantage of the kernel function to map the input features to a high-dimensional space to form linear solutions. SVR is derived as follows:

First, the model function corresponding to the large interval division hyperplane is given

$$f(x) = w^T x + b \quad (16)$$

where w is a vector of weights and b is a constant.

Given a training sample $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, $y_i \in R$. For sample (x, y) , traditional regression models usually calculate the loss based on model outputs $f(x)$ and y being as close as possible, and their loss is only zero when $f(x)$ and y are equal. However, the SVR assumes a deviation of at most ϵ between $f(x)$ and y , and the loss is only calculated when the deviation is greater than ϵ . That is, an interval band of width 2ϵ is constructed with $f(x)$ as the center. The prediction results are correct when the training samples fall within the interval band of width 2ϵ . Thus, the regression problem can be transformed into

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \ell_\delta(f(x_i) - y_i) \quad (17)$$

where C is the regularization factor, which determines the balance between the loss function and the regularization part. ℓ_ϵ is a ϵ -insensitive loss function, which can represent the decision function in terms of sparse points. ℓ_ϵ is calculated as follows

$$\ell_\delta(z) = \begin{cases} 0, & \text{if } |z| \leq \delta \\ |z| - \delta, & \text{otherwise} \end{cases} \quad (18)$$

$$\begin{aligned} \text{s.t. } & f(x_i) - y_i \leq \delta + \xi_i \\ & y_i - f(x_i) \leq \delta + \hat{\xi}_i \\ & \xi_i \geq 0, \hat{\xi}_i \geq 0, i = 1, 2, \dots, m \end{aligned}$$

where $z = y - f(x)$, ξ is the slack variable.

Introducing the Lagrangian multiplier $\mu_i \geq 0, \hat{\mu}_i \geq 0, \alpha_i \geq 0, \hat{\alpha}_i \geq 0$, the Lagrangian function of Eq. (18) is:

$$\begin{aligned} & L(w, b, \alpha, \hat{\alpha}, \xi, \hat{\xi}, \mu, \hat{\mu}) \\ & = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i + \hat{\xi}_i) - \sum_{i=1}^m \mu_i \xi_i - \sum_{i=1}^m \hat{\mu}_i \hat{\xi}_i \\ & + \sum_{i=1}^m \alpha_i (f(x_i) - y_i - \delta - \xi_i) + \sum_{i=1}^m \hat{\alpha}_i (y_i - f(x_i) - \delta - \hat{\xi}_i) \end{aligned} \quad (19)$$

Assuming that the partial derivatives of $L(w, b, \alpha, \hat{\alpha}, \xi, \hat{\xi}, \mu, \hat{\mu})$ with respect to w , b , ξ_i and $\hat{\xi}_i$ are zero, and substitute Eq. (16) to obtain

$$w = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) x_i \quad (20)$$

$$0 = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) \quad (21)$$

$$C = \alpha_i + \mu_i = \hat{\alpha}_i + \hat{\mu}_i \quad (22)$$

Substituting Eqs. (20) ~ (22) into Eq. (19), the Wolfe dual problem of SVR is as follows:

$$\max_{\alpha, \hat{\alpha}} \sum_{i=1}^m y_i (\hat{\alpha}_i - \alpha_i) - \delta (\hat{\alpha}_i + \alpha_i) - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m (\hat{\alpha}_i - \alpha_i) (\hat{\alpha}_j - \alpha_j) x_i^T x_j \quad (23)$$

$$\begin{aligned} \text{s.t. } & \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) = 0 \\ & 0 \leq \alpha_i, \hat{\alpha}_i \leq C \end{aligned}$$

The KKT conditions that satisfy the above process are shown below:

$$\begin{cases} \alpha_i (f(x_i) - y_i - \epsilon - \xi_i) = 0 \\ \hat{\alpha}_i (y_i - f(x_i) - \epsilon - \hat{\xi}_i) = 0 \\ \alpha_i \hat{\alpha}_i = 0, \xi_i \hat{\xi}_i = 0 \\ (C - \alpha_i) \xi_i = 0, (C - \hat{\alpha}_i) \hat{\xi}_i = 0 \end{cases} \quad (24)$$

From Eq. (24), α_i can take a non-zero value when $f(x_i) - y_i - \epsilon - \xi_i = 0$, and $\hat{\alpha}_i$ is the same. At this moment, the sample (x_i, y_i) is outside the ϵ -interval band. In addition, constraints $f(x_i) - y_i - \epsilon - \xi_i = 0$ and $y_i - f(x_i) - \epsilon - \hat{\xi}_i = 0$ cannot hold simultaneously. Therefore, at least one of α_i and $\hat{\alpha}_i$ is zero. Substituting Eq. (17) into Eq. (20), the SVR solution is obtained as follows:

$$f(x) = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) x_i^T x_i + b \quad (25)$$

The SVs of the SVR model are the samples that enable Eq. (25) to satisfy $\hat{\alpha}_i - \alpha_i \neq 0$, which necessarily fall outside the ϵ -spacing band. SVs are only a part of the training samples, i.e., the SVs solution has sparsity.

From the KKT conditions of Eq. (24), for each sample there are $(C - \alpha_i) \xi_i = 0$ and condition $\alpha_i (f(x_i) - y_i - \epsilon - \xi_i) = 0$. Thus, after α_i is obtained, if $0 < \alpha_i < C$ is satisfied, then $\xi_i = 0$. Further, b is calculated as follows:

$$b = y_i + \delta - \sum_{j=1}^m (\hat{\alpha}_j - \alpha_j) x_j^T x_j \quad (26)$$

Multiple samples satisfying $0 < \alpha_i < C$ are selected to obtain the solution and the average value of b . Then the feature mapping form of the model function is as follows:

$$f(x) = w^T \phi(x) + b \quad (27)$$

Then Eq. (20) is converted into:

$$w = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) \phi(x_i) \quad (28)$$

Substituting Eq. (28) into Eq. (27), the SVR can be presented as

$$f(x) = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) \kappa(x, x_i) + b \quad (29)$$

where $\kappa(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ is the kernel function that satisfies the Mercer condition. The types of kernel function are usually PF, LF, RBF, etc. In this paper, RBF is chosen as the kernel function, because it is suitable for handling nonlinear mappings. The specific form of RBF is as follows:

$$\kappa(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (30)$$

where σ is the width factor of the function, which is used to determine the radial action area of the function.

Since SVR describes a convex programming problem, any one of its solutions is globally optimal, which helps to reduce the cases where ML models are trapped in a locally optimal solution. From Eq. (17) and (30), the regularization factor C and the width parameter σ are important parameters for the SVR. Where C is the penalty coefficient. The higher the value of C , the larger the error and the easier the model is to overfit; the smaller C is, the model tends to be underfitted and the generalization ability becomes worse. The width parameter σ determines the data distribution after mapping to a new feature space. The larger the σ , the smaller the standard deviation, which only acts near the SVs sample, resulting in fewer SVs and the model tends to overfit. Therefore, the proper selection of these two parameters is crucial. In this paper, a grid search method is used to select and optimize the regularization coefficient C and width parameter σ to achieve more accurate and robust SOH estimation.

3.2.3 Nonlinear constraint optimization

Nonlinear constraint optimization is used to solve the minimum of the objective function with constraints. The standard form of the fmincon solver is as follows:

$$\begin{aligned} \min f(x) & \quad (31) \\ \text{s.t. } A \cdot x & \leq b \\ Aeq \cdot x & = beq \\ c(x) & \leq 0 \\ Ceq(x) & = 0 \\ lb & \leq x \leq ub \end{aligned}$$

where $A \cdot x \leq b$ is a linear inequality constraint, $Aeq \cdot x = beq$ is a linear equation constraint, $c(x) \leq 0$ is a nonlinear inequality constraint, $Ceq(x) = 0$ is a nonlinear equation constraint, and $lb \leq x \leq ub$ is a variable

constraint.

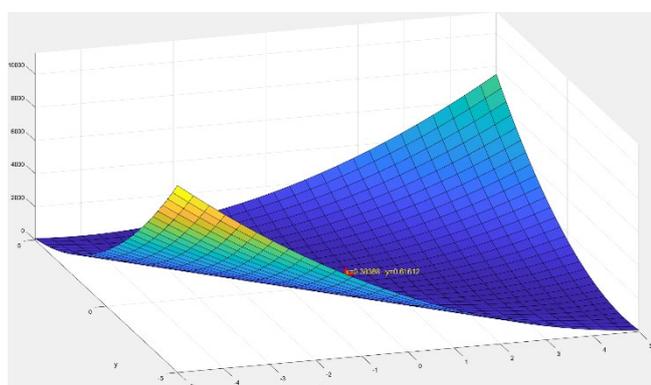
In this paper, the optimization option `OPTIONS` in the `fmincon` solver is specified as an `OPTIMSET` structure.

The selected objective function is as follows:

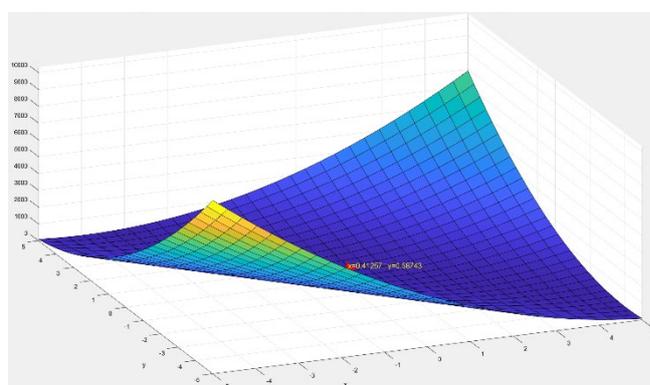
$$f(x) = \sum_{i=1}^n (O_{real} - (\omega_1 \cdot O_{LSTM} + \omega_2 \cdot O_{SVR}))^2 \quad (32)$$

where n is the training step, O_{real} is the reference value, O_{LSTM} is the test output of the LSTM network, O_{SVR} is the test output of the SVR, ω_1 and ω_2 are the weights of individual models, which satisfy $\omega_1 + \omega_2 = 1$.

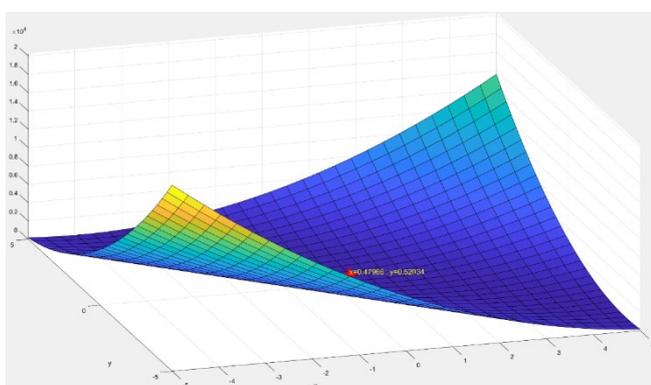
The combined model weights obtained from the nonlinear constraint optimization are shown in Fig. 7. The combinatorial model weights corresponding to each experimental result in Chapter 4 are shown in Tab. 4.



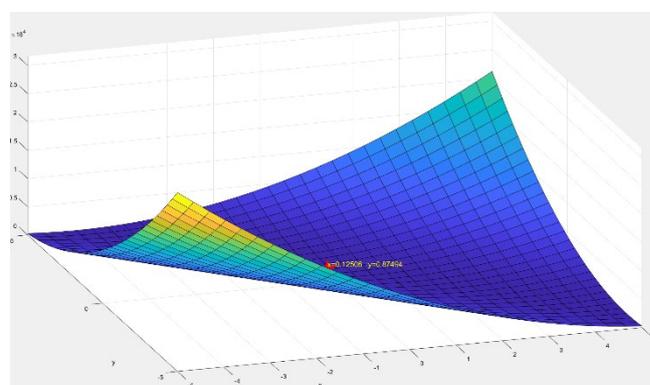
(a) Combined model weights for results in Fig. 9.



(b) Combined model weights for results in Fig. 10.



(c) Combined model weights for results in Fig. 11. (a)



(d) Combined model weights for results in Fig. 11. (a)

Fig. 7. Partial model weights obtained from the `fmincon` solver

Tab. 4. Combined model weights corresponding to the experimental results

Corresponding experiments	Model weights		Corresponding experiments	Model weights	
	LSTM	SVR		LSTM	SVR
Fig. 8.	0.38388	0.61612	Fig. 12. (c)	0.30486	0.69514
Fig. 9.	0.41257	0.58743	Fig. 14. (a)	0.32876	0.67124
Fig. 11. (a)	0.47966	0.52034	Fig. 14. (c)	0.40443	0.59557
Fig. 11. (c)	0.12506	0.87494	Fig. 15. (a)	0.13991	0.86009
Fig. 12. (a)	0.27123	0.72877	Fig. 15. (c)	0.40964	0.59039

From Table 4, the LSTM-SVR combined model can dynamically adjust the weights of the two models under

different operating conditions, which makes the model with advantages account for a larger proportion, and thus the accuracy and robustness of SOH estimation are effectively improved under different input feature lengths and different operating conditions. Meanwhile, the LSTM-SVR combined model is not a complex strong estimator, and it has the advantages of simplicity of use, less computation and good stability, which can be considered for practical applications.

3.3 SOH prediction framework

The overall block diagram of the LSTM-SVM combined model of SOH estimation for lithium-ion batteries is shown in Fig. 8.

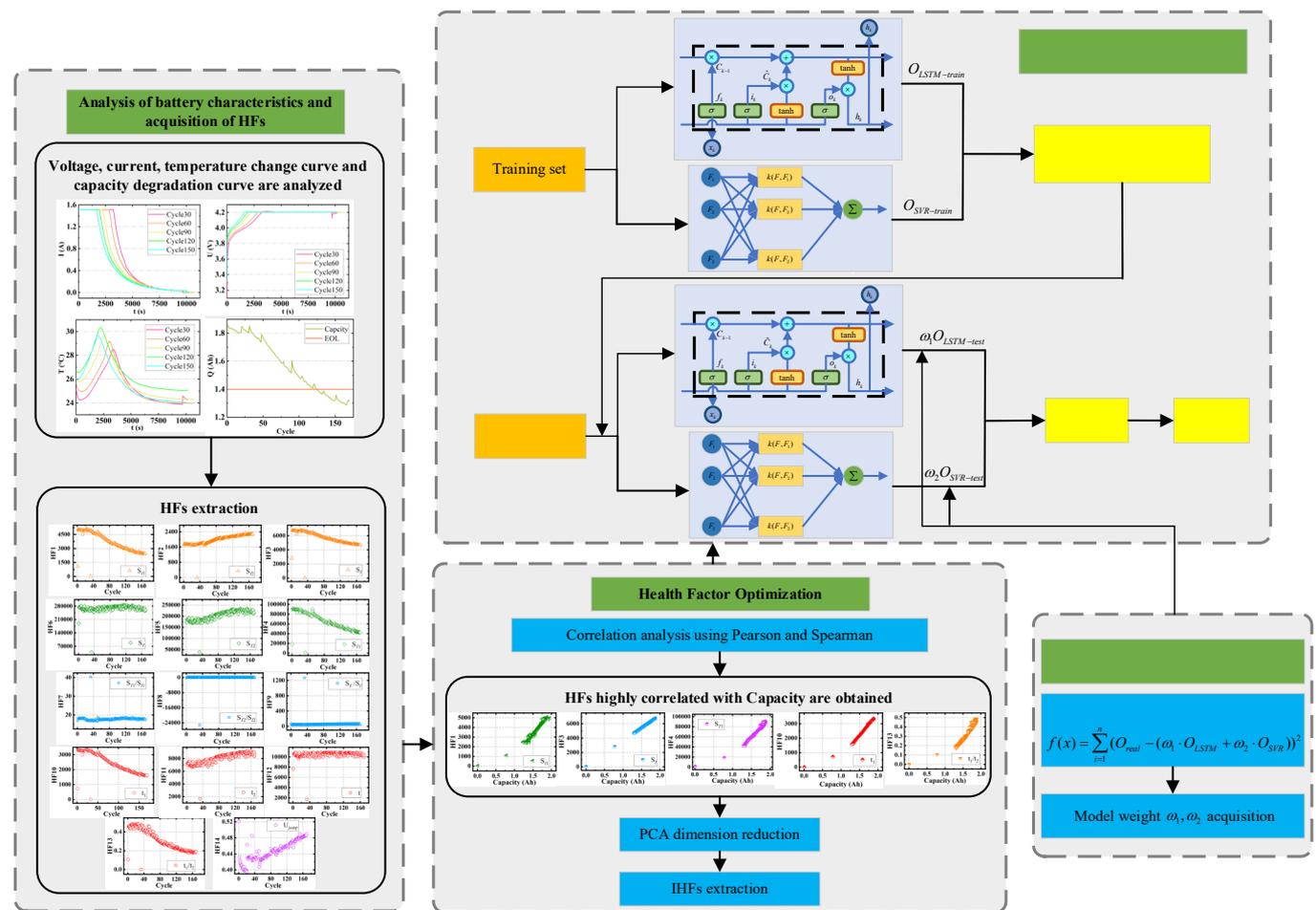


Fig. 8. Flowchart of LSTM-SVM combined model for lithium-ion batteries SOH estimation

As shown in Fig. 8, the proposed SOH estimation method for lithium-ion batteries includes HFs selection and LSTM-SVR combined model construction. Among them, HFs are obtained by analyzing the current, voltage and temperature change curves of the lithium-ion battery charging process and the overall capacity degradation curve. The final IHFs were selected by Pearson and Spearman correlation analysis and PCA dimensionality reduction. The LSTM-SVR combined model is an efficient combination of the LSTM network and SVR. The weights of individual models are determined by the nonlinear constraint optimization to minimize the sum of squared errors of the

1
2
3 combined model.

4
5 However, the proposed method has some limitations in SOH estimation. First, the SOH estimation does not
6
7 consider the effect of RUL, and the reasonableness of the estimation results needs to be improved. Second, The SOH
8
9 estimation does not consider the partial charging curves as input. In practice, SOH estimation using the partial charge
10
11 curves can increase the time that the driver starts charging from the scheduled charging state. However, this
12
13 experiment is a simulation attempt to investigate the combined model with strong input characteristics, less
14
15 computational burden and multi-dominance combination for accurate and robust SOH estimation of lithium-ion
16
17 batteries. In actual vehicle operation, lithium-ion batteries are exposed to complex environmental conditions that
18
19 require the ability to quickly and accurately assess the SOH. The proposed SOH estimation approach is expected to
20
21 improve the accuracy of SOH estimation for different input feature lengths and different operating conditions by
22
23 combining ML models with different characteristics.

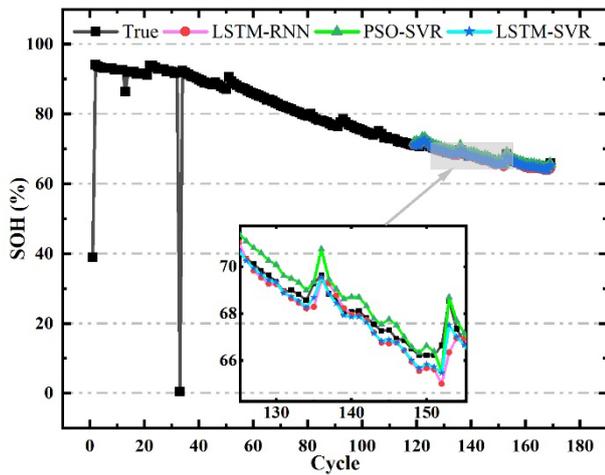
24 25 **4 SOH experiment results and analysis**

26
27 This section further analyzes and discusses the accuracy and robustness of the proposed method under different
28
29 operating conditions. LSTM-RNN network [57] and PSO-SVR [58] are chosen as the comparison benchmark. For
30
31 fair comparative verification, the LSTM network and SVR model have the same structure and parameters except for
32
33 the RNN network and the PSO algorithm. The lithium-ion batteries data used are the six cells in Tab. 1. Two sets of
34
35 cell data at room temperature (B0005 and B0006) are used to test the accuracy and robustness of the LSTM-SVR
36
37 combined model for SOH estimation, two sets of cell data at high temperature (B0030 and B0032) and two sets of
38
39 cell data at low temperature (B0045 and B0047) are used to test the accuracy of the LSTM-SVR combined model
40
41 for SOH estimation at different temperatures.

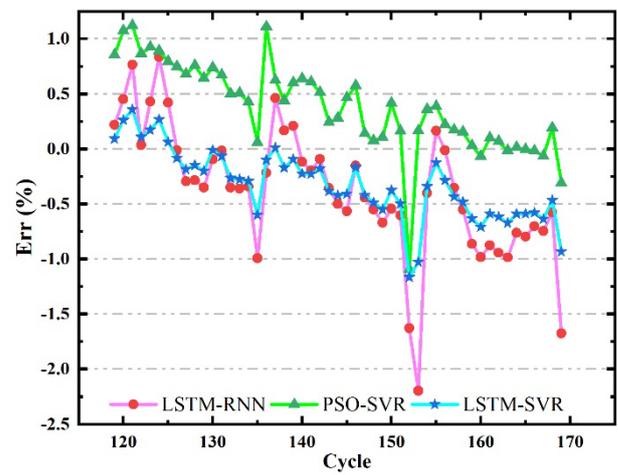
42
43 Although capacity can indicate the degree of battery aging during battery life decay, it is not easily measured
44
45 directly by the sensor. Therefore, this paper utilizes the IHFs selected in Section 2.2 as the input features and SOH
46
47 as the output features. In addition, to better evaluate the SOH estimation effect of the LSTM-SVR combined model,
48
49 MaxAE, MAE, RMSE, and R-squared coefficient are introduced as the evaluation metrics of the model performance.

50 51 *4.1 SOH accuracy analyses*

52
53 To evaluate the accuracy of the LSTM-SVR combined model for SOH estimation, the first 70% of IHFs data
54
55 from B0005 and B0006 cells were used as model training, and the last 30% of IHFs data were used as model testing.
56
57 The SOH estimation results of B0005 and B0006 cells are plotted in Fig. 9 and Fig. 10, respectively, and their error
58
59 properties are evaluated as shown in Fig. 11.
60

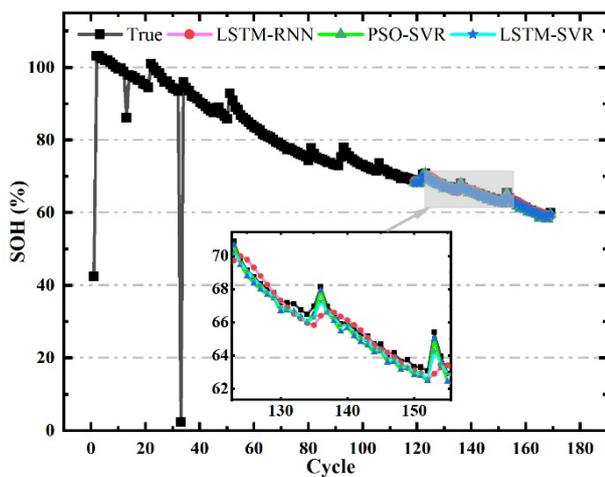


(a) SOH estimated change curves with 70% test set

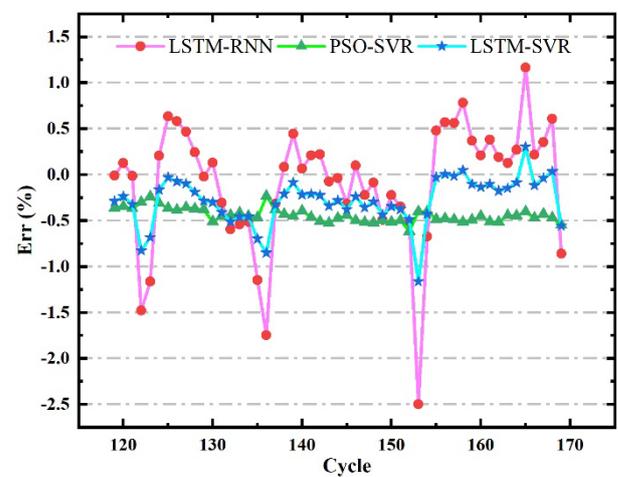


(b) SOH estimation errors with 70% test set

Fig. 9. SOH estimation results of B0005 cell with 70% test set



(a) SOH estimated change curves with 70% test set



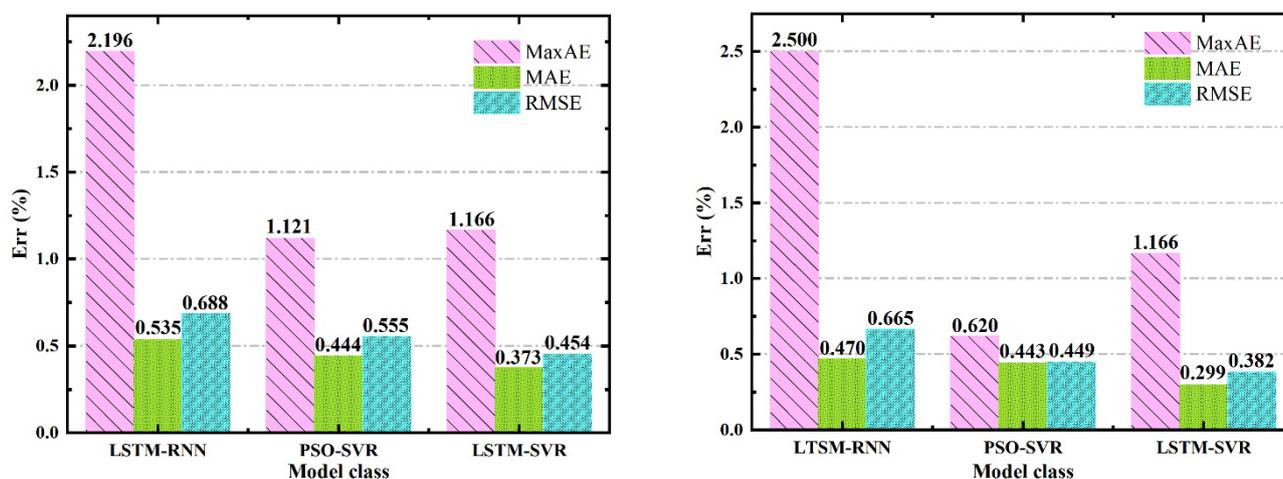
(b) SOH estimation errors with 70% test set

Fig. 10. SOH estimation results of B0006 cell with 70% test set

As shown in Fig. 9 (a) and Fig. 10 (a), the initial capacity values of the B0005 and B0006 cells show low capacity, and the capacity curve shows a missing sample (zero value). NASA has not yet explained well the reason for this extreme situation. However, the overall trend of battery capacity is reasonable, which can be used to evaluate the proposed SOH estimation method.

The LSTM-RNN network can adequately handle the problem of time series prediction to effectively learn the long short-term dependencies of battery capacity decay trends from historical data. However, the RNN network suffers from the problem of exploding or vanishing gradients. In addition, the LSTM-RNN network is not suitable for effectively characterizing the capacity regeneration phenomenon. Because the LSTM-RNN network relies heavily on historical data and large sample data, which are less adaptive to outliers (e.g., capacity regeneration phenomena, fluctuations caused by external disturbances, etc.). The PSO algorithm has better global optimization ability for the penalty parameters and kernel parameters, and a suitable kernel function can effectively solve the regression problem with high-dimensional features and nonlinearities. The PSO-SVR model has a good estimation

effect for small samples and reflects the regeneration phenomenon of battery capacity well. However, the PSO-SVR model is a weak estimator with poor temporal ordering and is not as effective as the LSTM-RNN network in tracking actual SOH changes. Notably, the LSTM-SVR combined model synthesizes the advantages of the two models while compensating for the shortcomings, which not only captures the SOH changes well but also effectively reflects the capacity regeneration phenomenon of the battery and noise-induced fluctuations. As the cycle number increases, although the bias of the PSO-SVR model gradually decreases, its initial predicted value deviates more from the actual value. Although the SOH estimation bias of the LSTM-RNN network and the LSTM-SVR combined model increases slightly with the cycle number, they are within reasonable limits. Among them, the initial bias of the LSTM-SVR combined model is small, and the increase in bias with the increase in the cycle number is also small, which has a good SOH estimation effect.



(a) SOH estimation error properties of B0005 cell

(b) SOH estimation error properties of B0006 cell

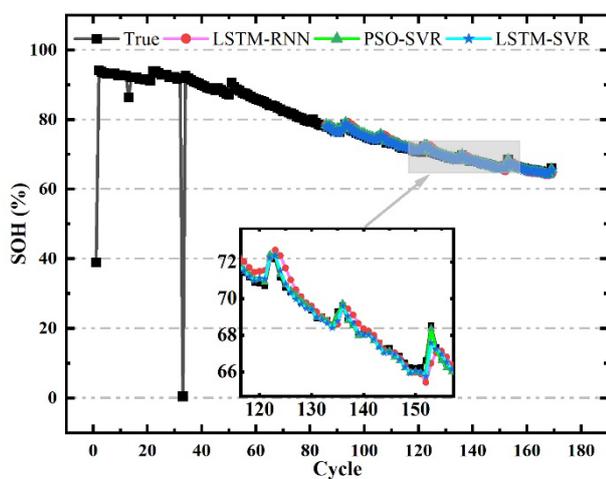
Fig. 11. SOH estimation error properties of two cells with 70% test set

From Fig. 11, since the LSTM-RNN network is not suitable for small samples and is poorly adapted to outliers, its overall estimation effect is poor and has large error characteristics. Among the three models, the PSO-SVR model has the smallest MaxAE, because the PSO algorithm finds the optimal values of the penalty parameter and kernel parameter of the SVR model faster than the grid search method. However, the PSO-SVR model is weak in handling time series data with a high ratio of MAE to RMSE. The LSTM-SVR combined model incorporates the advantages of the two models while compensating for the limitations of the single model, which has the lowest MaxAE, MAE, and RMSE among the three. In addition, the SOH estimation errors of the LSTM-SVR combined model are all less than 1.166%, with good prediction accuracy and robustness.

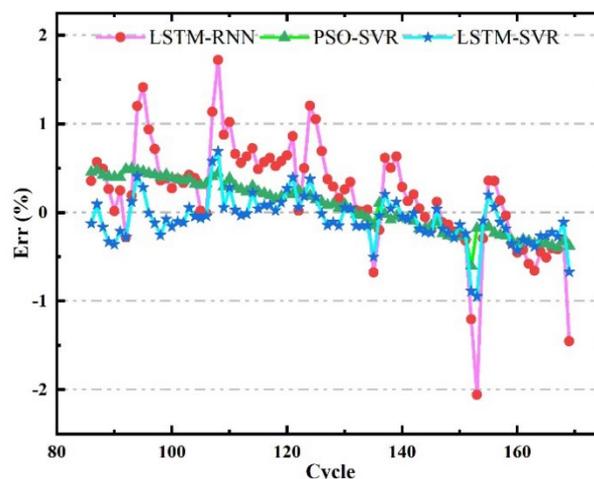
4.2 SOH robustness analysis

To analyze the robustness of the LSTM-SVR combined model under different cycle counts, the first 50% and first 30% of IHFs data were used as the training set of the proposed model for B0005 and B0006 cells, respectively.

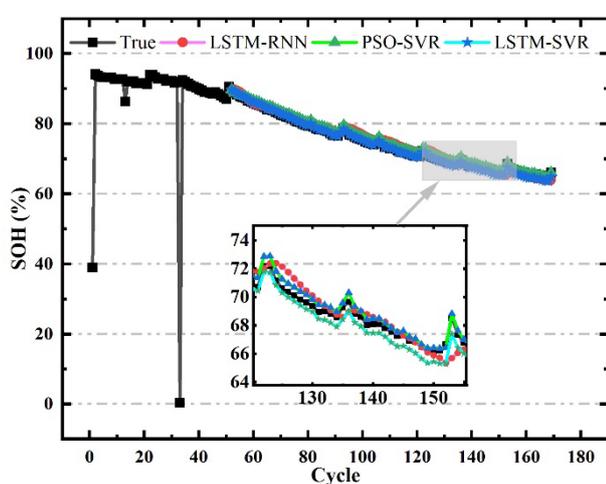
The results of the robustness evaluation for B0005 and B0006 cells are present in Fig. 12 and Fig. 13, respectively, and their error properties are shown in Fig. 14.



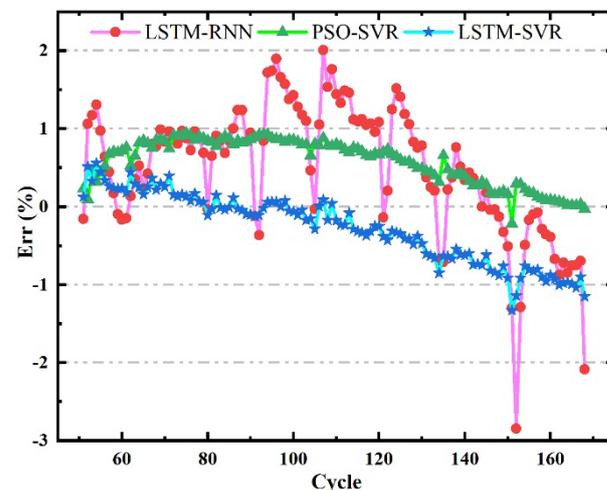
(a) SOH estimated change curves with 50% test set



(b) SOH estimation errors with 50% test set

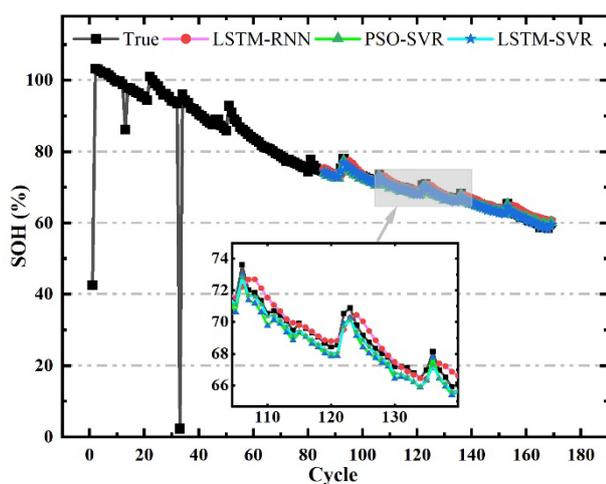


(c) SOH estimated change curves with 30% test set

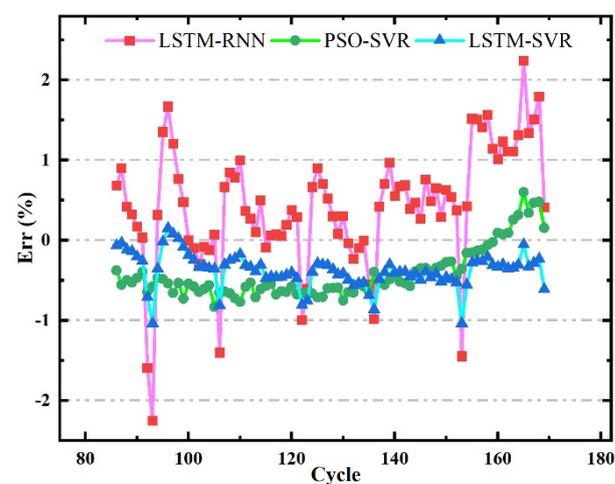


(d) SOH estimation errors with 30% test set

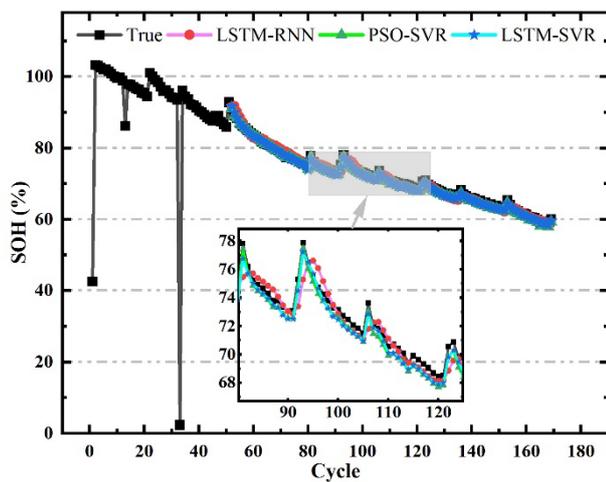
Fig. 12. SOH estimation results of B0005 cell with different cycles



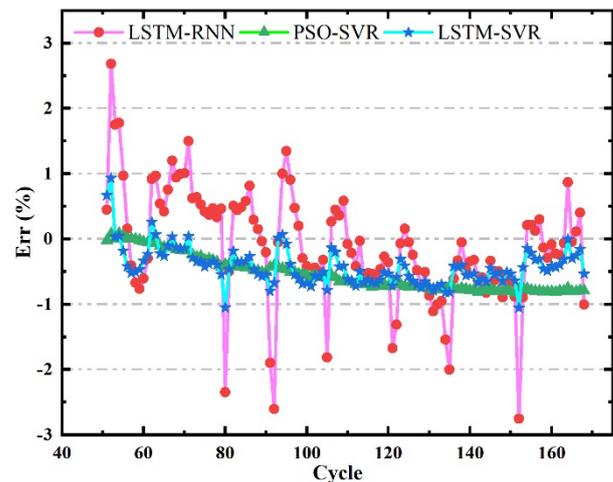
(a) SOH estimated change curves with 50% test set



(b) SOH estimation errors with 50% test set



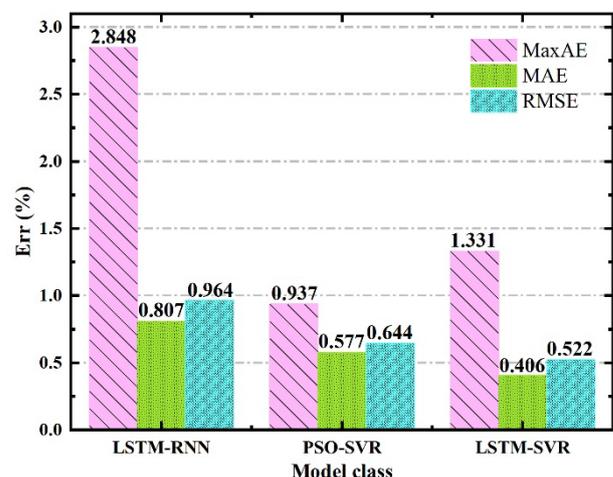
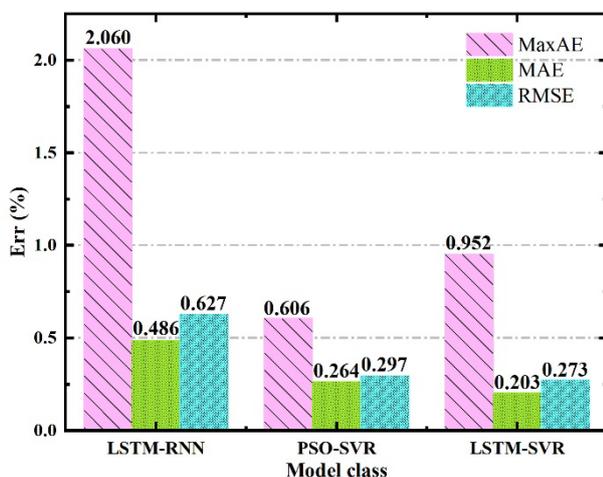
(c) SOH estimated change curves with 30% test set



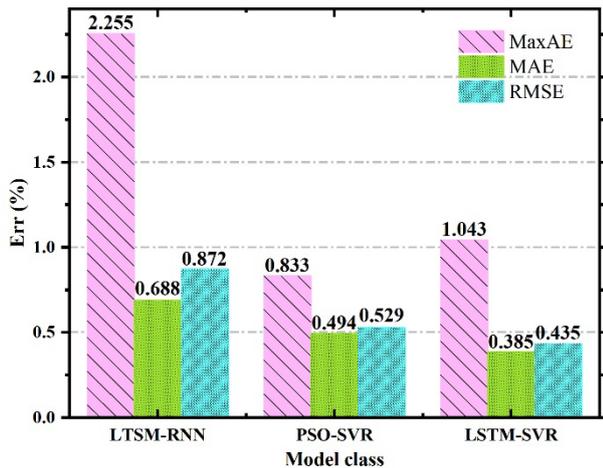
(d) SOH estimation errors with 30% test set

Fig. 13. SOH estimation results of battery B0006 with different cycles

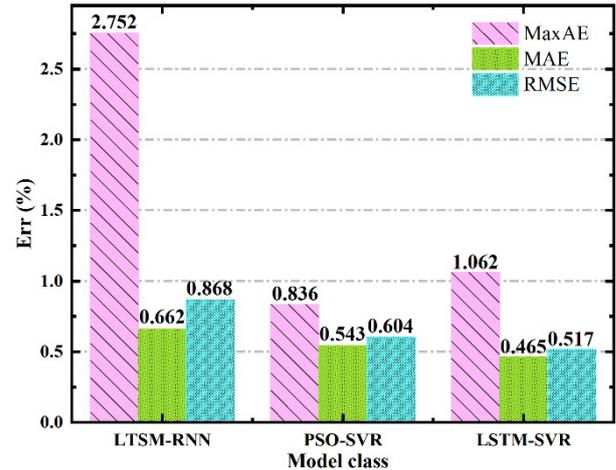
From Figs. 12 and 13, the cyclic aging predictions of all three methods exhibit good uniformity. The LSTM-RNN network can effectively utilize historical data to obtain the changing pattern of battery aging. However, for the capacity regeneration phenomenon and fluctuations caused by external perturbations, the PSO-SVR model has better adaptability than the LSTM-RNN network. In addition, as the sample size decreases, the advantage of the LSTM-RNN network gradually diminishes, and the PSO-SVR model, which has strong generalization ability in the face of small samples, gradually dominates. The LSTM-SVR network combined model can dynamically adjust the weights of different models under the conditions of different cutoff voltages and training set data, which can effectively combine the advantages of the two models. It has the advantages of both LSTM network with strong temporal order and full utilization of historical data, as well as SVR model nonlinear mapping and small-sample learning, which enables accurate and robust SOH estimation. Tab. 4 can further support the above analysis. For B0005 and B0006 cells, the weight of the SVR model in the combined model gradually increases as the training set data decreases. In addition, the bias of SOH estimation of the three methods gradually increases as the cycle number increases and the training set decreases. Among them, the LSTM-SVR combined model has small bias increases and good robustness.



(a-1) SOH estimation error properties of B0005 with 50% test set



(a-2) SOH estimation error properties of B0005 with 30% test set



(b-1) SOH estimation error properties of B0006 with 50% test set

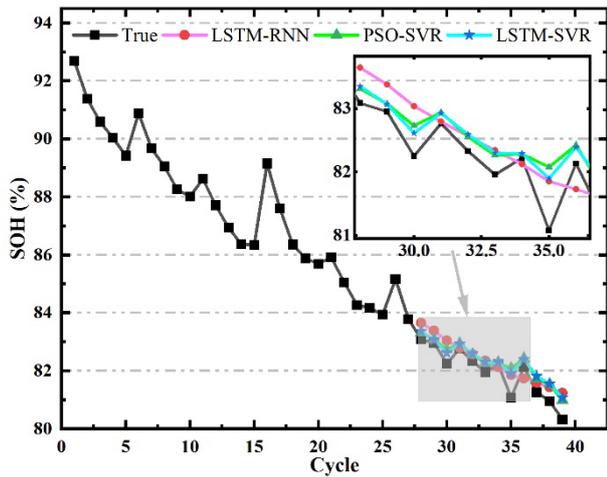
(b-2) SOH estimation error properties of B0006 with 30% test set

Fig. 14. SOH estimation error properties of two cells with different test set

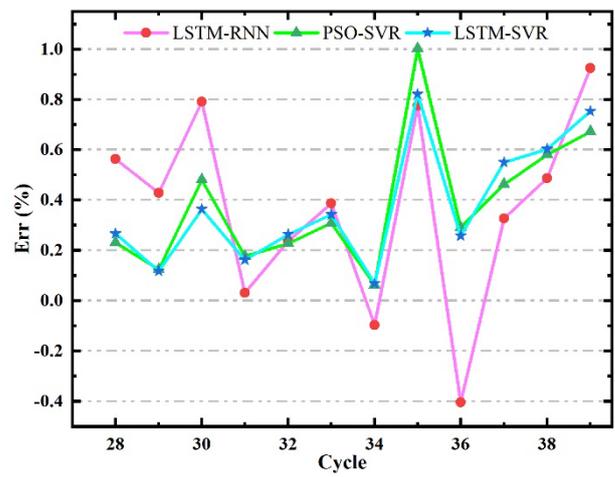
From Fig. 14, all three methods have better SOH estimation despite the reduced training sample size. The LSTM-RNN network has a large MaxAE due to the limitations of small sample inapplicability and outlier insensitivity. However, the LSTM-RNN network is more capable of processing time series data than the PSO-SVR model, with a lower ratio of MAE to RMSE and better following ability. The PSO-SVR model has a strong generalization ability when dealing with small samples with the smallest MaxAE among the three models. Notably, the LSTM-SVR combined model has the lowest MAE and RMSE, as well as a reasonable range of MaxAE, which ensures the accuracy and robustness of SOH estimation under different training sample sizes.

4.3 SOH evaluation under different working conditions

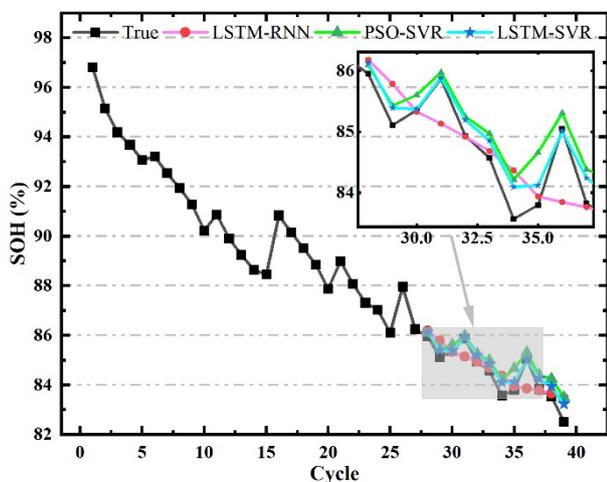
In practical applications, lithium-ion batteries are exposed to a wide range of temperature variations due to seasonal and temporal changes. Both high and low can affect the accurate SOH estimation. To evaluate the SOH estimation effect of the LSTM-SVR combined model at high and low temperatures, the B0030 and B0032 cell data from the battery cycle aging experiment at 43 °C and the B0045 and B0047 cell data from the battery cycle aging experiment at 4 °C were used to test the temperature adaptation of the model. Among them, the first 70% IHFs data were used as model training and the last 30% IHFs data were used as model testing. The experimental results of the high-temperature and low-temperature are plotted in Fig. 15 and Fig. 16, respectively, and their error properties are shown in Fig. 17.



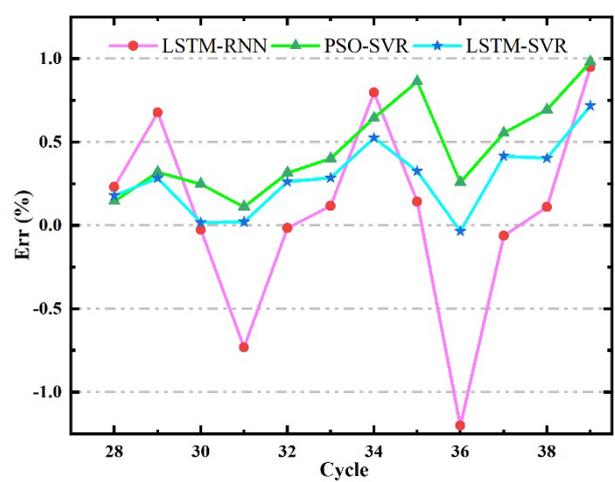
(a) SOH estimated change curves of B0030 cell at 43°C



(b) SOH estimation errors of B0030 cell at 43°C

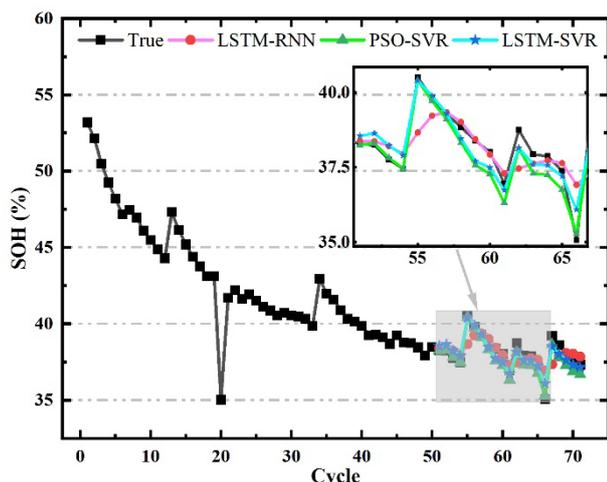


(c) SOH estimated change curves of B0032 cell at 43°C

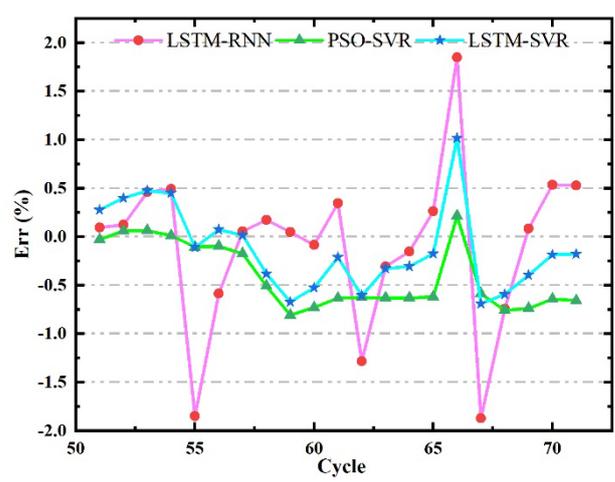


(d) SOH estimation errors of B0032 cell at 43°C

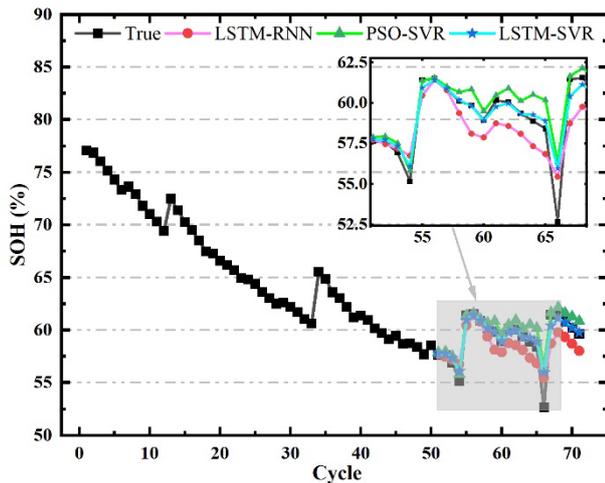
Fig. 15. SOH estimation results of two cells at 43°C



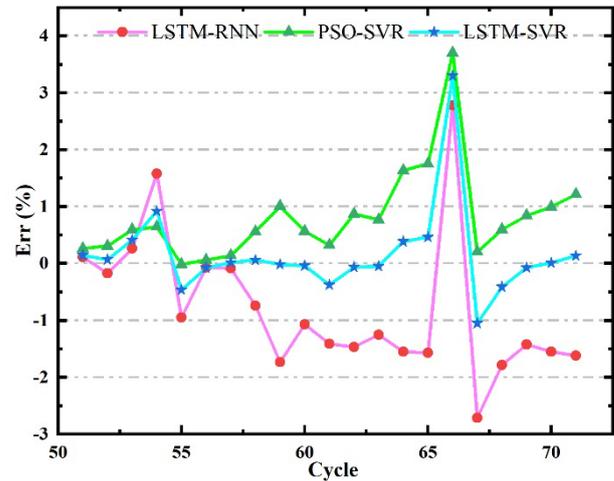
(a) SOH estimated change curves of B0045 cell at 4°C



(b) SOH estimation errors of B0045 cell at 4°C



(c) SOH estimated change curves of B0047 cell at 4°C



(d) SOH estimation errors of B0047 cell at 4°C

Fig. 16. SOH estimation results of two cells at 4°C

As shown in Fig. 16 (a) and (c), the maximum observed capacity of the B0045 and B0047 cells is below 1.4 Ah, indicating severe performance degradation. In addition, except for B0045, the cyclic aging trends of the other three batteries were relatively regular, with no extreme cases of low capacity and lack of samples. Although NASA has not yet explained the cause of this extreme situation, the overall battery capacity decline trend is reasonable and can be used to validate the proposed SOH estimation method. Battery aging involves complex electrochemical reactions, and the aging rate often shows a strong sensitivity.

From Fig. 15 and Fig.16, the capacity of the B0030 and B0032 cells decayed faster than the B0045 and B0047 cells for the same cycle number. Because the increased discharge rate accelerates the decay rate of the battery capacity. Compared to other cells, B0045 cells are characterized by faster aging, lower SOH estimation, and significant capacity degradation. Because the lower discharge cut-off voltage leads to an increase in internal resistance, further accelerating battery aging and capacity degradation. Ambient temperature is also one of the most important factors affecting battery aging. Due to the effects of high and low temperatures, the capacity of these four cells degrades faster and with fewer cycles compared to B0005 and B0006 cells. At an ambient temperature of 43°C, some irreversible chemical reactions inside the battery are accelerated, resulting in decreases in the active substance of the battery, and accelerated battery aging and capacity degradation. At an ambient temperature of 4°C, some chemical reactions within the battery slow down, and the discharge capacity and discharge platform of the battery decrease, resulting in the power and energy output of the battery being affected. However, all three methods showed a good SOH estimation effect under different operating conditions. Among them, the LSTM-SVR combined model can not only follow the change of battery capacity degradation better but also reflect the capacity regeneration phenomenon and fluctuations caused by external perturbations better, which has a good SOH estimation performance.

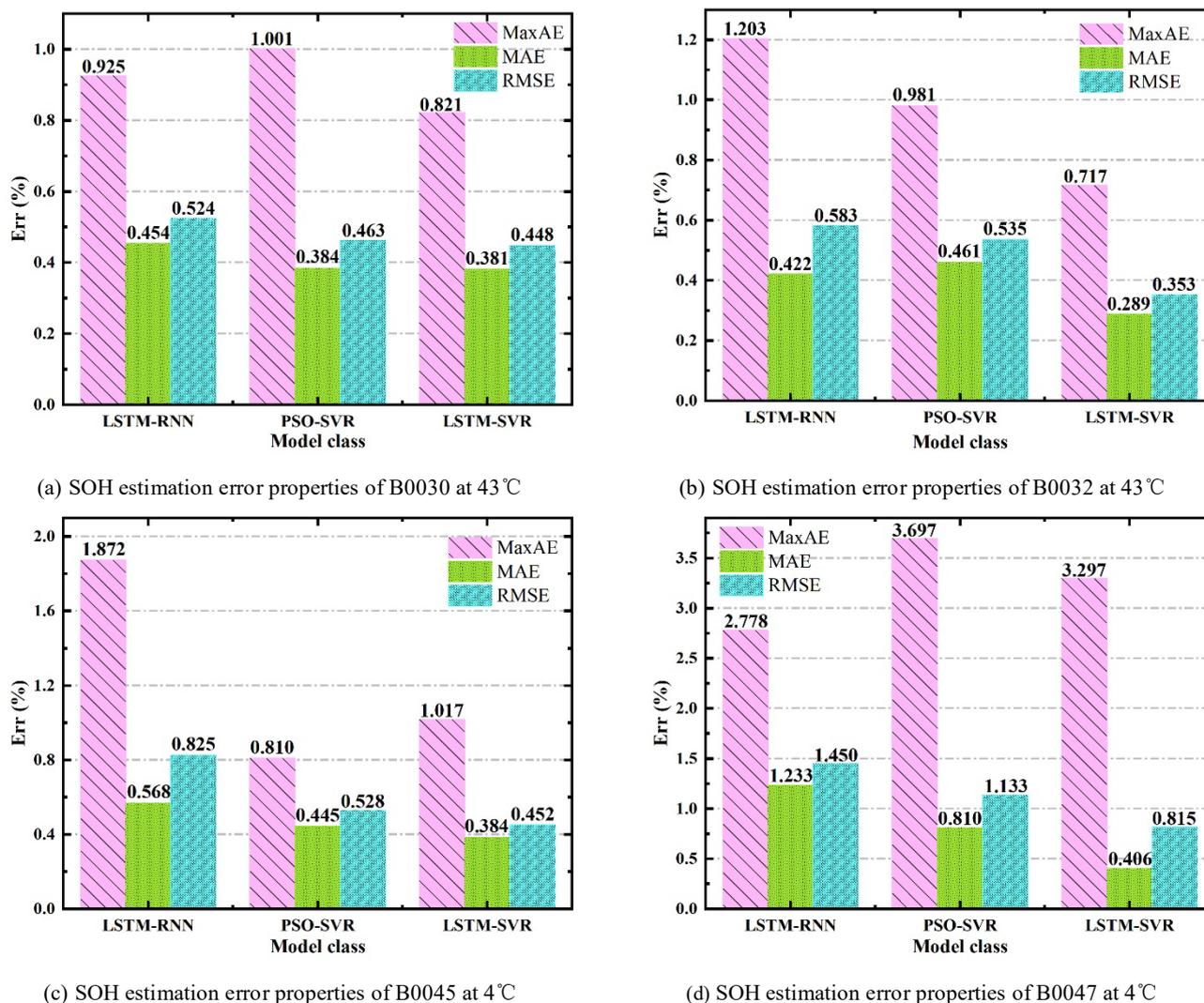


Fig. 17. SOH estimation error properties of four cells at different temperature

From Fig. 17, a good training set allows the data-driven model to effectively track the decaying changes in battery capacity. Since the LSTM-RNN network does not apply to small-sample estimation, its MAE and RMS are generally higher than those of the PSO-SVR model. The overall SOH estimation effect of the LSTM-SVR combined model is the best among the three methods, with MAE and RMSE for the four cells were less than 0.406% and 0.815%, respectively. The PSO-SVR model has good generalization ability for small-sample estimation and can track the outliers during battery aging well, so the MaxAE of the PSO-SVR model is generally lower than that of the LSTM-SVR combined model. However, the MaxAE of the LSTM-SVR combined model is still within reasonable range. In addition, the LSTM-SVR combined model not only effectively combines the models with different merits, but also better adapts to different working conditions, which makes the SOH estimation more accurate and robust.

The R-squared coefficients of the LSTM-RNN network, PSO-SVR, and LSTM-SVR combined model corresponding to all the above experiments are shown in Fig. 18.

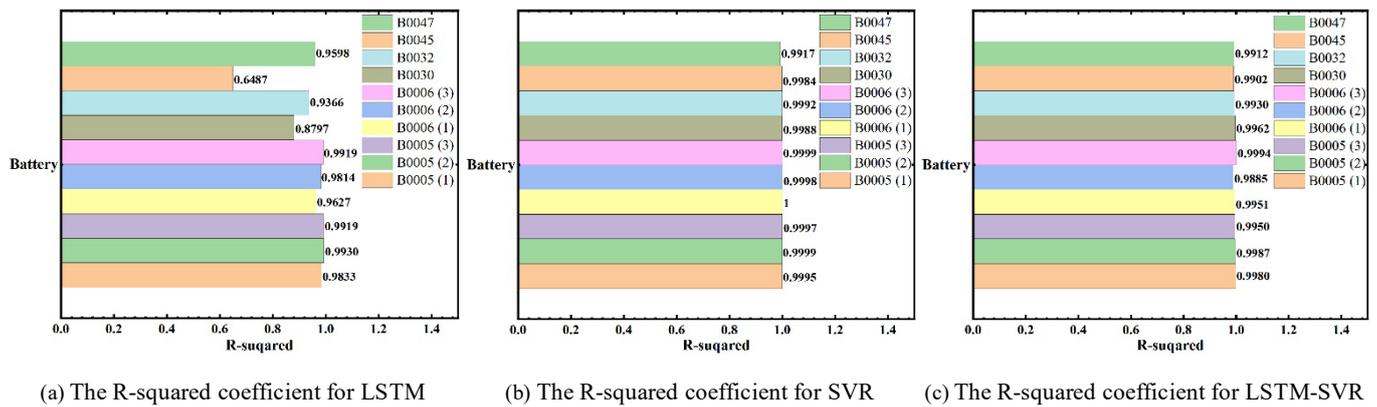


Fig. 18. The R-squared coefficients for the three models

The R-squared coefficient is the ratio of the regression sum of squares to the total deviation sum of squares, which is between 0 and 1. The larger the value, the more accurate the model and the better the regression fit. B0005 (1) ~ (3) denote experiments with 70%, 50% and 30% of IHFs data as the model training sets, respectively, and B0006 (1) ~ (3) similarly. From Fig. 18, the R-squared coefficient of the LSTM-SVR combined model is approximately 1, which has a good accuracy and fitting effect. Although the R-squared coefficient increases as the counts of independent variables increase, the increase of low or even uncorrelated independent variables does not reflect the fitting ability and estimation accuracy of the model. The input features of the LSTM-SVR combined model are IHFs that are highly correlated with capacity and in appropriate quantities. Therefore, the feasibility of the proposed combined model can be confirmed by the R-squared coefficient.

5 Conclusions

Considering the realistic issues of economic costs, computational costs and overall benefits, it is crucial to balance weak and strong machine learning models to make them efficient in real vehicle applications. To obtain a good SOH estimation effect in actual life, the two vital directions of feature selection and model optimization must receive more attention. In this paper, a combined ML model with strong input features, small computational burden and multi-dominance combination is proposed for accurate and robust lithium-ion batteries SOH estimation. This method improves effectively the accuracy and robustness of SOH estimation under different input feature lengths and different operating conditions by combining ML models with different characteristics. First, the current, voltage and temperature change curves during the battery charging process were analyzed, and fourteen HFs related to capacity decay were extracted. In addition, the Pearson and Spearman correlation analysis was used to select the HFs that are highly correlated with battery capacity degradation. Furthermore, a principal component analysis based on eigenvalue decomposition was used to downscale the selected high-dimensional features. Thus, the IHFs that effectively characterize battery aging are obtained as input features to the model. Second, the LSTM network with strong time-permissibility and the SVR model with the advantages of nonlinear mapping and small sample learning

1
2 are chosen for combination. The LSTM network is used to compensate for the disadvantages of the SVR model, such
3 as poor timing and difficulty in preserving information for a long time. The SVR model is used to compensate for
4 the disadvantages of the LSTM network, which is not suitable for small sample estimation and has randomness.
5
6 Finally, a novel LSTM-SVR combined model is constructed by using a nonlinear constraint optimization to assign
7 the weights of individual models in terms of minimizing the sum of squared errors of the combined model. The
8 proposed SOH estimation method is validated using different batteries, test sets of different proportions and different
9 operating conditions. The results present that the LSTM-SVR combined model has good accuracy and robustness in
10 SOH estimation. In the validation of accuracy and robustness, the MAE and RMSE of the LSTM-SVR combined
11 model are all less than 0.75% and 0.91%, respectively; in the validation of the adaptability for different working
12 conditions, the MAE and RMSE of the LSTM-SVR combined model are all less than 0.53% and 0.97%, respectively.
13
14 The LSTM-SVR combined model has the advantages of the individual model, which is not only simple to use, less
15 computational burden and has good stability but also can adapt to different input feature lengths and different working
16 conditions with high accuracy and robustness. In addition, the rapid advances in computer technology and cloud
17 computing hold great promise for the widespread application of ML technologies to real life. Therefore, the LSTM-
18 SVR combined model could be an option for accurate and robust SOH estimation of lithium-ion batteries in practical
19 applications in the future.
20
21
22
23
24
25
26
27
28
29
30
31
32

33 **Acknowledgments**

34 The work is supported by the National Natural Science Foundation of China (No. 62173281).
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- [1]. Dai, H., et al., Advanced battery management strategies for a sustainable energy future: Multilayer design concepts and research trends. *Renewable & Sustainable Energy Reviews*, 2021. **138**.
- [2]. Ge, M.-F., et al., A review on state of health estimations and remaining useful life prognostics of lithium-ion batteries. *Measurement*, 2021. **174**.
- [3]. Wang, P., et al., Stochastic management of hybrid AC/DC microgrids considering electric vehicles charging demands. *Energy Reports*, 2020. **6**: p. 1338-1352.
- [4]. Song, K., et al., Remaining life prediction of lithium-ion batteries based on health management: A review. *Journal of Energy Storage*, 2023. **57**.
- [5]. Pan, D., et al., Evaluating the accuracy of electro-thermal coupling model in lithium-ion battery via altering internal resistance acquisition methods. *Journal of Power Sources*, 2020. **463**.
- [6]. Vanem, E., et al., Data-driven state of health modelling-A review of state of the art and reflections on applications for maritime battery systems. *Journal of Energy Storage*, 2021. **43**.
- [7]. Preger, Y., et al., Degradation of Commercial Lithium-Ion Cells as a Function of Chemistry and Cycling Conditions. *Journal of the Electrochemical Society*, 2020. **167**(12).
- [8]. Copley, R.J., et al., Measurements and modelling of the response of an ultrasonic pulse to a lithium-ion battery as a precursor for state of charge estimation. *Journal of Energy Storage*, 2021. **36**.
- [9]. Semeraro, C., et al., Battery monitoring and prognostics optimization techniques: Challenges and opportunities. *Energy*, 2022. **255**.
- [10]. Basia, A., et al., Review on State of Health estimation methodologies for lithium-ion batteries in the context of circular economy. *Cirp Journal of Manufacturing Science and Technology*, 2021. **32**: p. 517-528.
- [11]. Remmlinger, J., et al., State-of-health monitoring of lithium-ion batteries in electric vehicles by on-board internal resistance estimation. *Journal of Power Sources*, 2011. **196**(12): p. 5357-5363.
- [12]. Meddings, N., et al., Application of electrochemical impedance spectroscopy to commercial Li-ion cells: A review. *Journal of Power Sources*, 2020. **480**.
- [13]. Li, X.Y., et al., State of health estimation for Li-Ion battery using incremental capacity analysis and Gaussian process regression. *Energy*, 2020. **190**.
- [14]. Barai, A., et al., A study of the influence of measurement timescale on internal resistance characterisation methodologies for lithium-ion cells. *Scientific Reports*, 2018. **8**.
- [15]. Pradhan, S.K. and B. Chakraborty, Battery management strategies: An essential review for battery state of health monitoring techniques. *Journal of Energy Storage*, 2022. **51**.
- [16]. Vyroubal, P. and T. Kazda, Equivalent circuit model parameters extraction for lithium ion batteries using electrochemical impedance spectroscopy. *Journal of Energy Storage*, 2018. **15**: p. 23-31.
- [17]. Choi, W., et al., Modeling and Applications of Electrochemical Impedance Spectroscopy (EIS) for Lithium-ion Batteries. *Journal of Electrochemical Science and Technology*, 2020. **11**(1).
- [18]. Villuri, R.T., M. Singh, and Y. Beck, Experimental analysis of electric vehicle's Li-ion battery with constant pulse and constant voltage charging method. *International Journal of Energy Research*, 2022. **46**(15): p. 22365-22385.
- [19]. Ansean, D., et al., Lithium-Ion Battery Degradation Indicators Via Incremental Capacity Analysis. *Ieee Transactions on Industry Applications*, 2019. **55**(3): p. 2992-3002.
- [20]. Zhu, J., et al., Investigation of lithium-ion battery degradation mechanisms by combining differential voltage analysis and alternating current impedance. *Journal of Power Sources*, 2020. **448**.
- [21]. Wang, Z., et al., State of health estimation of lithium-ion batteries based on the constant voltage charging curve. *Energy*, 2019. **167**: p. 661-669.
- [22]. Stroe, D.-I. and E. Schaltz, Lithium-Ion Battery State-of-Health Estimation Using the Incremental Capacity Analysis Technique. *Ieee Transactions on Industry Applications*, 2020. **56**(1): p. 678-685.

- 1
2 [23]. Zhang, S., et al., A rapid online calculation method for state of health of lithium-ion battery based on coulomb counting
3 method and differential voltage analysis. *Journal of Power Sources*, 2020. **479**.
4
5 [24]. Ruan, H., et al., Artificial Intelligence-based health diagnostic of Lithium-ion battery leveraging transient stage of
6 constant current and constant voltage charging. *Applied Energy*, 2023. **336**.
7
8 [25]. Wang, Z., et al., A review on online state of charge and state of health estimation for lithium-ion batteries in electric
9 vehicles. *Energy Reports*, 2021. **7**: p. 5141-5161.
10
11 [26]. Tamilselvi, S., et al., A review on battery modelling techniques. 2021. **13**(18): p. 10042.
12
13 [27]. Tran, M.-K., et al., A comprehensive equivalent circuit model for lithium-ion batteries, incorporating the effects of state
14 of health, state of charge, and temperature on model parameters. *Journal of Energy Storage*, 2021. **43**.
15
16 [28]. Gao, Y., et al., Co-Estimation of State-of-Charge and State-of- Health for Lithium-Ion Batteries Using an Enhanced
17 Electrochemical Model. *Ieee Transactions on Industrial Electronics*, 2022. **69**(3): p. 2684-2696.
18
19 [29]. Cen, Z. and P. Kubiak, Lithium-ion battery SOC/SOH adaptive estimation via simplified single particle model.
20 *International Journal of Energy Research*, 2020. **44**(15): p. 12444-12459.
21
22 [30]. Ma, L., et al., Co-estimation of state of charge and state of health for lithium-ion batteries based on fractional-order
23 model with multi-innovations unscented Kalman filter method. *Journal of Energy Storage*, 2022. **52**.
24
25 [31]. Liu, B., X. Tang, and F. Gao, Joint estimation of battery state-of-charge and state-of-health based on a simplified pseudo-
26 two-dimensional model. *Electrochimica Acta*, 2020. **344**.
27
28 [32]. Li, X., et al., State-of-charge estimation tolerant of battery aging based on a physics-based model and an adaptive
29 cubature Kalman filter. *Energy*, 2021. **220**.
30
31 [33]. Ling, L. and Y. Wei, State-of-Charge and State-of-Health Estimation for Lithium-Ion Batteries Based on Dual Fractional-
32 Order Extended Kalman Filter and Online Parameter Identification. *Ieee Access*, 2021. **9**: p. 47588-47602.
33
34 [34]. Wang, X., et al., A review of modeling, acquisition, and application of lithium-ion battery impedance for onboard battery
35 management. *Etransportation*, 2021. **7**.
36
37 [35]. Zhou, W., et al., Review on the battery model and SOC estimation method. 2021. **9**(9): p. 1685.
38
39 [36]. Vidal, C., et al., Machine Learning Applied to Electrified Vehicle Battery State of Charge and State of Health Estimation:
40 State-of-the-Art. *Ieee Access*, 2020. **8**: p. 52796-52814.
41
42 [37]. Zhou, Y., et al., Hybrid genetic algorithm method for efficient and robust evaluation of remaining useful life of
43 supercapacitors. *Applied Energy*, 2020. **260**.
44
45 [38]. Cheng, G., X. Wang, and Y. He, Remaining useful life and state of health prediction for lithium batteries based on
46 empirical mode decomposition and a long and short memory neural network. *Energy*, 2021. **232**.
47
48 [39]. Shang, Y., et al., A multi-fault diagnosis method based on modified Sample Entropy for lithium-ion battery strings.
49 *Journal of Power Sources*, 2020. **446**.
50
51 [40]. Li, X., C. Yuan, and Z. Wang, State of health estimation for Li-ion battery via partial incremental capacity analysis based
52 on support vector regression. *Energy*, 2020. **203**.
53
54 [41]. Chen, L., et al., Battery state-of-health estimation based on a metabolic extreme learning machine combining
55 degradation state model and error compensation. *Energy*, 2021. **215**.
56
57 [42]. Pepe, S., et al., Neural ordinary differential equations and recurrent neural networks for predicting the state of health of
58 batteries. *Journal of Energy Storage*, 2022. **50**.
59
60 [43]. Wu, J., et al., State of health estimation of lithium-ion battery with improved radial basis function neural network. *Energy*,
2023. **262**.
[44]. Chang, C., et al., Lithium-ion battery state of health estimation using the incremental capacity and wavelet neural
networks with genetic algorithm. *Journal of Energy Storage*, 2021. **38**.
[45]. Han, X.J., Z.R. Wang, and Z.X. Wei, A novel approach for health management online-monitoring of lithium-ion batteries
based on model-data fusion. *Applied Energy*, 2021. **302**.
[46]. Jiang, Y., et al., State of Health Estimation for Lithium-Ion Battery Using Empirical Degradation and Error

- 1
2 Compensation Models. *Ieee Access*, 2020. **8**: p. 123858-123868.
- 3 [47]. Mao, L., et al., A LSTM-STW and GS-LM Fusion Method for Lithium-Ion Battery RUL Prediction Based on EEMD.
4 *Energies*, 2020. **13**(9).
- 5 [48]. Hu, X., et al., Battery Health Prediction Using Fusion-Based Feature Selection and Machine Learning. *Ieee Transactions*
6 *on Transportation Electrification*, 2021. **7**(2): p. 382-398.
- 7 [49]. Khaleghi, S., et al., Online health diagnosis of lithium-ion batteries based on nonlinear autoregressive neural network.
8 *Applied Energy*, 2021. **282**.
- 9 [50]. Tan, Y. and G. Zhao, Transfer Learning With Long Short-Term Memory Network for State-of-Health Prediction of
10 Lithium-Ion Batteries. *Ieee Transactions on Industrial Electronics*, 2020. **67**(10): p. 8723-8731.
- 11 [51]. Tian, H., et al., A review of the state of health for lithium -ion batteries: Research status and suggestions. *Journal of*
12 *Cleaner Production*, 2020. **261**.
- 13 [52]. Liu, K., et al., Towards Long Lifetime Battery: AI-Based Manufacturing and Management. *Ieee-Caa Journal of*
14 *Automatica Sinica*, 2022. **9**(7): p. 1139-1165.
- 15 [53]. Ng, M.-F., et al., Predicting the state of charge and health of batteries using data-driven machine learning. *Nature*
16 *Machine Intelligence*, 2020. **2**(3): p. 161-170.
- 17 [54]. Wang, S., et al., A critical review of improved deep learning methods for the remaining useful life prediction of lithium-
18 ion batteries. *Energy Reports*, 2021. **7**: p. 5562-5574.
- 19 [55]. Sui, X., et al., A review of non-probabilistic machine learning-based state of health estimation techniques for Lithium-
20 ion battery. *Applied Energy*, 2021. **300**.
- 21 [56]. Yang, Y.R., et al., State of Health Prediction of Lithium-Ion Batteries Based on the Discharge Voltage and Temperature.
22 *Electronics*, 2021. **10**(12).
- 23 [57]. Liu, K., L. Kang, and D. Xie, Online State of Health Estimation of Lithium-Ion Batteries Based on Charging Process
24 and Long Short-Term Memory Recurrent Neural Network. *Batteries-Basel*, 2023. **9**(2).
- 25 [58]. Li, R., et al., On-Line Estimation Method of Lithium-Ion Battery Health Status Based on PSO-SVM. *Frontiers in Energy*
26 *Research*, 2021. **9**.
- 27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60