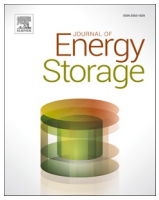
Journal of Energy Storage 99 (2024) 113376

Contents lists available at ScienceDirect

Journal of Energy Storage

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



Research papers

A novel state of health estimation method for lithium-ion battery based on forward-broad learning system

Haoyu Ming, Dong Lu, Hui Zhang, Guangfeng Wang, Dongyu Liu, Naxin Cui *

School of Control Science and Engineering, Shandong University, Jinan 250061, China

ARTICLE INFO

Keywords:
Lithium-ion battery
Health features
SOH estimation
Broad learning system

ABSTRACT

Accurate battery state estimation is crucial for optimizing performance, enhancing safety, and prolonging battery life. To improve predictive accuracy of State of Health (SOH) and enhance the generalization capability, this paper proposes a novel framework for SOH estimation based on Forward-Broad Learning System (F-BLS). Firstly, health features are extracted from the charging data, and correlation analysis is conducted to select health metrics highly correlated with battery life degradation. Secondly, in order to achieve efficient training of the model, an adaptive parameter optimization algorithm Forward-AdaBound (FAdaBound) is, integrated into the Broad Learning System (BLS) to create the F-BLS. In addition, the proposed F-BLS integrates regularization techniques to improve its generalization performance. Finally, three datasets are employed to evaluate the performance of the proposed method, which is compared it with two commonly used data-driven methods. The test result demonstrates that the proposed SOH estimation method accurately tracks the capacity degradation of the battery, with RMSE less than 0.02 and MAE less than 0.03. Importantly, the F-BLS not only achieves excellent training metrics but also sustains good prediction accuracy on unforeseen test samples, showcasing strong generalization capability.

1. Introduction

To address the depletion of fossil fuels and environmental challenges posed by global industrial development, it is imperative to focus on renewable energy [1]. Lithium-ion batteries are extensively used in electric vehicles and mobile devices due to their low pollution, recyclability, and high energy efficiency. They offer high energy density, long lifespan, rapid charging, and low self-discharge rates [2]. Lithium-ion batteries undergo internal changes as they age, including lithium ion loss, electrode material degradation, diaphragm aging, and electrolyte decomposition. These factors contribute to battery aging and a subsequent decline in cycling capacity [3,4]. State-of-health (SOH) estimation plays a crucial role in accurately predicting battery life, optimizing battery performance, ensuring secure operation, and is one of the most vital aspects of battery management. Accurate SOH estimation enables the prediction of remaining battery life, ensures the reliable operation of battery systems [5].

The complexity of internal aging mechanisms in lithium-ion batteries poses a significant challenge for precise SOH estimation [6]. Based on comprehensive research, both data-driven and model-based [7] approaches are currently employed for SOH estimation. Model-based approach enables monitoring of the internal state of lithium-ion batteries and accurate assessment of battery condition [8]. Model-based approaches, including equivalent circuit modeling (ECM) and electrochemical modeling (EM), are widely used. Li et al. [9] proposed an improved ECM, including an additional capacitor, to estimate the SOH of lithium-ion batteries. This method significantly enhances SOH estimation accuracy by extracting health features from electrochemical impedance spectroscopy. Gao et al. [10] proposed a reduced-order EM that simplifies the full-order pseudo-two-dimensional model using the Padé approximation method. This approach enhances the accuracy of state estimation. Lyu et al. [11] developed a Particle Filtering (PF) framework combined with EM. This framework defines the model parameters degraded with the battery performance as state variables, achieving higher estimation accuracy compared to conventional PF. Zhu et al. [12] proposed a hybrid approach that combines multiple filtering methods, not only improving the estimation accuracy of SOH but also enhancing the estimation speed, enabling fast and accurate prediction. Li et al. [13] introduced a state observer based on the extended single-particle model. The observer utilizes reduced-order electromagnetism

* Corresponding author.
E-mail address: cuinx@sdu.edu.cn (N. Cui).

https://doi.org/10.1016/j.est.2024.113376
Received 16 April 2024; Received in revised form 22 July 2024; Accepted 13 August 2024
Available online 24 August 2024

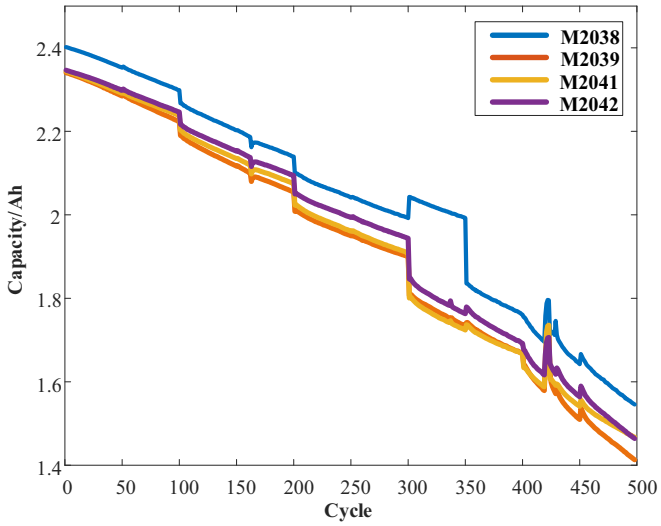


Fig. 1. Li-ion battery cyclic aging experiment result.

to monitor real-time state changes in the battery, thereby achieving accurate SOH estimation. By modeling batteries and simulating their behavior under different conditions, it is possible to identify aging characteristics and achieve SOH estimation. However, model-based methods encounter challenges such as parameter uncertainty, model complexity, extended computation times, and limited dynamic adaptability, etc.

Data-driven approaches demonstrate greater flexibility and adaptability across diverse battery types and usage conditions compared to model-based approaches [24]. This method enables accurate identification of performance decline trends and estimation of SOH through the analysis of battery cycling data. Currently, predominant methods used for SOH estimation include neural networks [14–16], Random Forests (RF) [17], Long Short-Term Memory (LSTM) [18,41], and Gaussian Process Regression (GPR) [19], etc. Among them, Fan et al. [20] proposed a combination of gated recurrent units and convolutional neural networks. This combination effectively captures spatio-temporal features in charging data by learning battery charging curves, resulting in enhanced SOH prediction accuracy. Zhang et al. [21] introduced a hybrid model that integrates RF, the artificial bee colony algorithm (ABC), and the generalized regression neural network (GRNN). This model performs feature selection through RF and refines GRNN parameters using the ABC algorithm, achieving high-precision estimation of SOH even before the battery undergoes significant aging. Deep neural networks, as a subset of neural networks, have found wide-ranging applications. Ma et al. [22] integrated enhanced LSTM with the differential evolutionary gray wolf optimizer for hyper-parameter optimization, enhancing the accuracy of predictions. Peng et al. [23] developed a battery SOH estimation method based on multi-health features extraction and an improved LSTM, which reduces the error caused by a single health feature and improves the estimation accuracy. However, deep neural networks face issues such as overfitting, demanding extensive time and memory for training, and being susceptible to gradient vanishing and explosion. The overfitting problem results in reduced generalization ability, while the gradient problem leads to instability in model training and affects the reliability of predictions. These issues negatively impact the accurate estimation of SOH.

In recent years, the Broad Learning System (BLS) has emerged as an effective method [25], achieving excellent results in SOH estimation. Gu et al. [26] proposed a Polak-Ribière-Polyak conjugate gradient algorithm optimized broad learning system (BLS) for accurate lithium-ion battery SOH estimation, demonstrating high accuracy with a mean absolute error below 1 % and improved model generalization through enhanced training data. Chen et al. [31] proposed a hybrid algorithm combining the BLS and relevance vector machine for accurately estimating the SOH of lithium-ion batteries, which solves the problems of capacity regeneration and stochasticity. Despite the research that has been conducted to improve and optimize BLS, the mainstream BLS method still relies on ridge regression. While this approach effectively addresses vanishing and exploding gradients, its limitations in prediction accuracy and generalization performance constrain its widespread adoption. To address these issues, this paper introduces the Forward-Broad Learning System (F-BLS) and proposes a SOH estimation framework aimed at enhancing the accuracy of SOH estimation.

The primary contributions and innovations of this paper are outlined below:

- A novel adaptive optimization algorithm, Forward-AdaBound (FAdaBound), based on AdaBound and forward automatic differentiation, is proposed. This algorithm enables faster and more efficient handling and analysis of battery data without inverse matrix calculations. By eliminating the need for backpropagation, FAdaBound simplifies the optimization process and enhances the robustness of the SOH estimation model. Consequently, it provides more accurate and reliable SOH predictions.
- A novel SOH estimation method based on F-BLS is proposed. This method integrates the FAdaBound optimization algorithm and regularization techniques into BLS, representing a significant advancement in SOH estimation for various battery types and usage conditions. This improvement optimizes BLS efficiently and accurately, reducing overfitting risk and ensuring a more precise assessment of battery SOH.
- The proposed method has been validated on multiple datasets, showcasing consistent high estimation accuracy and generalizability across various types of batteries.

The remainder of the paper is organized as follows: Section II conducts battery cycle aging experiments to process the acquired data along with data from available datasets. Health features are then extracted, followed by correlation analysis. Section III explains the F-BLS as well as the optimization algorithm proposed in this paper and presents the SOH estimation steps. In Section IV, the performance of F-BLS is validated and two publicly available datasets are used to compare the efficacy of F-BLS with other methods.

2. Battery data processing

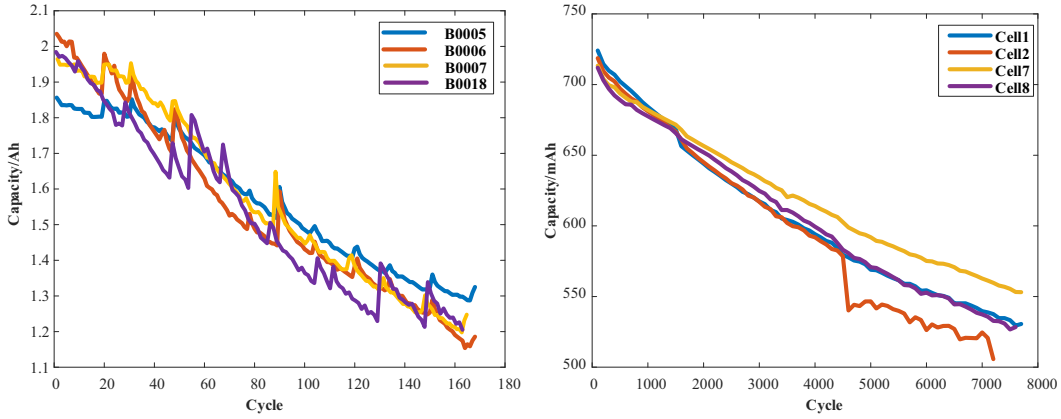
In this section, experimental datasets are utilized to extract health features. The extracted health features are further analyzed using Pearson correlation analysis.

2.1. Cyclic aging experiment

The experimental platform comprises the battery test instrument ARBIN BT-5HC, a temperature chamber, and a host computer. Four LR18650SZ lithium-ion batteries, labeled as M2038, M2039, M2041, and M2042, are utilized in this experiment to conduct the cyclic aging experiment. The temperature was set to 25 °C in the temperature chamber, and the experimental program was configured and experimental data were stored using the host computer. The experimental program was configured for constant current and constant voltage (CC-CV) charging mode. It charged at 1C until reaching 4.2 V during the CC phase, and then transitioned to CV charging mode until the current decreased to 0.05 A. After 1 h of resting, the batteries were discharged at a 1C until the voltage was reduced to 3.0 V. Each battery was subjected to 500 charge/discharge cycles.

The aging data are presented in Fig. 1.

In addition to the cyclic aging dataset obtained from above experiments, this paper also utilizes NASA battery aging dataset [33] and Oxford University battery aging dataset. NASA Battery aging dataset



(a) (b)

Fig. 2. Li-ion battery aging dataset: (a) Capacity of NASA dataset. (b) Capacity of Oxford dataset.

Table 1
Comparison of the battery aging dataset.

Battery number	Capacity/Ah	Q range/Ah	Voltage/V	Cutoff voltage/V
B05	2.03	2.03 → 0.80	4.2	2.7
B06	1.85	1.85 → 0.80	4.2	2.5
B07	1.95	1.95 → 0.80	4.2	2.2
B018	1.98	1.98 → 0.80	4.2	2.2
Cell1	0.74	0.74 → 0.43	4.2	2.7
Cell2	0.73	0.73 → 0.43	4.2	2.7
Cell7	0.72	0.72 → 0.43	4.2	2.7
Cell8	0.72	0.72 → 0.43	4.2	2.7
M2038	2.4	2.4 → 1.4	4.2	3.0
M2039	2.35	2.35 → 1.4	4.2	3.0
M2041	2.35	2.35 → 1.4	4.2	3.0
M2041	2.35	2.35 → 1.4	4.2	3.0

cycled with LG Chem 18,650 batteries. These batteries have a rated capacity of 2.1 Ah and an operating voltage range of 3.2 V–4.2 V. The aging data were obtained through CC-CV charge/discharge cycles at 24 °C. B05, B06, B07 and B18 were selected as the research objects. The Oxford University battery aging dataset consists of aging data from eight lithium cobaltate pouch batteries. The battery has a rated capacity of 740mAh and the eight cells are labeled Cell1, Cell2, …, Cell8, respectively. The aging data of this dataset was obtained by CC charging at 40 °C and discharging it under simulated urban driving conditions. The capacity was calibrated with a 1C current discharge every 100 charge/discharge cycles. Cell1, Cell 2, Cell 7, and Cell 8 were selected for the study, and the aging data are shown in Fig. 2.

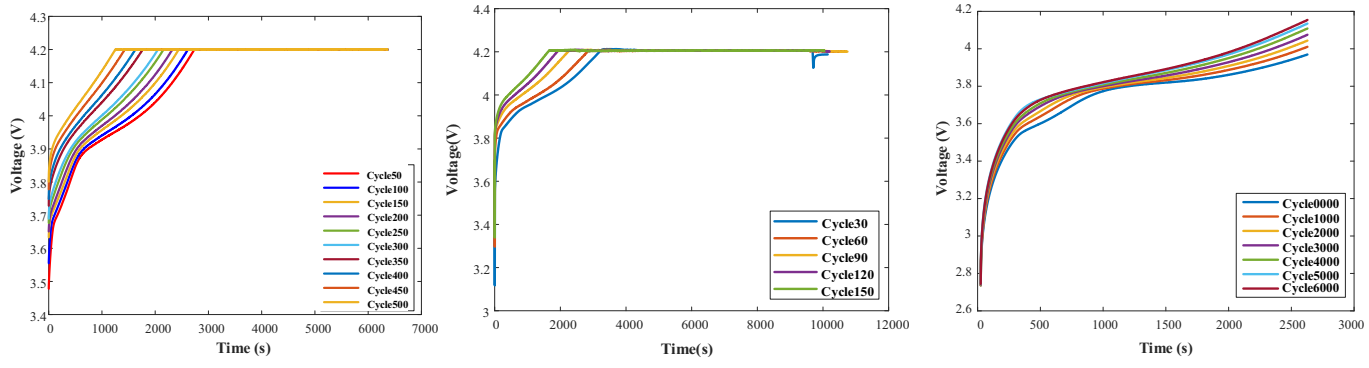
The electrochemical characteristics and experimental conditions of the batteries used in the three datasets were different, which are presented in Table 1.

2.2. Health feature extraction

The charge and discharge curves of a battery change as it ages. In practical scenarios, the discharge curve shape varies due to uncertain working conditions. In contrast, charging conditions are typically fixed, often using the CC-CV charging mode. The charging curve is relatively smoother compared to the discharge curve, making feature extraction easier. The dataset comprises information on charging time, current, voltage, temperature, and capacity. This serves as a crucial source of information for estimating the SOH of batteries. The voltage curves for batteries undergoing CC charging at various aging levels often exhibit similar and flat characteristics, which makes it difficult to identify the degree of battery aging. Consequently, in addition to the health features obtained through direct measurement, Incremental Capacity Analysis (ICA) is employed as a more effective approach to handle charging data [32]. The ICA method can transform seemingly similar and smooth voltage curves into a series of Incremental Capacity (IC) curves that exhibit peaks and valleys, which makes it convenient to identify batteries with different levels of aging, and the health features are obtained from IC curves. In this paper, the SOH can be expressed as:

$$SOH = \frac{Q_{now}}{Q_{new}} \times 100\% \tag{1}$$

In this equation, Q_{new} represents the nominal capacity of the battery



(a) (b) (c)

Fig. 3. Charging voltage curves of different datasets: (a) Charging voltage curves of cyclic aging experiment. (b) Charging voltage curves of NASA. (c) Charging voltage curves of Oxford.

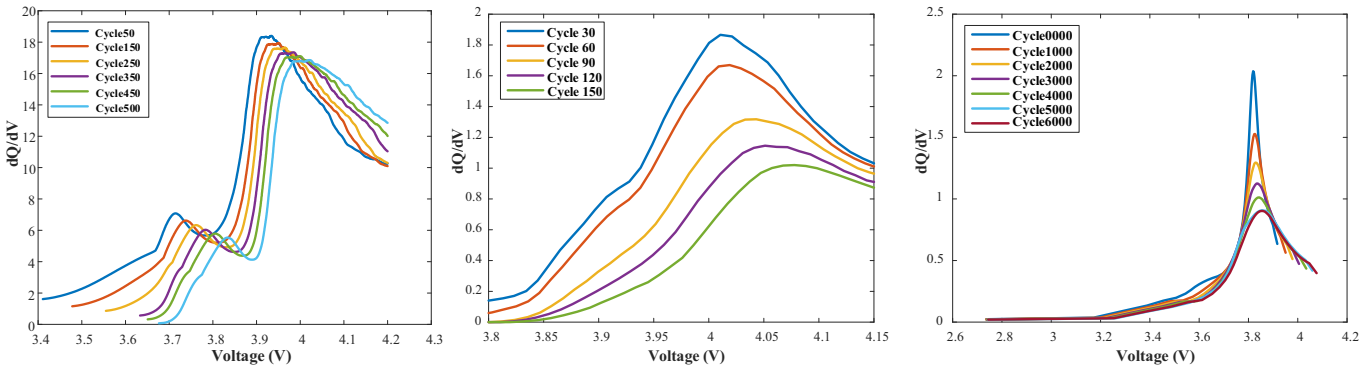


Fig. 4. IC curves: (a) IC curves of cycling aging experiment. (b) IC curves of NASA dataset. (c) IC curves of Oxford dataset.

Table 2
HF extracted from charging curves.

HF name	Illustration of HF	Type
EVD	The time required for the same interval of voltage change	Charging time
t_{cc}	CC charging time	Charging time
ICP	peak of the IC curve	IC curve
ICPL	IC curve peak location	IC curve

Table 3
The result of Pearson correlation analysis.

HF	*EVD*	t_{cc}	*ICP*	*ICPL*
Coefficient	0.9987	0.9992	−0.9558	−0.9622

and Q_{now} represents the capacity obtained by measurement after aging.

2.2.1. HF related to charging time

Changes in charging time can indicate alterations in battery capacity, and charging time features are easily measurable. Partial charging and discharging modes are often applied to batteries because full charging or discharging takes a long time and negatively impacts battery lifespan. Partial charge/discharge data still contain crucial information about the battery's health, allowing the extraction of health features without losing the key characteristics of a complete cycle.

The charging voltage curves for batteries at various aging stages are depicted in Fig. 3. With battery aging, the charging curve gradually shifts to the left, and the time needed for the voltage to reach the same value decreases. This parameter provides an effective reflection of capacity decline. *EVD* is defined as the time required for the same voltage change interval, the definition equation is as follows:

$$\Delta V = V_{i+1} - V_i \tag{2}$$

$$EVD = |t_{i+1} - t_i| \tag{3}$$

where t_{i+1} is the time when the voltage is V_{i+1}, and t_i is the time when the voltage is V_i. The charging curves of the three different types of batteries show the most significant variation in the voltage range of 3.55 V–4 V. Therefore, in this study, the time required to charge from 3.55 V to 4 V is selected as the health feature *EVD*, as this voltage range is applicable to all three types of batteries. In this paper, V_i is selected as 3.55 V and V_{i+1} as 4 V.

Capacity decline is a common indicator of battery aging, resulting in a reduction in the maximum energy that can be stored. Consequently, the time to reach the preset voltage decreases, shortening the CC charging time (t_{cc}). This feature directly indicates changes in the battery's capacity, is easy to measure, and allows for real-time SOH monitoring.

Since *EVD* and t_{cc} can better indicate the capacity decline and the features are easy to obtain. Therefore, in this paper, *EVD* and t_{cc} are chosen as health features related to charging time.

2.2.2. HF related to incremental capacity curves

The incremental capacity curve is shown in Fig. 4. By using the ICA technique, it is possible to observe the changes in the IC curve at different aging levels [35]. The ICA calculation formula is shown below:

$$\frac{dQ}{dV} = \frac{\Delta Q}{\Delta V} = \frac{I_c \cdot dt}{dV} \tag{4}$$

$$Q = \int I_c dt \tag{5}$$

where, Q represents the current capacity, dV/dt is the rate of voltage change, and I_c is the charging current.

During the aging process, the IC curves under different aging states have different locations with different sizes of wave peaks, the peak of the IC curve (*ICP*) rises, the location of the peak (*ICPL*) increases, and the IC curve shifts downward gradually. These changes can be used to reflect the battery capacity decline. Therefore, the *ICP* and the *ICPL* are selected as the health features. The health features selected in this paper are detailed in Table 2.

2.3. Pearson correlation analysis method

Health features correlated with SOH are selected as model inputs to reduce model complexity and improve prediction performance. In this paper, Pearson correlation analysis is employed for a rapid examination of multiple health features [34]. The correlation coefficient between the two variables $a = (a_1, a_2, \ldots, a_n)$ and $b = (b_1, b_2, \ldots, b_n)$ is shown in Eq. (6):

$$r = \frac{\sum_{i=1}^{n}(a_i - \overline{a})(b_i - \overline{b})}{\sqrt{\sum_{i=1}^{n}(a_i - \overline{a})^2}\sqrt{\sum_{i=1}^{n}(b_i - \overline{b})^2}} \tag{6}$$

Among them, a_i and b_i represent the i-th observation values for two variables, $\overline{a}$ and $\overline{b}$ is the mean of two variables, n represents the number of observations.

The correlation analysis was performed on the selected health features in this section, four parameters, *EVD*, t_{cc}, *ICP*, *ICPL* were selected as the health features. The correlation result is presented in Table 3.

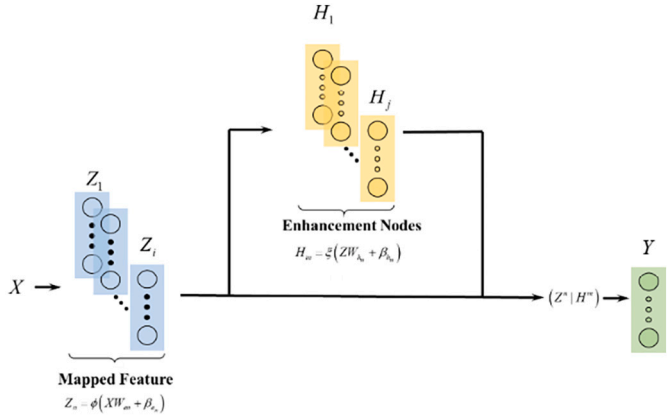


Fig. 5. Broad learning system network structure diagram.

3. Methodology

This section begins with an exposition of the basic BLS. On the basis of BLS, in order to obtain a satisfactory SOH estimation method, this paper proposes an F-BLS model integrating the regularization technique and trains the model using FAdaBound, an adaptive optimization algorithm based on AdaBound [36] and forward auto-differentiation.

3.1. Broad learning system

BLS is a neural network constructed based on the traditional Random Vector Functional Link Neural Network (RVFLNN). RVFLNN excels in overcoming long training times and showcasing high generalization ability during function approximation. It directly employs input data to construct the enhancement nodes. Building on this, BLS initially maps the input data to a set of features and then utilizes both the input data and the mapped features to construct the enhancement nodes. These adjustments enable BLS to perform more intricate feature extraction and data dimensionality reduction, thereby bolstering the model's flexibility and generalization ability. The network structure of BLS is illustrated in Fig. 5.

In BLS, it is assumed that the input data X has N samples, each with dimension M, and $Y \in \mathbb{R}^{N\times C}$ is the output matrix. For n feature mappings, each mapping generates k mapping nodes, and the mapping features can be expressed as Eq. (7):

$$Z_i = \phi_i\left(XW_{e_i} + \beta_{e_i}\right), \quad i = 1, 2, \ldots, n \tag{7}$$

where W_{ei}, β_{ei} are randomly generated weights and biases with appropriate dimensions. The mapped features form a feature mapping layer $Z \equiv [Z_1, Z_2 \ldots, Z_i]$, containing i sets of mapped features. These features Z_i are mapped by a nonlinear transformation to generate enhancement nodes H_j:

$$H_j = \xi_j\left(Z_i W_{hj} + \beta_{hj}\right), \quad j = 1, 2, \ldots, m \tag{8}$$

where W_{hj} and β_{hj} are the weights and biases associated with the enhancement nodes. The enhancement nodes are combined into an enhancement node layer $H \equiv \left[H_1, H_2, \ldots, H_j\right]$, including j groups of enhancement nodes. i and j can be selected based on the specific modeling task and the complexity of the problem. Different numbers of mapping features and groups of augmentation nodes can be flexibly selected as needed to accommodate different data and tasks.

Thus, the BLS model can be represented in the following form:

$$\begin{aligned} Y &= [Z_1, \ldots, Z_n | H_1, \ldots, H_m] W_m \\ &= [Z|H] W_m \\ &= AW_m \end{aligned} \tag{9}$$

where $W_m = [Z_n|H_m]^+ Y$ is the connection weight of the output layer, and the broad network connection weights are obtained by solving the pseudo-inverse using ridge regression approximation [25]. The solution formula is as follows:

$$A^+ = \lim_{\lambda \to 0}\left(\lambda I + AA^T\right)^{-1} A^T \tag{10}$$

3.2. Forward-broad learning system

BLS is normally optimized using ridge regression[27–30], which still suffers from the problem of easy overfitting and unsatisfactory accuracy. The model proposed in this paper integrates the forward auto-differentiation technique and no longer needs the additional back-propagation computational process. In addition, a parameter update strategy [36] based on FAdaBound is applied to the proposed F-BLS during the forward propagation process. The F-BLS consists of three main parts, the base BLS, the regularized loss evaluation, and the forward gradient evaluation. The structure of the proposed F-BLS is shown in Fig. 6.

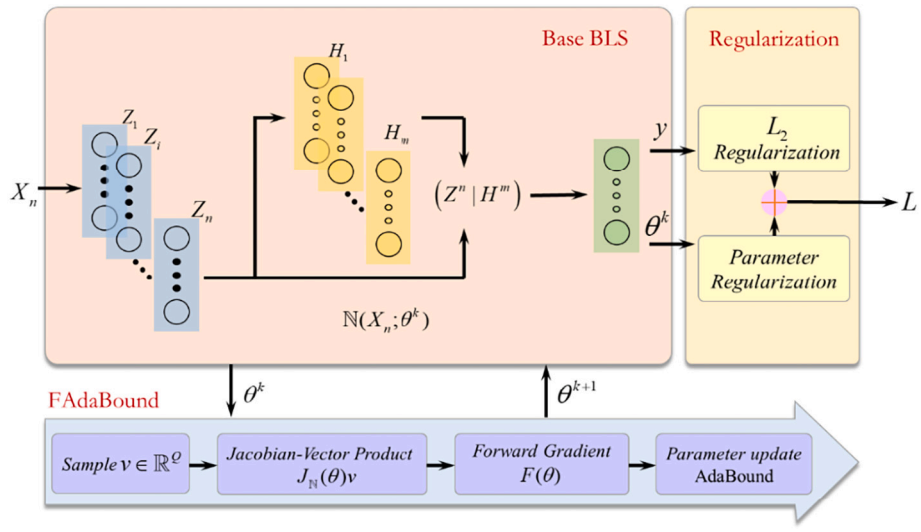


Fig. 6. The structure of the proposed F-BLS.

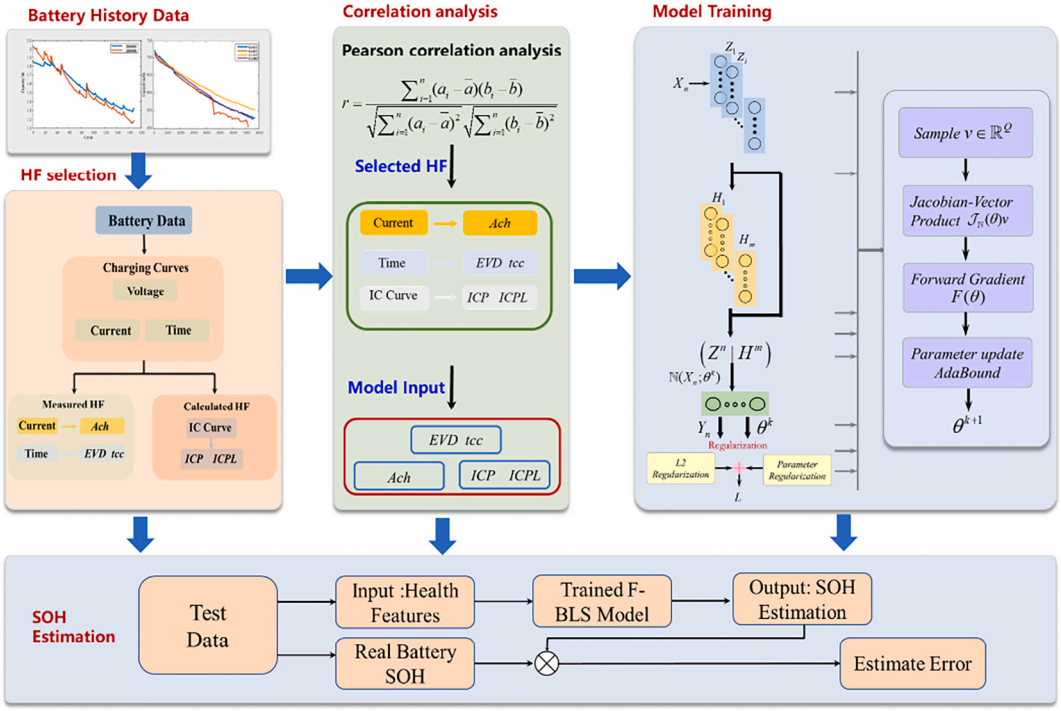


Fig. 7. Flow chart for SOH estimation.

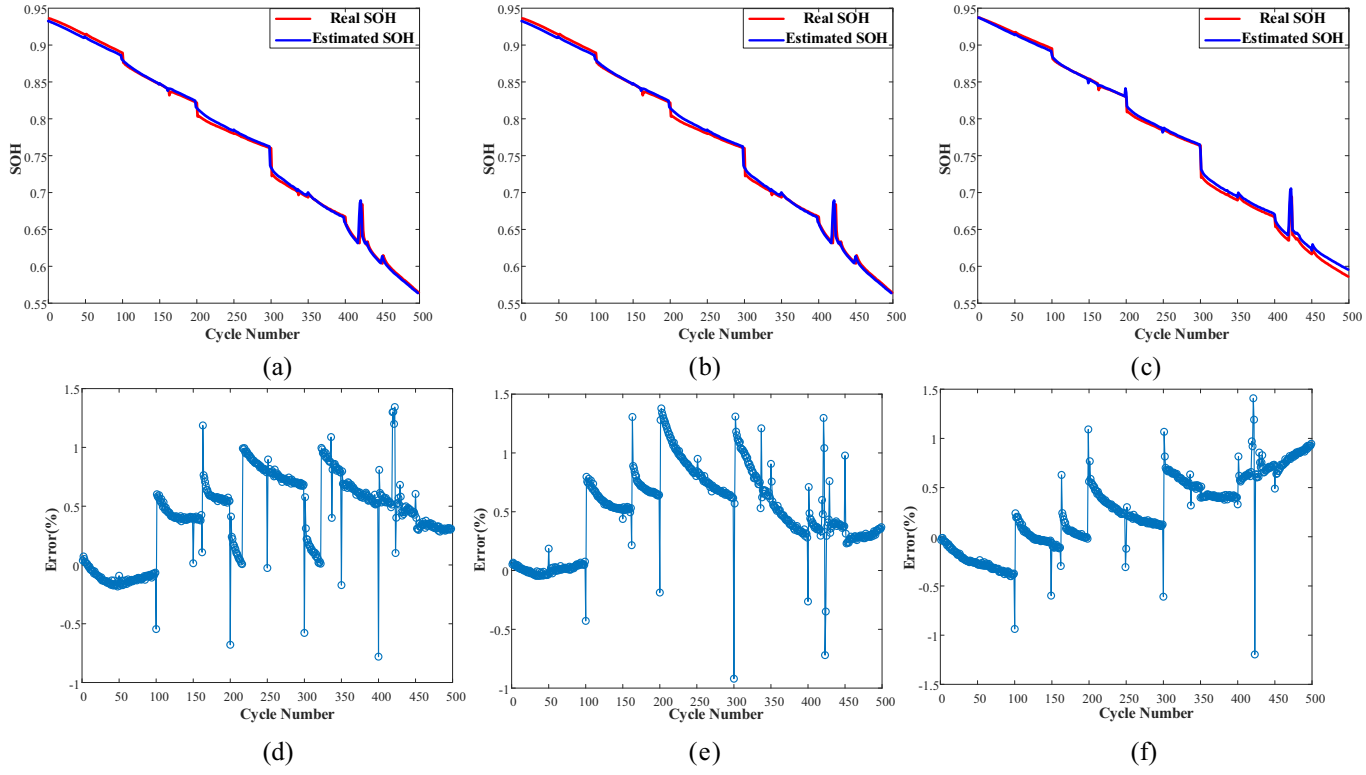


Fig. 8. The SOH estimation results and error: (a) SOH estimation result of 2#. (b) SOH estimation result of 3#. (c) SOH estimation result of 4 #. (d) SOH estimation error of 2#. (e) SOH estimation error of 3#. (f) SOH estimation error of 4#.

3.2.1. *Base BLS*

Assuming that the proposed F-BLS mathematical model can be represented as $\mathbb{N}(\cdot;\theta)$, Then the mapping relationship between its input and output can be expressed as:

$$y = \mathbb{N}(x;\theta) \tag{11}$$

In the equation, θ is the model parameter vector for BLS, $\theta \in \mathbb{R}^Q$, where θ includes the weight parameters W and bias parameters β for the feature mapping layer, enhanced node layer, and output layer.

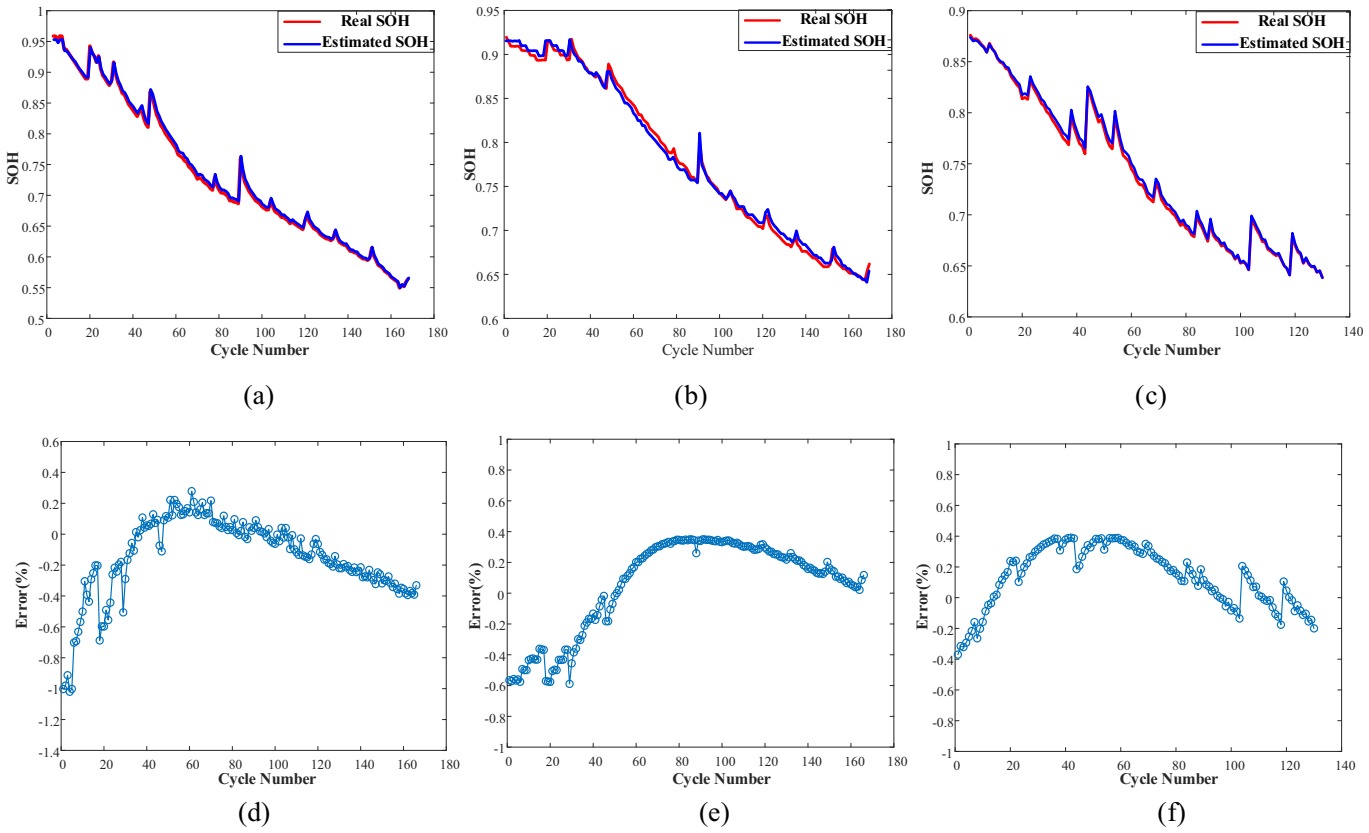


Fig. 9. The SOH estimation results and error: (a) SOH estimation result of 6#. (b) SOH estimation result of 7#. (c) SOH estimation result of 8#. (d) SOH estimation error of 6#. (e) SOH estimation error of 7#. (f) SOH estimation error of 8#.

3.2.2. *Regularization*

In machine learning, regularization techniques are often used to reduce the risk of overfitting [37] and improve model generalization. In this paper, the *L*2 regularization loss term is embedded in the proposed F-BLS:

$$L = \frac{1}{2}\sum_{n=1}^{N}\left(y_n - \mathbb{N}(x_n;\theta)\right)^2 + \frac{1}{\lambda}\sum_{q}^{Q}\theta_q^2 \tag{12}$$

where N represents the number of training samples, λ represents the regularization coefficient, and Q represents the number of model parameters.

3.2.3. *FAdaBound*

In this paper, we propose the FAdaBound algorithm for forward gradient evaluation and parameter adaptive updating. In fact, nearly all current gradient-based optimization algorithms require backpropagation to compute the partial derivatives of the losses with respect to the parameters. The backward mode of automatic differentiation techniques is widely successful in deep learning and is often used to compute Jacobian matrices [38], thus enabling fast backpropagation.

Forward propagation enables simultaneous estimation of parameter gradients during computation, facilitating parameter updates. This approach offers superior parameter optimization efficiency compared to backpropagation. As shown in Fig. 6, the FAdaBound can evaluate the gradient of the parameters $\partial L/\partial\theta_q$ while the F-BLS forward propagation is computed $\mathbb{N}(x;\theta)$.

Firstly, FAdaBound samples a perturbation vector while the input vector X_n is accepted by Base BLS:

$$v \sim \mathcal{N}(0,\mathrm{I}), v \in \mathbb{R}^Q \tag{13}$$

where $\mathcal{N}(0,\mathrm{I})$ denotes a normal distribution and the scalar components v_i of v are independent of each other and have a mean and statistics of 0 for all i.

Secondly, the computation of intermediate variables (including Z, H, etc.) and partial derivatives was then performed simultaneously at each step of the F-BLS forward propagation. To improve optimization efficiency beyond backpropagation, only one forward auto-differentiation is performed per iteration of the optimization, instead of Q times. In this process, the Jacobi matrix is not given exactly, but the Jacobi vector product is obtained by direct computation:

$$\mathcal{J}_{\mathbb{N}}(\theta)v = \left[\frac{\partial L}{\partial\theta_1},\ldots,\frac{\partial L}{\partial\theta_Q}\right]\begin{bmatrix} v_1 \\ \ldots \\ v_Q \end{bmatrix} \tag{14}$$

The partial derivatives can be taken as the directional derivative along a given perturbation vector v:

$$\Delta L(\theta)v = \sum_{q=1}^{Q}\frac{\partial L}{\partial\theta_q}v_q \tag{15}$$

where $\Delta L(\theta)v$ is a scalar.

Finally, the directional derivative $\Delta L(\theta)v$ is multiplied with the perturbation vector to obtain the forward gradient:

$$F(\theta) = \Delta L(\theta)v\cdot v = \begin{bmatrix} \frac{\partial L}{\partial\theta_1}v_1^2 + \frac{\partial L}{\partial\theta_2}v_1v_2 + \ldots + \frac{\partial L}{\partial\theta_Q}v_1v_Q \\ \frac{\partial L}{\partial\theta_1}v_1v_2 + \frac{\partial L}{\partial\theta_2}{v_2}^2 + \ldots + \frac{\partial L}{\partial\theta_Q}v_2v_Q \\ \cdots \\ \frac{\partial L}{\partial\theta_1}v_1v_Q + \frac{\partial L}{\partial\theta_2}v_2v_Q + \ldots + \frac{\partial L}{\partial\theta_Q}{v_Q}^2 \end{bmatrix} \tag{16}$$

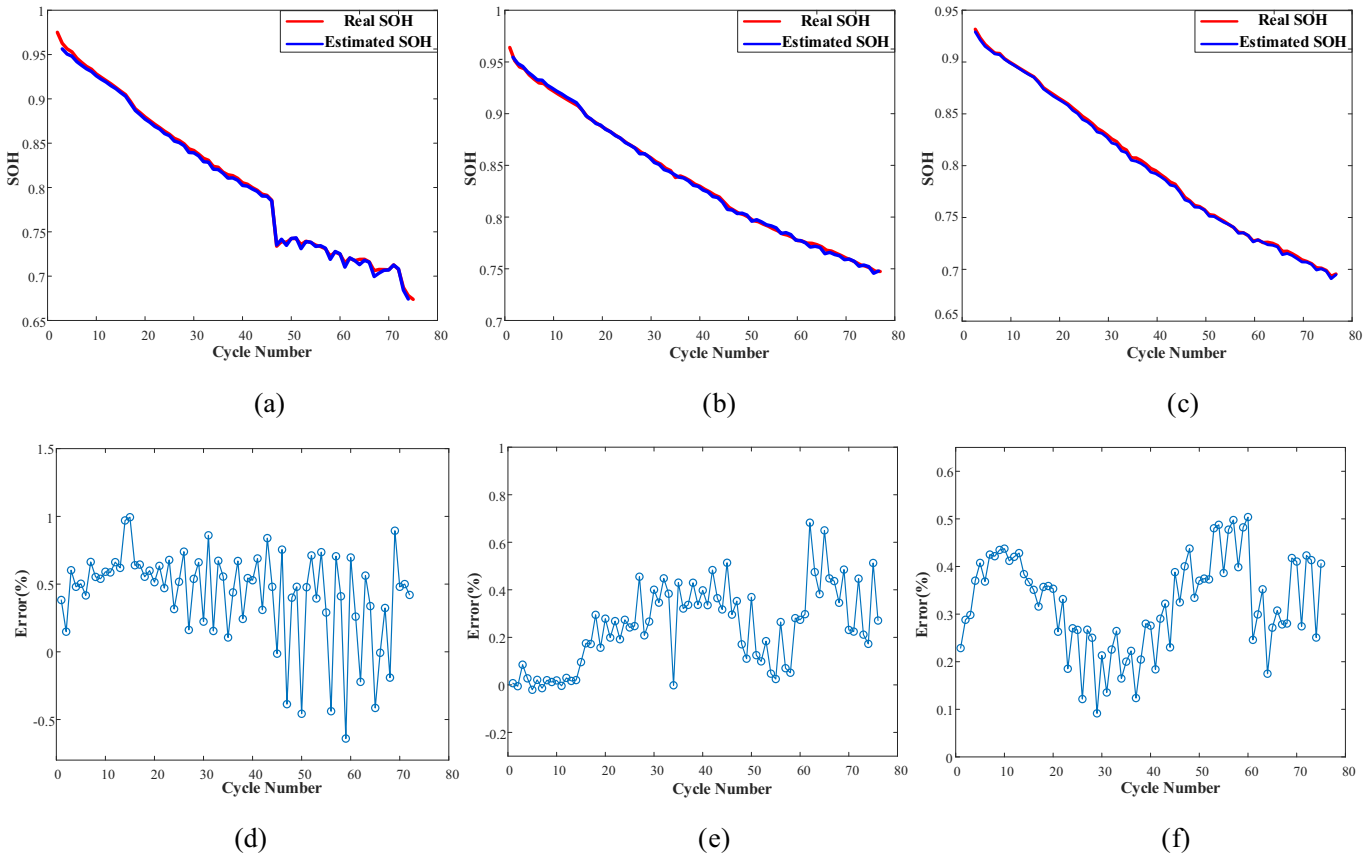


Fig. 10. The SOH estimation results and error: (a) SOH estimation result of 10#. (b) SOH estimation result of 11#. (c) SOH estimation result of 12#. (d) SOH estimation error of 10#. (e) SOH estimation error of 11#. (f) SOH estimation error of 12#.

where, $F(\theta) \in \mathbb{R}^Q$.

The adaptive learning strategy [36] is integrated into our proposed forward optimization algorithm FAdaBound for updating the model parameters of F-BLS. The proposed FAdaBound can compute an adaptive learning rate for each model parameter $\theta_q \Delta L(\theta) v$ from the forward gradient estimates at the first and second moments, $q \in [1, 2, \ldots, Q]$. The pseudo-code of the FAdaBound algorithm is shown in Algorithm 1.

Algorithm 1. Train a F-BLS by FAdaBound.

Require:
θ_0: Initial parameter vector of F-BLS;
α: Initial learning rate of FAdaBound;
$\beta_1, \beta_1 \in [0,1)$: Exponential decay rates of moment estimates in FAdaBound;
ϵ: A small positive number in FAdaBound;
$\ell(t)$, $u(t)$: The lower and upper bound functions in FAdaBound;
Initialize:
$\boldsymbol{m}_0=0$, $\boldsymbol{v}_0=0$;
$t \leftarrow 0$;
Compute:
while θ_t not converged **do**
  $t \leftarrow t+1$, gain a input vector x_t;
  $v_t \sim \mathcal{N}(\boldsymbol{\theta}, \boldsymbol{I})$; ▷Sample perturbation
  Note: the following computes Y_n and $F(\theta_t)$ simultaneously and without having to compute
  ∇L in F-BLS forward process;
  $Y_n, D(\theta_t) \leftarrow \mathrm{N}(x_t; \theta_t), \nabla L(\theta_t)\, v_t$; ▷Forward AD
  $F(\theta_t) \leftarrow D(\theta_t) v_t$;
  $\boldsymbol{m}_t \leftarrow \beta_1 \boldsymbol{m}_{t-1} + (1-\beta_1) F(\theta_t)$;
  $\boldsymbol{v}_t \leftarrow \beta_2 \boldsymbol{v}_{t-1} + (1-\beta_2)\, F(\theta_t)^2$;
  $\hat{\boldsymbol{m}}_t \leftarrow \frac{\boldsymbol{m}_t}{1-\beta_1^t}$; $\hat{\boldsymbol{v}}_t \leftarrow \frac{\boldsymbol{v}_t}{1-\beta_2^t}$;
  $\hat{\alpha} \leftarrow max\left[\ell(t), min\left(u(t), \frac{\alpha}{\sqrt{\hat{\boldsymbol{v}}_t}+\epsilon}\right)\right]$
  $\theta_{t+1} \leftarrow \theta_t - \hat{\alpha} \odot \hat{\boldsymbol{m}}_t$;
end while
Return θ_{t+1}

3.3. SOH estimation

In this section, SOH estimation is performed based on the method of F-BLS proposed in Section 3.2. The estimation framework is shown in Fig. 7.

The SOH estimation framework is divided into three parts: feature extraction, correlation analysis, and SOH estimation based on F-BLS.

Step I: Extract four HFs related to current, charging time and capacity increment from the Experimental dataset, NASA battery dataset and the Oxford dataset.

Step II: The four HFs were analyzed by Pearson's correlation analysis, the selected HFs were all highly correlated with SOH, namely *EVD*, t_{cc}, *ICP* and *ICPL*, were screened as inputs to F-BLS.

Step III: F-BLS initialization includes setting the number of feature mapping nodes and enhancement nodes. The input features are mapped using the feature mapping nodes to generate the feature mapping layer. The feature mapping layer is used as the input of the enhancement nodes and the enhancement nodes nonlinearly transform the feature mapping layer to obtain the enhancement node layer. The feature mapping layer are combined with the enhancement node layer into the input matrix. The weights of F-BLS are initialized, and the FAdaBound algorithm is employed for forward gradient evaluation and parameter adaptive updating. The weights of F-BLS are obtained through training data. After completing the training of F-BLS on the training set, SOH estimation is performed based on the trained model and validate the proposed SOH estimation method's performance using the test set.

4. Results and discussions

This section provides thorough comparative tests, accompanied by detailed results and analysis. Firstly, a brief description of the

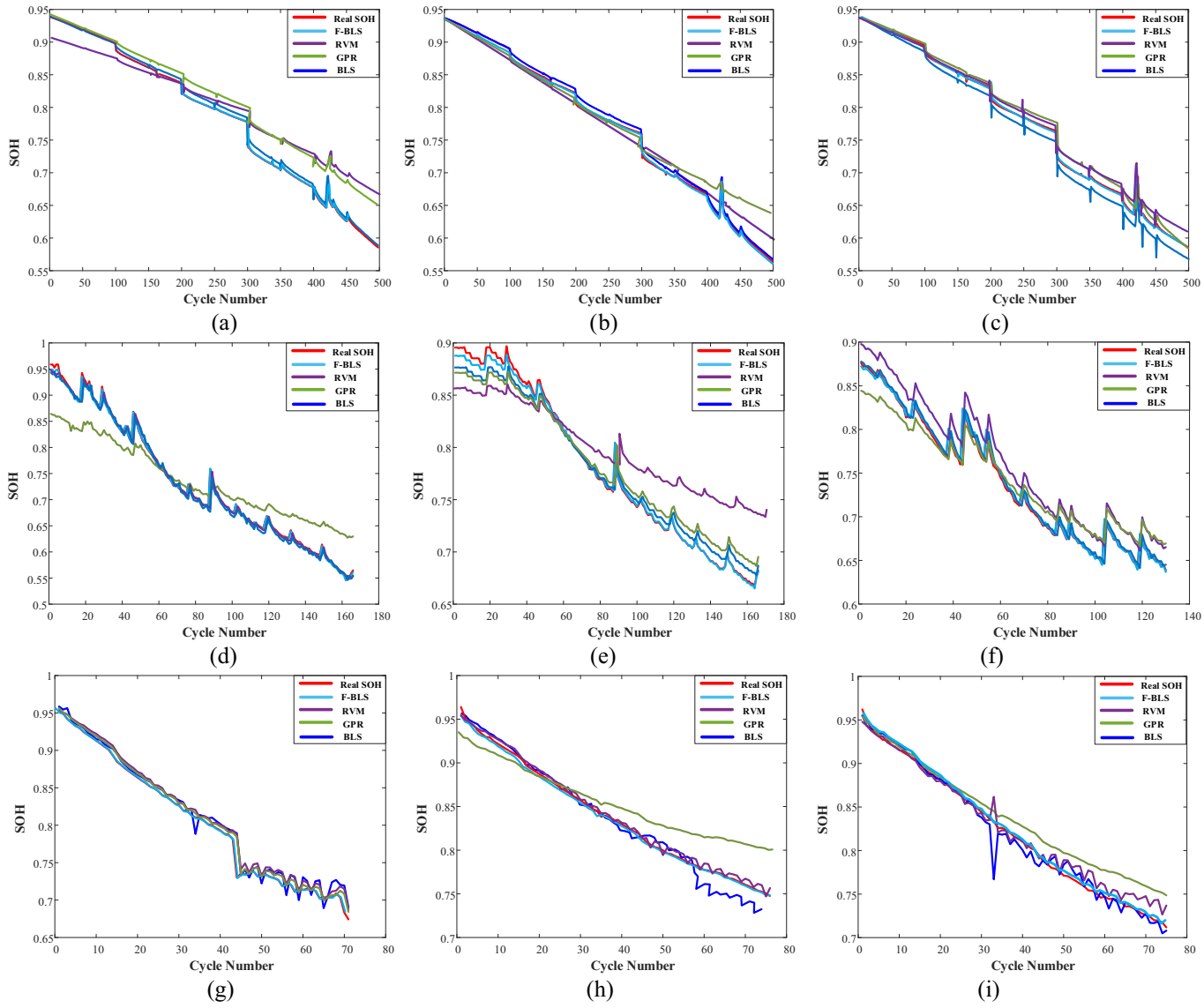


Fig. 11. The SOH estimation compared with different methods: (a) SOH estimation results of 2#, (b) SOH estimation results of 3#, (c) SOH estimation results of 4#, (d) SOH estimation of 6#. (e) SOH estimation results of 7#, (f) SOH estimation results of 8#, (g) SOH estimation results of 10#, (h) SOH estimation of 11#. (i) SOH estimation of 12#.

performance metrics used to evaluate the different methods will be presented. Subsequently, the performance of F-BLS is evaluated, and error analysis is performed using multiple batteries from three datasets. Lastly, the comparison methods, RVM [39], and GPR [40], are introduced.

4.1. Evaluation criteria

To assess the estimation accuracy of the algorithm for SOH, this paper employs mean absolute error (MAE), root mean square error (RMSE), and R^2 (R-square) to quantify the estimation error. The formula is shown below:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|C_i - \widehat{C}_i| \tag{17}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(C_i - \widehat{C}_i)^2} \tag{18}$$

$$R^2 = 1 - \frac{\sum_i (\widehat{C}_i - C_i)^2}{\sum_i (\overline{C}_i - \widehat{C}_i)^2} \times 100\% \tag{19}$$

In this equation, n represents the size of the sample, C_i represents the true capacity of the i-th cycle, $\widehat{C}_i$ is the actual capacity, and $\overline{C}$ is the average capacity. The diminution of MAE and RMSE corresponds to a reduction in the estimation error of SOH, indicating that the estimation method performs well. Additionally, the greater the magnitude of the coefficient of R^2, the heightened the degree of fit exhibited by the model, suggesting a more precise estimation approach.

4.2. Estimation results

In this section, the proposed SOH estimation method is tested on three datasets. The batteries labeled M2038, M2039, M2041, and M2042 are denoted as 1#, 2#, 3#, and 4# respectively. The batteries from the NASA dataset B05, B06, B07 and B018 are labeled as 5#, 6#, 7#, and 8#, while those from the Oxford dataset Cell1, Cell2, Cell3 and

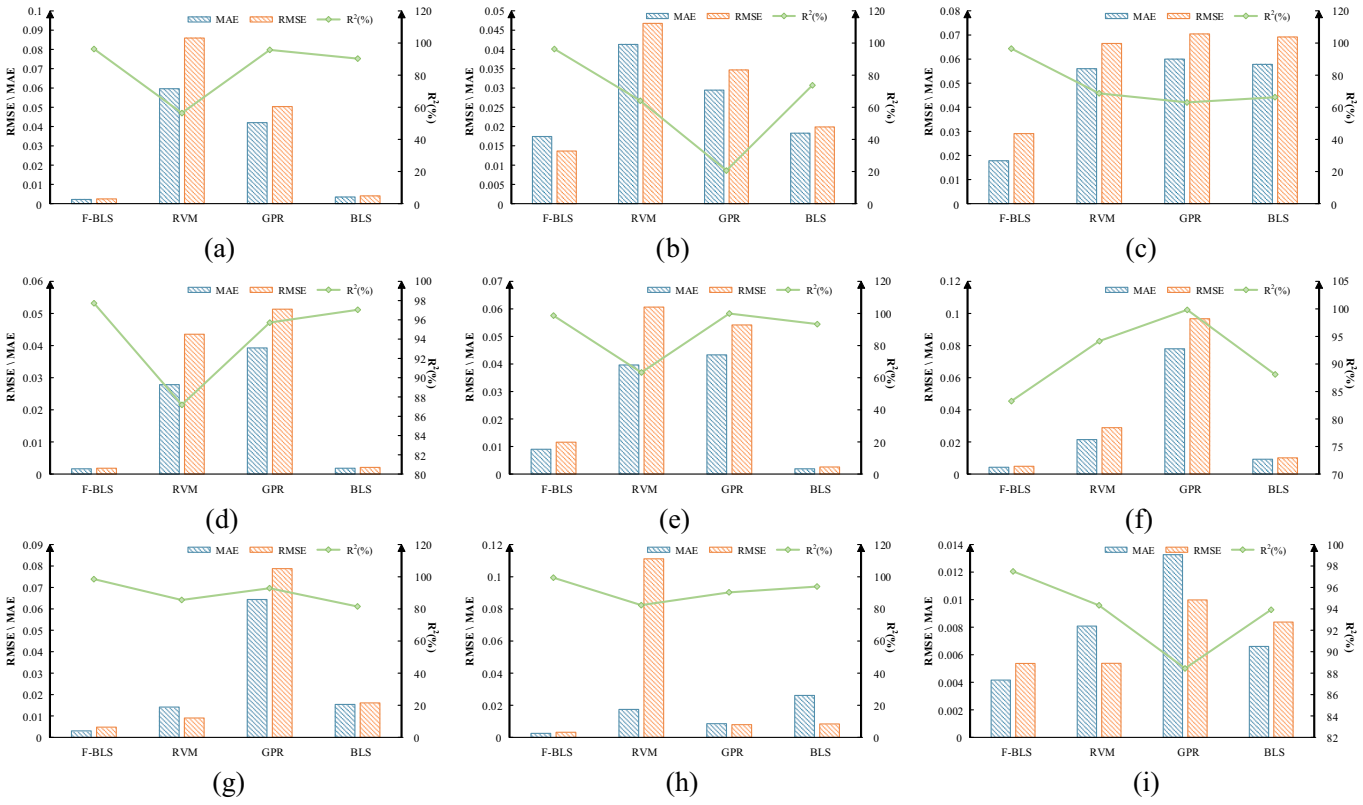


Fig. 12. The evaluation result for different methods: (a) evaluation result of 2#, (b) evaluation result of 3#, (c) evaluation result of 4#, (d) evaluation result of 6#. (e) evaluation result of 7#, (f) evaluation result of 8#, (g) evaluation result of 10#, (h) evaluation result of 11#. (i) evaluation result of 12#.

Cell4 are labeled as 9#, 10#, 11#, and 12#. The feature mapping nodes of F-BLS are set to 7, and the enhancement nodes are set to 50.

In the cyclic aging dataset obtained from the experiments, the model was trained using battery 1#, and its performance was verified using batteries 2#, 3#, and 4#. As depicted in Fig. 8, the SOH estimates exhibit the same declining trend as the actual data, with estimation errors consistently within ±1.5 %. This indicates that the method possesses the capability to accurately estimate the SOH and demonstrates superior reliability and stability.

Fig. 9 displays the full-life SOH prediction results for battery units 6#, 7#, and 8# using battery 5# for training, while Fig. 10 illustrates the SOH prediction results for battery units 10#, 11#, and 12# using battery 9# for training. In comparison with the actual data, the proposed F-BLS method demonstrates exceptionally high prediction accuracy for various datasets, with the error between estimated and actual values consistently within ±1 %.

To further validate the effectiveness and accuracy of F-BLS, commonly used SOH estimation data-driven methods, including RVM and GPR, are employed for comparison. Fig. 11 displays the results and estimation errors of SOH estimation using various methods. It is evident that F-BLS holds a significant advantage over other methods in terms of estimation accuracy and effectively minimizes bias in the estimation results. This paper uses MAE, RMSE, and R^2 to quantify estimation errors. The corresponding values are presented in Fig. 12. The results demonstrate that, compared to other methods, the F-BLS proposed in this paper significantly reduces the SOH estimation error, with MAE and RMSE values below 0.02. Overall, the test results indicate that F-BLS exhibits high estimation accuracy and exceptional generalization ability.

5. Conclusion

This paper introduces an F-BLS-based framework for estimating the SOH of lithium-ion batteries. The methodology employs forward auto-differentiation to optimize the BLS. SOH-related features are extracted from datasets sourced from three datasets. After Pearson correlation analysis, features strongly correlated with capacity decline are selected as inputs for the F-BLS. The F-BLS incorporates regularization techniques and leverages FAdaBound, an adaptive optimization algorithm based on forward auto-differentiation, significantly enhancing prediction accuracy and model generalization. To evaluate the method's efficacy, we validate it using the cyclic aging dataset from the experiment, along with NASA and Oxford battery datasets. The results are compared with two commonly used data-driven methods, indicating superior estimation accuracy of the F-BLS. Furthermore, it maintains robust prediction performance across various battery datasets, with RMSE below 0.03 and MAE values below 0.02, underscoring its high generalization ability. This research contributes to achieving rapid, real-time estimation of battery SOH.

Future work will include conducting aging experiments under various working conditions using diverse batteries to validate further the accuracy and generalization of F-BLS.

CRediT authorship contribution statement

Haoyu Ming: Writing – original draft, Validation, Software, Methodology. **Hui Zhang:** Writing – review & editing, Resources, Investigation, Conceptualization. **Guangfeng Wang:** Writing – review & editing, Visualization, Investigation, Formal analysis. **Dongyu Liu:** Writing – review & editing, Investigation, Conceptualization. **Naxin Cui:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (U23A20327)and the Special Funds of the Central Government Guiding Local Science and Technology Development (YDZX2022002).

References

[1] R. Kumar, E. Joanni, R.K. Singh, D.P. Singh, S.A. Moshkalev, Recent advances in the synthesis and modification of carbon-based 2D materials for application in energy conversion and storage[J], Prog. Energy Combust. Sci. 67 (2018) 115–157.
[2] R. Kumar, S. Sahoo, E. Joanni, R.K. Singh, W.K. Tan, K.K. Kar, A. Matsuda, Recent progress in the synthesis of graphene and derived materials for next generation electrodes of high performance lithium ion batteries[J], Progress in Energy and Combustion Science 75 (2019) 100786.
[3] X. Hu, L. Xu, X. Lin, M. Pecht, Battery lifetime prognostics[J], Joule 4 (2) (2020) 310–346.
[4] C. Pastor-Fernández, T.F. Yu, W.D. Widanage, J. Marco, Critical review of non-invasive diagnosis techniques for quantification of degradation modes in lithium-ion batteries[J], Renew. Sustain. Energy Rev. 109 (2019) 138–159.
[5] K. Liu, Y. Shang, Q. Ouyang, W.D. Widanage, A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery[J], IEEE Trans Ind Electron 68 (4) (2020) 3170–3180.
[6] A. Eddahech, O. Briat, N. Bertrand, J.Y. Delétage, J.M. Vinassa, Behavior and state-of-health monitoring of Li-ion batteries using impedance spectroscopy and recurrent neural networks[J], Int. J. Electr. Power Energy Syst. 42 (1) (2012) 487–494.
[7] S. Yang, C. Zhang, J. Jiang, W. Zhang, Y. Gao, L. Zhang, A voltage reconstruction model based on partial charging curve for state-of-health estimation of lithium-ion batteries[J], Journal of Energy Storage 35 (2021) 102271.
[8] D. Wang, Key-term separation based hierarchical gradient approach for NN based Hammerstein battery model[J], Appl. Math. Lett. 157 (2024) 109207.
[9] C. Li, L. Yang, Q. Li, Q. Zhang, Z. Zhou, Y. Meng, X. Zhao, L. Wang, S. Zhang, Y. Li, F. Lv, SOH estimation method for lithium-ion batteries based on an improved equivalent circuit model via electrochemical impedance spectroscopy[J], Journal of Energy Storage 86 (2024) 111167.
[10] Y. Gao, K. Liu, C. Zhu, D. Zhang, Co-estimation of state-of-charge and state-of-health for lithium-ion batteries using an enhanced electrochemical model[J], IEEE Trans Ind Electron 69 (3) (2021) 2684–2696.
[11] C. Lyu, Q. Lai, T. Ge, H. Yu, L. Wang, N. Ma, A lead-acid battery's remaining useful life prediction by using electrochemical model in the particle filtering framework [J], Energy 120 (2017) 975–984.
[12] F. Zhu, J. Fu, A novel state-of-health estimation for lithium-ion battery via unscented Kalman filter and improved unscented particle filter[J], IEEE Sensors J. 21 (22) (2021) 25449–25456.
[13] W. Li, Y. Fan, F. Ringbeck, D. Jöst, X. Han, M. Ouyang, D.U. Sauer, Electrochemical model-based state estimation for lithium-ion batteries with adaptive unscented Kalman filter[J], J. Power Sources 476 (2020) 228534.
[14] M. Rezvani, S. Lee, J. Lee, A Comparative Analysis of Techniques for Electric Vehicle Battery Prognostics and Health Management (PHM)[R], SAE Technical Paper, 2011.
[15] D. Andre, A. Nuhic, T. Soczka-Guth, D.U. Sauer, Comparative study of a structured neural network and an extended Kalman filter for state of health determination of lithium-ion batteries in hybrid electric vehicles[J], Eng. Appl. Artif. Intel. 26 (3) (2013) 951–961.
[16] A. Widodo, M.C. Shim, W. Caesarendra, B.S. Yang, Intelligent prognostics for battery health monitoring based on sample entropy[J], Expert Syst. Appl. 38 (9) (2011) 11763–11769.
[17] Z. Chen, M. Sun, X. Shu, J. Shen, R. Xiao, On-board state of health estimation for lithium-ion batteries based on random forest[C], in: 2018 IEEE International Conference on Industrial Technology (ICIT), IEEE, 2018, pp. 1754–1759.
[18] H. Xu, L. Wu, S. Xiong, W. Li, A. Garg, L. Gao, An improved CNN-LSTM model-based state-of-health estimation approach for lithium-ion batteries[J], Energy 276 (2023) 127585.
[19] X. Li, C. Yuan, Z. Wang, Multi-time-scale framework for prognostic health condition of lithium battery using modified Gaussian process regression and nonlinear regression[J], J. Power Sources 467 (2020) 228358.
[20] Y. Fan, F. Xiao, C. Li, G. Yang, X. Tang, A novel deep learning framework for state of health estimation of lithium-ion battery[J], Journal of Energy Storage 32 (2020) 101741.
[21] Y. Zhang, Z. Peng, Y. Guan, L. Wu, Prognostics of battery cycle life in the early-cycle stage based on hybrid model[J], Energy 221 (2021) 119901.
[22] Y. Ma, C. Shan, J. Gao, H. Chen, A novel method for state of health estimation of lithium-ion batteries based on improved LSTM and health indicators extraction[J], Energy 251 (2022) 123973.
[23] S. Peng, Y. Sun, D. Liu, Q. Yu, J. Kan, M. Pecht, State of health estimation of lithium-ion batteries based on multi-health features extraction and improved long short-term memory neural network[J], Energy 282 (2023) 128956.
[24] Z. Li, L. Li, J. Chen, D. Wang, A multi-head attention mechanism aided hybrid network for identifying batteries' state of charge[J], Energy 286 (2024) 129504.
[25] C.L.P. Chen, Z. Liu, Broad learning system: an effective and efficient incremental learning system without the need for deep architecture[J], IEEE Transactions on Neural Networks and Learning Systems 29 (1) (2017) 10–24.
[26] T. Gu, D. Wang, Y. Li, A Polak-Ribiere-Polyak conjugate gradient algorithm optimized broad learning system for lithium-ion battery state of health estimation [J], J. Electrochem. Soc. 169 (9) (2022) 090512.
[27] L. Zhang, J. Li, G. Lu, P. Shen, M. Bennamoun, S.A.A. Shah, Q. Miao, G. Zhu, P. Li, X. Lu, Analysis and variants of broad learning system[J], IEEE Trans Syst Man Cybern Syst 52 (1) (2020) 334–344.
[28] P.Q. Huang, B. Chen, Bidirectional broad learning system[C], in: 2020 IEEE 7th International Conference on Industrial Engineering and Applications (ICIEA), IEEE, 2020, pp. 963–968.
[29] C.L.P. Chen, Z. Liu, S. Feng, Universal approximation capability of broad learning system and its structural variations[J], IEEE Transactions on Neural Networks and Learning Systems 30 (4) (2018) 1191–1204.
[30] H. Huang, C. Yang, C.L.P. Chen, Optimal robot–environment interaction under broad fuzzy neural adaptive control[J], IEEE Transactions on Cybernetics 51 (7) (2020) 3824–3835.
[31] Z. Chen, N. Shi, Y. Ji, M. Niu, Y. Wang, Lithium-ion batteries remaining useful life prediction based on BLS-RVM[J], Energy 234 (2021) 121269.
[32] H. Dai, G. Zhao, M. Lin, J. Wu, G. Zheng, A novel estimation method for the state of health of lithium-ion battery using prior knowledge-based neural network and Markov chain[J], IEEE Trans Ind Electron 66 (10) (2018) 7706–7716.
[33] B. Sahaand, K. Goebel, Battery Data Set, NASA Ames Prognostics Data Repository [J], NASA Ames Research Center, 2007.
[34] Z. Li, Y. Yang, L. Li, D. Wang, A weighted Pearson correlation coefficient based multi-fault comprehensive diagnosis for battery circuits[J], Journal of Energy Storage 60 (2023) 106584.
[35] B. Jiang, H. Dai, X. Wei, Incremental capacity analysis based adaptive capacity estimation for lithium-ion battery considering charging condition[J], Appl. Energy 269 (2020) 115074.
[36] L. Luo, Y. Xiong, Y. Liu, X. Sun, Adaptive gradient methods with dynamic bound of learning rate[J], arXiv (2019) preprint arXiv:1902.09843.
[37] Y. LeCun, Y. Bengio, G. Hinton, Deep learning[J], Nature 521 (7553) (2015) 436–444.
[38] A.G. Baydin, B.A. Pearlmutter, A.A. Radul, J.M. Siskind, Automatic differentiation in machine learning: a survey[J], J. Mach. Learn. Res. 18 (153) (2018) 1–43.
[39] H. Li, D. Pan, C.L.P. Chen, Intelligent prognostics for battery health monitoring using the mean entropy and relevance vector machine[J], IEEE Trans Syst Man Cybern Syst 44 (7) (2014) 851–862.
[40] X. Li, C. Yuan, X. Li, Z. Wang, State of health estimation for Li-ion battery using incremental capacity analysis and Gaussian process regression[J], Energy 190 (2020) 116467.
[41] E. Chemali, P.J. Kollmeyer, M. Preindl, R. Ahmed, A. Emadi, Long short-term memory networks for accurate state-of-charge estimation of Li-ion batteries[J], IEEE Trans Ind Electron 65 (8) (2017) 6730–6739.