

Article

Deep Fusion Modeling for Lithium Batteries SOC-SOH Joint Estimation

Jian Wang¹, Shu Cheng¹ and Lulin Zhang^{2,*} 

¹ School of Traffic & Transportation Engineering, Central South University, Changsha 410075, China; wangjian277013@csu.edu.cn (J.W.); chengshu@csu.edu.cn (S.C.)

² School of Information and Electrical Engineering, Hunan University of Science and Technology, Xiangtan 411201, China

* Correspondence: llzhang@hnust.edu.cn

Abstract

This paper proposes a SOC-SOH joint estimation method based on adaptive weighted multi-channel LSTM Transformer fusion network (MLTA-Net). The proposed method constructs a battery health factor set which covers multi-level features, and the aging trend of batteries can be characterized from multiple dimensions. The MLTA-Net model adopts a multi-channel parallel architecture, which can analyze the different types of battery data characteristics. Short-term temporal dependencies are captured by LSTM encoder, and global operating characteristics are analyzed using Transformer multi head self-attention mechanism. Based on adaptive weighted fusion layer for feature fusion, high-precision estimation of battery state can be achieved. Experimental results on CATL-1 and CATL-2 datasets show that the proposed MLTA-Net achieves superior SOH estimation accuracy, with RMSE values of 0.286 and 0.287, MAE values of 0.151 and 0.162, MAPE values of 0.053 and 0.056, and R^2 values of 0.997 and 0.997, respectively. Compared with CNN-GRU, MLP-Attention, Transformer, MLP, RNN, and SVR models, the proposed method exhibits lower prediction errors and better robustness.

Keywords: SOC-SOH joint estimation; LSTM; transformer; battery

MSC: 93B30

1. Introduction

Lithium-ion batteries are widely used in power and energy storage systems due to their high energy density, long cycle life, and low self-discharge rate [1,2]. State of Charge (SOC) reflects the remaining energy of a battery [3,4], while State of Health (SOH) indicates its health status and degradation degree [5,6]. The accuracy of the estimation directly impacts energy management strategies and operational safety [7]. Existing SOC and SOH estimation methods can be broadly categorized as experimental [8,9], model-driven [10,11], and data-driven approaches [12–14]. Experimental methods such as electrochemical impedance spectroscopy can achieve high accuracy, but they are costly and difficult to deploy on-line. Model-driven methods provide physical interpretability, yet they rely on accurate parameter identification and may degrade when batteries age or operating conditions vary. Data-driven methods have become the most promising, as they can adapt to the nonlinear dynamic characteristics of batteries under complex operating conditions.

Typically, the Multi-Layer Perceptron (MLP) network can be applied for SOC estimation [15]. However, because the MLP network is unable to effectively capture the



Academic Editor: Alberto Fernández

Received: 15 May 2026

Revised: 23 June 2026

Accepted: 3 July 2026

Published: 11 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](#)

[Attribution \(CC BY\)](#) license.

temporal information in time-series data, it is limited when analyzing the time-dependent characteristics of lithium-ion battery charging and discharging processes. Considering the time-varying dynamic processes, SOC can be viewed as a time-series problem, and Recurrent Neural Networks (RNNs) can be employed for SOC estimation [16,17]. However, traditional RNNs are prone to issues such as “gradient vanishing” and “gradient explosion” during training, which can adversely affect the model’s convergence and prediction accuracy. In order to address this issue, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks have been adopted [18,19]. These two types of networks can effectively process sequential data and capture long-range dependencies. However, LSTM and GRU rely on step-by-step iteration when processing longer sequences, resulting in lower computational efficiency and potential loss of global features. Through the self-attention mechanism, Transformer can attend to all time steps in the entire time series simultaneously during a single forward pass, making it suitable for handling long time series data [20,21]. In addition, by integrating Convolutional Neural Networks with the Transformer, the model’s feature extraction and fusion capabilities can be further enhanced [22].

Although data-driven methods have achieved significant advances in SOC and SOH estimation, SOC and SOH are modeled independently in most research. Hence, the coupling relationship between SOC and SOH cannot be effectively analyzed. Battery capacity degradation can directly affect SOC estimation, especially during long-term operation or under deep charge–discharge cycles. If SOC estimation is only based on nominal capacity, systematic errors will be introduced. Conversely, the accumulation of SOC estimation errors can also lead to SOH prediction bias, further exacerbating the uncertainty in battery state estimation [23]. In order to solve this problem, dual Kalman filtering is used to update the internal parameters of the battery and SOC estimation. SOC-SOH joint estimation can be achieved, which improves the estimation accuracy to a certain extent. However, this method highly relies on the accuracy of the battery model, and parameter instability can still lead to a decrease in estimation performance when facing complex operating conditions and battery aging [24]. In addition, deep learning models such as support vector machines, RNNs, and LSTMs are combined and applied for SOC-SOH joint estimation [25,26]. On this basis, recent studies have further introduced cloud-based deep learning, attention mechanisms, adaptive filtering, and hybrid Transformer architectures into battery state estimation. In ref. [27], a cloud-based deep learning framework is developed for SOC-SOH co-estimation. In ref. [28], a BiLSTM-SA-based joint estimation method is proposed to improve temporal feature extraction. A dual adaptive central difference H-infinity filtering method is presented in [29] for coupled SOC and SOH estimation. In ref. [30], a Transformer–LSTM fusion model is used to capture battery degradation features, while a Transformer-GRU parallel architecture is constructed in [31] to enhance SOH estimation accuracy. However, these methods still have drawbacks such as high computational complexity and poor real-time performance.

Regarding the joint estimation of SOC-SOH, existing data-driven methods are limited in feature extraction and temporal modeling. Therefore, in relatively complex working conditions, the model estimation accuracy is not high and the stability is poor. In order to address these limitations, this paper proposes an MLTA-Net-based joint estimation framework for LFP batteries. The main contributions are as follows:

- (1) A multi-level battery health-factor set is constructed by integrating initialization features, statistical distribution features, and dynamic aging features. Mutual information is further used to select representative health indicators, which improves the effectiveness of SOH characterization.
- (2) An adaptive multi-channel LSTM-Transformer fusion architecture is designed. The multi-channel unit learns heterogeneous battery features independently, the LSTM

encoder captures local temporal dependencies, and the Transformer encoder models global long-range relationships. The adaptive weighting mechanism further enhances the representation of SOC-SOH coupling characteristics.

- (3) The proposed method is validated under battery aging and complex operating conditions. Compared with typical machine learning and deep learning models, MLTA-Net achieves lower estimation errors and better robustness.

2. Materials and Methods

2.1. Architecture of the MLTA-Net Model

This paper proposes a multi-channel adaptive network architecture based on LSTM and Transformer (MLTA-Net), as shown in Figure 1. The proposed model mainly includes three parts: a multi-channel feature extraction layer, a dual-encoder architecture, and a feature-weighted fusion module.

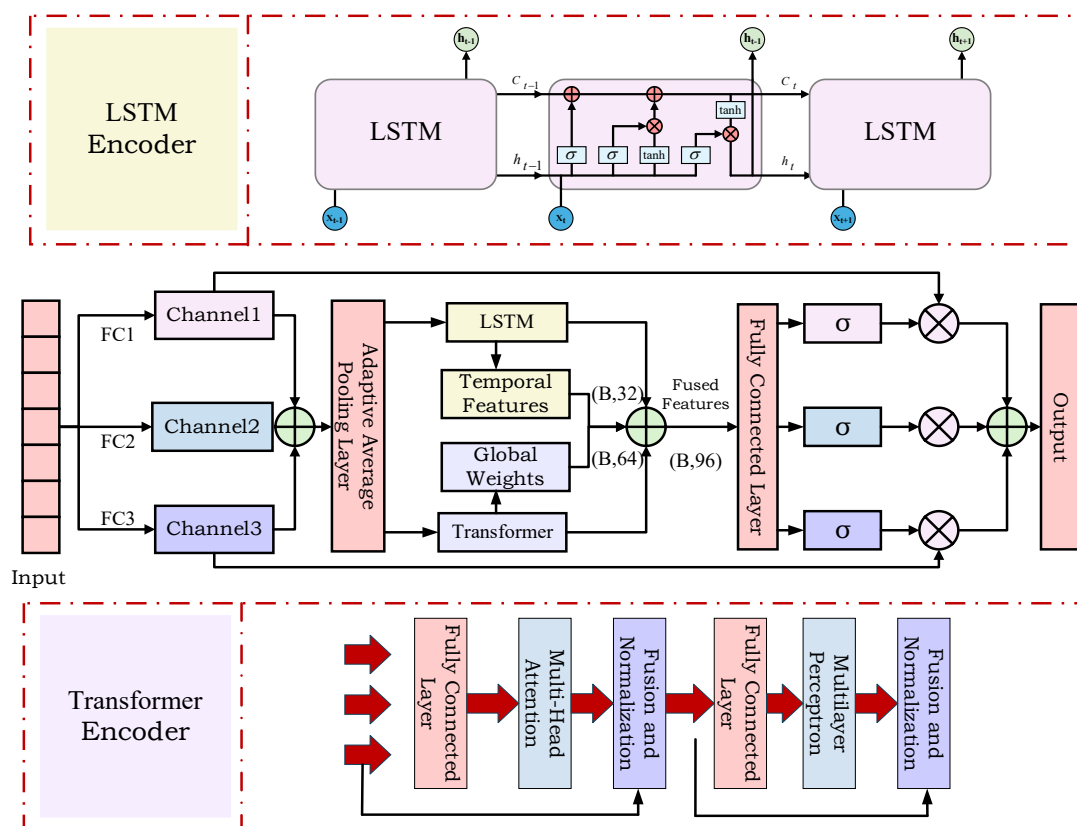


Figure 1. MLTA-Net model framework.

Figure 2 illustrates the overall workflow of the proposed SOC-SOH joint estimation framework. First, multimodal battery aging data, including voltage, current, temperature, and capacity information, are collected from battery degradation experiments. Then, aging-related features are extracted from three levels, including initialization features, statistical distribution features, and dynamic aging features. Mutual-information-based feature selection is further used to identify representative health factors that are highly related to battery degradation.

The selected health factors are then fed into different feature encoding channels. Each channel maps the input features through an independent fully connected layer and nonlinear activation function. Subsequently, the mapped features are encoded by the LSTM encoder and Transformer encoder to capture local temporal dependencies and global long-

range degradation characteristics, respectively. On this basis, an adaptive weighted fusion strategy is introduced to optimize the combination of key features and suppress irrelevant information. The fused features are used for SOH prediction. Meanwhile, SOH information is introduced into the SOC estimation process as long-term degradation information, enabling SOC prediction to reflect the actual dynamic state of the battery during aging. Therefore, the proposed framework effectively utilizes the nonlinear coupling relationship between SOC and SOH and improves battery state estimation through multimodal feature fusion and coupled state modeling.

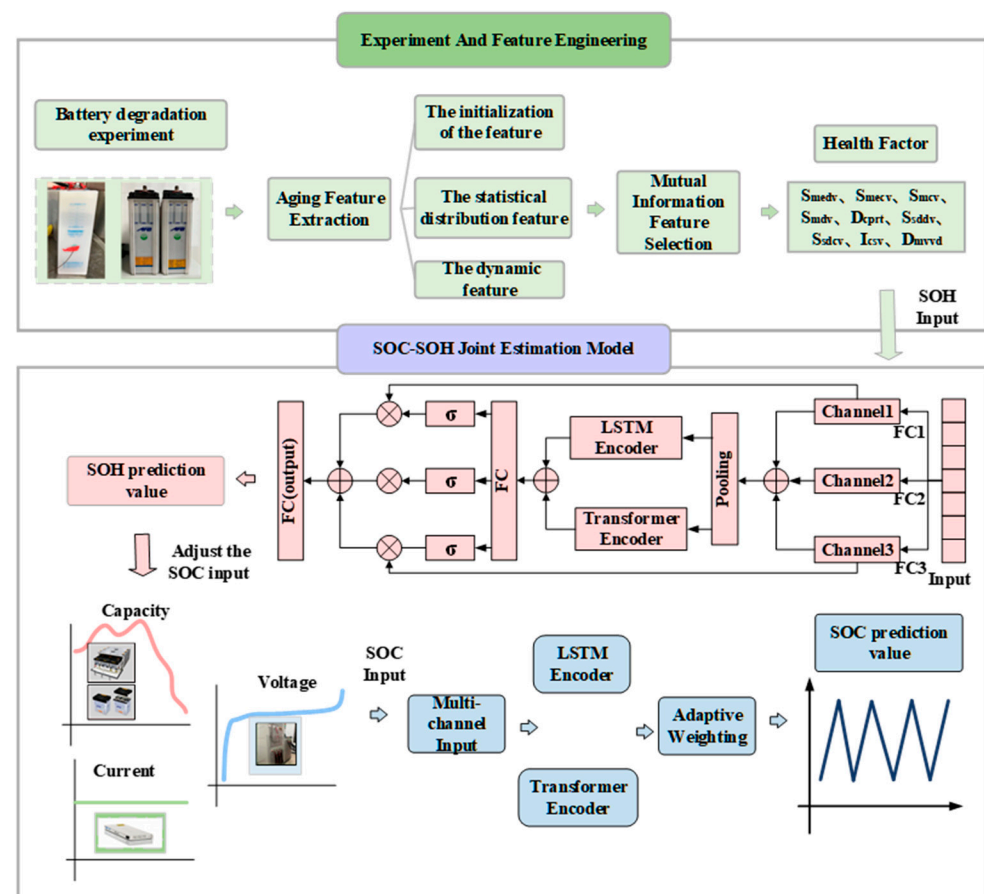


Figure 2. Overall workflow of the SOC-SOH joint estimation framework.

The proposed MLTA-Net can be mathematically described as a hierarchical and integrated feature representation framework consisting of multi-channel mapping, temporal feature encoding, and adaptive weighted fusion, rather than a simple serial combination of LSTM and Transformer modules. First, the input feature vector is mapped into multiple independent feature channels through fully connected layers and nonlinear activation functions, which can be expressed as $(x_i = \text{ReLU}(W_i x_t + b_i))$, where $(i = 1, 2, 3)$. This multi-channel structure decomposes heterogeneous battery information into different feature spaces and enables independent learning of different types of battery features. Then, the extracted channel features are encoded by the LSTM and Transformer modules with different functional roles. The LSTM encoder mainly models local temporal dependencies and short-term sequential evolution through the gate-state update mechanism, while the Transformer encoder captures global long-range degradation characteristics through the query-key-value self-attention mechanism. Finally, the encoded features are dynamically integrated through the adaptive weighted output unit, which assigns different importance to feature channels and generates the final SOC or SOH estimation result. Therefore, MLTA-

Net provides a structured feature representation mechanism for SOC-SOH joint estimation rather than a direct stacking of existing neural network modules.

2.2. Multi-Channel Input Unit

The input data is mapped to different channels through three fully connected layers: FC1, FC2, and FC3. The output of each fully connected layer undergoes nonlinear mapping by the ReLU activation function, which can be expressed as

$$\begin{cases} x_1 = \text{ReLU}(W_1x_t + b_1) \\ x_2 = \text{ReLU}(W_2x_t + b_2) \\ x_3 = \text{ReLU}(W_3x_t + b_3) \end{cases} \quad (1)$$

where W_i and b_i represent the weight matrix and bias term of each fully connected layer, respectively.

Each channel undergoes independent linear transformation and feature extraction, facilitating the capture of independent features from different input information. Subsequently, the features are fused, which can be expressed as

$$x_{\text{combined}} = x_1 + x_2 + x_3 \quad (2)$$

The proposed design leverages the heterogeneity of the data. By employing multi-channel inputs to achieve effective decomposition and independent learning of features, the model's expressive capacity is enhanced, and its ability to identify critical features of battery SOC and SOH is improved. Consequently, it can more accurately capture the dynamic changes and degradation trends of the battery.

The reason for using three independent channels is related to the composition and physical meaning of the input features. The input features mainly include three types of information: raw operating features, statistical distribution features, and aging-sensitive health factors. These features have different physical meanings, time scales, and relationships with battery degradation. Therefore, three independent channels are introduced to learn latent representations from different feature spaces separately. This design reduces the interference among heterogeneous features in a unified mapping space and improves the model's ability to represent the complex nonlinear coupling relationship between SOC and SOH. A two-channel structure may merge features with different physical meanings, while a four-channel structure may increase model complexity and introduce redundant feature decomposition. Therefore, the three-channel structure is adopted according to the feature composition in this study.

2.3. LSTM and Transformer Encoder Unit

The LSTM encoder includes an input gate i_t , a forget gate f_t , and an output gate O_t , which can be written as

$$\begin{cases} i_t = \sigma(\mathbf{W}_i x_t + \mathbf{U}_i h_{t-1} + \mathbf{b}_i) \\ f_t = \sigma(\mathbf{W}_f x_t + \mathbf{U}_f h_{t-1} + \mathbf{b}_f) \\ o_t = \sigma(\mathbf{W}_o x_t + \mathbf{U}_o h_{t-1} + \mathbf{b}_o) \end{cases} \quad (3)$$

where σ denotes the sigmoid function. $\mathbf{W}$ are weight matrices from the current input to the corresponding gates. $\mathbf{U}$ are weight matrices from the previous hidden state to the gates. $\mathbf{b}$ are bias vectors.

The cell state C_t is updated by combining the outputs of the forget gate and input gate, which can be derived as

$$\begin{cases} \widetilde{C}_t = \tanh(\mathbf{W}_c x_t + \mathbf{U}_c h_{t-1} + \mathbf{b}_c) \\ C_t = f_t \odot C_{t-1} + i_t \odot \widetilde{C}_t \end{cases} \quad (4)$$

Hidden state h_t is generated by nonlinear transformation of the output gate and current cell state. Pooled features x_{pooled} serve as LSTM input. Combined with the hidden state update formula, it is expressed as

$$\begin{cases} h_t = o_t \cdot \tanh(C_t) \\ h_t = LSTM(x_{pooled}, h_{t-1}) \end{cases} \quad (5)$$

The Transformer encoder captures global dependencies among elements by computing their correlations. Before input X_{pooled} , positional encoding PE is added to feature vectors at each time step. The positional encoding input can be expressed as

$$X_{input} = X_{pooled} + PE \quad (6)$$

For each input X_{input} , it is mapped to query Q , key K , and value V spaces by 3 linear transformations, which can be expressed as

$$\begin{cases} Q = X_{input} \mathbf{W}_Q \\ K = X_{input} \mathbf{W}_K \\ V = X_{input} \mathbf{W}_V \end{cases} \quad (7)$$

Attention scores can be computed as

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (8)$$

The final multi head self-attention output can be derived as

$$h_{\text{multi-head}} = \text{Concat}(h_1, h_2, \dots, h_h) \mathbf{W}_o \quad (9)$$

In the proposed design, the LSTM encoder retains historical information in temporal features, while the Transformer encoder captures global dependencies through self-attention mechanisms.

2.4. Adaptive Weighted Output Unit

By using 3 independent Sigmoid layers for nonlinear mapping, the feature weight values of each input channel can be obtained, which can be expressed as

$$\begin{cases} y_{\text{sigmoid1}} = \sigma(\mathbf{W}_1 x_1 + \mathbf{b}_1) \\ y_{\text{sigmoid2}} = \sigma(\mathbf{W}_2 x_2 + \mathbf{b}_2) \\ y_{\text{sigmoid3}} = \sigma(\mathbf{W}_3 x_3 + \mathbf{b}_3) \end{cases} \quad (10)$$

where $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3 \in \mathbf{R}^{1 \times 32}$ are the weight matrices of the Sigmoid layer. $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3 \in \mathbf{R}$ are bias vectors. σ is the Sigmoid activation function.

The final feature x_{final} can be calculated as

$$x_{final} = \text{concat}(\sigma_1 \times x_1, \sigma_2 \times x_2, \sigma_3 \times x_3) \quad (11)$$

The weighted features will be compressed through a fully connected layer, and the final prediction result can be generated, which can be expressed as

$$y_{final} = \mathbf{W}_{out} x_{final} + \mathbf{b}_{out} \quad (12)$$

In the adaptive weighted fusion module, each channel is assigned a learnable weight through an independent Sigmoid mapping layer. The obtained weight reflects the relative contribution of the corresponding feature channel to the final battery state representation. After weighting, the channel features are concatenated and passed through the final fully connected layer to generate the SOC or SOH prediction. This mechanism enables the model to dynamically enhance important feature channels and suppress redundant or weakly correlated information, thereby improving the robustness and accuracy of battery state estimation.

2.5. Model Training and Hyperparameter Settings

To verify the effectiveness and advantages of the proposed method, the charge–discharge process data of two batteries, namely CATL-1 and CATL-2, are selected for feature extraction and processing. To ensure scale consistency among different variables, all features and labels are normalized using MinMaxScaler and mapped to the interval of [0, 1]. To ensure scale consistency among different variables, all features and labels are normalized using MinMaxScaler and mapped to the interval of [0, 1]. The detailed dataset partitioning strategy is described in Section 4.1. The detailed dataset partitioning strategy is described in Section 4.1. This strategy improves the randomness and representativeness of the test data.

During data augmentation, a random-noise-based augmentation strategy is adopted to improve the robustness of the model to noise and data variations. Specifically, the training samples are augmented fivefold, and each augmented sample is generated by adding normally distributed Gaussian noise. The noise intensity is constrained within 1–3% to ensure that the noise level is moderate and does not distort the original data distribution. In the model training process, the hidden dimension of the LSTM encoder is set to 32, and the hidden dimension of the Transformer encoder is set to 64. A single-layer multi-head attention mechanism with two attention heads is used to model global features. The Adam optimizer is adopted, and the learning rate is set to 0.001. These settings provide a stable basis for model training and performance improvement.

3. Battery Aging Experiments and Feature Analysis

In this study, lithium iron phosphate battery cells with a rated capacity of 280 Ah and a rated voltage of 3.2 V are selected for charge–discharge aging experiments.

3.1. Health Factors of Lithium Batteries

The first contribution of this study is the construction of a multi-level health-factor set for battery degradation characterization. Since battery aging is reflected not only in capacity decay but also in voltage evolution, current distribution, and dynamic charging behavior, the proposed feature set integrates initialization features, statistical distribution features, and dynamic features. This design enables the model to describe battery degradation from multiple perspectives and provides physically meaningful inputs for the subsequent SOH estimation and SOC correction.

Three feature sets (a total of 18 features) are selected and extracted from the charging and discharging data, as shown in Table 1.

In order to obtain the optimal voltage interval for constant pressure rise time, a moving voltage window method is adopted. By adjusting the voltage interval, the proposed method evaluates the correlation between the constant pressure rise time and SOH within each interval to identify the voltage range that maximizes this correlation, as shown in Figure 3.

The common charging voltage interval for batteries is 2.74–3.65 V. In addition, the initial voltage window size is 0.01 V, and the update step size of the voltage window is

0.01 V. During the sliding process of the voltage window, a constant pressure rise time matrix under various SOH conditions is calculated. Pearson correlation coefficient analysis is performed on the matrices corresponding to each voltage window to identify the window with the strongest SOH correlation. As the voltage window size increased, the Pearson correlation coefficient exhibited a trend of rapid initial growth followed by stabilization, where the Elbow Point served as the optimal turning point on the curve. Based on the Kneedle algorithm, the elbow position was determined, and a voltage window of 0.26 V was ultimately selected. The corresponding voltage interval was 3.30 V to 3.56 V, as shown in Figure 4.

Table 1. Feature set.

Set	Feature
Set of initialize feature I	Charging Starting Voltage I_{CSV}
	Discharging Starting Voltage I_{DSV}
Set of statistical distribution feature S	Median Charging Voltage S_{mcv}
	Mean Charging Voltage S_{mecv}
	Standard Deviation of Charging Voltage S_{sdcv}
	Median Discharging Voltage S_{mdv}
	Mean Discharging Voltage S_{medv}
	Standard Deviation of Discharging Voltage S_{sddv}
	Mean Charging Current S_{mecc}
	Standard Deviation of Charging Current S_{sdcc}
	Mean Discharging Current S_{medc}
	Standard Deviation of Discharging Current S_{sddc}
	Skewness of Charging Current S_{scc}
	Kurtosis of Discharging Current S_{kdc}
	Skewness of Discharging Current S_{sdc}
Set of dynamic feature D	Constant Pressure Rise Time D_{cprt}
	Maximum Capacity Increment Voltage D_{mciv}
	Maximum Voltage Value During Charging D_{mvvc}
	Maximum Voltage Value During Discharging D_{mvvd}

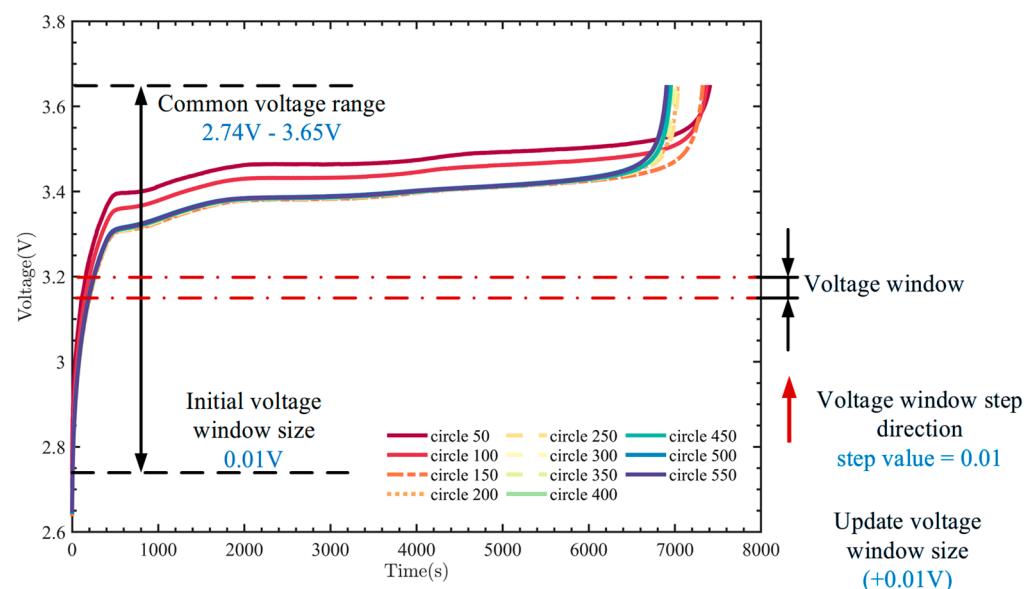


Figure 3. Constant pressure rise time.

It should be noted that the proposed framework is not a strict multi-task learning structure with two completely parallel output branches. Instead, the proposed method

follows a sequentially coupled estimation mechanism. MLTA-Net first extracts deep temporal features and degradation-related features from multi-source battery operation data for SOH estimation. Then, the estimated SOH information is introduced into the SOC estimation process as a long-term degradation representation to dynamically correct the SOC mapping relationship.

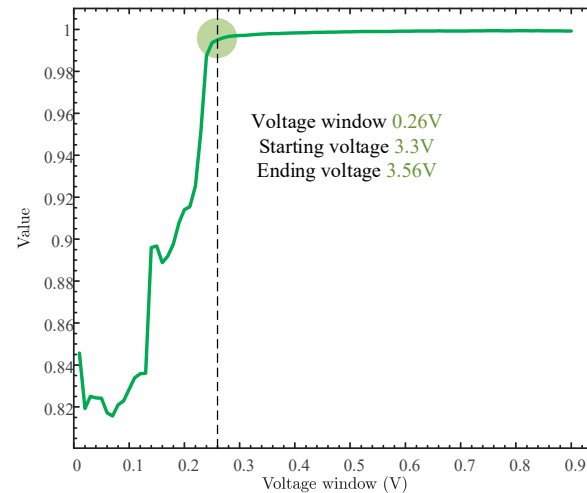


Figure 4. Pearson correlation coefficient.

Therefore, the network does not simply output SOC and SOH simultaneously through two independent parallel heads. SOH is used as auxiliary health information and is fed back into the SOC estimation module. As battery aging progresses, capacity fading and internal resistance growth change the voltage response and charge–discharge behavior of the battery, making the mapping relationship between SOC and operating features time-varying. If SOC is estimated only from instantaneous operating data, systematic bias may occur at different aging stages. By introducing SOH into the SOC estimation process, the proposed framework can better reflect battery degradation status and improve SOC estimation accuracy under different aging stages and operating conditions.

Incremental Capacity Analysis (ICA) can be utilized to evaluate battery performance. The ICA curves of the battery at different cycles are shown in Figure 5. It can be seen that as the number of cycles increases, the peak differential capacity decreases, indicating a decrease in the activity of the electrode material. In addition, the rightward shift in the peak voltage indicates an increase in the polarization degree inside the battery and a delay in the electrochemical reaction process.

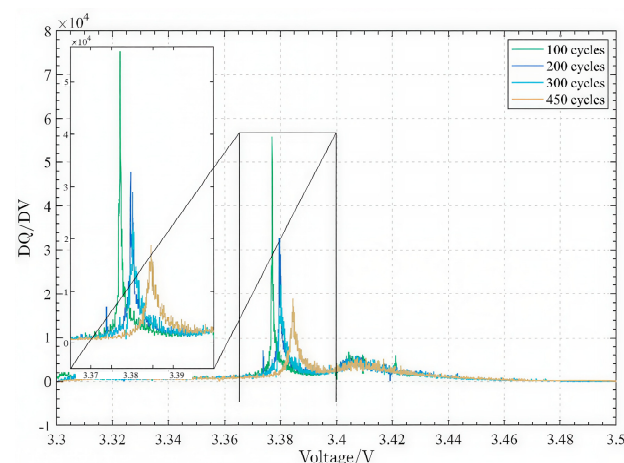


Figure 5. ICA curves of batteries with different cycle number.

3.2. Feature Selection Method Based on Mutual Information

In order to clearly characterize battery degradation and reduce redundant input variables, the battery health-factor selection process is divided into three steps. First, multi-source aging data, including voltage, current, and capacity information, are collected from battery charging and discharging experiments. Second, degradation-related features are extracted from three levels, including initialization features, statistical distribution features, and dynamic aging features. The initialization features describe the initial voltage state during charging and discharging, the statistical distribution features reflect the overall voltage and current distribution characteristics, and the dynamic aging features characterize the electrochemical degradation behavior during battery aging. Third, mutual information and correlation heatmap analysis are jointly used to select representative health indicators for SOH estimation.

Mutual information is first adopted to evaluate the statistical dependency between each extracted feature and the battery capacity label. A higher mutual information value indicates that the corresponding feature contains more degradation-related information and has stronger relevance to SOH estimation. As shown in Figure 6, features with higher mutual information values are selected as candidate health indicators.

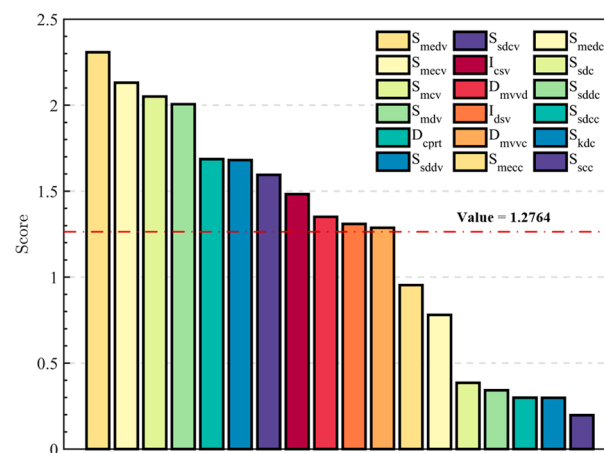


Figure 6. Mutual Information filtering results.

However, feature selection should not only consider the dependency between each feature and the capacity label, but also avoid redundant information caused by multicollinearity among features. Therefore, after calculating the mutual information values, a correlation heatmap is further used to analyze the correlation relationships among the candidate features. As shown in Figure 7, some features, such as I_{dsv} and D_{mvvc} , show relatively high mutual information values, but they also have strong correlations with other selected features, indicating potential redundancy. Including these highly correlated features may increase model complexity without providing additional effective degradation information.

Based on the mutual information ranking and correlation heatmap analysis, features with high relevance to battery capacity and relatively low redundancy are finally selected as battery health indicators. In this study, nine representative features are selected for subsequent SOH estimation, including S_{medv} , S_{mecv} , S_{mcv} , S_{mdv} , D_{cpvt} , S_{sddv} , S_{sdcv} , I_{csv} , and D_{mvvd} . This selection strategy helps improve the representativeness of battery health characterization while reducing redundant input variables.

Although mutual information and correlation heatmap analysis can effectively identify degradation-related and less redundant features, a systematic comparison with alternative feature selection methods and a detailed feature-importance stability analysis are still

limited in this study. Future work will further compare mutual information with correlation-based selection, recursive feature elimination, and SHAP-based feature importance analysis to evaluate the stability and robustness of selected health indicators.

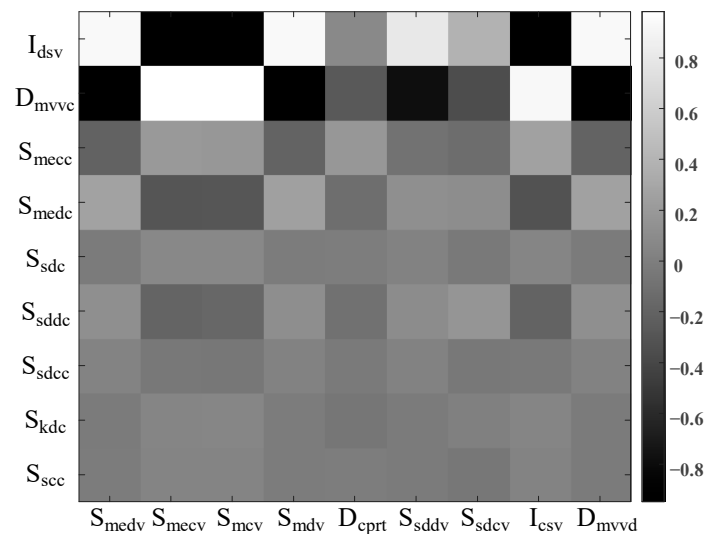


Figure 7. Feature correlation heatmap.

4. Experiment Verification

4.1. SOH Estimation Performance Evaluation

In this study, two CATL lithium iron phosphate battery cells, denoted as CATL-1 and CATL-2, are selected for charge–discharge aging experiments and subsequent SOC–SOH estimation. The tested battery cells have a rated capacity of 280 Ah and a rated voltage of 3.2 V. The experimental platform consists of a high-current charge–discharge test system, a high-low temperature chamber, and a host computer for data acquisition and storage. During the experiments, the batteries are placed in the temperature chamber, and the experimental ambient temperature is maintained at 25 ± 5 °C.

The operating conditions include deep charge–discharge and pulse charge–discharge experiments. For the deep charge–discharge condition, the battery is charged at 240 A until the cut-off voltage reaches 3.65 V, followed by constant-voltage charging at 3.65 V until the cut-off current decreases to 14 A. During discharge, the battery is discharged at 240 A until the cut-off voltage reaches 2.5 V, followed by constant-voltage discharging at 2.5 V until the cut-off current decreases to 14 A. For the pulse charge–discharge condition, constant-power charging with 448 W/44.8 W is used until the battery reaches the cut-off voltage of 3.65 V, while constant-power discharging with 448 W/89.6 W is used until the battery reaches the cut-off voltage of 2.5 V.

The collected data include voltage, current, capacity, and other charge–discharge process information. Based on these data, initialization features, statistical distribution features, and dynamic aging features are extracted for battery health characterization. To ensure scale consistency among different variables, all features and labels are normalized using MinMaxScaler and mapped to the interval of [0, 1]. The dataset is partitioned using an interval-based sampling strategy rather than a complete random shuffle. Specifically, 2 samples out of every 10 consecutive samples are selected as the test set, while the remaining samples are used as the training set. In this way, both the training and test sets cover the early, middle, and late degradation stages. During feature extraction, future-cycle features or future capacity labels are not used to predict historical battery states. In addition, a Gaussian-noise-based data augmentation strategy is adopted for the training set, and the training samples are augmented fivefold with a noise intensity of 1–3%.

To improve the reproducibility of the experimental results, all experiments in this study are conducted with a fixed random seed of 42. The training parameters, data preprocessing procedures, and data partitioning strategy are kept consistent across all comparison models. Therefore, the reported RMSE, MAE, and MAPE results can be reproduced under the same experimental configuration. However, this study mainly reports the results under a fixed and reproducible experimental setting, and confidence intervals or variance across multiple repeated runs are not provided. In future work, repeated experiments with different random seeds will be conducted to further evaluate the statistical variability and confidence intervals of the proposed model.

Gaussian noise is introduced during the training stage to simulate sensor measurement errors and random disturbances in practical operating environments. The purpose of this strategy is to improve the adaptability of the model to input uncertainty and reduce overfitting to ideal experimental data. However, the current study does not provide a systematic robustness evaluation under wider noise levels or more complex real-world disturbances.

As shown in Figure 8, the proposed algorithm exhibits a high degree of fitting in SOH prediction, with results that are generally consistent with the ground truth and a more concentrated error distribution. The error curve demonstrates a narrow width and a relatively symmetrical shape overall, indicating that the prediction errors are uniformly distributed with low volatility. These characteristics further validate the effectiveness of the proposed algorithm in capturing both global dependencies and local dynamic variations in SOH time-series data, thereby ensuring the accuracy and stability of the prediction results.

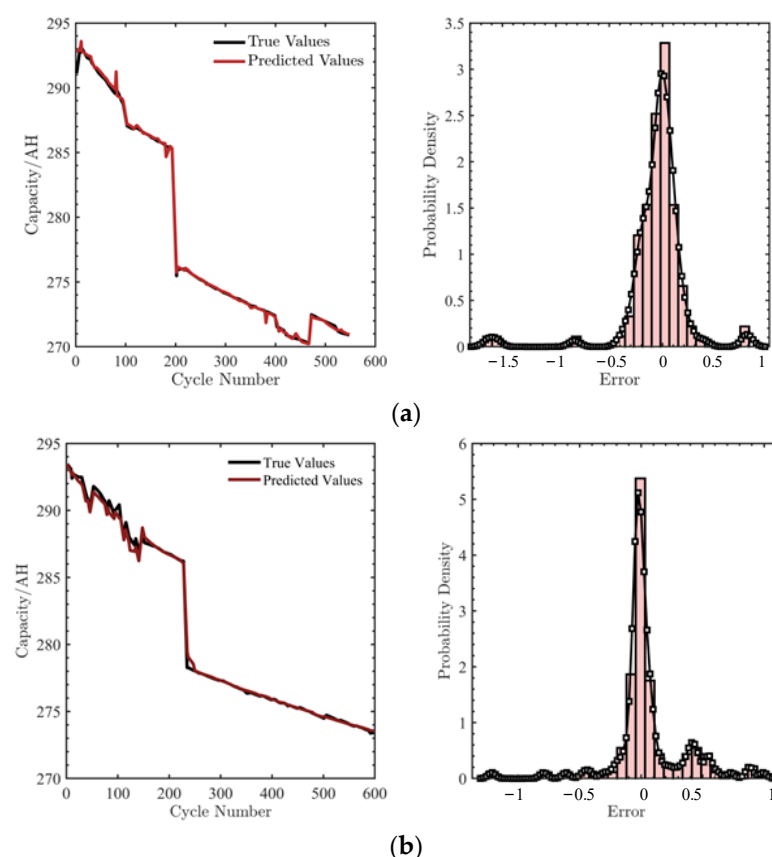


Figure 8. SOH prediction analysis. (a) CATL-1. (b) CATL-2.

In order to comprehensively evaluate the performance of different models in battery state estimation tasks, six typical machine learning and deep learning models are selected for comparative experiments, including CNN-GRU, MLP-Attention, Transformer, MLP, RNN, and SVR. The comparative experimental results are shown in Figure 9.

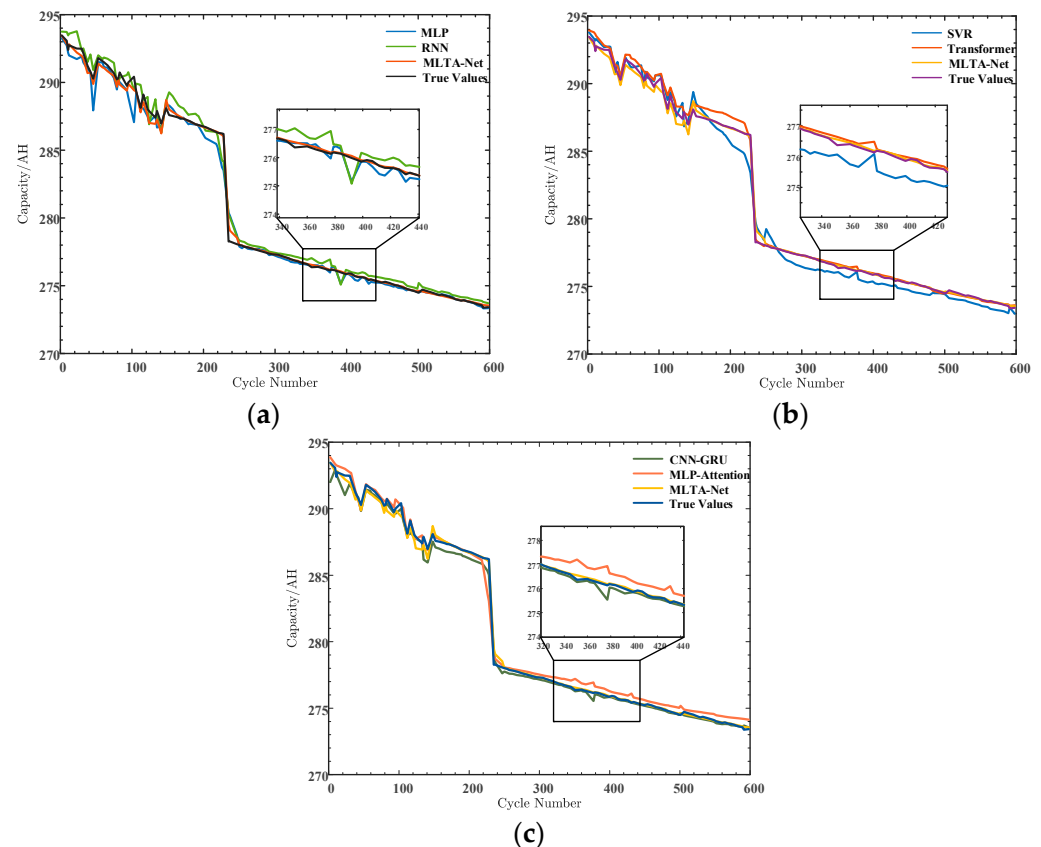


Figure 9. Comparison results of different models. (a) Comparison result 1. (b) Comparison result 2. (c) Comparison result 3.

The boxplots depicting the prediction error distributions of the seven models for battery capacity prediction are shown in Figure 10. The vertical axis represents the residual error between the model predictions and the true capacity. Each box corresponds to the error statistics of a model over the full cycle life, encompassing the median, interquartile range, and outliers. It can be seen that the proposed MLTA-Net exhibits the most concentrated error distribution, with a median close to zero and the smallest interquartile range, indicating that its prediction results are more precise and stable. Furthermore, MLTA-Net has fewer outliers, suggesting that its robustness under abnormal operating conditions is also superior to that of other models.

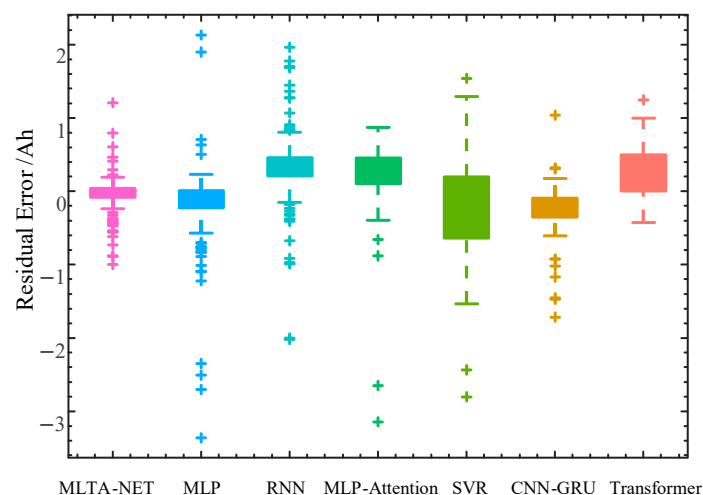


Figure 10. Box plots of performance errors estimated by different models.

The Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) of the seven models are listed in Table 2. It can be seen that the prediction accuracy of MLTA-Net is higher.

Table 2. Error analysis of model performance comparison.

	CATL-1			CATL-2		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE
MLTA-Net	0.286	0.151	0.053	0.287	0.162	0.056
CNN-GRU	0.439	0.305	0.110	0.412	0.278	0.098
MLP-Attention	0.421	0.257	0.091	0.545	0.380	0.136
Transformer	0.687	0.569	0.205	0.410	0.281	0.098
MLP	0.651	0.444	0.158	0.667	0.346	0.121
RNN	0.639	0.454	0.162	0.644	0.488	0.172
SVR	1.184	0.864	0.310	0.721	0.600	0.214

To further evaluate the contribution of the main temporal encoding components in MLTA-Net, ablation experiments are conducted by removing the LSTM encoder and the Transformer encoder separately. The complete MLTA-Net and the two ablation models are evaluated on the same CATL-1 and CATL-2 datasets under the same preprocessing procedure and evaluation metrics. The corresponding prediction curves are shown in Figure 11.

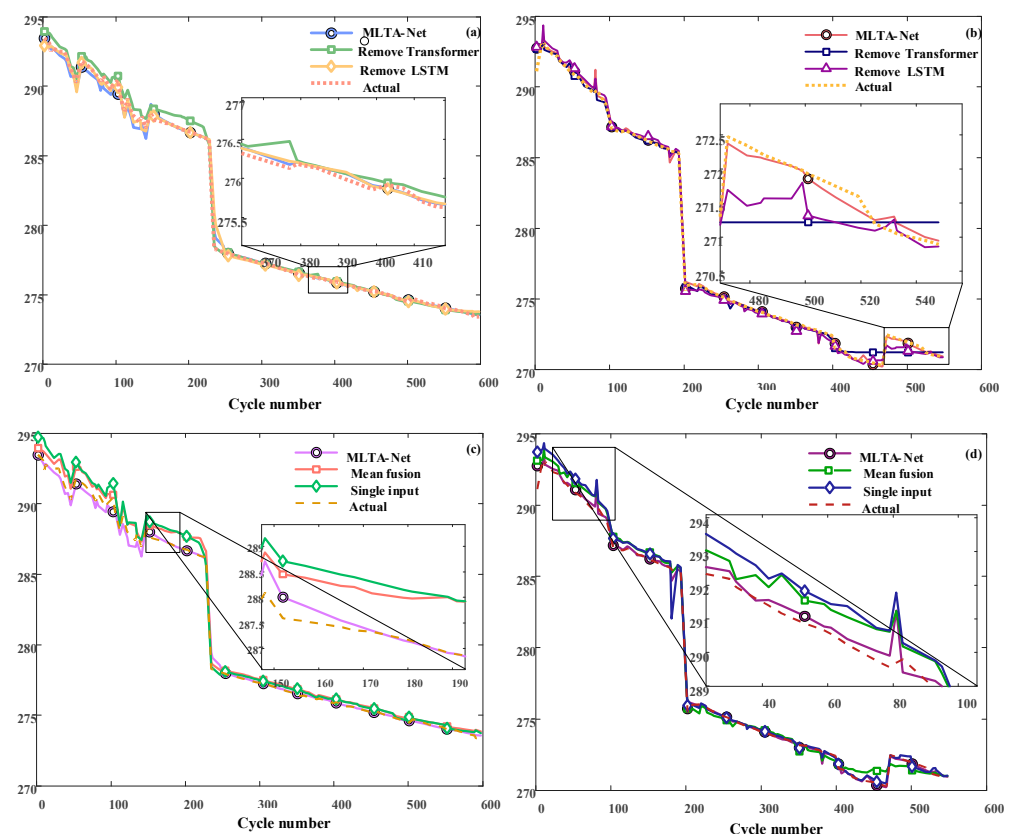


Figure 11. Experimental analysis of MLTA-Net ablation. (a) CATL-1. (b) CATL-2. (a) Ablation study analysis of CATL-1; (b) ablation study analysis of CATL-2; (c) structural optimization comparison of CATL-1; (d) structural optimization comparison of CATL-2.

The ablation results further demonstrate the necessity of the LSTM encoder and Transformer encoder in the proposed architecture. As shown in Figure 10, the complete

MLTA-Net tracks the true capacity degradation trajectory more accurately than the two ablation models. When the LSTM encoder is removed, the model shows larger deviations in local fluctuation regions, indicating that the LSTM module contributes to capturing short-term temporal dependencies and local dynamic variations. When the Transformer encoder is removed, the prediction curve shows weaker consistency with the long-term degradation trend, indicating that the Transformer module is important for modeling global degradation characteristics through the self-attention mechanism.

The quantitative results in Table 3 further confirm this conclusion. Taking CATL-1 as an example, removing the LSTM encoder increases the RMSE from 0.286 to 0.403, corresponding to an increase of 40.9%. Removing the Transformer encoder increases the RMSE from 0.286 to 0.385, corresponding to an increase of 34.6%. The MAE and MAPE values also increase after removing either module. These results indicate that both temporal encoding modules are necessary for the proposed MLTA-Net. The LSTM encoder enhances local temporal feature extraction, while the Transformer encoder improves global long-range dependency modeling. Therefore, the performance improvement of MLTA-Net is achieved by the complementary design of LSTM-based local temporal modeling and Transformer-based global feature representation.

Table 3. Comparative error analysis of ablation experiments.

	CATL-1				CATL-2			
	RMSE	MAE	MAPE	R ²	RMSE	MAE	MAPE	R ²
MLTA-Net	0.286	0.151	0.053	0.997	0.287	0.162	0.056	0.997
Remove LSTM	0.403	0.255	0.092	0.995	0.342	0.154	0.054	0.996
Remove Transformer	0.385	0.271	0.097	0.994	0.379	0.301	0.106	0.994

4.2. SOC Estimation Performance Evaluation

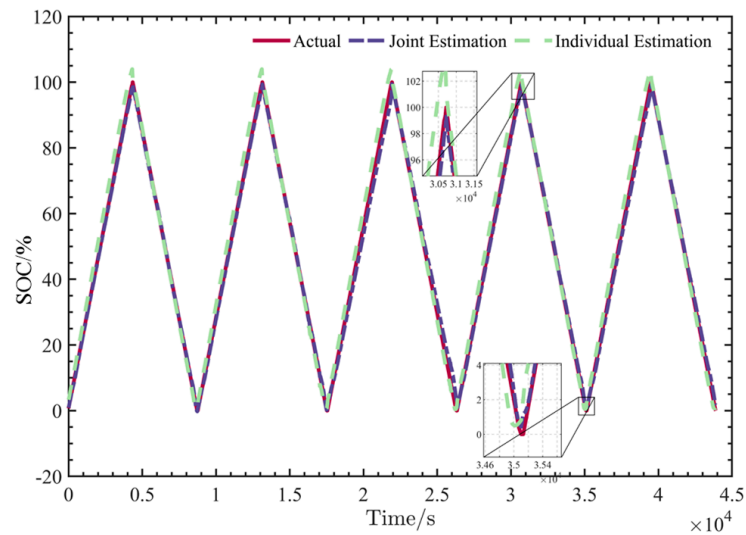
In order to validate the effectiveness of joint estimation, this paper conducts a comparative analysis between joint and individual estimation strategies for SOC prediction at different stages of the lifecycle (cycles 46–50, 346–350, and 546–550). As shown in Figure 12, it can be seen that the deviation between the joint estimation and the actual value is relatively low.

The error results are shown in Figure 13 and Table 4. It can be seen that joint estimation can significantly improve prediction accuracy in each cycle stage, especially under the aging conditions of later cycles.

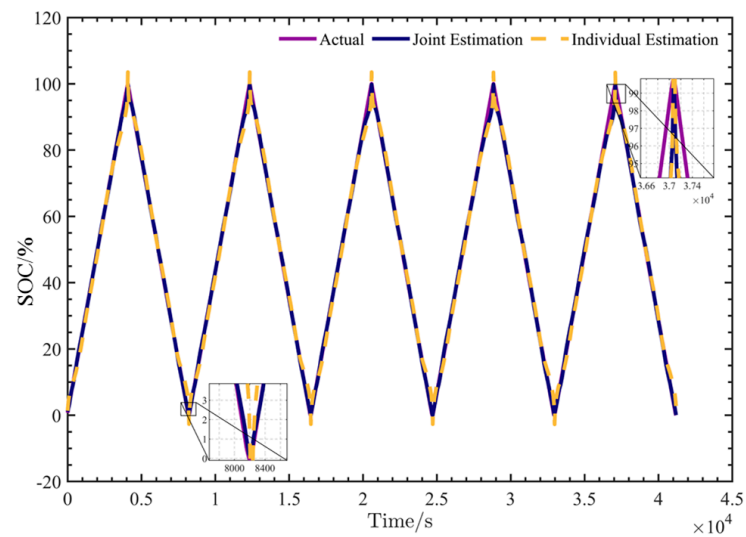
Table 4. SOC estimation performance error under deep charging and discharging conditions.

Cycle Number	Joint Estimation			Individual Estimation		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE
46–50	1.364	1.015	0.068	3.286	2.652	0.116
346–350	1.012	0.526	0.022	2.003	1.515	0.139
546–550	0.821	0.629	0.075	2.540	2.202	0.298

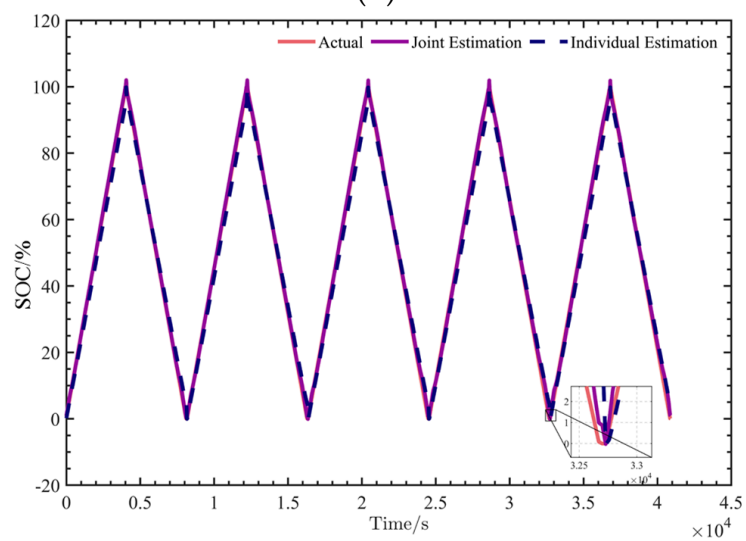
The results of joint and individual SOC estimation by the MLTA-Net model under pulse conditions for the CATL-2 battery at 450 and 600 cycles are shown in Figure 14 and Table 5. It can be seen that the joint estimation method provides more accurate SOC predictions under various operating conditions and adapts better to complex dynamic variations. In contrast, the individual estimation method generates larger errors near SOC mutation points.



(a)



(b)



(c)

Figure 12. Comparison of SOC estimation under deep charging and discharging conditions. (a) 46–50 life cycles. (b) 346–350 life cycles. (c) 546–550 life cycles.

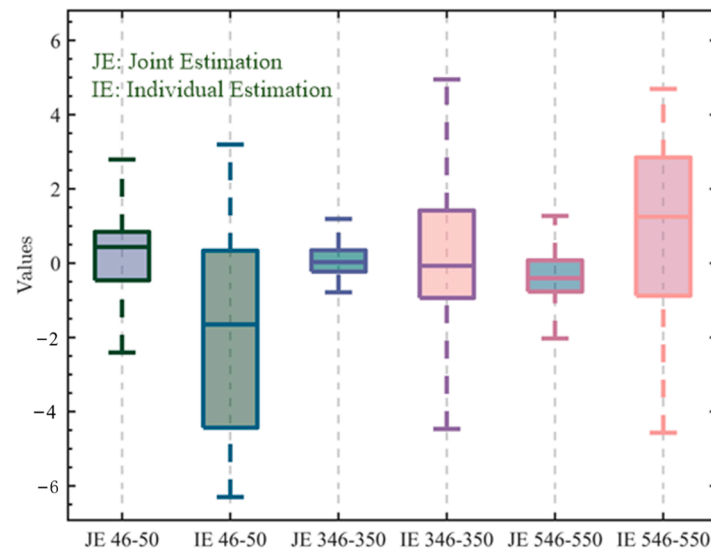


Figure 13. SOC estimation performance error box plot under deep charging and discharging conditions.

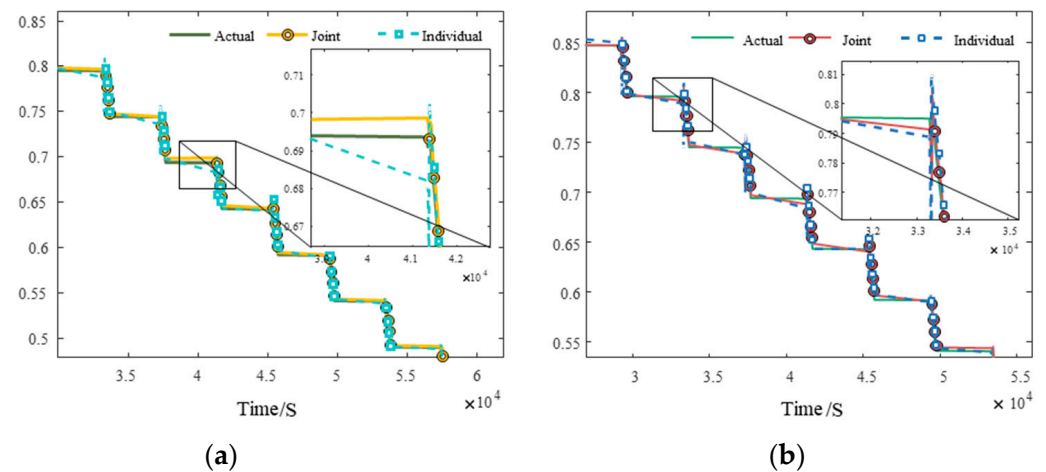


Figure 14. Comparison of SOC estimation under pulse conditions. (a) 450 life cycles. (b) 600 life cycles.

Table 5. SOC estimation performance error under pulse conditions.

Cycle Number	Joint Estimation			Individual Estimation		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE
450	0.0042	0.0034	0.0173	0.0063	0.0042	0.019
600	0.0023	0.0018	0.0075	0.0061	0.0043	0.0262

5. Conclusions

This paper proposes an MLTA-Net model for SOC-SOH joint estimation of lithium batteries, which can analyze the complex nonlinear coupling relationship between SOC and SOH in lithium batteries. The proposed method combines a multi-channel feature extraction module with a multimodal information fusion mechanism. Short-term temporal dependencies are captured through LSTM encoder, and global temporal features are mined using the self-attention mechanism of Transformer encoder. Therefore, the state characterization of lithium batteries is more comprehensive. Through experimental comparison, the proposed method demonstrates superior performance in terms of prediction error, stability, and computational efficiency. High precision estimation of battery status has been achieved.

Although the proposed MLTA-Net achieves promising SOC-SOH joint estimation performance, several limitations remain. First, the model is validated mainly on lithium iron phosphate batteries, and its scalability to other battery chemistries and different manufacturing batches requires further investigation. Second, the proposed method relies on high-quality aging data and extracted health factors; therefore, incomplete or noisy measurements may affect estimation accuracy. Third, although data augmentation is used to improve robustness, further validation under wider temperature ranges, different loading profiles, and real-world vehicle operating conditions is still necessary. These issues will be addressed in future work.

From the perspective of practical battery management systems, the proposed MLTA-Net provides a data-driven solution for aging-aware battery state estimation. By jointly considering SOC and SOH, the method can reduce the influence of capacity degradation on SOC estimation and provide more reliable state information for energy management, safety warning, and maintenance scheduling. Therefore, the proposed framework has potential application value in electric vehicles, energy storage systems, and intelligent battery management platforms.

Author Contributions: Conceptualization, J.W. and L.Z.; Methodology, J.W.; Software, J.W.; Validation, J.W., S.C. and L.Z.; Formal analysis, J.W.; Investigation, J.W.; Resources, J.W.; Data curation, J.W.; Writing—original draft preparation, J.W.; Writing—review and editing, J.W.; Visualization, J.W.; Supervision, J.W.; Project administration, J.W.; Funding acquisition, J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 52572392, and the National Key Research and Development Program of China 2023YFB4301600.

Data Availability Statement: The original contributions presented in the study are included in the paper, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sun, Q.; Lv, H.; Wang, S.; Gao, S.; Wei, K. Optimized State of Charge Estimation of Lithium-Ion Battery in SMES/Battery Hybrid Energy Storage System for Electric Vehicles. *IEEE Trans. Appl. Supercond.* **2021**, *31*, 1–6. [\[CrossRef\]](#)
2. Xu, J.; Mei, X.; Wang, X.; Fu, Y.; Zhao, Y.; Wang, J. A Relative State of Health Estimation Method Based on Wavelet Analysis for Lithium-Ion Battery Cells. *IEEE Trans. Ind. Electron.* **2020**, *68*, 6973–6981. [\[CrossRef\]](#)
3. Han, J.; Dong, Y.; Wang, W. Combined Framework for State of Charge Estimation of Lithium-Ion Batteries: Optimized LSTM Network Integrated with IAOA and AUKF. *Mathematics* **2025**, *13*, 2590. [\[CrossRef\]](#)
4. Fang, W.; Su, Z.-J.; Shao, Y.-T.; Wu, G.-P.; Liu, P. State of Charge Estimation of Lithium-Ion Batteries Based on Hidden Markov Factor Graphs. *Mathematics* **2025**, *13*, 2922. [\[CrossRef\]](#)
5. Al-Rahamneh, A.; Izco, I.; Serrano-Hernandez, A.; Faulin, J. Machine Learning-Based State-of-Health Estimation of Battery Management Systems Using Experimental and Simulation Data. *Mathematics* **2025**, *13*, 2247. [\[CrossRef\]](#)
6. Wei, M.; Yuan, S.; Chen, L.; Liu, Y.; Hu, J. State of Health Prediction of Pure Electric Vehicle Batteries Based on Patch Wavelet Transformer. *Mathematics* **2025**, *13*, 2099. [\[CrossRef\]](#)
7. Yu, Q.; Nie, Y.; Peng, S.; Miao, Y.; Zhai, C.; Zhang, R.; Han, J.; Zhao, S.; Pecht, M. Evaluation of the safety standards system of power batteries for electric vehicles in China. *Appl. Energy* **2023**, *349*, 121674. [\[CrossRef\]](#)
8. Hu, W.; Peng, Y.; Wei, Y.; Yang, Y. Application of electrochemical impedance spectroscopy to degradation and aging research of lithium-ion batteries. *J. Phys. Chem. C* **2023**, *127*, 4465–4495. [\[CrossRef\]](#)
9. Pang, Z.; Yang, K.; Song, Z.; Niu, P.; Chen, G.; Meng, J. A new method for determining SOH of lithium batteries using the real-part ratio of EIS specific frequency impedance. *J. Energy Storage* **2023**, *72*, 108693. [\[CrossRef\]](#)
10. Hu, F.; Xie, B. Multi-Model Fusion of Lithium Battery SOC Estimation Based on Bayesian Principle. *Mathematics* **2026**, *14*, 1642. [\[CrossRef\]](#)
11. Li, X.; Huang, Z.; Tian, J.; Tian, Y. State-of-charge estimation tolerant of battery aging based on a physics-based model and an adaptive cubature Kalman filter. *Energy* **2021**, *220*, 119767. [\[CrossRef\]](#)

12. Qin, P.; Zhao, L. A Data-Driven-Based SOC Estimation Method Using Newly Constructed Features for Different Application Environments. *IEEE Trans. Intell. Veh.* **2025**, *10*, 2963–2976. [\[CrossRef\]](#)
13. Song, Y.; Liu, D.; Liao, H.; Peng, Y. A hybrid statistical data-driven method for on-line joint state estimation of lithium-ion batteries. *Appl. Energy* **2020**, *261*, 114408. [\[CrossRef\]](#)
14. Khaleghi, S.; Hosen, S.; Karimi, D.; Behi, H.; Beheshti, S.H.; Van Mierlo, J.; Berecibar, M. Developing an online data-driven approach for prognostics and health management of lithium-ion batteries. *Appl. Energy* **2022**, *308*, 118348. [\[CrossRef\]](#)
15. Park, M.; Seo, M.; Song, Y.; Kim, S.W. Capacity estimation of Li-ion batteries using constant current charging voltage with multilayer perceptron. *IEEE Access* **2020**, *8*, 180762–180772. [\[CrossRef\]](#)
16. Jiao, M.; Wang, D.; Qiu, J. A GRU-RNN based momentum optimized algorithm for SOC estimation. *J. Power Sources* **2020**, *459*, 228051. [\[CrossRef\]](#)
17. Lu, J.; Xiong, R.; Tian, J.; Wang, C.; Hsu, C.-W.; Tsou, N.-T.; Sun, F.; Li, J. Battery degradation prediction against uncertain future conditions with recurrent neural network enabled deep learning. *Energy Storage Mater.* **2022**, *50*, 139–151. [\[CrossRef\]](#)
18. Chai, X.; Li, S.; Liang, F. A novel battery SOC estimation method based on random search optimized LSTM neural network. *Energy* **2024**, *306*, 132583. [\[CrossRef\]](#)
19. Ren, X.; Liu, S.; Yu, X.; Dong, X. A method for state-of-charge estimation of lithium-ion batteries based on PSO-LSTM. *Energy* **2021**, *234*, 121236. [\[CrossRef\]](#)
20. Shen, H.; Zhou, X.; Wang, Z.; Wang, J. State of charge estimation for lithium-ion battery using Transformer with immersion and invariance adaptive observer. *J. Energy Storage* **2022**, *45*, 103768. [\[CrossRef\]](#)
21. Hannan, M.A.; How, D.N.T.; Lipu, H.M.S.; Mansor, M.; Ker, P.J.; Dong, Z.Y.; Sahari, K.S.M.; Tiong, S.K.; Muttaqi, K.M.; Mahlia, I.T.M.; et al. Deep learning approach towards accurate state of charge estimation for lithium-ion batteries using self-supervised transformer model. *Sci. Rep.* **2021**, *11*, 19541. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Gu, X.; See, K.; Li, P.; Shan, K.; Wang, Y.; Zhao, L.; Lim, K.C.; Zhang, N. A novel state-of health estimation for the lithium-ion battery using a convolutional neural network and transformer model. *Energy* **2023**, *262*, 125501. [\[CrossRef\]](#)
23. Gao, Y.; Liu, K.; Zhu, C.; Zhang, X.; Zhang, D. Co-estimation of state-of-charge and state-of- health for lithium-ion batteries using an enhanced electrochemical model. *IEEE Trans. Ind. Electron.* **2022**, *69*, 2684–2696. [\[CrossRef\]](#)
24. Ye, L.; Peng, D.; Xue, D.; Chen, S.; Shi, A. Co-estimation of lithium-ion battery state-of-charge and state of-health based on fractional-order model. *J. Energy Storage* **2023**, *65*, 107225. [\[CrossRef\]](#)
25. Tian, J.; Xiong, R.; Shen, W.; Lu, J. State of-charge estimation of LiFePO₄ batteries in electric vehicles: A deep-learning enabled approach. *Appl. Energy* **2021**, *291*, 116812. [\[CrossRef\]](#)
26. Zheng, M.; Luo, X. Joint estimation of state of charge (SOC) and state of health (SOH) for lithium-ion batteries using support vector machine (SVM), convolutional neural network (CNN) and long sort term memory network (LSTM) models. *Int. J. Electrochem. Sci.* **2024**, *19*, 100747. [\[CrossRef\]](#)
27. Shi, D.; Zhao, J.; Wang, Z.; Zhao, H.; Eze, C.; Wang, J.; Lian, Y.; Burke, A.F. Cloud-Based Deep Learning for Co-Estimation of Battery State of Charge and State of Health. *Energies* **2023**, *16*, 3855. [\[CrossRef\]](#)
28. Wu, L.; Chen, C.; Li, Z.; Chen, Z.; Li, H. The Joint Estimation of SOC-SOH for Lithium-Ion Batteries Based on BiLSTM-SA. *Electronics* **2025**, *14*, 97.
29. Sang, B.; Wu, Z.; Yang, B.; Wei, J.; Wan, Y. Joint Estimation of SOC and SOH for Lithium-Ion Batteries Based on Dual Adaptive Central Difference H-Infinity Filter. *Energies* **2024**, *17*, 1640. [\[CrossRef\]](#)
30. Cai, X.; Liu, T. State of Health Prediction for Lithium-Ion Batteries Using Transformer-LSTM Fusion Model. *Appl. Sci.* **2025**, *15*, 3747. [\[CrossRef\]](#)
31. Chen, B.; Zhang, Y.; Wu, J.; Yuan, H.; Guo, F. Lithium-Ion Battery State of Health Estimation Based on Feature Reconstruction and Transformer-GRU Parallel Architecture. *Energies* **2025**, *18*, 1236. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.