*Robotica* (2026), **44**, pp. 2160–2176
doi:10.1017/S026357472510297X

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

# Transformer and self-supervised learning-based state-of-charge prediction method for lithium-ion batteries

Wei Zhang[1], Letian Niu[1], Shaojie Yang[2] and Xiang Li[3]

[1]Shenyang Aerospace University, China
[2]University of California San Diego, USA
[3]Xi'an Jiaotong University, China
**Corresponding author:** Xiang Li; Email: lixiang@xjtu.edu.cn

**Received:** 6 July 2025 UTC; **Revised:** 6 September 2025 UTC; **Accepted:** 13 November 2025 UTC;
**First published online:** 21 September 2026 UTC

**Keywords:** lithium-ion battery; state of charge; self-supervised learning; transformer; battery management system

## Abstract

The state-of-charge (SOC) estimation of lithium-ion batteries is essential for the performance, safety, and longevity of electric vehicles. While traditional physics-based models face difficulties in predicting SOC under complex conditions, data-driven methods demand large volumes of labeled data, limiting their practical applicability. This paper introduces a Transformer-based SOC estimation method that incorporates a self-supervised learning framework. By combining a masking recovery mechanism with a self-attention mechanism, the approach effectively captures SOC's long-term dependencies and identifies global patterns under complex operating conditions. Experimental results on the Nissan Leaf and University of Maryland datasets demonstrate significant performance improvements. Compared to traditional deep learning methods, this approach delivers higher prediction accuracy and robustness, offering an innovative solution for efficient SOC estimation in complicated environments.

## 1. Introduction

Li-ion batteries have emerged as the preferred energy solution for new energy electric vehicles (EVs), new energy electric aircraft, and various electric transportation modes due to their high energy density, prolonged lifespan, and low self-discharge rates [1–3]. As these electric transportation options gain global traction, the accurate monitoring and management of lithium-ion battery states have become increasingly crucial. In this context, precise prediction of the state of charge (SOC) not only enhances the energy efficiency and performance of EVs but also significantly aids in trip planning and the optimization of energy management systems. Furthermore, advanced SOC prediction technology provides users with more accurate range estimates, effectively alleviating what is commonly referred to as "range anxiety."

Lithium-ion battery SOC prediction methods mainly include physics-based models and deep learning-based approaches [4–6]. Smith et al. [7] developed an extended Kalman filter (EKF) algorithm based on Thevenin's theorem, accounting for driving modes and external factors, ensuring reliable SOC estimation for EVs. Zhao et al. [8] proposed a method combining EKF and ampere-hour integration to quickly correct sensor faults and parameter errors, enhancing SOC accuracy and battery management reliability. Plett et al. [9] proposed an improved weighted least squares method for battery capacity estimation, addressing noise-related limitations and enabling recursive computation for better embedded system applicability. It preserves robust error bound estimation and goodness-of-fit metrics. Yang et al. [10] introduced a piecewise-fitting EKF algorithm to enhance SOC estimation accuracy, demonstrating strong self-correction across the entire SOC range, with validation through MATLAB simulations and experiments. Sun et al. [11] developed an SOC estimation method using a variable window adaptive Kalman Filter, which dynamically adjusts the error covariance matrix to minimize estimation errors.

© The Author(s), 2026. Published by Cambridge University Press.

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

Zhou et al. [12] introduced an extended Kalman particle filter algorithm for simultaneous SOC and health status estimation, incorporating battery aging and historical data to improve accuracy across the battery's lifecycle. Simulation results show the method's high accuracy and feasibility under the New European Driving Cycle, effectively reducing SOC estimation errors. Buchicchio et al. [13] developed an SOC estimation method using electrochemical impedance spectroscopy (EIS) and equivalent circuit models. This approach effectively differentiates SOC values through EIS measurements, enabling efficient model training and the creation of a low-dimensional SOC classification model for real-time estimation. Despite these advances, challenges persist in dynamic environments where both model precision and data abundance are constrained.

Traditional physics-based model methods rely on the physical and chemical properties of the battery, using mathematical models to describe the battery's state [14–16]. However, these methods struggle to fully capture the changes in battery performance over time and under varying operating conditions in dynamic usage environments. To address this challenge, researchers have increasingly adopted data-driven prediction methods, using historical data and machine learning to uncover patterns in battery SOC variations. With advances in sensor technology and computational power, these methods more effectively capture dynamic changes in battery performance, providing flexible and accurate SOC estimates. Qian et al. [17] developed a one-dimensional convolutional neural network (CNN) method for lithium-ion battery capacity estimation, validated with experimental data. They found that longer or lower initial SOC segments led to more accurate estimates. How et al. [18] introduced an improved deep neural network (DNN)-based SOC prediction model for EVs, demonstrating its effectiveness across different driving cycles. Li et al. [19] proposed a CNN framework combining transfer learning and network pruning for online battery capacity estimation. By applying a pre-trained model to small datasets and pruning the model, they reduced size and complexity, improving efficiency. Zhang et al. [20] introduced an algorithm combining generative adversarial networks (GAN) and bidirectional long short-term memory (LSTM) networks for battery state prediction. They used GAN to generate data from charge/discharge images, training the LSTM model for accurate and efficient battery state forecasting. Chen et al. [21] proposed an LSTM–Self-Attention model combining LSTM and self-attention mechanisms to improve SOC estimation accuracy under varying temperatures and aging levels. It achieved faster convergence and higher accuracy than traditional LSTM models and machine learning methods. Yang et al. [22] introduced a temporal convolution network method for SOC prediction, enhancing accuracy with an attention mechanism. They also developed a capacity degradation model to correct cloud-provided SOC values, achieving more stable and accurate predictions on real datasets compared to other neural network models. Takyi-Aninakwa et al. [23] proposed a multi-layer LSTM network-based SOC estimation method that incorporates battery domain knowledge. First, an adaptive multi-timescale identification strategy and frequency feature decomposition were used to decouple the current and voltage sequences, identifying the dynamic parameters of a lithium iron phosphate battery. Then, these identified parameters were combined with the current and voltage data and fed into an LSTM network for SOC estimation. Finally, an adaptive squared-gain unscented Kalman filter was applied to denoise and optimize the SOC estimate, mitigating the impact of temperature uncertainties. This method significantly enhances the accuracy and robustness of SOC estimation. Gong et al. [24] developed a novel SOC estimation method combining voltage and stress measurements with an LSTM network for accurate prediction of pouch-type lithium-ion battery SOC. Their experiments showed a non-monotonic relationship between SOC and static stress, similar to SOC and open-circuit voltage. Stress was used as supplementary data to enhance accuracy. Hong et al. [25] proposed an SOC model based on a denoising autoencoder, which restores noisy data to its original form, effectively improving SOC prediction accuracy by compensating for data quality issues.

SOC prediction methods typically rely on traditional physics-based models or data-driven approaches, each with limitations in handling complex dynamics, nonlinearities, and missing data [26–28]. Physics-based models capture battery dynamics but struggle with nonlinearities, while data-driven methods require large datasets and fail under long-term dependencies or data loss.

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

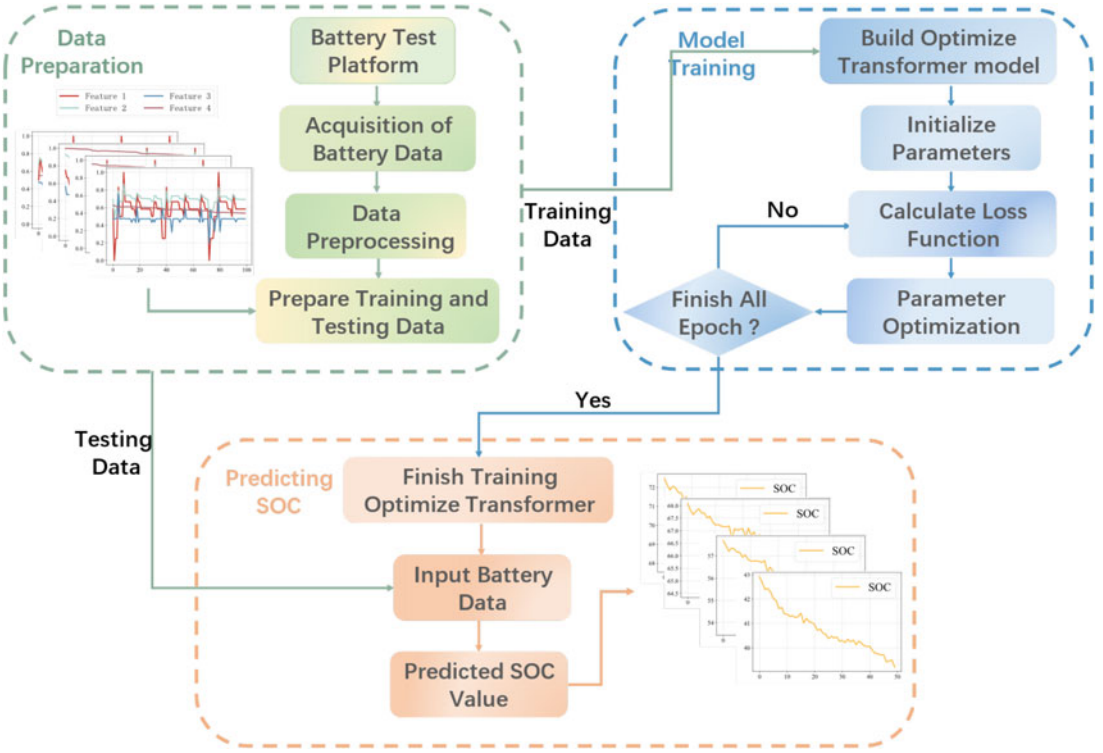



***Figure 1.*** *SOC prediction process diagram of lithium battery.*

To address these challenges, we propose a Transformer-based SOC prediction method with self-attention mechanisms to capture global dependencies. However, existing Transformer or self-supervised methods still face two critical limitations:

1. Computational inefficiency complexity hinders onboard deployment;
2. Poor few-shot adaptation conventional self-supervision generalizes poorly with scarce labels.

The Transformer captures long-term dependencies and adapts to complex dynamics. For missing data, our self-supervised mask-recovery autoencoder ensures robustness. Experiments on Nissan Leaf and University of Maryland (UMD) datasets demonstrate high accuracy under diverse conditions, proving adaptability for real-world applications. This Transformer–self-supervised learning (SSL) fusion advances SOC prediction for battery management and route planning, and the SOC prediction diagram is shown in Figure 1.

## 2. Methodology

This study proposes an SOC prediction method for electric vehicle lithium-ion batteries, combining the Transformer [29–31] model with a mask-recovery self-supervised learning mechanism. The steps include:

1. Data Collection and Preprocessing: Acquire battery data, including parameters such as voltage and current. Process the data through steps like time-windowing, cleaning, and normalization to generate a standardized dataset optimized for training neural networks.
2. Model Construction and Training: Develop a Transformer-based model designed to capture long-term dependencies. Implement a mask-recovery self-supervised learning approach, using

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

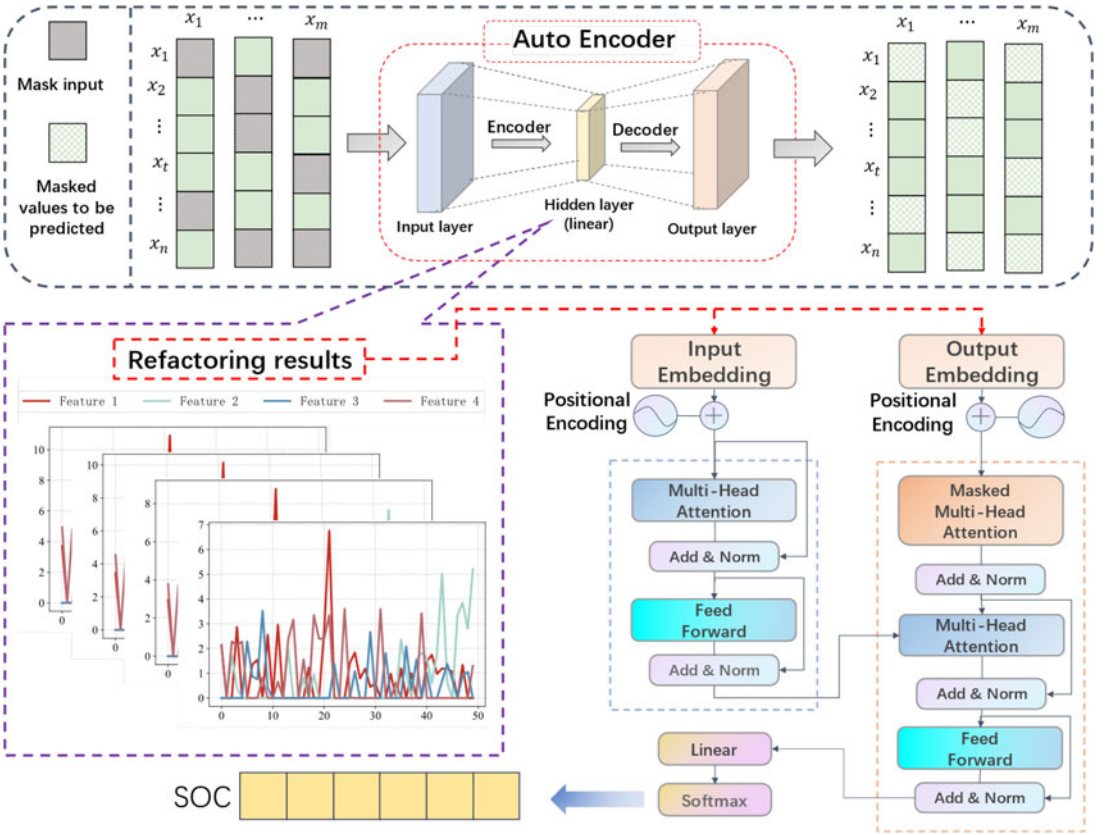



***Figure 2.*** *The structure of the proposed Transformer-based SOC prediction model.*

an autoencoder to mask portions of the input data, enabling the model to learn recovery. This enhances the model's robustness when handling incomplete or noisy data.

3. SOC Prediction and Result Generation: Utilize the trained model to predict battery data, generate SOC degradation curves, and analyze SOC variations across different operating conditions.

This paper proposes a Transformer-based SOC prediction model for lithium batteries, as shown in Figure 2. The model integrates multi-head attention and mask-recovery self-supervised learning with an autoencoder to efficiently handle complex temporal dependencies in long sequences. Key components include position encoding, multi-head attention, feedforward networks, and residual connections. A self-supervised learning mechanism is introduced to handle missing or noisy data by masking parts of the input during training, allowing the model to learn the underlying data structure. This strategy enhances the model's robustness and improves SOC prediction accuracy.

### *2.1. Self-supervised learning*

Traditional supervised learning requires a large amount of labeled data, but manual labeling is costly and time-consuming. Self-supervised learning addresses this by allowing the model to learn from unlabeled data through pretext tasks. This method is especially useful for tasks with limited labeled data, enabling the model to extract useful features from large unlabeled datasets for better initialization in downstream tasks.

This paper proposes a novel SOC prediction method combining self-supervised learning and the Transformer model. We employ a mask-recovery task as the pre-training objective, where the model

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

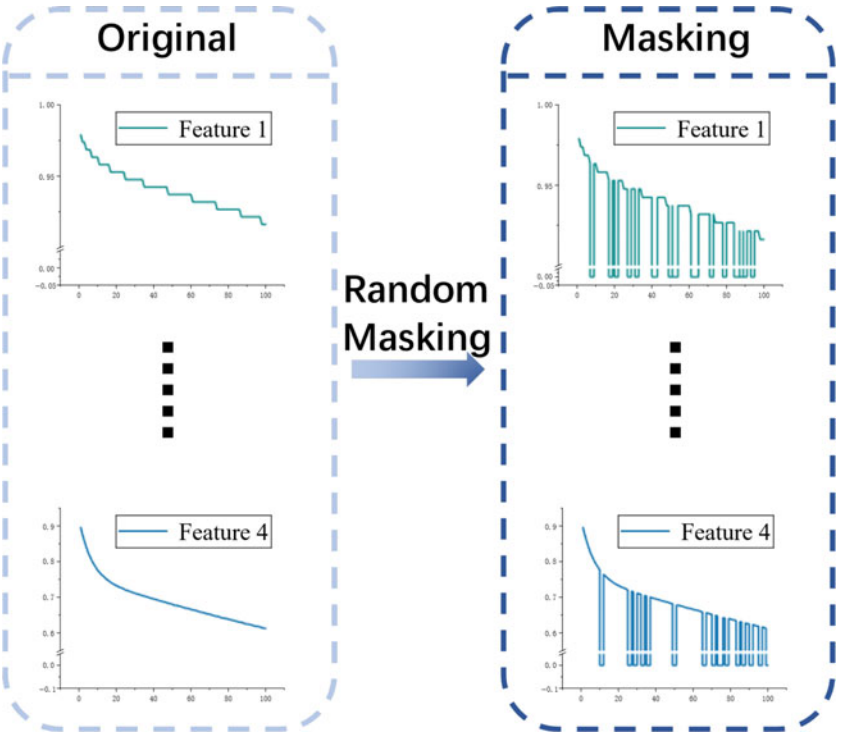



***Figure 3.*** *Random mask of raw data.*

randomly masks part of the data and learns to reconstruct it. This enhances the model's understanding of data patterns and improves robustness to missing or noisy data.

Raw Data Processing: Battery data are recorded 1–6 times per second. For simplicity, the first data point of each second is selected as the battery's state, and every 10 s, a data point is chosen to form a time-consistent dataset with current, voltage, open-circuit voltage, and SOC.

Before pre-training, the data are normalized and preprocessed. Laboratory-collected features, such as voltage, current, and open-circuit voltage, undergo cleaning and preparation to ensure consistency and quality. Each sample consists of $X \in \mathbb{R}^{a\times b}$, where a is the number of features and b is the time series length. At time $t$, the feature vector $x_t$ is normalized as follows:

$$x = \frac{x_{IV} - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

where $x_{IV}$ is the initial battery data and $x_{\max}$ and $x_{\min}$ are the maximum and minimum values for that dimension, respectively.

In self-supervised learning, the model randomly masks parts of the input data, hiding certain features or time steps to prevent direct access to the complete input. The model relies on the unmasked portions to learn and reconstruct the missing data. This enhances the model's ability to uncover data patterns and improves robustness to missing or noisy data. Figure 3 illustrates this transformation from original to masked data.

The normalized input sequence is $X = [x^1, x^2, \ldots, x^t]$, where t is the time series length, and each time point contains multiple features. We randomly select feature dimensions $D_t \in \{1, 2, \ldots, D\}$, where $D$ represents the number of features at each time step. The masking operation is as follows:

$$X_{masked} = \begin{cases} 0, & \text{if } d \in D_t \\ x_t, & \text{otherwise} \end{cases} \tag{2}$$

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

where $X_{masked}$ is the mask value, $x_t$ is the original value at time $t$, and $D_t$ is the set of masked feature dimensions at time $t$.

In pre-training, an autoencoder learns effective data representations by compressing the input into a low-dimensional representation via the encoder and reconstructing it with the decoder. The encoder maps the masked data $X_{masked} \in \mathbb{R}^n$ to a lower-dimensional latent space representation $Z \in \mathbb{R}^m$ denoted as:

$$Z = f_{enc}(X_{masked}; \theta_{enc}) \tag{3}$$

where $f_{enc}$ is the encoder's mapping function and $\theta_{enc}$ are its parameters. The decoder maps the latent representation $Z \in \mathbb{R}^m$ back to the original data space to reconstruct the input, as:

$$\hat{X} = f_{dec}\,(Z; \theta_{dec}) \tag{4}$$

where $f_{dec}$ is the decoder's mapping function and $\theta_{dec}$ are its parameters. The autoencoder's training objective is to minimize the difference between the input and the reconstructed data, with mean squared error (MSE) used as the loss function to measure the reconstruction error:

$$L\,(X, \hat{X}) = ||X - \hat{X}||_2^2 \tag{5}$$

where $||X - \hat{X}||_2^2$ is the Euclidean distance between the input $X$ and the reconstructed data $\hat{X}$.

After pre-training, the encoder becomes a feature extractor. The model has learned the latent features of the input data and is capable of mapping it to a low-dimensional latent space. At this stage, the encoder's role is to generate representative features, which are then passed to the Transformer model for SOC prediction.

### *2.2. Transformer*

In this model, the reconstructed features are input into the encoding layer and combined with positional encoding, which adds position-specific information to each time step's features, enabling the model to capture temporal relationships. It is represented as:

$$\begin{aligned} PE\,(pos, 2i) &= \sin\left(\frac{pos}{10000^{2i/d}}\right) \\ PE\,(pos, 2i+1) &= \cos\left(\frac{pos}{10000^{2i+1/d}}\right) \end{aligned} \tag{6}$$

where $PE(pos, 2i)$ and $PE(pos, 2i+1)$ represent the even and odd positions in the positional encoding matrix, $pos$ is the position, and $i$ is the dimension index. The encoding maps discrete features to a continuous vector space, as:

$$Emb(x) = Linear(x) + PE(x) \tag{7}$$

where $Emb(x)$ is the embedded sequence, $x$ is the input sequence, and $Linear(x)$ is the linear transformation. Positional encoding and embedding provide the model with valuable sequence information, enabling it to better process the data and lay the groundwork for training and prediction.

Self-Attention is a central part of the model, consisting of Query, Key, and Value. The calculation for these components is as follows:

$$Q = \mathrm{X} \cdot W_Q, K = \mathrm{X} \cdot W_K, V = \mathrm{X} \cdot W_V \tag{8}$$

where $W_Q$, $W_K$, and $W_V$ are learnable weight matrices. The attention weight matrix $A$ is calculated by the dot product of the query vector $Q$ and the key vector $K$, followed by scaling and softmax normalization:

$$A = softmax\left(\frac{Q \cdot K^T}{\sqrt{d_k}}\right) \tag{9}$$

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

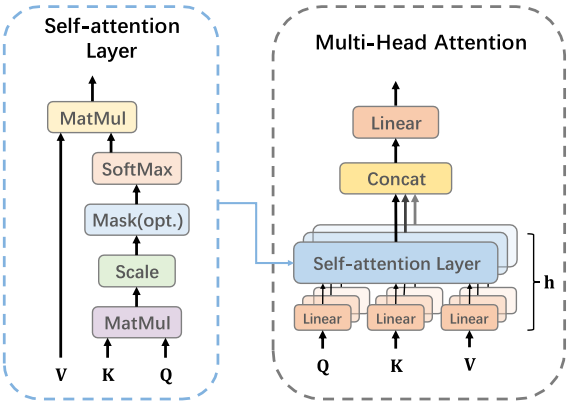



***Figure 4.*** *The multi-head attention mechanism.*

where $d_k$ is the dimension of the key vector. The attention mechanism output $Z$ is obtained by the weighted sum of the value vector $V$ using the attention weight matrix $A$:

$$Z = A \cdot V \tag{10}$$

The input sequence is mapped into h different vector subspaces through linear transformations, producing $h$ sets of $Q_i$, $K_i$, and $V_i$. Each attention head independently computes its attention weights and applies them to the sequence, resulting in independent attention outputs. Multi-head attention as shown in Figure 4, improves the model's ability to capture complex relationships in the data.

Feedforward Network: Comprising two fully connected layers, the feedforward network follows the multi-head attention mechanism and applies further nonlinear transformations to enhance the model's expressive power.

Residual Connection: The input is directly added to the output, helping prevent the vanishing gradient problem and improving training performance.

Loss Function: To evaluate the proposed method, we use MSE as the loss function, defined as:

$$L = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 \tag{11}$$

where $n$ is the number of samples, $y_i$ is the true value, and $\hat{y}_i$ is the predicted value.

Optimizer: The Adam optimizer is used to update model parameters, combining momentum and adaptive learning rates for faster convergence during training.

The model is trained over multiple iterations with masked input data. Loss is calculated, parameters updated, and performance evaluated using a validation set to fine-tune hyperparameters for optimal results.

## 3. Experiment and result

### *3.1. Datasets*

There are two datasets used in this paper to assess the model's generalization. Dataset A is the INR 18650-20R battery dataset from the University of Maryland, which includes four typical driving conditions [32]:

DST (Dynamic Stress Test): Evaluates extreme condition performance (3C rapid charge/discharge, 45°C ambient)

FUDS (Federal Urban Driving Schedule): Characterizes urban driving patterns with frequent <2C load variations

US06 (Supplemental FTP Cycle): Simulates highway driving with sustained >2C discharge rates

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

***Table I.*** *Main battery parameters.*

| **Manufacturer** | **INR 18650-20R** | **Nissan Leaf** |
|---|---|---|
| Nominal capacity | 2Ah | 56Ah |
| Electrode active materials (cathode/anode) | LiNiMnCo/graphite | $LiMn_2O_4$ with $LiNiO_2$/graphite |
| Voltage range | 2.5–4.2V | 2.6–4.2V |
| SOC range | 0–80% | 0–96% |

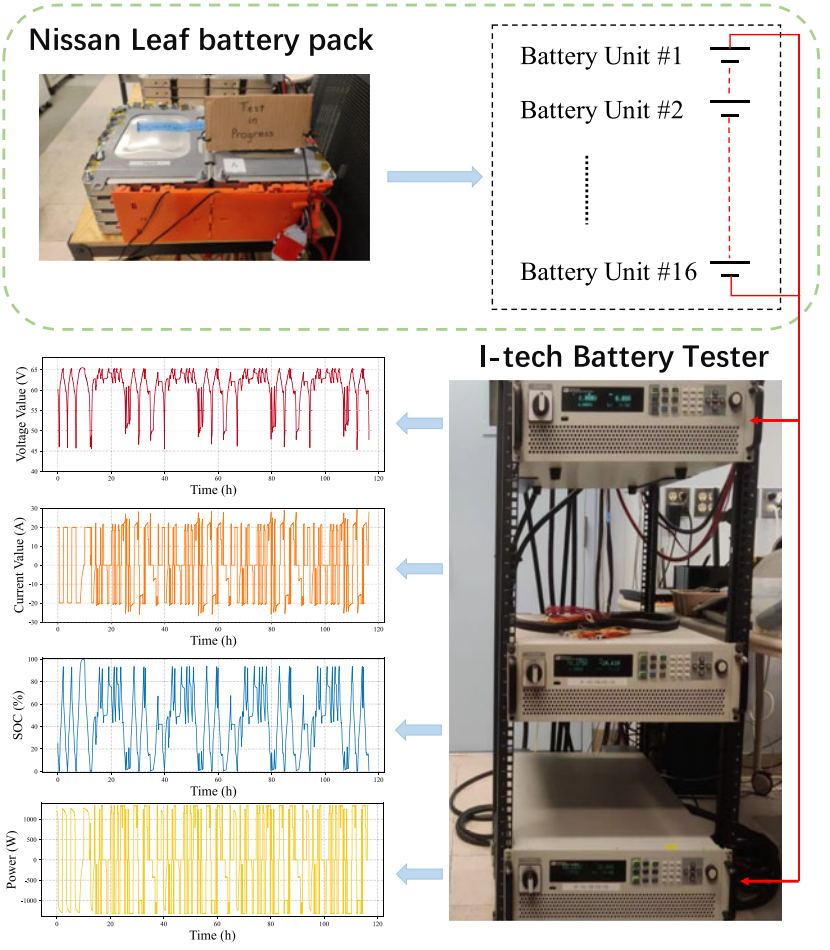



***Figure 5.*** *Schematic of battery testing with Nissan Leaf.*

BJDST (Beijing Driving Schedule Test): Represents composite urban–rural driving with stochastic 0.5–2.5C load fluctuations

Dataset B is established by collecting data from Nissan Leaf batteries. The experimental setup is shown in Figure 5. Retired Nissan Leaf batteries were used to collect charge and discharge data. The setup includes a battery pack of 16 cells, with the I-tech tester recording key parameters such as voltage, current, and temperature. To assess the model's generalization, the publicly available INR 18650-20R dataset from the University of Maryland was also used for comparison. Table I lists the main parameters of both battery types.

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

The charge and discharge cycles followed a 7-day cycle, adjusted using the Day-Ahead and Real-Time Energy Timing Shift [33]. The charging strategy was dynamically modified based on day-ahead and real-time energy demand fluctuations to simulate real-world usage. This approach ensured the experimental data's relevance and applicability.

### *3.2. Comparative methods*

An Abstract is required for every paper; it should succinctly summarize the reason for the work. This study applied various SOC prediction methods to highlight the effectiveness and superiority of the proposed approach. The methods used were:

1. Deep neural networks (DNN)

This study uses a DNN with four hidden layers to predict the battery SOC, with a single neuron at the output for SOC prediction.

2. Convolutional neural networks (CNN)

A CNN is used for SOC prediction, consisting of three convolutional layers with 32, 64, and 128 filters, two fully connected layers with 256 and 128 neurons, and three pooling layers with a pool size of 2.

3. Long short-term memory network model

An LSTM model with two 128-neuron layers and a fully connected layer is used for SOC prediction, leveraging its ability to capture long-term dependencies in time series data.

4. Transformer

A Transformer-based model, comprising an input layer, position encoding, multi-head attention, feedforward network, and output layer, is designed to predict the battery SOC trend accurately.

### *3.3. Nissan Leaf battery data tests*

To ensure accurate and reliable predictions, results are averaged over multiple trials to reduce randomness. The study uses five tasks to evaluate different methods on the Nissan Leaf lithium-ion battery dataset. Each task involves 10 batteries, with 20 discharge segments in total. Four segments are randomly selected for testing, and the rest are used for training. Tasks T1 through T5 independently assess the methods' performance across various battery discharge stages. Each task undergoes five independent trials per battery, with results averaged for consistency and reliability.

Figure 6 illustrates that the lithium-ion battery SOC prediction model proposed in this study outperforms the DNN, CNN, LSTM, and Transformer models in terms of prediction accuracy. In the red box, the proposed method achieves the smallest prediction error, with a root mean square error (RMSE) of 0.0012, whereas the RMSEs for the other methods are 0.0011, 0.0064, 0.0026, and 0.0016, respectively. In the blue box, the proposed method shows the largest prediction error at 0.0042, while the RMSEs for the other models are 0.0072, 0.0044, 0.0060, and 0.0069. Overall, the proposed model demonstrates superior accuracy in SOC prediction, outperforming the other methods in both the smallest and largest error cases.

Figure 7 presents the comparison of different networks in the T4-2 task. As shown in the red box, the proposed method achieves the smallest prediction error, with an RMSE of just 0.0008, significantly better than the other four methods, whose RMSEs are 0.0097, 0.0098, 0.0055, and 0.0030, respectively. In the blue box, although the proposed method exhibits the highest prediction error with an RMSE of 0.0035, it still outperforms all models except for the Transformer, whose RMSE is 0.0041. The RMSEs for the other methods are 0.0063, 0.0083, and 0.0127, respectively. This further confirms the superiority of the proposed method for SOC prediction.

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

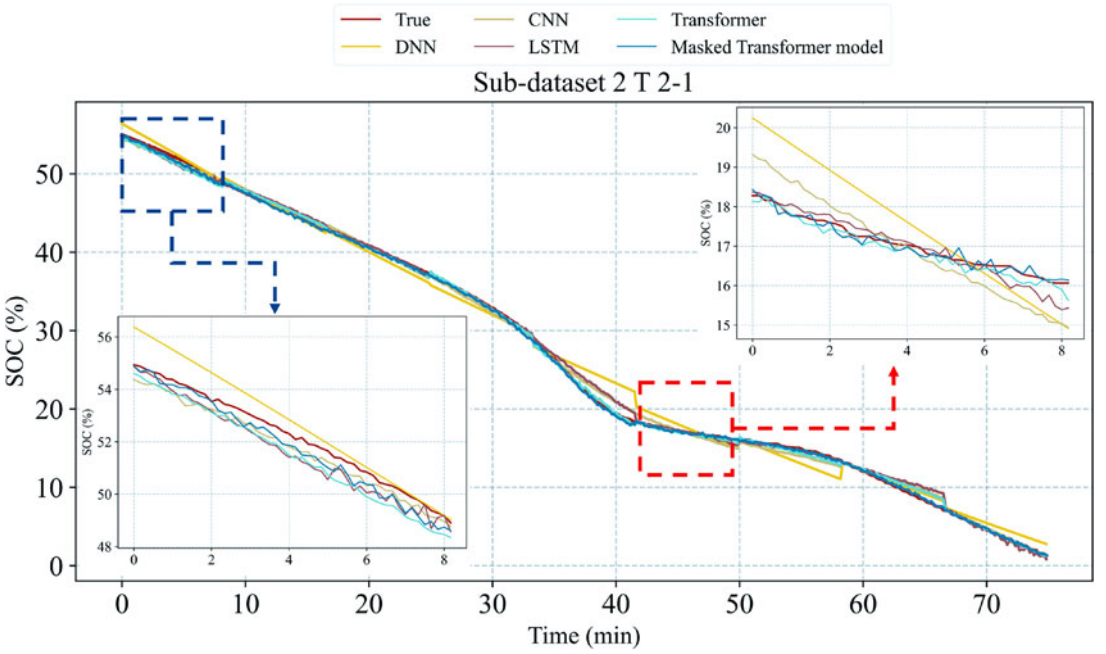



***Figure 6.*** *Comparison of prediction results from different networks in task T2-1.*

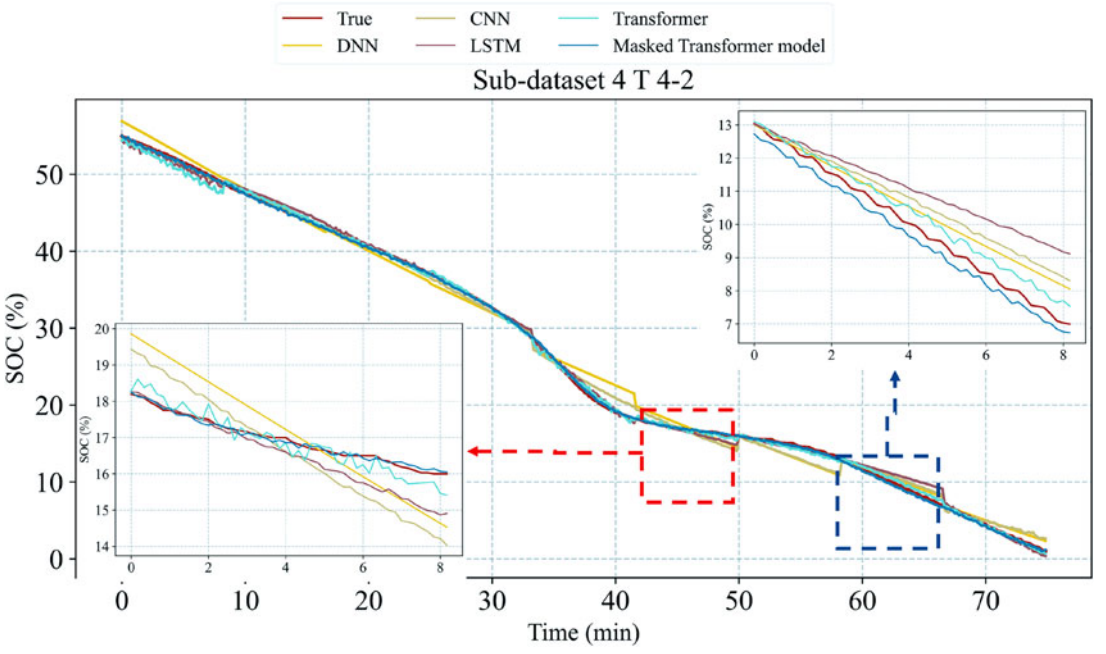



***Figure 7.*** *Comparison of prediction results from different networks in task T4-2.*

Figure 8 highlights the model's performance in SOC prediction. In the T5-4 task, the proposed method achieves the smallest prediction error, with an RMSE of 0.0017, significantly outperforming the other four methods, whose RMSEs are 0.012, 0.0067, 0.0135, and 0.0038, respectively, as shown in the red box. In the blue box, although the proposed method reaches the highest prediction error, it still

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

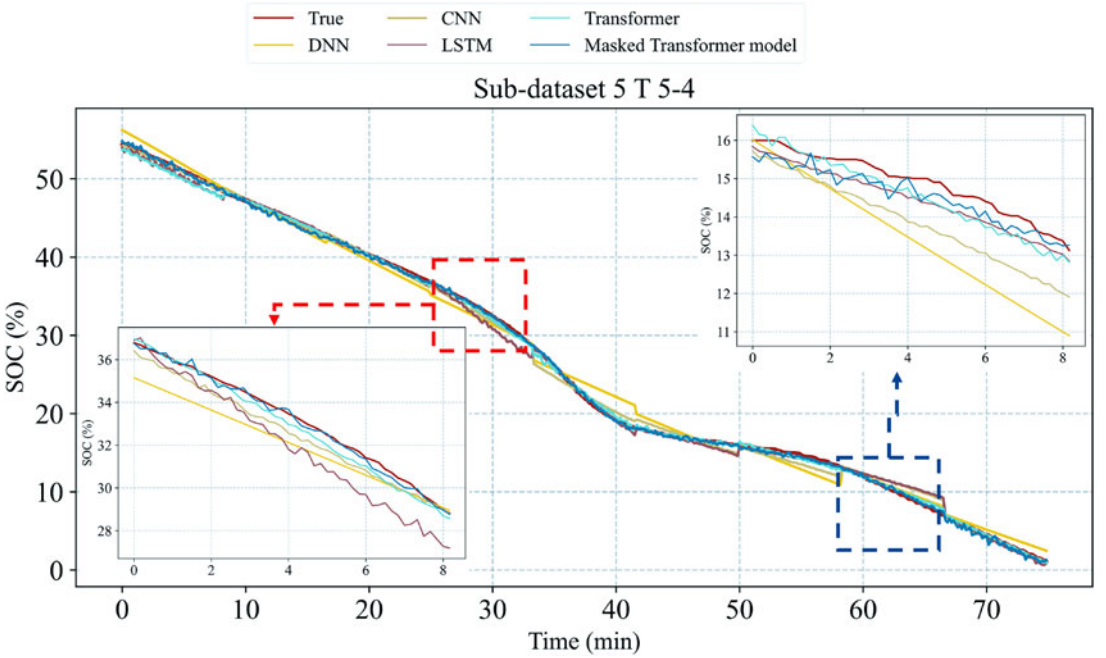



***Figure 8.*** *Comparison of prediction results from different networks in task T5-4.*

outperforms the other methods, whose RMSEs are 0.0162, 0.0107, 0.0045, and 0.0041. These results further confirm the effectiveness and superiority of the proposed method in SOC prediction.

Figure 9 illustrates the RMSE comparison across different tasks. Except for tasks T1-3, T2-4, T3-3, and T4-1, where the RMSE is slightly higher than that of the Transformer model, the proposed method outperforms the DNN, CNN, LSTM, and Transformer models in the remaining 16 tasks. These findings further highlight the superior prediction performance of the proposed method in the majority of tasks.

In summary, the proposed model demonstrates superior performance compared to baseline methods across all evaluation metrics, with consistently lower RMSE values indicating enhanced prediction accuracy. These results suggest that our Transformer-based SOC estimation approach represents a robust solution for the target application scenarios. However, the prediction accuracy exhibits relative degradation within the 10–30% SOC range, highlighting a critical challenge for future research to address. This limitation may be attributed to the characteristic voltage plateau effects in lithium-ion batteries during low-charge states.

### *3.4. INR 18650-20R battery data tests*

This section defines four tasks to assess the performance of various methods on the INR 18650-20R lithium-ion battery dataset. Each task consists of six discharge data segments: three with SOC ranges from 0% to 80%, and three with SOC ranges from 0% to 50%. For each task, two segments from the 0–80% SOC range and two from the 0–50% SOC range are used for training, while the remaining data are used for testing. By independently predicting using test data from different operating conditions, the performance and predictive capabilities of the methods on this dataset are thoroughly evaluated. To better illustrate the results, we selected four different sub-intervals from the entire discharge curve for each operating condition and plotted them in four separate graphs. Each graph displays the performance of different models within the corresponding interval, allowing for a clearer comparison of the RMSE values across models in different intervals.

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

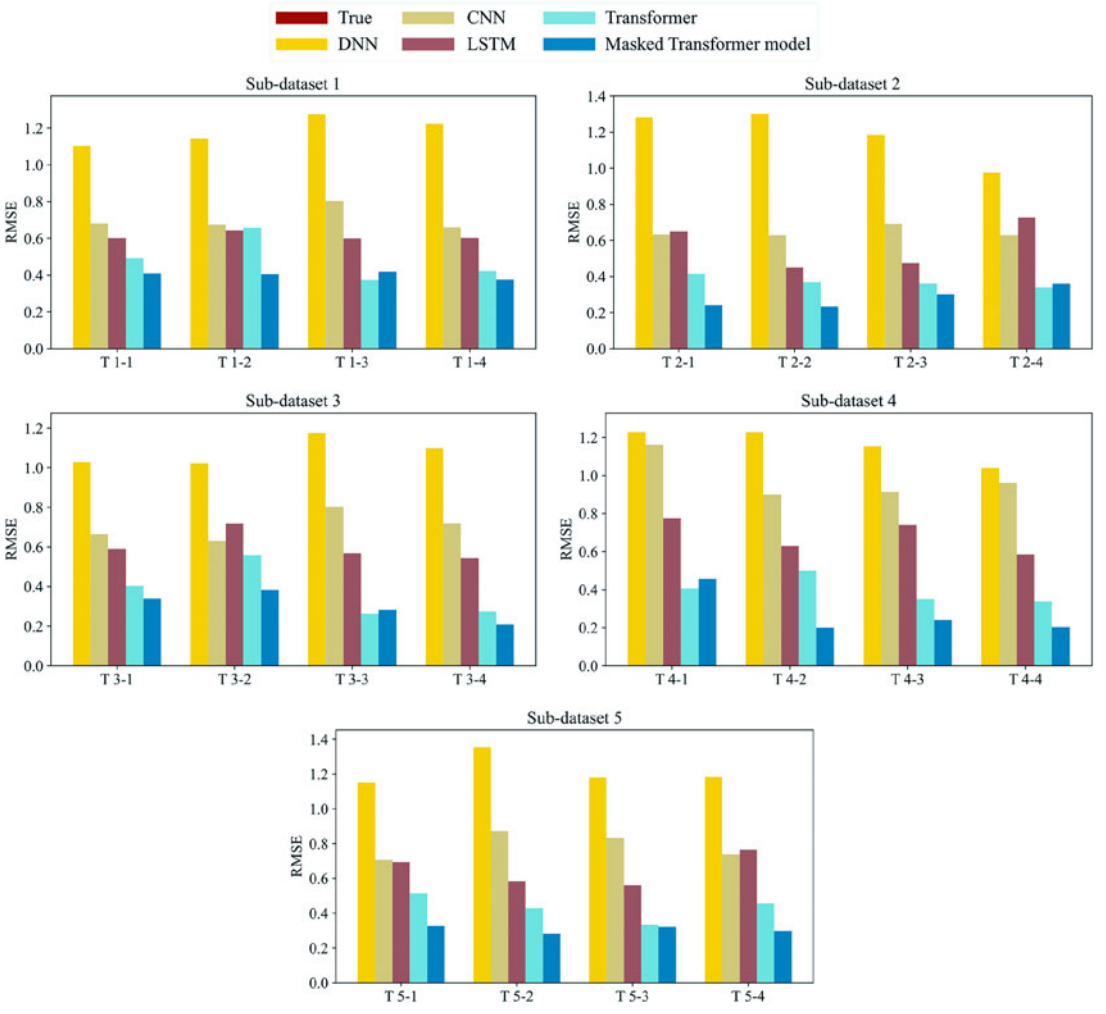



***Figure 9.*** *RMSE of SOC prediction results under different tasks.*

The BJDST condition simulates the battery's charging and discharging in real-world driving, covering a range of scenarios like low-speed cruising and dynamic events such as rapid acceleration and braking. This makes it an ideal case for testing SOC prediction models under diverse load conditions. As shown in Figure 10, the proposed method achieves the smallest RMSE of 0.0010, outperforming other models with RMSEs of 0.0045, 0.0025, 0.0035, and 0.0031. Even at its largest RMSE of 0.0028, it still outperforms the other models, which have RMSEs of 0.0072, 0.0035, 0.0038, and 0.0054.

The DST condition simulates battery performance under high-load, fast charging, and discharging conditions to assess its stability and reliability in extreme operations. It involves frequent high-power cycles, leading to significant SOC changes. As shown in Figure 11, the proposed method achieves the smallest RMSE of 0.0007, outperforming the other models with RMSEs of 0.0029, 0.0061, 0.0030, and 0.0027. In contrast, the proposed method's maximum RMSE of 0.0035 is better than the others, with RMSEs of 0.0080, 0.0041, 0.0041, and 0.0031. These results confirm the proposed method's exceptional performance under extreme conditions.

The US06 condition simulates highway driving and evaluates the battery's performance under high-speed driving and aggressive acceleration. These conditions cause significant load changes and rapid SOC fluctuations, requiring precise prediction models. As shown in Figure 12, the proposed method achieves the lowest RMSE of 0.0011, outperforming the other models with RMSEs of 0.0021, 0.0059,

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

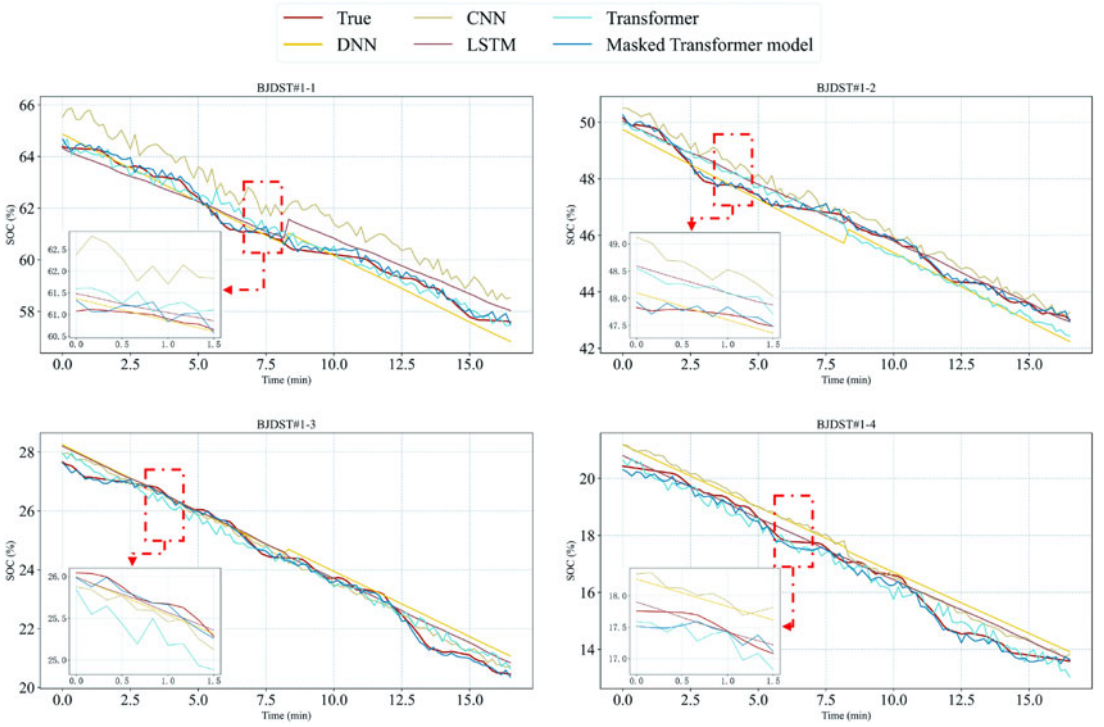



***Figure 10.*** *SOC prediction results under BJDST working condition.*

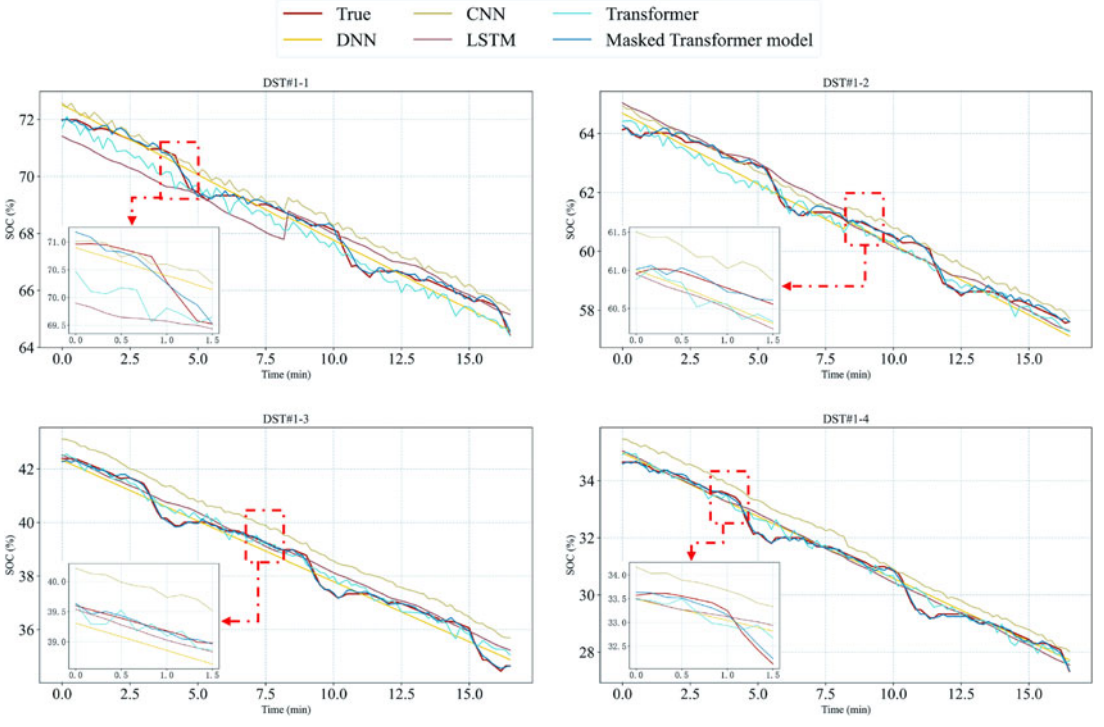



***Figure 11.*** *SOC prediction results under DST conditions.*

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

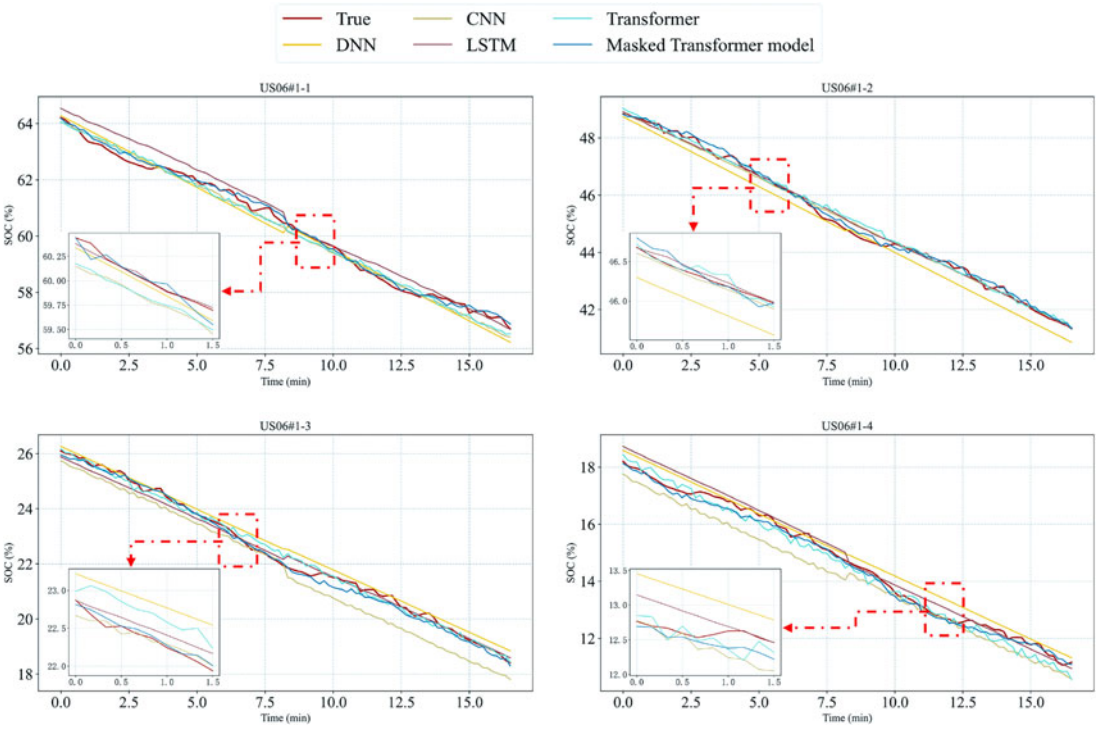



***Figure 12.*** *SOC prediction results under US06 conditions.*

0.0025, and 0.0021. Even at its maximum RMSE of 0.0019, it outperforms the others, with RMSEs of 0.0030, 0.0042, 0.0068, and 0.0029. These results validate the proposed method's superiority and reliability under high-load conditions.

The FUDS condition simulates urban driving, mainly testing the battery's SOC changes during frequent acceleration, braking, and low-speed driving. These conditions cause frequent load fluctuations, requiring SOC prediction models to respond dynamically to rapid changes. As shown in Figure 13, the proposed method achieves the smallest RMSE of 0.0011, outperforming the other models with RMSEs of 0.0040, 0.0094, 0.0058, and 0.0037. Even at its maximum RMSE of 0.0031, it still outperforms the other models, which have RMSEs of 0.0101, 0.0065, 0.0078, and 0.0063. These results confirm the proposed method's superiority and stability in urban driving conditions with frequent load changes.

As shown in Figure 14, the RMSE values for various methods under different conditions are compared. The self-supervised model proposed in this paper outperforms the other four models in all conditions, consistently achieving lower RMSE values. It demonstrates the highest prediction accuracy, particularly in complex load and extreme conditions, effectively responding to SOC fluctuations and delivering precise predictions. In comparison, the DNN and CNN models exhibit higher RMSE values, indicating suboptimal performance. Although the Transformer and LSTM models perform well, they are still outperformed by the proposed self-supervised model. These results confirm the exceptional performance of the self-supervised model, proving its effectiveness and reliability in battery SOC prediction.

## 4. Conclusion

The self-supervised SOC prediction model proposed in this study performs outstandingly under various conditions in both the Nissan Leaf lithium-ion battery dataset and the INR 18650-20R battery dataset, offering significant advantages over traditional methods such as DNN, CNN, LSTM, and Transformer.

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

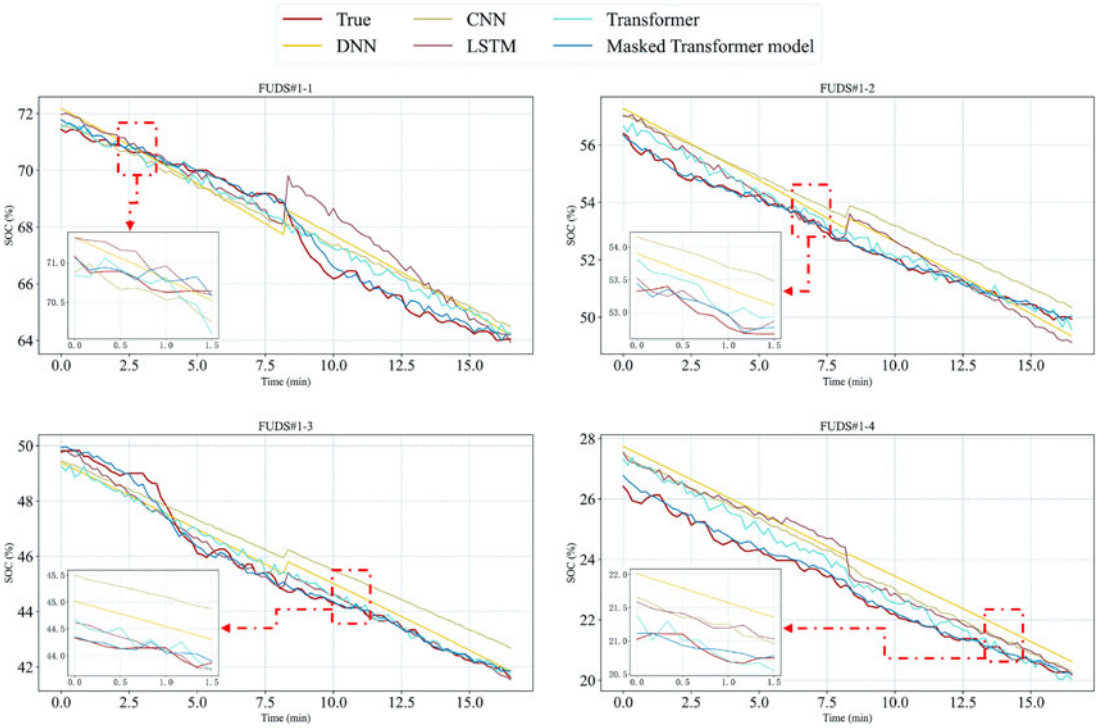



***Figure 13.*** *SOC prediction results under FUDS conditions.*

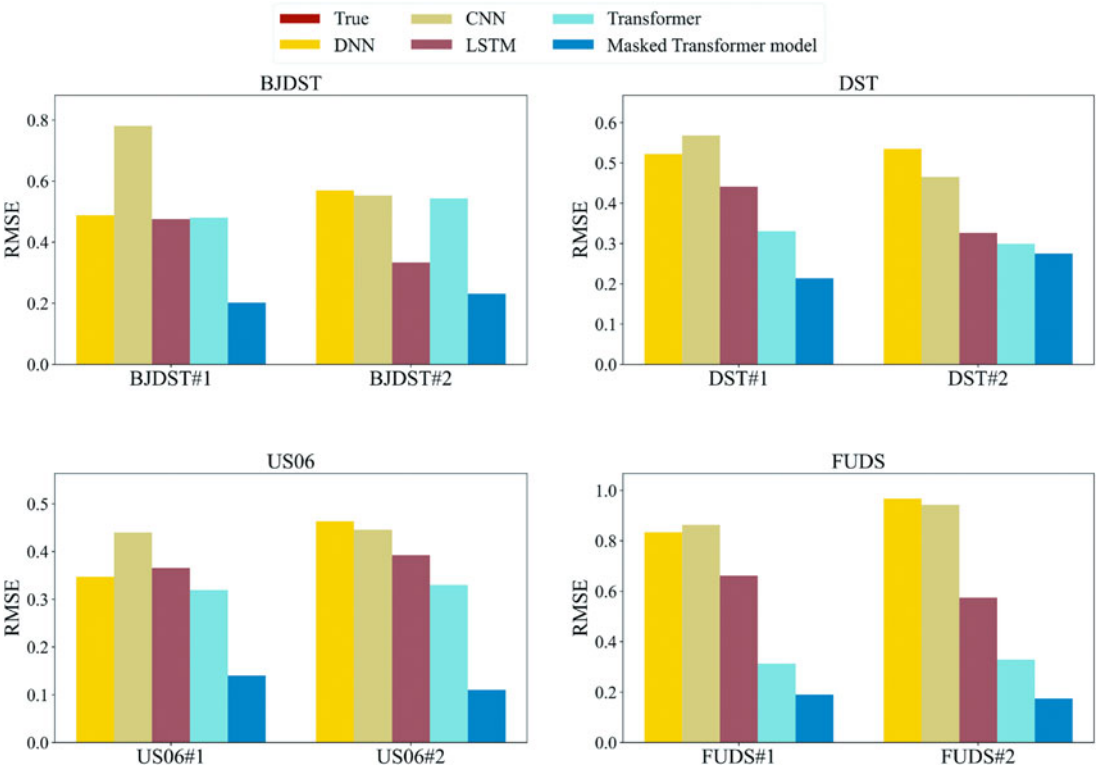



***Figure 14.*** *RMSE of SOC prediction results under different working conditions.*

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

In tasks involving the Nissan Leaf lithium-ion battery dataset, the proposed method consistently achieves low RMSE, demonstrating superior prediction accuracy. In both minimum and maximum error cases, the proposed method outperforms the other models, particularly in handling complex SOC variations, offering precise predictions.

Similarly, in the INR 18650-20R battery dataset, the proposed model excels under all four test conditions. It consistently maintains low RMSE, particularly under complex and extreme conditions, effectively predicting SOC changes and demonstrating its robustness and stability in diverse scenarios.

Overall, the proposed method demonstrates superior predictive capabilities across a range of conditions, confirming its effectiveness and reliability for battery SOC prediction.

**Author contribution.** Wei Zhang contributed to conceptualization, methodology, and formal analysis and provided supervision; Letian Niu developed the software, performed validation and formal analysis, and wrote the original draft; Shaojie Yang was responsible for data curation, investigation, and resources; Xiang Li handled writing review and editing, project administration, and funding acquisition.

**Financial support.** The material in this paper is based on work supported by the Ministry of Education of China under Grant (HKZY20220429).

**Competing interests.** The authors declare no conflicts of interest.

**Ethical standards.** Not applicable.

## References

[1] W. Zhang, N. Jiang, S. Yang and X. Li, "Federated transfer learning for remaining useful life prediction in prognostics with data privacy," *Measur. Sci. Technol.* **36**, 076107 (2025).
[2] C. Liu, X. Li, X. Chen and S. Khan, "Neuromorphic computing-enabled generalized machine fault diagnosis with dynamic vision," *Adv. Eng. Inform.* **65**, 103300 (2025).
[3] W. Zhang, M. Xu, H. Yang, X. Wang, S. Zheng and X. Li, "Data-driven deep learning approach for thrust prediction of solid rocket motors," *Measurement* **225**, 114051 (2024).
[4] W. Zhang, Z. Wang and X. Li, "Blockchain-based decentralized federated transfer learning methodology for collaborative machinery fault diagnosis," *Reliab. Eng. Syst. Saf.* **229**, 108885 (2023).
[5] X. Li, C. Zhang, X. Li and W. Zhang, "Federated transfer learning in fault diagnosis under data privacy with target self-adaptation," *J. Manuf. Syst.* **68**, 523–535 (2023).
[6] W. Zhang, X. Li, H. Ma, Z. Luo and X. Li, "Transfer learning using deep representation regularization in remaining useful life prediction across operating conditions," *Reliab. Eng. Syst. Saf.* **211**, 107556 (2021).
[7] E. P. Sangeetha, N. Subashini, T. K. Santhosh, S. Augusti Lindiya and D. Uma, "Validation of EKF based SoC estimation using vehicle dynamic modelling for range prediction," *Electr. Power Syst. Res.* **226**, 109905 (2024).
[8] X. Zhao, B. Sun, W. Zhang, X. He, S. Ma, J. Zhang and X. Liu, "Error theory study on EKF-based SOC and effective error estimation strategy for Li-ion batteries," *Appl. Energ.* **353**, 121992 (2024).
[9] G. L. Plett, "Recursive approximate weighted total least squares estimation of battery cell total capacity," *J. Power Sour.* **196**(4), 2319–2331 (2011).
[10] B. Yang, G. Li, W. Tang and H. Li, "Research on optimized SOC estimation algorithm based on extended Kalman filter," *Front. Energy Res.* **10**, 1027439 (2022).
[11] D. Sun, X. Yu, C. Wang, C. Zhang, R. Huang, Q. Zhou, T. Amietszajew and R. Bhagat, "State of charge estimation for lithium-ion battery based on an Intelligent Adaptive Extended Kalman Filter with improved noise estimator," *Energy* **214**, 119025 (2021).
[12] H. Zhou, J. Luo and Z. Yu, "Co-estimation of SOC and SOH for Li-ion battery based on MIEKPF-EKPF fusion algorithm," *Energy Rep.* **10**, 4420–4428 (2023).
[13] E. Buchicchio, A. De Angelis, F. Santoni, P. Carbone, F. Bianconi and F. Smeraldi, "Battery SOC estimation from EIS data based on machine learning and equivalent circuit model," *Energy* **283**, 128461 (2023).
[14] B. Yang, Y. Lei, X. Li, N. Li, X. Si and C. Chen, "A dynamic barycenter bridging network for federated transfer fault diagnosis in machine groups," *Mech. Syst. Signal Process.* **230**, 112605 (2025).
[15] Z. Wang, B. He, Y. Zhou, K. Liu and C. Zhang, "Design and implementation of a cable inspection robot for cable-stayed bridges," *Robotica* **39**(8), 1417–1433 (2021).
[16] W. Zhang and X. Li, "Federated transfer learning for intelligent fault diagnostics using deep adversarial networks with data privacy," *IEEE/ASME Trans. Mechatron.* **27**(1), 430–439 (2021).
[17] C. Qian, B. Xu, L. Chang, B. Sun, Q. Feng, D. Yang, Y. Ren and Z. Wang, "Convolutional neural network based capacity estimation using random segments of the charging curves for lithium-ion batteries," *Energy* **227**, 120333 (2021).

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press

[18] D. N. T. How, M. A. Hannan, M. S. H. Lipu, K. S. M. Sahari, P. J. Ker and K. M. Muttaqi, "State-of-charge estimation of Li-ion battery in electric vehicles: A deep neural network approach," *IEEE Trans. Ind. Appl.* **56**(5), 5565–5574 (2020).
[19] Y. Li, K. Li, X. Liu, Y. Wang and L. Zhang, "Lithium-ion battery capacity estimation—A pruned convolutional neural network approach assisted with transfer learning," *Appl. Energy* **285**, 116410 (2021).
[20] H. Zhang, W. Tang, W. Na, P.-Y. Lee and J. Kim, "Implementation of generative adversarial network-CLS combined with bidirectional long short-term memory for lithium-ion battery state prediction," *J. Energy Stor.* **31**, 101489 (2020).
[21] G. Chen, W. Peng and F. Yang, "An LSTM-SA model for SOC estimation of lithium-ion batteries under various temperatures and aging levels," *J. Energy Stor.* **84**, 110906 (2024).
[22] X. Yang, J. Hu, G. Hu and X. Guo, "Battery state of charge estimation using temporal convolutional network based on electric vehicles operating data," *J. Energy Stor.* **55**, 105820 (2022).
[23] P. Takyi-Aninakwa, S. Wang, G. Liu, A. N. Bage, E. D. Bobobee, E. Appiah and Q. Huang, "Enhanced extended-input LSTM with an adaptive singular value decomposition UKF for LIB SOC estimation using full-cycle current rate and temperature data," *Appl. Energy* **363**, 123056 (2024).
[24] L. Gong, Z. Zhang, Y. Li, X. Li, K. Sun and P. Tan, "Voltage-stress-based state of charge estimation of pouch lithium-ion batteries using a long short-term memory network," *J. Energy Stor.* **55**, 105720 (2022).
[25] S. Hong, M. Kang, J. Kim and J. Baek, "Investigation of denoising autoencoder-based deep learning model in noise-riding experimental data for reliable state-of-charge estimation," *J. Energy Stor.* **72**, 108421 (2023).
[26] X. Li, S. Yu, Y. Lei, N. Li and B. Yang, "Dynamic vision-based machinery fault diagnosis with cross-modality feature alignment," *IEEE/CAA J. Autom. Sin.* **11**(10), 2068–2081 (2024).
[27] W. Zhang and X. Li, "Data privacy preserving federated transfer learning in machinery fault diagnostics using prior distributions," *Struct. Health Monit.* **21**(4), 1329–1344 (2022).
[28] W. Zhang, X. Li and X. Li, "Deep learning-based prognostic approach for lithium-ion batteries with adaptive time-series prediction and on-line validation," *Measurement* **164**, 108052 (2020).
[29] J. W. Chan and C. K. Yeo, "A transformer based approach to electricity load forecasting," *The Electr. J.* **37**(2), 107370 (2024).
[30] S. Hazmoune and F. Bougamouza, "Using transformers for multimodal emotion recognition: Taxonomies and state of the art review," *Eng. Appl. Artif. Intell.* **133**, 108339 (2024).
[31] X. Guo, X. Lin, X. Yang, L. Yu, K.-T. Cheng and Z. Yan, "UCTNet: Uncertainty-guided CNN-transformer hybrid networks for medical image segmentation," *Pattern Recog.* **152**, 110491 (2024).
[32] INR 18650-20R battery testing data, " Univ. Maryland," (2021) [Online]. Available: https://calce.umd.edu/battery-data
[33] S. Yang, L. Ling, X. Li, J. Han and S. Tong, "Industrial battery state-of-health estimation with incomplete limited data toward second-life applications," *J. Dyn. Monitor. Diag.* **3**(4), 246–257 (2024).

**Cite this article:** W. Zhang, L. Niu, S. Yang and X. Li (2026). "Transformer and self-supervised learning-based state-of-charge prediction method for lithium-ion batteries", Robotica **44**, 2160–2176. https://doi.org/10.1017/S026357472510297X

https://doi.org/10.1017/S026357472510297X Published online by Cambridge University Press