

Article

Dual-Branch Bidirectional Long Short-Term Memory Network for Battery State of Health Estimation Under Incomplete Data

Le Ke , Xiangbo Zhang and Lujuan Dang * 

State Key Laboratory of Human-Machine Hybrid Augmented Intelligence, Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China; 24kele@stu.xjtu.edu.cn (L.K.); xiangbo@stu.xjtu.edu.cn (X.Z.)

* Correspondence: danglj@xjtu.edu.cn

Abstract

Accurate estimation of the state of health (SOH) of lithium-ion batteries is crucial for battery management systems (BMS), as it directly affects system safety, operational reliability, and remaining useful life prediction. Traditional SOH estimation methods generally rely on complete charge–discharge cycle data, which limits their practicality in real-time and online applications. To address this limitation, this paper proposes a novel voltage-charge increment curve-based dual-branch bidirectional long short-term memory network (VCIC-DB-BiLSTM) for battery SOH estimation under incomplete data. First, non-uniformly sampled battery current-voltage data are processed into standardized sequences with equal voltage intervals via voltage-charge increment curves based on ampere-hour integration and cubic spline interpolation. Subsequently, sliding window segmentation is applied to extract fixed-length curve segments from continuous voltage intervals as input features, while the corresponding complete voltage interval curves are used as labels. Finally, the VCIC-DB-BiLSTM network is designed, which uses a dual-branch structure to integrate feature extraction from both patch-processed and raw data, combined with bidirectional sequential modeling. Experimental validation on four benchmark datasets, CALCE, Oxford, XJTU, and TJU, demonstrates that the proposed method achieves competitive performance in SOH estimation under incomplete discharge data conditions, confirming its effectiveness and practical applicability.

Keywords: state-of-health (SOH); incomplete data; bidirectional long short-term memory (BiLSTM)



Academic Editor: JongHoon Kim

Received: 11 May 2026

Revised: 9 July 2026

Accepted: 16 July 2026

Published: 20 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Lithium-ion batteries, known for their high voltage, high energy density, and low self-discharge, have become the standard power source for electric vehicles, energy storage systems, and portable electronics [1–3]. As these batteries are deployed in increasingly demanding applications, monitoring their degradation becomes paramount. The state of health (SOH) is a key metric for this purpose. Accurately estimating SOH is essential for battery management systems (BMS) to ensure operational safety, reliability, and longevity. By identifying performance loss, SOH estimation contributes to preventing sudden failures and informs optimal maintenance or replacement schedules, thereby extending the useful life of battery systems [4,5]. Generally, SOH is most commonly defined as the ratio of the battery's present maximum available capacity to its original rated capacity [6]:

$$SOH = \frac{Q_{\max}}{Q_{\text{nom}}}, \quad (1)$$

where $Q_{\max}$ is the current maximum available capacity of the battery, and Q_{nom} is its original nominal capacity.

Traditional methods for estimating the SOH of batteries can be broadly categorized into resistance-based and capacity-based approaches. While resistance-based methods, such as internal resistance tracking or electrochemical impedance spectroscopy (EIS), do not necessarily require full cycles, they often demand specific operational conditions, stable environments, or specialized measurement equipment [7,8]. On the other hand, traditional capacity-based estimation methods typically require data from complete charge–discharge cycles [9]. However, in practical applications such as electric vehicles and smartphones, obtaining such full-cycle data is challenging due to time constraints and the complex operating conditions. Operational data are often fragmented, and battery cycles can begin or terminate at arbitrary points [10]. Taking smartphone batteries as an example of practical implementation, traditional SOH estimation typically relies on Coulomb counting (using a coulometer) or pre-calibrated discharge curve matching [11]. However, Coulomb counting requires periodic full charge–discharge cycles to recalibrate and eliminate accumulated integration errors, which fundamentally contradicts modern users' fragmented charging habits [12]. Similarly, traditional discharge curve matching assumes stable and continuous operational profiles, whereas actual smartphone usage involves highly dynamic and unpredictable loads [13]. Therefore, methods requiring full-cycle data or specific continuous conditions face significant implementation bottlenecks in such real-world applications [14].

To address this challenge, research has shifted toward utilizing incomplete cycle data for SOH estimation. Some studies have also applied such data to voltage prediction.

Bavand et al. [15] pioneered a rapid online SOH estimation technique capable of functioning with partial charge/discharge cycles. This innovative method employs only two consecutive partial discharge intervals to estimate battery parameters and open-circuit voltage (OCV). Xiong et al. [16] explored a method that leverages data collected from a fixed starting state of charge (SOC) up to 100% for predicting SOH of lithium-ion batteries.

Chen et al. [17] proposed an estimation method of SOH for lithium ion batteries based on data from the partial constant voltage (CV) charging phase and a convolutional neural network (CNN). The proposed method requires only the first 1000 s of data from the CV charging phase to achieve a highly accurate estimate of SOH. Gao et al. [18] use the incremental capacity (IC) curve in the partial charging data to extract aging features, and establish a linear aging state space model based on these features. Despite these advances, a common limitation of the aforementioned methods is their reliance on data from specific, predefined intervals (e.g., two consecutive discharges, the initial CV phase, or a fixed SOC range) to extract features. This dependency on particular operational segments restricts the applicability of these SOH prediction methods across diverse, real-world usage scenarios where such specific intervals may not be consistently available.

To address these limitations, this study proposes a novel method termed voltage-charge increment curve-based dual-branch bidirectional long short-term memory (VCIC-DB-BiLSTM) for accurate SOH estimation of lithium-ion batteries using only incomplete discharge data. The method begins by preprocessing raw current-voltage measurements to generate the voltage-charge increment curve (VCIC) through ampere-hour integration. Cubic spline interpolation is then applied to convert these non-uniformly sampled curves into standardized data with fixed voltage intervals. Next, fixed-length curve segments extracted from incomplete discharge intervals via a sliding window serve as input features, while complete voltage-interval curves are used as training labels. Finally, a dual-branch bidirectional LSTM (DB-BiLSTM) network is constructed to simultaneously process the segmented local features and the complete raw

data streams, reconstructing full voltage-capacity curves. The SOH is subsequently derived from the capacity value corresponding to the upper cutoff voltage.

To comprehensively evaluate the effectiveness, experiments were conducted on four publicly available battery datasets, the CALCE database from the University of Maryland, the Oxford University battery dataset, the Xi'an Jiaotong University (XJTU) dataset, and the Tongji University (TJU) dataset. These datasets cover diverse battery types, capacities, and cycling conditions, providing sufficient variety for robust validation. Experimental results demonstrate that the proposed method achieves high SOH estimation accuracy using only incomplete discharge data across batteries of varying capacities.

The main contributions of this work are summarized as follows:

- A novel VCIC-DB-BiLSTM framework for SOH estimation of lithium-ion batteries based on incomplete discharge data is proposed, encompassing a complete pipeline from raw data preprocessing and feature construction to model training and testing.
- The VCIC is innovatively adopted as an input feature, which more clearly reflects battery aging characteristics. Furthermore, a DB-BiLSTM network is designed to integrate local feature extraction with global sequence modeling, enabling accurate mapping from incomplete segments to complete SOH estimation.
- The performance of the proposed method is systematically validated on four datasets, CALCE, Oxford, XJTU, and TJU. Experiments show that the method achieves high estimation accuracy across all datasets, with errors below 7 mAh for CALCE, 6 mAh for Oxford, 7 mAh for XJTU, and 17 mAh for TJU, demonstrating its reliability and generalization capability in practical applications.

The remainder of this paper is organized as follows: Section 2 describes the structure of the proposed VCIC-DB-BiLSTM. Section 3 presents the experimental results evaluating the performance of the proposed method. Finally, the conclusion is provided in Section 4.

2. Methodology

This section describes our proposed method for estimating the SOH of lithium-ion batteries. The overall workflow, illustrated in Figure 1, encompasses three key stages: data processing, model training, and performance evaluation.

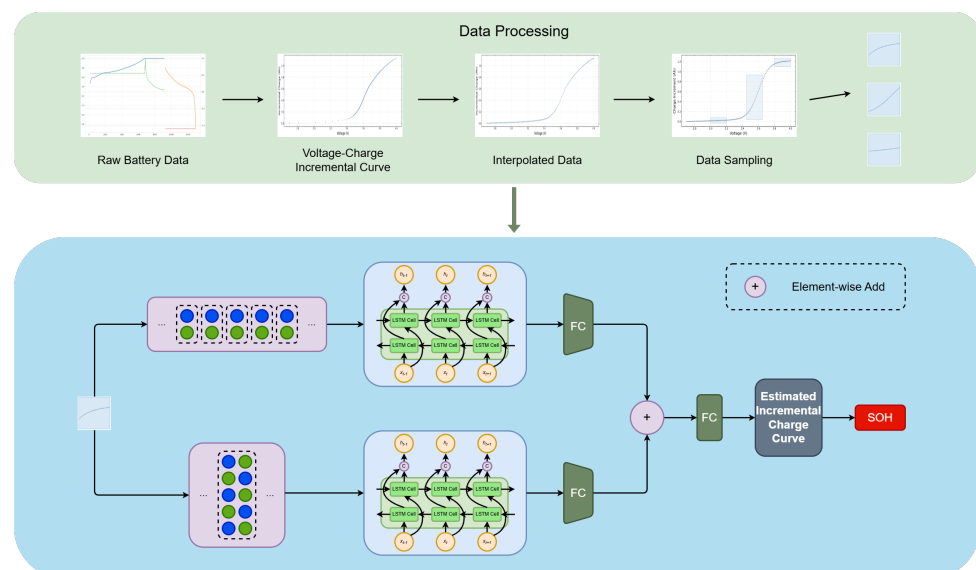


Figure 1. Overall Workflow of the Method.

2.1. Data Processing

A well-designed data processing pipeline is essential for preparing battery performance data before training the DB-BiLSTM model. The processing steps include voltage-capacity extraction, data interpolation, and feature selection. A robust preprocessing method ensures the accuracy and consistency of the input data, which directly influences the performance and reliability of the SOH estimation model.

During the cleaning stage, outliers and noise-corrupted points are removed to improve data reliability. Sensor signals collected during battery operation are often affected by environmental noise and hardware interference, leading to anomalies in the recorded time series [19]. To address this, a local-statistics-based filtering method is applied. The raw data sequence of length N is divided into segments using a fixed window length. For each segment, the mean (μ) and standard deviation (σ) are computed independently. Since an approximately normal distribution is a common characteristic of Additive White Gaussian Noise (AWGN) widely recognized in battery measurement and filtering systems [20,21], the local high-frequency sensor noise is assumed to follow this distribution, thereby establishing a dynamic threshold range $[\mu - 2\sigma, \mu + 2\sigma]$. Data points within each segment are retained only if they fall inside the corresponding local threshold; otherwise, they are discarded. The final valid points form a cleaned index set. When the last segment contains insufficient data, the window length is adjusted adaptively to preserve continuity. As shown in Figure 2a,b, the filtered data exhibit visibly reduced noise compared to the original signals, confirming the effectiveness of the cleaning procedure.

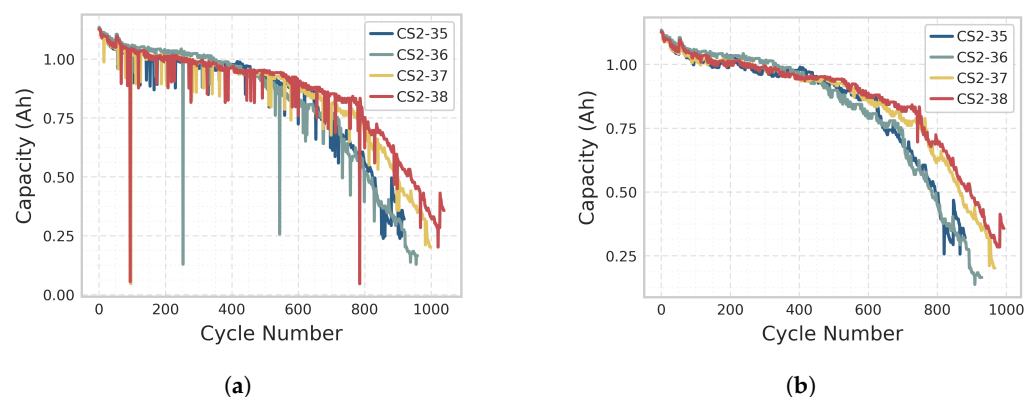


Figure 2. Comparison of Data Filtering Effects: (a) Unfiltered Data; (b) Adaptive Filtered Data.

2.1.1. Voltage and Capacity Extraction

The data collected during each battery cycle include current and voltage measurements. However, directly using these raw signals as model inputs presents several challenges. Raw voltage and current measurements are often noisy, lack a structured representation, and do not explicitly capture the underlying battery degradation patterns. For instance, current readings are susceptible to external interference, while voltage alone does not directly reflect cumulative capacity loss, a key indicator of SOH. These limitations hinder the extraction of discriminative features for accurate SOH estimation. To address these issues, we employ VCIC as an alternative representation. VCIC characterizes the battery's energy storage and release behavior during discharge, providing a more informative signature of battery health than raw time-series data. The incremental charge, computed by integrating current between the lower and upper cutoff voltages within each discharge cycle, encodes subtle electrochemical changes that may not be apparent from voltage or current alone. By integrating the current between the lower and upper cutoff voltages within each discharge cycle, we obtain a robust indicator of battery performance that enhances SOH estimation.

To accurately establish the mapping between voltage and incremental charge, the Ampere-hour (Ah) integration method is first applied [22]. This technique quantifies the capacity change during a discharge cycle by integrating current over time, as expressed in Equation (2)

$$\Delta Q = \int_{t_1}^{t_2} I(t) dt, \quad (2)$$

where t_1 and t_2 denote the start and end times of the integration interval, $I(t)$ represents the time-varying current, and ΔQ is the incremental charge accumulated over this period.

It is worth noting that traditional Ah integration is inherently susceptible to cumulative errors in practical applications due to current sensor noise and long-term measurement drift [22,23]. However, within the proposed framework, this error is mitigated. First, the integration in Equation (2) is strictly bounded within short, localized partial discharge intervals (e.g., 500 mV or 600 mV windows). Because the integration time is shortened, the temporal accumulation of sensor drift is minimized [14,15]. Second, the robust pre-filtering of raw data removes substantial hardware anomalies before integration [20,21]. Finally, the downstream DB-BiLSTM network is resilient to minor residual stochastic noise. It is designed to extract structural sequential degradation patterns (the shape of the VCIC) rather than relying solely on deterministic absolute values [24,25]. Consequently, the integration error is effectively contained and does not significantly propagate to the final SOH prediction.

Figure 3a shows the result after Ah integration. The obtained data points are distributed non-uniformly along the voltage axis, which necessitates further normalization and interpolation to obtain a consistent representation for subsequent analysis and model training.

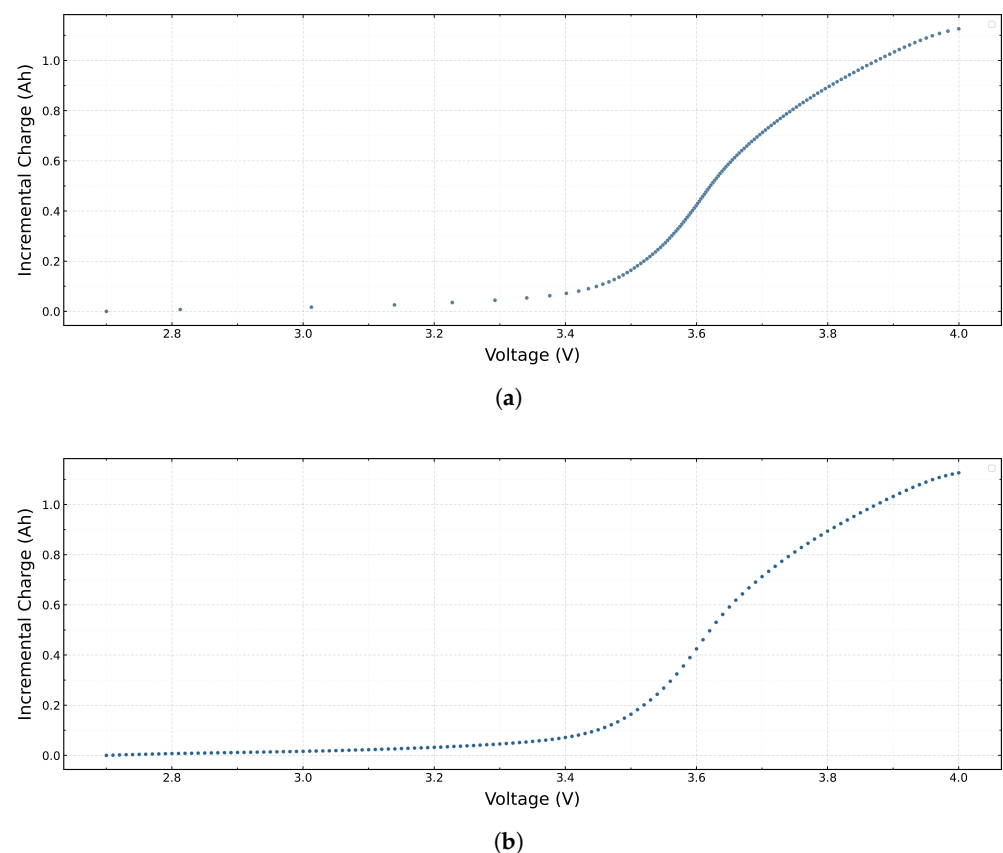


Figure 3. Comparison of Interpolation Effects: (a) Raw Data; (b) Data after Interpolation.

2.1.2. Spline Interpolation

To obtain uniformly distributed voltage data, cubic spline interpolation is applied [26], as illustrated in Figure 3b. This method smooths the raw data and fills in missing values, producing a continuous and uniformly sampled voltage profile. The interpolation works by constructing a piecewise cubic polynomial that passes through each data point, ensuring both continuity and smoothness across adjacent intervals. For a given set of data points $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$, the cubic spline interpolation uses the following polynomial for each interval $[x_i, x_{i+1}]$:

$$S_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3 \quad (3)$$

$$S_i(x_i) = y_i \quad \text{and} \quad S_i(x_{i+1}) = y_{i+1} \quad (4)$$

$$S'_i(x_{i+1}) = S'_{i+1}(x_{i+1}) \quad (5)$$

$$S''_i(x_{i+1}) = S''_{i+1}(x_{i+1}) \quad (6)$$

$$S''_0(x_0) = 0 \quad \text{and} \quad S''_{n-1}(x_n) = 0. \quad (7)$$

Here, $S_i(x)$ represents the cubic polynomial for the i -th interval, and a_i, b_i, c_i , and d_i are the coefficients to be determined. Equation (4) indicates that the spline must pass through all the data points. To guarantee global smoothness, the cubic spline enforces continuity of both the first derivative $S'_i(x)$ and the second derivative $S''_i(x)$ at each internal node, as expressed in Equations (5) and (6). Additionally, natural boundary conditions are applied by setting the second derivative to zero at the endpoints, as given in Equation (7). These constraints ensure that the interpolated curve passes through all data points while maintaining global smoothness and continuity. Figure 3b shows the resulting interpolated data using a uniform voltage interval of 0.01 V, illustrating a continuous and evenly sampled profile.

2.1.3. Feature Selection

For effective SOH estimation, the selection of appropriate features from the processed data is crucial, as it directly influences the model's ability to learn meaningful patterns and generalize well. In this study, sequence segments extracted via a sliding window are adopted as input features. Based on the uniform sampling interval of 0.01 V established during interpolation, we specifically define a voltage window size of 500 mV (0.5 V) for the CALCE, Oxford, and XJTU datasets, and a slightly larger window size of 600 mV (0.6 V) for the TJU dataset. Thus, each input sequence consists of 50 and 60 data points, respectively. The selection of these window sizes represents a critical trade-off: while enlarging the voltage window enables the extraction of richer feature information, it simultaneously causes a significant drop in the available sample size due to strict data boundary handling constraints during sequence extraction. Therefore, optimizing the voltage sampling interval is essential to obtain a representative yet manageable feature set that effectively captures battery performance characteristics.

In traditional machine learning, the SOH is often used directly as the training label. However, employing SOH as a direct target can cause convergence difficulties, especially when only incomplete discharge data is available. This is because accurately estimating a precise SOH value from incomplete or sparse data is inherently challenging, which hinders the model's ability to learn reliable underlying relationships. To overcome this limitation, we redefine the training label as the sequence of incremental charge values corresponding to the voltage range between the lower and upper cutoff voltages in each discharge cycle. Instead of predicting a single SOH value, the model outputs a complete incremental charge sequence. This representation provides richer information and allows the model to learn from a broader range of features, thereby improving stability and convergence.

during training. Although the ultimate objective remains SOH estimation, leveraging the entire incremental charge sequence introduces beneficial redundancy, offering additional contextual information that enhances the model generalization capability. This approach enables effective learning even with limited data, ultimately leading to more accurate and robust SOH predictions.

2.2. Model

In this work, the DB-BiLSTM model is introduced for estimating the SOH of lithium-ion batteries based on incomplete discharge data. Unlike conventional single-path architectures, VCIC-DB-BiLSTM processes the VCIC in two complementary forms, one branch handles the original continuous sequence, while the other operates on patched segments of the curve. After bidirectional feature extraction performed by separate BiLSTM layers, the outputs of both branches are fused and directed to an SOH estimator to generate the final prediction. The following subsections detail each component of the model.

BiLSTM enhances the standard LSTM by allowing information to flow in both forward and backward directions, thereby capturing temporal dependencies more completely [24,27]. This bidirectional design is particularly advantageous for modeling battery behavior over time, especially under the constraint of partial discharge observations. Each LSTM cell alleviates the vanishing-gradient problem common in recurrent neural networks through a memory cell governed by three gating mechanisms: the input gate, forget gate, and output gate. The internal computations of an LSTM unit are given by the following equations:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (8)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (9)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (10)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (11)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (12)$$

$$h_t = o_t \odot \tanh(c_t), \quad (13)$$

where x_t denotes the input at time step t , h_{t-1} is the hidden state from the previous time step, and c_{t-1} and c_t represent the previous and updated cell states, respectively. The intermediate variable $\tilde{c}_t$ represents the candidate cell state. The terms f_t , i_t , and o_t are the activation vectors of the forget, input, and output gates, respectively. The operator $\odot$ denotes the element-wise multiplication (Hadamard product), and the bracket notation $[h_{t-1}, x_t]$ represents the concatenation of the previous hidden state and the current input vector. The weight matrices W_f , W_i , W_o , and W_c , which correspond to the forget gate, input gate, output gate, and candidate cell state respectively, together with their corresponding biases b_f , b_i , b_o , and b_c , are parameters learned during training. The activation functions σ and $\tanh$ refer to the sigmoid and hyperbolic tangent functions, respectively.

Unlike conventional LSTM, which processes sequences only in the forward direction, BiLSTM incorporates both past and future contexts, making it particularly suitable for reconstructing the complete incremental charge curve from incomplete discharge data. This bidirectional capability enables the model to infer missing segments more effectively when working with incomplete cycles. Furthermore, instead of using only the final time step output, the proposed method aggregates outputs across all time steps. This feature representation retains the full temporal dynamics of the input sequence, thereby improving the accuracy of SOH estimation. The architecture of the BiLSTM is illustrated in Figure 4.

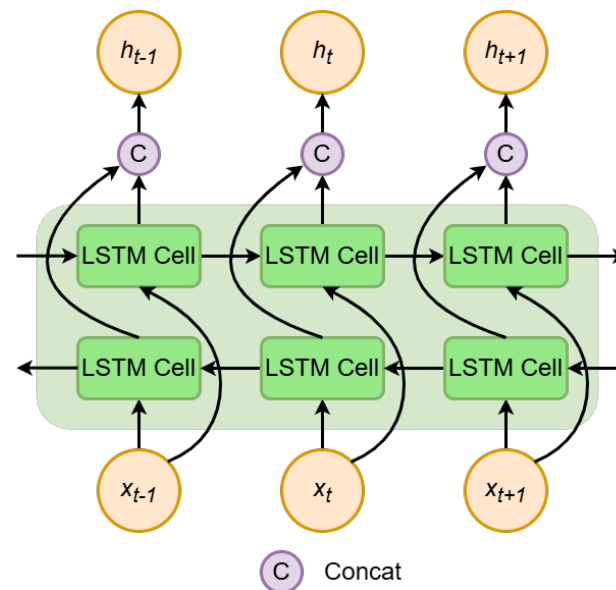


Figure 4. BiLSTM Architecture.

To enhance the model's ability to recognize local patterns, a patch mechanism is incorporated into the VCIC-DB-BiLSTM framework [28]. This mechanism partitions the continuous VCIC into fixed-length segments, which can be either non-overlapping or overlapping. As shown in Figure 5, each patch corresponds to a specific voltage interval and captures short-term behavioral changes within that local window, enabling finer detection of localized degradation features. Additionally, the patch-based representation improves robustness against noise and partial data loss while remaining compatible with interpolation methods, thereby supporting a smoother transition from partial to full-cycle reconstruction.

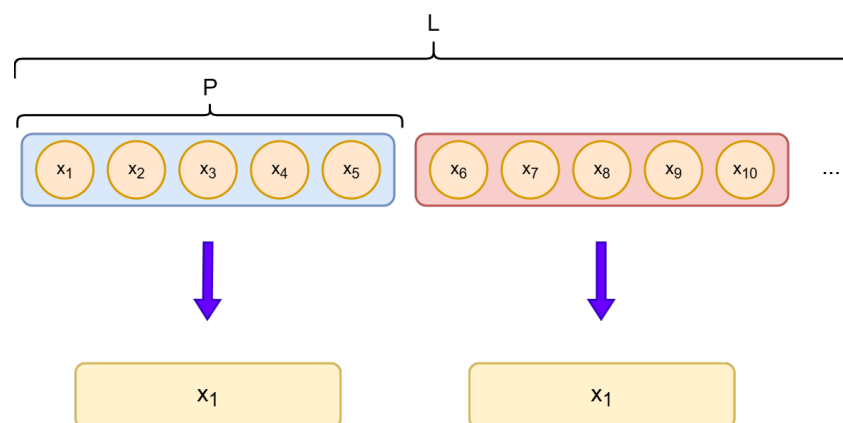


Figure 5. Patch Mechanism.

After extracting multi-scale temporal features, the outputs from the original-sequence branch, denoted as $h_{seq} \in \mathbb{R}^{d_1}$, and the patch-based branch, denoted as $h_{patch} \in \mathbb{R}^{d_2}$, are mapped to identical dimensions through independent fully connected layers. Let $h'_{seq} \in \mathbb{R}^{d_f}$ and $h'_{patch} \in \mathbb{R}^{d_f}$ represent these projected features. To preserve individual feature distributions and avoid excessive dimensionality, the features are fused via an element-wise addition operation. The fused multi-scale feature vector h_{fuse} is formulated as:

$$h_{fuse} = h'_{seq} + h'_{patch} \in \mathbb{R}^{d_f}, \quad (14)$$

where $+$ denotes element-wise addition, and d_f is the target dimension of the mapped outputs.

2.3. Detailed Configuration of the VCIC-DB-BiLSTM

Taking a 500 mV voltage window with a sequence length of $L = 50$ data points as an example, the architectural configuration and parameter settings of the proposed VCIC-DB-BiLSTM model are detailed as follows. The model processes the input sequence where each step initially contains two-dimensional features, specifically the voltage and its corresponding incremental charge (V, Q).

In the original-sequence branch, the raw data matrix is fed into a 2-layer BiLSTM network with a hidden state dimension of 6. Given the bidirectional structure, the output at each time step is concatenated to form a 12-dimensional feature vector. The sequence output is then flattened into a high-dimensional vector $h_{seq} \in \mathbb{R}^{600}$.

Simultaneously, the patch-based branch reorganizes the input sequence by grouping every 5 consecutive time steps into a single localized patch, transforming the sequence length to $L/5$ (10 steps) and expanding the input feature dimension to 10 per step. This patch-level representation is processed by a parallel 4-layer BiLSTM with a larger hidden state dimension of 12. The bidirectional output is subsequently flattened to generate the local feature representation $h_{patch} \in \mathbb{R}^{240}$.

In the feature fusion stage, h_{seq} and h_{patch} are routed through independent pathways. The h_{seq} vector is first passed through a linear layer mapping from 600 to 360, followed by a ReLU activation and another linear layer to reach the target output dimension. The h_{patch} vector is mapped from 240 to 240, passed through a ReLU activation, and then projected to the same target dimension. These aligned vectors are fused via element-wise addition to construct the integrated multi-scale descriptor h_{fuse} . Finally, h_{fuse} is mapped to the final incremental charge sequence through a fully connected output layer.

The specific execution process of the proposed VCIC-DB-BiLSTM model is summarized in Algorithm 1. It details the forward propagation pipeline, including the independent feature extraction across the two branches and the element-wise addition fusion strategy.

Algorithm 1 Forward Propagation of VCIC-DB-BiLSTM

Require: Processed input sequence $X \in \mathbb{R}^{L \times 2}$, where L is the sequence length (e.g., $L = 50$). d_f is the target output dimension.

Ensure: Estimated incremental charge sequence $\hat{Y} \in \mathbb{R}^{d_f}$.

```

1: // Original-Sequence Branch
2:  $H_{seq} \leftarrow \text{BiLSTM}_{seq}(X)$  % 2 layers, hidden size  $d_1 = 6$ 
3:  $h_{seq} \leftarrow \text{Flatten}(H_{seq})$  %  $h_{seq} \in \mathbb{R}^{600}$ 
4:  $h'_{seq} \leftarrow \text{Linear}_{12}(\text{ReLU}(\text{Linear}_{11}(h_{seq})))$  % Feature mapping to  $d_f$ 
5: // Patch-Based Branch
6:  $X_{patch} \leftarrow \text{Reshape}(X, L/5, 10)$  % Group every 5 time steps
7:  $H_{patch} \leftarrow \text{BiLSTM}_{patch}(X_{patch})$  % 4 layers, hidden size  $d_2 = 12$ 
8:  $h_{patch} \leftarrow \text{Flatten}(H_{patch})$  %  $h_{patch} \in \mathbb{R}^{240}$ 
9:  $h'_{patch} \leftarrow \text{Linear}_{22}(\text{ReLU}(\text{Linear}_{21}(h_{patch})))$  % Feature mapping to  $d_f$ 
10: // Feature Fusion and Output
11:  $h_{fuse} \leftarrow h'_{seq} + h'_{patch}$  % Element-wise addition
12:  $\hat{Y} \leftarrow \text{Linear}_{out}(h_{fuse})$  % Final prediction
13: return  $\hat{Y}$ 

```

As outlined in Algorithm 1, the input to the network is the processed sequence X . In the original-sequence branch, X is directly fed into BiLSTM_{seq} to extract global temporal dependencies, producing the sequence output H_{seq} . This output is then flattened into a high-dimensional vector h_{seq} . To project this feature into the target latent space, two fully connected layers (Linear_{11} and Linear_{12}) with an intermediate ReLU activation are applied, yielding the mapped feature $h'_{seq} \in \mathbb{R}^{d_f}$.

Concurrently, in the patch-based branch, X is restructured into X_{patch} by concatenating the features of every 5 consecutive time steps. This localized representation is processed by BiLSTM_{patch} to generate H_{patch} , which is subsequently flattened into h_{patch} . Similar to the first branch, h_{patch} undergoes independent fully connected transformations (Linear_{21} and Linear_{22}) to align its dimension, resulting in $h'_{patch} \in \mathbb{R}^{d_f}$.

Crucially, rather than concatenating the features from the two branches, h'_{seq} and h'_{patch} are integrated via element-wise addition to form the fused representation h_{fuse} . This fusion strategy effectively combines the macro-level temporal dynamics with the micro-level localized patch variations without unnecessarily expanding the feature dimensionality. Finally, h_{fuse} is passed through the output linear layer Linear_{out} to generate the complete estimated incremental charge sequence $\hat{Y}$.

To provide a clear and comprehensive overview of the experimental setup, all key inputs, outputs, and network hyperparameters of the VCIC-DB-BiLSTM architecture are summarized in Table 1.

Table 1. Inputs, Outputs, and Key Hyperparameters of the VCIC-DB-BiLSTM Model (500 mV Window Example).

Category	Configuration/Value
Inputs	Processed sequence $X \in \mathbb{R}^{50 \times 2}$
Outputs	Estimated incremental charge sequence $\hat{Y} \in \mathbb{R}^{d_f}$
Voltage Window Size	500 mV
Interpolation Interval	0.01 V
Sequence Length (L)	50
Original-Sequence Branch	2-layer BiLSTM, hidden size = 6
Patch-Based Branch	4-layer BiLSTM, hidden size = 12, patch size = 5
Learning Rate	0.001
Max Epochs	200
Batch Size	64

2.4. Training and Testing

The VCIC-DB-BiLSTM model was implemented on a workstation running Ubuntu 22.04 LTS, equipped with an NVIDIA GeForce RTX 3090 Ti GPU (24 GB VRAM). The software environment was configured with Python 3.10, PyTorch 2.1.0, and CUDA 11.8. During training, we employed a 4-fold cross-validation strategy to ensure robust model generalization and mitigate overfitting. For each fold, we randomly sampled voltage segments spanning 0.5 V or 0.6 V from the discharge curves, with their corresponding incremental charge values as input features. This voltage range selection was strategically designed to accommodate practical scenarios where full discharge cycles are unavailable. The model architecture combines a BiLSTM layer with a novel patching mechanism. Each input sequence was normalized to [0, 1] range using dataset-specific min-max scaling to ensure numerical stability. The training objective focuses on reconstructing complete voltage-incremental charge curves rather than direct SOH regression, forcing the model to learn fundamental battery degradation patterns. We employed the Adam optimizer with an initial learning rate of 0.001, combined with early stopping based on validation loss plateau detection [29]. The mean squared error (MSE) loss function is employed to measure the estimation accuracy, guiding the model towards minimizing the discrepancy between the estimated and actual voltage-incremental charge curves. Mini-batch training with size 64 was conducted over 200 epochs, showing stable convergence across all datasets.

During the testing phase, the model is presented with any voltage segments of voltage and incremental charge data. It then estimates the entire voltage-incremental charge curve for the cycle. From this estimated curve, the SOH for the cycle is determined by identifying

the incremental charge corresponding to the upper cut-off voltage. This method allows for accurate SOH estimation even with incomplete discharge data, demonstrating the model's practical applicability in real-world scenarios.

3. Results and Discussion

The experimental data used in this study are drawn from publicly available battery aging datasets, including the CALCE database [30], the Oxford University dataset [31], the Xi'an Jiaotong University (XJTU) dataset [25,32,33], and the Tongji University (TJU) dataset [34]. These datasets encompass batteries with diverse material systems and were collected under varied testing conditions.

CALCE Dataset: The CALCE dataset [30] comprises cycle data from cells CS35, CS36, CS37, and CS38. Each cell underwent a standard CC-CV charging protocol: constant-current (CC) charging at 0.5C until reaching 4.2 V, followed by CV charging until the current dropped below 0.05 A. The discharge cutoff voltage was set to 2.7 V. The cells are rated at 1100 mAh, employ a LiCoO₂ cathode with trace manganese content, and weigh approximately 21.1 g. Capacity degradation curves for the CALCE cells are shown in Figure 6a.

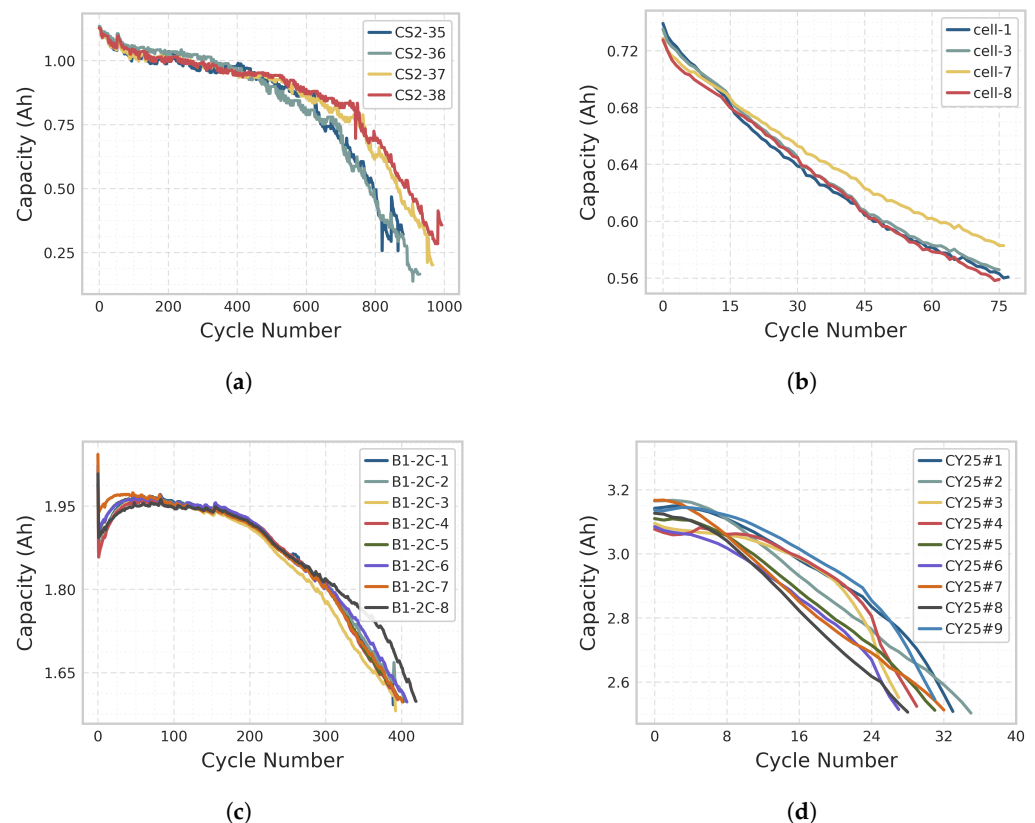


Figure 6. Battery Degradation Comparison Across Four Datasets: (a) CALCE Degradation; (b) Oxford Degradation; (c) XJTU Degradation; (d) TJU Degradation.

Oxford Dataset: The Oxford Battery [31] Degradation Dataset contains aging data from eight small lithium-ion pouch cells (Kokam CO LTD, Suwon-Si, Republic of Korea, model SLPB533459H4) employing a Nickel Cobalt Manganese (NCM) chemistry with a nominal capacity of 740 mAh. Tests were conducted in a thermal chamber at 40 °C. Cells were charged using a CC-CV profile and discharged according to a drive-cycle profile derived from the urban Artemis driving schedule. Characterization measurements were recorded every 100 cycles until end-of-life. For this study, cells numbered 1, 3, 7, and 8 were selected. The corresponding capacity degradation curves are presented in Figure 6b.

XJTU Dataset: The XJTU dataset [25,32,33] includes 55 LISHEN lithium-ion batteries based on the NCM chemistry with a nominal capacity of 2000 mAh and a nominal voltage of 3.6 V. Charge and discharge cutoff voltages are 4.2 V and 2.5 V, respectively. Batteries were cycled under six different charge/discharge protocols at room temperature until failure. The protocols include fixed-current charging/discharging, random discharging with fixed charging current across cycles, random-walk current profiles, and a geosynchronous earth orbit (GEO) satellite operational profile. For this study, a subset of 8 batteries from Batch 1, which underwent a fixed charging and discharging strategy at a 2C rate, was selected. In subsequent experiments, these are denoted as Batch-1-2C(1) through Batch-1-2C(8) to evaluate the model's performance under consistent high-rate degradation conditions. Capacity degradation curves are illustrated in Figure 6c.

TJU Dataset: The TJU dataset [34] covers three battery chemistries: Nickel Cobalt Aluminum (NCA) (3500 mAh, 2.65–4.2 V), NCM (3500 mAh, 2.5–4.2 V), and a blended NCM + NCA system (2500 mAh, 2.5–4.2 V). Cycling tests were performed in a temperature-controlled chamber with temperatures of 25 °C, 35 °C, and 45 °C, and charging current rates ranging from 0.25C to 4C. To ensure a rigorous validation, we specifically utilized 8 NCA batteries cycled under the baseline condition, which employs a constant charge/discharge rate of 1C at 25 °C (denoted as the CY-1 cycling condition protocol in the dataset). These batteries, referred to as CY-1-1(1) through CY-1-1(8), were selected to systematically benchmark the estimation accuracy across identically cycled units. The capacity degradation trends are shown in Figure 6d.

To provide a clear comparative overview of the experimental conditions, the core specifications of the selected batteries, including cathode chemistry, nominal capacity, and discharge profiles, are summarized in Table 2. This diversity in material systems (LiCoO₂, NCM, NCA) and operational stresses (from 1C/2C constant current to dynamic driving schedules) robustly validates the comprehensive engineering applicability of the proposed VCIC-DB-BiLSTM framework.

Table 2. Summary of Battery Specifications and Cycling Conditions across Four Datasets.

Dataset	Cathode Chemistry	Nominal Capacity	Discharge Profile (C-Rate)
CALCE	LiCoO ₂	1100 mAh	Constant Current (0.5C)
Oxford	NCM	740 mAh	Dynamic (Artemis Urban)
XJTU (Batch 1)	NCM	2000 mAh	Constant Current (2C)
TJU (CY-1 Baseline)	NCA	3500 mAh	Constant Current (1C)

To quantitatively assess the performance of the proposed method, the mean absolute error (MAE) and root mean square error (RMSE) between the estimated SOH values and the reference SOH values are calculated. These metrics are defined as follows:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (15)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (16)$$

where $\hat{y}_i$ denotes the estimated SOH value, y_i represents the corresponding actual reference SOH value, and N is the number of samples. Together, these metrics offer a comprehensive evaluation of the accuracy and reliability of the SOH estimation method.

While MAE and RMSE provide rigorous quantitative measurements of the absolute estimation error, evaluating the practical viability of these models requires contextualizing

the absolute errors against the battery's nominal capacity. In practical BMS for electric vehicles, an absolute SOH estimation error within 5% is generally considered sufficient for reliable implementation [35,36]. Consequently, the effectiveness of the proposed method is evaluated based on two objective criteria: first, its capability to maintain an absolute estimation error well below the rigorous 5% threshold across diverse battery chemistries; and second, its ability to consistently achieve lower MAE and RMSE compared to benchmark data-driven models under identical partial-discharge conditions.

3.1. Experiments with Different Datasets

To evaluate the effectiveness of the proposed DB-BiLSTM framework for SOH estimation using partial discharge data, comprehensive experiments were conducted on four benchmark datasets (CALCE, Oxford, XJTU, and TJU) covering a wide range of nominal capacities (from 1100 mAh to 3500 mAh). In the experiments, the Adam optimizer and MSE loss function were used, and a 4-fold cross-validation strategy was applied to ensure robust performance assessment. Each dataset was divided into four subsets; during each fold, the model was trained on three subsets and validated on the remaining one. The specific dataset partitions are detailed in Table 3.

Table 3. The Division of the Four Benchmark Datasets.

	Fold 1	Fold 2	Fold 3	Fold 4
CALCE	CS2-35	CS2-36	CS2-37	CS2-38
Oxford	Cell 1	Cell 3	Cell 7	Cell 8
XJTU	Batch-1-2C(1,2)	Batch-1-2C(3,4)	Batch-1-2C(5,6)	Batch-1-2C(7,8)
TJU	CY-1-1(1,2)	CY-1-1(3,4)	CY-1-1(5,6)	CY-1-1(7,8)

Figures 7a–d and 8a–d display comparisons between the actual and estimated SOH curves for selected batteries. The close alignment between the curves visually demonstrates the method's high accuracy and reliability, confirming its capability to effectively capture battery performance trends even when only partial discharge data are available.

To provide a quantitative assessment, Table 4 summarizes the performance metrics, including MAE, across the four datasets. The consistently low MAE values indicate high accuracy and stable SOH estimation. These results verify that the proposed DB-BiLSTM method is both effective and robust when applied to diverse battery types with varying characteristics.

Table 4. DB-BiLSTM MAE on Different Datasets.

	Fold 1	Fold 2	Fold 3	Fold 4
CALCE	0.0048	0.0056	0.0039	0.0067
Oxford	0.0054	0.0040	0.0041	0.0041
XJTU	0.0051	0.0069	0.0050	0.0070
TJU	0.0162	0.0099	0.0129	0.0141

To explicitly evaluate the model's effectiveness against the rigorous industrial standards established earlier, it is essential to contextualize the absolute MAE values presented in Table 4 as absolute percentage errors based on each dataset's nominal capacity. Considering the maximum (worst-case) MAE across all four folds for each dataset, the CALCE dataset (nominal capacity 1.1 Ah) exhibits a maximum MAE of 0.0067 Ah, which corresponds to an absolute error of 0.61%. Similarly, the maximum absolute errors for the Oxford (0.74 Ah nominal), XJTU (2.0 Ah nominal), and TJU (3.5 Ah nominal) datasets are 0.73% (0.0054 Ah), 0.35% (0.0070 Ah), and 0.46% (0.0162 Ah), respectively. Across all tested

scenarios and battery chemistries, the maximum estimation error remains strictly below 1%. This performance falls well below the 5% error threshold, firmly validating the reliability and industrial applicability of the proposed VCIC-DB-BiLSTM method.

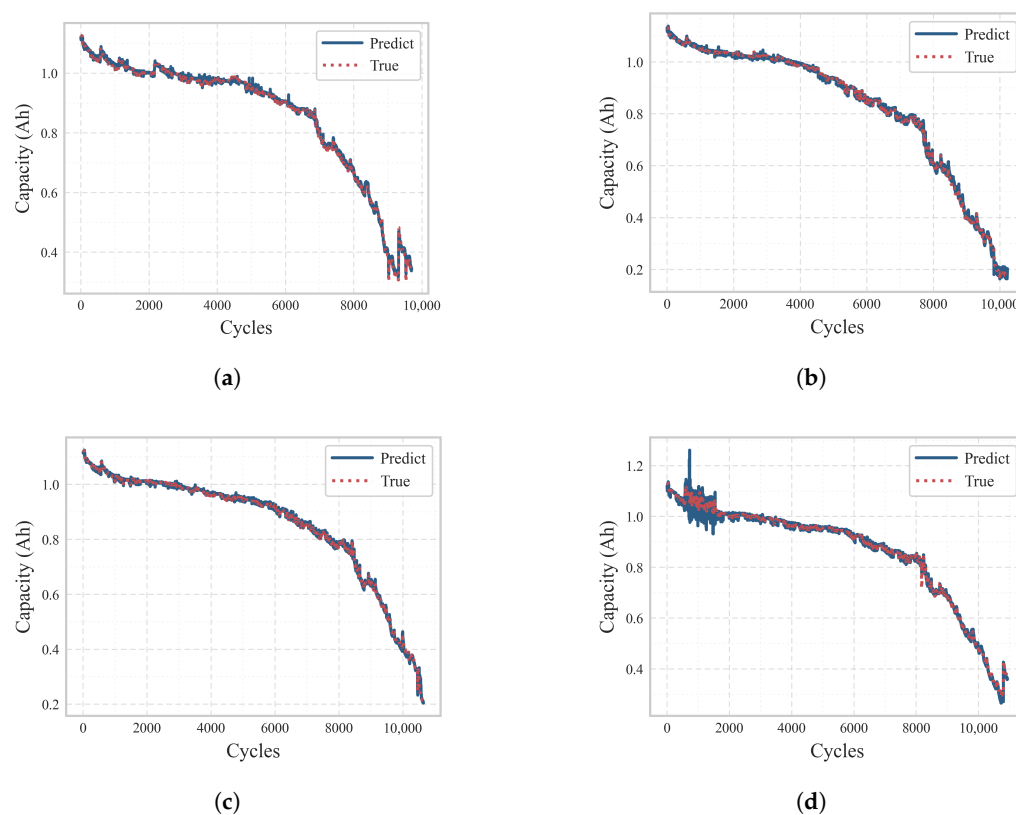


Figure 7. Comparison of CALCE Fold Test Results: (a) Fold 1; (b) Fold 2; (c) Fold 3; (d) Fold 4.

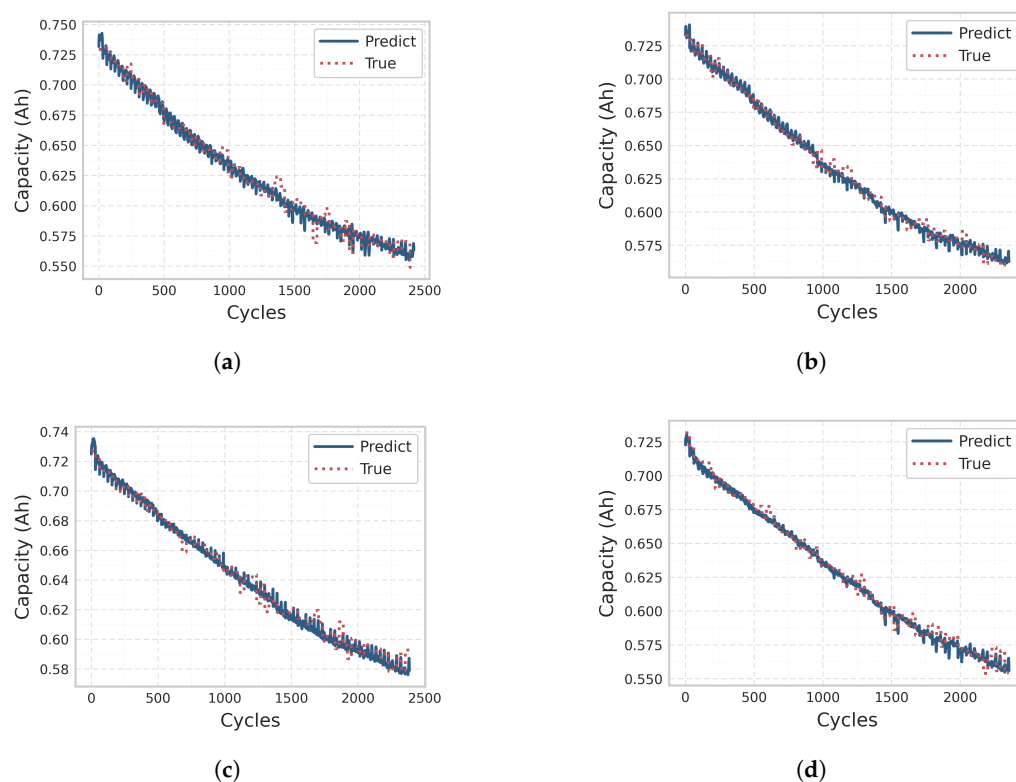


Figure 8. Comparison of Oxford Fold Test Results: (a) Fold 1; (b) Fold 2; (c) Fold 3; (d) Fold 4.

To validate the effectiveness of the proposed approach, a comparative analysis was conducted against several established models, including Attention LSTM [37], CNN [38], GRU [39], and RNN [40]. The cross-validation results, presented in Tables 5–8, show that the DB-BiLSTM method consistently outperforms these benchmark models across all four datasets (CALCE, Oxford, XJTU, and TJU), confirming its robustness and generalization capability.

Table 5. MAE of Different Methods on the CALCE Dataset.

Model	Fold 1	Fold 2	Fold 3	Fold 4
Attention	0.0128	0.0192	0.0121	0.0196
CNN	0.0117	0.0138	0.0101	0.0173
GRU	0.0080	0.0105	0.0070	0.0125
RNN	0.0090	0.0122	0.0077	0.0126
BiLSTM	0.0084	0.0101	0.0067	0.0106
DB-BiLSTM	0.0048	0.0056	0.0039	0.0067

Note: Bold values indicate the best performance.

Table 6. MAE of Different Methods on the Oxford Dataset.

Model	Fold 1	Fold 2	Fold 3	Fold 4
Attention	0.0084	0.0071	0.0109	0.0161
CNN	0.0078	0.0060	0.0053	0.0055
GRU	0.0063	0.0057	0.0052	0.0057
RNN	0.0104	0.0085	0.0075	0.0080
BiLSTM	0.0068	0.0057	0.0059	0.0056
DB-BiLSTM	0.0054	0.0040	0.0041	0.0041

Note: Bold values indicate the best performance.

Table 7. MAE of Different Methods on the XJTU Dataset.

Model	Fold 1	Fold 2	Fold 3	Fold 4
Attention	0.0081	0.0105	0.0060	0.0079
CNN	0.0079	0.0102	0.0071	0.0085
GRU	0.0073	0.0088	0.0065	0.0076
RNN	0.0093	0.0105	0.0085	0.0078
BiLSTM	0.0085	0.0100	0.0076	0.0068
DB-BiLSTM	0.0051	0.0069	0.0050	0.0070

Note: Bold values indicate the best performance.

Table 8. MAE of Different Methods on the TJU Dataset.

Model	Fold 1	Fold 2	Fold 3	Fold 4
Attention	0.0359	0.0245	0.0406	0.0369
CNN	0.0403	0.0314	0.0474	0.0509
GRU	0.0321	0.0185	0.0330	0.0382
RNN	0.0342	0.0305	0.0486	0.0340
BiLSTM	0.0267	0.0201	0.0331	0.0366
DB-BiLSTM	0.0162	0.0099	0.0129	0.0141

Note: Bold values indicate the best performance.

Moreover, the results highlight the advantage of using BiLSTM as the core architecture. Importantly, the proposed DB-BiLSTM model, which integrates a patching mechanism, achieves superior performance compared to the standard BiLSTM on every dataset. This demonstrates the added benefit of the patch-based design in enhancing local feature extraction and overall estimation accuracy.

Figure 9 compares the MAE and RMSE distributions of different models across the four datasets, visually illustrating the performance differences. The results consistently show that the DB-BiLSTM model achieves the best performance on both evaluation metrics. In particular, the MAE remains the lowest for DB-BiLSTM, reflecting its accurate alignment with the actual SOH values across all datasets. This indicates the model's high overall accuracy and strong generalization ability under varying battery conditions.

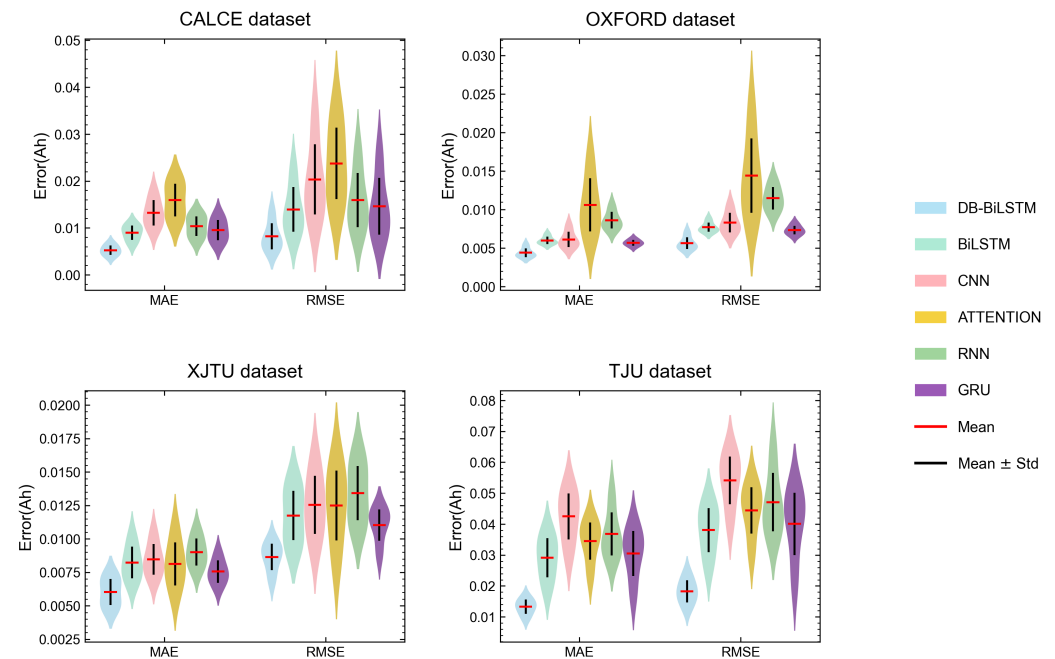


Figure 9. Comparison of MAE and RMSE for Different Models on the Oxford Dataset.

Moreover, the RMSE results demonstrate that DB-BiLSTM is especially effective in reducing large estimation errors, as evidenced by its consistently lower RMSE values compared to other models. This suggests that the proposed method not only captures the general degradation trend but also maintains stability and avoids significant deviations, even when handling diverse and complex battery data.

3.2. Experiments with Small Samples

To further examine the robustness of the proposed method under data-limited conditions, we conducted additional experiments to evaluate its performance in small-sample settings. This set of tests aims to verify the method's practicality and reliability in real-world scenarios where extensive data collection is often infeasible. Specifically, we selected two datasets, XJTU and TJU, both of which contain rich battery cycling records, for training and validation. The configuration of training and test subsets is detailed in Table 9.

Table 9. Training Set and Test Set Settings.

Sample Ratio	1	2	3	4	Test
XJTU(Batch-1-2C)	1	1,2	1,2,3	1,2,3,4	8
TJU(CY-1-1)	1	1,2	1,2,3	1,2,3,4	8

In these experiments, the training set size was intentionally restricted by gradually reducing the number of batteries used for training. Specifically, the model was trained successively using data from one, two, three, and finally four batteries, with each configuration evaluated on a separate test set. This stepwise reduction allowed us to systematically exam-

ine the model's performance under varying degrees of data availability, thereby assessing its generalizability and adaptability in data-constrained scenarios.

Figures 10 and 11 show the MAE of the model on the XJTU and TJU datasets as the training set size varies. The plots visually illustrate how model performance evolves with less training data. As expected, performance generally improves with more training samples, which is a typical trend in machine learning. Notably, however, even when trained on a very small dataset (e.g., data from only one battery), the proposed method consistently outperforms other benchmark methods commonly used for SOH estimation.

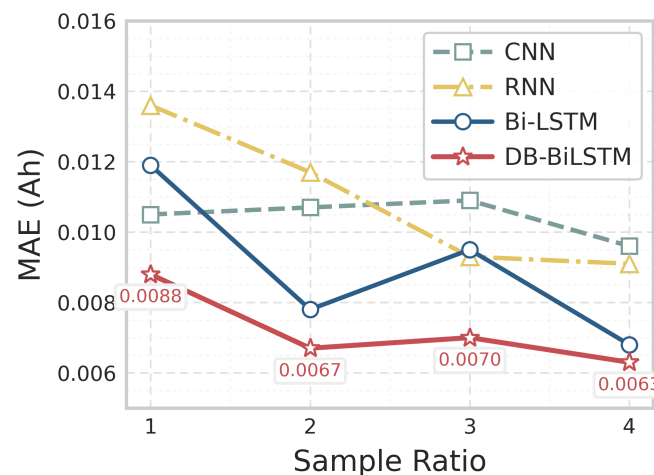


Figure 10. Experiments with Small Samples on the XJTU Dataset.

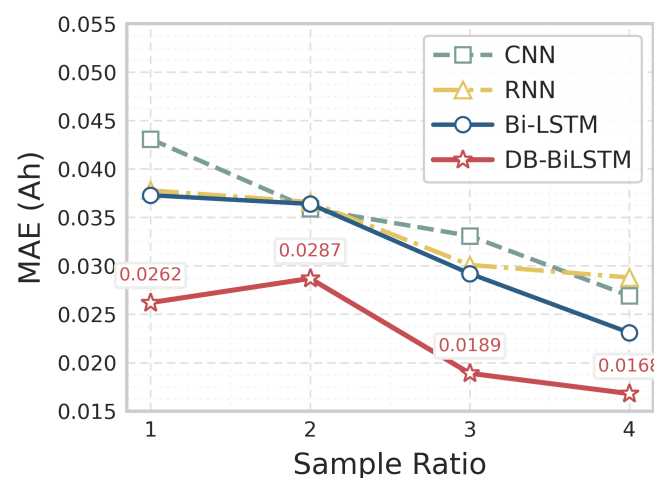


Figure 11. Experiments with Small Samples on the TJU Dataset.

Furthermore, the method's strong performance under small-sample conditions highlights its potential for practical deployment in settings where collecting large datasets is difficult or impractical. This capability is particularly valuable in industrial or field applications, where rapid and reliable SOH estimation is required despite limited available data.

3.3. Impact of Voltage Window Size on SOH Estimation Accuracy

The appropriate selection of the voltage sampling window is a key technical consideration in this study. An excessively large voltage window leads to a substantial increase in the amount of data per cycle, but critically, it causes a significant drop in the overall sample size due to strict data boundary handling requirements. This reduction in sample size can degrade model generalization and conflict with the goal of building a widely applicable framework.

To determine a suitable voltage interval, experiments were conducted with window sizes ranging from 200 mV to 600 mV for the CALCE, Oxford, and XJTU datasets, and from 200 mV to 700 mV for the TJU dataset. To ensure reliable evaluation, the data partitioning for these tests is strictly consistent with the experiments in Section 3.1, utilizing the same 4-fold cross-validation strategy. The estimation results are visualized using violin plots in Figure 12. In these plots, the individual data points represent the evaluation metric (e.g., MAE) obtained from each of the four distinct folds, while the central marker indicates the average performance across all folds. This visualization effectively demonstrates both the general trend and the variance in model robustness under different window sizes.

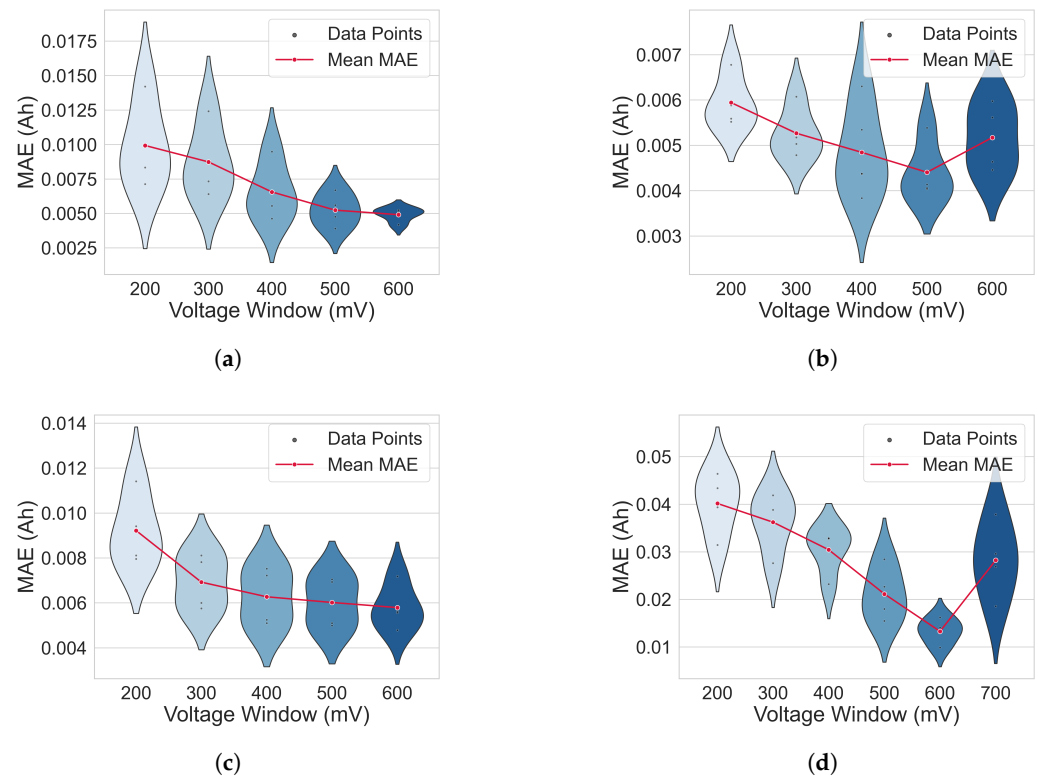


Figure 12. Impact of Voltage Window Size on SOH Estimation Accuracy of Different Datasets: (a) CALCE; (b) Oxford; (c) XJTU; (d) TJU.

As observed, as the window size continues to increase, the performance gains diminish and eventually drop (e.g., the Oxford dataset at 600 mV and the TJU dataset at 700 mV). This inflection point occurs because the loss of training samples heavily restricts the model's capability to learn robust representations. Based on this trade-off analysis, the optimal voltage intervals selected for input features are 500 mV for the CALCE, Oxford, and XJTU datasets, and 600 mV for the TJU dataset.

3.4. Impact of Voltage Window Starting Point on SOH Estimation Accuracy

In the feature selection process of this study, the voltage window length is fixed to maintain consistent input feature dimensions. This approach mitigates the influence of uneven data spacing while retaining the window's starting position as a tunable variable. As different voltage intervals convey varying amounts of information, the choice of the voltage window's starting point can significantly affect SOH estimation accuracy. Investigating how different starting positions impact performance provides practical guidance for selecting input features in online SOH applications.

Figure 13 illustrates the model's estimation accuracy across the four datasets at different voltage window starting points. Consistent with previous evaluations, these results are

derived using the same 4-fold cross-validation setup in Section 3.1, with the plotted values representing the average performance across all four folds to ensure statistical reliability. The horizontal axis represents the starting voltage value of the window. For example, for the CALCE dataset, a value of 3.48 V indicates that the input window spans from 3.48 V to 3.98 V. The plotted trends clearly show that the optimal starting point for achieving the highest estimation accuracy varies between datasets.

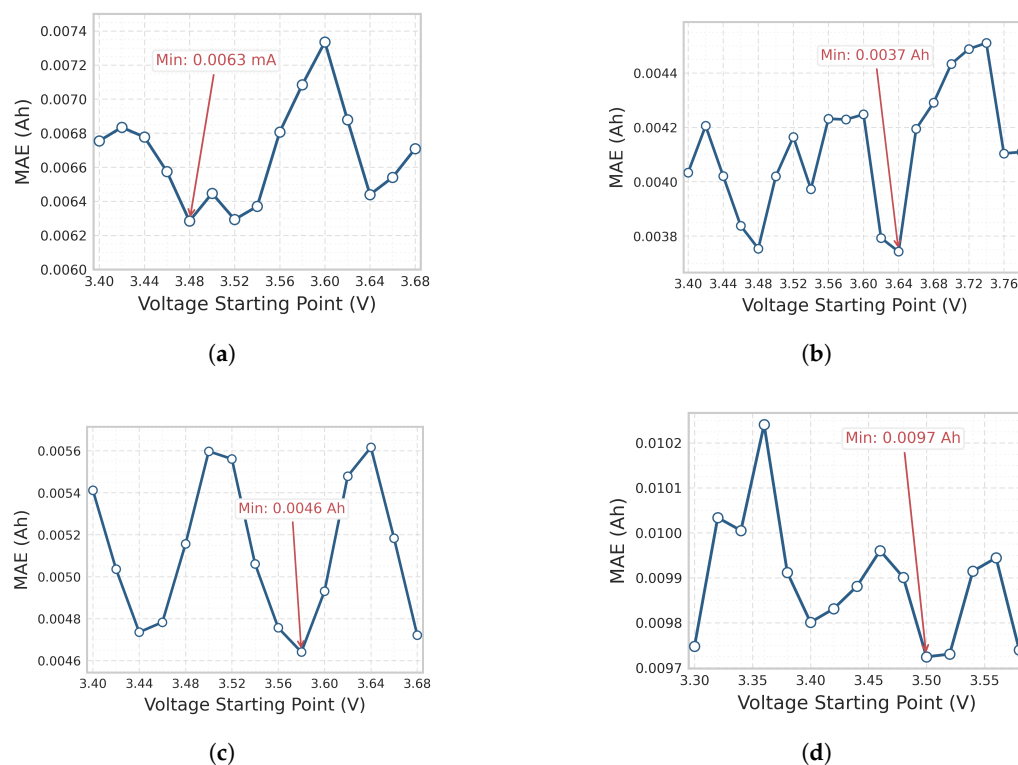


Figure 13. Impact of Voltage Window Starting Point on SOH Estimation Accuracy of Different Datasets: (a) CALCE; (b) Oxford; (c) XJTU; (d) TJU.

For instance, Figure 14 compares the estimation results on the CALCE dataset when the voltage window starts at 3.48 V versus 3.60 V. The estimates with a starting point of 3.48 V align closely with the ground truth, whereas those beginning at 3.60 V exhibit noticeable deviations in several cycles. This observation confirms the importance of the window's starting position. Therefore, in online SOH estimation, if multiple candidate voltage windows are available, selecting the one associated with the best estimation performance can effectively enhance model accuracy.

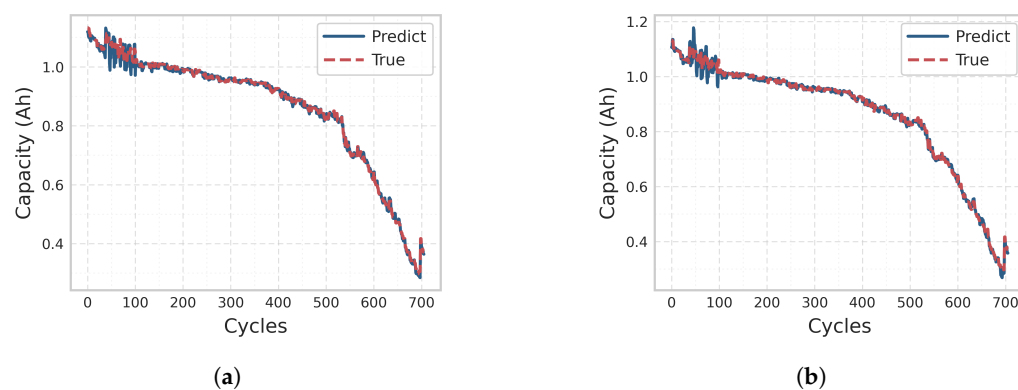


Figure 14. Cont.

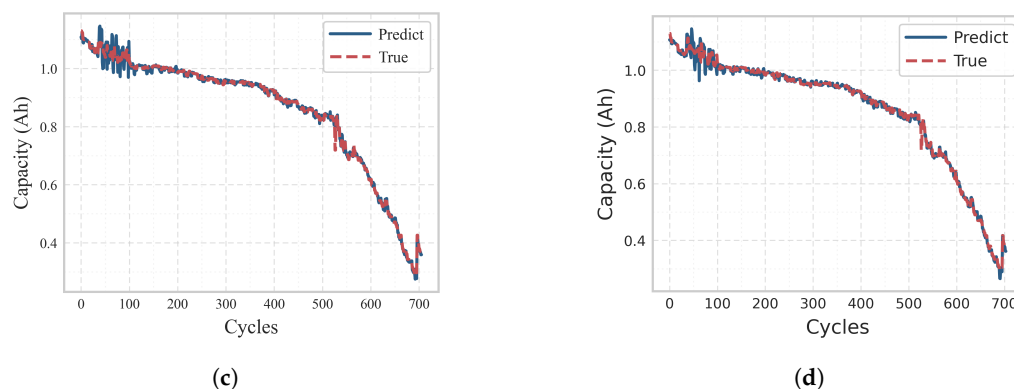


Figure 14. Model Estimation Performance at Different Voltage Starting Points: (a) 3.40 V; (b) 3.48 V; (c) 3.60 V; (d) 3.68 V.

3.5. Comprehensive Evaluation and Operational Boundaries

To clarify the practical applicability and operational boundaries of the proposed VCIC-DB-BiLSTM framework, its performance is evaluated across diverse battery chemistries, operational conditions, and lifecycle stages.

The framework demonstrates robustness across varying lithium-ion cathode materials. As detailed in Tables 5–8, the method maintains low estimation errors on datasets comprising LiCoO₂ (CALCE), NCM (Oxford and XJTU), and NCA (TJU) chemistries. This indicates that the VCIC feature can capture generalizable electrochemical degradation trends across the most prevalent commercial material systems.

Beyond material compatibility, the method accommodates a broad spectrum of discharge conditions. The evaluation encompasses low-to-moderate constant discharge rates (0.5C in CALCE, 1C in TJU), higher continuous stress (2C in XJTU), and a dynamic urban Artemis driving schedule (Oxford). The sustained accuracy under both stable and dynamic profiles suggests that the framework is not strictly constrained to constant-current operations.

Moreover, the estimation trajectories in Figures 7 and 8 illustrate that the VCIC-DB-BiLSTM framework effectively tracks capacity fade throughout the full battery lifecycle. The estimated SOH follows the degradation trend from the initial fresh state through the accelerated capacity drop at the end-of-life phase, capturing continuous variations without requiring stage-specific recalibration.

When assessed against practical accuracy requirements, the proposed method shows significant potential. Industrial BMS typically consider an absolute SOH estimation error within 5% as acceptable. The framework consistently maintains a maximum absolute error of less than 1% across the four datasets. Furthermore, under strictly identical partial-discharge constraints, the dual-branch architecture consistently achieves lower estimation errors compared to established data-driven baselines (CNN, GRU, RNN, and Attention LSTM).

Despite these promising results, the operational boundaries of this study must be acknowledged. The current evaluation is restricted to cell-level data with nominal capacities up to 3.5 Ah. Application to large-format cells or aggregated battery packs involving cell-to-cell variability has not been verified in this work. These aspects, along with performance under extreme fast-discharging and diverse thermal conditions, remain critical areas for future investigation.

4. Conclusions

This paper presents a novel DB-BiLSTM model for estimating the SOH of lithium-ion batteries using partial discharge data. The approach overcomes the common limitation of conventional methods that depend on complete charge–discharge cycles. By applying the

Ampere-hour integral to derive the voltage-incremental charge relationship and employing cubic spline interpolation to standardize voltage spacing, a structured input representation is obtained. Voltage data sampled at intervals of 0.5 V or 0.6 V, together with the full voltage-incremental charge curve as the training target, enable the model to learn detailed discharge patterns. Extensive experiments on multiple datasets, encompassing LiCoO₂, NCM, and NCA chemistries, demonstrate that the DB-BiLSTM model achieves high SOH estimation accuracy, with performance quantitatively confirmed using metrics such as MAE and RMSE. Additional experiments under limited-data conditions, conducted on the XJTU and TJU datasets with progressively reduced training samples, show that the proposed DB-BiLSTM maintains superior performance compared to other benchmark methods, even when very little data are available. This underscores the model's adaptability and effectiveness in data-constrained scenarios. The results indicate that the DB-BiLSTM model can enhance BMS by providing reliable SOH estimates from partial discharge data across the full range of the battery lifecycle, offering a practical solution for real-world applications in energy storage and electric vehicles. Future work will focus on investigating the impact of adverse environmental factors, particularly varying and extreme temperatures, on the model's performance. Because temperature significantly affects battery electrochemistry and accelerates degradation, incorporating temperature-compensated features into the VCIC-DB-BiLSTM framework will be a key step for further model refinement. Additionally, to address the operational boundaries identified previously, subsequent research will evaluate and optimize the framework under diverse power profiles, particularly in extreme high-power and fast-discharging scenarios, thereby ensuring its robust deployment in complex, real-world applications.

Author Contributions: Conceptualization, L.D.; methodology, L.K. and X.Z.; investigation, L.K. and X.Z.; data curation, L.K. and X.Z.; formal analysis, L.K. and X.Z.; visualization, L.K. and X.Z.; writing—original draft preparation, L.K.; writing—review and editing, L.K., X.Z. and L.D.; supervision, L.D.; project administration, L.D. 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 (No. 62401453), Fundamental Research Funds for Xi'an Jiaotong University (No. xzy012023134), and China Postdoctoral Science Foundation (No. 2023M732792).

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to their large size and complex structure.

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

Abbreviations

The following abbreviations are used in this manuscript:

Ah	Ampere-hour
AWGN	Additive White Gaussian Noise
BiLSTM	Bidirectional Long Short-Term Memory
BMS	Battery Management Systems
CALCE	Center for Advanced Life Cycle Engineering
CC	Constant Current
CNN	Convolutional Neural Network
CV	Constant Voltage
DB-BiLSTM	Dual-Branch Bidirectional Long Short-Term Memory
EIS	Electrochemical Impedance Spectroscopy
GEO	Geosynchronous Earth Orbit
GRU	Gated Recurrent Unit

IC	Incremental Capacity
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MSE	Mean Squared Error
NCA	Nickel Cobalt Aluminum
NCM	Nickel Cobalt Manganese
OCV	Open-Circuit Voltage
ReLU	Rectified Linear Unit
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
SOC	State of Charge
SOH	State of Health
TJU	Tongji University
VCIC	Voltage-Charge Increment Curve
XJTU	Xi'an Jiaotong University

References

1. Tian, H.; Qin, P.; Li, K.; Zhao, Z. A review of the state of health for lithium-ion batteries: Research status and suggestions. *J. Clean. Prod.* **2020**, *261*, 120813. [\[CrossRef\]](#)
2. Tian, J.; Liu, X.; Li, S.; Wei, Z.; Zhang, X.; Xiao, G.; Wang, P. Lithium-ion battery health estimation with real-world data for electric vehicles. *Energy* **2023**, *270*, 126855. [\[CrossRef\]](#)
3. Alsuwian, T.; Ansari, S.; Ammiral Atiqi Mohd Zainuri, M.; Ayob, A.; Hussain, A.; Hossain Lipu, M.; Alhawari, A.R.; Almasabi, S.; Taher Hindi, A. A review of expert hybrid and co-estimation techniques for SOH and RUL estimation in battery management system with electric vehicle application. *Expert Syst. Appl.* **2024**, *246*, 123123. [\[CrossRef\]](#)
4. Vaideeswaran, V.; Bhuvanesh, S.; Devasena, M. Battery Management Systems for Electric Vehicles using Lithium Ion Batteries. In Proceedings of the 2019 Innovations in Power and Advanced Computing Technologies (i-PACT), Tamil Nadu, India, 22–23 March 2019; Volume 1, pp. 1–9. [\[CrossRef\]](#)
5. Huang, S.; Liu, C.; Sun, H.; Liao, Q. State of health estimation of lithium-ion batteries based on the regional frequency. *J. Power Sources* **2022**, *518*, 230773. [\[CrossRef\]](#)
6. Ge, M.F.; Liu, Y.; Jiang, X.; Liu, J. A review on state of health estimations and remaining useful life prognostics of lithium-ion batteries. *Measurement* **2021**, *174*, 109057. [\[CrossRef\]](#)
7. Berecibar, M.; Gandiaga, I.; Villarreal, I.; Omar, N.; Van Mierlo, J.; Van den Bossche, P. Critical review of state of health estimation methods of Li-ion batteries for real applications. *Renew. Sustain. Energy Rev.* **2016**, *56*, 572–587. [\[CrossRef\]](#)
8. Remmlinger, J.; Buchholz, M.; Meiler, M.; Bernreuter, P.; Dietmayer, K. State-of-health monitoring of lithium-ion batteries in electric vehicles by on-board internal resistance estimation. *J. Power Sources* **2011**, *196*, 5357–5363. [\[CrossRef\]](#)
9. Barai, A.; Uddin, K.; Dubarry, M.; Somerville, L.; McGordon, A.; Jennings, P.; Bloom, I. A comparison of methodologies for the non-invasive characterisation of commercial Li-ion cells. *Prog. Energy Combust. Sci.* **2019**, *72*, 1–31. [\[CrossRef\]](#)
10. Song, L.; Zhang, K.; Liang, T.; Han, X.; Zhang, Y. Intelligent state of health estimation for lithium-ion battery pack based on big data analysis. *J. Energy Storage* **2020**, *32*, 101836. [\[CrossRef\]](#)
11. Tang, K.; Luo, B.; Chen, D.; Wang, C.; Chen, L.; Li, F.; Cao, Y.; Wang, C. The state of health estimation of lithium-ion batteries: A review of health indicators, estimation methods, development trends and challenges. *World Electr. Veh. J.* **2025**, *16*, 429. [\[CrossRef\]](#)
12. Balasingam, B.; Ahmed, M.; Pattipati, K. Battery management systems—Challenges and some solutions. *Energies* **2020**, *13*, 2825. [\[CrossRef\]](#)
13. Zhang, M.; Dai, X.; Zeng, Q.; Tian, Y.; Xu, X. SOC and SOH Joint Estimation of Lithium-Ion Batteries Under Dynamic Current Rates Based on Machine Learning. *Symmetry* **2026**, *18*, 623. [\[CrossRef\]](#)
14. Meng, Z.; Agyeman, K.A.; Wang, X. Multi-segment state of health estimation of lithium-ion batteries considering short partial charging. *IEEE Trans. Energy Convers.* **2023**, *38*, 1913–1923. [\[CrossRef\]](#)
15. Bavand, A.; Khajehoddin, S.A.; Ardakani, M.; Tabesh, A. Online Estimations of Li-Ion Battery SOC and SOH Applicable to Partial Charge/Discharge. *IEEE Trans. Transp. Electr.* **2022**, *8*, 3673–3685. [\[CrossRef\]](#)
16. Xiong, R.; Zhang, Y.; Wang, J.; He, H.; Peng, S.; Pecht, M. Lithium-Ion Battery Health Prognosis Based on a Real Battery Management System Used in Electric Vehicles. *IEEE Trans. Veh. Technol.* **2019**, *68*, 4110–4121. [\[CrossRef\]](#)
17. Chen, S.Z.; Liang, Z.; Yuan, H.; Yang, L.; Xu, F.; Fan, Y. A novel state of health estimation method for lithium-ion batteries based on constant-voltage charging partial data and convolutional neural network. *Energy* **2023**, *283*, 129103. [\[CrossRef\]](#)

18. Gao, R.; Zhang, Y.; Lyu, Z. A SOH estimation method of lithium-ion batteries based on partial charging data. *J. Energy Storage* **2024**, *103*, 114309. [CrossRef]
19. Xiong, R.; Li, L.; Tian, J. Towards a smarter battery management system: A critical review on battery state of health monitoring methods. *J. Power Sources* **2018**, *405*, 18–29. [CrossRef]
20. Plett, G.L. Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 1. Background. *J. Power Sources* **2004**, *134*, 252–261. [CrossRef]
21. Xiong, R.; Sun, F.; Chen, Z.; He, H. A data-driven multi-scale extended Kalman filtering based parameter and state estimation approach of lithium-ion polymer battery in electric vehicles. *Appl. Energy* **2014**, *113*, 463–476. [CrossRef]
22. Lu, L.; Han, X.; Li, J.; Hua, J.; Ouyang, M. A review on the key issues for lithium-ion battery management in electric vehicles. *J. Power Sources* **2013**, *226*, 272–288. [CrossRef]
23. Ng, K.S.; Moo, C.S.; Chen, Y.P.; Hsieh, Y.C. Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries. *Appl. Energy* **2009**, *86*, 1506–1511. [CrossRef]
24. Schuster, M.; Paliwal, K.K. Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* **1997**, *45*, 2673–2681. [CrossRef]
25. Wang, F.; Zhai, Z.; Zhao, Z.; Di, Y.; Chen, X. Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis. *Nat. Commun.* **2024**, *15*, 4332. [CrossRef] [PubMed]
26. Weng, C.; Cui, Y.; Sun, J.; Peng, H. On-board state of health monitoring of lithium-ion batteries using incremental capacity analysis with support vector regression. *J. Power Sources* **2013**, *235*, 36–44. [CrossRef]
27. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
28. Nie, Y.; Nguyen, N.H.; Sinthong, P.; Kalagnanam, J. A Time Series is Worth 64 Words: Long-term Forecasting with Transformers. *arXiv* **2022**, arXiv:2211.14730.
29. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
30. Pecht, M. CALCE Battery Data. 2021. Available online: <https://calce.umd.edu/battery-data> (accessed on 15 January 2025).
31. Birkel, C.R. Diagnosis and Prognosis of Degradation in Lithium-Ion Batteries. Ph.D. Thesis, Department of Engineering Science, University of Oxford, Oxford, UK, 2017.
32. Wang, F.; Zhi, Q.; Zhao, Z.; Zhai, Z.; Liu, Y.; Xi, H.; Wang, S.; Chen, X. Inherently Interpretable Physics-Informed Neural Network for Battery Modeling and Prognosis. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, *36*, 1145–1159. [CrossRef]
33. Wang, F.; Zhai, Z.; Liu, B.; Zheng, S.; Zhao, Z.; Chen, X. Open access dataset, code library and benchmarking deep learning approaches for state-of-health estimation of lithium-ion batteries. *J. Energy Storage* **2024**, *77*, 109884. [CrossRef]
34. Zhu, J.; Wang, Y.; Huang, Y.; Bhushan Gopaluni, R.; Cao, Y.; Heere, M.; Mühlbauer, M.J.; Mereacre, L.; Dai, H.; Liu, X.; et al. Data-driven capacity estimation of commercial lithium-ion batteries from voltage relaxation. *Nat. Commun.* **2022**, *13*, 2261. [CrossRef] [PubMed]
35. Venugopal, P.; T, V. State-of-health estimation of li-ion batteries in electric vehicle using indrnn under variable load condition. *Energies* **2019**, *12*, 4338. [CrossRef]
36. He, J.; Meng, S.; Li, X.; Yan, F. Partial charging-based health feature extraction and state of health estimation of lithium-ion batteries. *IEEE J. Emerg. Sel. Top. Power Electron.* **2022**, *11*, 166–174.
37. Qu, J.; Liu, F.; Ma, Y.; Fan, J. A Neural-Network-Based Method for RUL Prediction and SOH Monitoring of Lithium-Ion Battery. *IEEE Access* **2019**, *7*, 87178–87191. [CrossRef]
38. Tian, J.; Xiong, R.; Shen, W.; Lu, J.; Yang, X.G. Deep neural network battery charging curve prediction using 30 points collected in 10 min. *Joule* **2021**, *5*, 1521–1534. [CrossRef]
39. Sun, H.; Yang, D.; Du, J.; Li, P.; Wang, K. Prediction of Li-ion battery state of health based on data-driven algorithm. *Energy Rep.* **2022**, *8*, 442–449. [CrossRef]
40. Raman, M.; Champa, V.; Prema, V. State of Health Estimation of Lithium Ion Batteries using Recurrent Neural Network and its Variants. In Proceedings of the 2021 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Online, 9–1 July 2021; pp. 1–6. [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.