energies

Article

A Robust Lithium-Ion Battery Capacity Prediction Framework Using Multi-Point Voltage Temporal Features and an OOF-Trained Adaptive Gating Mechanism

Lun-Yi Lung, Bo-Hao Zhou and Cheng-Chien Kuo *

Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei 106335, Taiwan; d11107004@mail.ntust.edu.tw (L.-Y.L.); d11407010@mail.ntust.edu.tw (B.-H.Z.)
* Correspondence: cckuo@mail.ntust.edu.tw

Abstract

Accurate capacity prediction is paramount for ensuring the operational safety and reliability of lithium-ion battery management systems (BMS). Nevertheless, contemporary data-driven approaches often grapple with limited feature representation—frequently relying solely on aggregate charging duration or noise measures—which compromises the robustness of these approaches. To address these limitations, this study proposes a robust framework integrating multi-point voltage temporal sampling (MVTS) with an adaptive gated hybrid ensemble learning strategy. The MVTS method is first used to extract high-dimensional geometric features from the constant-current (CC) charging phase (3.9 V–4.15 V), effectively capturing subtle degradation patterns. Subsequently, an unsupervised isolation forest algorithm is incorporated for automated anomaly detection and rectification, thereby augmenting data stability prior to training. In the fusion stage, a heterogeneous hybrid model comprising eXtreme gradient boosting (XGBoost) and long short-term memory (LSTM) is constructed. An adaptive gating mechanism based on random forest (RF) is added to dynamically weight the base learners. To mitigate data leakage during the stacking process, this study employs an out-of-fold (OOF) training strategy based on leave-one-battery-out (LOBO) cross-validation to generate unbiased meta-features for the gating model. This mechanism dynamically modulates fusion weights contingent upon the multi-point voltage features and model discrepancies, thereby accommodating diverse aging stages and capacity degradation patterns. Experimental results from the NASA battery aging dataset demonstrate that the proposed framework significantly outperforms single-model baselines in terms of RMSE and R^2, exhibiting superior adaptability and predictive precision.

Keywords: lithium-ion battery; capacity prediction; multi-point voltage temporal features; adaptive gated hybrid ensemble learning; out-of-fold training

Check for updates

Academic Editor: JongHoon Kim

Received: 26 February 2026
Revised: 18 March 2026
Accepted: 31 March 2026
Published: 2 April 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Lithium-ion batteries, owing to their high energy density, long cycle life, and low self-discharge rate, have become the dominant energy storage components in electric vehicles, large-scale grid energy storage systems, and a wide range of consumer electronic devices. However, under long-term charge–discharge cycling and complex operating conditions, batteries inevitably experience irreversible electrochemical aging mechanisms, including solid electrolyte interphase (SEI) growth, loss of lithium inventory, and structural

degradation of active electrode materials. These aging processes ultimately lead to capacity fading and increased internal resistance, thereby degrading system efficiency, driving range, and operational safety.

Since the state of health (SOH) of a battery cannot be directly measured by onboard sensors, accurate SOH estimation from limited and noise-contaminated measurements has become a core challenge in battery management systems (BMS). Reliable SOH estimation is essential not only for ensuring safe operation, but also for remaining useful life prediction and the implementation of optimal energy management strategies [1–3].

Existing SOH estimation approaches can generally be classified into model-based and data-driven methods. Model-based approaches rely on equivalent circuits or electrochemical models to describe internal battery dynamics. Although physically interpretable, these methods require complex parameter identification procedures. Moreover, model parameters exhibit pronounced time-varying behavior as batteries age, resulting in high computational costs for online updating and limiting their applicability in real-time onboard BMS implementations [1,3].

In contrast, data-driven methods establish nonlinear mappings between measurable signals, such as voltage, current, and temperature, and SOH without explicitly modeling internal battery mechanisms. Owing to their modeling flexibility and potential for cross-condition generalization, data-driven approaches have attracted increasing attention in recent years [3,4]. Various algorithms, including support vector regression, Gaussian process regression, and neural networks, have been systematically investigated and shown to provide satisfactory performance across different application scenarios [5].

Due to the highly nonlinear and multi-stage nature of battery aging, a single model often fails to maintain consistent performance throughout the entire lifecycle. As a result, recent studies have increasingly explored hybrid models and ensemble learning frameworks that combine complementary strengths in feature representation and temporal sequence modeling.

Previous work has shown that combining convolutional neural networks (CNNs) with recurrent neural networks, such as bidirectional gated recurrent unit (GRU) or long short-term memory (LSTM), enables simultaneous extraction of local geometric characteristics from voltage curves and long-term temporal dependencies, achieving improved SOH estimation accuracy on public battery datasets [6]. In addition, attention mechanisms and Transformer-based architectures have been introduced to enhance the modeling of critical degradation regions and long-sequence dependencies [7,8].

Despite these advances, data-driven SOH estimation in practical BMS applications still faces several challenges. Feature engineering plays a crucial role in estimation accuracy. Incremental capacity analysis (ICA) has been widely studied for its ability to reflect electrochemical behavior, but requires differentiation of the capacity–voltage curve, making it sensitive to measurement noise. Furthermore, under partial charging conditions or fragmented datasets, the stability and reproducibility of ICA-based features are often compromised [9–11]. In contrast, methods based on partial charging voltage segments or voltage trajectory features avoid differential operations and thus provide improved noise robustness and practical feasibility [12–14]. However, even with the introduction of deep learning architectures such as CNNs for automated feature extraction [15,16], these approaches often lack physical interpretability, particularly regarding the global geometric evolution of voltage trajectories during rapid capacity degradation stages.

In practical BMS applications, operational data are inevitably affected by sensor noise, communication losses, and irregular operating conditions, resulting in outliers and anomalous segments. Without appropriate data cleaning or uncertainty handling, such noise can significantly degrade model training and compromise SOH estimation reliability [17,18].

Although some studies have attempted to enhance robustness through algorithmic improvements, integrated frameworks that jointly address data cleaning, anomaly detection, and model learning remain limited [19].

Most existing hybrid SOH estimation methods adopt fixed-weight or globally averaged fusion strategies, such as static combinations of XGBoost and LSTM models [20]. These approaches implicitly assume consistent model performance throughout the battery lifecycle. However, according to dynamic ensemble selection (DES) theory, when system states, or data distributions, evolve due to concept drift, fixed fusion strategies fail to adaptively adjust model emphasis, leading to delayed responses or amplified prediction errors [21].

To address these limitations, mixture-of-experts (MoE) and gating architectures provide a promising alternative by dynamically allocating model weights according to input characteristics [22,23]. Such concepts have recently been introduced into battery health estimation tasks and demonstrated advantages over static fusion strategies [24]. For example, ref. [25] proposed a hierarchical knee-point strategy combined with an adaptive gated sequential network integrating IndRNN and AST-LSTM. While this approach highlights the potential of gating mechanisms, its feature extraction still relies on predefined curvature point detection, and the model architecture remains largely based on homogeneous neural network structures.

Based on the above considerations, this study proposes a new capacity estimation framework that integrates multi-point voltage temporal sampling (MVTS) with an adaptive gated hybrid ensemble learning strategy. The main contributions of this work are summarized as follows:

- Design of MVTS:
  The constant-current (CC) charging voltage curve is formalized as a high-dimensional geometric feature vector by extracting the corresponding time information at multiple predefined voltage anchor points. Crucially, by aligning the temporal sequence at the initial anchor 3.9 V, the framework effectively eliminates the interference of inconsistent initial SOC often encountered in practical applications. Beyond simple sampling, the inclusion of voltage variance within the sampling segment captures the subtle geometric contractions and slope changes induced by increasing polarization resistance as the battery ages. This comprehensive feature construction significantly enhances the model's sensitivity to non-linear degradation behaviors and ensures high feature stability across diverse experimental cycles.
- Isolation Forest-Based Data Cleaning Mechanism:
  An unsupervised anomaly detection and correction procedure is introduced prior to model training to automatically identify and repair measurement noise and outliers. Unlike conventional methods that simply discard anomalous data, the proposed mechanism employs linear interpolation to reconstruct corrupted charging cycles. This is particularly crucial for mitigating the impact of 'capacity regeneration' events caused by extended rest periods, ensuring the continuity and physical consistency of the training trajectory.
- Adaptive Gated Hybrid Ensemble Strategy:
  To effectively integrate heterogeneous learning paradigms for battery degradation modeling, an adaptive gated hybrid ensemble strategy is proposed. Unlike conventional hybrid models that rely on fixed-weight averaging or globally optimized fusion coefficients, the proposed approach introduces a context-aware gating mechanism based on a random forest (RF) regressor to dynamically combine the predictions of two complementary base learners: XGBoost and LSTM. By leveraging the strong nonlinear mapping capability of RF in high-dimensional feature spaces, the gate adaptively evaluates the reliability of each base learner at different aging stages, incorporating

both the MVTS geometric features and the prediction discrepancy between models. To further ensure reliable gate training and prevent information leakage, an out-of-fold (OOF) strategy is adopted to generate unbiased meta-features. This design allows the gating model to learn from realistic prediction error distributions rather than overfitted training outputs, which is particularly important when dealing with small battery datasets. Specifically, the gate takes as inputs: (i) the instantaneous predictions of XGBoost and LSTM, (ii) the magnitude of prediction disagreement between the two models, and (iii) the corresponding MVTS feature vector for each cycle. The gate outputs a cycle-wise adaptive weight ω, and the final SOH prediction is computed as $\hat{C}_{\text{final}} = \omega\hat{C}_{\text{XGB}} + (1-\omega)\hat{C}_{\text{LSTM}}$. This mechanism enables the framework to exploit the temporal modeling strength of LSTM during smooth degradation phases, while automatically increasing the contribution of XGBoost during rapid capacity decline or highly nonlinear aging stages. Consequently, the proposed adaptive gating strategy allows the hybrid model to dynamically adjust its predictive emphasis according to the evolving degradation characteristics, leading to improved accuracy and robustness compared with static ensemble approaches.

Overall, the novelty of this work lies not in a single algorithmic component but in the integration of physically interpretable voltage–trajectory representation MVTS, anomaly-aware preprocessing, and leakage-free adaptive hybrid learning within a unified capacity prediction framework.

2. Materials and Methods

The overall architecture of the proposed adaptive gated hybrid ensemble framework for SOH prediction is illustrated in Figure 1. The framework takes raw battery aging data from the NASA Prognostics Center of Excellence (PCoE, NASA Ames Research Center, Mountain View, CA, USA) as input and extracts voltage–time curves during the CC charging stage [26]. A cycle-level alignment procedure is then applied to ensure a consistent representation of charging behavior across different cycles.

Subsequently, an MVTS scheme is employed to transform continuous charging curves into high-dimensional temporal feature vectors, enabling the characterization of geometric variations in voltage trajectories associated with battery degradation. Considering that practical measurement data are often affected by noise and operational uncertainties, an unsupervised anomaly detection mechanism based on isolation forest is further introduced. Detected outliers are repaired using linear interpolation rather than being directly removed, thereby preserving the temporal continuity of the capacity degradation sequence.

After data cleaning, the extracted features are fed into two heterogeneous base learners in parallel: an XGBoost model focusing on static feature mapping and an LSTM model designed for sequential modeling. To achieve adaptive model fusion, an RF-based gating model is constructed, in which an OOF training strategy is adopted to generate unbiased samples for learning fusion weights while avoiding the risk of data leakage. Finally, the gating model dynamically assigns fusion weights according to the current voltage features and the disagreement between base model predictions, producing the final capacity estimation.

All computational experiments were implemented in Python (version 3.9.12, Python Software Foundation, Wilmington, DE, USA) using standard scientific computing libraries, including NumPy (version 1.22.3) and Pandas (version 1.4.2).

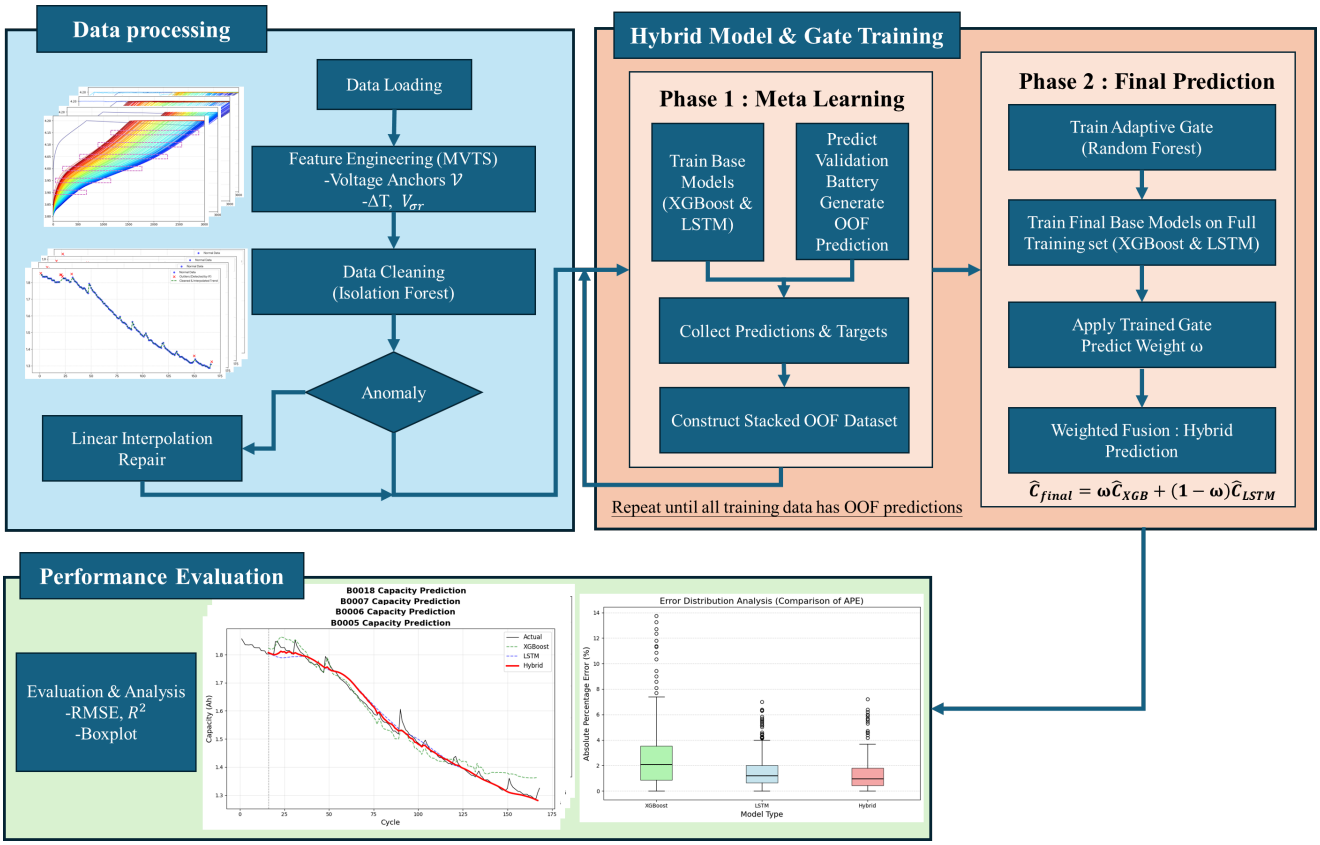


Figure 1. Overall workflow of the proposed adaptive gated hybrid ensemble framework for battery SOH prediction.

2.1. Feature Extraction via MVTS

During the CC charging stage of lithium-ion batteries, the geometric profile of the voltage–time curve evolves systematically as internal impedance increases. Compared with low-dimensional indicators such as total charging time or a single health indicator (HI), these coarse features are insufficient to characterize localized polarization effects and kinetic behaviors associated with different voltage regions.

To provide a more detailed representation of the geometric variations in voltage trajectories during battery aging, this study adopts an MVTS approach. By selecting multiple fixed voltage anchors within the CC charging region, the continuous voltage–time curve is transformed into a high-dimensional temporal feature vector. This representation enables the learning model to better capture nonlinear degradation behaviors reflected in different segments of the charging process.

Definition of Voltage Anchors

During the CC charging stage, the voltage–time trajectory of lithium-ion batteries exhibits systematic geometric evolution as aging progresses. Figure 2 illustrates the voltage curves of all charging cycles for the four batteries investigated in this study. Although the overall CC charging profile remains generally consistent, subtle yet progressive temporal expansion can be observed in the high-voltage region as cycle number increases. Specifically, the time required to reach the same voltage level gradually increases, reflecting the accumulation of internal resistance and polarization effects.

A few irregular traces appear in Figure 2. These curves originate from truncated or incomplete charge records in the raw NASA dataset, where the voltage sequence may start from an abnormally high level or exhibit abrupt voltage changes due to logging

interruptions. The formal data preprocessing and anomaly handling used for model development are performed separately in the subsequent stage described in Section 2.2.

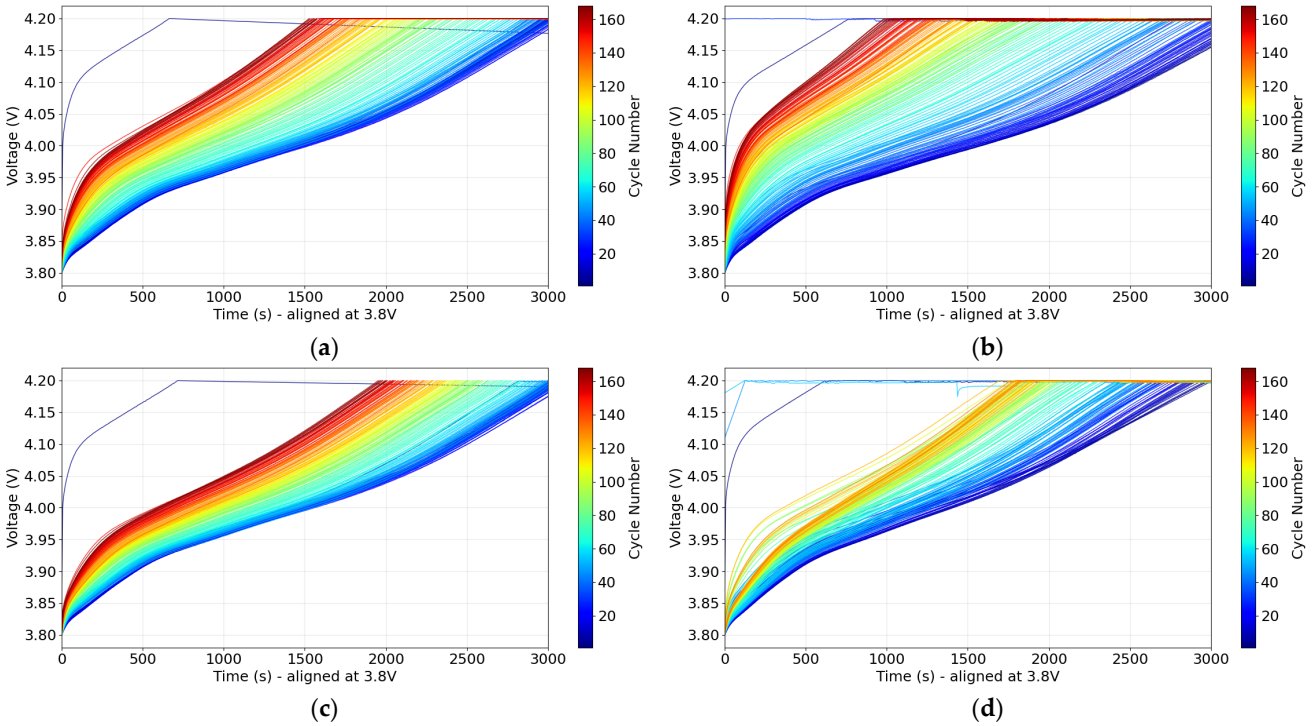


Figure 2. Raw voltage curves. (**a**) Battery: B0005; cycle 1-168. (**b**) Battery: B0006; cycle 1-168. (**c**) Battery: B0007; cycle 1-168. (**d**) Battery: B0018; cycle 1-132.

These observations indicate that the voltage trajectory contains rich degradation-related geometric information beyond coarse indicators such as total charging duration. Therefore, instead of representing each cycle using a single scalar feature, this study formalizes the voltage curve into multiple voltage-anchored temporal measurements through the proposed MVTS strategy.

Considering the high-voltage region of the charging curve in lithium-ion batteries, the voltage response in this interval tends to be more influenced by the combined effects of internal resistance growth and diffusion-related kinetic limitations, and thus may exhibit higher sensitivity to aging progression. Meanwhile, to avoid signal fluctuations and boundary data loss caused by the transition from CC to CV operation near the cutoff voltage of 4.2 V, a safety margin is introduced by setting the upper bound of the feature sampling range to 4.15 V, ensuring the robustness of feature extraction. The lower bound of 3.9 V is selected as a practical design choice to define a consistent sampling window within the CC charging stage. This region is close to the upper voltage limit, where variations in the voltage–time trajectory across cycles become more observable. In addition, empirical inspection of the NASA dataset indicates that the voltage segment between approximately 3.9 V and 4.15 V is consistently available across most cycles, providing a stable and reliable feature extraction window.

Accordingly, the voltage interval from 3.9 V to 4.15 V is selected as the sampling range, and a set of voltage anchors is defined with a fixed voltage step of $\Delta V = 0.05$ V:

$$\mathcal{V} = \{v_1, v_2, \ldots, v_k\} \tag{1}$$

where k denotes the total number of voltage anchors. Figure 3 illustrates the selected voltage anchors using the charging curve of battery B0005 at cycle 50 as an example.

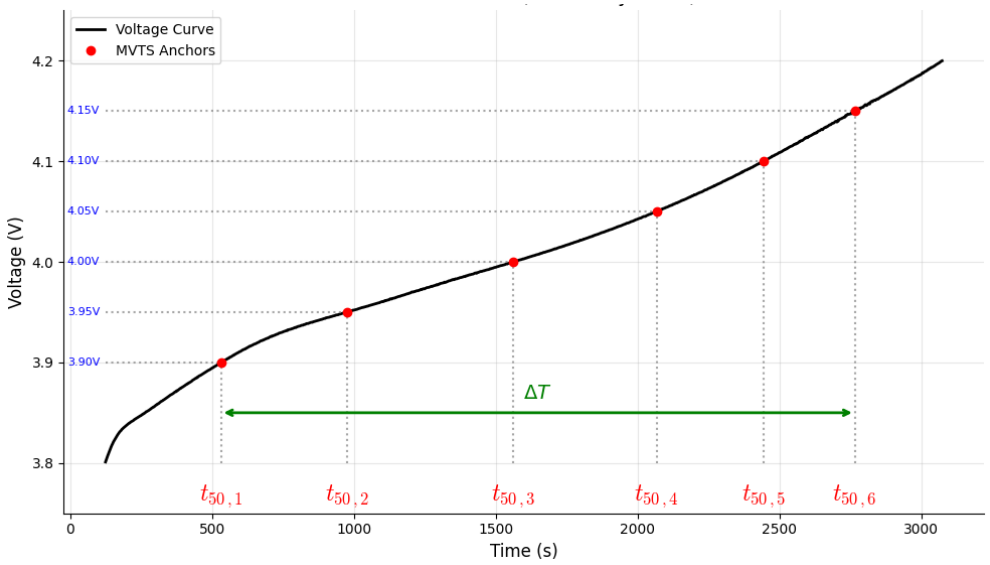


Figure 3. MVTS Schematic (B0005, cycle 50).

The selection of ΔV = 0.05 V is an optimized trade-off; while a larger step of 0.1 V may fail to capture sufficient geometric degradation features, a much finer step of 0.01 V can introduce redundant dimensionality and increase sensitivity to sensor noise, potentially leading to overfitting. A detailed sensitivity analysis comparing different voltage steps is provided in Section 3.6.

For each charging cycle i, a corresponding set of time stamps is extracted and denoted as

$$\mathrm{T}_i = \{t_{i,1}, t_{i,2}, \ldots, t_{i,k}\} \tag{2}$$

where $t_{i,k}$ represents the time at which the voltage in the i-th cycle first reaches the voltage anchor v_k. In practice, the measured voltage data are discretely sampled and often do not exactly coincide with the predefined voltage anchors. Simply selecting the nearest measured value may therefore lead to feature distortion or dimensional inconsistency. To ensure a consistent feature dimension across different cycles and sampling rates, linear interpolation is adopted to estimate the time stamp corresponding to each voltage anchor, which is defined as:

$$t_{i,k} = \mathrm{Interp}(v_k, \mathrm{V}_{i,measured}, \mathrm{T}_{i,measured}) \tag{3}$$

where $\mathrm{V}_{i,measured}$ and $\mathrm{T}_{i,measured}$ denote the measured voltage and time sequences of the i-th charging cycle, respectively. This interpolation procedure assumes a locally linear voltage variation between adjacent sampling points, allowing each charging cycle to be mapped into a fixed-length standardized temporal feature vector.

Based on the obtained time stamps, the following core features are further defined:

- Absolute time feature $t_{i,k}$:
  the time at which the voltage in the i-th charging cycle first reaches the anchor v_k.
- Interval time feature ΔT_i:
  defined as the total charging duration within the observed voltage interval.

$$\Delta T_i = t_{i,k} - t_{i,1} \tag{4}$$

- Voltage fluctuation feature $V_{\sigma r,i}$:
  defined as the variance of the measured voltage values within the specified voltage interval, which characterizes the dispersion of the voltage trajectory. A larger value indicates a less smooth voltage curve with stronger fluctuations.

$$V_{\sigma r,i} = \frac{1}{M_i} \sum_{j=1}^{M_i} \left(V_{i,measured,j} - \overline{v}_i \right)^2 \quad (5)$$

M_i denotes the total number of sampling points within the voltage range of 3.9–4.15 V for the i-th cycle, and $\overline{v}_i$ is the mean voltage over this interval.

To exclude abnormal charging behaviors caused by instrumentation faults or prolonged idle periods, any cycle with $\Delta T_i \geq 3000$ s is regarded as physically unreasonable and removed from the feature set.

Finally, each valid charging cycle i is represented by a multi-dimensional MVTS feature vector,

$$\text{MVTS}_i = [T_{i,1}, T_{i,2}, \ldots, T_{i,k}, \Delta T_i, V_{\sigma r,i}] \quad (6)$$

which serves as the input for subsequent model training and prediction.

2.2. Data Preprocessing and Anomaly Detection via Isolation Forest

In practical BMS operation, measurement data are inevitably affected by sensor noise, communication loss, and operational disturbances, which may introduce outliers in either the extracted features or the corresponding capacity labels. Directly using such contaminated data for model training can adversely affect learning stability and compromise prediction reliability, as anomalous samples may distort the regression surface and bias parameter updates in data-driven models.

To address this issue, an unsupervised anomaly detection method based on Isolation Forest is employed. This approach recursively partitions the feature space using randomly constructed trees, such that anomalous samples can be isolated with shorter average path lengths. Compared with distance- or density-based approaches, Isolation Forest does not rely on explicit distributional assumptions and is therefore well suited for high-dimensional MVTS feature representations. In this study, the contamination ratio is set to 0.04, and the anomaly detection model is trained independently at the individual battery level. Specifically, the algorithm evaluates both the high-dimensional MVTS features and the capacity labels concurrently, ensuring a comprehensive and aligned assessment across all data dimensions. This joint evaluation helps prevent the algorithm from mistakenly identifying physically meaningful degradation phenomena as anomalies. For example, gradual shifts in MVTS temporal features associated with internal resistance growth remain consistent with the degradation trajectory and therefore are not isolated by the algorithm.

This design enables effective discrimination between random measurement noise and battery-intrinsic physical behaviors. For instance, capacity regeneration phenomena caused by prolonged rest periods may appear as local capacity recoveries, as illustrated in Figure 4. Although such behavior manifests as short-term upward deviations, it preserves temporal continuity and electrochemical consistency across cycles. As a result, these samples are not erroneously identified as anomalies by the Isolation Forest, allowing physically meaningful aging dynamics to be retained for subsequent modeling.

For samples identified as anomalous, they are not directly removed from the dataset. Instead, the corresponding values are marked as missing and subsequently repaired through linear interpolation along the cycle axis. The choice of linear interpolation is justified by the generally smooth and monotonic nature of battery capacity degradation over short, localized cycle windows, thereby avoiding the introduction of artificial computational artifacts or abrupt discontinuities. This strategy suppresses the influence of measurement noise while preserving the temporal continuity of the capacity degradation sequence, which is particularly important for subsequent time-series modeling.

Ultimately, this robust preprocessing framework prevents the feature-driven XGBoost model from overfitting to extreme outliers and ensures that the sequence-driven LSTM model receives stable, uninterrupted historical trajectories. By stabilizing the data distribution prior to hybrid learning, the proposed cleaning mechanism enhances both model generalization capability and predictive robustness under practical operating conditions.

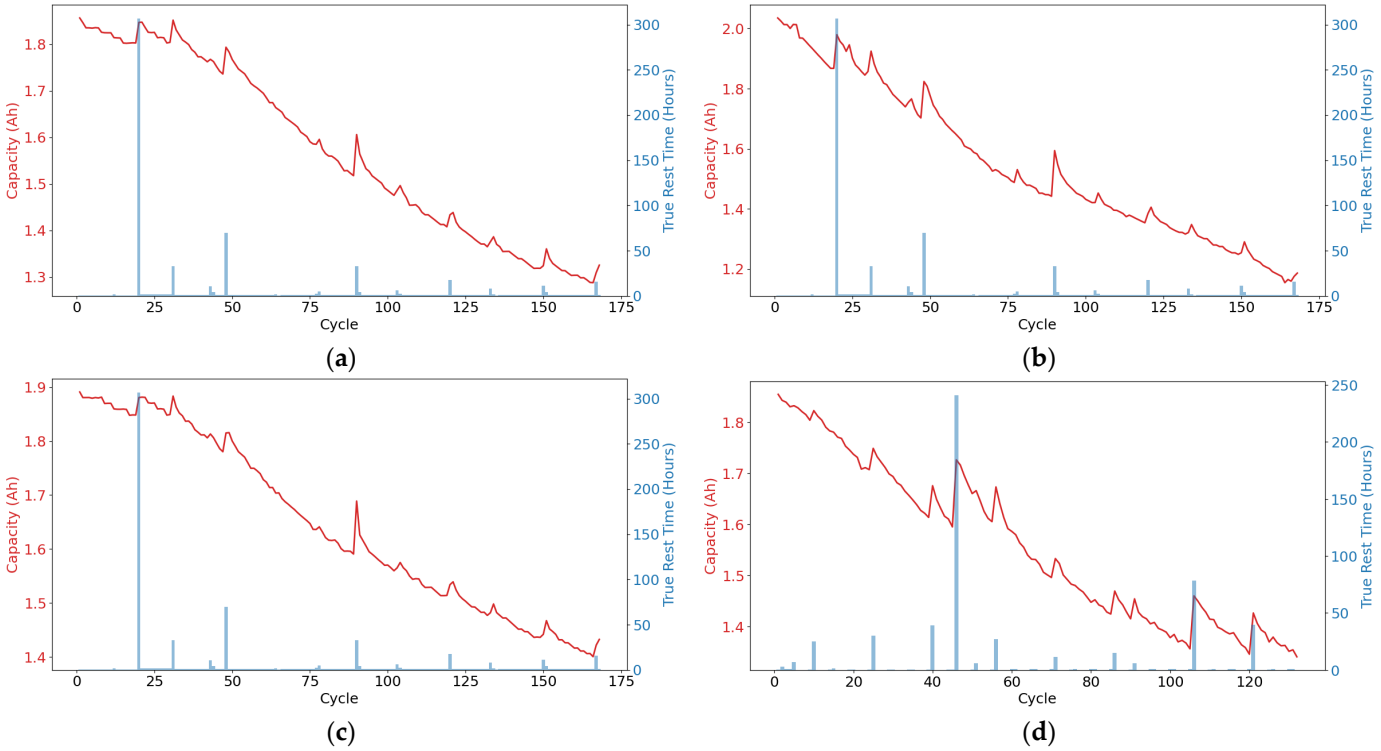


Figure 4. Capacity Regeneration and Rest Time. (**a**) Battery: B0005. (**b**) Battery: B0006. (**c**) Battery: B0007. (**d**) Battery: B0018.

2.3. Adaptive Gated Hybrid Ensemble Learning Strategy

To characterize lithium-ion battery aging behavior across different temporal scales, an ensemble learning framework that integrates feature space and temporal space modeling is adopted. Two heterogeneous base learners, XGBoost and LSTM, are employed to capture complementary aspects of the degradation process. Specifically, XGBoost focuses on decoding the instantaneous physical state reflected by the extracted features, whereas LSTM is used to model the long-term degradation trend through sequential dependency learning.

2.3.1. Feature-Driven Learner: XGBoost

The evolution of internal resistance in lithium-ion batteries directly induces geometric changes in the CC charging voltage curve, such as local slope variations and shifts in characteristic inflection points. These localized geometric alterations are closely associated with capacity degradation and often manifest over short time scales. To effectively capture such instantaneous, feature-level physical characteristics, XGBoost is adopted as a feature-driven learner in this study.

XGBoost is an ensemble learning algorithm developed within the gradient boosting framework, in which multiple decision tree weak learners are trained sequentially in an additive manner. Rather than directly re-estimating the target variable at each iteration, the model focuses on fitting the residuals generated by the preceding ensemble, thereby progressively correcting prediction bias. By incorporating second-order gradient information into the objective function and introducing regularization terms to control model

complexity, XGBoost achieves a balanced trade-off between bias and variance, leading to stable generalization performance in nonlinear feature spaces.

In the present implementation, XGBoost takes as input a high-dimensional voltage temporal feature vector extracted from the current cycle using the proposed MVTS scheme. The input features include the interval time feature (ΔT_i), the voltage fluctuation feature ($V_{\sigma r,i}$), and the time stamps corresponding to multiple voltage anchors ($t_{i,k}$). Regarding hyperparameter configuration, the number of estimators is set to 500 with a relatively small learning rate ($\eta = 0.01$), allowing the model to refine prediction errors in a gradual and stable manner. In addition, the maximum tree depth is limited to four to mitigate overfitting to noise in the training data and to enhance robustness on unseen cycles.

Compared with sequence-oriented models such as LSTM, XGBoost inherently operates as a static mapping model. Its primary strength lies in its sensitivity to localized nonlinear relationships in high-dimensional feature spaces, enabling effective detection of abrupt capacity drops and irregular degradation behaviors that may be associated with nonlinear aging mechanisms, such as lithium plating or micro-short phenomena. However, due to the absence of recurrent structures for retaining historical information across cycles, XGBoost is not well suited for directly modeling long-term temporal dependencies, resulting in relatively less smooth predictions along the time axis. This characteristic—high feature sensitivity with limited temporal memory—naturally complements the strengths of LSTM, providing the motivation for the adaptive gated fusion strategy introduced in the subsequent section.

2.3.2. Sequence-Driven Learner: LSTM

Battery aging is an accumulative process that strongly depends on historical operating states. Features extracted from a single cycle are often insufficient to characterize the long-term evolution of capacity degradation. To explicitly model the temporal dependencies embedded in the aging process, an LSTM network is employed as the sequence-driven learner in this study.

Compared with conventional recurrent neural networks (RNNs), which are prone to vanishing gradient issues when trained on long sequences, LSTM introduces a cell state (C_t) as the primary information carrier. This structure allows degradation-related information to be retained over extended sequences spanning hundreds of charge–discharge cycles. The internal structure of the LSTM unit adopted in this work is illustrated in Figure 5.

An LSTM cell consists of three gating mechanisms regulated by sigmoid (σ) and hyperbolic tangent (tanh) functions, which collectively control the direction and magnitude of information flow. The forget gate selectively preserves or discards information from the previous cell state C_{t-1}, thereby suppressing short-term fluctuations and non-informative noise. The input gate evaluates the current input features X_t to determine which degradation-related information should be incorporated into the cell state. Subsequently, the output gate generates the hidden state H_t based on the updated cell state C_t, which serves as the model output at the current time step. Through the coordinated operation of these gates, LSTM is able to attenuate short-term irregular variations while retaining long-term capacity degradation trends.

In the model implementation, a sliding look-back window with length $L = 15$ is adopted, where the capacity sequence from the previous 15 cycles is used as the input to the LSTM network, resulting in an input dimension of 15×1. The number of hidden units is set to 50, allowing the univariate capacity sequence to be projected into a higher-dimensional latent feature space to facilitate temporal representation learning. In addition, a dropout layer with a rate of 0.2 is applied after the LSTM layers to reduce the risk of overfitting and to improve generalization performance. This architectural configuration enables the

model to maintain prediction smoothness while capturing the overall evolution trend of capacity degradation.

Compared with the feature-driven XGBoost model, LSTM typically exhibits better continuity and prediction stability during phases of gradual and steady degradation. However, because its internal dynamics rely heavily on historical memory, the model response may become less sensitive when encountering abrupt capacity drops (e.g., capacity dives) or high-frequency oscillatory behaviors. This limitation provides a key motivation for introducing the adaptive gating mechanism in the subsequent hybrid ensemble framework.

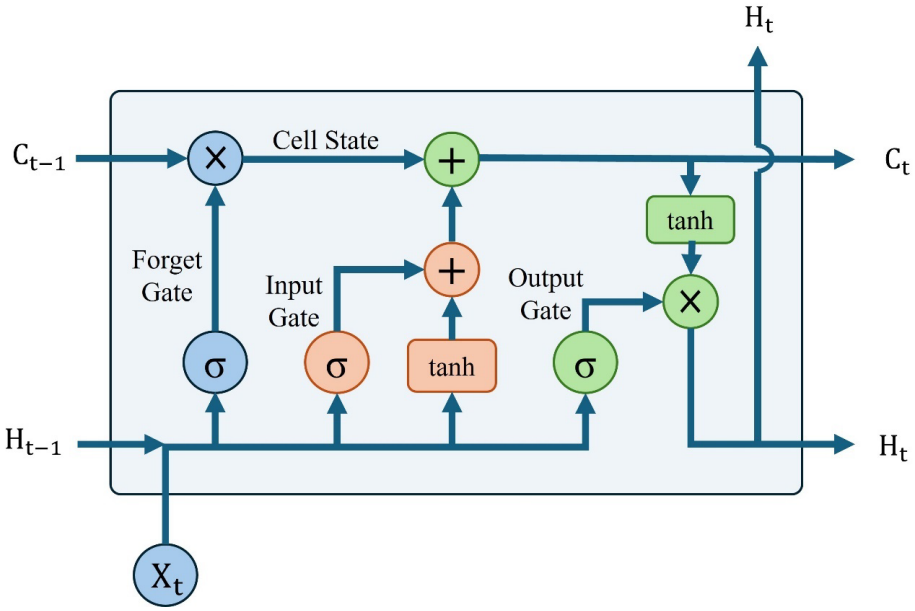


Figure 5. The internal structure of the LSTM unit.

2.3.3. Complementarity of Base Learners

XGBoost and LSTM exhibit complementary predictive characteristics at different stages of battery aging. XGBoost is more sensitive to instantaneous feature variations and local nonlinear patterns, whereas LSTM is more effective in capturing long-term temporal dependencies and gradual degradation trends. This heterogeneity provides the theoretical basis for the adaptive gated fusion mechanism, which dynamically adjusts the contribution of each base learner according to the current operating context.

The final capacity prediction is obtained through a weighted combination of the two base learners, expressed as

$$\hat{C}_{\text{final}} = \omega \hat{C}_{\text{XGB}} + (1 - \omega)\hat{C}_{LSTM} \tag{7}$$

where $\omega \in [0, 1]$ denotes the adaptive fusion weight generated by the gating model. The weight is estimated as

$$\omega = F_{\text{RF}}(X_{\text{meta}}) \tag{8}$$

with F_{RF} representing an RF regressor and X_{meta} denoting the meta-feature vector used for gating.

To train the RF-based gating model, predictions from the base learners are required as input features. However, directly using predictions obtained from data already seen during base learner training would result in overly optimistic performance estimates, leading to severe overfitting and data leakage. In such cases, the gating model would fail to learn how to handle realistic prediction errors.

To address this issue, an OOF strategy is adopted to generate unbiased meta-features. Specifically, an internal leave-one-battery-out (LOBO) cross-validation scheme is employed, following the procedure outlined below:

- Internal Data Splitting:
  Given a training set consisting of N batteries, one battery is iteratively held out as an internal validation set, while the remaining $N-1$ batteries are used to train the base learners.
- Unbiased Prediction Generation:
  XGBoost and LSTM models trained on the $N-1$ batteries are used to generate predictions for the held-out validation battery. Since this battery is excluded from the training process, the resulting predictions provide an unbiased estimate of model performance on unseen data.
- Stacked Dataset Construction:
  This process is repeated until all training samples obtain corresponding OOF predictions. The collected OOF predictions, together with the associated MVTS features, are then assembled to form a clean and leakage-free dataset for training the RF gating model.

Through this strategy, the feature distribution observed by the gating model during training remains consistent with that encountered during testing, thereby ensuring reliable generalization under practical deployment conditions.

Based on the data generated using the OOF strategy, the input vector (meta-features) of the gating model, denoted as X_{meta}, is defined to enable context-aware selection between base learners as

$$X_{\text{meta}} = \left[\hat{C}_{\text{XGB}}, \hat{C}_{\text{LSTM}}, \left|\hat{C}_{\text{XGB}} - \hat{C}_{\text{LSTM}}\right|, \text{F}_{\text{raw}}\right] \tag{9}$$

where $\left|\hat{C}_{\text{XGB}} - \hat{C}_{\text{LSTM}}\right|$ serves as an uncertainty indicator reflecting the prediction discrepancy between the two base learners. Unlike simple disagreement-based fusion strategies that rely solely on prediction discrepancies, the proposed gating mechanism also incorporates the MVTS feature vector describing the current voltage temporal characteristics of the battery. These features contain degradation-related geometric information from the charging trajectory and therefore provide additional physical context for model selection. As a result, the gating model is able to differentiate between transient prediction disagreements caused by measurement noise and systematic changes in degradation behavior associated with evolving aging mechanisms. A large discrepancy suggests that the data may lie in a non-stationary regime, prompting the gating mechanism to reallocate fusion weights accordingly. The term F_{raw} corresponds to the MVTS feature set, including $t_{i,k}$, ΔT_i, and $V_{\sigma r,i}$.

Since the gating model is trained in a supervised manner, an ideal fusion weight must be defined for each training sample. Using the known ground-truth capacity C_{true} and the prediction errors of the base learners, the optimal weight ω^* is derived by minimizing the fusion error, expressed as

$$\omega^* = \text{clip}\left(\frac{C_{\text{true}} - \hat{C}_{\text{LSTM}}}{\hat{C}_{\text{XGB}} - \hat{C}_{\text{LSTM}}}, 0, 1\right) \tag{10}$$

where the clipping operation constrains the weight to the interval $[0, 1]$. During training, the RF-based gating model F_{RF} is fitted to ω^* as the regression target. The final SOH prediction is then obtained using Equation (7).

An RF regressor is selected as the gating model due to its ability to handle high-dimensional feature inputs, such as MVTS features, while maintaining effective control over overfitting. Unlike simple linear weighting schemes, the RF-based gate learns a

context-dependent decision logic. For example, when the voltage fluctuation feature $V_{\sigma r,i}$ indicates highly unstable behavior near the end of discharge, the gating model may reduce the contribution of LSTM—whose predictions are more sensitive to short-term noise—and instead place greater emphasis on XGBoost, which is more responsive to localized geometric feature variations. In this framework, the RF acts as a meta-learner that does not directly estimate SOH but rather determines which base learner is more reliable under the current physical operating condition.

3. Results

3.1. Dataset and Experimental Configuration

To evaluate the effectiveness and cross-cell generalization capability of the proposed Adaptive Gated Hybrid Ensemble framework for capacity prediction, experiments were conducted using the lithium-ion battery aging dataset released by the NASA Prognostics Center of Excellence (PCoE). This dataset has been widely adopted as a benchmark for SOH estimation studies and is therefore considered representative for comparative evaluation.

Four lithium-ion batteries, namely B0005, B0006, B0007, and B0018, were selected for analysis. The dataset records multiple charge–discharge cycles for each battery until the capacity reaches the predefined end-of-life (EoL) threshold. All aging experiments were performed under a controlled ambient temperature of 24 °C.

According to the experimental protocol provided by the NASA PCoE dataset documentation, the charge–discharge procedure can be summarized as follows:

- Charge: Each battery was charged under a CC of 1.5 A until the terminal voltage reached 4.2 V, followed by a CV phase until the current decreased to 20 mA.
- Discharge: Batteries were discharged under a CC of 2 A to a cutoff voltage of 2.7 V for B0005, B0006, and B0007, and 2.5 V for B0018.

Although the NASA dataset has been widely used as a benchmark for battery health estimation research, it contains only a limited number of cells within each experimental group and relatively consistent operating conditions. In addition, the NASA repository includes several battery groups tested under different experimental protocols. However, these groups are typically subject to substantially different operating conditions, such as variations in ambient temperature, current profiles, and cycling strategies, which can lead to significantly different degradation patterns.

To avoid the confounding effects introduced by heterogeneous experimental conditions, this study focuses on four batteries from the same experimental group. This design ensures consistent testing conditions and allows the proposed framework to be evaluated under controlled laboratory aging scenarios.

3.2. Implementation Details

All experiments in this study were conducted using Python 3.8. The deep learning models were implemented based on the TensorFlow 2.x framework, while machine learning models were developed using the Scikit-learn and XGBoost libraries. To prevent data leakage, all normalization parameters were computed exclusively from the training set and subsequently applied to the corresponding test set. The detailed hyperparameter settings of each model are summarized in Table 1.

3.3. Evaluation Metrics

Referencing SOH estimation methods reported in [5,20], this study adopted the root mean square error (RMSE) and the coefficient of determination (R^2) as the primary evaluation metrics. RMSE is employed to quantify the magnitude of prediction errors with quadratic penalization, thereby reflecting the overall deviation between predicted and

measured capacities. R^2, on the other hand, is used to assess the goodness of fit and to indicate the proportion of variance in capacity degradation explained by the model. The mathematical definitions of these metrics are given as follows:

$$RMSE = \sqrt{\frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}\left(C_{\mathrm{true},i} - \hat{C}_{\mathrm{final},i}\right)^2} \tag{11}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(C_{\mathrm{true},i} - \hat{C}_{\mathrm{final},i}\right)^2}{\sum_{i=1}^{n}\left(C_{\mathrm{true},i} - \overline{C}_{\mathrm{true}}\right)^2} \tag{12}$$

$C_{true,i}$ denotes the measured capacity of the battery at the i-th cycle, $\hat{C}_{\mathrm{final},i}$ represents the corresponding predicted capacity, $\overline{C}_{\mathrm{true}}$ is the mean value of the measured capacities, and n denotes the total number of test samples. To ensure fair comparison, performance evaluation was conducted after temporal alignment, preventing bias arising from the look-back window used in sequence-based models.

Table 1. Detailed hyperparameter settings of each model.

Model Component	Parameter Description	Value
Data Preprocessing	Anomaly detection method	Isolation Forest
	Contamination rate	0.04
	Number of estimators	100
MVTS Feature Design	Voltage anchor range	3.90–4.15 V
	Voltage step size (ΔV)	0.05 V
	Feature composition	Cycle + ΔT + $V_{\sigma r}$ + T
Feature-driven Learner (XGBoost)	Number of boosted trees	500
	Maximum tree depth	4
	Learning rate (η)	0.01
	Subsample ratio	0.7
	Feature subsampling ratio	0.7
Sequence-driven Learner (LSTM)	Look-back window length (L)	15
	Hidden layer structure	2 × LSTM (50 units)
	Dropout probability	0.2
	Optimizer learning rate	α = 0.0005
	Batch size	32
Adaptive Gating Network (RF)	Number of trees	600
	Maximum depth constraint	7
	Minimum samples per leaf	10
	Input feature dimension	4 + KMVTS
Validation and Evaluation	Validation strategy	LOBO
	Evaluation metrics	RMSE, R^2

3.4. Performance Evaluation

3.4.1. Overall Performance Comparison

Table 2 summarizes the performance comparison between the proposed adaptive gated hybrid model and the individual base learners (XGBoost and LSTM) across the four batteries. As illustrated in Figure 6, the hybrid model achieves the lowest RMSE and the highest R^2 values in all test cases.

For the relatively complex aging pattern exhibited in B0006, which exhibited pronounced capacity drop, and B0018, characterized by relatively noisy degradation behavior, the hybrid model reduces RMSE by 21.32% and 12.01%, respectively, compared with the best-performing individual model. These results indicate that the adaptive gating mech-

anism enables effective integration of complementary predictive characteristics under complex degradation conditions. Even for B0005, where the degradation trajectory is comparatively smooth, the hybrid model provides a slight improvement in prediction accuracy without degrading performance stability.

Table 2. Comparison of prediction performance for different models on NASA battery dataset.

Test Battery	Model	RMSE	R^2	Improvement (vs. Best Single)
B0005	XGBoost	0.0307	0.9708	-
	LSTM	0.0184	0.9895	-
	Hybrid	0.0148	0.9933	19.57%
B0006	XGBoost	0.0689	0.9034	-
	LSTM	0.0394	0.9683	-
	Hybrid	0.0310	0.9804	21.32%
B0007	XGBoost	0.0336	0.9501	-
	LSTM	0.0250	0.9723	-
	Hybrid	0.0211	0.9804	15.60%
B0018	XGBoost	0.0415	0.8978	-
	LSTM	0.0308	0.9436	-
	Hybrid	0.0271	0.9565	12.01%

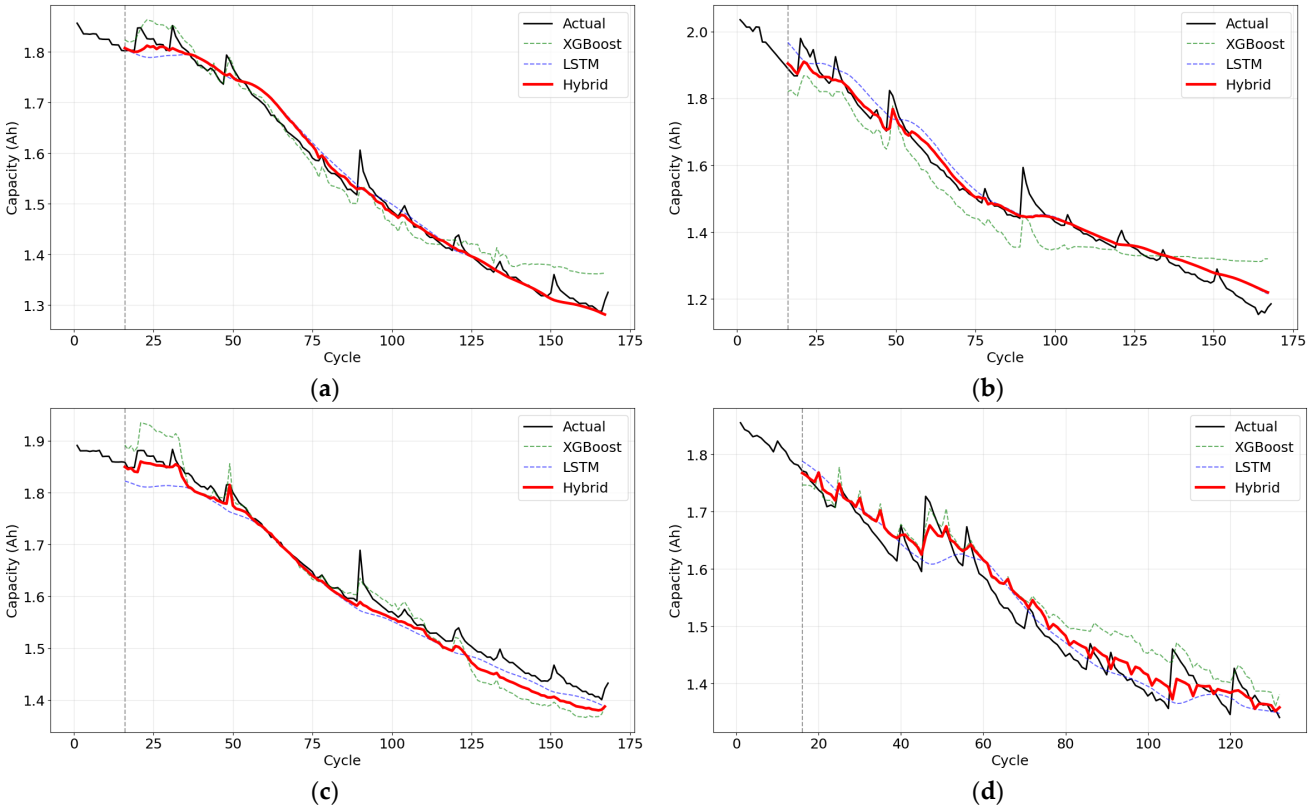


Figure 6. Prediction results for batteries. (**a**) Battery: B0005. (**b**) Battery: B0006. (**c**) Battery: B0007. (**d**) Battery: B0018.

To further examine the behavior of the proposed gating mechanism, the dynamic weight allocation curves for each battery are presented in Figure 7. The results show that the fusion weight ω assigned to XGBoost varies across cycles rather than remaining constant. For batteries with relatively smooth degradation patterns (e.g., B0005), the

weight gradually shifts toward the sequence-driven learner during stable aging phases. In contrast, for batteries exhibiting abrupt capacity changes or higher noise levels (e.g., B0006 and B0018), the weight distribution fluctuates more noticeably, reflecting the model's adaptive response to changing degradation characteristics. This observation confirms that the proposed gating network performs context-dependent weighting rather than static averaging, thereby allowing the hybrid model to adjust its reliance on each base learner according to the evolving physical state of the battery.

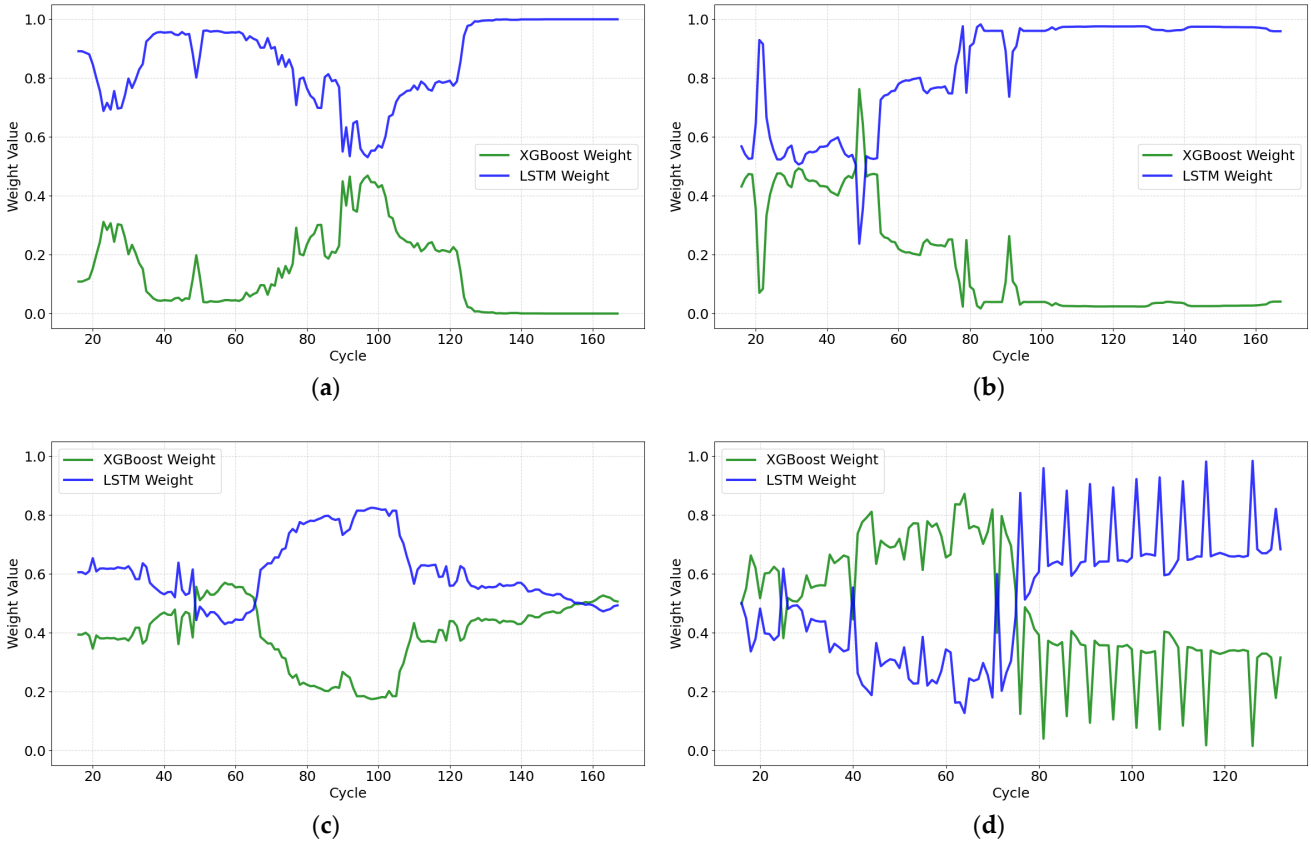


Figure 7. Dynamic weight allocation of the adaptive gating mechanism. (**a**) Battery: B0005. (**b**) Battery: B0006. (**c**) Battery: B0007. (**d**) Battery: B0018.

3.4.2. Analysis of Adaptive Gating Mechanism

For battery B0007, pronounced high-frequency oscillations are observed after cycle 120. Examination of the weight evolution indicates that the adaptive gating mechanism increases the contribution of XGBoost within this interval (approximately 0.4–0.5), suggesting that the gating model assigns greater reliance to the feature-driven learner when local fluctuations become more evident.

However, comparison with the prediction curves shows that the hybrid model does not consistently outperform the standalone LSTM in this region. In several later cycles, slight over-adjustment can be observed. This behavior reflects the inherent trade-off in the hybrid design when balancing sensitivity to local feature variations and temporal smoothness.

While XGBoost is capable of capturing non-monotonic changes in capacity, its prediction amplitude may occasionally be amplified by small perturbations in the input features. Although the adaptive gating mechanism correctly identifies the potential underfitting risk of LSTM during rapid fluctuations, achieving an optimal balance between the smooth bias of LSTM and the higher variance of XGBoost remains challenging under strongly oscillatory conditions.

Nevertheless, in terms of overall performance, the hybrid model achieves an RMSE of 0.0211 Ah on B0007, compared with 0.0250 Ah for LSTM. This result indicates that, despite localized discrepancies in highly oscillatory segments, the dynamic weighting strategy effectively limits extreme prediction errors at the global scale and maintains overall predictive stability.

3.4.3. Error Distribution Analysis

To further examine the reliability and stability of the proposed framework, the distribution of the absolute percentage error (APE) was analyzed. Figure 8 presents the boxplots of APE for XGBoost, LSTM, and the hybrid model on the test set, while Table 3 summarizes the corresponding statistical indicators.

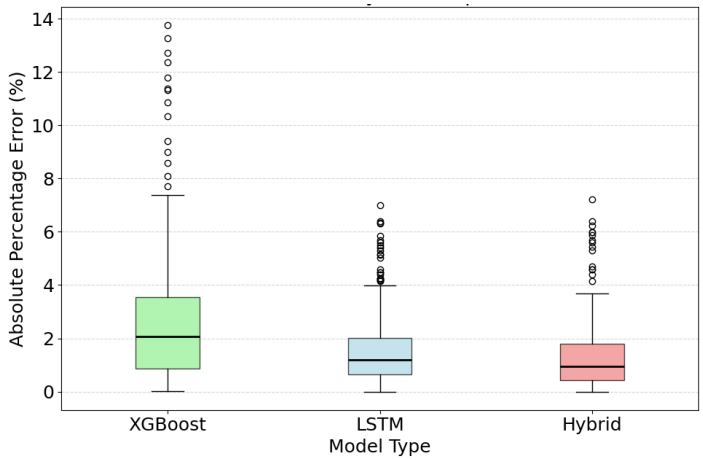


Figure 8. Error distribution analysis (comparison of APE).

Table 3. Descriptive statistics of absolute percentage error.

Statistic	XGBoost (%)	LSTM (%)	Proposed Hybrid (%)
Number of Samples	573	573	573
Mean	2.503	1.479	1.245
Standard Deviation	2.133	1.185	1.101
Minimum	0.004	0.001	0
25th Percentile (Q1)	0.853	0.639	0.42
Median (Q2)	2.07	1.198	0.947
75th Percentile (Q3)	3.54	2.014	1.805
Maximum	13.747	7.002	7.213

The results indicate that XGBoost exhibits relatively large variability in prediction errors. Its maximum APE reaches 13.75%, and multiple outliers are observed beyond the interquartile range. This suggests that the feature-driven model is more sensitive to nonlinear degradation behavior or high-noise segments.

In comparison, the hybrid model constrains the maximum error within 7.21%, reducing the likelihood of large prediction deviations during practical BMS operation. In terms of average accuracy, the hybrid model achieves a mean APE of 1.245%, compared with 1.479% for the best-performing standalone LSTM model, corresponding to an improvement of approximately 15.8%.

More importantly, the median APE of the hybrid model decreases to 0.947%, indicating that in more than half of the operational cycles, the prediction error remains below 1%. The standard deviation of APE is 1.101, which is the lowest among the three models. These

results suggest that the adaptive gating mechanism mitigates the error dispersion observed in individual learners and provides more stable SOH and capacity degradation estimation.

3.5. Impact of Data Cleaning

To examine the necessity of the proposed Isolation Forest-based anomaly detection scheme, this section compares model performance under two data processing strategies: without cleaning and with Isolation Forest cleaning. The results shown in Figure 9 illustrate the influence of data cleaning under different operating conditions.

- Noise mitigation effect:
  For batteries with relatively high measurement noise and observable outliers, such as B0006 and B0018, the introduction of the cleaning mechanism reduced RMSE by 50.9% and 38.3%, respectively. This indicates that identifying anomalous samples and repairing them via interpolation can effectively reduce the influence of random measurement noise and operational disturbances. As a result, the models are less susceptible to erroneous patterns, leading to improved prediction stability.
- Robustness under stable conditions:
  Even for B0005, whose degradation trajectory is comparatively smooth, the application of Isolation Forest resulted in a 2.9% reduction in RMSE. This suggests that the chosen anomaly detection setting maintains a reasonable balance between sensitivity and preservation of genuine degradation characteristics. The preprocessing scheme therefore serves not only as a corrective measure for noisy data but also as a stabilizing step under relatively clean operating conditions, demonstrating consistent behavior across different degradation patterns.

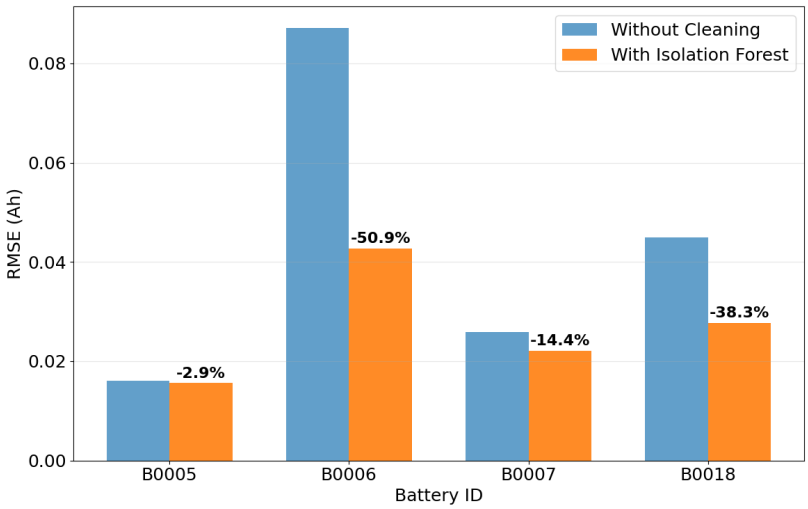


Figure 9. Comparison of RMSE with and without Isolation Forest data cleaning.

3.6. Sensitivity Analysis of Voltage Step Size

To justify the selection of the voltage step size $\Delta V = 0.05$ V and to ensure that the MVTS method is not based on an arbitrary configuration, a sensitivity analysis was conducted. The performance of the proposed hybrid framework was evaluated under three different sampling granularities: a coarse interval of $\Delta V = 0.1$ V, the proposed interval of $\Delta V = 0.05$ V, and a finer interval of $\Delta V = 0.01$ V. The results, averaged across all test batteries, are summarized in Table 4.

Table 4. Performance comparison of the hybrid model under different voltage step sizes.

Voltage Step ΔV	Feature Dimension	Mean RMSE (Ah)	Mean R^2
0.1 V	3	0.0253	0.9722
0.05 V	6	0.0235	0.9777
0.01 V	26	0.0236	0.9776

When a coarse step of ΔV = 0.1 V is employed, the mean RMSE increases to 0.0253 Ah and the mean R^2 decreases to 0.9722. This performance degradation suggests that a limited number of voltage anchors may not sufficiently capture the complex geometric variations in the voltage temporal trajectories during battery aging.

Conversely, reducing the step size to ΔV = 0.01 V increases the feature dimensionality from 6 to 26. However, this finer sampling does not provide a corresponding improvement in prediction accuracy. The mean RMSE slightly increases to 0.0236 Ah. This observation indicates that excessively dense sampling may introduce redundant information and amplify high-frequency measurement noise, which can increase the risk of overfitting and slightly affect the model's generalization capability.

Overall, the results indicate that ΔV = 0.05 V provides a favorable balance between feature resolution and model robustness. This configuration captures sufficient aging-related information while maintaining computational efficiency and noise tolerance. Therefore, ΔV = 0.05 V is adopted as the standard configuration for all experiments in this study.

4. Conclusions and Future Work

4.1. Conclusions

This study proposed a robust SOH prediction framework for lithium-ion batteries by integrating MVTS features with an adaptive gated hybrid ensemble learning strategy. To address limitations in existing data-driven approaches—particularly simplified feature representation and rigid fusion mechanisms—the main contributions and findings of this work are summarized as follows:

- Effective geometric feature representation:
  By transforming the CC charging voltage curve into a high-dimensional MVTS feature vector, the proposed method is able to capture geometric variations in the voltage trajectory induced by internal impedance evolution, such as slope changes and nonlinear distortions, occurring in the battery degradation process. This feature design enhances the model's ability to characterize degradation-related patterns.
- Noise-resilient data preprocessing scheme:
  An unsupervised anomaly detection procedure based on Isolation Forest was introduced, followed by linear interpolation repair. This approach mitigates the influence of outliers and measurement noise commonly observed in practical BMS data. Experimental results indicate that this preprocessing step improves the stability of subsequent model training under irregular operating conditions.
- Heterogeneous complementarity and context-aware fusion:
  At the model level, XGBoost and LSTM were employed as complementary base learners to capture instantaneous feature-based information and long-term degradation trends, respectively. An RF-based adaptive gating mechanism was further developed to enable cycle-level dynamic weight adjustment between the two complementary base learners.
- Consistent performance across test cases:
  Experimental evaluation on the NASA battery aging dataset shows that the proposed

adaptive gated model achieved the lowest RMSE and highest R^2 in most test cases, outperforming individual base models and conventional fixed-weight fusion strategies.

4.2. Future Work

Although the proposed framework demonstrates consistent performance, several aspects warrant further investigation with the aim of practical BMS deployment.

- Extension to diverse operating conditions:
  The current MVTS design is primarily based on the complete CC charging stage. In real-world electric vehicle applications, charging behaviors are often irregular and incomplete. Future work will explore extending the proposed feature extraction strategy to partial charging segments and constant-voltage phases, thereby improving applicability under incomplete data conditions. In addition, validating the model under dynamic discharge profiles remains an important direction for assessing generalization capability.
- Lightweighting model and edge deployment:
  While the hybrid architecture improves prediction accuracy, it also increases computational complexity. Future studies will investigate model compression techniques, such as pruning, quantization, and knowledge distillation, to facilitate deployment of the adaptive gating framework on resource-constrained embedded BMS platforms. The objective is to maintain predictive performance while satisfying real-time inference requirements.
- Diagnostic potential and interpretability of gating weights:
  The evolution of adaptive gating weights was observed to correlate with specific degradation phenomena, such as capacity regeneration and abrupt capacity decline. Future research may further examine the feasibility of utilizing weight transitions as diagnostic indicators. Significant shifts in model weighting could potentially reflect changes in internal electrochemical states, thereby providing complementary health-related information beyond scalar SOH estimation.
- Multi-physics integration:
  Battery degradation is strongly influenced by temperature. Incorporating temperature-related variables into the MVTS feature space, or treating them as contextual inputs to the gating network, may enhance model adaptability across varying thermal conditions. Developing a temperature-aware capacity prediction framework constitutes a meaningful extension of the present work.
- Distinction between sensor noise and meaningful transient behaviors:
  Isolation Forest is used as a statistical tool and does not explicitly distinguish sensor noise from physically meaningful transient behaviors. Detected anomalies are repaired via interpolation to preserve temporal continuity. This step is intended for robustness rather than physical interpretation. Future work will incorporate physical knowledge to improve this distinction.
- Robustness of the gating mechanism under degradation shifts:
  The RF-based gating model does not explicitly distinguish between temporary prediction disagreement and systematic degradation changes. Instead, it learns a data-driven reliability mapping based on prediction discrepancy and MVTS features. Future work will explore more interpretable gating strategies to better handle such distinctions under concept drift.

Author Contributions: Conceptualization, L.-Y.L. and B.-H.Z.; methodology, L.-Y.L.; software, L.-Y.L.; validation, B.-H.Z. and C.-C.K.; formal analysis, L.-Y.L. and B.-H.Z.; writing—original draft preparation, L.-Y.L. and B.-H.Z.; writing—review and editing, C.-C.K.; visualization, L.-Y.L.; supervision, C.-C.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data available in a publicly accessible repository.

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

Abbreviations

The following abbreviations are used in this manuscript:

SOH	State of health
BMS	Battery management systems
MVTS	Multi-point voltage temporal sampling
OOF	Out-of-fold training
SEI	Solid electrolyte interphase
CNNs	Convolutional neural networks
RNNs	Recurrent neural networks
GRU	Gated recurrent unit
DES	Dynamic ensemble selection
LSTM	Long short-term memory
XGBoost	eXtreme gradient boosting
CC	Constant-current
CV	Constant-voltage
HI	Health indicator
LOBO	Leave-one-battery-out
RF	Random forest
APE	Absolute percentage error
EoL	End-of-life
MoE	Mixture-of-experts
RMSE	Root mean square error
R^2	Coefficient of determination

References

1. Yao, L.; Xu, S.; Tang, A.; Zhou, F.; Hou, J.; Xiao, Y.; Fu, Z. A Review of Lithium-Ion Battery State of Health Estimation and Prediction Methods. *World Electr. Veh. J.* **2021**, *12*, 113. [CrossRef]
2. Shu, X.; Shen, S.; Shen, J.; Zhang, Y.; Li, G.; Chen, Z.; Liu, Y. State of Health Prediction of Lithium-Ion Batteries Based on Machine Learning: Advances and Perspectives. *iScience* **2021**, *24*, 103265. [CrossRef] [PubMed]
3. Ren, Z.; Du, C. A Review of Machine Learning State-of-Charge and State-of-Health Estimation Algorithms for Lithium-Ion Batteries. *Energy Rep.* **2023**, *9*, 2993–3021. [CrossRef]
4. Cai, L.; Lin, J.; Liao, X. A Data-Driven Method for State of Health Prediction of Lithium-Ion Batteries in a Unified Framework. *J. Energy Storage* **2022**, *51*, 104371. [CrossRef]
5. Kim, E.; Jung, S. Battery State of Health Estimation Methods: Implementation and Comparison of 11 Algorithms. *IEEE Access* **2025**, *13*, 55599–55615. [CrossRef]
6. Mazzi, Y.; Ben Sassi, H.; Errahimi, F. Lithium-Ion Battery State of Health Estimation Using a Hybrid Model Based on a Convolutional Neural Network and Bidirectional Gated Recurrent Unit. *Eng. Appl. Artif. Intell.* **2024**, *127*, 107199. [CrossRef]
7. Zhao, H.; Chen, Z.; Shu, X.; Shen, J.; Lei, Z.; Zhang, Y. State of Health Estimation for Lithium-Ion Batteries Based on Hybrid Attention and Deep Learning. *Reliab. Eng. Syst. Saf.* **2022**, *15*, 5599. [CrossRef]
8. Zhang, Y.; Li, Z.; Kong, L.; Xu, H.; Shen, H.; Chen, M. Multi-Step State of Health Prediction of Lithium-Ion Batteries Based on Multi-Feature Extraction and Improved Transformer. *J. Energy Storage* **2025**, *105*, 114538. [CrossRef]
9. Li, X.; Yuan, C.; Li, X.; Wang, Z. State of Health Estimation for Li-Ion Battery Using Incremental Capacity Analysis and Gaussian Process Regression. *Energy* **2020**, *190*, 116467. [CrossRef]

10. Huang, S.; Wang, X.; Kang, L.; Xie, D.; Zhang, X. State of Health Estimation for Lithium-Ion Battery Using Partial Incremental Capacity Curve and Transfer Learning. *Batteries* **2024**, *10*, 324. [CrossRef]
11. Xu, Z.; Chen, Z.; Yang, L.; Zhang, S. State of Health Estimation for Lithium-Ion Batteries Based on Incremental Capacity Analysis and Transformer Modeling. *Appl. Soft Comput.* **2024**, *165*, 112072. [CrossRef]
12. Wu, J.; Liu, Z.; Zhang, Y.; Lei, D.; Zhang, B.; Cao, W. Data-Driven State of Health Estimation for Lithium-Ion Battery Based on Voltage Variation Curves. *J. Energy Storage* **2023**, *73*, 109191. [CrossRef]
13. Shu, X.; Yang, H.; Liu, X.; Feng, R.; Shen, J.; Hu, Y.; Chen, Z.; Tang, A. State of Health Estimation for Lithium-Ion Batteries Based on Voltage Segment and Transformer. *J. Energy Storage* **2025**, *108*, 115200. [CrossRef]
14. Feng, X.; Weng, C.; He, X.; Han, X.; Lu, L.; Ren, D.; Ouyang, M. Online State-of-Health Estimation for Li-Ion Battery Using Partial Charging Segment Based on Support Vector Machine. *IEEE Trans. Veh. Technol.* **2019**, *68*, 8583–8593. [CrossRef]
15. 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]
16. Liu, S.; Chen, Z.; Yuan, L.; Xu, Z.; Jin, L.; Zhang, C. State of Health Estimation of Lithium-Ion Batteries Based on Multi-Feature Extraction and Temporal Convolutional Network. *J. Energy Storage* **2024**, *75*, 109658. [CrossRef]
17. Akbar, K.; Zou, Y.; Awais, Q.; Baig, M.J.A.; Jamil, M. A Machine Learning-Based Robust State of Health (SOH) Prediction Model for Electric Vehicle Batteries. *Electronics* **2022**, *11*, 1216. [CrossRef]
18. Rasul, M.J.M.A.; Abbas, A.; Baek, J.; Kim, J. A Hybrid Ensemble Learning Framework with Uncertainty Quantification for State-of-Health Estimation in Lithium-Ion Batteries. *Measurement* **2026**, *266*, 120528. [CrossRef]
19. Shang, S.; Xu, Y.; Zhang, H.; Zheng, H.; Yang, F.; Zhang, Y.; Wang, S.; Yan, Y.; Cheng, J. Enhancing Lithium-Ion Battery State-of-Health Estimation via an IPSO-SVR Model: Advancing Accuracy, Robustness, and Sustainable Battery Management. *Sustainability* **2025**, *17*, 6171. [CrossRef]
20. Khan, W.; Aziz, A.; Yousaf, M.Z.; Renhai, F.; Mao, Y.; Latif, M.W.; Tonghao, W.; Khan, B.; Mabunda, N.E.; Ali, A. Hybrid XGBoost–LSTM Model for State-of-Health Prediction of Lithium-Ion Batteries under Different Thermal Conditions. *Int. J. Electrochem. Sci.* **2026**, *21*, 101218. [CrossRef]
21. Cruz, R.M.O.; Sabourin, R.; Cavalcanti, G.D.C. Dynamic Classifier Selection: Recent Advances and Perspectives. *Inf. Fusion* **2018**, *41*, 195–216. [CrossRef]
22. Shazeer, N.; Mirhoseini, A.; Maziarz, K.; Davis, A.; Le, Q.; Hinton, G.; Dean, J. Outrageously Large Neural Networks: The Sparsely-Gated Mixture-of-Experts Layer. *arXiv* **2017**, arXiv:1701.06538.
23. Fedus, W.; Zoph, B.; Shazeer, N. Switch Transformers: Scaling to Trillion Parameter Models with Simple and Efficient Sparsity. *J. Mach. Learn. Res.* **2022**, *23*, 1–40. [CrossRef]
24. Chen, D.; Zhou, X. AttMoE: Attention with Mixture of Experts for Remaining Useful Life Prediction of Lithium-Ion Batteries. *J. Energy Storage* **2024**, *84*, 110780. [CrossRef]
25. Wang, K.; Gao, Q.; Pang, X.; Li, H.; Liu, W. Estimating the Health State of Lithium-Ion Batteries Using an Adaptive Gated Sequence Network and Hierarchical Feature Construction. *Batteries* **2024**, *10*, 278. [CrossRef]
26. Saha, B.; Goebel, K. Battery data set. In *NASA AMES Prognostics Data Repository*; NASA Ames Research Center: Mountain View, CA, USA, 2007. Available online: https://www.nasa.gov/intelligent-systems-division/discovery-and-systems-health/pcoe/pcoe-data-set-repository/ (accessed on 5 December 2025).

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.