mathematics

MDPI

Article

State of Health Prediction of Pure Electric Vehicle Batteries Based on Patch Wavelet Transformer

Min Wei [1,2,3,†], Siquan Yuan [1,2,†], Lin Chen [1,2], Yuhang Liu [1,2] and Jie Hu [1,2,*]

[1] School of Automotive Engineering, Wuhan University of Technology, Wuhan 430070, China; minwei@whut.edu.cn (M.W.); siquan_yuan@whut.edu.cn (S.Y.); 333929@whut.edu.cn (L.C.); 13050182710@163.com (Y.L.)
[2] Hubei Key Laboratory of Advanced Technology for Automotive Components, Wuhan University of Technology, Wuhan 430070, China
[3] Technical Development Center, SAIC, General Wuling Automobile Co., Ltd., Liuzhou 545007, China
* Correspondence: auto_hujie@whut.edu.cn
† These authors contributed equally to this work as co-first authors.

Abstract

The accuracy of onboard power battery capacity prediction is often limited due to excessive reliance on laboratory data and the neglect of complex usage environments. To address this issue, a Transformer-based model, named PWT, which integrates a patching strategy with wavelet decomposition, is proposed. By utilizing multi-scale temporal feature extraction and attention mechanisms, the model effectively enhances the capability of battery degradation modeling. To tackle the challenge of limited capacity label availability, a fuzzy Kalman filtering model based on ampere-hour integration is designed, reducing the relative error in SOH estimation to 0.906% and significantly improving label accuracy. Furthermore, a charging behavior scoring mechanism based on fuzzy membership functions and a current–temperature interaction feature matrix is constructed to enhance the model's sensitivity to degradation factors. Experimental results show that the proposed method outperforms LSTM, Transformer, and PatchTST under various real-world operating conditions, achieving a worst-case RMSE of 0.0226, MAPE of 0.0725, and R^2 of 0.903, demonstrating higher accuracy, robustness, and computational efficiency. In conclusion, the proposed method exhibits promising prospects in both theoretical research and engineering applications, providing a novel and effective approach to power battery health management.

Keywords: SOH prediction; real-world vehicle data; data-driven approach; deep learning

MSC: 68T07; 68T05

check for updates

Academic Editor: Jonathan Blackledge

Received: 28 May 2025
Revised: 20 June 2025
Accepted: 23 June 2025
Published: 26 June 2025

Citation: Wei, M.; Yuan, S.; Chen, L.; Liu, Y.; Hu, J. State of Health Prediction of Pure Electric Vehicle Batteries Based on Patch Wavelet Transformer. *Mathematics* **2025**, *13*, 2099. https://doi.org/10.3390/math13132099

Copyright: © 2025 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 (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Under the global carbon neutrality strategy, lithium-ion batteries, as core components of electric vehicles and energy storage systems, are projected to reach an installed capacity of 4.7 TWh by 2030. The accurate prediction of their SOH has become a focal point of interest for both academia and industry. The progressive degradation of battery performance—such as capacity fade and impedance rise—directly affects system reliability and economic efficiency. Studies have shown that, when the SOH deviation among cells in a power battery pack exceeds 15%, the available system capacity may drop by over 22%, and battery replacement costs in energy storage stations can account for up to 34% of the total life-cycle cost. Therefore, accurately predicting the SOH of batteries is not only key to optimizing

maintenance strategies but also a core technical challenge for enabling battery second-life applications and carbon footprint tracking.

SOH prediction methods for lithium-ion batteries mainly fall into two paradigms: mechanism-driven and data-driven approaches. Early studies were centered on electrochemical models and statistical learning methods. Li et al. [1] improved the resampling strategy of the particle filter algorithm, reducing SOH prediction error to within 5%; however, the model relies on predefined capacity degradation equations, making it difficult to adapt to the dynamic charging strategies encountered in real-world scenarios. Wang et al. [2] proposed a particle flow filtering framework combined with a gray model, which compensates for the capacity regeneration phenomenon and demonstrated superior performance over traditional Kalman filtering under laboratory cycling data. Nevertheless, its parameters must be calibrated for specific battery chemistries, limiting its generalizability across different materials. The fundamental bottleneck of such methods lies in the multi-scale coupled processes of battery aging, such as SEI film growth and lithium dendrite formation. Constructing precise physical models faces dual constraints of computational complexity and parameter identifiability [3]. Moreover, frequent shallow charge–discharge cycles in real-world vehicle operations obscure features in ICA curves, rendering feature extraction methods based on complete charge–discharge cycles ineffective [4].

The emergence of deep learning has opened new avenues for data-driven approaches. Ren et al. [5] were among the first to apply DNN to RUL prediction, establishing an end-to-end mapping between voltage–current time series and capacity degradation. Their model achieved an MSE of 0.0023 on the NASA dataset. Zhao et al. [6] proposed a hybrid architecture that further enhanced feature extraction capabilities: CNNs were used to capture local aging patterns, while BiLSTM modeled long-term dependencies. An attention mechanism dynamically weighted key time-step features, resulting in a 37% reduction in prediction error compared to conventional LSTM models. Hou et al. [7] constructed a prediction model based on the ANN, but its performance was suboptimal in handling nonlinear temporal degradation under dynamic operating conditions. Mamo et al. [8] demonstrated the effectiveness of LSTM in capturing long-term trends in capacity fade but noted its limited sensitivity to local transient features such as fast-charging pulses. Although these methods performed well on controlled laboratory datasets, they often overlooked the complexities of real-world scenarios—such as current fluctuations caused by dynamic charging strategies and internal temperature gradients within battery packs.

Zhang et al. [9] developed a cross-cell deep learning framework that shared hidden layer parameters to enable knowledge transfer across different degradation paths, achieving an MAE of 1.8% under ten charging/discharging protocols. However, their model lacked sufficient representation of nonlinear effects from charging rates and electro-thermal coupling, with significantly increased prediction errors under local overheating conditions caused by fast charging.

To address the challenges of data scarcity and operating condition diversity, transfer learning and novel feature engineering methods have become research hotspots. Iftikhar et al. [10] built a domain-adversarial deep model that maximized feature distribution alignment between source and target domains, reducing prediction error by 21% in cross-temperature experiments. Sun et al. [11] proposed a TrAdaBoost-based transfer learning framework that leveraged laboratory data to support real-vehicle prediction. However, their feature alignment strategy failed to account for the asymmetric impact of the charging rate across different SOC ranges, leading to unquantified aging acceleration in the high-SOC fast-charging region. Wang et al. [12] developed a cross-domain prediction model via degradation trend consistency analysis, and Ma et al. [13] employed deep transfer learning for personalized health prediction. Nevertheless, these methods still face limitations in

feature decoupling under dynamic conditions—particularly in modeling the interaction between ambient temperature fluctuations and charging/discharging strategies.

On the feature engineering front, Tang et al. [3] used ICA to extract peak voltage shifts as degradation features and optimized nonlinear inflection point prediction with Gaussian kernel functions. However, this method depends on high-resolution voltage data and exhibits poor stability under real-world noise. Fan et al. [14] constructed aging indicators using relaxation voltage time constants and successfully predicted degradation inflection points but required several hours of rest-state data, which limits real-time applicability. Krishna et al. [15] combined IoT sensor data with machine learning, using surface strain signals to infer lithium plating levels—offering a novel approach for multiphysics modeling. However, their experiments focused solely on cylindrical cells, without validation for prismatic or pouch cells. Hu et al. [4] were the first to build an SOH estimation model based on real-vehicle data, revealing correlations between behavior features such as the number of rapid accelerations and the initial SOC of charging with capacity fade. Yet, their linear regression model failed to capture the nonlinear effects of charging rates across SOC ranges, leading to reduced prediction stability under high-dynamic conditions.

Probabilistic prediction and real-time optimization have become prominent research directions in recent years. Li et al. [16] developed an interpretable online prediction method that enhances model transparency through feature importance ranking. However, the improvement in interpretability comes at the cost of reduced prediction accuracy. Ly et al. [17] proposed a T-shaped data structure and a quantile LSTM framework to achieve non-crossing quantile prediction, effectively quantifying RUL uncertainty. Nevertheless, its computational complexity limits its applicability in onboard systems. Huang et al. [18] innovatively adopted an unsupervised learning method without labeled data by constructing a loss function based on the self-similarity of degradation curves, alleviating the high cost of data annotation. However, the reliability of predictions during the early degradation stage still requires validation. Jafari et al. [19] combined the Harris Hawk Optimization algorithm with Random Forest/LightGBM models to enable automatic hyperparameter tuning, reducing prediction errors by 12% on public datasets, though they did not address the temporal non-stationarity present in real-world vehicle data. Addressing the phenomenon of capacity regeneration, Pang et al. [20] proposed a detection algorithm based on transient relaxation features, effectively distinguishing between true degradation and capacity fluctuations. However, its threshold settings rely on empirical parameters, limiting its adaptability. Cheng et al. [21] combined empirical mode decomposition with LSTM to enhance prediction stability through signal denoising, but the issue of mode mixing resulted in the incomplete suppression of high-frequency noise. Kwon et al. [22] constructed a multi-level LSTM framework to simultaneously predict SOH and RUL while accounting for inconsistencies in cell voltage, yet they did not incorporate spatial distribution features such as temperature gradients, leaving the characterization of pack-level aging incomplete.

These developments highlight three core bottlenecks:

1. Existing models struggle with feature representation under real-world dynamic operating conditions. Traditional LSTM/CNN architectures lack temporal sensitivity to local degradation patterns, making it difficult to capture abnormal aging caused by transient behaviors such as fast-charging pulses or sudden temperature changes;
2. In cross-domain transfer learning, the distribution gap between laboratory and real-world vehicle data has not been effectively addressed through feature space reconstruction. In particular, the randomness in charging strategies leads to feature shifts that remain unmodeled;
3. The nonlinear modulation mechanisms of aging trajectories caused by multiphysics interactions—such as the dynamic coupling of current, temperature, and mechanical

stress—remain insufficiently explored. Existing feature engineering mainly focuses on one-dimensional signal analysis and lacks the quantitative modeling of cross-scale effects such as electrothermal coupling.

To address the challenges of accurate SOH estimation and prediction under real-world operating conditions, this study proposes a novel framework, PWT, that integrates multi-scale time-frequency analysis and multiphysics-aware feature modeling. The approach significantly enhances the model's ability to capture complex degradation dynamics, improve predictive accuracy, and generalize across diverse scenarios. The key innovations are as follows:

1. The proposed PWT architecture reorganizes temporal aging features into local patches and applies wavelet decomposition to multi-dimensional degradation signals. This enables multi-scale time-frequency analysis and reconstruction, strengthening local temporal correlations and enhancing the model's capacity to learn from non-stationary sequences common in real-world battery data.
2. To capture the nonlinear effects of varying charging rates across different SOC intervals, a novel charging behavior scoring mechanism is designed using fuzzy membership functions. This quantifies degradation sensitivity under diverse charging conditions and embeds domain-specific knowledge into the model.
3. An interaction feature matrix of current and temperature is constructed to reflect the electrothermal coupling effects. Tensor decomposition is then applied to extract dynamic modulation patterns, allowing the model to incorporate multiphysics information and improving its generalization across complex operating environments.

2. Real Vehicle Data Processing

To address issues such as noise interference, outliers, and data incompleteness in real-world vehicle operation data, this chapter proposes a systematic data processing framework. The framework includes anomaly detection based on RRCF, a multi-dimensional abnormal data cleaning strategy, and a dynamic evaluation method for battery SOH labels based on ampere-hour integration and fuzzy Kalman filtering. This process aims to improve data quality and provide a reliable data foundation for battery SOH prediction models. The overall framework for SOH prediction is illustrated in Figure 1.

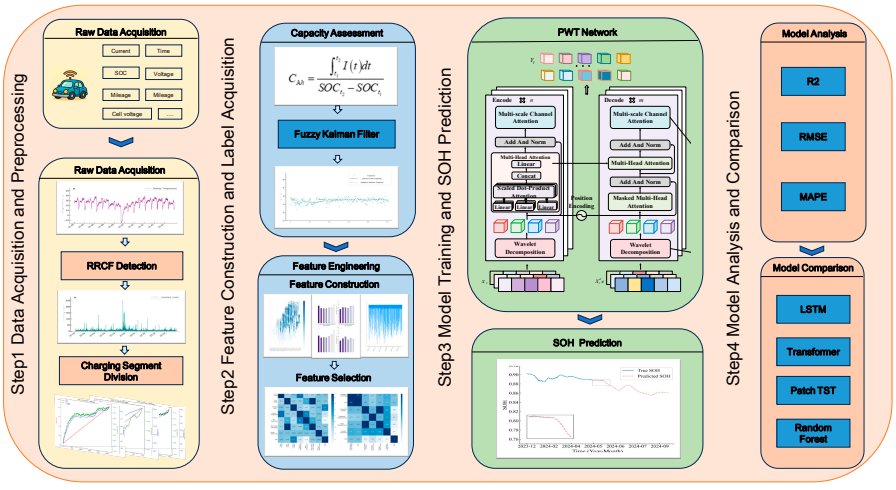


Figure 1. Structure of informer SOH prediction framework.

2.1. Anomaly Data Detection Based on RRCF

Battery data collected from real vehicles—including voltage, current, temperature, and vehicle mileage—are often affected by complex operating conditions, sensor drift,

and communication interference, resulting in nonlinear anomalous features. Traditional threshold-based methods and Isolation Forest algorithms have limited adaptability to dynamic time-series data. This study employs the RRCF algorithm for unsupervised anomaly detection. The core advantages of RRCF lie in

- High-dimensional time-series adaptability: By randomly projecting multi-dimensional time series (such as combined voltage–current–temperature sequences) into a lower-dimensional space, it avoids the curse of dimensionality;
- Dynamic robustness: An anomaly scoring mechanism based on local data density changes effectively identifies sudden anomalies caused by abrupt capacity drops or sensor failures during battery degradation;
- Incremental computational efficiency: The tree structure supports online updates, enabling the real-time processing of streaming vehicle data.

Both Isolation Forest and RRCF algorithms evaluate anomaly scores based on the path length needed to isolate a data point. The path length refers to the number of steps from the root node to the leaf node where the data point is isolated. Specifically, the shorter the path length, the easier it is to isolate the point, and, thus, the higher the likelihood of it being an anomaly. For a given data point, the isolation path length is computed as follows:

$$h(x) = \frac{H(x)}{c(n)} \tag{1}$$

where $H(x)$ is the path length required to isolate a data point x in a single tree. $c(n)$ is a constant used to normalize the path length, representing the maximum possible path length of the tree. For a tree containing n samples, $c(n)$ is approximately

$$c(n) = 2(ln(n-1) + \gamma) \tag{2}$$

where γ is a constant used to normalize the expected path length of the tree. This normalization constant ensures that the path lengths across all trees in the Isolation Forest are on a consistent scale.

$$\begin{cases} s(x) = \frac{2^{-E(h(x))}}{H(n)} \\ E(h(x)) = \frac{1}{T}\sum_{i=1}^{T} h_i(x) \end{cases} \tag{3}$$

Here, T is the total number of trees, and $h_i(x)$ is the path length of data point x in the i tree. $H(n)$ is the normalization constant, representing the maximum possible path length of the tree. $E(h(x))$ denotes the expected isolation path length of the data point across all trees, representing its average path length over multiple trees.

The specific implementation process is as follows: First, the raw data is segmented using a sliding window (with the window length set to a single charge–discharge cycle), extracting statistical features within each window (mean, variance, extreme values) as well as time series features (differenced sequences, rate of change). Next, the RRCF model is constructed, and anomaly scores are determined by calculating the collusive displacement distance between each sample point and other samples in the forest. A dynamic threshold (set at the fifth percentile) is used to mark abnormal segments. Using part of the data from a particular vehicle as columns, anomaly detection is performed via the RRCF algorithm, yielding the results shown in Figure 2. Specifically, Figure 2A shows the original data curve, while Figure 2B displays the anomaly scores after detection. The higher the anomaly score, the greater the likelihood that the data point is an outlier.

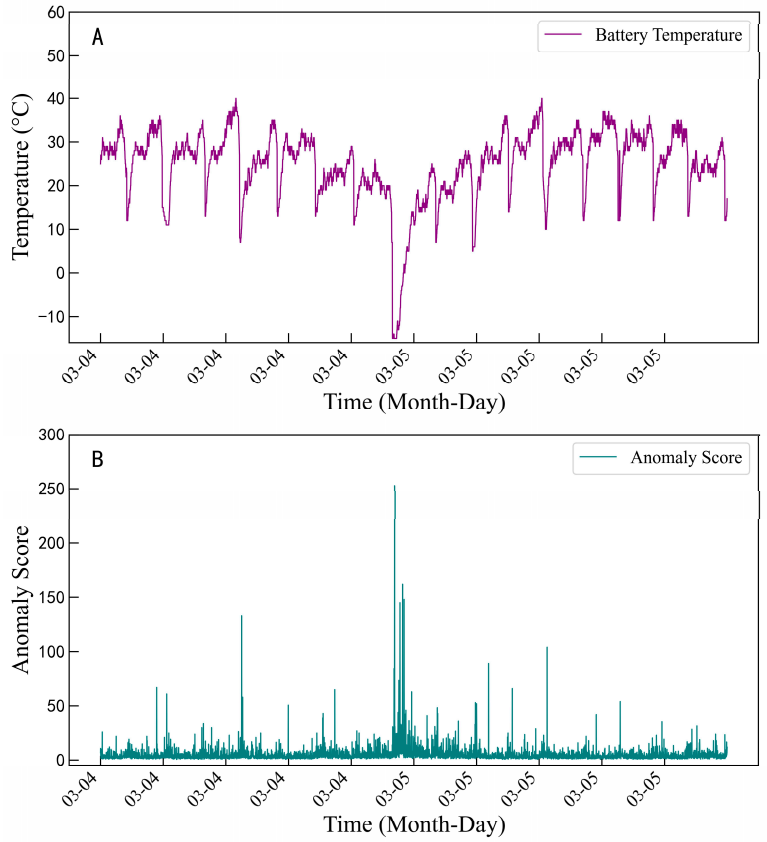


Figure 2. Example of RRCF anomaly detection results. (**A**) Temperature curve.; (**B**) Temperature anomaly detection score.

2.2. Multi-Dimensional Anomalous Data Cleaning

For the anomalous data detected by RRCF, a set of multi-dimensional cleaning rules is designed based on the battery system mechanism and vehicle operation logic:

1. Battery SOC abnormality correction: Sudden jumps or prolonged stagnation in the battery SOC values are usually caused by Kalman filter estimation errors or sensor failures. Since SOC values have a significant impact on battery capacity assessment, abnormal SOC values must be corrected before capacity estimation. By combining the vehicle's charging status and speed data, the reasonableness of SOC values can be evaluated. As shown in Figure 3a, when the battery is in a non-charging state (charging status code 03) and the vehicle speed accelerates from zero, if the SOC value suddenly drops sharply to zero, it is considered abnormal and requires correction. Abnormal SOC values can be corrected using a three-point constraint method: using the SOC values at the start and end of the charging process as boundaries and applying smoothing interpolation combined with the ampere-hour integration results.
2. Charging status logical anomaly: The battery charging status directly affects SOC values, so abnormal charging states need to be corrected before further analysis. As shown in Figure 3b, when the vehicle is in the start state (status code 01) and the vehicle speed is greater than or equal to zero, if the battery charging state shows 01 (indicating parking charging), the data is considered abnormal. At this point, corrections should be made based on vehicle speed, vehicle status, and the charging state information before and after.
3. Cumulative mileage anomaly handling: Mileage data during actual driving may exhibit non-monotonic increments due to CAN bus packet loss or odometer resets. As shown in Figure 3c, segmented linear interpolation is applied to correct abnormal

jump points, and the reasonableness of the interpolation is verified by integrating vehicle speed over time. If the deviation between the interpolation result and the integral calculation exceeds 5%, the data segment is considered invalid and should be discarded.

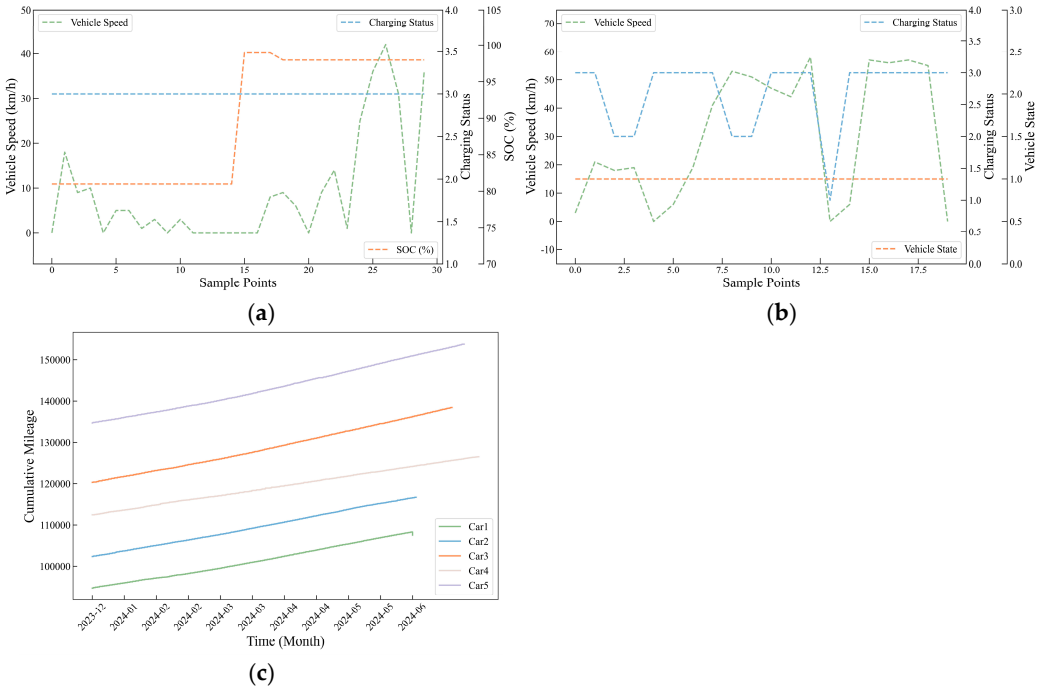


Figure 3. Examples of anomalous data: (**a**) SOC anomaly, (**b**) charging logic anomaly, (**c**) mileage anomaly.

2.3. True Capacity Label Evaluation Based on Real Vehicle Data

This study uses real vehicle data. Due to the lack of laboratory-grade capacity testing conditions in real vehicles, multiple methods were preliminarily compared. The results indicate that the ampere-hour integration method demonstrates more stable performance in battery capacity assessment. Therefore, this paper proposes a dynamic capacity estimation method that combines the ampere-hour integration method with fuzzy Kalman filtering:

Ampere-hour Integration Estimation: Within a complete charging cycle (SOC 20–100%), the capacity is calculated by integrating the charging current:

$$C_{Ah} = \frac{\int_{t_1}^{t_2} I(t)dt}{SOC_{t_2} - SOC_{t_1}} \tag{4}$$

where $I(t)$ is the current as a function of time, t is time, SOC_{t_1} and SOC_{t_2} represent the state of charge at the start and end of the charging cycle, respectively, and C_{Ah} is the battery capacity. Figure 4 shows scatter plots of battery capacity calculated using the ampere-hour integration method for six selected vehicles. The results indicate considerable dispersion in capacity estimates, thus necessitating further correction using fuzzy Kalman filtering.

Fuzzy Kalman Filter Correction: The ampere-hour integration method is susceptible to errors from current sensor inaccuracies and SOC estimation errors, requiring further evaluation of battery capacity. Several methods were compared, as shown in Figure 5, including LSTM, Kalman Filter, and Fuzzy Kalman Filter. Since the Kalman Filter results better capture the battery degradation trend, we introduce a fuzzy logic–enhanced Kalman filtering framework, in which the process noise covariance Q and the measurement noise covariance RRR are dynamically adjusted according to the system's real-time estimation performance. Instead of using fixed values for Q and R, which often fail to reflect varying

sensor quality or noise levels in real-world conditions, our method treats these parameters as outputs of a fuzzy inference system. This system uses the estimation residual e (i.e., the difference between the measured capacity and the Kalman-predicted capacity) and its first-order difference Δe as inputs, thereby enabling the filter to respond adaptively to different operating conditions.

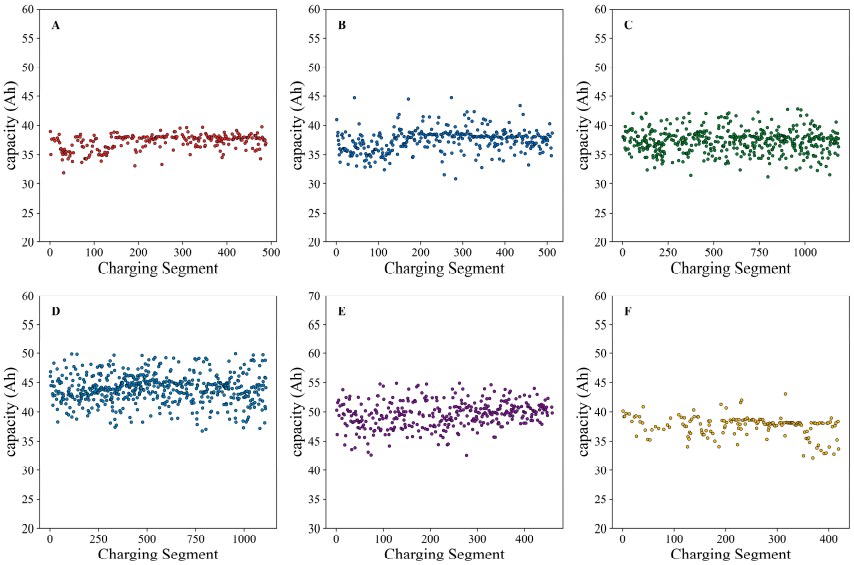


Figure 4. Capacity scatter plot. (**A**–**F**) are the plots showing the relationship between capacity and cycle number for the six selected vehicles (2, 13, 27, 34, 46, 59) in the experiment.

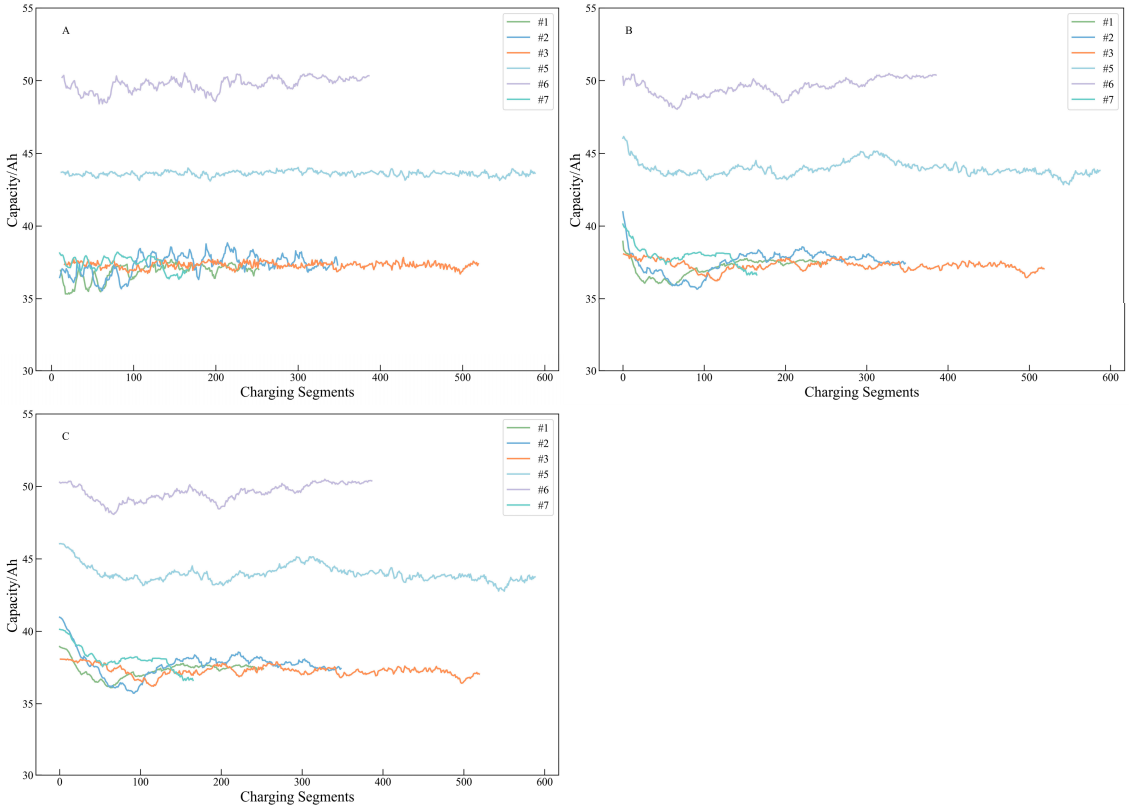


Figure 5. Battery capacity correction result comparison chart. (**A**) LSTM fitting results; (**B**) Traditional Kalman filter results; (**C**) Fuzzy Kalman filter results.

The selection of fuzzy rules and membership functions is grounded in the expert knowledge of Kalman filter behavior and battery estimation dynamics. Specifically, trian-

gular membership functions are used to categorize both e and Δe into linguistic variables such as "poor", "average", and "good", reflecting the magnitude and stability of estimation errors. When large residuals and sharp fluctuations (i.e., poor estimation quality) are detected, the fuzzy logic controller increases the process noise covariance Q to allow the filter to respond more flexibly while decreasing R to reduce the reliance on potentially unreliable measurements. Conversely, when the estimation residual is small and stable, indicating high confidence in both the model and measurements,Q is reduced, and R is increased accordingly to improve filtering precision. The fuzzy inference system employs a Mamdani-type structure and uses a centroid defuzzification method to produce crisp outputs for Q and R at each time step. This continuous adaptation enables the Kalman filter to maintain the accurate tracking of battery capacity despite the presence of noise, sensor drift, or model mismatch. Experimental results demonstrate that the fuzzy Kalman filter achieves superior performance compared to both traditional Kalman filtering and data-driven approaches such as LSTM.

This study achieves dynamic optimization of the observation noise covariance matrix in the SOH estimation process by integrating a fuzzy inference system with the Kalman filtering framework. To validate the effectiveness of the proposed method, multi-source data fusion was employed.

Validation Strategy: Laboratory-measured data from retired vehicle batteries were used to obtain baseline capacity values through standard charge–discharge tests. In addition, lithium-ion battery aging data from the publicly available dataset of the University of Maryland were integrated [23]. As shown in Table 1, compared with the traditional Kalman filter method, the proposed approach reduces the root RMSE by 9.6% across both datasets. This indicates that the fuzzy Kalman filter method, through its dynamic noise adjustment mechanism, significantly improves the robustness of capacity label estimation under complex operating conditions, thereby providing high-confidence supervisory signals for SOH prediction models.

Table 1. Comparison of results from two methods.

Data Type	Kalman Filter Estimation (Ah)	Fuzzy Kalman Filter (Ah)	Experimental Value (Ah)
Public Dataset	1.43	1.04	1.19
Real Vehicle Disassembly Data	43.67	42.30	41.92

3. Feature Engineering

This chapter focuses on the feature engineering framework for battery SOH prediction, establishing a multi-dimensional parameter correlation-based feature selection approach. The charging behavior and voltage consistency of the battery are evaluated, along with an in-depth analysis of the interaction between battery temperature and current, in order to extract features relevant to SOH prediction. To reduce computational complexity, the Pearson correlation coefficient method is employed to select features, ensuring both efficiency and conciseness in feature extraction.

3.1. Feature Construction

3.1.1. Evaluation of Charging Behavior

The charging conditions of real-world vehicles are influenced by driver behavior and differ from standard cycle tests conducted in laboratories. As a result, actual charging behavior has a significant impact on the health status of the battery. By analyzing the performance of power batteries under real charging conditions and properly evaluating their charging behavior, valuable insights can be gained for predicting the SOH of the battery.

To comprehensively characterize charging behavior and uncover patterns, it is necessary to define parameters that describe the actual charging conditions of the battery. During charging, the vehicle remains stationary, and parameters such as the charging current and the starting and ending SOC are closely related to the battery's charging behavior. The specific parameters selected in this study are listed in Table 2.

Table 2. Selection of parameters for charging behavior evaluation.

Parameters	Special Parameter
Charging Current	Step Current: Current magnitude
	Step Current: Duration
	Step Current: Starting SOC
Battery SOC	Starting SOC
	Ending SOC
	Depth of Charge

Due to the highly complex internal mechanisms of batteries, it is difficult to directly establish a model to evaluate charging behavior. Therefore, membership functions are employed to describe and assess the fuzzy relationship between charging parameters and charging behavior. The introduction of membership functions enables the scoring system to effectively handle uncertainty and fuzziness, making the evaluation of charging behavior more consistent with real-world conditions. Moreover, membership functions exhibit strong robustness in the presence of a small number of abnormal behaviors or deviations, helping to reduce the impact of noise and errors and thereby generating more stable scoring results. In addition, the relationship between battery charging behavior and battery health status has been experimentally validated. Thus, using membership functions to score charging behavior and incorporating the results as features into SOH prediction models is a reasonable approach [24]. To quantify and evaluate the battery charging behavior in a structured and health-oriented manner, a novel assessment algorithm is proposed. This method combines multi-dimensional discretization of the charging state space with fuzzy membership weighting, enabling a more interpretable and health-aware scoring of charging patterns. The entire process involves five key steps, including state discretization, frequency-based mapping, fuzzy function design, matrix fusion, and score normalization.

To evaluate charging behavior, a discretized state-space framework is constructed using dual-dimensional partitioning over SOC and current. The continuous state space is mapped to an $n \times m$ grid:

$$SOC_i = \frac{i-1}{n-1} \times 100\%,\ I_j = \frac{j-1}{m-1} \times I_{\text{rated}} \tag{5}$$

The cumulative time spent in each grid cell forms the charging matrix A_{I-SOC}, and its normalized form approximates the charging trajectory's probability distribution:

$$P(s_i, i_j) = \frac{T_{s_i i_j}}{\sum_{i=1}^{n} \sum_{j=1}^{m} T_{s_i i_j}} \tag{6}$$

Fuzzy membership functions are defined to reflect battery health preferences. The SOC membership is modeled by a Gaussian function:

$$A_{SOC}(x) = \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),\ \mu = 50,\ \sigma = 20 \tag{7}$$

The current membership function uses a piecewise linear form:

$$A_I(x) = \begin{cases} 1, & x \leq 100 \\ \frac{180-x}{80}, & 100 < x < 180 \\ 0, & x \geq 180 \end{cases} \tag{8}$$

A two-dimensional weight matrix A_{SOC-I} is generated by the outer product of the two membership functions. The final behavior score is computed as follows:

$$\text{Score} = \sum_{i=1}^{n} \sum_{j=1}^{m} A_{SOC-I}(i,j) \cdot A_{I-SOC}(i,j) \cdot W_{\text{norm}} \tag{9}$$

where W_{norm} is a normalization factor for temporal comparability. Monthly scoring is performed by averaging valid membership products over time, enabling the robust, health-aware assessment of charging patterns.

The detailed evaluation process is illustrated in Figure 6, and the evaluation results, using four vehicles as examples, are shown in Figure 7.

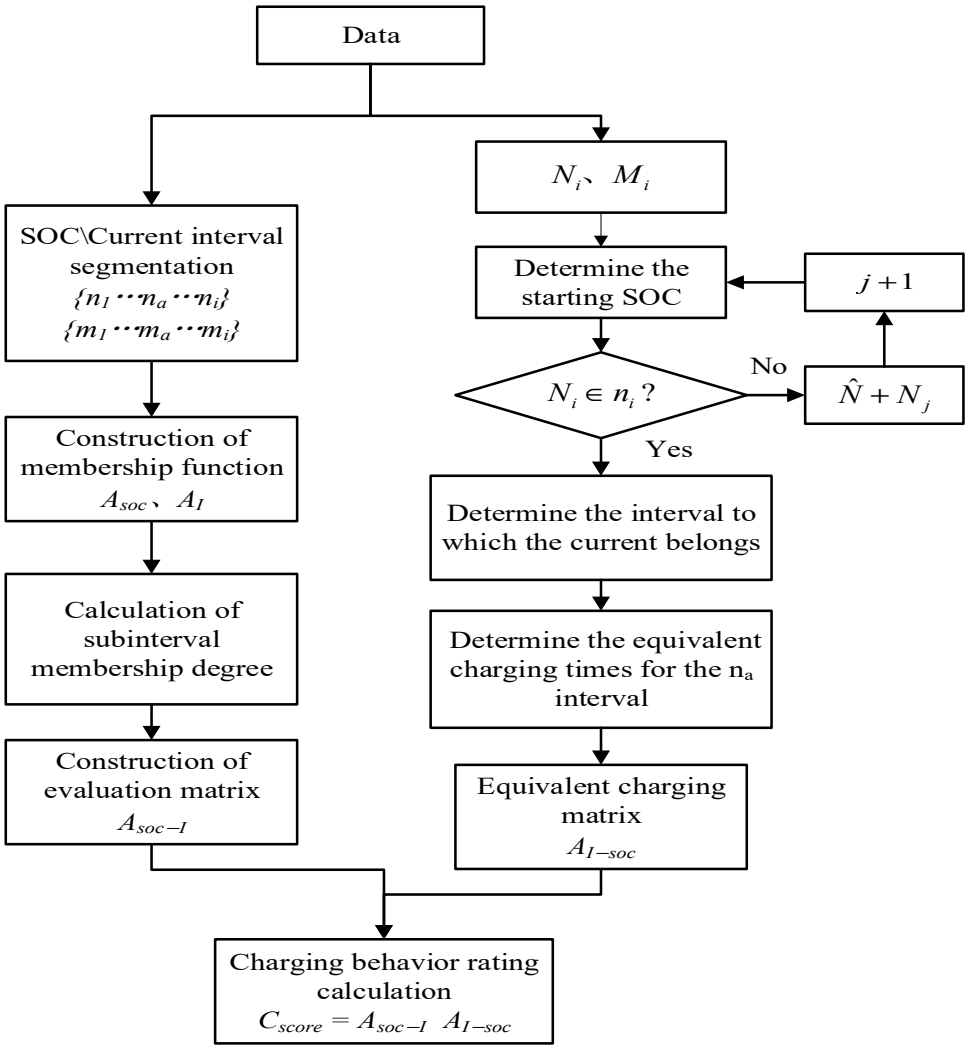


Figure 6. Charging behavior evaluation process.

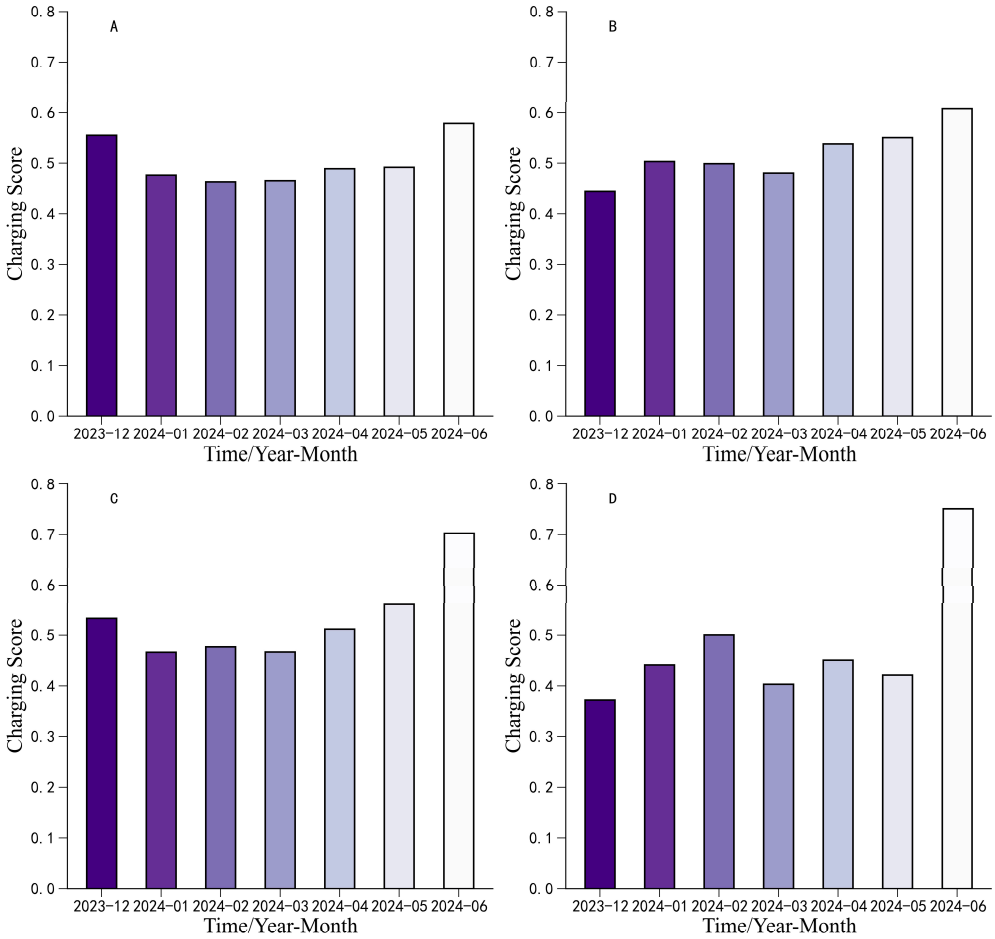


Figure 7. Example of evaluation results. (**A**) Charging behavior score of Vehicle 2; (**B**) Charging behavior score of Vehicle 13; (**C**) Charging behavior score of Vehicle 27; (**D**) Charging behavior score of Vehicle 34.

3.1.2. Battery Voltage Consistency

Voltage consistency refers to the similarity and uniformity of the voltages of individual battery cells within a battery pack of a pure electric vehicle during charging and discharging processes. A battery pack is typically composed of multiple individual cells, each cell's voltage variation reflecting its charge state and health condition. High voltage consistency indicates small voltage differences among the cells, suggesting that the cells operate in similar states and that the overall performance of the battery pack is stable. Conversely, low voltage consistency may imply significant voltage fluctuations in certain cells, which could lead to the overcharging or over-discharging of individual cells, thereby accelerating battery aging or causing premature failure.

The voltage consistency score effectively reflects the degree of voltage uniformity among the individual cells in the battery pack. This score primarily measures the magnitude of voltage differences between cells, thus representing the overall balance of the battery pack. The calculation formula for the voltage consistency score is as follows:

$$V_{score} = 1 - \frac{\sigma(V)}{\overline{V}} \tag{10}$$

In the formula, $\sigma(V)$ represents the standard deviation of the individual cell voltages, reflecting the degree of voltage fluctuation among the cells. It is the mean voltage of the individual cells, indicating the overall voltage level of the battery pack. The results of the voltage consistency score are shown in Figure 8. As observed from Figure 7, over the

course of one year, the voltage consistency score shows a general declining trend. This indicates that, as the battery usage time increases, the voltage differences among individual cells gradually widen, resulting in reduced voltage consistency. The decline in voltage consistency reflects the gradual deterioration of the battery's health status and indicates the degradation of the battery pack's performance.

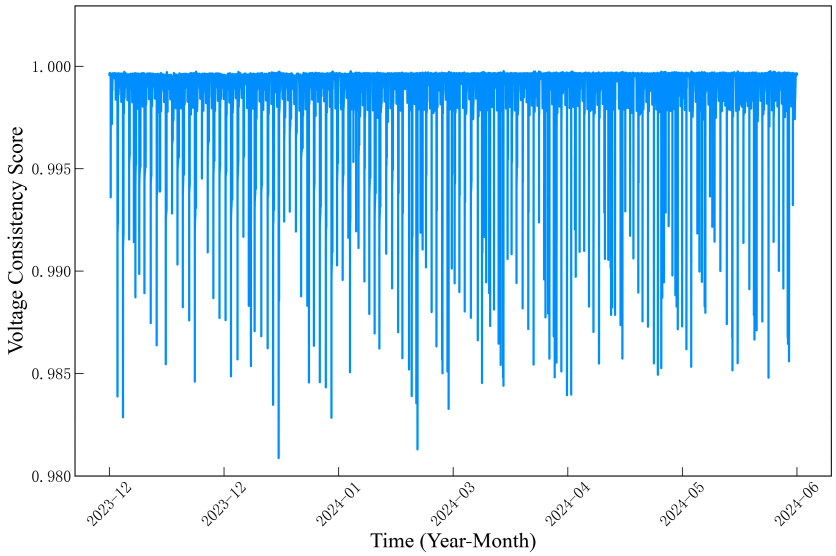


Figure 8. Example of voltage consistency scoring.

3.1.3. Interaction Between Current and Temperature

Battery capacity degradation is influenced not only by the individual effects of the current and temperature but also by their interaction, which accelerates capacity fade and internal resistance increase, significantly shortening battery life [25,26]. To analyze the variation trends of battery capacity under different temperature and current conditions, a pivot table aggregation method was used. By grouping data according to the maximum temperature and total current and calculating the mean capacity value for each combination, Figure 9 reveals the adaptability performance under varying temperature and current conditions. It can be seen from the figure that the charging current magnitude is affected by temperature, decreasing as the temperature drops.

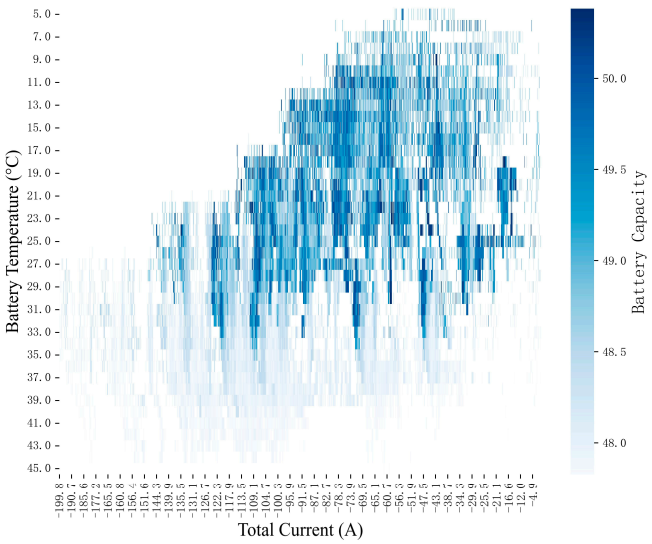


Figure 9. Current–temperature interaction diagram.

To incorporate the interaction between current and temperature into the battery SOH prediction model, polynomial interaction features combining current and temperature were constructed to capture nonlinear relationships. Considering that the relationship between current, temperature, and battery SOH is not simply linear, a gradient descent-based optimization method was used to determine the polynomial parameters. The specific formula is as follows [27]:

Interaction feature $F(I, T)$ represents the nonlinear combination of current I and temperature T and is expressed as follows:

$$F(I, T) = \sum_{m=0}^{M} \sum_{n=0}^{M} a_{m,n} I^m T^n \tag{11}$$

Here, M and N represent the highest polynomial degrees of current and temperature, respectively, and $a_{m,n}$ are the polynomial coefficients to be optimized. The loss function ℓ is expressed as follows:

$$\ell = \frac{1}{2} \sum_{k=1}^{K} [y_k - F(I_k, T_k)]^2 \tag{12}$$

The equation variables are as follows: y_k represents the battery capacity of the k sample, $F(I_k, T_k)$ is the interaction feature value of current and temperature for the kk-th sample, and K is the total number of samples.

The coefficients $a_{m,n}$ are optimized using gradient descent, with the gradient calculated as follows:

$$\frac{\partial \ell}{\partial a_{m,n}} = -\sum_{k=1}^{K} [y_k - F(I_k, T_k)] {I_k}^m {T_k}^m \tag{13}$$

The parameters are iteratively updated to minimize the loss function, with the update formula given by the following:

$$a_{m,n}^{(t+1)} = a_{m,n}^{(t)} - \eta \frac{\partial \ell}{\partial a_{m,n}} \tag{14}$$

In the formula, t represents the iteration number, and η is the learning rate, which determines the step size for parameter updates.

To avoid gradient vanishing, regularization is introduced, resulting in the improved loss function:

$$\ell_{reg} = \ell + \lambda \sum_{m=0}^{M} \sum_{n=0}^{M} a_{m,n}^2 \tag{15}$$

The parameter update formula is as follows:

$$a_{m,n}^{(t+1)} = a_{m,n}^{(t)} - \eta \left(\frac{\partial \ell}{\partial a_{m,n}} + \lambda a_{m,n} \right) \tag{16}$$

3.2. Feature Selection

In high-dimensional feature data, there may exist redundant, duplicate, or irrelevant features. These not only increase the computational complexity of the model but also may introduce noise, reduce the model's predictive performance and stability, and even lead to overfitting or the curse of dimensionality. Therefore, performing feature selection before feature extraction is a crucial step. Its purpose is to filter out features closely related to the prediction task from the initially constructed feature set while removing ineffective features, thereby providing a more concise and high-quality input foundation for subsequent feature extraction and model training.

This study employs the Pearson correlation coefficient method to perform correlation analysis on the initially constructed features, quantitatively evaluating the degree of linear correlation between each pair of features. The Pearson correlation coefficient is a classic statistical measure used to assess the correlation between two variables, with values ranging from -1 to 1. A coefficient close to 1 indicates a strong positive correlation, meaning that the two features exhibit similar variation trends; a coefficient close to -1 indicates a strong negative correlation, meaning that the two features show opposite variation trends; and a coefficient close to 0 indicates almost no linear correlation between the two features.

During feature selection, if certain feature pairs exhibit high correlation (e.g., the absolute value of the correlation coefficient exceeds a set threshold), they are considered redundant. Typically, one important feature is retained while the others are removed to reduce input dimensionality. At the same time, features with very low correlation to the battery's SOH are considered ineffective and are eliminated. Through this process, the retained feature set becomes more compact and efficient, helping to improve the model's learning ability and generalization performance.

To better illustrate the correlation between features and the target variable, a correlation heatmap between the battery SOH and the remaining features was plotted, as shown in Figure 10. The heatmap uses a color gradient to represent the degree of correlation between each pair of features, where darker colors indicate stronger correlations and lighter colors indicate weaker correlations. By observing the heatmap, highly correlated feature pairs as well as features unrelated to the target variable can be quickly identified, providing a basis for feature selection. This method not only significantly reduces feature dimensionality but also effectively alleviates redundancy and noise interference among features before model training. The selected features after screening are listed in Table 3.

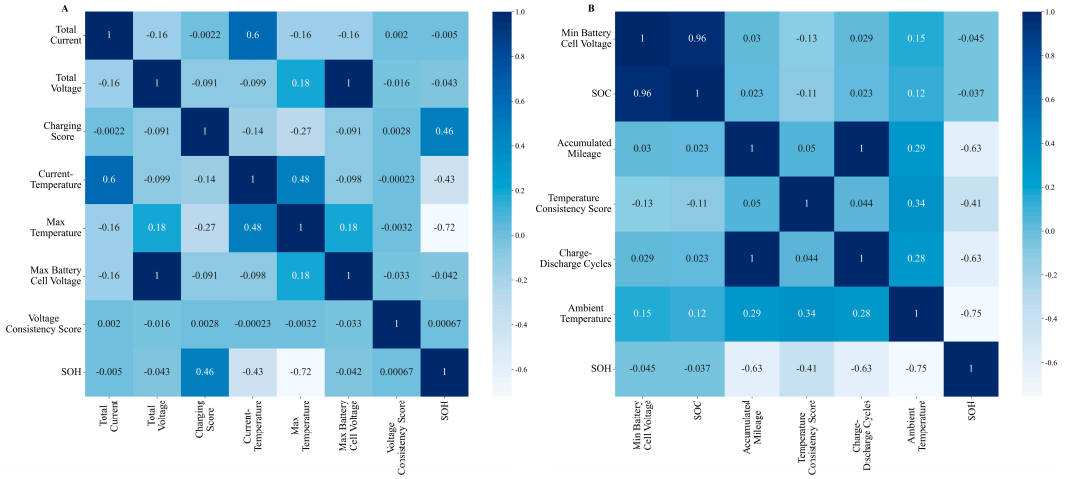


Figure 10. Feature importance heatmap. (**A**) Includes correlation analysis between SOH and total current, total voltage, charging behavior score, maximum temperature, maximum cell voltage, and voltage consistency score.; (**B**) Includes correlation analysis between SOH and minimum cell voltage, SOC, accumulated mileage, temperature consistency score, number of charge-discharge cycles, and ambient temperature..

Table 3. Feature selection results.

Feature	Description	Correlation Coefficient
Ambient Temperature	Daily average ambient temperature	−0.75
Temperature Consistency Score	Battery cell temperature consistency score	−0.41
Charging Behavior Score	Battery charging behavior score	0.46
Cumulative Mileage	Vehicle cumulative mileage	−0.63

Table 3. *Cont.*

Feature	Description	Correlation Coefficient
Charge–Discharge Cycle Count	Number of charge–discharge cycles	−0.63
Current–Temperature Interaction	Interaction feature of current and temperature	−0.43
Maximum Temperature	Maximum temperature of battery cells	−0.72

4. Patch Wavelet Transformer

To simultaneously capture both temporal and frequency-domain information in time series forecasting and preserve the strong local correlations inherent in time sequences, we designed a hybrid Transformer framework driven by two core concepts: frequency-domain enhancement and local attention. This framework retains the original Encoder–Decoder structure of the Transformer, with similar architectures for both.

In the Encoder, the original sequence first undergoes wavelet decomposition. The reconstructed signals containing rich information across various frequency bands are then fused with the original time series along the channel dimension, forming a multi-scale feature representation that captures both trends and fine-grained details.

Next, the fused long sequence is divided into multiple short segments (patches), and self-attention is applied at the patch level. This helps preserve the local dependencies within the sequence while significantly reducing the computational complexity of attention operations.

Finally, a carefully designed multi-scale spatial attention module is introduced to integrate features from different high-dimensional representations, thereby enhancing the model's prediction accuracy. This design fully leverages the frequency-domain insights brought by wavelet decomposition and ensures model scalability and efficiency through patch-level attention (Figure 11).

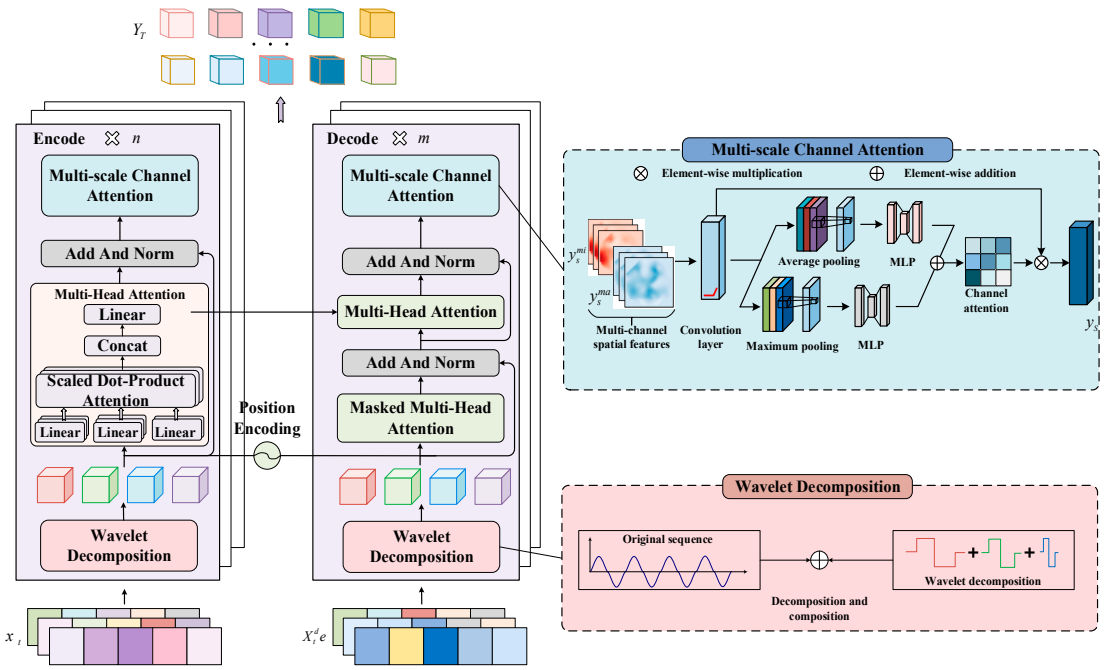


Figure 11. PWT model structure diagram.

4.1. Wavelet Decomposition

The real-world operating data of battery electric vehicles exhibits significant non-stationarity and multi-scale characteristics. Traditional Transformer models often struggle

to simultaneously capture both local details and long-term trends when modeling such complex time-series signals. In recent years, to address complex trends and local variations in time-series data, some studies have introduced decomposition mechanisms into the Transformer architecture to improve its long-term modeling capabilities [28].

Inspired by this, we designed a wavelet decomposition-based module that decomposes the original time-series data into high-frequency and low-frequency sub-signals. This allows for the extraction of both instantaneous variations and overall trend features, enhancing the representation of temporal and frequency-domain information and enabling the fine-grained modeling of potential temporal dependencies [29].

To strengthen the representation of temporal and frequency-domain information in the sequence data and to better capture subtle temporal dependencies, we propose a Wavelet Decomposition and Feature Concatenation Module. This module introduces multi-scale feature representations through decomposition and reconstruction of the original time series, thereby improving the model's ability to capture both short-term details and long-term trends.

Specifically, let the original input time series be the following:

$$X \in R^{\wedge}(B \times T \times F) \tag{17}$$

where B denotes the batch size, T represents the time step length, and F indicates the number of features.

First, the wavelet decomposition method is used to decompose the original sequence X into a high-frequency component X_h and a low-frequency component X_l, which, respectively, capture instantaneous variations and overall trends. The decomposition process is as follows:

$$(X_l, X_h) = WaveletDecompose(X) \tag{18}$$

where X_l and X_h both have the shape $R^{B \times T \times F}$.

To fully integrate the original information with the sub-signals at different scales, X, X_l, and X_h are concatenated along the feature dimension (i.e., the last dimension), resulting in the enhanced feature representation X_{cat}:

$$X_{cat} = Concat(X, X_l, X_h) \tag{19}$$

The dimension of the concatenated tensor is as follows:

$$X_{cat} \in R^{\wedge}(B \times T \times 2F) \tag{20}$$

The overall information flow process is as follows:

$$X \rightarrow WaveletDecompose \rightarrow (X_l, X_h) \rightarrow Concat \rightarrow X_{cat} \tag{21}$$

The Wavelet Decomposition and Feature Concatenation Module effectively integrates the original time series features with multi-scale fine-grained variation features, providing richer input information for subsequent modeling layers (such as the Transformer module). This contributes to enhancing the model's performance in handling complex, non-stationary, and multi-scale time series data.

4.2. Patch-Based Time Series Segmentation

To effectively handle non-stationary and multi-scale time series data, this paper proposes a Patch-based time series segmentation method. This approach decomposes the time series into multiple small segments using a sliding window, enabling the Transformer

model to more accurately capture both local details and long-term trends. By employing the sliding window technique, the data is divided into multiple patches. Each patch has a length of L and a stride of S. The p patch can be represented as follows:

$$P_p = X_{cat}[(p-1)S + 1 : (p-1)S + L] \tag{22}$$

where P_p has the size (L, 3d), meaning that each patch contains L time steps, and the feature dimension is 3d. The specific time series patching process includes the following steps:

- Input data: Assume the input data is X_{cat} with dimensions (T, 3d).
- Window partition: Use a sliding window method with fixed length L and stride S to segment X_{cat} into multiple patches.
- Patch representation: Each patch p consists of L consecutive time steps of data, represented by Equation (22).
- Patch input: All generated time series patches are fed into subsequent models for processing.

4.3. Temporal Attention Mechanism

The core advantage of the Transformer architecture in battery life prediction lies in its efficient modeling of long-range temporal dependencies. Through the self-attention mechanism, the model can flexibly capture cross-cycle correlation features within time series data, effectively integrating key degradation information during the charge–discharge process. Moreover, compared to recurrent structures, the Transformer is more efficient in parallel computation and handling long sequences, making it well-suited for high-frequency, long-duration real vehicle data.

In this model, both the encoder and decoder are constructed based on the Transformer architecture. Taking the standard "Scaled Dot-Product Attention" as an example, its calculation formula is as follows [30]:

$$Attention(Q, K, V) = \mathrm{so}ftmax(\frac{QK^T}{\sqrt{d_k}})V \tag{23}$$

where Q, K, and V represent the Query, Key, and Value matrices, respectively, all obtained by applying linear transformations to the input features. Here, d_{k} denotes the dimension of the key vectors.

To enhance the model's expressive capability, the Multi-Head Attention mechanism is introduced, which computes multiple attention subspaces in parallel and then concatenates and fuses the results:

$$MultiHead(Q, K, V) = Concat(head_1, \ldots, head_h)W^O \tag{24}$$

Each attention head is computed as follows:

$$head_i = Attention\left(QW_i^Q, KW_i^K, VW_i^V\right) \tag{25}$$

In this way, the model can learn diverse features in the sequence from different representation subspaces.

4.4. Multi-Scale Channel Attention Mechanism

Considering that the importance of components at different scales after wavelet transform varies, this paper further introduces a multi-scale channel attention mechanism to adaptively weight the feature responses at different scales. This mechanism draws on the

design concept of the SE (Squeeze-and-Excitation) module and mainly includes global pooling, nonlinear activation, and channel recalibration processes:

$$\alpha_c = \sigma(W_2 \cdot ReLU(W_1 \cdot Pool(F_c))) \tag{26}$$

Here, F_c represents the feature of the c channel, and α_c is the weight of that channel. This mechanism effectively enhances the model's sensitivity to key frequency band features, further improving its feature extraction capability.

5. Model Validation

5.1. Validation Across Different Vehicles

This study uses real-world data from 86 vehicles. The real-world dataset used for model validation consists of operational data collected from 86 electric vehicles over approximately six months, spanning from December 2023 to June 2024. All vehicles are equipped with lithium iron phosphate ($LiFePO_4$) battery packs, widely adopted in commercial electric vehicles for their safety and longevity. To ensure geographic and environmental diversity, the vehicles are distributed roughly evenly across six major cities in China Mainland: Kunming (14 vehicles), Leshan (10 vehicles), Guangzhou (22 vehicles), Changchun (6 vehicles), Tianjin (14 vehicles), and Hangzhou (20 vehicles). This distribution covers a broad spectrum of climatic conditions—from the subtropical climate in Guangzhou to the colder temperate environment in Changchun—and different traffic and usage patterns. In this study, the tested battery system is a lithium iron phosphate ($LiFePO_4$) battery pack composed of prismatic cells, each with a nominal voltage of 3.2 V. The entire battery pack has a rated mass of approximately 200 kg and is structurally divided into nine regions. To monitor thermal behavior, each region is equipped with one temperature probe, enabling distributed temperature measurement across the pack during real-world vehicle operation.

Meanwhile, data from six vehicles operating in different environments (as shown in Figure 12) were selected as the validation dataset to evaluate the model's performance in predicting the battery SOH. The validation results of the model are presented in Table 4. From the data in Table 4, it can be seen that there are some differences in the model's performance across different vehicles. However, based on the RMSE, MAPE, and R^2 metrics, the model used in this study demonstrates lower prediction errors on these six vehicles compared to other models, indicating that the model can effectively adapt to battery characteristics under different vehicles and operating conditions. In the experimental process, the parameters of the proposed PWT and their settings are as follows: the number of attention heads is 4, the patch window size is 3, and the batch size is 64. The Adam algorithm is adopted for optimization, with an initial learning rate set to 0.001 and decay based on 0.0001. All these parameters are tuned via grid search. All experiments, including model training and evaluation, were conducted on a workstation equipped with an Intel® Xeon® Gold 6130 CPU @ 2.10 GHz, 96 GB of RAM, and an NVIDIA Tesla T4 GPU with 16 GB of memory. This computational environment ensured efficient model development and the fair comparison of runtime performance across different methods.

Table 4. Comparison of evaluation metrics on validation data.

Evaluation Metric	Vehicle 2	Vehicle 13	Vehicle 27	Vehicle 34	Vehicle 46	Vehicle 59
RMSE	0.0226	0.0152	0.0242	0.0150	0.0161	0.0134
R^2	0.914	0.943	0.903	0.991	0.993	0.998
MAPE	0.0502	0.0111	0.0725	0.0032	0.0025	0.0017

To further analyze the effect of proposed method, this study conducted a visual comparison between the prediction results of the Transformer model and the Wavelet Transformer model. Figure 13 shows the prediction results of the Transformer model, while Figure 14 presents the results of the PWT model.

Compared with the Transformer model, the PWT model demonstrates higher prediction accuracy and stronger adaptability. In particular, in regions with abrupt changes, the PWT model is able to rapidly adjust its predictions to closely follow the variations in the actual values. At the same time, it also shows more accurate long-term trend capture. Under different time periods and data distributions, the model exhibits greater stability, fully reflecting its significant advantages in the SOH prediction task. It can accurately track short-term fluctuations while maintaining high overall prediction accuracy.

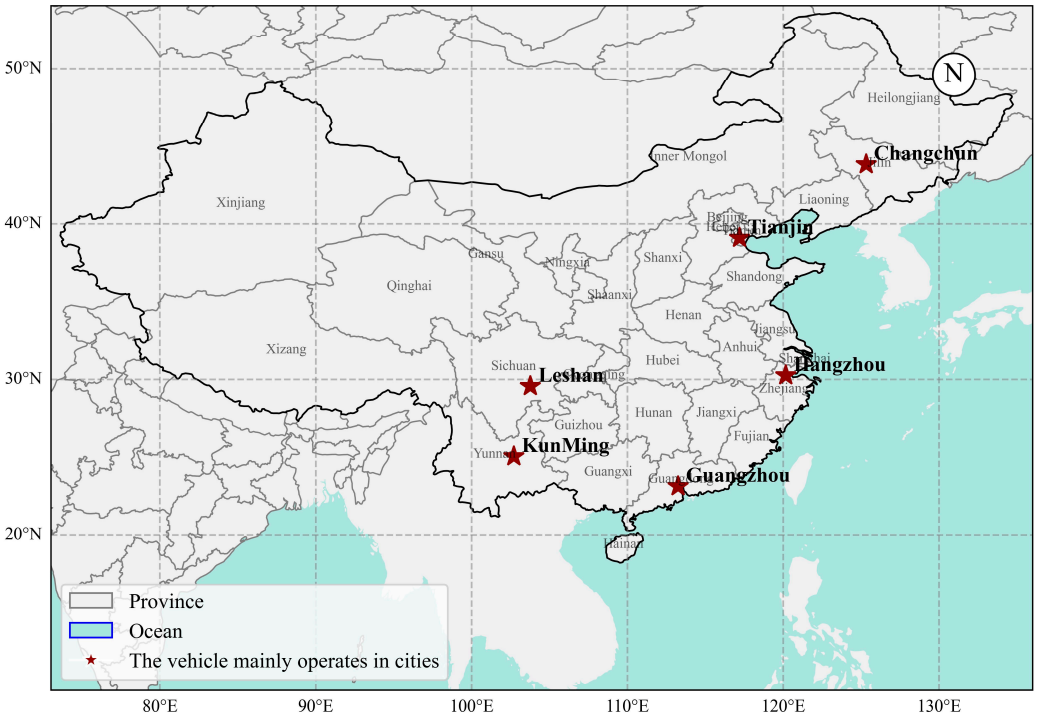


Figure 12. Main operating city in China Mainland of the vehicle.

5.2. Comparative Validation of Models

To validate the predictive performance of the proposed PWT model, several commonly used models—including LSTM, Random Forest, Transformer, and PatchTST—were selected as baselines for comparison. Table 5 presents the RMSE, MAPE, R^2, and computation time of each model on the test set, offering a comprehensive evaluation of their prediction accuracy and computational efficiency.

Table 5. Evaluation metrics of different prediction models.

Model	RMSE	MAPE (%)	Time (s)
LSTM	1.357	2.458	1.324
Random Forest	1.543	3.344	2.342
Transformer	0.112	1.202	0.653
PatchTST	0.0432	0.088	0.453
PWT	0.0177	0.0235	0.523

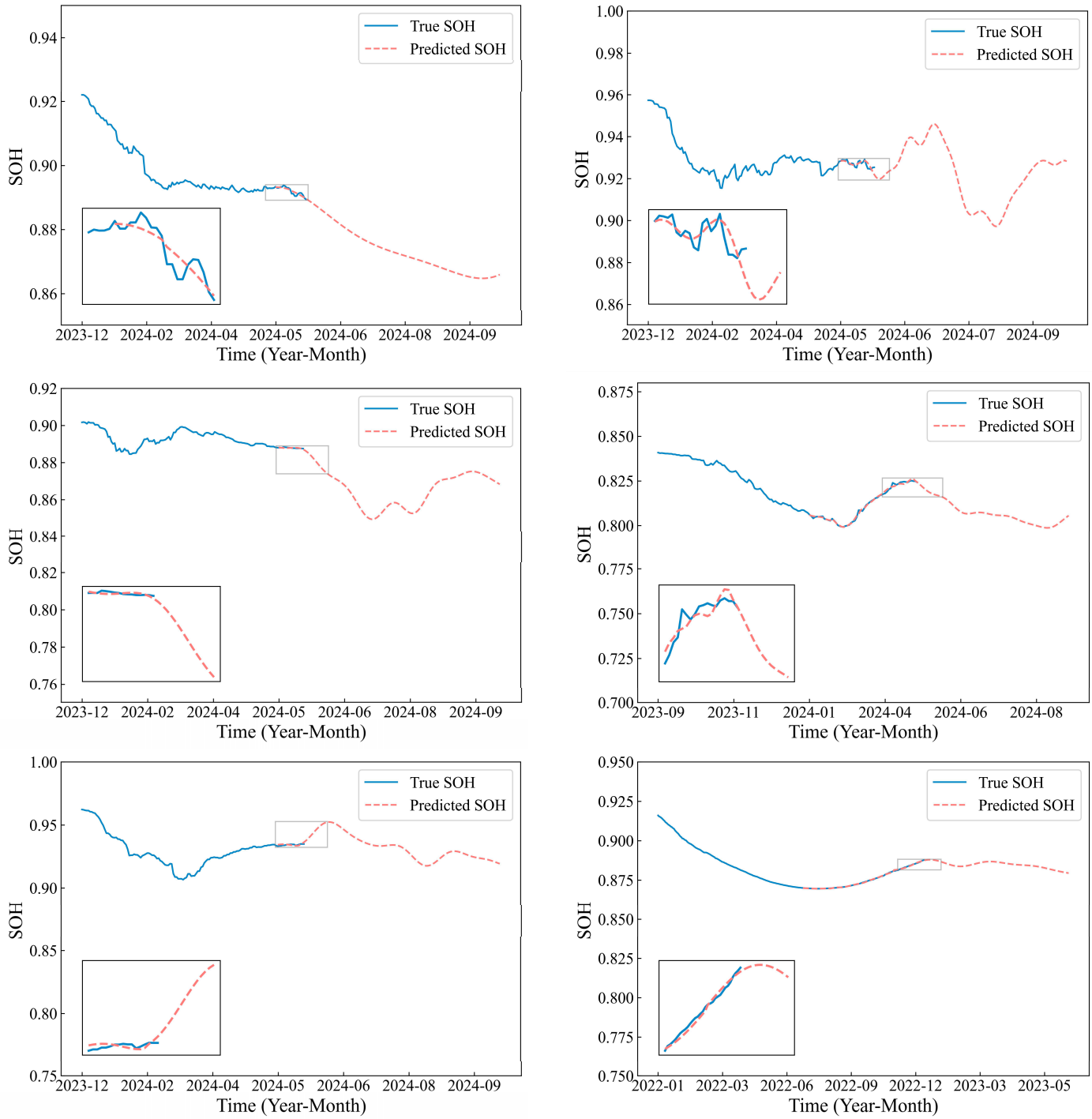


Figure 13. The prediction results of the Transformer model.

The experimental results clearly demonstrate that PWT outperforms all other models across all metrics. Compared with the Transformer, the RMSE and MAPE of PWT are reduced by approximately 83% and 96%, respectively, and the computation time is reduced by about 20%. Compared with the strong-performing PatchTST, PWT further reduces the error metrics by over 50% while maintaining a similar computation time. Overall, PWT achieves high prediction accuracy with notable computational efficiency, demonstrating strong comprehensive performance and practical application potential.

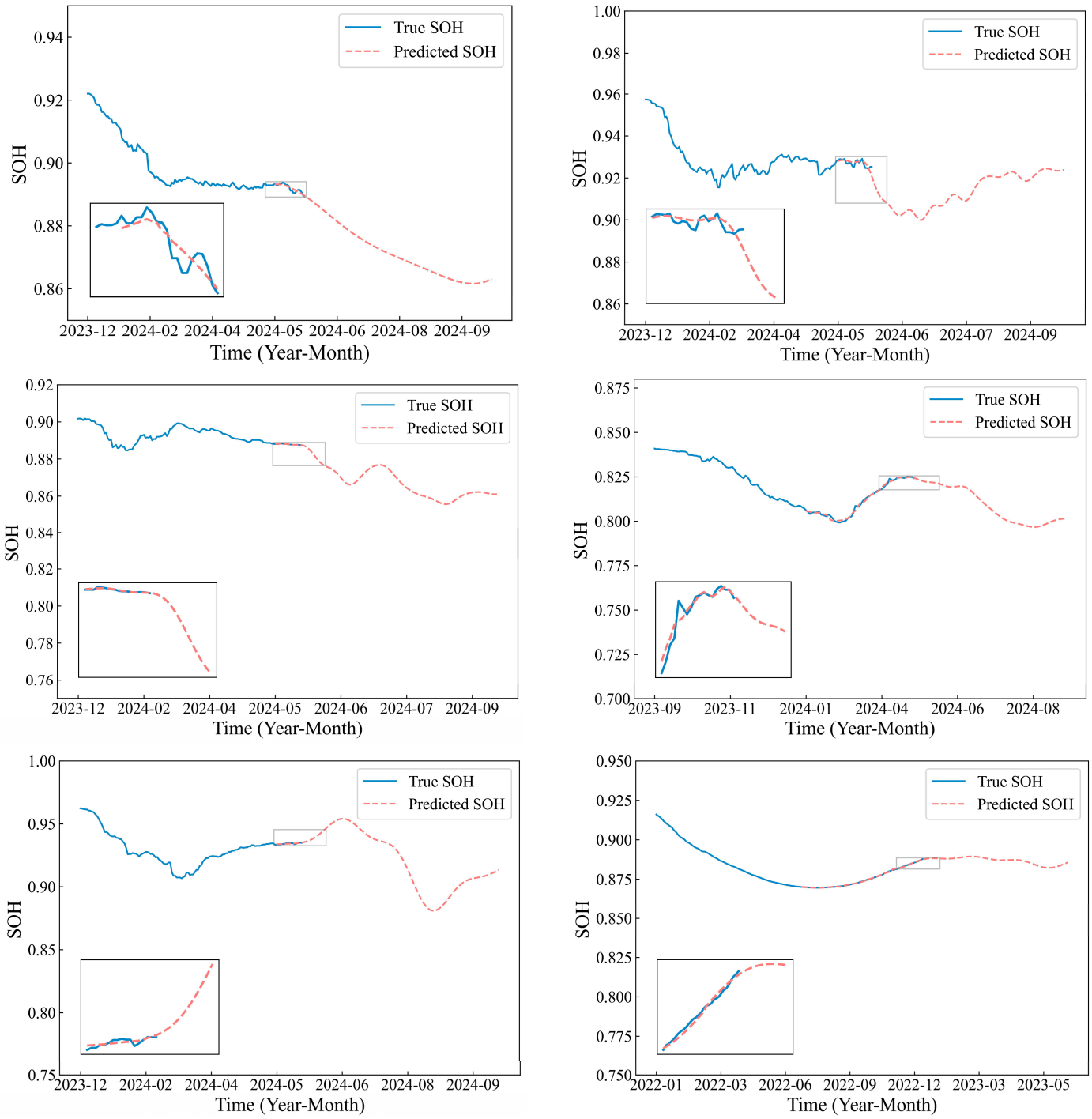


Figure 14. Prediction results of the PWT model.

5.3. Experimental Data Validation

To further verify the practical effectiveness of the proposed battery SOH prediction model, a series of experiments were designed in this study. The experiments focused on disassembling the battery from a real vehicle that had been operating under various working conditions for several years. Full charge/discharge tests were conducted using a bench testing system to obtain more accurate battery state information than that collected directly via onboard sensors. This experiment is conducted independently of the earlier validation on the 86 vehicles. Unlike the previous dataset—collected over a six-month period (December 2023 to June 2024) from vehicles in the field—this battery pack was

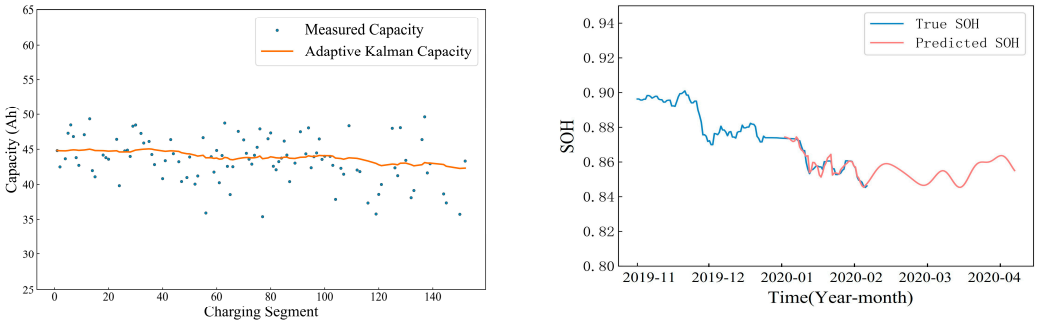
recovered from a different vehicle outside the training and validation set, featuring a different usage history, operating environment, and time span.

The experimental setup is shown in Figure 15. It consists of a battery testing and simulation system used for battery performance validation and battery management system (BMS) research and testing. The black rectangular device in the image is the battery pack under test. The two devices in the background are battery testers, which charge and discharge the battery while monitoring parameters such as voltage, current, and temperature. With this equipment, this study could comprehensively evaluate battery performance and provide more reliable validation data for the SOH prediction model.



Figure 15. Experimental equipment.

After collecting the experimental data, the actual battery data were compared with the SOH predicted by the model. In Figure 16a, the experimental curve of the battery capacity evaluation method proposed in this study shows an estimated value of 42.3 Ah, while the experimentally calculated value is 41.92 Ah. Figure 16b displays the prediction results based on experimental data: the actual collected SOH values (red curve) and the model-predicted SOH values (blue curve) exhibit a similar overall trend. The actual data show a gradual decline, and the overall trend closely matches the true values. According to model evaluation metrics, the RMSE is 0.0254, MAPE is 0.0204, and R^2 is 0.935, further demonstrating the accuracy of the prediction results.



(**a**) (**b**)

Figure 16. Experimental data curve. (**a**) Capacity degradation curves of the test vehicles; (**b**) SOH prediction results of the test vehicles.

6. Conclusions

This paper proposes a PWT model that integrates wavelet decomposition with Transformer architecture, aiming to address the limitations of existing battery capacity prediction

models that overly rely on laboratory data and overlook the impact of real-world operating conditions. The proposed model utilizes wavelet decomposition for multi-scale temporal data processing, combined with attention mechanisms to effectively enhance the modeling of battery degradation patterns.

To address the challenge of hard-to-obtain capacity labels, a fuzzy Kalman filtering method based on ampere-hour integration is proposed. Validation results show a relative error of only 0.906%, significantly improving the accuracy of capacity labeling. In terms of feature engineering, a charging behavior scoring mechanism based on fuzzy membership functions and a current–temperature interaction feature matrix are introduced to further enhance the model's sensitivity to battery degradation. Experimental results demonstrate that the PWT model maintains strong stability and generalization capabilities, with a worst-case RMSE of 0.0226, MAPE of 0.0725, and R^2 of 0.903 under various operating conditions. Compared with models such as LSTM, Transformer, and PatchTST, the proposed approach achieves superior prediction accuracy, robustness, and computational efficiency.

Despite the promising results, this study has certain limitations. One key constraint is the limited diversity of the dataset, which currently includes only a single vehicle model, restricting the generalizability of the proposed method across various vehicle platforms and operating conditions. Future work will focus on building a more diverse dataset encompassing different vehicle types, usage conditions, and degradation patterns to improve the scalability and robustness of the model.

In addition, the current SOH estimation is primarily based on capacity-related features, without fully incorporating other health indicators such as internal resistance and power output consistency. A more comprehensive characterization of battery degradation that integrates multiple dimensions of information could further enhance the accuracy and stability of SOH predictions.

To improve the adaptability and deployment efficiency of the model in practical scenarios, future research will also explore the integration of pre-trained foundation models and fine-tuning techniques. By leveraging transfer learning, it is expected that the model can be quickly adapted to individual vehicle characteristics within large fleets, without the need for extensive retraining. This approach offers a promising path toward efficient, scalable, and personalized SOH prediction across real-world electric vehicle populations.

In summary, the method presented in this study exhibits promising potential in both theoretical innovation and practical application, providing an effective solution for power battery health management.

Author Contributions: Original draft preparation, M.W.; validation, S.Y. and Y.L.; software, L.C.; writing—review and editing, supervision and funding acquisition, J.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Guangxi Major Science and Technology Project, grant number AA23062062.

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

Conflicts of Interest: Min Wei was employed by the SAIC, General Wuling Automobile Co., Ltd. We declare that we have no financial or personal relationships with any other individuals or organizations that could improperly influence our work; there are no professional or other personal interests of any nature or kind in any products, services, and/or companies that could be interpreted as influencing the positions presented in this article or the review of this article's title.

Abbreviations

The following abbreviations are used in this manuscript:

PWT	Patch Wavelet Transformer
SOH	State of Health
ANN	Artificial Neural Network
LSTM	Long Short-Term Memory
PatchTST	Patch Time Series Transformer
MAPE	Mean Absolute Percentage Error
MSE	Mean Squared Error
MAE	Mean Absolute Error
BiLSTM	Bidirectional Long Short-Term Memory Network
RRCF	Robust Random Cut Forest
ICA	Incremental Capacity Analysis
CNN	Convolutional Neural Network
RMSE	Root Mean Squared Error
RUL	Remaining Useful Life
DNN	Deep Neural Network

References

1. Li, H.; Liu, Z.; Su, H. Prediction of remaining useful life of the lithium-ion battery based on improved auxiliary particle filter. In Proceedings of the 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, 22–24 May 2021; pp. 1267–1272.
2. Wang, S.; Li, Y.; Zhou, S.; Chen, L.; Michael, P. Remaining useful life prediction of lithium-ion batteries using a novel particle flow filter framework with grey model. *Sci. Rep.* **2025**, *15*, 3311.
3. Tang, Y.; Zhong, S.; Wang, P.; Zhang, Y.; Wang, Y. Remaining useful life prediction of high-capacity lithium-ion batteries based on incremental capacity analysis and Gaussian kernel function optimization. *Sci. Rep.* **2024**, *14*, 23524. [CrossRef] [PubMed]
4. Hu, J.; Weng, L.; Gao, Z.; Yang, B. State of health estimation and remaining useful life prediction of electric vehicles based on real-world driving and charging data. *IEEE Trans. Veh. Technol.* **2023**, 382–394. [CrossRef]
5. Ren, L.; Zhao, L.; Hong, S.; Zhao, S.; Wang, H.; Zhang, L. Remaining useful life prediction for lithium-ion battery: A deep learning approach. *IEEE Access* **2018**, *6*, 50587–50598. [CrossRef]
6. Zhao, F.; Gao, D.; Cheng, Y.; Yang, Q. Application of state of health estimation and remaining useful life prediction for lithium-ion batteries based on AT-CNN-BiLSTM. *Sci. Rep.* **2024**, *14*, 29026. [CrossRef]
7. Hou, E.; Qiao, X.; Liu, G. Remaining useful life prediction of power lithium-ion battery based on artificial neural network model. In Proceedings of the International Conference on Mechanical, Electronic, Control and Automation Engineering (MECAE), Beijing, China, 25–26 March 2017; pp. 371–374.
8. Mamo, T.; Wang, F. Attention-Based Long Short-Term Memory Recurrent Neural Network for Capacity Degradation of Lithium-Ion Batteries. *Batteries* **2021**, *7*, 66. [CrossRef]
9. Zhang, H.; Li, Y.; Zheng, S.; Lu, Z.; Gui, X.; Xu, W.; Bian, J. Battery lifetime prediction across diverse ageing conditions with inter-cell deep learning. *Nat. Mach. Intell.* **2025**, *7*, 270–277. [CrossRef]
10. Iftikhar, M.; Shoaib, M.; Altaf, A.; Iqbal, F.; Villar, S.; Lopez, L.; Ashraf, I. A deep learning approach to optimize remaining useful life prediction for Li-ion batteries. *Sci. Rep.* **2024**, *14*, 25838. [CrossRef]
11. Sun, C.; Lu, T.; Li, Q.; Liu, Y.; Yang, W.; Xie, J. Predicting the Future Capacity and Remaining Useful Life of Lithium-Ion Batteries Based on Deep Transfer Learning. *Batteries* **2024**, *10*, 303. [CrossRef]
12. Wang, F.; Zhao, Z.; Ren, J.; Zhai, Z.; Wang, S.; Chen, X. A transferable lithium-ion battery remaining useful life prediction method from cycle-consistency of degradation trend. *J. Power Sources* **2022**, *521*, 230975. [CrossRef]
13. Ma, G.; Xu, S.; Jiang, B.; Cheng, C.; Yang, X.; Shen, Y.; Yang, T.; Huang, Y.; Ding, H.; Yuan, Y. Real-time personalized health status prediction of lithium-ion batteries using deep transfer learning. *Energy Environ. Sci.* **2022**, *15*, 4083–4094. [CrossRef]
14. Fan, W.; Zhu, J.; Qiao, D.; Jiang, B.; Wang, X.; Wei, X.; Dai, H. Prediction of nonlinear degradation knee-point and remaining useful life for lithium-ion batteries using relaxation voltage. *Energy* **2024**, *294*, 130900. [CrossRef]
15. Krishna, G.; Singh, R.; Gehlot, A.; Almogren, A.; Altameem, A.; Rehman, A.U.; Hussen, S. Advanced battery management system enhancement using IoT and ML for predicting remaining useful life in Li-ion batteries. *Sci. Rep.* **2024**, *14*, 30394. [CrossRef] [PubMed]
16. Li, Z.; Shen, S.; Ye, Y.; Cai, Z.; Zhen, A. An interpretable online prediction method for remaining useful life of lithium-ion batteries. *Sci. Rep.* **2024**, *14*, 12541. [CrossRef] [PubMed]

17. Ly, S.; Xie, J.; Wolter, F.; Nguyen, H.; Weng, Y. T-shape data and probabilistic remaining useful life prediction for Li-ion batteries using multiple non-crossing quantile long short-term memory. *Appl. Energy* **2023**, *349*, 121355. [CrossRef]
18. Huang, Z.; Zhou, X.; Gao, D.; Zhang, X.; Jiang, F.; Chen, B.; Yang, Y.; Wu, M.; Peng, J. A novel label-free supervision learning method for lithium-ion battery RUL prediction. In Proceedings of the 2019 IEEE Energy Conversion Congress and Exposition (ECCE), Baltimore, MD, USA, 29 September–3 October 2019; pp. 3150–3156.
19. Jafari, S.; Byun, Y. Optimizing battery RUL prediction of lithium-ion batteries based on Harris Hawk optimization approach using random forest and LightGBM. *IEEE Access* **2023**, *11*, 87034–87046. [CrossRef]
20. Pang, X.; Huang, R.; Wen, J.; Shi, Y.; Jia, J.; Zeng, J. A lithium-ion battery RUL prediction method considering the capacity regeneration phenomenon. *Energies* **2019**, *12*, 2247. [CrossRef]
21. Cheng, G.; Wang, X.; He, Y. Remaining useful life and state of health prediction for lithium batteries based on empirical mode decomposition and a long and short memory neural network. *Energy* **2021**, *232*, 121022. [CrossRef]
22. Kwon, S.; Han, D.; Park, J.; Lee, P.; Kim, J. Joint state-of-health and remaining-useful-life prediction based on multi-level long short-term memory model prognostic framework considering cell voltage inconsistency reflected health indicators. *J. Energy Storage* **2022**, *55*, 105731. [CrossRef]
23. Severson, K.A.; Attia, P.M.; Jin, N.; Perkins, N.; Jiang, B.; Yang, Z.; Chen, M.H.; Aykol, M.; Herring, P.K.; Fraggedakis, D.; et al. Data-driven prediction of battery cycle life before capacity degradation. *Nat. Energy* **2019**, *4*, 383–391. [CrossRef]
24. Yang, A.; Zhang, G.; Tian, C.; Peng, W.; Liu, Y. Charging Behavior Portrait of Electric Vehicle Users Based on Fuzzy C-Means Clustering Algorithm. *Energies* **2024**, *17*, 1651. [CrossRef]
25. Fleischhammer, M.; Waldmann, T.; Bisle, G.; Hogg, B.; Wohlfahrt-Mehrens, M. Interaction of cyclic ageing at high-rate and low temperatures and safety in lithium-ion batteries. *J. Power Sources* **2015**, *274*, 432–439. [CrossRef]
26. Werner, D.; Sabine, P.; Achim, W. Inhomogeneous temperature distribution affecting the cyclic aging of Li-ion cells. Part I: Experimental Investigation. *Batteries* **2020**, *6*, 13. [CrossRef]
27. Taghavian, H.; Andersson, M.; Johansson, M. Model-Free Fast Charging of Lithium-Ion Batteries by Online Gradient Descent. *arXiv* **2024**, arXiv:2405.10623.
28. Wu, H.; Xu, J.; Wang, J.; Long, M. Autoformer: Decomposition Transformers with Auto-Correlation for Long-Term Series Forecasting. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 22419–22430.
29. Mallat, S. *A Wavelet Tour of Signal Processing*, 2nd ed.; Academic Press: Cambridge, MA, USA, 1999.
30. Liu, S.; Wang, X. An improved transformer based traffic flow prediction model. *Sci. Rep.* **2025**, *15*, 8284. [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.