Advances in Computer, Signals and Systems (2026)
Clausius Scientific Press, Canada

DOI: 10.23977/acss.2026.100204
ISSN 2371-8838 Vol. 10 Num. 2

PL-LSTM-Based State of Health Estimation and TTE Prediction for Lithium-Ion Batteries

Yeqi Zhang[1], Zekai Yu[1], Wenyue Wang[1], Yanyan Wu[2,3,4,*]

[1]Big Data College, Fuzhou College of Foreign Studies and Trade, Fuzhou, China
[2]College of Artificial Intelligence, Ningbo University of Finance and Economics, Ningbo, China
[3]Faculty of Data Science, City University of Macau, Macau, China
[4]Key Laboratory of Data Science and Intelligent Computing, Fuzhou University of International Studies and Trade, Fuzhou, Fujian, China
**Corresponding Author*

Keywords: Lithium-ion battery, SOH estimation, TTE prediction, PINN, LSTM, physical constraints, battery management system

Abstract: This study proposes a hybrid model, PL-LSTM, that integrates a Physics-Informed Neural Network (PINN) with a Long Short-Term Memory (LSTM) network for estimating the State of Health (SOH) and predicting the Time to End-of-Life (TTE) of lithium-ion batteries. The model extracts dynamic temporal features within an extremely short time window and embeds physical constraints—derived from the first-order RC equivalent circuit and SOC evolution equation—into a composite loss function, thereby achieving an organic integration of data-driven learning and mechanistic modeling. Experimental results demonstrate that PL-LSTM significantly outperforms traditional LSTM models on the CALCE battery aging dataset and the GreenHub dynamic load dataset. For SOH estimation, PL-LSTM effectively reduces MAE and RMSE, accurately capturing nonlinear degradation and abrupt capacity changes; for TTE prediction, the model reveals the dual-driving mechanism of capacity decay and internal resistance growth, validating the amplifying effect of aging on lifespan reduction under high-load scenarios. Overall, PL-LSTM enhances the accuracy and robustness of battery state estimation while providing interpretable physical insights, offering a new technical pathway for intelligent and reliable battery management in electric vehicles, energy storage systems, and mobile devices.

1. Introduction

Predicting the State of Health (SOH) and Time to End-of-Life (TTE) of lithium-ion batteries is a core task of battery management systems. Traditional methods rely on ampere-hour integration and equivalent circuit modeling, which can provide high accuracy during complete charge-discharge cycles. However, in real-world state estimation scenarios—particularly during the cold-start phase at the beginning of discharge—these methods are often limited by measurement lag and noise interference. These limitations make it difficult for traditional methods to meet the real-time prediction requirements in dynamic load environments.

To address this issue, this study proposes a hybrid model PL-LSTM that integrates a Physical

Information Neural Network (PINN) with a Long Short-Term Memory (LSTM) network. This model utilizes LSTM to extract the battery's dynamic temporal features within an extremely short time window. Simultaneously, by constructing a battery load estimation model based on the superposition of component power consumption and introducing a first-order RC equivalent circuit to characterize voltage relaxation effects, the model embeds the battery's physical constraints into the composite loss function of the deep temporal network. This ensures that the prediction results possess both data-driven high fitting accuracy and compliance with the battery's physical laws. Experimental results demonstrate that PL-LSTM significantly reduces SOH estimation errors on the CALCE dataset.

2. SOC Theory and Battery Physical Modeling

2.1 Battery Load Estimation Based on Component Power Consumption

The State of Charge (SOC) of a battery is defined as the ratio of the remaining dischargeable capacity $Q(t)$ to the current maximum available capacity $C(t)$. Its mathematical definition is given by [1]:

$$SOC(t) = \frac{Q(t)}{C(t)} \tag{1}$$

In the equation, $Q(t)$ represents the remaining charge at time t, and $C(t)$ represents the available capacity of the battery at time t, calculated as $C(t) = C_n \cdot SOH(t)$, where C_n is the battery's initial rated capacity and $SOH(t)$ is the state of health at time t.

Since the current $I(t)$ represents the time derivative of the charge, substituting it into the SOC definition yields the most basic continuous-time state equation:

$$\frac{dSOC(t)}{dt} = \frac{1}{C(t)} \cdot \frac{dQ(t)}{dt} = -\frac{I(t)}{C(t)} \tag{2}$$

To achieve an efficient and accurate estimation of battery load on mobile devices, this study proposes a Power Summation Model (PCM) based on the superposition of power consumption from multiple components. The total power of the device, $P(t)$, can be decomposed into the sum of the linear superposition of power consumption from each independent hardware module and a nonlinear correction term:

$$P(t) = P_{base} + P_{screen}(B) + P_{cpu}(C) + P_{net}(type) + P_{gps}(G) + \epsilon \tag{3}$$

The specific analytical expression is as follows:

$$P(t) = \theta_0 + \theta_{scr}B(t) + \theta_{cpu}C(t) + \sum_{k \in \mathcal{N}} \theta_k s_k(t) + \theta_{gps}G(t) + \theta_{app}N_{app}(t) \tag{4}$$

$B(t)$ and $C(t) \in [0,1]$ represent the normalized screen brightness and CPU utilization, respectively. "wifi", "4g", and "5g" denote network communication protocol types; $s_k(t)$ is an indicator function representing the on/off state of the corresponding network; $G(t)$ and $N_{app}(t)$ represent the activation state of the GPS module and the number of background resident applications, respectively; $\theta_{(\cdot)}$ is the physical power consumption coefficient to be identified.

Each coefficient θ_i intuitively reflects the inherent power consumption characteristics of the corresponding hardware module; when deploying across device platforms, rapid migration can be achieved simply by recalibrating each θ coefficient.

In terms of parameter identification, this study employs OLS for coefficient estimation. Compared to traditional black-box machine learning models, the physical coefficients derived from OLS

solutions offer clear mechanistic interpretability and can intuitively map the energy consumption contributions of each hardware component. The experiment selected complete discharge trajectory data for three representative mobile devices—the Pixel 5, Galaxy S21, and iPhone 13—from the GreenHub public dataset [2], and divided the data into training and validation sets in an 8:2 ratio.

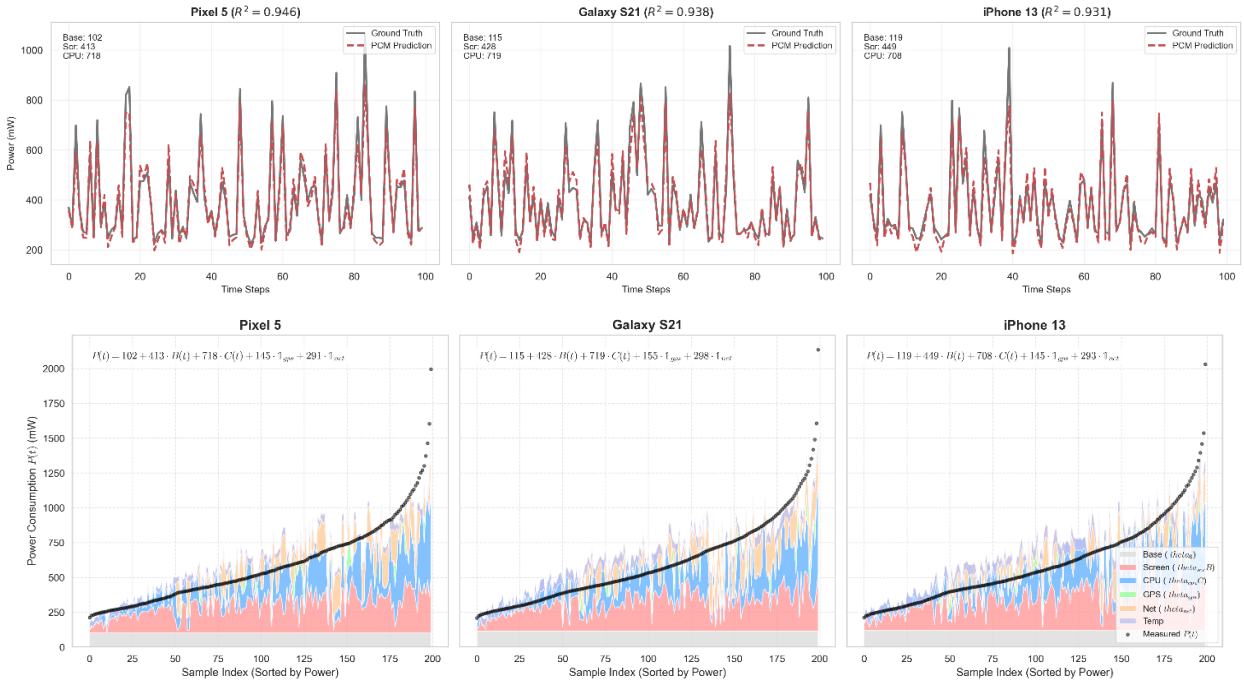

Figure 1: PCM Model Verification.

The upper half of Figure 1 shows the R ²values for the three models on the training and test sets; all test results exceed 0.9, confirming the high fidelity of the models. The lower half compares the time-response curves of the actual measured power and the power predicted by the models. The results indicate that the PCM model can accurately track and reproduce the dynamic fluctuations in terminal power.

2.2 First-Order RC Equivalent Circuit

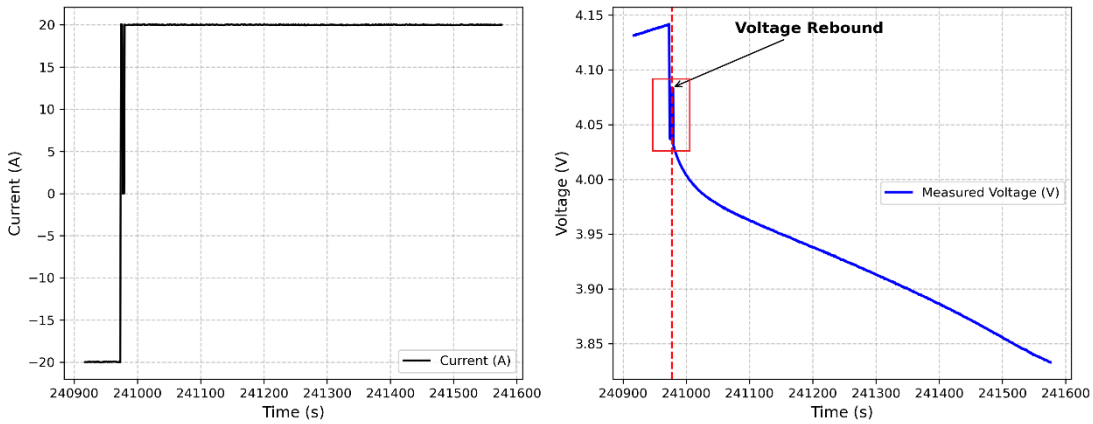

Figure 2: Voltage Rebound Effect.

In addition to obtaining the transient load power $P(t)$, it is necessary to further incorporate the dynamically calculated terminal voltage $V_t(t)$ response characteristics. Through an in-depth

analysis of the GreenHub and CALCE battery aging datasets, this study observed a significant relaxation effect.[3]

As shown in Figure 2, when a step change occurs in the load current, the battery terminal voltage does not exhibit an ideal instantaneous jump but rather a gradual, exponential rebound characteristic. This transient process is primarily attributed to the complex electrochemical kinetic mechanisms within lithium-ion batteries, particularly concentration polarization and electrochemical polarization effects.

Given that traditional R-int equivalent circuit models struggle to accurately capture the transient polarization characteristics described above, this study introduces a first-order RC equivalent circuit model. The core physical definition of this model is as follows:

$V_{OCV}(SOC)$: An ideal voltage source controlled by a SOC represents the thermodynamic equilibrium potential.

R_0: Internal resistance, which represents the contact resistance between the electrolyte and the electrode, causes an instantaneous voltage drop.

$R_p//C_p$: Polarization-parallel circuit, used to characterize the transient processes of concentration polarization and electrochemical polarization.

In particular, the $V_{OCV}(SOC)$ curve data is taken from Jonas A. Braun's experimental results[4,5], as shown in Figure 3:

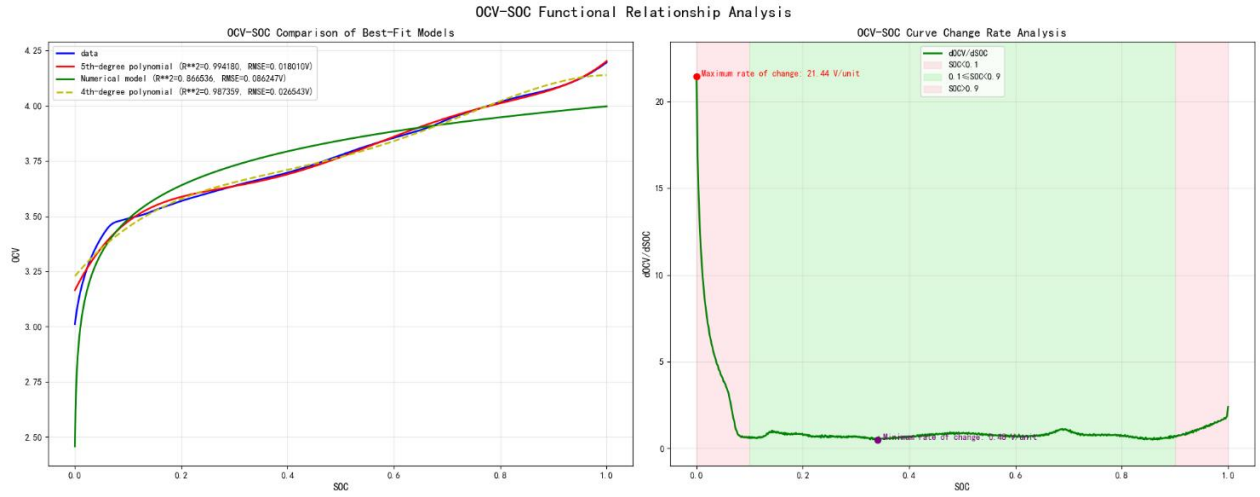

Figure 3: OCV-SOC Curve.

As shown on the left side of Figure 3, the study employed polynomial regression to fit the nonlinear mapping relationship between OCV and SOC. The results indicate that a fifth-order polynomial exhibits the best accuracy and is capable of accurately capturing the nonlinear characteristics in the low-charge range; the specific fitting coefficients are shown in Table 1.

Table 1: Polynomial coefficients of the OCV–SOC fitting model.

Term	Coefficient
SOC_5	16.2152
SOC_4	-44.8389
SOC_3	45.6969
SOC_2	-20.8191
SOC_1	4.7842

As shown in the sensitivity analysis on the right, within the extremely low or high SOC ranges, the terminal voltage is highly sensitive to even minor changes in SOC, resulting in a significant

amplification of estimation errors; in contrast, within the intermediate charge ranges, the voltage response is relatively flat, making this range more suitable for high-precision continuous SOC prediction.

The internal dynamics of a mobile device battery can be described by a coupled system consisting of the SOC integration equation and the polarization voltage differential equation:

$$\begin{cases} \dot{SOC}(t) = \dfrac{dSOC(t)}{dt} = -\dfrac{I(t)}{C(t)} \\ \dot{V}_p(t) = -\dfrac{V_p(t)}{R_p C_p} + \dfrac{I(t)}{C_p} \end{cases} \tag{5}$$

Combining the SOC evolution equation with the differential equation for the polarization voltage yields the following observation equation for the terminal voltage $V_t(t)$:

$$V_t(t) = V_{OCV}\big(SOC(t)\big) - V_p(t) - I(t)R_0 \tag{6}$$

Compared to the R-int model, which lacks a memory effect, the first-order RC model introduces the polarization voltage state variable V_p, thereby enabling accurate tracking of the gradual hysteresis characteristics of the voltage following a load step. At the moment the load is removed, the terminal voltage does not immediately return to steady state but instead undergoes two sequential phases: a "transient ohmic jump" followed by a "gradual polarization recovery". This behavior closely matches the voltage dynamic response data obtained from the CALCE experimental platform.[6]

Based on the terminal voltage observation equation, the equivalent dynamic current can be derived as $I_t(t) = \frac{P(t)}{V_t(t)}$, where $P(t)$ is the measured transient load power and $V_t(t)$ is the observed terminal voltage.

Substituting into Equation 2 yields the final SOC continuous-time differential equation and its time-domain integral form in terms of t':

$$\frac{dSOC(t)}{dt} = -\frac{1}{C(t)} \cdot \frac{P(t)}{V_t(t)} \tag{7}$$

$$SOC(t) = SOC(0) - \frac{1}{C(t)} \int_0^t \frac{P(t')}{V_t(t')} dt' \tag{8}$$

3. Transient SOH Estimation via PL-LSTM

Traditional SOH estimation methods primarily rely on ampere-hour integration, requiring a complete charge-discharge cycle or a prolonged period of charge accumulation to produce results with high confidence. However, in real-world dynamic applications, such purely mechanistic methods suffer from observation lag during the "cold start" phase; due to the extremely small initial charge accumulation and low signal-to-noise ratio, they are easily affected by high-frequency sensor noise. To address this issue, this study proposes an LSTM-based PINN framework for transient SOH estimation. The overall architecture is shown in Figure 4.

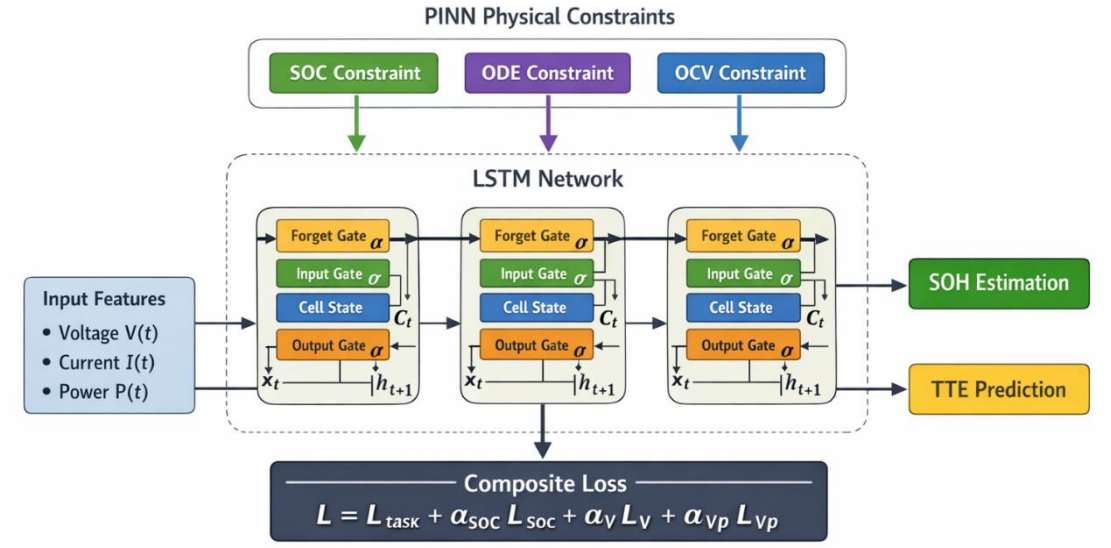

Figure 4: PL-LSTM Architecture Diagram.

3.1 LSTM Sequential Networks

To capture the complex temporal characteristics of batteries, this study employs a Long Short-Term Memory (LSTM) network as its core architecture. [7] Compared to traditional RNNs, LSTMs effectively mitigate the vanishing gradient problem through their gating mechanism and can retain the ability to remember long-term temporal dependencies even in noisy environments. Their core computational process is as follows:

$$\begin{gathered} f_t = \sigma(W_f[h_{t-1}, x_t] + b_f), i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \\ \tilde{C}_t = tanh(W_c[h_{t-1}, x_t] + b_c), C(t) = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \\ o_t = \sigma(W_o[h_{t-1}, x_t] + b_o), h_t = o_t \odot tanh(C(t)) \end{gathered} \tag{9}$$

In the equation, σ denotes the sigmoid activation function, tanh is used for state updates, and $\odot$ denotes the Hadamard product. Through the combination of the forget gate, input gate, and output gate, the network is able to extract stable SOH evolution features under dynamic loads and provide a reliable temporal representation for subsequent physical constraint coupling.

The model inputs are voltage $V(t)$, current $I(t)$, power $P(t)$, cycle count N, and time step t. These features form a time series of length T, which serves as the input to the PL-LSTM network for extracting battery aging and dynamic response features, with the final output being $SOH(t)$ as the prediction target.

3.2 Design of the PINN Composite Loss Function

Physics-Informed Neural Networks are a class of models designed for supervised learning tasks. Unlike traditional, purely data-driven neural networks, PINNs not only fit the distribution patterns of sample data during training but also explicitly incorporate physical laws or intrinsic constraints described by partial differential equations and other mathematical forms. By imposing physical constraints on the optimization objective, PINNs can obtain models with greater generalization capabilities even under limited data conditions. [8][9][10]

To ensure that the network not only fits the observed data during training but also adheres to the underlying physical laws governing the battery, this study incorporates the continuous-time SOC equation derived earlier and the first-order RC equivalent circuit model into the optimization

objective function of the PINN[10], thereby establishing a composite loss framework with joint physical constraints. The specific form is as follows:

The continuous-time differential equation for the SOC is given by Equation (8):

$$L_{SOC} = \| \frac{dSOC(t)}{dt} + \frac{P(t)}{C(t)V(t)} \|_2^2 \tag{10}$$

This constraint ensures that the SOC evolution predicted by the network complies with the law of charge conservation, thereby preventing the occurrence of "phantom charge" or drifts that violate physical laws.

The terminal voltage equation can be derived from Equation (6) and the fifth-order polynomial fit shown in Table 1:

$$L_V = \| V(t) - \left[V_{OCV}\left(\hat{SOC}(t)\right) - \hat{V}_p(t) - I(t)R_0 \right] \|^2 \tag{11}$$

Ensure that the predicted voltage matches the actual measured value and conforms to the first-order RC circuit model, thereby avoiding voltage response distortion caused by overfitting.

Equation (5) yields the dynamic equation for the polarization voltage:

$$L_{Vp} = \| \frac{d\hat{V}_p(t)}{dt} + \frac{1}{R_p C_p}\hat{V}_p(t) - \frac{1}{C_p} I(t) \|^2 \tag{12}$$

This constraint ensures that the network can accurately capture the relaxation effect of the battery following a sudden load change, thereby preventing voltage prediction lag caused by neglecting the polarization process.

The final composite loss function is the weighted sum of the above constraints:

$$L = L_{task} + \alpha_{SOC} L_{SOC} + \alpha_V L_V + \alpha_{Vp} L_{Vp} \tag{13}$$

Here, L_{task} represents the loss function for the prediction task itself. A weighting coefficient α is introduced to balance the influence of each physical constraint term during the overall optimization process, ensuring that the network maintains a high degree of fit to the observed data while strictly adhering to the physical mechanisms of the battery.

3.3 Network Hyperparameter Settings

This study systematically tuned the core hyperparameters of PINN-LSTM. The configuration is shown in Table 2 and includes a two-layer LSTM with 64 hidden units, Tanh and Sigmoid activation functions, the Adam optimizer, an initial learning rate of 0.005, a physical constraint weight of 0.8, an ODE constraint weight of 0.1, a maximum of 500 iteration rounds, and a 30-minute observation time window.

Table 2: Hyperparameter Configuration.

Hyperparameter	Value
Network layers	2 layers LSTM
Activation function	Tanh (State), Sigmoid (Gate)
Optimizer	Adam
Learning Rate	0.005
λ_{phy}	0.8
λ_{ode}	0.1
Epochs	500

4. Analysis of Experimental Results

4.1 SOH Estimation Results and Error Analysis

The study employed a hybrid dataset combining the CALCE battery aging dataset with dynamic load data from GreenHub mobile devices, and divided the dataset into training, validation, and test sets in a 7:2:1 ratio.

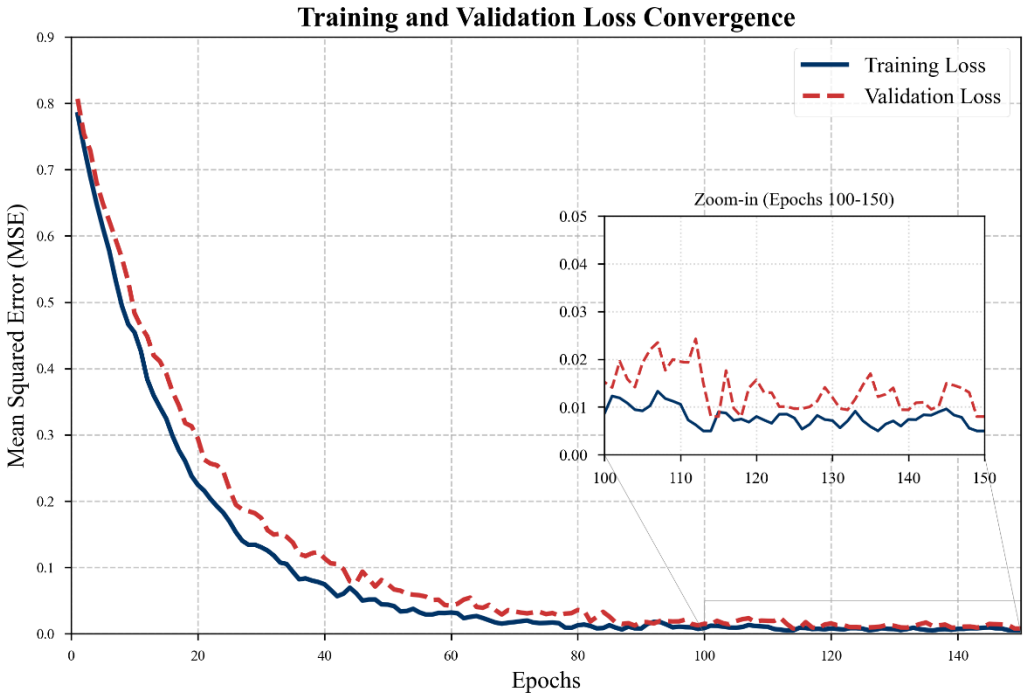

Figure 5: Convergence curve of the PL-LSTM loss.

Figure 5 illustrates the convergence of the PL-LSTM model during training. The model's error decreases rapidly in the initial stages and stabilizes around 80 epochs. The validation set loss curve closely follows that of the training set, with no signs of significant overfitting, and the MSE remains stable at an extremely low level.

4.2 Comparison of SOH Estimation Errors Based on the CALCE Dataset

To further validate the model's predictive and generalization capabilities during actual charge-discharge cycles, the experiment selected two batteries—CS2_35 and CS2_36—from the CALCE dataset. This dataset documents the aging process of battery SOH and is characterized by a large number of high-frequency capacity fluctuations and nonlinear accelerated aging phenomena.

As shown in Figures 6 and 7, the LSTM model is able to fit the overall degradation trend reasonably well, but exhibits significant prediction errors during periods of high-frequency capacity fluctuations. This is particularly evident during the later accelerated aging phase, where the predicted values are significantly overestimated and exhibit severe fluctuations. In contrast, the PL-LSTM model is able to track the actual capacity degradation path with high fidelity; specifically, during the "plunge" phase, the model continues to accurately capture and closely follow the actual cliff-like drop in performance.

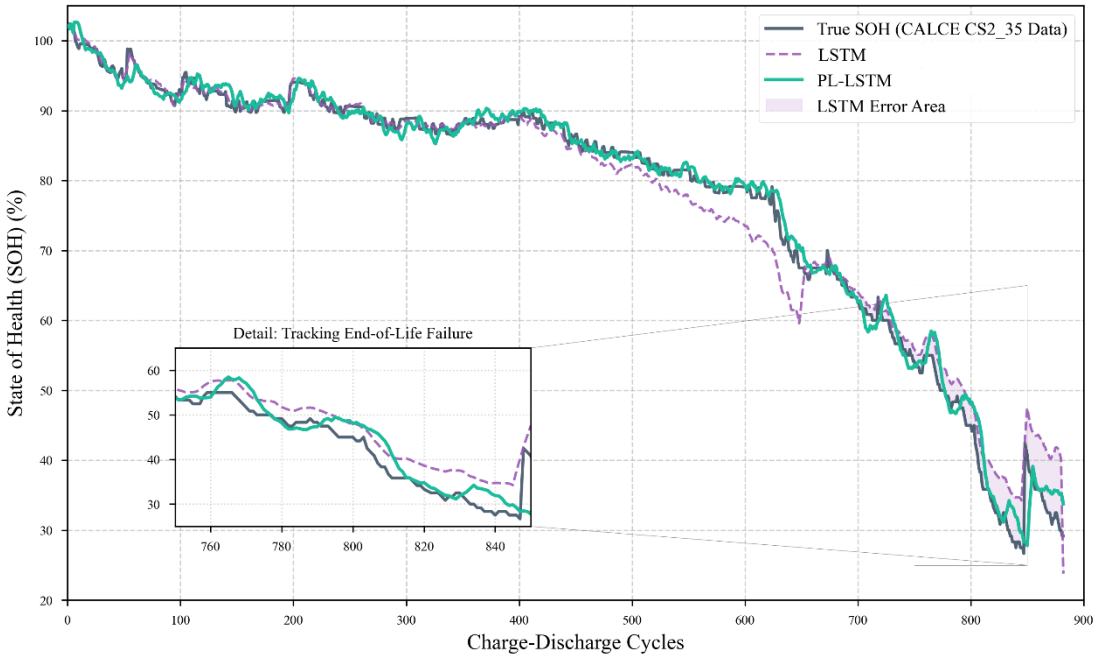

Figure 6: Comparison of SOH Prediction Results for CS2_35.

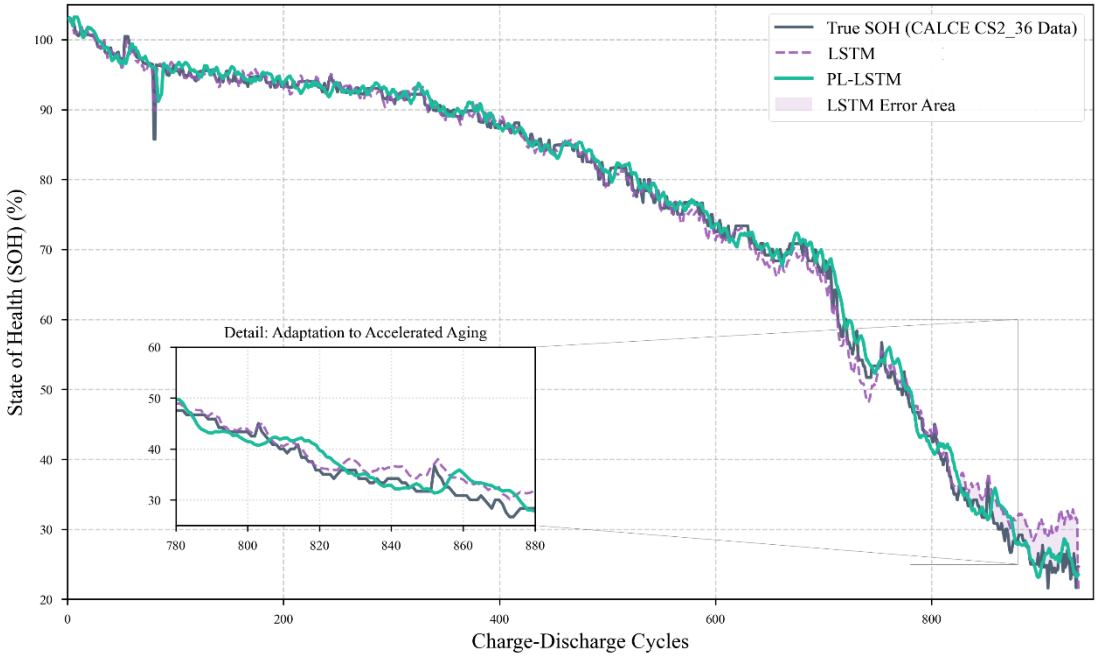

Figure 7: Comparison of SOH Prediction Results for CS2_36.

To evaluate the predictive accuracy of each model, this study selected MAE and RMSE as the primary evaluation metrics. Tables 3 and 4 present the error results for the LSTM and PL-LSTM models on the CS2_35 and CS2_36 datasets, respectively.

As shown in Tables 3 and 4, compared to LSTM, PL-LSTM significantly reduced both MAE and RMSE for both batteries. Taking CS2_35 as an example, the MAE decreased from 0.0201 to 0.0124, and the RMSE decreased from 0.0310 to 0.0190. In the cross-battery validation scenario, the error of the LSTM increased with distribution shifts, while the PL-LSTM maintained a smaller fluctuation in error, indicating that the introduction of physical constraints can effectively suppress prediction drift and enhance robustness.

Table 3: CS2_35 SOH estimation error.

Indicators	LSTM	PL-LSTM
MAE	0.0201	0.0124
RMSE	0.0310	0.0190

Table 4: CS2_36 SOH estimation error.

Indicators	LSTM	PL-LSTM
MAE	0.0250	0.0122
RMSE	0.0444	0.0184

4.3 TTE Prediction Methods and Load Scenario Analysis

Based on the continuous-time SOC dynamic model derived earlier, the terminal voltage solution V_{base} in the steady state can be derived as follows:

$$V_{base}(SOC, P(t)) = \frac{V^0(SOC) + \sqrt{(V^0(SOC))^2 - 4P(t)R}}{2} \tag{14}$$

In the equation, $V^0(SOC)$ is the pre-calibrated open-circuit voltage function $V_{OCV}(SOC)$.

By substituting V_{base} into the SOC time-domain differential equation and integrating with respect to the time variable, the TTE prediction formula can be derived:

$$TTE_{base} = \int_0^{z(t_0)} \frac{C_{sim}^n \cdot V_{base}(\zeta, P(t))}{P(t)} dz \tag{15}$$

To improve prediction accuracy, this study uses the SOH estimates output by the PL-LSTM network as input to correct the capacity and internal resistance. The internal resistance term is corrected using the following equation:

$$R_{aging} = R_{new} \cdot \left(1 + \alpha(1 - SOH(t))\right) \tag{16}$$

The $SOH(t)$ is derived from the transient awareness results of the PL-LSTM, where α is an empirical coefficient. The final TTE prediction formula is:

$$TTE_{final} = \int_0^{z(t_0)} \frac{C_n \cdot SOH(t) \cdot V_{aging}(SOC, P(t), SOH(t))}{P(t)} dz \tag{17}$$

In particular, the terminal voltage V_{aging} associated with the coupling aging effect is given by:

$$V_{aging}(SOC, P(t), SOH(t)) = \frac{V^0(SOC) + \sqrt{(V^0(SOC))^2 - 4P(t) \cdot R_{aging}}}{2} \tag{18}$$

To provide a visual analysis of the mechanism by which battery aging affects TTE, this study designed three typical interactive load scenarios to simulate real-world user habits. As shown in Figure 8, the predicted TTE boundaries for a new battery with an SOH of 100% and an aged battery with an SOH of 83.2% under the same load conditions are overlaid and compared.

As shown in Figure 8, under high-load conditions, the terminal voltage of aged batteries drops significantly due to increased internal resistance, causing the device to reach the cutoff voltage threshold earlier. The reduction in TTE is approximately 28%, far exceeding the capacity decay rate of about 16.8%. Under low-load conditions, however, the reduction in TTE is primarily determined by capacity decay, amounting to approximately 20%.

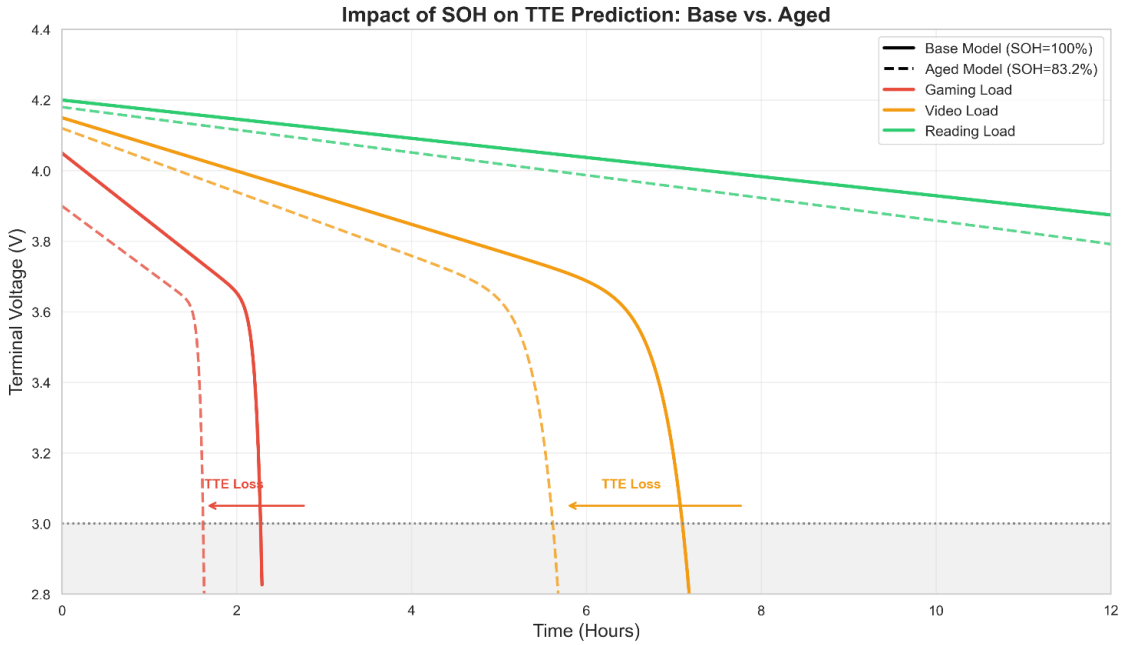

Figure 8: Analysis of the Impact of SOH Decay on TTE Prediction Boundaries.

These experimental results indicate a coupling effect between battery aging and load intensity. The ohmic voltage drop under high power dissipation conditions accelerates the failure process, leading to a significant amplification of the actual reduction in TTE. This validates the predictive mechanism proposed in this study, namely the “dual-driven mechanism of capacity decay and internal resistance increase.”

5. Conclusion

This study proposes a hybrid model PL-LSTM that integrates a Physical Information Neural Network (PINN) with LSTM network for estimating the SOH and predicting the Time to Failure (TTE) of lithium-ion batteries. This model effectively extracts dynamic temporal features of the battery within an extremely short time window. By embedding a first-order RC equivalent circuit and the SOC evolution equation, it incorporates physical constraints into the composite loss function, achieving an organic integration of data-driven learning and mechanistic constraints.

Experimental results demonstrate that PL-LSTM achieves significantly higher prediction accuracy than traditional LSTM models on the CALCE battery aging dataset and the GreenHub dynamic load dataset. Regarding SOH estimation, PL-LSTM effectively reduces MAE and RMSE, accurately capturing the nonlinear degradation and abrupt changes in capacity; regarding TTE prediction, the model reveals the dual-driving mechanism of capacity decay and internal resistance increase, verifying the amplifying effect of aging on lifespan reduction under high-load scenarios.

In summary, PL-LSTM not only enhances the accuracy and robustness of battery state estimation but also provides interpretable physical insights for real-time prediction in dynamic load environments. This method offers a new technical pathway for improving the intelligence and reliability of BMS systems. In the future, it can be further extended to multi-scenario battery applications and cross-platform transfer, supporting the safe and efficient operation of new energy vehicles, energy storage systems, and mobile devices.

Acknowledgments

This research was supported by the Scientific Research Project of Zhejiang Provincial Department of Education [GrantNo. Y202559174]; Open Research Project of Key Laboratory of Data Science and Intelligent Computing, Fuzhou University of International Studies and Trade [Grant No.SZ2025001].

References

[1] Yao Fang, Yang Xiaona, Li Qian, et al. Cooperative Estimation of SOC and SOH in Lithium-Ion Batteries Considering Coupling and Correlation Characteristics [J]. Power Electronics Technology, 2024, 58(11): 85–90.

[2] Pereira, R., Matalonga H ,Couto M , et al. GreenHub: a large-scale collaborative dataset to battery consumption analysis of android devices[J].Empirical Software Engineering,2021,26(3):1-55.

[3] DOYLE M, FULLER T F, NEWMAN J. Modeling of Galvanostatic Charge and Discharge of the Lithium/Polymer/Insertion Cell[J]. Journal of The Electrochemical Society, 1993, 140(6): 1526-1533.

[4] S. Liu et al., "State of Charge Estimation of Lithium Battery Based on Improved Unscented Kalman filter," 2024 IEEE 10th International Conference on Underwater System Technology: Theory and Applications (USYS), Xi'an, China, 2024, pp. 1-6.

[5] A. J B ,Ren é B ,David S , et al. State of charge and state of health diagnosis of batteries with voltage-controlled models[J].Journal of Power Sources,2022,544.

[6] Khan M S, Maddipatla S, Pecht M. A review of float charging in lithium-ion batteries: Degradation mechanisms, influencing factors, and optimization strategies[J]. Journal of Power Sources, 2026, 616: 235861.

[7] Mei Yuankun. A Study on Vehicle Intention Recognition on Expressways Using PL-LSTM [J]. Modern Information Technology, 2026, 10(01): 81-85.

[8] Guo Shengpin, Wang Huiming. Deep Learning Research on Concrete Failure Criteria Based on Physical Information Neural Networks [J]. Concrete, 2024, (09): 28–34.

[9] An Peng, Cao Danping, Zhao Baoyin, et al. A Study on Reservoir Property Parameter Prediction Methods Based on LSTM Recurrent Neural Networks [J]. Advances in Geophysics, 2019, 34(5): 1849–1858.

[10] WANG Q, GAO T, LI X C. SOC estimation of Lithium-ion battery based on equivalent circuit model with variable parameters[J]. Energies, 2022, 15(16):5829.