The RUL prediction of Li-Ion batteries based on adaptive LSTM

Samrat Koirala Thakuri[1], Huibo Li[2], Diwang Ruan[2]*, Xianyu Wu[2]

[1]Technische Universität Berlin, School of Electrical Engineering and Computer Science, Germany, 10625

[2]Hypersonic Technology Laboratory, National University of Defense Technology, Changsha, 410073, P. R. China

*Corresponding author: ruandiwang607@163.com

Received Month X, XXXX | Accepted Month X, XXXX | Posted Online Month X, XXXX

Abstract: With the widespread adoption of electric vehicles and energy storage systems, predicting the remaining useful life (RUL) of lithium-ion batteries (LIBs) is critical for enhancing system reliability and enabling predictive maintenance. Traditional RUL prediction methods often exhibit reduced accuracy during the nonlinear aging stages of batteries and struggle to accommodate complex degradation processes. This paper introduces a novel adaptive long short-term memory (LSTM) approach that dynamically adjusts observation and prediction horizons to optimize predictive performance across various aging stages. The proposed method employs principal component analysis (PCA) for dimensionality reduction on publicly available NASA and Mendeley battery datasets to extract health indicators (HIs) and applies K-means clustering to segment the battery lifecycle into three aging stages (run-in, linear aging, and nonlinear aging), providing aging-stage-based input features for the model. Experimental results show that, in the NASA dataset, the adaptive LSTM reduces the MAE and RMSE by 0.042 and 0.043, respectively, compared to the CNN, demonstrating its effectiveness in mitigating error accumulation during the nonlinear aging stage. However, in the Mendeley dataset, the average prediction accuracy of the adaptive LSTM is slightly lower than that of the CNN and Transformer. These findings indicate that defining aging-stage-based adaptive observation and prediction horizons for LSTM can effectively enhance its performance in predicting battery RUL across the entire lifecycle.

Keywords: RUL prediction, Li-Ion battery, Battery degradation mechanism, Adaptive LSTM

1. Introduction

With the rapid development of green energy and low-carbon technologies, LIBs have become core components of energy storage systems (ESSs) [1-3]. However, prolonged usage leads to the gradual aging of LIBs due to external factors such as temperature, discharge rate, electrochemical reactions, and physical and chemical

changes in materials [4-6]. If not replaced in time before failure, these aging processes can result in abnormal device operation or even severe safety incidents. Consequently, accurately predicting the RUL of LIBs is crucial for ensuring the stability of systems and the reliability of devices.

RUL prediction methods for LIBs can generally be divided into model-based and data-driven approaches. Model-based methods are developed based on empirical and mathematical models that explain the physical and chemical degradation processes and the underlying mechanisms of LIBs. Zhang et al. [7] proposed an RUL prediction method that leverages an exponential model and PF, addressing LIB capacity degradation's nonlinear and non-Gaussian nature. The paper details the prediction error across various prediction starting points. When compared to other methods, including the auto-regressive integrated moving average (ARIMA) model, the fusion nonlinear degradation auto regressive model, and the regularized particle filter (RPF) algorithm, the proposed method shows superior prediction performance with lower, root mean squared error (RMSE). To overcome the limitations of PF, that is, sample degeneracy and impoverishment, Li et al. [8] explored particle distribution optimization (PDO) and ensured particle diversity. M. Ahwiadi and W. Wang [9] have proposed an enhanced mutated particle filter (EMPF) technique. The effectiveness of the proposed EMPF technique is demonstrated through simulation tests and its application in predicting the RUL of LIB using the battery data base at the NASA PCoE. Due to the complex internal mechanisms of lithium-ion batteries, building precise degradation models is highly challenging [6, 7].

Data-driven methods offer enhanced flexibility and accuracy by extracting key health indicators—such as voltage, current, resistance and capacity—that reflect the degradation trends of batteries during operation. These indicators serve as inputs for intelligent algorithms used in learning and predictive analysis [6]. Selina et al. [10] proposed a naive Bayes model for predicting the RUL of Lithium-ion batteries under constant operating conditions. The paper demonstrates that under constant discharge environments, the NB method can accurately predict the RUL of LIB, regardless of specific operating condition values. Zhang et al. [11] presented an LSTM-RNN model tailored to capture capacity degradation patterns in lithium-ion batteries. This model effectively captures long-term dependencies in the degradation process, enabling highly precise RUL prediction. Lei et al. [12] proposed a deep learning framework combining autoencoders with DNNs to predict the RUL of LIBs. This method employs autoencoders for multidimensional feature extraction to characterize battery degradation, followed by DNN training for RUL prediction. Validation on the NASA PCoE lithium-ion battery dataset for battery B7 showed that this approach improved prediction accuracy by 4-5% compared to Bayesian regression, linear regression and SVM. Wang et al. [13] introduced an ensemble approach incorporating local tangent space alignment (LTSA) for feature extraction and adaptive sliding window long short-term memory (ASW LSTM) to enhance RUL estimation accuracy. LTSA automatically extracts health indicators with a high Spearman correlation to unmeasured battery capacity, while the ASW-LSTM model applies these HIs in RUL estimation under standard conditions. By dynamically adjusting inputs with a variable-length sliding horizon, this model captures both long-term dependencies and local variations.

Although the aforementioned techniques have provided valuable insights and advances for LIBs RUL prediction, many challenges still remain to be addressed. Due to the nonlinear aging process and variable operating conditions of LIBs, most data-driven methods typically require extensive charge–discharge cycling data collected under multi-condition and long-term scenarios. For instance, when the prediction starting point exceeds half of the cycle life, the methods described in [14-17] can achieve relatively satisfactory prediction results. However, during the early stages of battery cycling, RUL predictions tend to exhibit a wide prediction range, which increases the negative impact of cumulative errors; in the later stages, if the same observation and prediction horizons used in the early stages are applied, the rapid decay characteristics of the RUL in the battery's more pronounced nonlinear aging phase cannot be effectively captured, leading to a significant decline in prediction accuracy. Therefore, achieving high-precision full life-cycle RUL prediction under limited data conditions has become a core issue in both theoretical research and practical engineering applications [18-24]. Moreover, in response to these challenges, recent studies have incorporated adaptive mechanisms into LSTM, primarily focusing on the dynamic optimization of parameters such as the learning rate or weights [25,26]. For example, Chong et al. [25] accelerated and stabilized network training by adaptively adjusting the learning rate. Prajith Pillai et al. [26] dynamically updated weights based on real-time data variations to enhance model performance across different aging stages. However, these approaches still encounter considerable challenges when applied to longer lifetimes and more complex usage scenarios.

To address these challenges, this paper proposes an adaptive LSTM-based RUL prediction method for LIBs. This method integrates the aging mechanisms of LIBs and dynamically adjusts both the observation and prediction horizons, providing a more flexible and accurate solution for RUL prediction.

The primary contributions of this work are summarized as follows:

- Dimensionality reduction of features using PCA to construct battery HIs. Based on HIs and K-means clustering, the battery's lifecycle is divided into three distinct aging stages.
- Different observation and prediction horizons are defined for each aging stage, thus creating an adaptive LSTM that enhances the accuracy of RUL prediction.
- Validated on NASA and Mendeley datasets and compared with general LSTM, CNN, and Transformer that have fixed observation and prediction horizons.

The structure of this paper is as follows: Section 2 describes the design and training of the adaptive LSTM. Section 3 introduces the datasets and data processing. Section 4 presents the aging stage classification. Section 5 provides experimental results and analysis. Section 6 introduces the ablation study. Finally, the paper concludes with the main findings and discusses future research directions.

2. Methodology

2.1. Theory introduction of LSTM

LSTM is a variant of recurrent neural networks (RNNs) that is widely used for time series forecasting. The memory cell is the core component of LSTM architecture, allowing the network to selectively store and retrieve information across different time steps. The structure of an LSTM cell is shown in Figure 1.

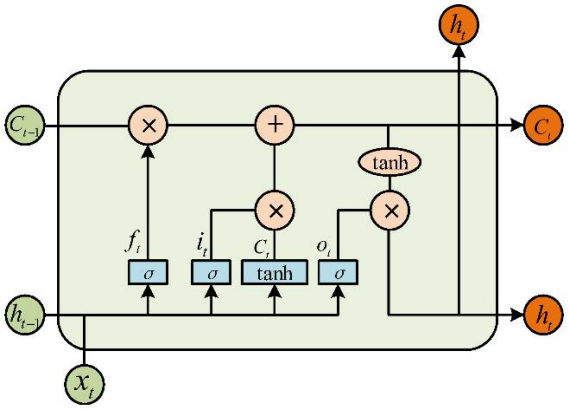


Fig. 1. Structure of LSTM cell [27].

In LSTM networks, the forget gate, input gate and output gate work together to manage the storage, retrieval and forgetting of information. The detailed calculation process is as follows (1) - (6).

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

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (3)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (4)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t * \tanh(C_t) \quad (6)$$

where x_t is the input, h_t is the output, i, o and f represent the input gate, output gate, and forget gate. C_t denotes the cell state. σ and *tanh* are activation functions, with W indicating the weight matrix and b representing the bias vector.

2.2. Definition of adaptive observation and prediction horizons in LSTM

In traditional LSTM for RUL prediction, a fixed-length observation horizon for input and a fixed-length prediction horizon for output is typically employed [11, 13]. However, due to the pronounced nonlinear characteristics of battery degradation—especially as degradation accelerates in later stages—a fixed-horizon LSTM struggle to adapt to this complexity, bringing reduced prediction accuracy. To address this, this paper proposes an adaptive LSTM that incorporates two novel model parameters: the adaptive observation horizon and the adaptive prediction horizon. These parameters can be adjusted independently, allowing the model to generate input and output of varying sizes based on specific parameter configurations. The definition of these two parameters and the implementation of adaptive LSTM will be explained in the following.

Adaptive observation horizon: This parameter controls the input size of the LSTM, with its length dynamically adjustable according to the requirements of different aging stages. Given the varying degradation rates across stages, selecting an appropriate observation horizon helps to maximize prediction accuracy. In the run-in and linear aging, where degradation is gradual, a longer observation horizon captures long-term trends. In contrast, in the nonlinear aging, where degradation accelerates, a shorter observation horizon enables the model to respond quickly to changes in battery health.

Adaptive prediction horizon: This parameter controls the output size of the LSTM. Given the varying dynamics of different aging stages, this parameter should correspond to the specific aging phase. For instance, shorter prediction horizons are more suitable for nonlinear aging, whereas longer horizons are preferable during the linear aging. The horizon length for the run-in may be positioned between these two extremes.

Figure 2 illustrates a schematic of the adaptive LSTM technique with adaptive observation and prediction horizons. After processing, the dataset is scaled and divided into three clusters using PCA and K-means, where cluster 0 represents the run-in stage, cluster 1 indicates the linear aging stage, and cluster 2 is the nonlinear aging stage. Observation Lo and prediction Lp horizons are independently configured according to the three aging stages, indexed as 0, 1 and 2. Additionally, clustering the data provides sufficient data points for each stage, effectively capturing both localized and overall fluctuations. The clustered samples are used as input data for the LSTM to train and validate the model. Compared to a general LSTM with only two degrees of freedom, this model offers six degrees of freedom, providing greater flexibility and superior predictive performance than the traditional fixed-horizon LSTM.

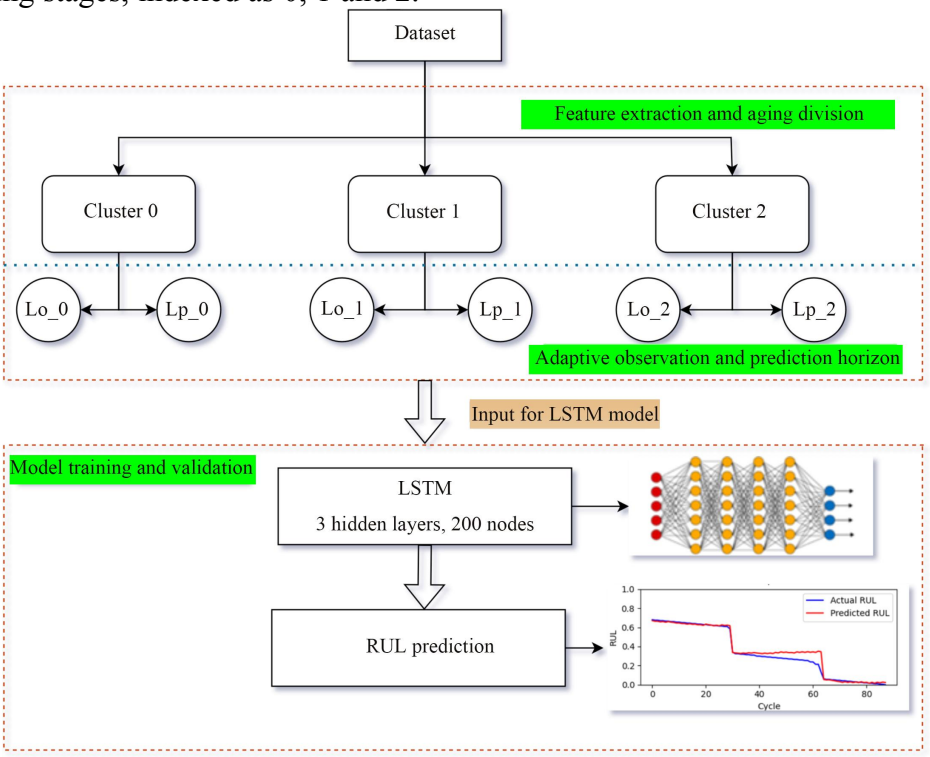


Fig. 2. A schematic of the constructed LSTM structure.

3. Datasets and data preprocessing

3.1. Introduction to datasets

This study utilizes two publicly available lithium-ion battery datasets: the NASA PCoE dataset [28] and the Mendeley dataset [29]. These datasets provide comprehensive data to support battery health assessment and RUL prediction. The NASA PCoE dataset includes multicycle test data for three representative batteries (Nos. 33, 34 and 36) tested at a controlled room temperature of 24°C. These batteries were tested under three distinct operational modes: charging, discharging, and impedance measurement. The batteries gradually aged through repeated charge and discharge cycles, and impedance measurements were used to analyze internal parameters evolving with aging. Testing was terminated upon the battery capacity reaching 70% of its nominal 2 Ah.

The Mendeley dataset is derived from nominally identical high-energy 18650 lithium-ion batteries, each with a rated capacity of 2.4 Ah and a nominal voltage of 3.7 V. This experiment consists of two stages. In the initial stage, 20 preliminary cycles were used to simulate the primary usage of the batteries and facilitate early battery life prediction. These preliminary cycles involved charging using a constant-current-constant-voltage (CC-CV) method at 0.5C and discharging at a constant current of 2C. In the subsequent stage, the focus was on studying degradation characteristics under different operating conditions. The 77 batteries were divided into Group I and Group II for further cyclic degradation tests. Group I, consisting of 22 batteries, underwent cyclic degradation under fixed charge and discharge currents (1C, 2C or 3C). Group II, comprising 55 batteries, was subjected to cyclic degradation under varying operating conditions. In this group, the charge current was randomly changed every five cycles, selected uniformly among 1C, 2C and 3C, while the discharge current was consistently set at 3C. All batteries recorded 101 cycles in the second stage.

3.2. Data preprocessing

To meet the requirements of the LSTM network for time series data, this study standardized the NASA and Mendeley battery datasets and converted them into a cycle-based format to more accurately characterize the degradation trend of the batteries. Firstly, the StandardScaler method was used to standardize the data by removing the mean and scaling to a unit standard deviation, so as to eliminate the influence of different feature scales on model training and improve the convergence speed and prediction accuracy of the model. At the same time, to reduce the computational complexity, irrelevant parameters such as test time, date time, and entry points were removed, and only key features such as charging and discharging voltages, currents, and capacities were retained to ensure the consistency and usability of the data.

4. Aging phase identification

4.1. Health index construction based on PCA

Since the RUL is only obtainable after the battery has fully aged, it cannot be directly used for classifying the aging stages of the battery. To address this issue, this study constructs HIs by applying PCA to the raw feature parameters of the battery (such as voltage, current, capacity, etc.), which can effectively replace RUL for classifying the battery's aging stages.

Figure 3 illustrates the complete process from data processing to the classification of aging stages, including PCA for dimensionality reduction, construction of the HI, and the final classification of aging stages through steps such as K-means clustering and outlier removal. The goal of PCA is to extract the most representative principal components from the battery data set to reduce the dimensionality of the data features. Each principal component extracted by PCA corresponds to a different HI, reflecting distinct degradation characteristics of the battery. Specifically, PC_1 corresponds to HI_1, which is primarily driven by the battery's capacity and voltage; PC_2 corresponds to HI_2, which is closely related to current and temperature; and PC_3 corresponds to HI_3, which reflects the changes in temperature and energy.

To validate whether HIs can effectively capture degradation characteristics related to RUL, this study further conducted Pearson correlation analysis. For battery B0036 in the NASA dataset, the correlation analysis results are as follows:

- HI_1 and RUL: $\rho=-0.9941$ ($p = 0.0000$)
- HI_2 and RUL: $\rho=0.0032$ ($p = 0.9747$)
- HI_3 and RUL: $\rho=0.0883$ ($p = 0.3797$)

The results show a significant negative correlation between HI_1 and RUL, indicating that HI_1 effectively reflects the degradation trend driven by battery capacity and voltage. In contrast, HI_2 and HI_3 show weaker correlations with RUL, with larger corresponding p-values, suggesting that changes in current, temperature, and energy have a smaller direct impact on RUL prediction. These results demonstrate that classifying the battery's health stages based on HI_1 and HI_2 is reasonable, and that HIs can effectively replace RUL for classifying aging stages.

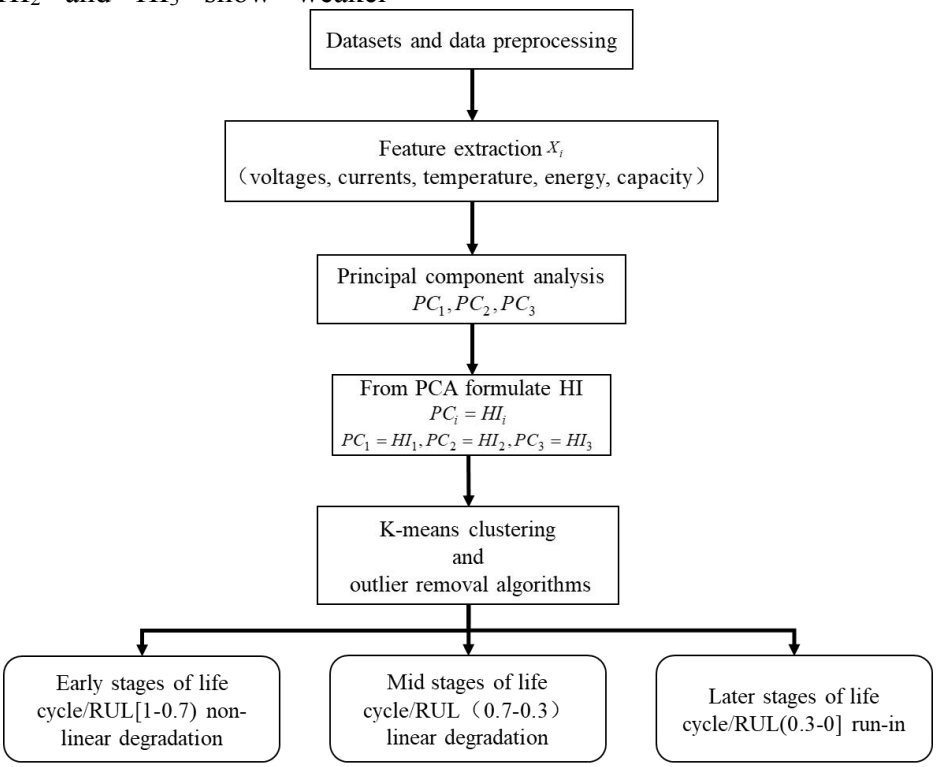


Fig. 3. Aging stage classification flowchart.

4.2. Aging phase division with K-means

The aging process of LIBs involves complex physicochemical changes, leading to a gradual decline in state of health over time. As shown in Figure 4, this process can be categorized into three main stages: run-in, linear aging, and nonlinear aging [34]. Each stage represents different aging mechanisms that the battery undergoes during its lifespan.

The K-means clustering algorithm was employed to classify HIs extracted via PCA, facilitating the identification of distinct battery aging stages. As illustrated in Figure 5, the battery life cycle data were divided into three clusters, each corresponding to a specific aging stage.

- Run-in stage: Corresponding to the run-in RUL changes within [1.0, 0.7), where degradation is gradual, performance

remains stable, and changes in health indicators are minimal.

- Linear aging stage: Corresponding to the linear aging RUL changes within [0.7, 0.3), where the health status declines rapidly, degradation accelerates, health indicators fluctuate significantly, and substantial changes occur within the battery's internal structure.
- Nonlinear aging stage: Corresponding to the nonlinear aging RUL changes within [0.3, 0.0), where the battery's health is near the failure point, degradation accelerates further, and health indicators exhibit a sharp decline.

Through K-means clustering, the characteristic changes at each life stage of the battery can be more precisely captured. Based on the features of each stage, the observation and prediction horizons of the adaptive LSTM are dynamically adjusted, thereby enhancing predictive accuracy in nonlinear aging phases.

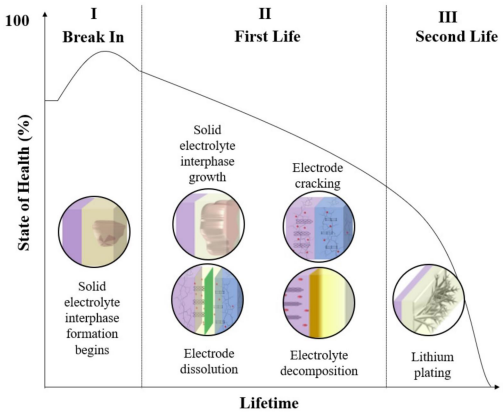
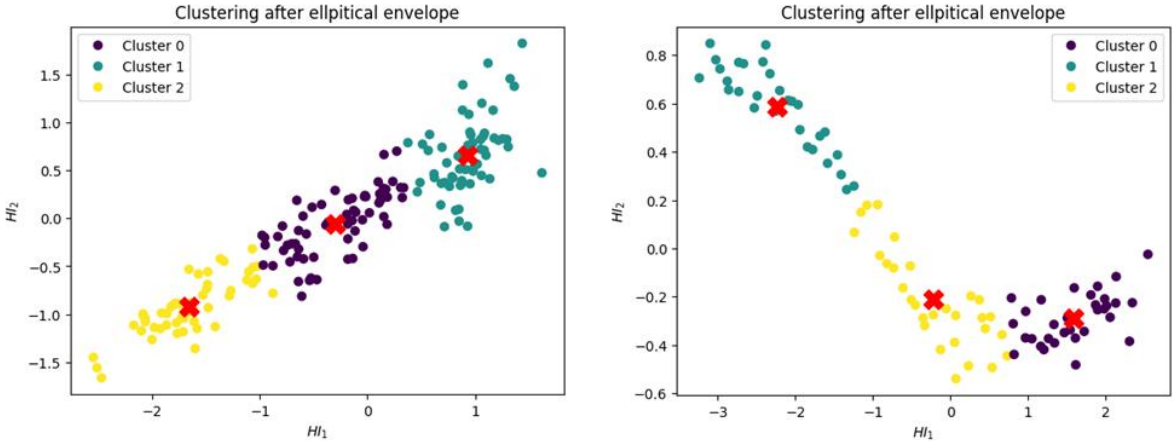


Fig. 4. Typical evolution of the state of health and aging mechanisms during battery lifetime [34].



(a) (b)

Fig. 5. K-means clustering after elliptical envelop method. (**a**) Battery B0036 from NASA dataset. (**b**) Battery 35 from Mendeley dataset.

5. Experiments and results

5.1. Model training and evaluation

The adaptive LSTM was trained on 70% of the data, with 20% used for testing and 10% for validation. Table 1 provides the segmentation of different battery datasets used for training, testing and validation. Table 2 lists the hyperparameters assigned to both the adaptive LSTM and other models, including the general LSTM, CNN, and Transformer. A model with three hidden layers was constructed, with a learning rate set to 0.001, 200 units in each hidden layer, an Adam optimizer, mean squared error for training loss evaluation, and a batch size of 16. The feature dimensions depend on the battery data. For the adaptive LSTM, the

observation range can be selected from $Lo \in [1,10]$ and the prediction range from $Lp \in [1,8]$ (with (Lo) and (Lp) both being integers). In contrast, other models use fixed values for observation and prediction. Additionally, an early stopping callback is integrated, terminating training if validation loss shows no improvement within 50 epochs.

The model evaluates the stability and accuracy of battery RUL predictions using two standard error metrics: mean absolute error (MAE) and root mean squared error (RMSE). Equation (7) presents the calculation for MAE, Equation (8) presents the calculation for RMSE.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{8}$$

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

Table 1. NASA and Mendeley battery datasets.

Dataset	Cluster	Training cycles	Testing cycles	Validation cycles
NASA	Cluster 0	50	14	7
	Cluster 1	72	22	11
	Cluster 2	14	4	2
Mendeley	Cluster 0	24	7	4
	Cluster 1	20	5	2
	Cluster 2	20	6	3

Table 2. Hyperparameters for model training and testing.

Models	Adaptive LSTM	General LSTM	CNN	Transformer
Batch size	16	16	16	16
Learning rate	0.001	0.001	0.001	0.001
Hidden layers	3	3	3	2
Epochs	100	100	100	100
Feature dimension	4	4	4	4
Neuronal units	200	200	64 and 100	32 (head size)
Observation horizon	1-10	Single value	Single value	Single value
Prediction horizon	1-8	Single value	Single value	Single value

5.2. RUL prediction for battery using adaptive LSTM

The input for the adaptive LSTM includes the formulated HI and a newly created RUL associated with each HI. The created RUL values facilitate the implementation of supervised learning. Based on the values of the adaptive observation and prediction horizons, the model predicts new RUL values. Specifically, assuming (Lo) and (Lp) are set to 4 and 2, respectively, the model takes the first four values of the processed data frame in the HI series as the input array and predicts the next two RUL values, which correspond to the fifth and sixth values in the RUL series. The RUL is defined as the number of cycles remaining until the battery's capacity decreases to 70%-80% of its initial value, which is typically considered the end of useful life for a battery. For example, in the case of the B0036

battery from the NASA dataset, the $Cycle_{\max}$ is 194 cycles, while for batteries in the Mendeley dataset, $Cycle_{\max}$ is 101 cycles. The RUL is calculated as shown in Equation (9):

$$RUL(i) = 1 - \frac{Cycle_i}{Cycle_{\max}} \tag{9}$$

where $Cycle_i$ represents the current cycle and $Cycle_{\max}$ denotes the maximum cycle life of the battery.

The adaptive LSTM can leverage adaptive observation and prediction horizons to forecast aging patterns for each cluster. This allows the model to test various sets of horizon values for each cluster and validate their accuracy. Consequently, the horizon values that yield optimal prediction results for a specific stage are adopted. By applying this process across all clusters, predictions with minimal average RUL prediction error can be achieved. Different values for Lo from $[1, 10]$ and for Lp from $[1, 8]$ were selected and the battery was tested in controlled experiments. For each combination of Lo and Lp, the RUL prediction that produced the best results was chosen.

Figure 6 illustrates the RUL predictions of batteries in various clusters from the NASA and Mendeley datasets using the adaptive LSTM method. It shows the relationship between predicted and actual RULs over cycle counts. The adaptive LSTM demonstrates strong RUL prediction performance across different batteries and clusters. Notably, for Clusters 0 and 1 of Battery 35, the predicted values closely match the actual values, with a MAE below 0.03 and a RMSE below 0.05, indicating high prediction accuracy. However, Figure 6(b) and Figure 6(f) reveal significant prediction deviations. Figure 6(b) shows the RUL predictions for Cluster 1 of B0036 battery from the NASA dataset. Predictions are accurate before the 15th cycle, but deviate notably after, especially for higher cycle values (e.g., around cycle 100), likely due to the model's sensitivity to nonlinear degradation behaviors when a larger prediction window is set. This cluster may be in a nonlinear degradation phase, warranting further study for better modeling. Similarly, Figure 6(f) presents the RUL predictions for Cluster 2 of Battery 35 from the Mendeley dataset. After 14 cycles, the predicted RUL diverges from the actual values, suggesting a nonlinear degradation stage for this cluster. Thus, additional model or data adjustments may be needed to improve predictions in nonlinear regions.

In summary, the adaptive LSTM shows excellent RUL prediction performance in most cases, especially for Cluster 0 and Cluster 1 of Battery 35. However, in specific cases like Cluster 1 of Battery B0036 and Cluster 2 of Battery 35, there are deviations in later cycles, indicating the need for further model optimization or data adjustments to enhance prediction accuracy. These results highlight the adaptive LSTM's ability to improve RUL prediction accuracy by adjusting observation and prediction horizons according to battery aging stages and characteristics, which is a significant advantage over other methods.

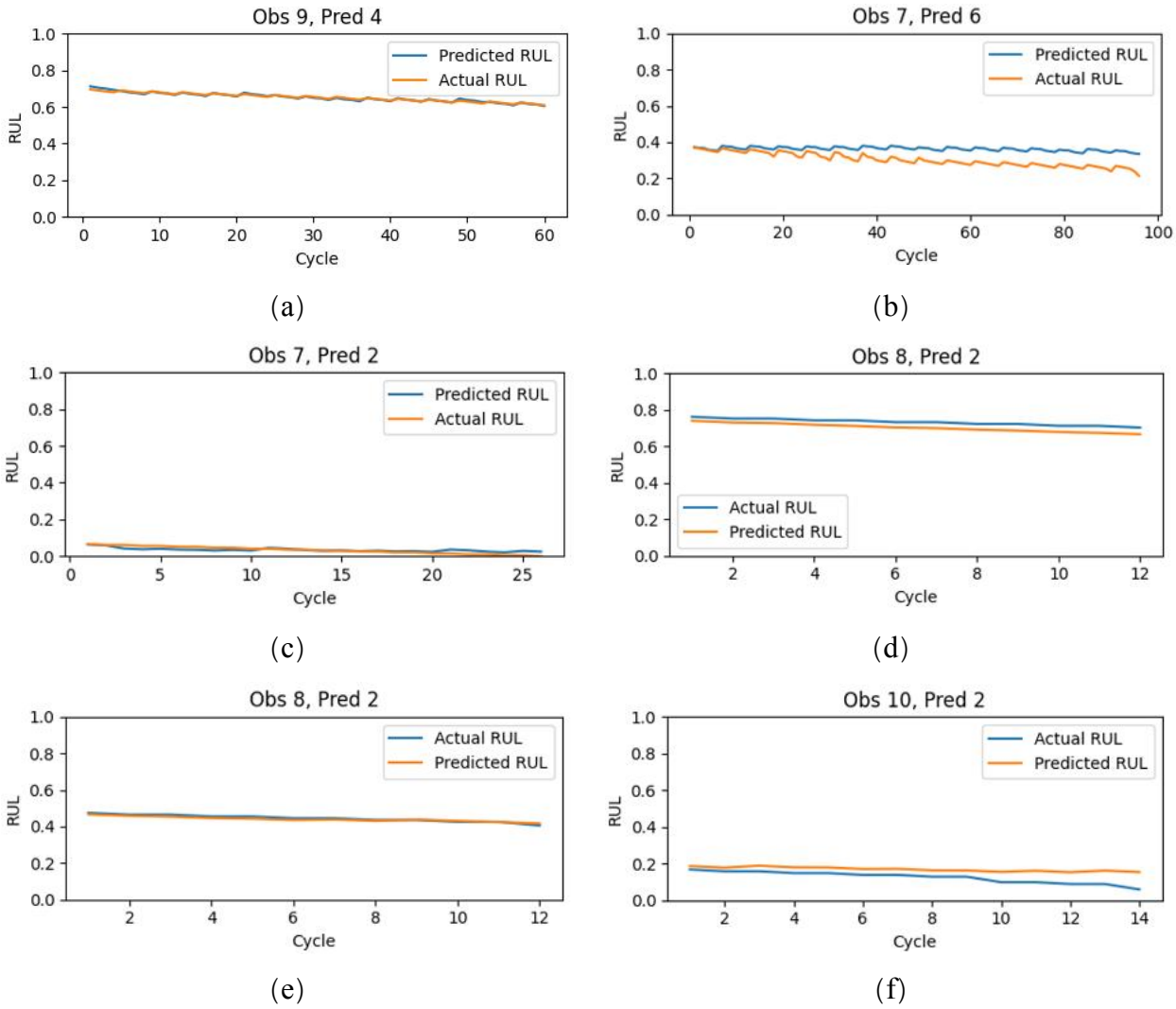


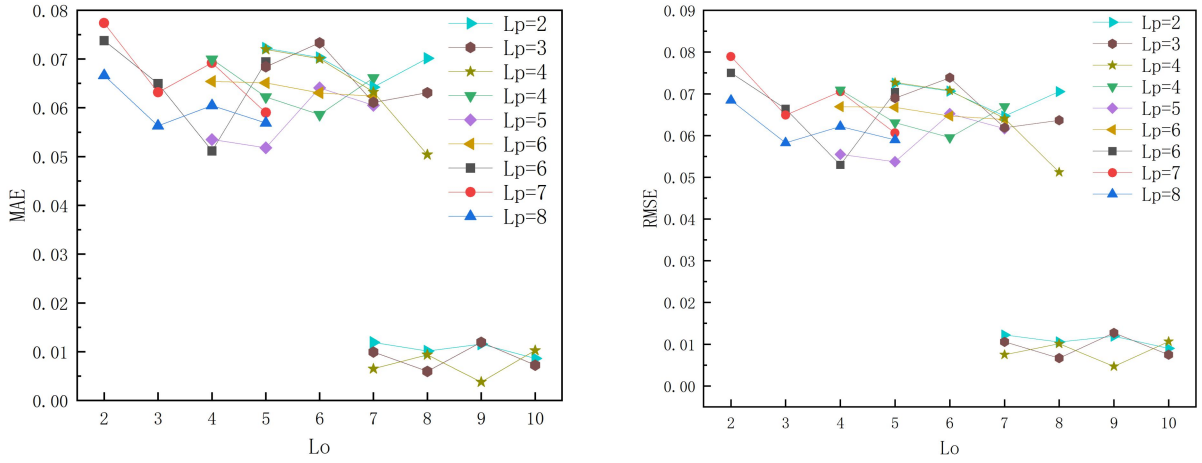
Fig. 6. RUL prediction for adaptive LSTM. (**a**) Cluster 0 for battery B0036 from MASA dataset. (**b**) Cluster 1 for battery B0036 from MASA dataset. (**c**) Cluster 2 for battery B0036 from MASA dataset. (**d**) Cluster 0 for battery 35 from Mendeley dataset. (**e**) Cluster 1 for battery 35 from Mendeley dataset. (**f**) Cluster 2 for battery 35 from Mendeley dataset.

5.3. Comparison and analysis between adaptive LSTM and other models

This section provides an in-depth analysis of the performance of adaptive LSTM versus other methods in battery RUL prediction tasks. By conducting experiments on different aging stages using NASA and Mendeley datasets, this subsection focuses on examining the impact of the Lo and Lp parameters on predictive accuracy.

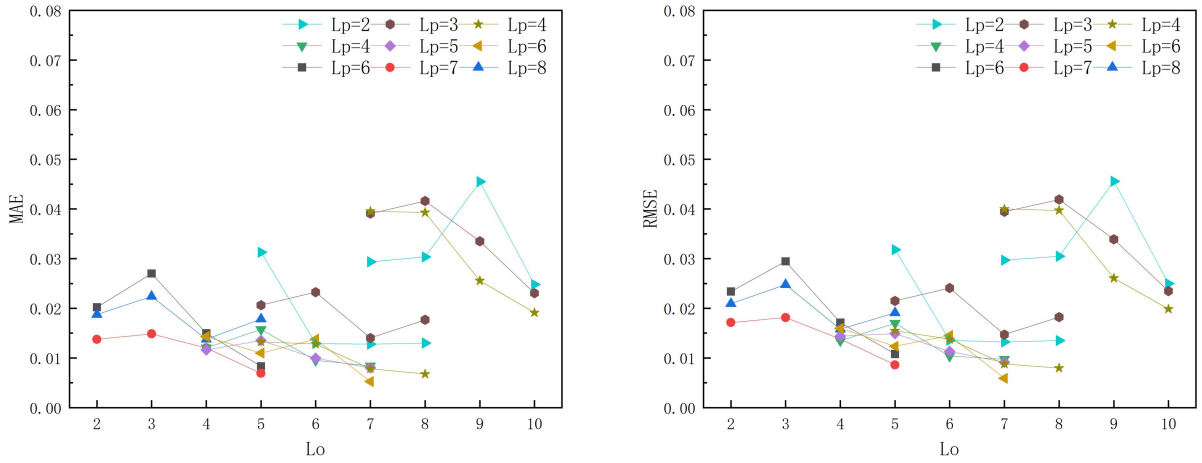
Taking battery B0036 from the NASA dataset as an example, Figure 7 shows how prediction accuracy changes under different Lo and Lp combinations in Cluster 0. When Lo ranges from 2 to 6, the overall error remains relatively high. However, once Lo exceeds 6, both MAE and RMSE decrease significantly. When Lo reaches between 8 and 10, the error values stabilize at a low level, regardless of the Lp setting. This phenomenon indicates that a smaller Lo is

insufficient for capturing the features of the initial linear aging stage, while increasing *Lo* allows the model to more comprehensively learn the linear capacity aging trend of the battery, thereby substantially improving prediction accuracy.



(a) (b)

Fig. 7. Illustration of the effects of different *Lo* and *Lp* combinations on prediction accuracy during the Cluster 0 stage. (**a**) MAE trends for different *Lo* and *Lp* combinations in Cluster 0 of battery B0036 from the NASA dataset. (**b**) RMSE trends for different *Lo* and *Lp* combinations in Cluster 0 of battery B0036 from the NASA dataset.

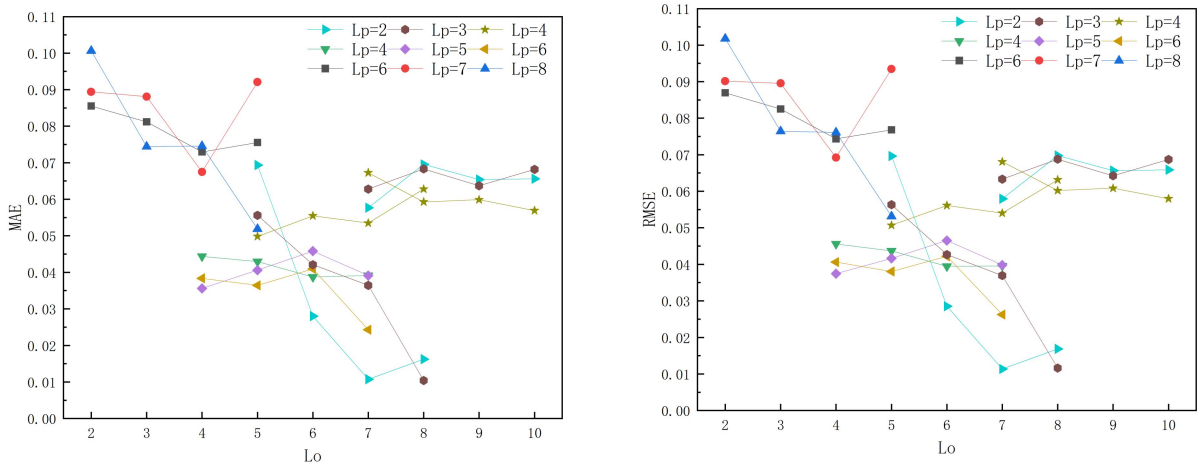


(a) (b)

Fig. 8. Illustration of the effects of different (**a**) combinations on prediction accuracy during the Cluster 1 stage. (**b**) MAE trends for different *Lo* and *Lp* combinations in Cluster 1 of battery B0036 from the NASA dataset.

Figure 8 presents the experimental results for the Cluster 1 stage. At this stage, the battery's aging trend exhibits nonlinear inflection points, making the improvement in accuracy from simply increasing *Lo* less pronounced than in Cluster 0. An *Lo* in the moderate range (approximately 5-7) can significantly reduce MAE and RMSE but

extending *Lo* beyond this range does not further decrease the errors. This is because the data in Cluster 1 contain fluctuations and nonlinear characteristics, and an overly long observation window may average out these variations, thereby reducing the model's sensitivity to recent changes and limiting its potential performance improvements. Figure 9 illustrates the results for Cluster 2 stage. At this stage, the battery aging rate and characteristics become unstable, exhibiting pronounced nonlinear behavior. Increasing *Lo* at this stage offers limited reduction in MAE and RMSE. An excessively large *Lo* incorporates extensive historical data, which does not significantly aid in capturing rapid recent changes; instead, it may impede the model's ability to identify abrupt variations. Experimental results indicate that *Lo* values within the range of 5 to 7 achieve relatively optimal prediction performance.



(a) (b)

Fig. 9. Illustration of the effects of different *Lo* and *Lp* combinations on prediction accuracy during the Cluster 2 stage. (**a**) MAE trends for different *Lo* and *Lp* combinations in Cluster 2 of battery B0036 from the NASA dataset. (**b**) RMSE trends for different *Lo* and *Lp* combinations in Cluster 2 of battery B0036 from the NASA dataset.

By comparing Figures 7, 8 and 9, it is evident that *Lp* affects prediction errors differently across various aging stages. In Cluster 0, small to medium *Lp* values exert minimal impact on prediction errors. However, excessively large *Lp* values slightly increase the error due to the uncertainty introduced by longer prediction horizons. In Cluster 1, smaller *Lp* values, such as $Lp = 2$ or $Lp = 3$, are more conducive to enhancing prediction accuracy. When *Lp* increases to 6, 7 or 8, both MAE and RMSE rise significantly, attributable to increased nonlinearity, which renders long-term predictions more uncertain. Smaller *Lp* values enable the model to more effectively track performance changes, thereby mitigating long-term prediction bias. In Cluster 2, the benefits of smaller *Lp* values are even more pronounced. Larger *Lp* values hinder the model's ability to accurately capture abrupt future changes, resulting in significantly increased errors.

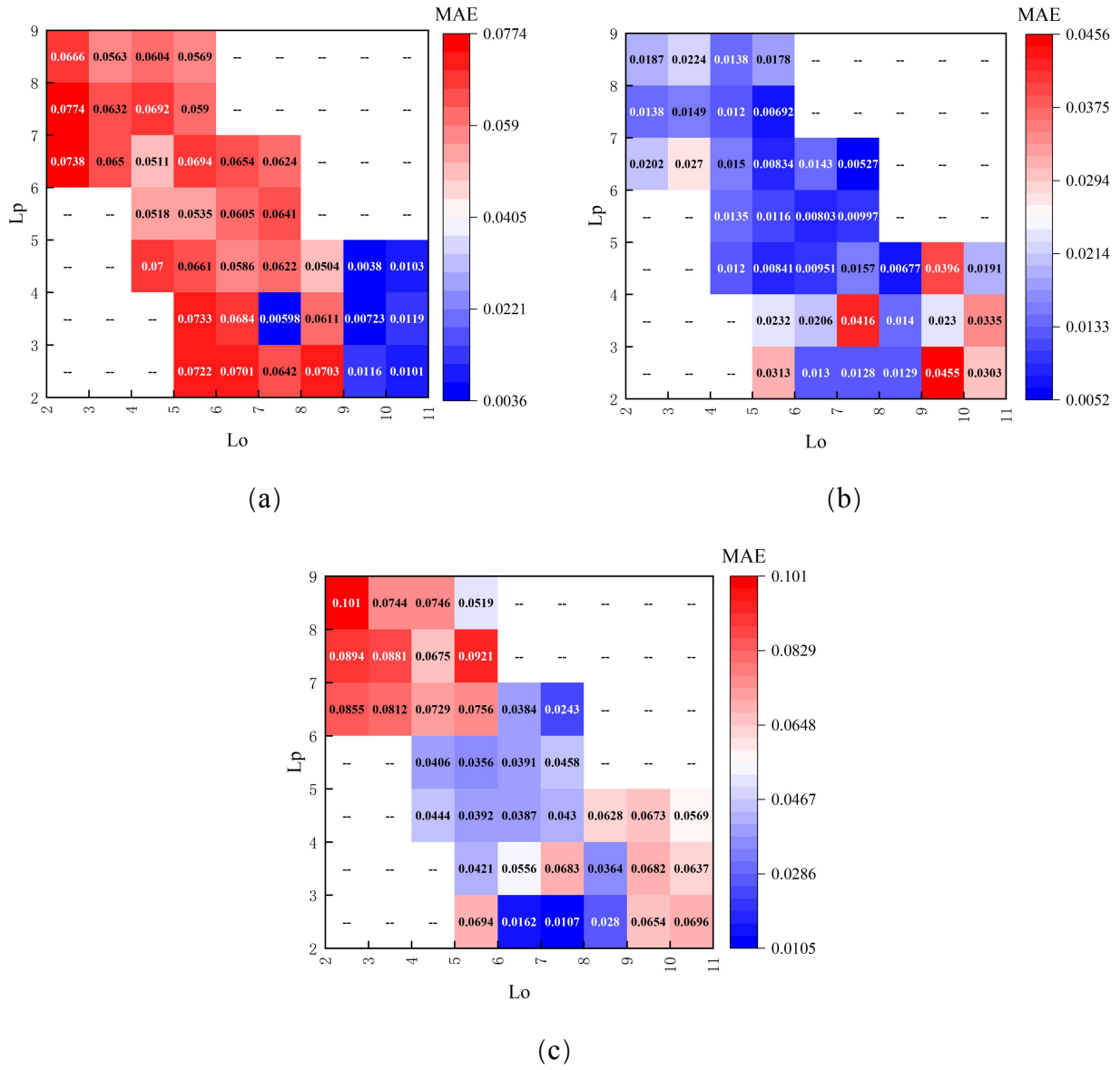


Fig. 10. Optimal parameter combinations for each aging stage in the NASA dataset for battery B0036. (**a**) Cluster 0 stage. (**b**) Cluster 1 stage. (**c**) Cluster 2 stage.

Overall, while the impact of *Lp* on performance is less substantial than that of *Lo* smaller *Lp* values are more advantageous for accurate predictions during nonlinear aging stages. Conversely, in the linear aging initial stage, medium *Lp* values also achieve relatively good performance. To further elucidate the optimal *Lo* and *Lp* combinations for each aging stage, Figure 10 presents the best parameter combinations based on MAE. The following optimal combinations were identified. In Cluster 0, the optimal combination is $Lo = 9$ and $Lp = 4$, with an error of 0.003796. In Cluster 1, the optimal combination is $Lo = 7$ and $Lp = 6$, with an error of 0.005271. In Cluster 2, the optimal combination is $Lo = 7$ and $Lp = 2$, with an error of 0.010416. These findings indicate that parameter requirements significantly differ across aging stages. Leveraging this characteristic, an adaptive

LSTM can be constructed to dynamically adjust Lo and Lp based on the current aging stage, thereby achieving superior prediction performance.

To validate its effectiveness, this study conducts a comparative analysis of the adaptive LSTM and other models with fixed parameters, including general LSTM, CNN, and Transformer, using both the NASA and Mendeley datasets. The results indicate (table 3) that in the NASA dataset, the adaptive LSTM outperforms all other models across various aging stages, particularly in the nonlinear degradation phase, where it effectively mitigates error accumulation caused by increasing nonlinearity. However, in the Mendeley dataset, the average prediction accuracy of CNN and Transformer surpasses that of the adaptive LSTM. Further analysis reveals that while the adaptive LSTM achieves the best performance in Cluster 0, its predictive capability in Cluster 1 and Cluster 2 is inferior to that of CNN and Transformer. Possible reasons for this discrepancy include the Mendeley dataset exhibiting stronger local temporal patterns, where CNN benefits from convolutional operations to capture short-term features, and Transformer leverages self-attention mechanisms to extract global patterns within short time windows, making them more effective for this dataset. Additionally, the smaller scale of the Mendeley dataset may limit the adaptive LSTM's ability to learn optimal Lo and Lp adjustments, as its dynamic tuning strategy typically requires a larger dataset for effective generalization.

Table 3. Comparison of MAE and RMSE errors for adaptive LSTM and other models on NASA and Mendeley datasets.

Type		NASA dataset		Mendeley dataset	
		MAE	RMSE	MAE	RMSE
Adaptive LSTM	Cluster 0	$Lo \mid Lp = 9,4$ 0.003796	$Lo \mid Lp = 9,4$ 0.004688	$Lo \mid Lp = 8,3$ 0.004355	$Lo \mid Lp = 8,3$ 0.005055
	Cluster 1	$Lo \mid Lp = 7,6$ 0.005271	$Lo \mid Lp = 7,6$ 0.005915	$Lo \mid Lp = 8,2$ 0.030128	$Lo \mid Lp = 8,2$ 0.030177
	Cluster 2	$Lo \mid Lp = 7,2$ 0.010416	$Lo \mid Lp = 7,2$ 0.011352	$Lo \mid Lp = 10,2$ 0.029218	$Lo \mid Lp = 10,2$ 0.029332
	Average error	0.019483	0.021955	0.063701	0.064564
General LSTM	Average error	$Lo \mid Lp = 10,4$ 0.072832	$Lo \mid Lp = 10,4$ 0.075085	$Lo \mid Lp = 10,3$ 0.08424	$Lo \mid Lp = 10,3$ 0.086133
CNN	Average error	$Lo \mid Lp = 8,4$ 0.060709	$Lo \mid Lp = 8,4$ 0.064199	$Lo \mid Lp = 8,2$ 0.026112	$Lo \mid Lp = 8,2$ 0.032015
Transformer	Average error	$Lo \mid Lp = 4,6$ 0.069821	$Lo \mid Lp = 4,6$ 0.076491	$Lo \mid Lp = 10,3$ 0.048651	$Lo \mid Lp = 10,3$ 0.059121

6. Ablation study

To investigate the impact of various hyperparameters on model performance, an ablation experiment was conducted, systematically varying the number of LSTM layers, hidden units, and activation functions. Table 4 summarizes the experimental configurations, including the baseline model and several variations. Table 5 reports the MAE and RMSE values for each experimental configuration. Based on the results, the baseline model configuration

exhibits lower error values and superior performance. Altering the number of LSTM layers does not necessarily lead to performance improvements. In ablation experiments with hidden units ranging from 50 to 250, prediction errors were minimized when the number of hidden units was between 100 and 200. Notably, the configuration employing the softmax activation function yielded lower error values compared to the baseline.

The ablation experiments indicate that 3 LSTM layers paired with 200 hidden units and the use of a relu activation function can achieve relatively high prediction accuracy. However, replacing the activation function with softmax can further enhance performance, thereby providing valuable insights for subsequent network architecture design and hyperparameter selection.

Table 4. Ablation experiment configurations.

Experiment name	LSTM layers	Hidden units	Activation function
Baseline	3	200	relu
LSTM layers	2,4,5	200	relu
Hidden units	3	50,100,150,250	relu
Activation function	3	200	tanh, sigmoid, softmax

Table 5. Ablation experiment results.

Experiment	MAE	RMSE
Baseline	0.075213	0.075697
LSTM layers_2	0.084807	0.085356
LSTM layers_4	0.111596	0.111727
LSTM layers_5	0.129514	0.129577
Hidden units_50	0.097496	0.097647
Hidden units_100	0.073599	0.074583
Hidden units_150	0.075698	0.075909
Hidden units_250	0.078008	0.078221
Activation function _ tanh	0.063745	0.064195
Activation function _ sigmoid	0.148890	0.148929
Activation function _ softmax	0.031234	0.031503

7. Conclusions and outlook

To investigate the impact of aging characteristics at different stages of LIB aging on prediction accuracy, this study proposes a RUL prediction method for LIBs based on an adaptive LSTM. A HI is constructed using PCA based on battery state monitoring data. Considering the aging characteristics of LIBs, an unsupervised clustering method is used to classify the HI, dividing the battery aging process into distinct stages: the run-in, linear aging, and nonlinear aging. Different observation and prediction horizons are defined for each stage, leading to the development of an adaptive LSTM to better capture and interpret the features of each aging stage and their influence on prediction accuracy.

Experimental results show that, in the NASA dataset, the adaptive LSTM reduces the MAE and RMSE by 0.042 and 0.043, respectively, compared to the CNN, demonstrating its effectiveness in mitigating error accumulation during the nonlinear aging stage. However, in the Mendeley

dataset, the average prediction accuracy of the adaptive LSTM is slightly lower than that of the CNN and Transformer models.

The adaptive LSTM effectively addresses the variations in LIB characteristics across different aging stages. At the same time, these findings provide deeper insights into the advantages and limitations of the adaptive LSTM. In future work, we will explore partitioning the HI into different aging stages for RUL prediction and integrate physical mechanisms with deep learning methods, as outlined in our previous work [35], to optimize the *Lo* and *Lp* parameters for enhanced performance across varying aging stages.

Acknowledgments

This research is supported by National Natural Science Foundation of China (Grant No. 62403475).

References

1. S. Zhou, J. Zhou, S. Zhang, P. Wang, "Optimized LSTM based on an improved sparrow search algorithm for power battery fault diagnosis in new energy vehicles," Int. J. Metrol. Qual. Eng. 15 (2024) 12.

2. S. Bamati, H. Chaoui, H. Gualous, "Enhancing battery thermal management with virtual temperature sensor using hybrid CNN-LSTM," IEEE Trans. Transp. Electrification (2024), doi:10.1109/TTE.2024.3376515.

3. S. Wang, S. Jin, D. Deng, C. Fernandez, "A critical review of online battery remaining useful lifetime prediction methods," Front. Mech. Eng. 7 (2021).

4. M.H. Lipu, M. Hannan, T.F. Karim, A. Hussain, M.H.M. Saad, A. Ayob, M.S. Miah, T.I. Mahlia, "Intelligent algorithms and control strategies for battery management system in electric vehicles: Progress, challenges and future outlook," J. Clean. Prod. 292 (2021) 126044.

5. S. Ansari, A. Ayob, M.H. Lipu, A. Hussain, M.H.M. Saad, "Remaining useful life prediction for lithium-ion battery storage system: A comprehensive review of methods, key factors, issues and future outlook," Energy Rep. 8 (2022) 12153–12185.

6. X. Hu, L. Xu, X. Lin, M. Pecht, "Battery lifetime prognostics," Joule 4(2) (2020) 310–346.

7. L. Zhang, Z. Mu, C. Sun, "Remaining useful life prediction for lithium-ion batteries based on exponential model and particle filter," IEEE Access 6 (2018) 17729–17740.

8. T. Li, S. Sun, T.P. Sattar, J.M. Corchado, "Fighting sample degeneracy and impoverishment in particle filters: A review of intelligent approaches," Expert Syst. Appl. 41(8) (2014) 3944–3954.

9. M. Ahwiadi, W. Wang, "An enhanced mutated particle filter technique for system state estimation and battery life prediction," IEEE Trans. Instrum. Meas. 68(3) (2019) 923–935.

10. S.S. Ng, Y. Xing, K.L. Tsui, "A naive bayes model for robust remaining useful life prediction of lithium-ion battery," Appl. Energy 118 (2014) 114–123.

11. Y. Zhang, R. Xiong, H. He, Z. Liu, "A LSTM-RNN method for the lithium-ion battery remaining useful life prediction," in: 2017 Prognostics and System Health Management Conference (PHM-Harbin), 2017, pp. 1–4.

12. L. Ren, L. Zhao, S. Hong, S. Zhao, H. Wang, "Battery remaining useful life prediction using deep learning and feature extraction," IEEE Trans. Transp. Electrification 5(3) (2018) 639–648.

13. Z. Wang, N. Liu, Y. Guo, "Adaptive sliding window LSTM NN based RUL prediction for lithium-ion batteries integrating LTSA feature reconstruction," Neurocomputing 466 (2021) 178–189.

14. F.K. Wang, Z.E. Amogne, J.H. Chou, et al., "Online remaining useful life prediction of lithium-ion batteries using bidirectional long short-term memory with attention mechanism," Energy (2022) 124344.

15. S. Li, H. Fang, B. Shi, "Remaining useful life estimation of lithium-ion battery based on interacting multiple model particle filter and support vector regression," Reliab Eng Syst Saf 210 (2021) 107542.

16. Y. Ma, M. Yao, H. Liu, et al., "State of Health estimation and Remaining Useful Life prediction for lithium-ion batteries by Improved Particle Swarm Optimization-Back Propagation Neural Network," J Energy Storage 52 (2022) 104750.

17. S. Wang, Y. Fan, S. Jin, et al., "Improved anti-noise adaptive long short-term memory neural network modeling for the robust remaining useful life prediction of lithium-ion batteries," Reliab Eng Syst Saf 230 (2023) 108920.

18. L. Yan, J. Peng, D. Gao, et al., "A hybrid method with cascaded structure for early-stage remaining useful life prediction of lithium-ion battery," Energy 243 (2022) 123038.

19. K.A. Severson, P.M. Attia, N. Jin, et al., "Data-driven prediction of battery cycle life before capacity degradation," Nat Energy 4(5) (2019) 383–391.

20. S.S. Afshari, S. Cui, X. Xu, et al., "Remaining useful life early prediction of batteries based on the differential voltage and differential capacity curves," IEEE Trans. Instrum. Meas. 71 (2021) 1–9.

21. F. Xu, F. Yang, Z. Fei, et al., "Life prediction of lithium-ion batteries based on stacked denoising autoencoders," Reliab Eng Syst Saf 208 (2021) 107396.

22. K. Liu, Y. Shang, Q. Ouyang, et al., "A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery," IEEE Trans. Ind Electron 68(4) (2021) 3170–3180.

23. S. Saxena, L. Ward, J. Kubal, et al., "A convolutional neural network model for battery capacity fade curve prediction using early life data," J Power Sources 542 (2022) 231736.

24. C.W. Hsu, R. Xiong, N.Y. Chen, et al., "Deep neural network battery life and voltage prediction by using data of one cycle only," Appl Energy 306 (2022) 118134.

25. Bian, C., He, H. L., & Yang, S. K. “Stacked bidirectional long short-term memory networks for state-of-charge estimation of lithium-ion batteries,” *Energy* 191 (2020) 116538.

26. Pillai, P., Pal, P., Chacko, R., Jain, D., & Rai, B, “Leveraging long short-term memory (LSTM)-based neural networks for modeling structure–property relationships of metamaterials from electromagnetic responses,” Sci. Rep. 11(1) (2021) 18629.

27. Q. Bao, W. Qin, Z. Yun, "A multi-stage adaptive method for remaining useful life prediction of lithium-ion batteries based on swarm intelligence optimization," Batteries 9(4) (2023) 224.

28. B. Saha, K. Goebel, "Battery data set," NASA AMES prognostics data repository, 2007.

29. J. Lu, R. Xiong, J. Tian, C. Wang, C.-W. Hsu, N.-T. Tsou, F. Sun, J. Li, "Battery degradation dataset (fixed current profiles and arbitrary uses profiles)," 2022–, [Online; accessed].

30. Z. Liu, H. Hongwen, "Model-based sensor fault diagnosis of a lithium-ion battery in electric vehicles," Energies 8 (2015) 6509–6527.

31. J. Lu, R. Xiong, J. Tian, C. Wang, C.-W. Hsu, N.-T. Tsou, F. Sun, J. Li, "Battery degradation prediction against uncertain future conditions with recurrent neural network enabled deep learning," Energy Storage Mater. 50 (2022) 139–151.

32. J. Tian, R. Xiong, W. Shen, J. Lu, X.-G. Yang, "Deep neural network battery charging curve prediction using 30 points collected in 10 min," Joule 5(6) (2021) 1521–1534.

33. J.A.M. Penna, C.L. Nascimento, L.R. Rodrigues, "Health monitoring and remaining useful life estimation of lithium-ion aeronautical batteries," in 2012 IEEE Aerospace Conference, 2012, pp. 1–12.

34. M.A. Ali, C.M. Da Silva, C.H. Amon, "Multiscale modelling methodologies of lithium-ion battery aging: A review of most recent developments," Batteries 9(9) (2023).

35. D. Ruan, J. Wang, J. Yan, C. Gühmann, "CNN parameter design based on fault signal analysis and its application in bearing fault diagnosis," Adv. Eng. Informatics 55 (2023) 101877.