

Research Article

A hybrid deep learning model for robust and data-efficient lithium-ion battery remaining useful life prediction

Mengkang Xu^{1,2}, Shihao Xing³, Jianhe Hong³, Xinpeng Tian⁴, Boyuan Huang^{1,2}, Hongyun Jin^{3,*}, Jiangyu Li^{1,2,*}

¹Department of Materials Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, Guangdong, China.

²Guangdong Provincial Key Laboratory of Functional Oxide Materials and Devices, Southern University of Science and Technology, Shenzhen 518055, Guangdong, China.

³Engineering Research Center of Nano-Geomaterials of Ministry of Education, Faculty of Materials Science and Chemistry, China University of Geosciences, Wuhan 430074, Hubei, China.

⁴Department of Engineering Mechanics, Xi'an University of Technology, Xi'an 710048, Shaanxi, China.

***Correspondence to:** Prof. Jiangyu Li, Department of Materials Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, Guangdong, China. E-mail: lijy@sustech.edu.cn; Prof. Hongyun Jin, Engineering Research Center of Nano-Geomaterials of Ministry of Education, Faculty of Materials Science and Chemistry, China University of Geosciences, Wuhan 430074, Hubei, China. E-mail: jinhongyun@cug.edu.cn

How to cite this article: Xu, M.; Xing, S.; Hong, J.; Tian, X.; Huang, B.; Jin, H.; Li, J. A hybrid deep learning model for robust and data-efficient lithium-ion battery remaining useful life prediction. *J. Mater. Inf.* 2025;5:[Accepted]. <https://dx.doi.org/10.20517/jmi.2025.80>

Received: 16 September 2025 | **Revised:** 9 December 2025 | **Accepted:** 18 December 2025

Abstract

Lithium-ion batteries are extensively utilized in applications like new energy vehicles and aerospace owing to their high energy density and safety, but their service life diminishes due to irreversible capacity degradation from repeated charge-discharge cycles, making accurate remaining useful life (RUL) prediction critical for reliability and operational safety. Current data-driven methods often struggle with long-range dependencies, noise from capacity regeneration, and efficient data utilization. To address these challenges, this study introduces a novel hybrid neural network architecture that integrates Complementary Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) for denoising preprocessing with a Transformer-Convolutional Neural Network (CNN)- Bi-directional Gated Recurrent Unit (BiGRU) model. The framework employs health indicators extracted from voltage and current profiles as inputs, where CEEMDAN mitigates interference effects, the Transformer captures global degradation trends via self-attention mechanisms, the CNN extracts localized short-term features, and the bidirectional GRU models temporal dependencies bidirectionally. Experimental validation on NASA and our own test datasets demonstrates that the proposed approach significantly outperforms other models in key metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), achieving high accuracy even with minimal training data (e.g., only 40% of cycles). Furthermore, cross-dataset validation demonstrates robust generalization of our model, achieving MAPE below 3.5% when transferring between NASA and our battery data without retraining. This hybrid model offers a robust, data-efficient solution for enhancing RUL prediction in practical battery management systems, with strong generalization across diverse battery types.

Keywords: Lithium-ion battery, deep learning model, RUL prediction, CEEMDAN denoising, Bi-GRU, transformer

INTRODUCTION

Owing to their long storage life, pollution-free, high-energy density, and light weight, lithium-ion batteries find widespread applications across various domains, including new energy vehicles, aerospace, and mobile communication systems^[1-3]. However, as

the number of charge-discharge cycles increases progressively, irreversible chemical reactions, such as oxidation decomposition, occur within the battery. This results in the loss of electrode active materials and a decrease in available capacity, thereby diminishing the service life of the battery^[4-7]. The remaining useful life (RUL) of a lithium-ion battery refers to the number of charge-discharge cycles that can be sustained before it is degraded from the current state to the end of life (EOL) threshold. Since RUL is a critical metric for characterizing lithium-ion battery health and performance, accurate RUL prediction is vital for enhancing reliability and operational safety^[8-10]. In this regard, current work mainly focuses on two major categories of RUL prediction approaches: model-based methods and data-driven methods.

Common model-based approaches include electrochemical models (EMs) and equivalent circuit models (ECMs), both of which aim to establish physical frameworks describing battery degradation trends^[11-14]. Among them, EMs primarily focus on the internal chemical reaction mechanisms within different battery cells^[13,15], whereas ECMs typically utilize various filtering algorithms to process resistance or capacitance data^[11,12,14]. However, due to the difficulty in acquiring precise model parameters, these approaches typically exhibit limited generalization capability and incur notably high computational costs.

From the data-driven perspective, machine learning developments demonstrate considerable promise for RUL prediction. Unlike traditional modelling approaches, data-driven methods circumvent the need to account for the internal mechanisms of lithium-ion batteries. Instead, they analyze the historical operational data of these batteries to uncover the underlying patterns of battery degradation. Among which, recent studies have reported methods for predicting RUL based on physics-informed neural networks (PINNs)^[16-19]. Such approaches incorporate physical laws governing battery degradation, thereby enabling RUL prediction with only a small amount of training data. However, these methods are typically designed for a single battery chemistry. To enable direct application across diverse battery chemistries, our focus is

on neural network methods that can accurately predict RUL using only externally measurable data. Commonly employed data-driven architectures encompass convolutional neural networks (CNN)^[20-23] and recurrent neural networks (RNN)^[24-26], among others^[27-29]. CNN, initially developed for image processing, possess robust self-adaptive feature extraction capabilities and demonstrate significant advantages in extracting localized features from data^[30,31], while RNNs are particularly effective in modeling sequential data and have found broad applications within the RUL prediction domain. However, original RNN architectures suffer from issues with long-term dependencies and gradient vanishing. To address these limitations, researchers have developed enhanced RNN model such as gated recurrent units (GRU)^[32] and long short-term memory (LSTM) networks^[33-35]. Subsequently, to more effectively capture contextual relationships across sequential data, bidirectional GRU (BiGRU) and Bidirectional LSTM (BiLSTM) networks were introduced. Recently, to address long-term dependencies limitations inherent in traditional deep learning models, researchers have introduced the Attention Mechanism (AM)^[22,36,37]. AM is a technique specifically designed for modeling dependencies within sequences. Building on this capability, researchers have developed hybrid neural network architectures that integrate AM with CNNs and RNNs^[20,22,37,38]. Furthermore, building upon the AM framework, Vaswani et al. from Google Brain proposed a novel architecture called Transformer^[39,40]. Due to its unique structure, exceptional ability to capture long-range dependencies, and efficient parallel computation, the Transformer model has achieved remarkable success in natural language processing and is steadily gaining traction in time series forecasting domains^[38,41-43]. Moreover, data preprocessing techniques are widely employed to address phenomena like capacity regeneration during battery degradation and measurement noise induced by random load fluctuations. These techniques include, but are not limited to, Empirical Mode Decomposition (EMD)^[44], Variational Mode Decomposition (VMD)^[33], Variational Filtering (VF)^[45], and Particle Filtering (PF)^[46]. The combination of data-driven methods with data preprocessing techniques offers a potent approach to further improve the accuracy of RUL

predictions^[33,38].

While data-driven methods have advanced battery RUL prediction, their practical application faces three interconnected challenges: (1) RNN-based models often fail to capture long-range dependencies in degradation data due to gradient issues^[47]; (2) sensor noise and capacity regeneration can mislead predictions, yet existing preprocessing techniques are not well integrated with deep learning models^[20]; (3) many models require large training datasets - often scarce in practice - and struggle to generalize across different battery types.

To address the aforementioned challenges holistically, this study proposes a novel hybrid framework that synergistically integrates advanced data preprocessing with a multi deep learning architecture. Unlike previous works that may focus on a single aspect, our work is designed to make concurrent breakthroughs in all three areas. The specific contributions and how they map to the challenges are as follows: To capture long-range dependencies, we leverage the Transformer module, whose self-attention mechanism is inherently powerful at modeling global contextual information across the entire battery lifecycle; To mitigate noise interference, we incorporate Complementary Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) as a robust preprocessing step to denoise the health indicators and disentangle the capacity regeneration phenomenon; To extract local features and model complex temporal dynamics for improved data efficiency and robustness, we integrate Convolutional Neural Networks (CNN) for local pattern extraction and a Bidirectional Gated Recurrent Unit (BiGRU) for bidirectional temporal dependency modeling. The CNN-BiGRU structure ensures that both short-term fluctuations and sequential trends are captured effectively, enhancing performance even with limited training data. The synergistic integration of these components enables unified modeling of short-term variations and global degradation patterns, significantly enhancing prediction accuracy, robustness, and deployment feasibility in practical engineering applications.

In the present work, experimental validation employs battery datasets from NASA and Lithium Nickel Cobalt Manganese Oxide (NCM) battery datasets from China University of Geosciences (CUG). Comparative analysis against traditional models (LSTM, GRU, BiGRU, and BiLSTM networks) demonstrates superior prediction accuracy without substantial computational overhead. Furthermore, cross-dataset validation demonstrates robust generalization capabilities of the proposed model across diverse battery types, including NASA and NCM types.

MATERIALS AND METHODS

To present the methodology more clearly, we organize the description into three parts: (1) data preprocessing, (2) model configuration, and (3) implementation details.

Data processing

In data preprocessing, health indicators (HIs) are first extracted from raw battery operational data. The validity of these HIs is rigorously evaluated to establish strong correlation with battery degradation trajectories. Subsequent application of CEEMDAN decomposes and reconstructs the HI time series, effectively mitigating noise interference from capacity regeneration and electromagnetic disturbances. This study utilizes two distinct battery datasets: the NASA B005 cell and the China University of Geosciences NCM-1 cell. Their representative capacity degradation curves are presented in Figure 1. Using the NASA battery dataset as an example, we first detail the HIs extraction methodology.

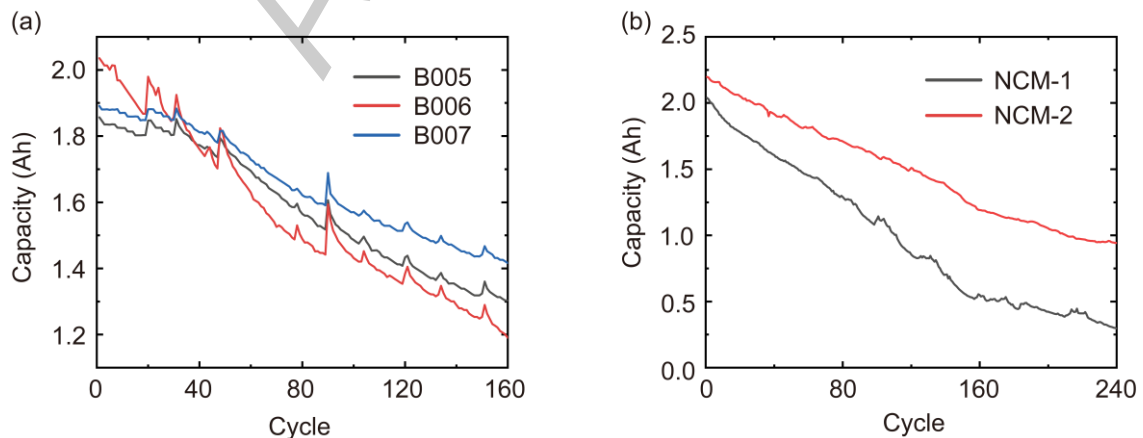


Figure 1. Representative battery capacity degradation curves. (a) NASA battery dataset. (b) NCM battery dataset.

Beyond direct capacity utilization, selecting appropriate HIs for indirect capacity prediction offers substantial cost savings in RUL prediction. Consequently, this study extracts indirect HIs from readily monitorable voltage and current profiles during charge/discharge cycles to assess LiB capacity degradation, as shown in Figure 2a-c for discharge voltage profiles, charge voltage profiles, and charge current profiles at uniform cycle intervals for NASA-B005 cells. These profiles reveal progressive deterioration of voltage response and current characteristics with cycling, indicating underlying material/structural alterations. Three temporally correlated features - effectively characterizing LiBs health - are identified, including H1 the discharge voltage decay duration, H2 the charge voltage rise duration, and H3 constant-current charge duration. The cycle-by-cycle evolution of these HIs in Figure 2 (d-f) exhibits significant correspondence with the capacity profiles shown in Figure 1a. To quantitatively validate the degradation relevance, linear (Pearson) and monotonic (Spearman) correlations between each HI and RUL are analyzed using their respective coefficients (for details see Supplementary Note 1). As a result, Figure 2g and Supplementary Figure 1 illustrates correlation analysis results for NASA-B005. Both coefficients demonstrate strong HI-RUL relationships, confirming the selected HIs from battery data as validated inputs for subsequent prediction, with capacity designated as the output variable.

Following the same methodology, we extract HIs from the NCM-1 battery data. Supplementary Figure 2a and b presents NCM-1's discharge and charge voltage profiles sampled at uniform cycle intervals. Given the constant-current charge/discharge protocol employed for NCM cells, only H1 (Discharge voltage decay duration) and H2 (Charge voltage rise duration) are utilized as HIs. The cycle-by-cycle evolution of these HIs is depicted in Supplementary Figure 2c and d, while Supplementary Figure 2e and f respectively displays their corresponding Pearson and

Spearman correlation coefficients. It is observed that both HIs exhibit strong correlations with the capacity degradation profile shown in Figure 1b, confirming their suitability as prognostic inputs.

To mitigate noise interference from capacity regeneration, electromagnetic disturbances, and varying charge/discharge rates, this study employs CEEMDAN for input data preprocessing. CEEMDAN resolves mode mixing through adaptive noise injection, facilitating effective separation of distinct modes. Additionally, it reduces pseudo-modes via ensemble averaging, smooths injected noise, and retains only significant components. After a series of decomposition processes (for details see Supplementary Note 2), the HI sequence $H(t)$ decomposes into k IMFs and a residual $R(t)$ ^[20]:

$$H(t) = R(t) + \sum_{k=1}^n IMF_k(t). \quad (1)$$

Figure 2h and Supplementary Figure 3 illustrate the decomposition results of HI sequences for NASA-B005 and NCM-1 batteries. Eventually, HI reconstruction via autocorrelation coefficients of IMFs yields denoised inputs for subsequent machine learning models.

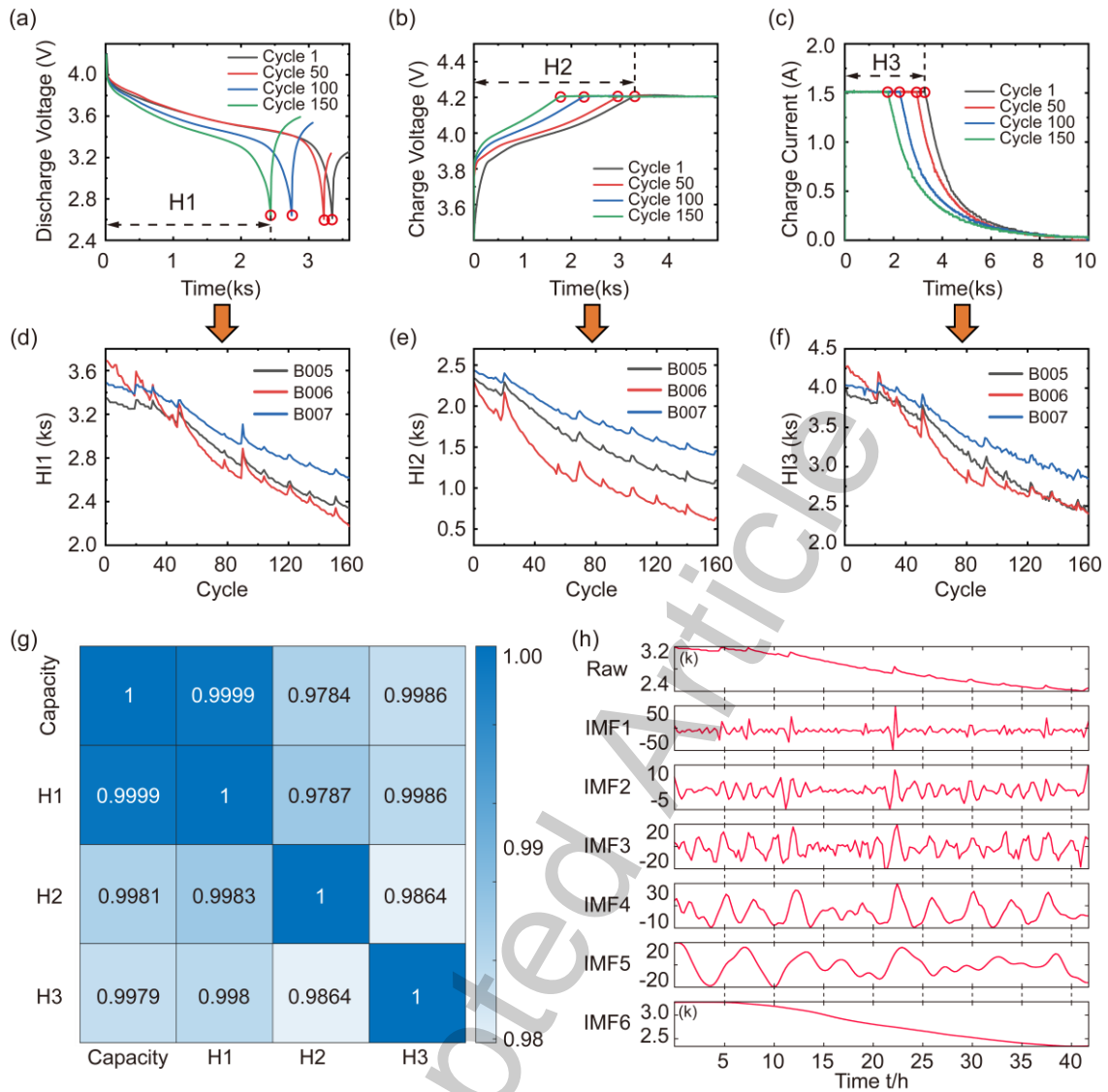


Figure 2. HIs extraction of NASA-B005 battery data. (a) Discharge voltage versus time curve. (b) Charge voltage versus time. (c) Charge current versus time. (d) The extracted H1 represents discharge voltage decay duration. (e) H2 is charge voltage rise duration. (f) H3 represents constant-current charge duration. (g) show the Spearman correlation analysis diagram. (h) The CEEMDAN decomposition curves for the extracted H1 for NASA-B005, where $h = 4$ is sampling interval.

Model training configuration

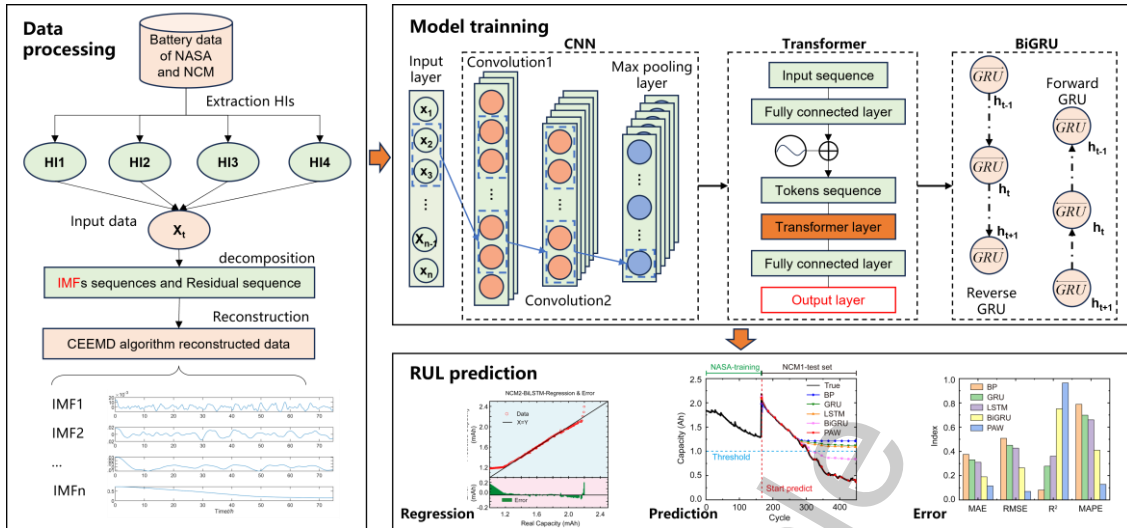


Figure 3. The architecture of our battery RUL prediction framework.

This study adopts an indirect prediction approach, with the lithium-ion battery RUL prediction framework illustrated in Figure 3. For model training, we propose a hybrid Transformer-CNN-BiGRU architecture for RUL estimation. Within this integrated framework, each component provides complementary capabilities. The CNN extracts localized transient patterns, the Transformer resolves long-range dependencies, and the BiGRU captures bidirectional temporal contexts, collectively enabling robust and accurate RUL prediction. The RUL prediction module employs the initial 40%, 50%, and 60% of capacity data as training sets to investigate model sensitivity to data volume. Reconstructed HI sequences serve as model inputs, with corresponding capacity measurements as outputs. The Transformer-CNN-BiGRU network undergoes pretraining using these varying-length datasets before performing real-time RUL prediction on the remaining data. Prediction quality is quantified through established metrics - Mean Absolute Error (MAE), Root Mean Square Error (RMSE), coefficient of determination (R^2), and Mean Absolute Percentage Error (MAPE). Implementation details for each module are elaborated in subsequent sections.

Transformer-CNN- BiGRU prediction model

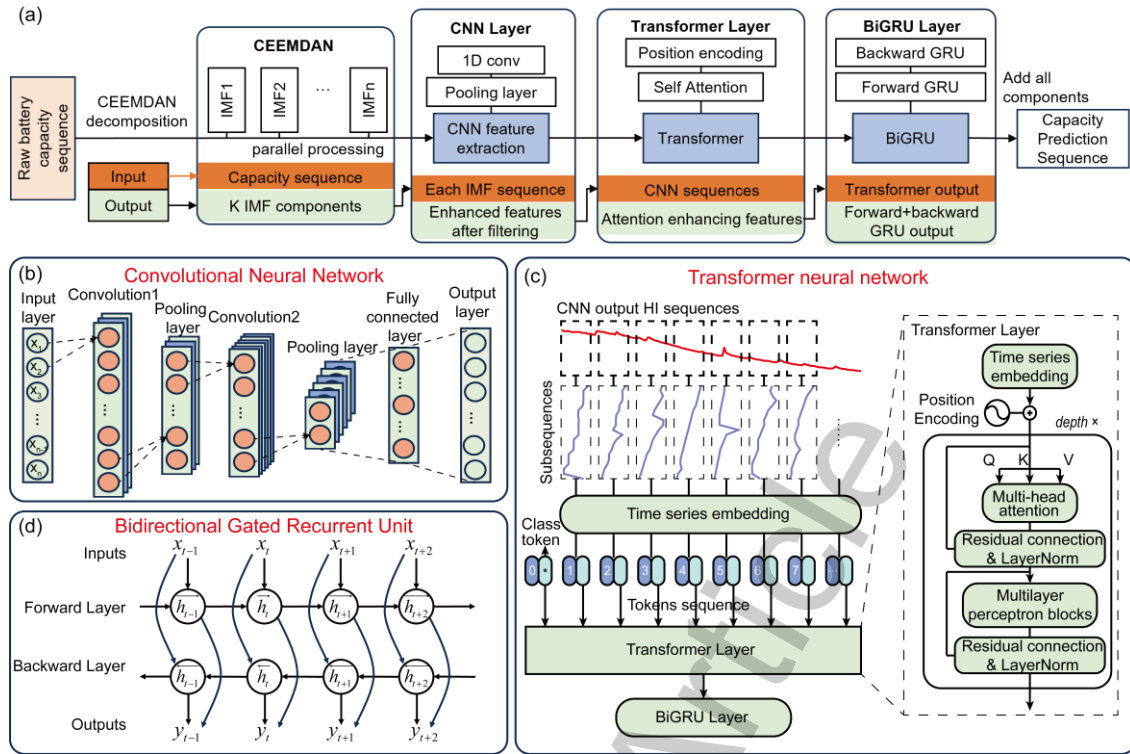


Figure 4. Transformer-CNN-BiGRU Model structure. (a) Model framework. Schematic diagram of (b) CNN structure, (c) Transformer structure, (d) BiGRU structure.

This section details the hybrid neural network model, Transformer-CNN-BiGRU, adopted in the present work. This model captures short-term local characteristics through CNN, processes long-range dependencies via Transformer, and models bidirectional temporal information using BiGRU, thereby generating a more robust and accurate predictor for battery health states. Overall, the four modules - CEEMDAN, Transformer, CNN, and BiGRU - work collaboratively as illustrated in Figure 4a. Their respective tasks and core functions are outlined below:

(1) CEEMDAN is responsible for signal decoupling and denoising. Its key role is to decompose the original multicomponent signal into several Intrinsic Mode Functions (IMFs), each corresponding to features at a specific frequency scale, thereby isolating a purified degradation signature. It takes the raw battery health indicator (HI) degradation curve as input and outputs k IMFs, each being a one-dimensional time series of the same length as the original signal.

(2) CNN performs spatio-temporal feature extraction. Its core function is to capture local frequency-domain features through convolution and, by sliding the convolutional window along the time dimension, to extract the temporal evolution patterns of battery degradation. It receives the denoised HI sequence (IMF time-frequency information) as input, processes each IMF sequence in parallel, uses past time steps (the input window) to predict the next time step, and outputs data after convolution and pooling.

(3) Transformer aims to extract global features and long-term dependencies from the sequence. It works by enabling any time point in the sequence to interact directly with and be weighted against all other time points. Positional encoding adds order information to the input sequence, allowing the model to understand the sequential arrangement of data points. Its input is the IMF sequence containing local features output by the CNN, and its output is a sequence that encodes long-range dependencies.

(4) BiGRU is designed to capture short-term temporal dynamics from both forward and backward directions. It operates through reset and update gates to effectively model short-term dependencies in the sequence and mitigate the vanishing gradient problem. It takes the Transformer's output as input and produces a prediction for each IMF component.

For the k VMD components, k independent Transformer-BiGRU models are trained. The prediction results for each component on the training and test sets are obtained, and the predictions from all k components are summed directly to reconstruct the final capacity prediction sequence.

To extract local spatial features from input data, we employ CNN to apply convolutional filters to the input time series. The primary CNN structure consists of an input layer, convolutional layer, pooling layer, fully connected layer, and output layer,

as illustrated in Figure 4b. For 1D convolution, the feature extraction operation is defined as^[22]:

$$F_j^n = f(w_k^n * Hcnn^{n-1} + C_k^n), \quad (2)$$

where $f(\cdot)$ denotes the convolutional filter, w_k^n represents the weight matrix of the k -th kernel in the n -th layer, $Hcnn^{n-1}$ is the output from the $(n-1)$ -th layer, F_j^n corresponds to the k -th feature in the n -th layer, ‘*’ is the convolution operator, and C_k^n denotes the bias term. Subsequently, the pooling layer extracts salient features from multiple feature matrices^[20]:

$$y_k^{n+1}(w) = \max\{Hcnn_j^n(x)\}, \quad x \in D, \quad (3)$$

where $Hcnn_j^n$ denotes elements within the pooling kernel region of the k -th feature matrix in the n -th layer, $y_k^{n+1}(w)$ represents elements of the k -th feature matrix in the $(n + 1)$ -th layer after pooling, and D signifies the region covered by the k -th pooling kernel. Finally, the fully connected layer integrates these features by mapping the pooled data to the output layer.

The Transformer architecture is a method for processing long sequential data based on self-attention mechanisms. By computing pairwise relationships across all timesteps, it enables the model to capture global dependencies. Its structure is illustrated in Figure 4c. For time series forecasting, the Transformer utilizes multi-head attention layers to compute multiple self-attention heads in parallel, concatenating their outputs for the final result. Using linearly projected Query (Q), Key (K), and Value (V) matrices from the input sequence, attention weights are computed via the attention function (for details see Supplementary Note 3). This process allows the model to learn

dependencies between distant time steps.

For the last module, GRU represents a streamlined variant of LSTM, incorporating only a reset gate and an update gate, as shown in Figure 4d. While achieving comparable prediction accuracy to LSTM, GRU's optimized structure features fewer trainable parameters and accelerated convergence speed. The computational flow within each GRU cell see Supplementary Note 4. Furthermore, BiGRU captures dependencies between sequence endpoints by processing information bidirectionally forward (start-to-end) and backward (end-to-start). The Transformer output features are processed by a BiGRU layer to model bidirectional temporal dependencies^[20]:

$$\vec{h}_t = \overline{GRU}(h_{trans,t}, \vec{h}_{t-1}). \quad (4)$$

The backward GRU Layer processes the reversed sequence:

$$\overleftarrow{h}_t = \overline{GRU}(h_{trans,q-t+1}, \overleftarrow{h}_{t-1}), \quad (5)$$

and the final output \mathbf{h}_t at each timestep t integrates both directions:

$$h_t = \alpha \vec{h}_t + \beta \overleftarrow{h}_t + c_t, \quad (6)$$

where α, β are weight coefficients controlling contributions of forward/backward states, and c_t is the bias term.

RESULTS AND DISCUSSION

This section focuses on experimental validation of the proposed hybrid model, structured through four pillars: dataset and experimental setup, validation of the method, and comparative analysis. The experiments aim to demonstrate significant improvements in reliability and prediction accuracy over standalone architectures and

baseline models (GRU, LSTM, CNN-GRU).

Dataset

Two distinct battery degradation datasets were utilized, including NASA Battery Prognostics Repository (B0005 cell) and laboratory-tested NCM batteries obtained at China University of Geosciences. The NCM battery testing protocol maintained ambient temperature at $25 \pm 1^\circ\text{C}$ within a 2.7-4.3V voltage window under half-cell configuration, implementing formation cycling at 0.1C for 5 cycles followed by 300 aging cycles at 1C, with electrode mass loading standardized at 12.8 mg/cm^2 on 15mm diameter substrates. We will open source this data for researchers' reference.

Experimental setting

The proposed lithium battery remaining useful life prediction model employs a hybrid Transformer-CNN-BiGRU architecture to capture both long-term dependencies and sequential degradation patterns. The specific experimental parameters are listed in Supplementary Table 1. Specifically, the Transformer module processes input sequences using a historical window of two cycles, with positional encoding supporting sequences up to 512 steps. It incorporates four parallel attention heads, each projecting inputs into 32-dimensional key spaces, and utilizes causal masking in the initial attention layer to enforce temporal constraints. The bidirectional GRU component consists of asymmetric forward and backward processing paths, with the forward GRU containing six hidden units and the reverse GRU containing ten hidden units. These bidirectional features are fused through concatenation before final regression output. The model processes normalized capacity data scaled to the [0,1] range using the Adam optimizer with an initial learning rate of 0.001, which decays by a factor of 0.1 after 50 epochs. Training implemented a batch size of 256 samples with L2 regularization ($\lambda = 0.001$) and gradient clipping at a threshold of 10 to ensure stable convergence. This configuration leverages the Transformer's strength in capturing complex temporal relationships across cycles while utilizing the BiGRU's capacity to model bidirectional degradation dynamics from limited historical data points. Moreover, to quantitative

evaluate our method. we adopt three statistical metrics - Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE), for details see Supplementary Note 5.

Model validation

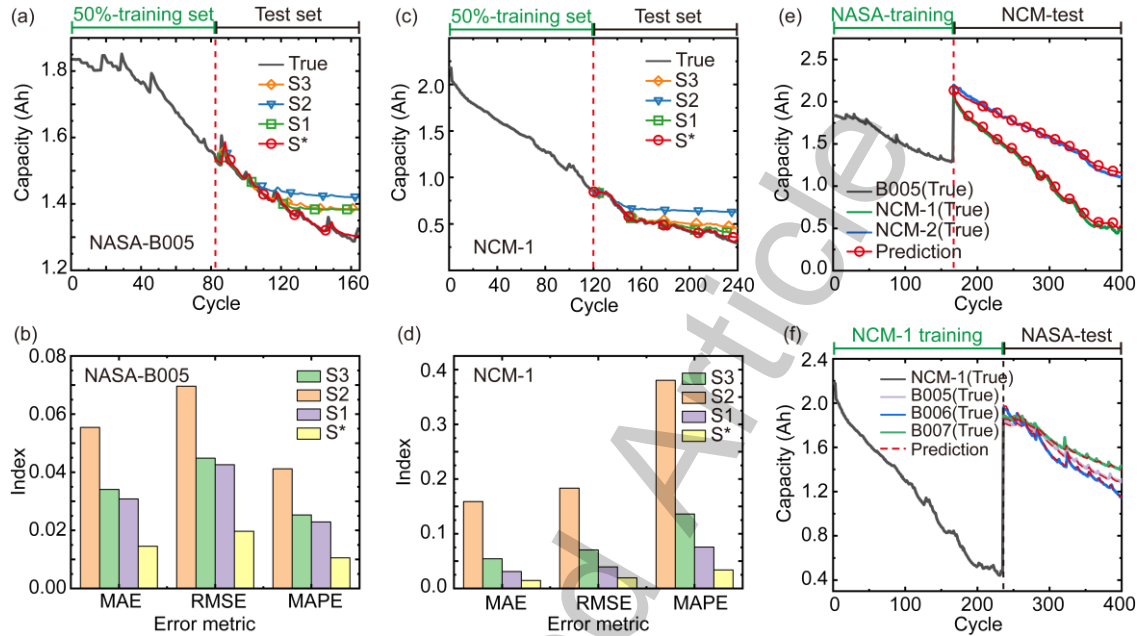


Figure 5. Validation of RUL prediction performance. (a) RUL prediction results and (b) Evaluation metrics for NASA-B005 battery cell using different schemes. (c) RUL prediction results for NCM-1 battery cell using different schemes. (d) Evaluation metrics for NCM-1 prediction results. (e) Predicted versus actual RUL curves for the NCM-1 and NCM-2 batteries using the model trained exclusively on NASA-B005 data. (f) Generalization results on NASA-B005, B006, and B007 batteries from the model trained on NCM-1 data.

To evaluate the individual contributions of each component module within the proposed Transformer-CNN-BiGRU hybrid model, combined with CEEMDAN data preprocessing, to the predictive performance, four comparative schemes were designed as presented in Table 1. Here, S* denotes the complete proposed scheme, while S1, S2, and S3 represent the prediction schemes where the CNN module, the CEEMDAN preprocessing, and the Transformer module were removed, respectively. These four schemes were applied to both NASA and NCM batteries. The corresponding prediction

results and evaluation metrics are illustrated in Figure 5a-d. Specifically, Figure 5a and b display the results for the NASA battery, and Figure 5 c and d show the results for the NCM battery. For both experimental datasets, the first 50% of the data was utilized as the training set during the train-test split.

Table 1. Method statement

Symbol	Method
S^*	CEEMDAN+Transformer-CNN-BiGRU
$S1$	CEEMDAN+Transformer-BiGRU
$S2$	CEEMDAN+CNN-BiGRU
$S3$	Transformer-CNN-BiGRU

As evident from the Figure 5a-d, S^* (the full model) achieved the best performance on both battery datasets, with specific error metrics provided in Supplementary Table 2. It achieved the lowest MAPE on the NASA dataset (approximately 1.05%) and an MAPE below 3.3% on the NCM dataset. The predicted curve closely follows the true values, remaining stable even during capacity regeneration phases (e.g., curve fluctuations), indicating that all modules work together effectively to capture short-term fluctuations, long-term trends, and noise robustness. $S1$ (removing CNN)'s performance declined noticeably, with significantly higher MAE and MAPE than S^* . For example, on the NASA data [Figure 5a and b], the prediction curve shows smoothed deviations in the mid-term and fails to track local capacity fluctuations accurately. This highlights the critical role of the CNN: it extracts local features (e.g., short-term changes in voltage/current) and provides essential context for subsequent modules. Without the CNN, the Transformer and BiGRU lack local contextual information, leading to reduced sensitivity to transient variations. $S2$ (removing Transformer) performed the worst, with the prediction curve deviating substantially from the true values toward the end of the cycle [Figure 5a and c], and the MAPE on the NASA data increased sharply.

This indicates that the Transformer is central to handling long-range dependencies: its self-attention mechanism captures the global degradation trend, preventing the model from overfitting to short-term noise. Without the Transformer, the model (CNN-BiGRU only) overfits locally and fails to model the slow degradation over the full life cycle effectively. The ablation experiments in Figure 5a-d confirm that the Transformer, CNN, BiGRU, and CEEMDAN modules interact through tightly coupled sequential collaboration, achieving a clear division of labor and mutual reinforcement. The superiority of the full model (S*) stems from the complementary roles of the modules: CNN-extracted local features support the global analysis of the Transformer, the Transformer's output improves the bidirectional modeling of the BiGRU, and CEEMDAN pre-processing supplies a clean input to the entire pipeline.

Moreover, the denoising efficacy of CEEMDAN and its parameter sensitivity are pivotal to the model's performance. Supplementary Note 6 provides a comprehensive analysis, showing that the prediction error (MAE) reaches a minimum at a noise amplitude of 0.15 (see Supplementary Figure 4a and Supplementary Table 3) and stabilizes with approximately six IMFs (see Supplementary Figure 4b and Supplementary Table 4). This parameter optimization directly contributes to the high accuracy observed in our main results. Moreover, as shown in Supplementary Figure 4c and the associated error metrics provided in Supplementary Table 5, comparative studies with VMD and wavelet-based denoising demonstrate that CEEMDAN more effectively captures the underlying capacity regeneration trends, which is critical for reducing error propagation in the prognostic model.

To validate the generalization capability of the proposed model across diverse battery types, we conducted a cross-dataset experiments. As illustrated in Figure 5e, the model trained exclusively on the entire NASA-B005 dataset was deployed to predict RUL for the NCM cells, while Figure 5f demonstrates predictions on NASA-B005, B006, and B007 batteries using a model trained solely on NCM-1 data. Furthermore, the

evaluation metrics [Supplementary Table 6] reveal that predicted capacity curves exhibit remarkable consistency with ground-truth measurements across all test cases. Specifically, MAE values remained below 0.04 and MAPE under 3.5% despite significant differences in battery materials and cycling protocols. These results conclusively demonstrate the model's capacity for robust cross-dataset RUL predictions, highlighting its potential for deployment in practical battery management systems with diverse cell types.

Furthermore, to validate the accuracy of health indicator (HI) extraction, we conducted a comparative experiment under the same model configuration, using either the extracted HI or the direct capacity as input. The corresponding prediction results are shown in Supplementary Figure 5, and the associated error metrics are provided in Supplementary Table 7. It can be observed that the results obtained using HI as an indirect input exhibit only minimal discrepancy compared to those using the direct capacity values. This small deviation can be attributed to the strong correlation between the HI and the actual capacity, as supported by the high Spearman and Pearson correlation coefficients discussed earlier. Therefore, it can be concluded that the use of HI as an indirect metric adequately meets the requirement for high-precision prediction.

Comparative analysis

To further validate the performance improvement of the proposed hybrid model over standalone architectures and baseline models, comparative studies were conducted against LSTM, GRU, BiLSTM, and CNN-BiGRU. As depicted in Figure 6, these five models were evaluated using NASA-B005 data, with training sets comprising the initial 40%, 50%, and 60% of the data, respectively. This allowed for an assessment of model performance relative to the amount of available training data.

Figure 6a, c and e demonstrate that the predicted capacity curves of all five models progressively converge towards the true curve as the volume of training data increases. Notably, the proposed hybrid model exhibited superior predictive performance even

with only 40% training data. As Figure 6b reveals, its MAE, RMSE, and MAPE values were substantially lower than those of the four baseline models across all training data fractions. Besides, among the standalone models, performance was unsatisfactory for LSTM, GRU, and BiLSTM. Except for our hybrid model, only the CNN-BiGRU model achieved marginally satisfactory performance when trained on 60% of the data [Figure 6e and f]. Supplementary Table 8 presents detailed error metrics, which further support the aforementioned results. An identical comparative analysis was performed on the NCM-1 battery dataset, with results presented in Supplementary Figure 6. The proposed hybrid model consistently maintained excellent predictive performance on this dataset, unambiguously demonstrating its generalization capability.

These results collectively indicate that the proposed hybrid model achieves high-precision predictions with only a small amount of training data. Compared to the baseline models, the hybrid model offers a significant enhancement in prediction accuracy while exhibiting robust generalization capability.

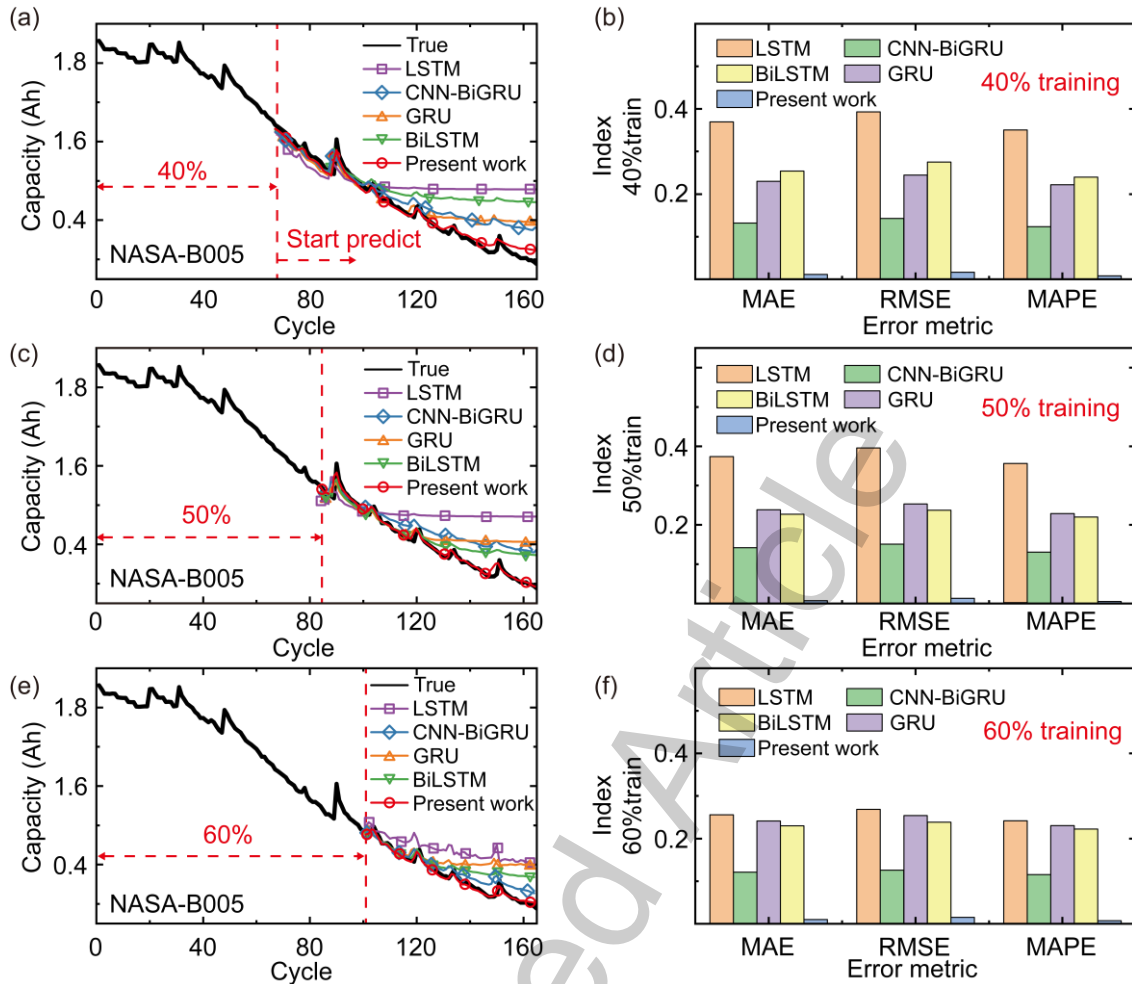


Figure 6. The RUL prediction results for the NASA-B005 battery. (a, c and e) respectively show the prediction results using the first 40%, 50%, and 60% of the data as the training set, while (b, d and f) display the corresponding metric analyses.

CONCLUSIONS

In summary, this work presents a highly effective data-driven framework for battery RUL prognostics. The innovative integration of CEEMDAN denoising with a Transformer-CNN-BiGRU architecture successfully addresses key challenges in capturing both long-term degradation trends and short-term capacity variations. Besides, by leveraging readily monitorable voltage and current data to derive predictive HIs, the framework offers a cost-effective alternative to direct capacity measurement, enhancing feasibility for deployment in battery management systems. Critically, cross-dataset validation demonstrates robust generalization, achieving MAPE below 3.5% when transferring between NASA and NCM batteries without retraining. Moreover, the

model's ability to deliver reliable long-term RUL forecasts with limited early-cycle data is particularly valuable for optimizing battery utilization, mitigating safety risks associated with unexpected failures. Future work will focus on validating the framework across an even broader range of battery types and operating scenarios, exploring adaptive mechanisms for continuous online learning within battery management system constraints, and further optimizing the architecture for embedded implementation.

DECLARATIONS

Authors' contributions

Data curation (lead); formal analysis (lead); investigation (lead); methodology (lead); visualization (lead); writing - original draft (lead); writing - review and editing (lead): Xu, M.

Data curation; formal analysis; investigation: Xing, S.; Hong, J.; Tian, X.; Huang, B.; Jin, H.

Conceptualization (lead); supervision (lead); visualization; writing - original draft (lead); writing - review and editing (lead): Li, J.

Availability of data and materials

Part of the data in this study is sourced from the NASA Open Science Data Repository. NCM battery data used in this study are available on GitHub website: <https://github.com/MTT-ginger/NCM-battery-dataset>. The machine-learning codes developed for the accelerated framework are available on GitHub website: <https://github.com/MTT-ginger/CEEMDAN-CNN-Transformer-BiGRU>.

Financial support and sponsorship

This work was supported by the National Natural Science Foundation of China (12192213, 92466201, 12202155), the Guangdong Basic and Applied Basic Research Foundation (2023B1515130003), the Financial Support for Outstanding Talents Training Fund in Shenzhen, the Guangdong Provincial Key Laboratory Program (Grant No. 2021B1212040001), the China Postdoctoral Science Foundation (2025M771862).

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2025.

REFERENCES

- [1] Goodenough, J. B.; Kim, Y. Challenges for Rechargeable Li Batteries. *Chemistry of Materials* 2010, 22 (3), 587-603. 10.1021/cm901452z
- [2] Huang, J.; Boles, S. T.; Tarascon, J.-M. Sensing as the key to battery lifetime and sustainability. *Nature Sustainability* 2022, 5 (3), 194-204. 10.1038/s41893-022-00859-y
- [3] Dembitskiy, A. D.; Humonen, I. S.; Eremin, R. A.; Aksyonov, D. A.; Fedotov, S. S.; Budenny, S. A. Benchmarking machine learning models for predicting lithium ion migration. *npj Computational Materials* 2025, 11 (1), 131. 10.1038/s41524-025-01571-z
- [4] Attia, P. M.; Grover, A.; Jin, N.; Severson, K. A.; Markov, T. M.; Liao, Y.-H.; Chen, M. H.; Cheong, B.; Perkins, N.; Yang, Z.; et al. Closed-loop optimization of fast-charging protocols for batteries with machine learning. *Nature* 2020, 578 (7795), 397-402. 10.1038/s41586-020-1994-5
- [5] Downie, L. E.; Hyatt, S. R.; Wright, A. T. B.; Dahn, J. R. Determination of the Time Dependent Parasitic Heat Flow in Lithium Ion Cells Using Isothermal Microcalorimetry. *The Journal of Physical Chemistry C* 2014, 118 (51), 29533-29541. 10.1021/jp508912z
- [6] 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. *Nature Energy* 2019, 4 (5), 383-391.
10.1038/s41560-019-0356-8

[7] Tian, H.; Qin, P.; Li, K.; Zhao, Z. A review of the state of health for lithium-ion batteries: Research status and suggestions. *Journal of Cleaner Production* 2020, 261, 120813. <https://doi.org/10.1016/j.jclepro.2020.120813>

[8] Lin, C.-P.; Cabrera, J.; Yang, F.; Ling, M.-H.; Tsui, K.-L.; Bae, S.-J. Battery state of health modeling and remaining useful life prediction through time series model. *Applied Energy* 2020, 275, 115338. <https://doi.org/10.1016/j.apenergy.2020.115338>

[9] Shu, X.; Yang, W.; Wei, K.; Qin, B.; Du, R.; Yang, B.; Garg, A. Research on capacity characteristics and prediction method of electric vehicle lithium-ion batteries under time-varying operating conditions. *Journal of Energy Storage* 2023, 58, 106334. <https://doi.org/10.1016/j.est.2022.106334>

[10] Wang, Y. Application-oriented design of machine learning paradigms for battery science. *npj Computational Materials* 2025, 11 (1), 89. 10.1038/s41524-025-01575-9

[11] Yang, L.; Cai, Y.; Yang, Y.; Deng, Z. Supervisory long-term prediction of state of available power for lithium-ion batteries in electric vehicles. *Applied Energy* 2020, 257, 114006. <https://doi.org/10.1016/j.apenergy.2019.114006>

[12] Ramadesigan, V.; Northrop, P. W. C.; De, S.; Santhanagopalan, S.; Braatz, R. D.; Subramanian, V. R. Modeling and Simulation of Lithium-Ion Batteries from a Systems Engineering Perspective. *Journal of The Electrochemical Society* 2012, 159 (3), R31. 10.1149/2.018203jes

[13] Deng, Z.; Yang, L.; Deng, H.; Cai, Y.; Li, D. Polynomial approximation pseudo-two-dimensional battery model for online application in embedded battery management system. *Energy* 2018, 142, 838-850. <https://doi.org/10.1016/j.energy.2017.10.097>

[14] Karimi, D.; Behi, H.; Van Mierlo, J.; Berecibar, M. Equivalent Circuit Model for High-Power Lithium-Ion Batteries under High Current Rates, Wide Temperature Range, and Various State of Charges. In *Batteries*, 2023; Vol. 9.10.3390/batteries9020101

[15] Larsson, C.; Larsson, F.; Xu, J.; Runesson, K.; Asp, L. E.

- Electro-chemo-mechanical modelling of structural battery composite full cells. *npj Computational Materials* 2025, 11 (1), 141. 10.1038/s41524-025-01646-x
- [16] Wang, F.; Zhai, Z.; Zhao, Z.; Di, Y.; Chen, X. Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis. *Nature Communications* 2024, 15 (1), 4332. 10.1038/s41467-024-48779-z
- [17] E, L.; Wang, J.; Yang, R.; Wang, C.; Li, H.; Xiong, R. A physics-informed neural network-based method for predicting degradation trajectories and remaining useful life of supercapacitors. *Green Energy and Intelligent Transportation* 2025, 4 (3), 100291. <https://doi.org/10.1016/j.geits.2025.100291>
- [18] E, L.; Wang, J.; Sun, Y.; Shen, W.; Xiong, R. Physics-Informed Hybrid Neural Architecture for Coupled Degradation Modeling and Remaining Useful Life Prediction of LiFePO₄ Batteries. *Journal of The Electrochemical Society* 2025, 172 (6), 060505. 10.1149/1945-7111/addd4e
- [19] Wen, P.; Ye, Z. S.; Li, Y.; Chen, S.; Xie, P.; Zhao, S. Physics-Informed Neural Networks for Prognostics and Health Management of Lithium-Ion Batteries. *IEEE Transactions on Intelligent Vehicles* 2024, 9 (1), 2276-2289. 10.1109/TIV.2023.3315548
- [20] Lv, K.; Ma, Z.; Bao, C.; Liu, G. Indirect Prediction of Lithium-Ion Battery RUL Based on CEEMDAN and CNN-BiGRU. In *Energies*, 2024; Vol. 17.10.3390/en17071704
- [21] Xu, Q.; Wu, M.; Khoo, E.; Chen, Z.; Li, X. A Hybrid Ensemble Deep Learning Approach for Early Prediction of Battery Remaining Useful Life. *IEEE/CAA Journal of Automatica Sinica* 2023, 10 (1), 177-187. 10.1109/JAS.2023.123024
- [22] Zheng, D.; Zhang, Y.; Guo, X.; Ning, Y.; Wei, R. Research on the remaining useful life prediction method for lithium-ion batteries based on feature engineering and CNN-BiGRU-AM model. *Ionics* 2025, 31 (6), 5717-5736. 10.1007/s11581-025-06293-2
- [23] Wang, C.; Jiang, W.; Yang, X.; Zhang, S. RUL Prediction of Rolling Bearings Based on a DCAE and CNN. In *Applied Sciences*, 2021; Vol. 11.10.3390/app112311516

- [24] Guo, L.; Li, N.; Jia, F.; Lei, Y.; Lin, J. A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing* 2017, 240, 98-109. <https://doi.org/10.1016/j.neucom.2017.02.045>
- [25] Zhang, H.; Zhang, Q.; Shao, S.; Niu, T.; Yang, X. Attention-Based LSTM Network for Rotatory Machine Remaining Useful Life Prediction. *IEEE Access* 2020, 8, 132188-132199. [10.1109/ACCESS.2020.3010066](https://doi.org/10.1109/ACCESS.2020.3010066)
- [26] Zeng, F.; Li, Y.; Jiang, Y.; Song, G. A deep attention residual neural network-based remaining useful life prediction of machinery. *Measurement* 2021, 181, 109642. <https://doi.org/10.1016/j.measurement.2021.109642>
- [27] Tan, R.; Lu, X.; Cheng, M.; Li, J.; Huang, J.; Zhang, T.-Y. Forecasting battery degradation trajectory under domain shift with domain generalization. *Energy Storage Materials* 2024, 72, 103725. <https://doi.org/10.1016/j.ensm.2024.103725>
- [28] Wei, M.; Ye, M.; Wang, Q.; Xinxin, X.; Twajamahoro, J. P. Remaining useful life prediction of lithium-ion batteries based on stacked autoencoder and gaussian mixture regression. *Journal of Energy Storage* 2022, 47, 103558. <https://doi.org/10.1016/j.est.2021.103558>
- [29] Xing, J.; Zhang, H.; Zhang, J. Remaining useful life prediction of – Lithium batteries based on principal component analysis and improved Gaussian process regression. *International Journal of Electrochemical Science* 2023, 18 (4), 100048. <https://doi.org/10.1016/j.ijoes.2023.100048>
- [30] Li, H.; Zhao, W.; Zhang, Y.; Zio, E. Remaining useful life prediction using multi-scale deep convolutional neural network. *Applied Soft Computing* 2020, 89, 106113. <https://doi.org/10.1016/j.asoc.2020.106113>
- [31] Wu, L.; Yoo, S.; Suzana, A. F.; Assefa, T. A.; Diao, J.; Harder, R. J.; Cha, W.; Robinson, I. K. Three-dimensional coherent X-ray diffraction imaging via deep convolutional neural networks. *npj Computational Materials* 2021, 7 (1), 175. [10.1038/s41524-021-00644-z](https://doi.org/10.1038/s41524-021-00644-z)
- [32] Zhou, J.; Qin, Y.; Chen, D.; Liu, F.; Qian, Q. Remaining useful life prediction of bearings by a new reinforced memory GRU network. *Advanced Engineering*

Informatics 2022, 53, 101682. <https://doi.org/10.1016/j.aei.2022.101682>

[33] Gao, K.; Huang, Z.; Lyu, C.; Liu, C. Multi-scale prediction of remaining useful life of lithium-ion batteries based on variational mode decomposition and integrated machine learning. *Journal of Energy Storage* 2024, 99, 113372. <https://doi.org/10.1016/j.est.2024.113372>

[34] Rouhi Ardeshiri, R.; Ma, C. Multivariate gated recurrent unit for battery remaining useful life prediction: A deep learning approach. *International Journal of Energy Research* 2021, 45 (11), 16633-16648. <https://doi.org/10.1002/er.6910>

[35] Wang, Z.; Liu, N.; Chen, C.; Guo, Y. Adaptive self-attention LSTM for RUL prediction of lithium-ion batteries. *Information Sciences* 2023, 635, 398-413. <https://doi.org/10.1016/j.ins.2023.01.100>

[36] Wang, F.-K.; Amogne, Z. E.; Chou, J.-H.; Tseng, C. Online remaining useful life prediction of lithium-ion batteries using bidirectional long short-term memory with attention mechanism. *Energy* 2022, 254, 124344. <https://doi.org/10.1016/j.energy.2022.124344>

[37] Tian, Y.; Wen, J.; Yang, Y.; Shi, Y.; Zeng, J. State-of-Health Prediction of Lithium-Ion Batteries Based on CNN-BiLSTM-AM. In *Batteries*, 2022; Vol. 8.10.3390/batteries8100155

[38] Wang, T.; Wang, K.; Pan, K. A Transformer-CNN-BiLSTM Architecture for SOH Prediction of Battery Storage Systems. In *2024 IEEE 8th Conference on Energy Internet and Energy System Integration (EI2)*, 29 Nov.-2 Dec. 2024, 2024; pp 4591-4596.

[39] Jin, Y.; Hou, L.; Chen, Y. A Time Series Transformer based method for the rotating machinery fault diagnosis. *Neurocomputing* 2022, 494, 379-395. <https://doi.org/10.1016/j.neucom.2022.04.111>

[40] Zhang, Q.; Qin, C.; Zhang, Y.; Bao, F.; Zhang, C.; Liu, P. Transformer-based attention network for stock movement prediction. *Expert Systems with Applications* 2022, 202, 117239. <https://doi.org/10.1016/j.eswa.2022.117239>

[41] Hu, W.; Zhao, S. Remaining useful life prediction of lithium-ion batteries based on wavelet denoising and transformer neural network. *Frontiers in Energy Research* 2022,

Volume 10 - 2022, Methods. 10.3389/fenrg.2022.969168

- [42] Huang, X.; Tang, J.; Shen, Y. Long time series of ocean wave prediction based on PatchTST model. *Ocean Engineering* 2024, 301, 117572. <https://doi.org/10.1016/j.oceaneng.2024.117572>
- [43] Tang, P.; Qiu, Z.; Yao, Z.; Pan, J.; Cheng, D.; Gu, X.; Sun, C. Lithium-ion battery RUL prediction based on optimized VMD-SSA-PatchTST algorithm. *Scientific Reports* 2025, 15 (1), 26824. 10.1038/s41598-025-11934-7
- [44] Li, X.; Zhang, L.; Wang, Z.; Dong, P. Remaining useful life prediction for lithium-ion batteries based on a hybrid model combining the long short-term memory and Elman neural networks. *Journal of Energy Storage* 2019, 21, 510-518. <https://doi.org/10.1016/j.est.2018.12.011>
- [45] Jiao, R.; Peng, K.; Dong, J. Remaining Useful Life Prediction of Lithium-Ion Batteries Based on Conditional Variational Autoencoders-Particle Filter. *IEEE Transactions on Instrumentation and Measurement* 2020, 69 (11), 8831-8843. 10.1109/TIM.2020.2996004
- [46] Ahwiadi, M.; Wang, W. An Enhanced Mutated Particle Filter Technique for System State Estimation and Battery Life Prediction. *IEEE Transactions on Instrumentation and Measurement* 2019, 68 (3), 923-935. 10.1109/TIM.2018.2853900
- [47] Rahimifard, S.; Ahmed, R.; Habibi, S. Interacting Multiple Model Strategy for Electric Vehicle Batteries State of Charge/Health/ Power Estimation. *IEEE Access* 2021, 9, 109875-109888. 10.1109/ACCESS.2021.3102607