RESEARCH ARTICLE

A Hybrid RUL Prediction Approach for Lithium-ion Batteries Based on CEEMDAN-SSA-SVR-BiGRU

Fei Zhao[1,*] **and Xinyu Dai**[2]

[1] School of Management, Northeastern University at Qinhuangdao, Qinhuangdao 066004, China
[2] School of Control Engineering, Northeastern University at Qinhuangdao, Qinhuangdao 066004, China

Abstract

The capacity regeneration phenomenon in lithium-ion batteries is inevitable and leads to non-monotonic fluctuations in capacity degradation trajectories, significantly complicating accurate remaining useful life (RUL) prediction. To address this challenge, this paper proposes a hybrid prediction model based on CEEMDAN-SSA-SVR-BiGRU. The method first employs Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) to decompose the original capacity sequence into multiple Intrinsic Mode Functions (IMFs) representing local regeneration fluctuations, and a residual component (RES) referring to the global degradation trend, thereby achieving effective signal decoupling. Subsequently, distinct prediction strategies are applied to different components after decomposition. Support Vector Regression (SVR) is utilized to capture nonlinear local fluctuations, while Bidirectional Gated Recurrent Unit (BiGRU) models long-term dependencies. To further enhance the model performance, the Sparrow Search Algorithm (SSA) is introduced to jointly optimize kernel parameters and penalty factors in SVR, as well as architectural hyperparameters in BiGRU. Experimental results on the NASA lithium battery dataset demonstrate that the proposed model achieves higher accuracy, with Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE) no more than 0.0067, 0.0049, and 0.0094, respectively, significantly outperforming the ablated models and some baseline models. This study validates that the integration of signal decomposition, component-specific modeling, and hyperparameter optimization yields a significant improvement in the accuracy and robustness of the RUL prediction for lithium-ion batteries under capacity regeneration.

Keywords: remaining useful life, capacity regeneration, CEEMDAN, sparrow search algorithm, support vector regression, bidirectional gated recurrent unit.

1 Introduction

Lithium-ion batteries serve as the primary power source for new energy vehicles, portable electronic devices, and energy storage systems. Owing to the significant nonlinear and time-varying nature of their performance degradation, the RUL prediction has become a pivotal task in battery health management. Currently, the RUL prediction methods for lithium-ion batteries can be broadly categorized into three

Submitted: 01 December 2025
Accepted: 09 December 2025
Published: 30 December 2025

Vol. 1, **No.** 2, 2025.
10.62762/TSSR.2025.657859

***Corresponding author:**
✉ Fei Zhao
zhaofei@neuq.edu.cn

Citation
Zhao, F., & Dai, X. (2025). A Hybrid RUL Prediction Approach for Lithium-ion Batteries Based on CEEMDAN-SSA-SVR-BiGRU. *ICCK Transactions on Systems Safety and Reliability*, 1(2), 136–148.

© 2025 ICCK (Institute of Central Computation and Knowledge)

types, i.e., physic-based, data-driven and hybrid approaches [1]. Model-based studies which construct models based on the internal physicochemical mechanisms of batteries, offer clear mechanistic interpretability [2]. Ma et al. [3] combined Particle Filter (PF) with the Mann-Whitney U test to identify capacity regeneration points, utilizing an autoregressive model and PF algorithm for lithium-ion battery RUL prediction, where the capacity predicted by the autoregressive model was used to update the degradation model parameters of the PF algorithm. Nevertheless, such methods typically entail complex modeling processes and rely heavily on expert prior knowledge, which restricts their generalizability in practical applications.

In contrast, data-driven methods leverage historical operational data and employ machine learning techniques to uncover underlying degradation patterns [4], thereby avoiding dependence on complex internal battery characteristics and enhancing their practical applicability in intricate real-world application scenarios. Fang et al. [5] proposed a least squares SVR algorithm for the prediction of voltage variations in individual cells, which resulted in a successful improvement of prediction accuracy. Liu et al. [6] developed an improved SVR method for predicting the self-discharge voltage drop in lithium-ion batteries and demonstrated the performance. Furthermore, Sun et al. [7] proposed a novel fault prediction method named CNN-LSTM based on Convolutional Neural Network, Long Short-Term Memory and correlation coefficients, where CNN extracts spatial features and LSTM learns long-term dependencies. The prediction results on the dataset demonstrated that the proposed CNN-LSTM method achieves higher accuracy and reliability compared to LSTM and BiLSTM models, exhibiting significant potential for vehicle battery management systems.

To combine the strengths of both physic-based and data-driven paradigms, hybrid approaches have been developed to enhance prediction accuracy and robustness. Tian et al. [8] developed an intelligent online RUL prediction method for lithium batteries based on the artificial fish swarm algorithm and PF, optimizing particle distribution to improve prediction accuracy and convergence. Similarly, Chang et al. [9] proposed a hybrid approach integrating Unscented Kalman Filter (UKF), CEEMD, and Relevance Vector Machine (RVM). The method generates both prediction results and raw error sequences via UKF, analyzes and constructs new error sequences with CEEMD, and finally employs RVM to predict and correct prognostic errors, with experimental results validating its high reliability. Although the aforementioned hybrid methods have made significant progress in the field of RUL prediction for lithium-ion batteries, there are still some limitations and challenges in addressing the capacity regeneration phenomenon.

In real-world battery degradation processes, capacity regeneration manifests as a temporary recovery of capacity following several cycles of rest or low-load operation [10]. Driven by intricate electrochemical mechanisms such as lithium-ion redistribution, SEI film repair, and electrode stress release, this phenomenon gives rise to non-monotonic fluctuations in the capacity trajectory. Thereby, it severely undermines the predictive performance of traditional prediction models built on the assumption of monotonic degradation behavior. Current research efforts targeting the capacity regeneration phenomenon in RUL prediction predominantly focus on models that integrate signal decomposition algorithms with machine learning methods. For example, Yang et al. [11] employed ensemble empirical mode decomposition (EEMD) to decompose capacity data sequences and used grey wolf optimizer (GWO) to optimize parameters in SVR. Hence, an improved GWO-SVR model is presented to reduce the interference of capacity regeneration and improve RUL prediction accuracy. Yao et al. [12] utilized particle swarm optimization (PSO) to optimize the parameters in extreme learning machine (ELM) and RVM, and then the optimized ELM and RVM models are used to predict diverse sequences respectively, aiming to mitigate the impact of the capacity regeneration. However, PSO is prone to falling into local optima, which impairs the accuracy of RUL prediction results. Furthermore, the performance of deep learning models is largely contingent on hyperparameter tuning, making the choice of an appropriate optimization algorithm crucial in enhancing predictive performance. Xue et al. [13] tested SSA on 19 benchmark functions and compared its performance with other optimization algorithms like GWO and PSO, and the results demonstrated that SSA outperforms GWO and PSO in terms of accuracy, convergence speed, and stability.

To some extent, these above studies have reduced the prediction uncertainty incurred by the capacity regeneration in lithium-ion battery degradation

modeling and RUL prediction. However, such methods still suffer from certain limitations when confronted with the complex, variable, and stochastic characteristics of battery capacity fluctuations. For example, EEMD is prone to mode mixing issues, which hinder the accurate analysis of battery capacity data. This, in turn, leads to insufficient reliability and adaptability of prediction results in the face of complex real-world capacity regeneration scenarios. Therefore, to improve the accuracy and stability of RUL prediction in the context of capacity regeneration, it is imperative to explore more accurate capacity decomposition algorithms and design tailored model construction strategies for lithium-ion battery degradation analysis.

To overcome the above limitations, this paper proposes a novel lithium-ion battery RUL prediction method based on CEEMDAN-SSA-SVR-BiGRU, with the following specific contributions:

1) CEEMDAN is employed to decompose the original lithium-ion battery capacity sequence, which enables the effective decoupling of non-stationary capacity signals into global degradation components and local regeneration components, thereby separating capacity regeneration fluctuations from long-term degradation trajectories of lithium-ion batteries.

2) A hybrid SVR-BiGRU prediction model is constructed to achieve accurate prediction in accordance with the dynamic characteristics of distinct decomposed components. Specifically, SVR is specialized in capturing local nonlinear fluctuations, while BiGRU is harnessed to model the long-term temporal dependencies in the battery capacity sequence.

3) To address the local optima problem of traditional optimization algorithms, we incorporate the SSA for the global optimization of hyperparameters in both SVR and BiGRU models, which enhances model convergence efficiency and prediction accuracy of the hybrid model.

The remainder of this paper is structured as follows. Section 2 elaborates on the CEEMDAN-based capacity decomposition. Section 3 introduces the proposed CEEMDAN-SSA-SVR-BiGRU prediction framework. Section 4 provides comprehensive experimental validation and comparative analyses. Finally, conclusions are presented in Section 5.

2 Capacity Decomposition Based on CEEMDAN

Based upon EEMD [14], CEEMDAN is designed to more effectively address the mode mixing problem inherent in traditional empirical mode decomposition, enabling the stabilization of non-stationary signals [15]. In the CEEMDAN framework, the original time series is defined as $x(t)$, and E_i represents the i-th modal component obtained by decomposition using the EMD algorithm. The calculation steps are as follows:

Step 1: Add white noise $\delta_0\omega^i$ with a noise coefficient of δ_0 to $x(t)$:

$$x^i(t) = x(t) + \delta_0\omega^i(t) \quad i = 1, 2, ...N \tag{1}$$

Step 2: Perform EMD decomposition on $x^i(t)$ and calculate the average to obtain the first modal component IMF_1 :

$$IMF_1 = \frac{1}{N}\sum_{i=1}^{N} IMF_1^i(t) \tag{2}$$

Then the first RES $r_1(t)$ is calculated as follows:

$$r_1(t) = x(t) - IMF_1(t) \tag{3}$$

Step 3: Perform EMD decomposition for $r_1(t) + \delta_1 E_1[\omega^i(t)]$ and calculate the average to obtain the second modal component IMF_2

$$IMF_2(t) = \frac{1}{N}\sum_{i=1}^{N} E_1\{r_1(t) + \delta_1 E_1[\omega^i(t)]\} \tag{4}$$

Then we have the second RES $r_2(t)$, which is given by:

$$r_2(t) = r_1(t) - IMF_2(t) \tag{5}$$

Step 4: Repeat Steps 2 to 3. The m-th RES and the $(m+1)$-th modal component are expressed as:

$$r_m(t) = r_{m-1}(t) - IMF_m(t) \quad m = 1, 2, ..., M \tag{6}$$

$$IMF_{m+1}(t) = \frac{1}{N}\sum_{i=1}^{N} E_1\{r_m(t) + \delta_m E_m[\omega^i(t)]\} \tag{7}$$

Step 5: Repeat Step 4 until the RES shows a monotonic trend, and the original signal is decomposed into:

$$X(t) = \sum_{m=1}^{M} IMF_m(t) + R(t) \tag{8}$$

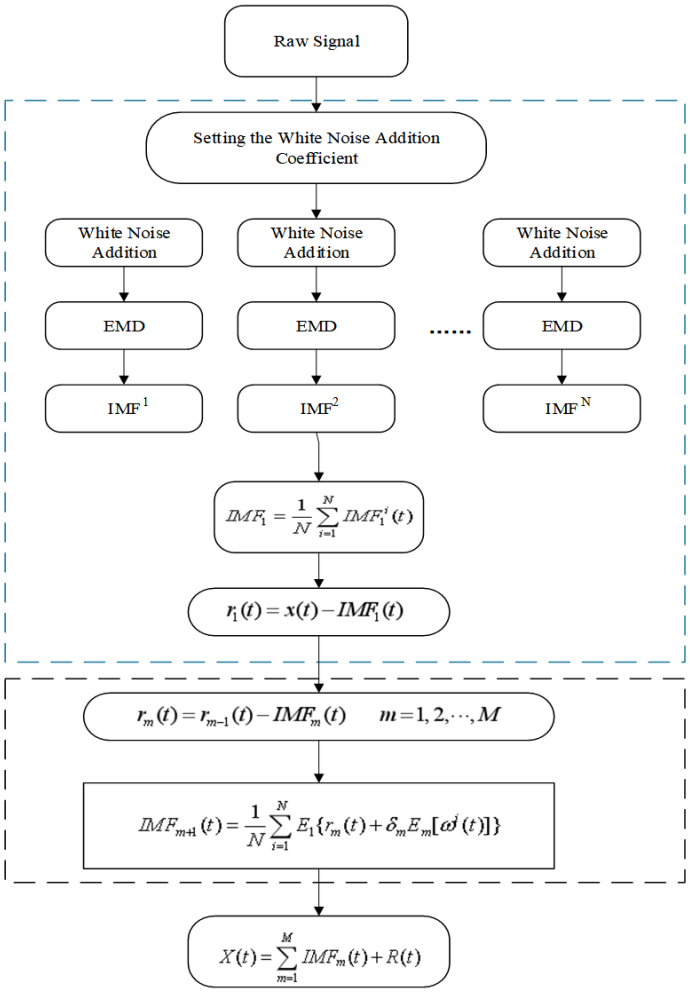


Figure 1. Flowchart of capacity decomposition based on CEEMDAN.

in which, M is the number of modal components, and $R(t)$ is the final RES. The specific process is shown in Figure 1.

To decompose the non-stationary capacity signals, we implemented the CEEMDAN algorithm on three specified lithium-ion battery cells from the NASA battery dataset. As illustrated in Figure 2, the capacity signals are decomposed into 6 IMFs and a RES. It is obviously seen that the RES shows a decreased curve that is consistent with the battery degradation trajectory. Figure 3 demonstrates that the RES captures the global degradation trend of battery capacity, while the IMFs reflect local regeneration characteristics. Consequently, the RES is defined as the global degradation indicator, and the IMF components are characterized as local regeneration indicators.

3 RUL Prediction Model Based on CEEMDAN-SSA-SVR-BiGRU

3.1 Support Vector Regression

SVR is a supervised learning algorithm based on support vector machine theory, designed for solving regression problems [16]. By adhering to the structural risk minimization principle, SVR exhibits significant advantages in addressing nonlinear problems. For a given training dataset, the SVR algorithm employs nonlinear mapping to project the data from a low-dimensional input space to a high-dimensional feature space [17]. Consider a dataset $D = ((x_i, y_i))_{i=1}^{n}$, where each input $x_i \in R^n$ and its associated output $y_i \in R$. An SVR function is characterized by a linear mapping, and defined as:

$$f(x_i) = \omega \cdot \phi(x_i) + b \tag{9}$$

where ω , and b are the weight vector and bias vector respectively. The nonlinear function ϕ maps the input x_i to a high-dimensional feature space. The SVR can be formulated as:

$$\min_{\omega,b} \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n} \ell_\varepsilon(f(x_i) - y_i) \tag{10}$$

where C is the penalty factor, $\varepsilon(\varepsilon > 0)$ is the maximum allowable regression error, and ℓ_ε is the sensitivity loss function, which is expressed as:

$$\ell_\varepsilon(z) = \begin{cases} 0, & |z| \le \varepsilon \\ |z| - \varepsilon, & otherwise \end{cases} \tag{11}$$

Then, the slack variables ξ_i and $\hat{\xi}_i$ are introduced and Eq.10 can be rewritten as:

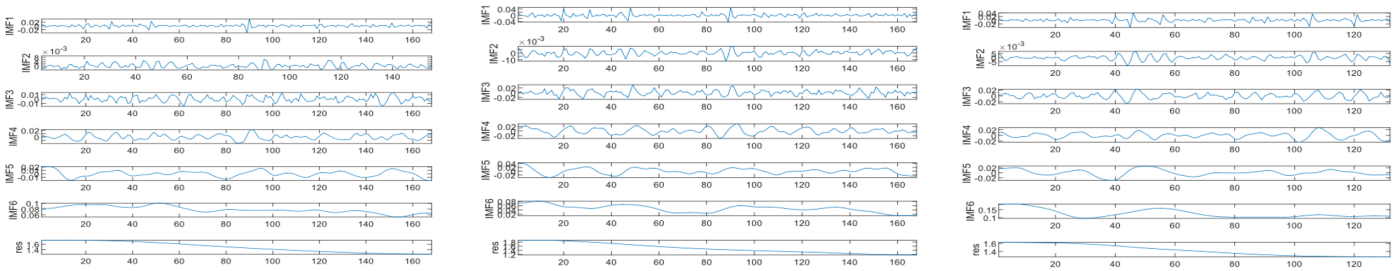
$$\min_{\omega,b,\xi_i,\hat{\xi}_i} \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}(\xi_i + \hat{\xi}_i) \tag{12}$$

$$\text{s.t.} \begin{cases} f(x_i) - y_i \le \varepsilon + \xi_i \\ y_i - f(x_i) \le \varepsilon + \hat{\xi}_i \\ \xi_i \ge 0, \hat{\xi}_i \ge 0, \quad i = 1, 2, ..., n \end{cases} \tag{13}$$

By introducing Lagrangian multipliers and kernel functions, the objective function is transformed into its dual form.

$$s.t. \begin{cases} \sum_{i=1}^{n}(a_i - \hat{a}_i) = 0 \\ 0 \le a_i, \hat{a}_i \le C, i = 1, 2, ..., n \end{cases} \tag{14}$$

where α_i and $\hat{\alpha}_i$ are Lagrange multipliers, and the kernel function can transform the inner product



(a) Decomposition results of B0005 (b) Decomposition results of B0006 (c) Decomposition results of B0018

Figure 2. Decomposition results based on CEEMDAN.

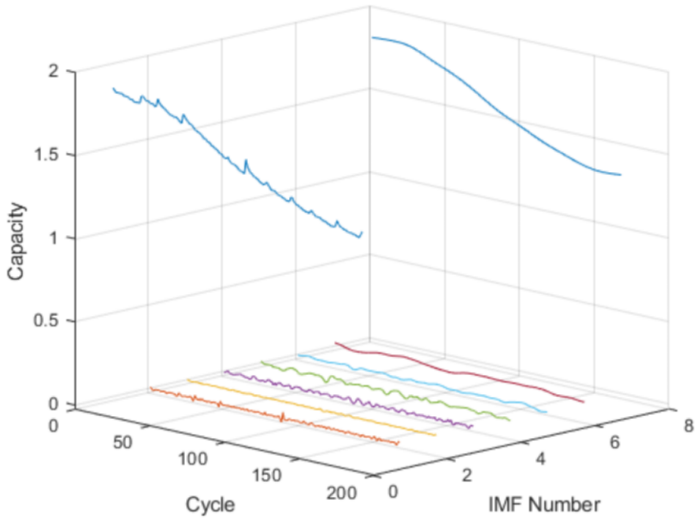


Figure 3. Decomposition diagram of B0005.

operations in the low-dimensional space into kernel function computations in the high-dimensional space. After minimizing the Lagrangian function, the SVR expression can be formulated as follows:

$$f(x) = \sum_{i=1}^{n} (a_i, \hat{a}_i) K(x_i, x) + b \tag{15}$$

3.2 Bidirectional Gated Recurrent Network

The gated recurrent unit (GRU), as a variant of recurrent neural network, comprises reset and update gates [18]. The reset gate determines the integration of new information with historical memory, while the update gate regulates the utilization of current state information relative to historical temporal data. The gating architecture enables efficient extraction of temporal features while maintaining lower parameter complexity and diminished computational demands [19]. The basic structure of GRU is shown in Figure 4 [20], and the unit output h_t at time t can be obtained as follows:

$$z_t = \sigma(W_z q_t' + U_z h_{t-1} + b_z) \tag{16}$$

$$r_t = \sigma(W_r q_t' + U_r h_{t-1} + b_r) \tag{17}$$

$$c_t = \tanh(W_c q_t' + r_t(U_c \cdot h_{t-1}) + b_c) \tag{18}$$

$$\mathrm{h}_t = c_t \odot (1 - z_t) + \mathrm{h}_{t-1} \odot z_t \tag{19}$$

where q_t' is the input capacity at time t, and z_t, r_t, c_t are the update gate output, reset gate output, and candidate variable, respectively. W_z, W_r and W_c denote the weights of the update gate, reset gate, and candidate variable with the input. U_z, b_z, U_r, b_r, U_c and b_c represent the weights and biases of the update gate, reset gate, and candidate variable, respectively. The tanh in Eq.18 is the hyperbolic tangent function, and $\odot$ in Eq.19 denotes the dot product operation.

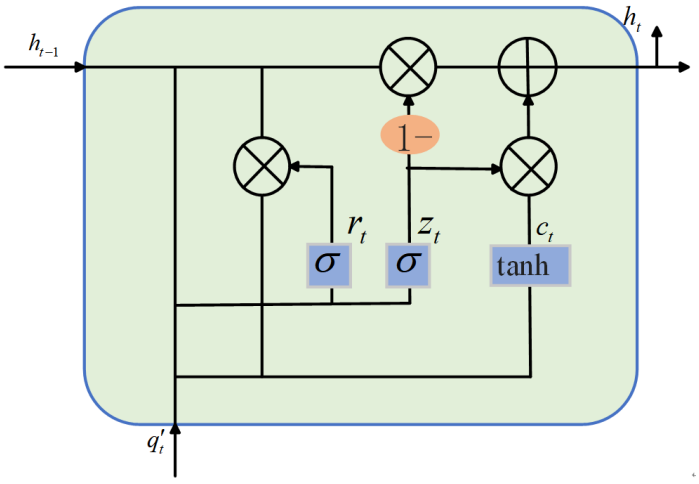


Figure 4. Gated recurrent unit.

To utilize the reverse capacity time series, a bidirectional GRU network is adopted, the forward and backward recurrent layers are used to acquire the hidden state through a splicing operation. On this basis, the input of the forward GRU is q_t', so that the forward output is obtained in the following Eq.20:

$$\vec{h}_t = \overrightarrow{GRU}(q_t', h_{t-1}), \quad t = 1, 2, ..., L \tag{20}$$

where $\overrightarrow{GRU}$ is the expression of the forward GRU, h_0 is initialized to 0. Correspondingly, the reverse output

is written as:

$$\overleftarrow{h_t} = \overleftarrow{GRU}(q'_t, h_{t-1}), \quad t = L, L-1, ..., 1 \tag{21}$$

where $\overleftarrow{GRU}$ is the expression of the backward GRU and the final hidden layer output $\hat{h}_t$ is defined as:

$$\widetilde{h_t} = \left[\overrightarrow{h_t}, \overleftarrow{h_t}\right] \tag{22}$$

Thus, the forward and backward lithium-ion battery capacity sequences are fused to promote the efficiency of information usage, improve the capacity prediction performance of the classical GRU. Therefore, the capacity $\hat{q}_{L+1}$ can be derived through hidden full-connected layers, which can be shown as:

$$\hat{q}_{L+1} = \eta\left(\tilde{h}_t\right) = \xi\left(W_u\left(...\xi\left(W_1\tilde{h}_t + b_1\right)...\right) + b_u\right) \tag{23}$$

where η is the mapping function of fully connected layers, ξ is the activation of each fully connected layer, W_1 and b_1 are the weight matrix and bias of the first fully connected layer, and W_u and b_u are the weight matrix and bias of the u-th fully connected layer, respectively.

3.3 Hyperparameter optimization with SSA

SSA is an intelligent optimization algorithm that simulates the foraging and anti-predation behaviors of sparrow populations. This algorithm classifies the population into three distinct roles i.e., discoverers, followers and vigilantes [21]. Additionally, it incorporates an alert mechanism to facilitate danger and population migration. Discoverers are highly adaptive individuals within the sparrow population whose primary function is to locate food sources and communicate the positions of these sources to follower individuals. Discoverers typically account for 10% and 20% of the total sparrow population. The position update for discovers is expressed as follows:

$$X_{ij}^{t+1} = \begin{cases} X_{ij}^t \cdot \exp\left(\frac{-i}{\alpha \cdot M}\right) & R < ST \\ X_{ij}^t + Q \cdot L & R \geq ST \end{cases} \tag{24}$$

where t denotes the current iteration number, X_{ij} is the position of the i-th sparrow in the j-th dimension, the variable α is defined as a uniformly distributed random variable over the interval $(0, 1]$. M is the maximum number of iterations, Q is a positive random number, L is a $1 \times j$ matrix. $R \in [0, 1]$ and $ST \in [0.5, 1]$ are defined as warning or safety value, respectively.

The position update of followers is as follows:

$$X_{ij}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{wj}^t - X_{ij}^t}{i^2}\right) & i > \frac{n}{2} \\ X_{pj}^{t+1} + \left|X_{ij}^t - X_{pj}^{t+1}\right| \cdot A^+ \cdot L & i \leq \frac{n}{2} \end{cases} \tag{25}$$

where X_{wj} denoting the iterated global worst position, X_{pj} is the optimal position in iteration. A is a $1 \times j$ matrix with values of either 1 or -1 and we have $A^+ = A^T\left(AA^T\right)^{-1}$.

Upon detecting danger, a sparrow initiates an alarm signal to summon the vigilante individuals within the population. Accordingly, its position is updated as:

$$X_{ij}^{t+1} = \begin{cases} X_{bj}^t + \beta\left(X_{ij}^t - X_{bj}^t\right) & f_i > f_g \\ X_{ij}^t + k \cdot \left(\frac{X_{ij}^t - X_{wj}^t}{(f_i - f_\omega) + \xi}\right) & f_i = f_g \end{cases} \tag{26}$$

where X_{bj} is the global optimal position, β is the step control parameter, k is a uniform random number and $k \in [-1, 1]$, ξ is the minimum actual number, and f_i and f_ω are the best and worst fitness values of the current sparrow population, respectively. The specific process is shown in Figure 5.

SVR and BiGRU models exhibit strong dependence on their hyperparameter configurations to achieve the optimal prediction performance. The most common way is to manual hyperparameter tuning according to empirical knowledge, which is time-consuming. Moreover, it is largely difficult to ensure the accuracy since there are some model parameters related to its architecture. Thus, the optimization for hyperparameters presents a nontrivial challenge. We propose the hyperparameter optimization of SVR and BiGRU models independently through SSA, with the objective of obtaining optimal models to enhance the prediction accuracy [22].

The hyperparameter optimization process can be divided into the following steps:

Step 1: Determine the hyperparameters in SVR and BiGRU to be optimized by SSA, which involve the penalty factor C and radial basis kernel function parameter σ in the SVR model, as well as parameters such as the number of network layers, number of neurons, and dropout rate in the BiGRU model. Meanwhile, initialize the SSA parameters.

Step 2: Select RMSE as the fitness function. RMSE (Root Mean Square Error) is defined as the square root of the average of the squared differences between predicted and actual values, as shown in Eq.27, where

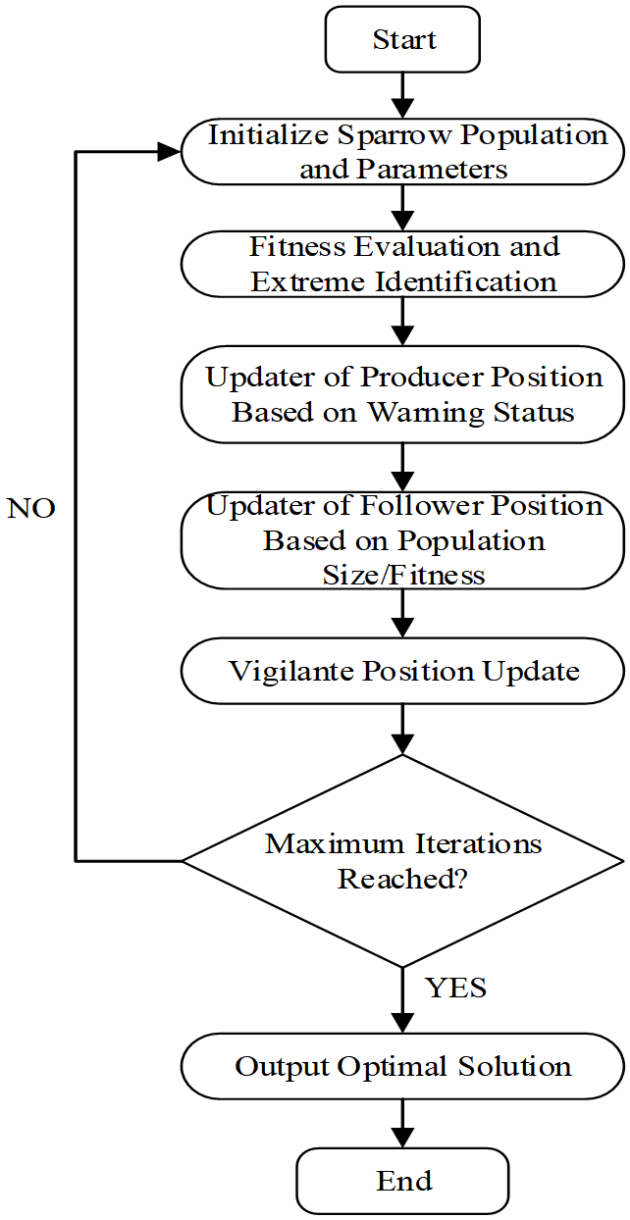


Figure 5. Flowchart of SSA.

C_t and $\tilde{C}_t$ represent the actual and predicted capacity values, respectively, and n is the length of the test sample. The initial fitness of each sparrow was computed, with the optimal fitness value and its corresponding sparrow position being preserved.

$$fitness = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(c_t - \tilde{c}_t)^2} \tag{27}$$

Step 3: The positions of discoverers, followers, and vigilantes are updated using Eq.24, 25, and 26, respectively.

Step 4: The fitness value of each individual is computed to evaluate its quality. The optimal individual from the current population is subsequently selected as the best solution for the present iteration.

Step 5: A termination check is performed to determine whether the optimal fitness has been achieved or the maximum number of iterations has been reached. If either condition is satisfied, the optimal hyperparameters are returned. Otherwise, return to Step 3.

3.4 CEEMDAN-SSA-SVR-BiGRU Prediction Model

The RUL prediction flowchart of lithium-ion battery based on the CEEMDAN-SSA-SVR-BiGRU is shown in Figure 6, and expressed as follows.

(1) During data preprocessing, the dataset is first partitioned as the training and test sets, and then normalized. The proposed CEEMDAN is adopted to decompose NASA battery training datasets and generate multiple IMFs and RESs.

The lithium-ion battery dataset is chronologically partitioned, with the first 50% allocated as the training set and the remaining 50% as the test set. The raw capacity data of the training set is normalized to the interval [0,1] using Eq.28:

$$x_{norm} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{28}$$

where x represents the raw data, while $x_{\min}$ and $x_{\max}$ denote the minimum and maximum values of the dataset, respectively. The training set was decomposed by employing the CEEMDAN method from the PyEMD library. This process yields multiple IMF components (IMF1, IMF2,..., IMFN) and one RES.

(2) Initialize SSA parameters, determine the hyperparameters with their value ranges for both SVR and BiGRU models and then optimize them via SSA.

The SSA parameters are initialized with a population size of 30, maximum iterations of 50, discoverer proportion of 10%, follower proportion of 80%, vigilante proportion of 10%, and warning value of 0.5. SSA is employed to optimize SVR and BiGRU respectively. For SVR, the radial basis function is selected as the kernel, with the kernel parameter and penalty factor defined as tunable parameters. For BiGRU, the number of hidden units, learning rate, batch size, dropout rate, and time steps are designated as the optimizable hyperparameters. The specific hyperparameter ranges are provided in Table 1.

(3) The IMF and RESs are fed into SVR and BiGRU models with the optimal hyperparameters for training,

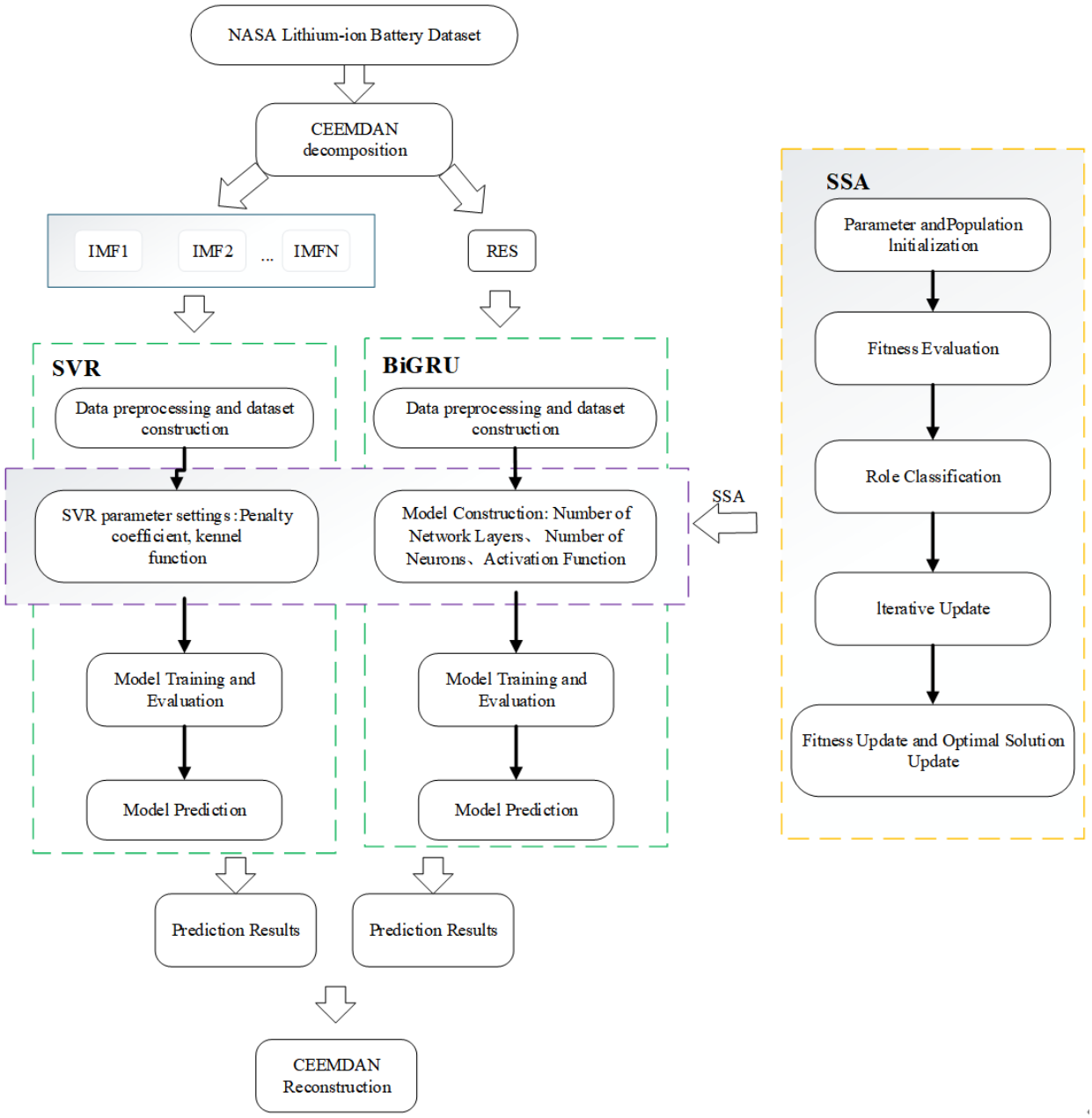


Figure 6. Prediction flowchart of CEEMDAN-SSA-SVR-BiGRU.

Table 1. Hyperparameter ranges for SSA optimization.

Hyperparameter	Range
Penalty Factor	[0.01, 100]
Kernel Parameter	[1e-4, 1]
Number of Hidden Layer Units	[16, 32, 64, 128]
Number of Layers	[1, 2, 3, 4]
Learning Rate	[1e-4, 5e-4, 1e-3, 5e-3, 1e-2]
Batch Size	[10, 20, 50, 100]
Dropout rate	[0.05, 0.10, 0.15, 0.20]
Time Step	[5, 10, 15, 20]

respectively. The training data for the SVR model is structured into supervised learning format using a sliding window approach with a window size of 10, where the previous 10 time steps are utilized to predict the next value. Each IMF component is normalized individually. The SVR model is trained with the optimal hyperparameters obtained by SSA to forecast IMF values from the historical sequences, effectively capturing nonlinear degradation characteristics while balancing fitting capability and generalization. For the BiGRU model, the training utilizes SSA-optimized hyperparameters, with input being three-dimensional residual sequences constructed by sliding windows (samples $\times$ time steps $\times$ features). The model employs bidirectional GRU layers for temporal feature extraction, Dropout layers for overfitting prevention, and the Adam optimizer to minimize the MAE

with adaptive learning rates. Early stopping is implemented, which implies that the training halts if the validation loss fails to be improved for 10 consecutive epochs, and GPU acceleration is applied. The final model captures long-term dependencies in RESs through multiple iterations.The predicted IMFs and RESs are reconstructed via CEEMDAN and denormalized to generate the final capacity prediction.

4 Experiment analyses and Comparisons

4.1 Data description

The proposed CEEMDAN-SSA-SVR-BiGRU prediction method was evaluated on the publicly available NASA battery dataset. To capture the aging process of lithium-ion batteries, the internal parameters such as capacity, along with operational data including current, voltage, and resistance, were measured across various operating conditions and charge-discharge cycles. The present study employs a publicly available NASA battery dataset to validate the effectiveness of the proposed prediction method. Table 2 summarizes the experimental operating conditions for all NASA batteries.

Here, we select the first group dataset and this dataset comprises four lithium-ion batteries (B0005, B0006, B0007, and B0018) with a rated capacity of 2Ah. All cycling tests were conducted under a constant ambient temperature of 24°C, with the specific charge-discharge processes. In the charging stage, a constant current-constant voltage (CC-CV) regime was implemented, commencing with a 1.5A constant current charge until the battery voltage reached 4.2V, followed by a constant voltage charge maintained until the current decayed to 20mA. In the discharging stage, discharge was performed at a constant current of 2 A until the voltage fell to the cut-off threshold. Since cell B0007 did not reach failure threshold during testing, this study exclusively utilizes data from B0005, B0006, and B0018 for experimental analysis.

4.2 Ablation Experiment

The RUL prediction experiment utilizes three lithium-ion battery datasets (B0005, B0006, and B0018) from the NASA public dataset. The end-of-life threshold is defined as 70% of the nominal capacity and equals to 1.4 Ah, with the prediction starting point set at 50% of the data length. To evaluate the contributions of different submodules, the proposed model is compared against four baseline models through ablation studies, i.e., SSA-SVR, SSA-BiGRU, CEEMDAN-SSA-SVR, and CEEMDAN-SSA-BiGRU. Figure 7 presents the RUL prediction results for the three battery units, where M1 denotes SSA-SVR, M2 represents SSA-BiGRU, M3 corresponds to CEEMDAN-SSA-SVR, M4 indicates CEEMDAN-SSA-BiGRU, and M0 refers to the proposed CEEMDAN-SSA-SVR-BiGRU model with the failure threshold set at 1.4 Ah. As shown in Figure 7, the prediction curve of model M0 achieves a closer alignment with the actual battery capacity values in comparison with the curves of other models.

4.3 Model Evaluation

To comprehensively evaluate the proposed method, MAE, MAPE, and RMSE are employed as evaluation metrics. Smaller values for these metrics indicate better predictive performance [23]. The corresponding calculation of these metrics are as follows:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|C_t - \widetilde{C_t}\right| \tag{29}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{C_t - \widetilde{C_t}}{C_t}\right| \tag{30}$$

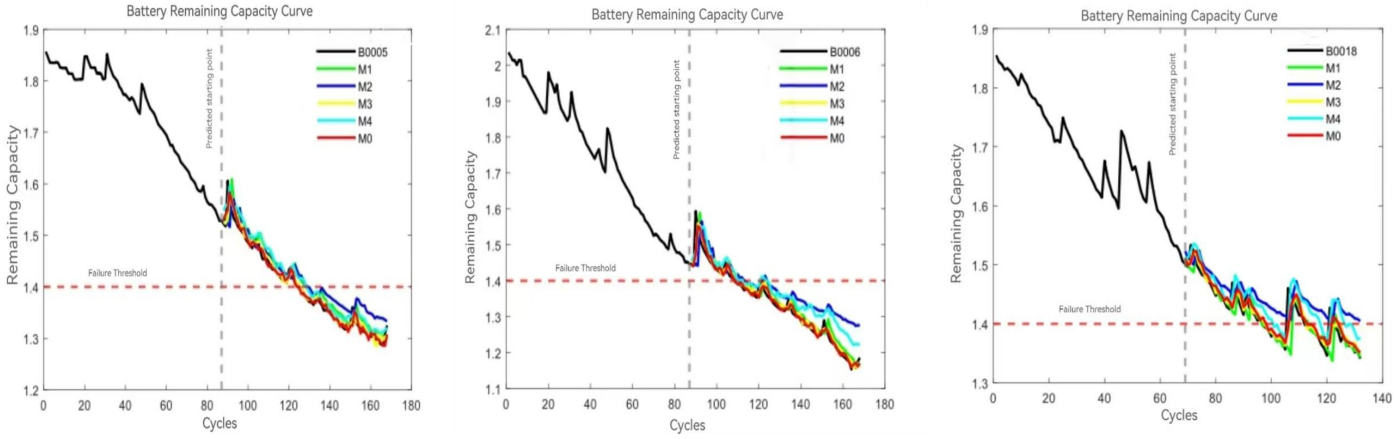
$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(C_t - \widetilde{C_t}\right)^2} \tag{31}$$

where C_i and $\overrightarrow{C_i}$ represent the actual and predicted capacity values, respectively; and n denotes the length of the test samples.

To conduct a quantitative assessment of prediction performance, the MAE, MAPE, and RMSE were calculated for all five models, and the corresponding results are given in Table 3. It can be noted that the CEEMDAN-SSA-SVR-BiGRU model achieves superior performance across all evaluation metrics compared to the other four models.

(1) Analysis of prediction results based on capacity decomposition

In Table 3, neither Model M1 nor Model M2 considers capacity decomposition, while the remaining three models perform RUL prediction based on capacity decomposition with CEEMDAN. It is evident that the prediction errors of the models M3, M4 and M0 are generally smaller than those of models without CEEMDAN. For the B0005 battery, model M3 achieves a relative reduction of 0.0090 in MAE, 0.0064 in MAPE and 0.0097 in RMSE compared with model M1. The underlying reason for this is that the capacity decomposition algorithm not only achieves effective



(a) Prediction outcome of B0005 (b) Prediction outcome of B0006 (c) Prediction outcome of B0018

Figure 7. Results of ablation experiments.

Table 2. Experimental conditions for NASA dataset.

Battery	Temperature/$^{\circ}C$	Charge current/A	Discharge current/A	Cut-off voltage/V
B05, B06, B07, B18	24	1.5	2	2.7, 2.5, 2.2, 2.5
B25, B26, B27, B28	24	1.5	4	2, 2.2, 2.5, 2.7
B29, B30, B31, B32	43	1.5	4	2, 2.2, 2.5, 2.7
B33, B34, B36	24	1.5	4, 2	2, 2.2, 2.7
B38, B39, B40	24	1.5	1, 2, 4	2.2, 2.5, 2.7

decomposition of lithium battery capacity data but also counteracts the impact of the capacity regeneration phenomenon on the accuracy of prediction results.

(2) Analysis of prediction results based on different methods

Table 3 further illustrates significant performance discrepancies among three models in consideration of CEEMDAN. In relation to battery B0005, model M0 attains a reduction of 0.0010, 0.0008, and 0.0008 across the three evaluation metrics when compared to M3, and a reduction of 0.0090, 0.0066, and 0.0088 in comparison with model M4. This superiority stems from the complementary strengths of SVR in predicting local capacity regeneration components and BiGRU in modeling global capacity degradation trends, thereby demonstrating the effectiveness of the integrated CEEMDAN-SSA-SVR-BiGRU framework for accurate lithium-ion battery RUL prediction.

4.4 Comparative Experiments

To further validate the effectiveness of the proposed model, a series of comparative experiments were conducted with three established RUL prediction methods(CNN-GRU [24], BiGRU-Transformer [25] and EEMD-GWO-SVR [11]) under the same datasets and consistent prediction starting points. The prediction results and corresponding errors are shown in Figures 8 and 9, respectively.

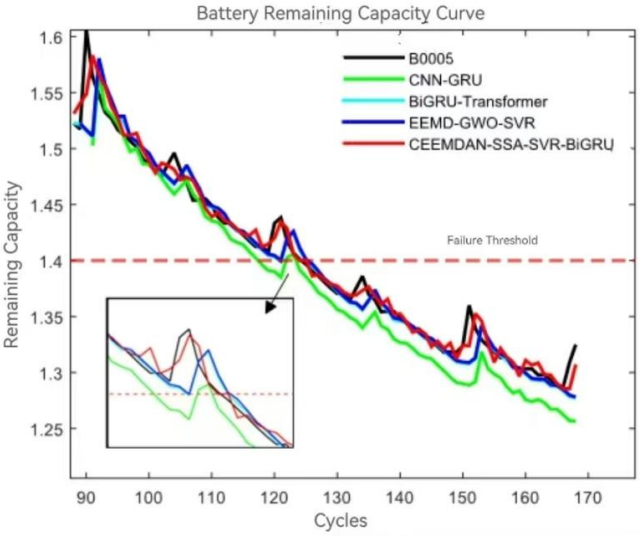


Figure 8. RUL prediction results of various models.

As visually depicted in Figures 8 and 9, the proposed model manifests significantly diminished prediction errors (the error is defined as the difference between the actual value and the predicted value) in handling capacity regeneration phenomena, as compared to other methods. This performance advantage arises from the CEEMDAN decomposition technique,

Table 3. Evaluation metrics of different models.

Battery	Model	MAE	MAPE	RMSE
B0005	M1:SSA-SVR	0.0157	0.0111	0.0184
	M2:SSA-BiGRU	0.0232	0.0169	0.0279
	M3:CEEMDAN-SSA-SVR	0.0067	0.0047	0.0087
	M4:CEEMDAN-SSA-BiGRU	0.0147	0.0105	0.0167
	M0:CEEMDAN-SSA-SVR-BiGRU	0.0057	0.0039	0.0079
B0006	M1:SSA-SVR	0.0143	0.0105	0.0223
	M2:SSA-BiGRU	0.0443	0.0347	0.0561
	M3:CEEMDAN-SSA-SVR	0.0071	0.0052	0.0091
	M4:CEEMDAN-SSA-BiGRU	0.0337	0.0261	0.0391
	M0:CEEMDAN-SSA-SVR-BiGRU	0.0067	0.0049	0.0093
B0018	M1:SSA-SVR	0.0104	0.0073	0.0219
	M2:SSA-BiGRU	0.0317	0.0226	0.0348
	M3:CEEMDAN-SSA-SVR	0.0071	0.0050	0.0102
	M4:CEEMDAN-SSA-BiGRU	0.0185	0.0132	0.0212
	M0:CEEMDAN-SSA-SVR-BiGRU	0.0067	0.0047	0.0094

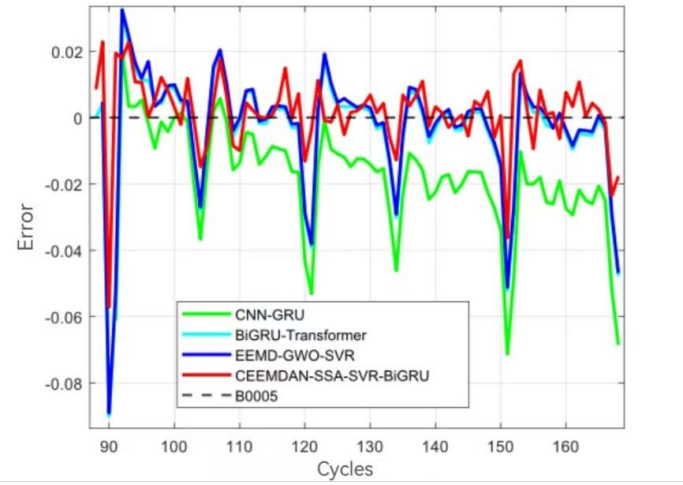


Figure 9. Error curves of various models.

which mitigates the negative impacts of capacity regeneration on prediction accuracy effectively. Although the EEMD-based decomposition strategy is taken into account in the EEMD-GWO-SVR model, it depends on a single predictive model SVR that fails to achieve optimal performance for all components. In contrast, our framework strategically integrates SVR for capturing local capacity regeneration trends with BiGRU for characterizing global capacity degradation trajectories. This synergistic integration enables substantial improvement in RUL prediction accuracy for lithium-ion batteries. Furthermore, the proposed methodology not only tackles the issues posed by capacity regeneration phenomena but also enhances the RUL prediction accuracy. A comparative evaluation conducted against five models–namely CNN-GRU, BiGRU-Transformer, EEMD-GWO-SVR,Temporal and Differential Attention Network (TDANet) [26], and Gaussian Process Regression Model Based on Ant-Lion Optimization Algorithm(ALO-GPR) [27]–confirms the superior predictive performance of our developed approach, with detailed MAE and RMSE results shown in Table 4. As is clearly illustrated in Table 4, the proposed model attains markedly lower MAE and RMSE values than the other five benchmark models. These results highlight the enhanced capability of our approach in dealing with datasets characterized by capacity regeneration phenomena, thereby confirming the effectiveness of the proposed model.

Table 4. MAE and RMSE of predictive models.

Prediction model	MAE	RMSE
CNN-GRU	0.0066	0.0087
BiGRU-Transformer	0.0080	0.0143
EEMD-GWO-SVR	0.0076	0.0096
TDANet	0.0101	0.0145
ALO-GPR	0.0060	0.0093
The proposed model	0.0057	0.0079

5 Conclusions

To address the challenge of improving RUL prediction accuracy for lithium-ion battery in the presence of capacity regeneration scenario, this paper develops a novel CEEMDAN-SSA-SVR-BiGRU

framework. The methodology first decomposes battery capacity degradation time-series data with CEEMDAN technique to extract distinct local capacity regeneration components and global capacity degradation components. These decomposed sub-components are subsequently predicted individually. An SSA-optimized SVR is employed for local capacity regeneration prediction, while a BiGRU model is utilized for global capacity degradation prediction. The final RUL prediction is obtained by reconstructing the predictive outputs of these individual components. Extensive comparisons assessments in ablation experiments demonstrate the superior performance of the proposed model across three battery units, which attains MAE, MAPE, and RMSE values of less than 0.0067, 0.0049, and 0.0094, respectively. Furthermore, the proposed model also exhibits a competitive advantage when compared against the previously well-established methods. The experimental findings confirm that the proposed CEEMDAN-SSA-SVR-BiGRU framework markedly elevates RUL prediction accuracy and provides a viable and effective strategy for the estimation of lithium-ion battery lifespan in practical applications.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported by the Hebei Natural Science Foundation under Grant F2023501011, and the National Natural Science Foundation of China under Grant 72271049.

Conflicts of Interest

The authors declare no conflicts of interest.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Jia, Z., Li, Z., Wang, Z., Liu, Z., & Vong, C. M. (2025). Joint prediction of SOH and RUL for lithium batteries considering capacity self recovery and model drift. *IEEE Internet of Things Journal*. [Crossref]

[2] Zhang, S., Liu, Z., Xu, Y., & Su, H. (2024). A physics-informed hybrid multitask learning for lithium-ion battery full-life aging estimation at early lifetime. *IEEE Transactions on Industrial Informatics*. [Crossref]

[3] Ma, Q., Zheng, Y., Yang, W., Zhang, Y., & Zhang, H. (2021). Remaining useful life prediction of lithium battery based on capacity regeneration point detection. *Energy, 234*, 121233. [Crossref]

[4] Zhang, J., Chen, J., Liu, D., He, L., Yang, K., Du, F., ... & Zhang, X. (2025). Multi-state joint prediction algorithm for lithium battery packs based on data-driven and physical models. *Energy, 322*, 135641. [Crossref]

[5] Fang, W., Chen, H., & Zhou, F. (2022). Fault diagnosis for cell voltage inconsistency of a battery pack in electric vehicles based on real-world driving data. *Computers and Electrical Engineering, 102*, 108095. [Crossref]

[6] Liu, Z., He, H., Xie, J., Wang, K., & Huang, W. (2022). Self-discharge prediction method for lithium-ion batteries based on improved support vector machine. *Journal of Energy Storage, 55*, 105571. [Crossref]

[7] Sun, J., Ren, S., Shang, Y., Zhang, X., Liu, Y., & Wang, D. (2023). A novel fault prediction method based on convolutional neural network and long short-term memory with correlation coefficient for lithium-ion battery. *Journal of Energy Storage, 62*, 106811. [Crossref]

[8] Tian, Y., Lu, C., Wang, Z., & Tao, L. (2014). Artificial fish swarm algorithm-based particle filter for Li-ion battery life prediction. *Mathematical Problems in Engineering, 2014*(1), 564894. [Crossref]

[9] Chang, Y., Fang, H., & Zhang, Y. (2017). A new hybrid method for the prediction of the remaining useful life of a lithium-ion battery. *Applied energy, 206*, 1564-1578. [Crossref]

[10] Chen, X., Liu, Z., Sheng, H., Wu, K., Mi, J., & Li, Q. (2024). Transfer learning based remaining useful life prediction of lithium-ion battery considering capacity regeneration phenomenon. *Journal of Energy Storage, 76*, 109798. [Crossref]

[11] Yang, Z., Wang, Y., & Kong, C. (2021). Remaining useful life prediction of lithium-ion batteries based on a mixture of ensemble empirical mode decomposition and GWO-SVR model. *IEEE Transactions on Instrumentation and Measurement, 70*, 1-11. [Crossref]

[12] Yao, F., He, W., Wu, Y., Ding, F., & Meng, D. (2022). Remaining useful life prediction of lithium-ion batteries using a hybrid model. *Energy, 248*, 123622. [Crossref]

[13] Xue, J., & Shen, B. (2020). A novel swarm intelligence optimization approach: sparrow search algorithm. *Systems science & control engineering, 8*(1), 22-34. [Crossref]

[14] Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., ... & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences,*

454(1971), 903-995. [Crossref]

[15] Zhang, C., Wang, S., Yu, C., Wang, Y., & Fernandez, C. (2023). A complete ensemble empirical mode decomposition with adaptive noise deep autoregressive recurrent neural network method for the whole life remaining useful life prediction of lithium-ion batteries. *Ionics, 29*(10), 4337-4349. [Crossref]

[16] Liu, Y., Sun, J., Shang, Y., Zhang, X., Ren, S., & Wang, D. (2023). A novel remaining useful life prediction method for lithium-ion battery based on long short-term memory network optimized by improved sparrow search algorithm. *Journal of Energy Storage, 61*, 106645. [Crossref]

[17] Balogun, A. L., Rezaie, F., Pham, Q. B., Gigović, L., Drobnjak, S., Aina, Y. A., ... & Lee, S. (2021). Spatial prediction of landslide susceptibility in western Serbia using hybrid support vector regression (SVR) with GWO, BAT and COA algorithms. *Geoscience Frontiers, 12*(3), 101104. [Crossref]

[18] Jia, S., Ma, B., Guo, W., & Li, Z. S. (2021). A sample entropy based prognostics method for lithium-ion batteries using relevance vector machine. *Journal of Manufacturing Systems, 61*, 773-781. [Crossref]

[19] Zhao, F., Gao, W., Lu, J., Jiang, H., & Shi, J. (2024). Real-time concentration detection of Al dust using GRU-based Kalman filtering approach. *Process Safety and Environmental Protection, 189*, 154-163. [Crossref]

[20] Zhang, J., Huang, C., Chow, M. Y., Li, X., Tian, J., Luo, H., & Yin, S. (2023). A data-model interactive remaining useful life prediction approach of lithium-ion batteries based on PF-BiGRU-TSAM. *IEEE Transactions on Industrial Informatics, 20*(2), 1144-1154. [Crossref]

[21] Wang, Z., Wang, Q., Liu, Z., & Wu, T. (2024). A deep learning interpretable model for river dissolved oxygen multi-step and interval prediction based on multi-source data fusion. *Journal of hydrology, 629*, 130637. [Crossref]

[22] Wen, J., Jia, C., & Xia, G. (2025). State of health prediction of lithium-ion batteries for driving conditions based on full parameter domain sparrow search algorithm and dual-module bidirectional gated recurrent unit. *arXiv preprint arXiv:2505.17405*.

[23] Fan, T. E., Wang, K., & Qu, B. (2025). Joint estimation of Li-ion battery states based on parallel TCN-Transformer multi-task learning model. *Journal of Energy Storage, 134*, 118077. [Crossref]

[24] Gong, Q., Wang, P., Cheng, Z., & Zhang, J. A. (2021). A method for estimating state of charge of lithium-ion batteries based on deep learning. *Journal of The Electrochemical Society, 168*(11), 110532. [Crossref]

[25] Jia, C., Tian, Y., Shi, Y., Jia, J., Wen, J., & Zeng, J. (2023). State of health prediction of lithium-ion batteries based on bidirectional gated recurrent unit and transformer. *Energy, 285*, 129401. [Crossref]

[26] Wang, T., Ma, Z., & Zou, S. (2023). Remaining useful life prediction of lithium-ion batteries: A temporal and differential guided dual attention neural network. *IEEE Transactions on Energy Conversion, 39*(1), 757-771. [Crossref]

[27] Liu, S., Sun, C., Sun, B., Fang, L., & Li, D. (2025). A two-layer full data-driven model for state of health estimation of lithium-ion batteries based on MKRVM-ELM hybrid algorithm with ant-lion optimization. *Journal of Energy Storage, 114*, 115716. [Crossref]



Fei Zhao is currently an associate professor from Northeastern University at Qinhuangdao. Her research interests include system reliability, maintenance, and operations research and optimization. (Email: zhaofei@neuq.edu.cn)



Xinyu Dai is currently pursuing a degree in Control Science and Engineering at Northeastern University at Qinhuangdao. His research focuses on the prediction of the remaining useful life of lithium-ion batteries. (Email: 15165077158@163.com)