Heliyon 9 (2023) e17754

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Early prediction of remaining useful life for lithium-ion batteries based on CEEMDAN-transformer-DNN hybrid model

Yuxiang Cai [a,b,1], Weimin Li [b,1], Taimoor Zahid [c], Chunhua Zheng [b], Qingguang Zhang [a,b,1], Kun Xu [b,*]

[a] Department of Materials Science and Engineering, Southern University of Science and Technology, 518055, Shenzhen, China
[b] Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, 518055, Shenzhen, China
[c] College of Electrical and Mechanical Engineering, National University of Sciences and Technology, Pakistan

ARTICLE INFO

Keywords:
Early prediction
RUL
Lithium-ion batteries
Capacity regeneration
CEEMDAN
Transformer
Deep neural networks

ABSTRACT

A reliable and safe energy storage system utilizing lithium-ion batteries relies on the early prediction of remaining useful life (RUL). Despite this, accurate capacity prediction can be challenging if little historical capacity data is available due to the capacity regeneration and the complexity of capacity degradation over multiple time scales. In this study, data decomposition, transformers, and deep neural networks (DNNs) are combined to develop a model of RUL prediction for lithium-ion batteries. Complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) is used for battery capacity sequential data to account for the capacity regeneration effect. The transformer networks are leveraged to predict each component of capacity regeneration thus improving the model's ability to handle long sequences while reducing the amount of data. The global degradation trend is predicted using a deep neural network. We validated the early prediction performance of the model using two publicly available battery datasets. Results show that the prediction model only uses 25%-30% data to achieve high accuracy. In the two public data sets, the RMSE errors were 0.0208 and 0.0337, respectively. A high level of accuracy is achieved with the model proposed in this study, which is based on fewer capacity data.

1. Introduction

Lithium-ion batteries have attracted the new energy industry due to their higher energy density, longer cycle life, and no memory effect. There are many electronic devices and many types of equipment that use lithium-ion batteries, such as cell phones, computers, electric vehicles, and intelligent power grids, which already use lithium-ion batteries to store energy [1–3]. However, lithium-ion batteries face performance degradation in practical applications because the internal chemical reaction of batteries causes increased internal resistance [4], capacity decay, and other performance-degradation problems [5]. Thus, battery health gradually declines along with capacity degradation [6]. For electric vehicles, batteries must be replaced or repaired when their capacity drops below a certain value, e.g., 80%. Otherwise, a battery can fail and even lead to a serious accident [7]. So advanced battery health and life management are essential for ensuring battery safety and reliability [8]. RUL is an essential health indicator for battery life prediction

* Corresponding author.
E-mail address: kun.xu@siat.ac.cn (K. Xu).
[1] These authors contributed to the work equally and should be regarded as co-first authors.

https://doi.org/10.1016/j.heliyon.2023.e17754
Received 10 April 2023; Received in revised form 22 June 2023; Accepted 27 June 2023

Available online 3 July 2023

that is described as the number of cycles from the current time to the failure threshold [9–11]. Generally, battery aging models are developed to simulate the decay of battery life and to access the internal relationship between the input sequence and life cycles to accurately determine the life remaining of the battery.

In the existing studies, the prediction approaches of RUL are divided into two types: model-based and data-driven [12].

Model-based approach studies the physicochemical phenomena and degradation factors during the operation of the batteries and builds interpretable mathematical prediction models to capture the aging trend of batteries. The filtering algorithm is commonly used as the modeling approach. Yang et al. [13] proposed an optimal particle resampling strategy combined with an unscented particle filter for RUL prediction. Tang et al. [14] developed a prediction model using the unscented Kalman filter with varying weights calculated by the particle filter. By combining an exponential model and particle filter, Zhang et al. [15] achieved better prediction performance than the autoregressive moving average model. Using models-based approaches does not require extensive historical degradation data. However, batteries' complex and dynamic electrochemical mechanisms make it difficult to model the aging performance accurately during multiple charge and discharge cycles. Many electrochemical parameters are difficult to obtain for building a model with high accuracy, leading to a high model complexity. This means that the model needs to be improved while being simplified for real-world applications.

The data-driven methods mainly rely on artificial intelligence and statistical theory to capture the hidden information of life aging and predict the remaining life cycles by utilizing the battery's historical charge and discharge data. The data-driven methods don't apply the electrochemical properties of the lithium-ion battery, making them more suitable for practical applications. Several data-driven methods have been studied for RUL prediction, such as naive Bayes [16], relevance vector machine [17], gaussian process regression [18,19], and deep learning [20,21], etc. Data-driven RUL models can achieve high accuracy but rely on large amounts of historical data to train long-term prediction models. The data-driven approach described above typically requires over 50% of the total capacity data for accurate RUL prediction. Because of the lack of capacity data, these methods may not predict the RUL accurately at early stages. Another significant challenge affecting the model performance arises from the capacity regeneration phenomenon. This phenomenon of transient capacity regeneration makes data-driven models sensitive to input data, leading to substantial deviations from the actual capacity degradation curve. Therefore, developing an efficient early RUL prediction model using fewer data can significantly facilitate achieving high accuracy in actual applications.

Yang et al. [22] develop an RUL predictive model which combines signal decomposition and artificial intelligence. By combining the prediction model with EEMD, the capacity regeneration phenomenon is better predicted. However, it requires a large volume of historical data. Based on a hybrid model, Tong et al. [23] propose a model to predict RUL that uses both adaptive drop-outs long short-term memory and Monte Carlo simulation. Although this method achieves accurate prediction with only 25 percent degradation data, it only identifies the overall trend of capacity degradation, leaving out points where capacity can be restored. An integrated two-step method is suggested by Ma et al. to estimate the RUL [24]. First, the cycle life of selected LIBs was estimated on the basis of a double exponential model (DEM) using a convolutional neural network. As a second step, the determined DEM is used as an initial mean function to GPR for early RUL prediction. In comparison to CNN-based and GPR-based RUL prediction methods, the results show superior performance. However, the model's performance needs to be improved.

This research proposes a hybrid model for RUL prediction to strengthen the model's ability to deal with the transient capacity regeneration phenomenon and improve prediction accuracy at an early stage. By decomposing the input capacity sequences with CEEMDAN, we obtain several components representing capacity regeneration and a residual signal representing global capacity degradation. Then, the prediction networks are developed to predict the components separately using transformer networks and the residual trend using deep neural networks, respectively. Lastly, we combine all the predictions and sum them up to get the prediction result of RUL.

2. The CEEMDAN-transformer-DNN hybrid model

2.1. The structure and workflow of the hybrid model

The proposed hybrid model combines a signal decomposition unit using CEEMDAN, multiple prediction networks for capacity regeneration intrinsic mode functions (IMFs) based on the transformer, and a prediction model based on deep neural networks (DNNs) for the degradation trend residue (RES). The structure of the CEEMDAN-Transformer-DNN model is presented in Fig. 1. The hybrid RUL prediction model obeys the following steps.

Step 1: Extracting the capacity sequences from the battery charge and discharge cycling data.

Step 2: Implementing the signal decomposition of the battery capacity sequences using CEEMDAN and obtaining the IMFs for local capacity regeneration and the RES for global capacity degradation.

Step 3: Building the prediction models for the IMFs and the RES, respectively. Transformer networks are used for local prediction models of IMFs, and the DNN is used for the prediction model of the global RES. These data are divided into two batches for training and testing. Afterward, the model is updated using the optimal parameters that were obtained through model training.

Step 4: Combining the prediction results from the transformer networks of IMFs and the DNN of RES to acquire the RUL prediction result on the basis of the CEEMDAN-Transformer-DNN hybrid model.

2.2. Decomposing capacity sequences using CEEMDAN

The CEEMDAN is based on empirical mode decomposition (EMD), an improved signal decomposition methodology. The EMD method was proposed to handle time series data that are non-stationary [25]. In EMD, a complex signal is decomposed into an

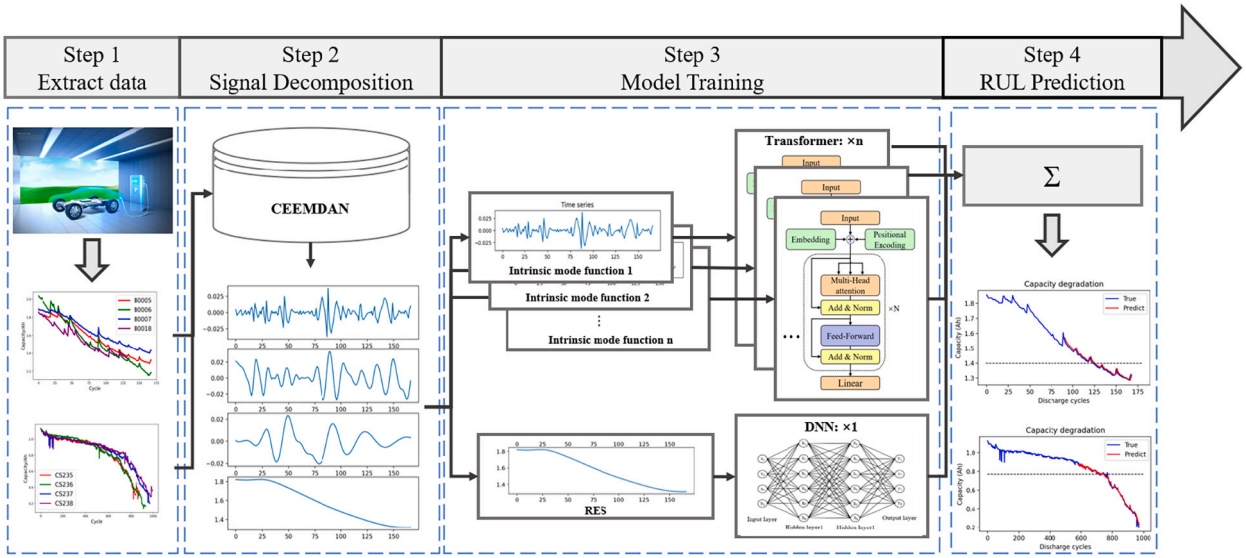

Fig. 1. RUL prediction framework.

infinite number of basis functions, called IMFs, together with a convex signal to indicate the global trend, known as RES. The IMFs derived from the decomposition of real-world signals may reflect the natural processes from various causes that may occur at different time intervals. Thus, EMD is adaptive and can be applied to analyze nonlinear and nonstationary time series.

In practical applications, due to noise, interference, or intermittent signals, the IMFs obtained from EMD may encounter the difficulty of mode mixing, where one IMF contains vastly different time scales or the same time scale is distributed in multiple IMFs. To solve the mode mixing issue, EEMD was proposed that incorporate white noise to collect comparable scales into a single IMF, leveraging the frequency domain characteristics of white noise [26]. However, the decomposition of EEMD is incomplete, and the noise is difficult to eliminate. The CEEMDAN is proposed to address the above problems in EEMD [27], which adds adaptive white noise to the decomposition.

As the capacity sequences contain many local capacity regeneration behaviors with different time-scale features, we leverage the signal decomposition algorithm, i.e., the CEEMDAN, to acquire the sequences representing the capacity regeneration phenomenon and a RES. With CEEMDAN, mode mixing caused by EMD can be solved, and noise generated by EEMD can be effectively removed. The battery capacity sequences, i.e., the input data, are fed into the CEEMDAN block for signal decomposition as shown in the steps below.

Step 1: Reconstructing the original signal $y(t)$ by incorporating gaussian white noise. To obtain the first-order IMF C_1, EMD is applied to the reconstructed signal in the equation (1):

$$E(y(t)+(-1)^q \epsilon \upsilon^j(t)) = C_1^j(t) + r^j, \tag{1}$$

where $E(t)$ is the i-th IMF obtained after EMD, $y(t)$ is the original input signal that requires to be decomposed, υ^j is a gaussian white noise which satisfies the standard normal distribution, $q = 1, 2$, $j = 1, 2, 3...$, ϵ is the weight coefficient of the gaussian white noise. Using CEEMDAN as a decomposition method, we obtain the i-th intrinsic mode function as $C_i^j(t)$. r^j is the residual generated by EMD.

Step 2: The first IMF of CEEMDAN is obtained by averaging $C_i^j(t)$. For the second EMD, we calculate the signal $r_1(t)$ by subtracting the original signal $y(t)$ from the first IMF. It is shown in equation (2):

$$\begin{aligned} C_1(t) &= \frac{1}{N}\sum_{j=1}^{N} C_1^j(t), \\ r_1(t) &= y(t) - C_1(t). \end{aligned} \tag{2}$$

Step 3: The first-order mode component $D_1(t)$ of a sequential signal is obtained by incorporating gaussian white noise to $r_1(t)$ and using EMD to decompose the new signal. By averaging $D_1(t)$ and calculating the difference between the initial signal and the second IMF, we obtain $r_2(t)$ for the second IMF of the CEEMDAN. The detailed process is shown in the equation (3):

$$\begin{aligned} C_2(t) &= \frac{1}{N}\sum_{j=1}^{N} D_1^j(t), \\ r_2(t) &= r_1(t) - C_2(t). \end{aligned} \tag{3}$$

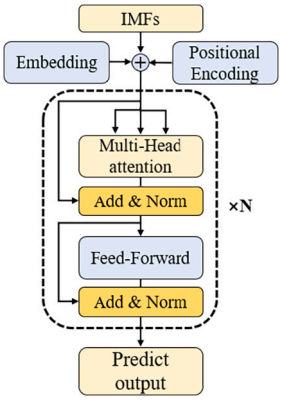

Fig. 2. Transformer encoder structure.

Step 4: Repeating the above steps until obtaining a monotonic function. Finally, the capacity sequence is decomposed into multiple IMFs and RES. We can build the prediction models for each component separately.

2.3. IMFs prediction model based on transformer

The RUL prediction is a problem of predicting long sequences based on the battery time-series data, i.e., the capacity sequences, which can be addressed by recurrent neural networks (RNNs) [28] and long short-term memory networks (LSTMs) [29]. Nevertheless, RNNs suffer from vanishing and exploding gradients when predicting long-time sequences. While LSTM is more capable of handling long-term sequences compared with RNN, it still requires a lot of historical data to train, which makes it unsuitable for early RUL prediction.

The transformer is a new deep learning framework proposed in 2017 [30]. When it comes to processing time series problems, the transformer is two times more efficient than traditional neural networks such as RNN and LSTM due to its self-attention and feed-forward layers. It significantly enhances the speed of the model operation and improves the precision of prediction. Due to the lack of historical capacity data, the model predicts the IMFs after CEEMDAN with an attention-mechanism-based transformer to improve long-term dependence and solve its early prediction problem. IMFs have a cyclical pattern of assignment fluctuations, so the transformer is a suitable choice. This study uses the encoder of the transformer to predict IMFs since it is a numerical sequence prediction. The main structure of the transformer encoder layer is shown in Fig. 2, which consists of the following parts.

Positional encoding: To encode each vector's position, sine, and cosine functions are added to the input data to produce a higher dimensional feature vector. The positional encoding function is shown in equation (4):

$$\begin{aligned} PE(t,2k) &= sin(t/(10000^{2k/m}), \\ PE(t,2k+1) &= cos(t/(10000^{2k/m}), \end{aligned} \tag{4}$$

where $PE(t,k)$ is the positional encoding function, t is the time, k is a constant number, and m is the dimension of the positional encoding.

Multi-Head attention: It is composed of multiple self-attention layers superimposed [31]. In multi-head attention layers, each vector will be multiplied with W_Q, W_K, and W_V, then acquire the $Query$, Key, and $Value$. Lastly, the similarity score of each vector will be calculated. Equation (5) shows how self-attention is calculated [30]:

$$Attention(Q,K,V) = softmax(\frac{QK^T}{\sqrt{d_K}})V, \tag{5}$$

among them, d_K is the dimension of Q matrix and K matrix. To avoid an excessively large inner product, QK^T is divided by $\sqrt{d_k}$, and the inner product is normalized using the softmax function and multiplied with the value matrix to obtain the output of self-attention.

Feed-Forward neural network: This network has a wide range of applications that contains two main layers. The first layer applies RELU as an activation function, while the second layer is linear. The expression of the function for this network is in the equation (6):

$$FFN(x) = Relu(xW_1 + b_1)W_2 + b_2, \tag{6}$$

where W_1, b_1 are the weights and deviations of the activation function layer, and the linear layer's weights and deviations are W_2 and b_2.

2.4. DNN-based RES prediction model

The deep neural networks (DNNs) can have more than one hidden layer besides inputs and outputs. DNNs can learn more complex and abstract high-level features than shallower networks thanks to the deep feature hierarchy from multiple neurons and

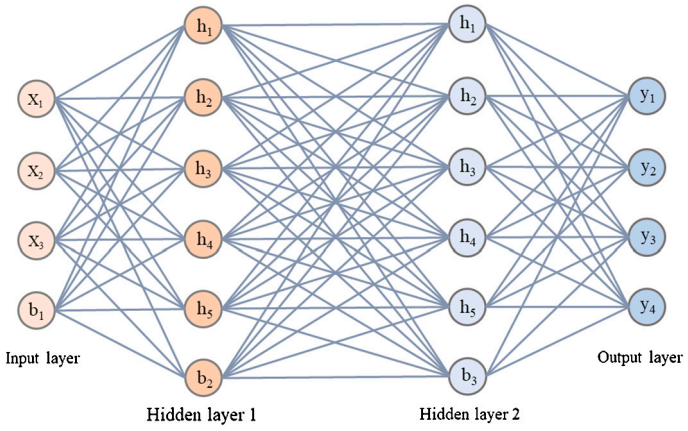

Fig. 3. Example of DNN model.

layers. Therefore, DNNs are capable of diverse tough tasks with superior performance. We choose DNN to predict the RES due to its excellent single-feature regression ability.

The number of neurons and hidden layers can be adjusted according to the task requirements. An example of a typical DNN model can be found in Fig. 3, in which $\{x_1, x_2, x_3\}$ represents the input, $\{b_1, b_2, b_3\}$ represents the biases between the input and hidden layers, $\{h_1, h_2, h_3, h_4, h_5\}$ represents each layer's neuron nodes, and $\{y_1, y_2, y_3, y_4\}$ represents the output.

The forward propagation process of DNN from the input layer to the hidden layer is calculated as equation (7):

$$h_j = \sum_{i=1}^{n} f(x_i \omega_{ij} + b_k) \quad i = 1, 2, ..., n \quad j = 1, 2, ..., n \quad , \tag{7}$$

where, F is the activation function of the network, x_i represents the input data, i represents the dimension of input features, ω_{ij} represents the weight between input layer and hidden layer, b_k is deviation of the layer, k is the layer number of the current network, h_j is the value of the nodes, and j is the total number of neurons in each layer.

The forward propagation from the hidden layer to the hidden layer is also similar to equation (7), in which the input is replaced with h_j in the hidden layer. This process is shown in equation (8):

$$y_m = \sum_{i=1}^{n} f(h_j \omega_{jm} + b_k) \, i = 1, 2, ..., n \, j = 1, 2, ..., n \quad , \tag{8}$$

among them, ω_{jm} is the weight between the hidden layer and the output layer, y_m is the output of the network, and m is the number of output features.

3. Experimental design

In this study, the capacity sequences were decomposed using EMD, EEMD, and CEEMDAN to compare the precision of different data decomposition methods. Then, the data after decomposition were used as the input to the hybrid model for RUL prediction. Further, we assessed the accuracy of the model under the same experimental conditions by comparing its predictions with those of other studies [32]. Finally, the model's performance was tested for RUL early prediction. The following sections discuss the datasets for the test, the neural network parameters, and the criteria for evaluating the neural network's performance.

3.1. Datasets

We used two public battery datasets as a basis for testing the feasibility of the model, namely those from the NASA Ames Predictive Center of Excellence and the Advanced Life Cycle Engineering Center of the Maryland University (CALCE) [33,34]. Scholars have widely leveraged the two battery datasets for verifying and evaluating the performance of the advanced battery algorithms because of their strong feasibility and applicability.

The NASA dataset provides nine batches of battery aging data sets from multiple aging tests using 18650 batteries with a standard capacity of 2.0 Ah. This study compares the precision of the proposed RUL prediction model against one of the batch numbers 5, 6, 7, and 18. The batteries were tested at a temperature of 24 °C, consisting of three steps: charging, discharging, and impedance testing. Detailed descriptions of the experimental steps follow.

Step 1: Continue charging batteries in the constant voltage till the voltage rises to 4.2 V, after charging in the constant current mode till the current falls to 20 mA.

Step 2: Using the constant current mode of 2A, discharge batteries till voltage with numbers 5, 6, 7, and 18 falls to 2.7 V, 2.5 V, 2.2 V, and 2.5 V.

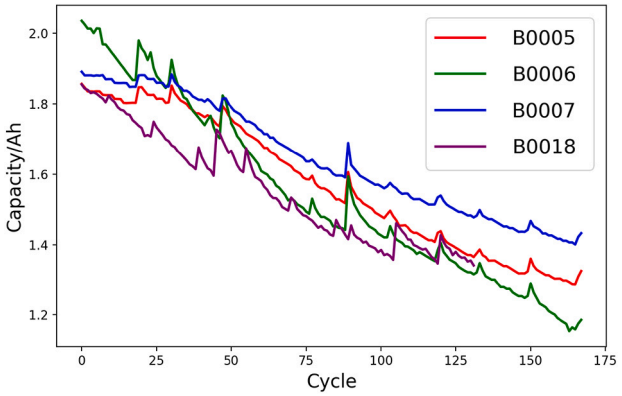

Fig. 4. Capacity curves of NASA dataset.

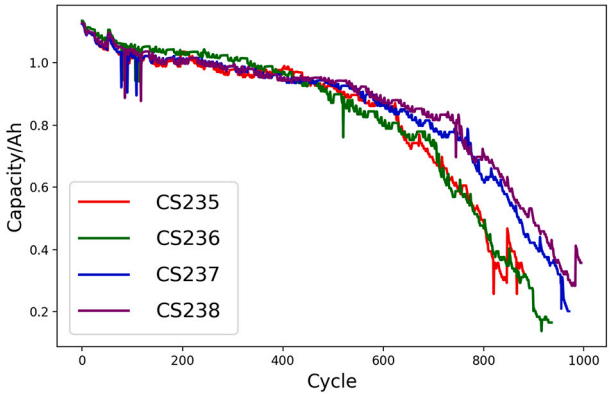

Fig. 5. Capacity curves of CALCE dataset.

Step 3: In the frequency range of 0.1 Hz to 5 kHz, electrochemical impedance spectroscopy was performed.

Step 4: Tests should be halted when the battery finishes its service life, e.g., when its rated capacity drops by 30%, or from 2.0 Ah to 1.4 Ah.

University of Maryland's CALCE battery dataset contains data on batteries with a standard capacity of 1.1 Ah. This study evaluates the performance of our proposed early prediction method with battery numbers 35, 36, 37, and 38. These batteries were tested at 1 ˚C. The main steps are the same as those described in the NASA dataset. The test's termination condition was a capacity drop from 1.1 Ah to 0.77 Ah.

The battery RUL curves for the NASA dataset are shown in Fig. 4, and the CALCE dataset is shown in Fig. 5.

3.2. Model settings

This prediction model uses historical capacity sequences as input and residual useful life as output. The data size of historical cycles is critical for the early prediction model. According to most existing studies, the starting point for predicting battery life is generally 50%-70% of the total cycles since the battery was first charged.

We use the same starting point and the same datasets as other models to compare the accuracy. Then, we use fewer input data than the existing RUL prediction methods to validate the early prediction performance. As input data, we use only 25% cycles from NASA and 30% cycles from CALCE. Table 1 shows the parameters of the transformer and DNN models.

There are usually two training methods in the time-series model training process: single-step and multi-step, as shown in Fig. 6. We define the sliding window size and then use the training sets to generate multiple sequential capacity data sets *train_x* according to the sliding window size. For the single-step prediction, we use the next capacity value after the sliding window data sequence as *train_y*. For the multi-step prediction, we use the future capacity sequence following the sliding window as *train_y*.

As the single-step method may bring large error accumulation in the long-time prediction, we use the multi-step (e.g., ten-step) method for training and testing. The prediction model uses the generated capacity data segments with the sliding window size as the input data to produce the predicted output capacity values ten steps ahead, which are then used to generate the input sliding window for the next iteration. As the sliding window moves forward, the multi-step prediction process continues iteratively until all the capacity values of the remaining cycles are predicted. Both neural networks are trained using root mean square error (RMSE) as their loss function.

Table 1
Model parameters.

Parameters	Transformer	DNN
Learning rate	0.0001	0.001
Hidden layer	2	4
Number of neurons	256	256
Encoder layer	2	none
Self-attention heads	10	none
Optimizer	ADAM	ADAM
Activation function	RELU	RELU
Sliding window size	20(NASA)	20(NASA)
	200(CALCE)	200(CALCE)
Train size	45(NASA)	45(NASA)
	300(CALCE)	300(CALCE)

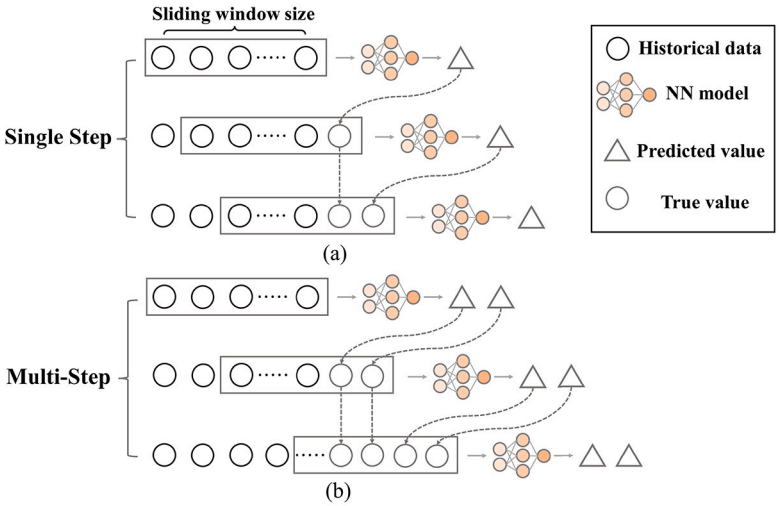

Fig. 6. Output method of the model. (a): single step (b): multi-step.

For the RUL prediction, it requires setting a starting point. Data from historical capacity with the defined sliding window size is the input to the prediction network, just as it is to the training process. The capacity data after the starting point can be predicted iteratively by forwarding the sliding window for each model as in equation (9):

$$\begin{aligned} y^j_{k+1:k+n} &= N^j(x_{k-m+1}, x_{k-m+2}, ..., x_k), \\ y^j_{k+2:k+n+1} &= N^j(x_{k-m+2}, x_{k-m+3}, ..., x_k, y_{k+1}), \\ &... \end{aligned} \tag{9}$$

where y is the output sequence, k represents the cycling number of the starting point, n is the future prediction steps, N^j is the $j-th$ prediction model, x is the actual input data, and m is the sliding window size.

3.3. Evaluation metrics

As performance metrics, absolute error (AE), absolute error rate (AER), and root mean square error (RMSE) are selected, which are extensively applied in the study of RUL prediction. The evaluation metrics are described below.

The metric AE is defined as equation (10):

$$AE = |Cycle_T - Cycle_P|, \tag{10}$$

the metric RMSE is defined as equation (11):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i}^{n} |CapT_i - CapP_i|^2}, \tag{11}$$

and the metric AER is defined as equation (12):

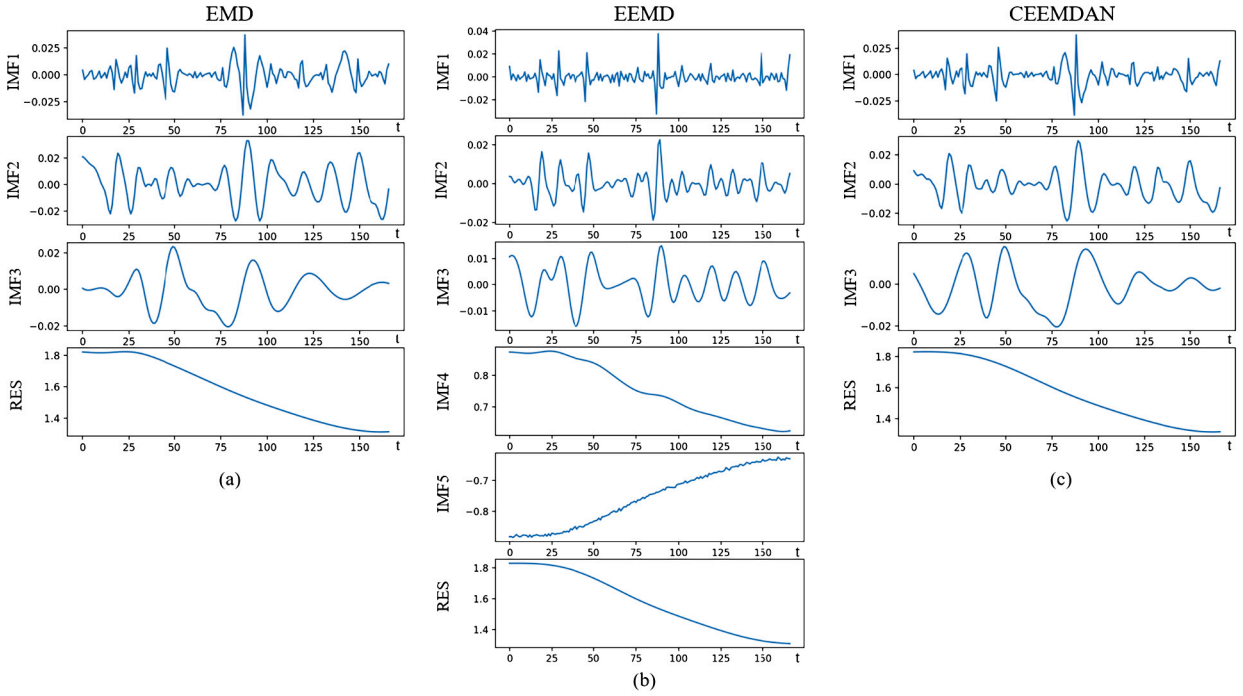

Fig. 7. Decomposition results of the capacity of B0005. (a): EMD (b): EEMD (c): CEEMDAN.

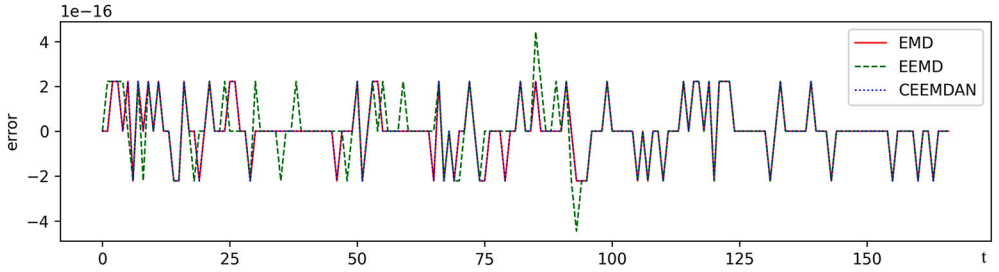

Fig. 8. Error curves of EMD, EEMD, CEEMDAN of B0005 capacity.

$$AER = \frac{|Cycle_T - Cycle_P|}{Cycle_T}, \tag{12}$$

where $Cycle_T$ represents the true number of cycles, $Cycle_P$ denotes the predicted number of remaining cycles, $CapT_i$ indicates the true capacity, and $CapP_i$ indicates the predicted value of the capacity.

4. Results and discussion

4.1. Comparison of the signal decomposition

The capacity sequence of battery B0005 is decomposed using three signal decomposition methods, i.e., EMD, EEMD, and CEEMDAN. Fig. 7 shows the signal decomposition results. Fig. 8 shows the error curves for the decomposition results of the three methods. From Fig. 7 (a) with the EMD method, we can see that mode mixing occurs between IMF1 and IMF2, where IMF2 contains components from the frequency band of IMF1. Another mode of mixing occurs between IMF3 and RES, where RES also contains components from the frequency band of IMF3. The above mode mixing renders the IMFs physically meaningless and leads to an inaccurate decomposition of the capacity data. The RMSE of the EMD method is 1.564×10^{-32}.

From Fig. 7 (b) with the EEMD method, the mode mixing phenomenon is greatly eliminated, as the frequency bands of IMFs are clearer than the EMD method. However, as the white noise was added several times during the decomposition, IMF4 and IMF5 could not be eliminated during the ensemble averaging, resulting in a redundant EEMD. Moreover, the noise is not eliminated in IMF5, which leads to a large error in EEMD. The RMSE of the EEMD is 1.919×10^{-32}.

Table 2
RUL prediction of lithium-ion batteries datasets.

Datasets	Method	Starting point	Real RUL	Predict RUL	AE	AER
	LSTM	90	34	50	16	0.4706
	EMD-LSTM	90	34	50	16	0.4706
B0005	CEEMDAN-LSTM	90	34	45	11	0.2444
	Our model	**90**	**34**	**33**	**1**	**0.0294**
	Our model(early)	**45**	**77**	**74**	**3**	**0.0389**
	LSTM	90	18	27	9	0.5000
	EMD-LSTM	90	18	24	6	0.3333
B0006	CEEMDAN-LSTM	90	18	23	5	0.2778
	Our model	**90**	**18**	**18**	**0**	**0**
	Our model (early)	**45**	**63**	**66**	**3**	**0.0476**
	LSTM	133	35	62	27	0.7714
	EMD-LSTM	133	35	67	32	0.9143
B0007	CEEMDAN-LSTM	133	35	51	16	0.4571
	Our model	**133**	**35**	**34**	**1**	**0.0285**
	Our model (early)	**45**	**123**	**123**	**0**	**0**
	LSTM	72	24	34	10	0.4167
	EMD-LSTM	72	24	30	6	0.2500
B0018	CEEMDAN-LSTM	72	24	22	2	0.0833
	Our model	**72**	**24**	**24**	**0**	**0**
	Our model (early)	**30**	**66**	**65**	**1**	**0.0151**
	LSTM	462	178	185	7	0.0393
	EMD-LSTM	462	178	180	2	0.0112
CS235	CEEMDAN-LSTM	462	178	184	6	0.0337
	Our model	**462**	**178**	**179**	**1**	**0.0056**
	Our model (early)	**300**	**340**	**339**	**1**	**0.0033**
	LSTM	599	46	52	6	0.1304
	EMD-LSTM	599	46	45	1	0.0217
CS236	CEEMDAN-LSTM	599	46	49	3	0.0655
	Our model	**599**	**46**	**45**	**1**	**0.0217**
	Our model (early)	**300**	**345**	**334**	**11**	**0.0290**
	LSTM	586	130	155	25	0.1923
	EMD-LSTM	586	130	144	14	0.1077
CS237	CEEMDAN-LSTM	586	130	148	18	0.1385
	Our model	**586**	**130**	**136**	**6**	**0.0461**
	Our model (early)	**300**	**415**	**393**	**22**	**0.0530**
	LSTM	581	175	182	7	0.0400
	EMD-LSTM	581	175	180	5	0.0286
CS238	CEEMDAN-LSTM	581	175	185	10	0.0571
	Our model	**581**	**175**	**177**	**2**	**0.0114**
	Our model (early)	**300**	**456**	**436**	**20**	**0.0438**

Fig. 7 (c) shows the decomposed results using the CEEMDAN, where the mode mixing phenomenon is effectively eliminated. At the same time, the problem of excessive noise in the EEMD has been solved. The RMSE of the CEEMDAN method is 1.535×10^{-32}, which is smaller than EEMD, thus resulting in an exact signal decomposition without mode mixing.

Table 2 lists the prediction results using different methods and their combinations. According to the AE metric in Table 2, prediction models combining mode decomposition (EMD, CEEMDAN) have smaller prediction errors compared to prediction models that do not use signal decomposition, confirming that the combined model is effective and feasible.

Moreover, the error using the CEEMDAN-LSTM model is much smaller than that using the EMD-LSTM model, which indicates that the decomposition part of the CEEMDAN method can achieve a more accurate prediction than the original EMD method.

4.2. RUL prediction results

In this research, several existing methods and their combinations in the field of battery RUL prediction, including LSTM, EMD-LSTM, and CEEMDAN-LSTM, are compared with the proposed CEEMDAN-Transformer-DNN hybrid model. The prediction model is validated by using battery data numbered 5, 6, 7, and 18 from the NASA dataset, as well as battery data CS235, CS236, CS237, and CS238 from the CALCE dataset.

Using the same starting point as suggested in the proposed hybrid model, Fig. 9 illustrates the prediction results for RUL. As can be seen, the proposed hybrid model is very accurate at predicting capacity regeneration phenomena, greatly improving RUL prediction accuracy. Table 2 shows the results using different models. Only using the LSTM model results in significant prediction errors, with the largest AER values not found in any other datasets. The EMD-LSTM combines the signal decomposition and acquires

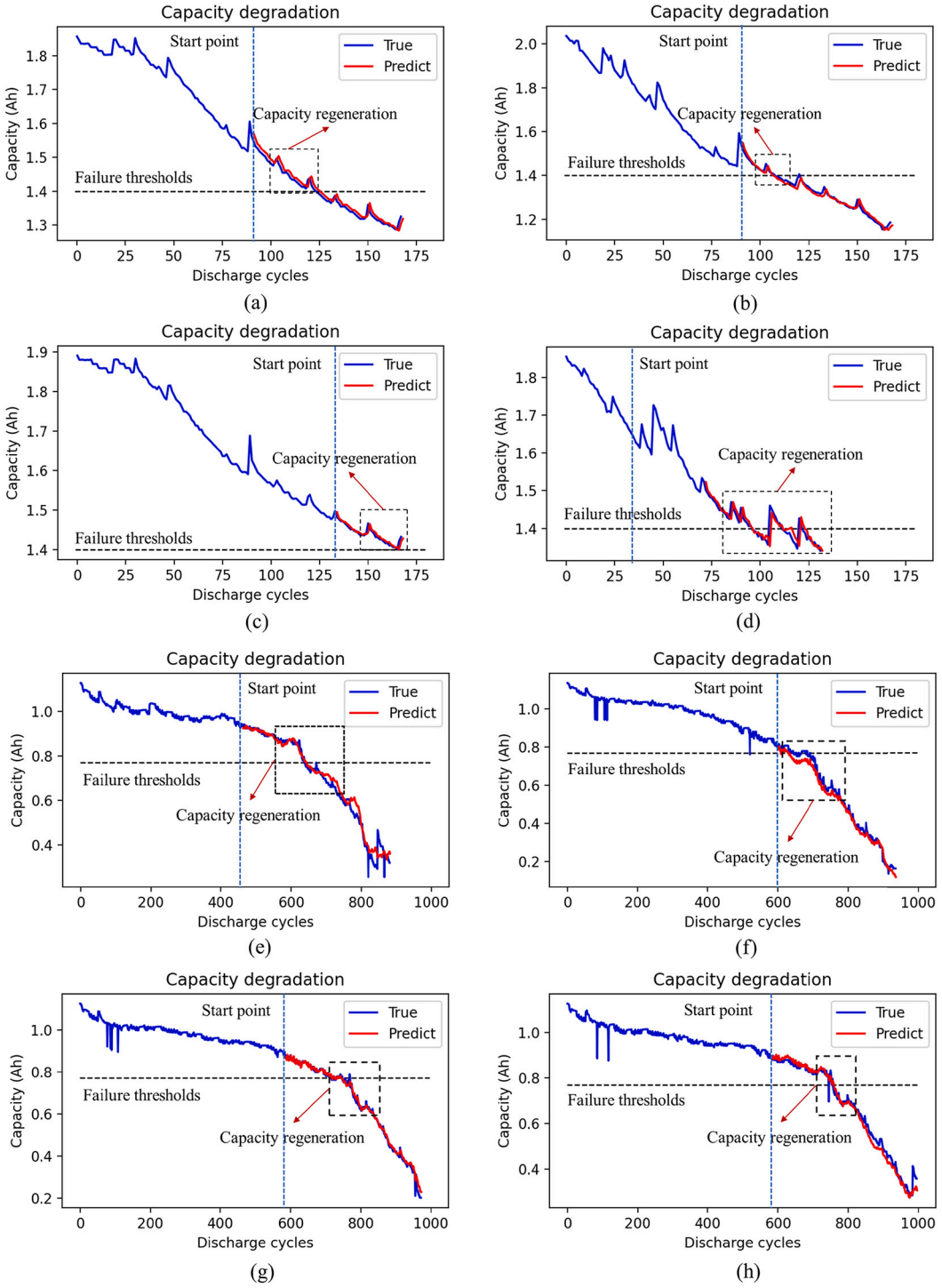

Fig. 9. RUL prediction results for the battery dataset used to validate the model. (a): B0005 (b): B0006 (c): B0007 (d): B0018 (e): CS235 (f): CS236 (g): CS237 (h): CS238.

decreased AERs. However, the improvement in accuracy is still limited, probably due to the extra errors caused by the mode mixing phenomenon introduced by the signal decomposition. The combination of CEEMDAN with LSTM substantially improves accuracy, proving the superiority of using CEEMDAN in the battery data.

For the NASA dataset, the RUL prediction for each battery is very accurate with our proposed method. The RMSEs are 0.0056, 0.0097, 0.018, and 0.007 for batteries 5, 6, 7, and 18. The proposed hybrid model shows better predictive accuracy than the combined

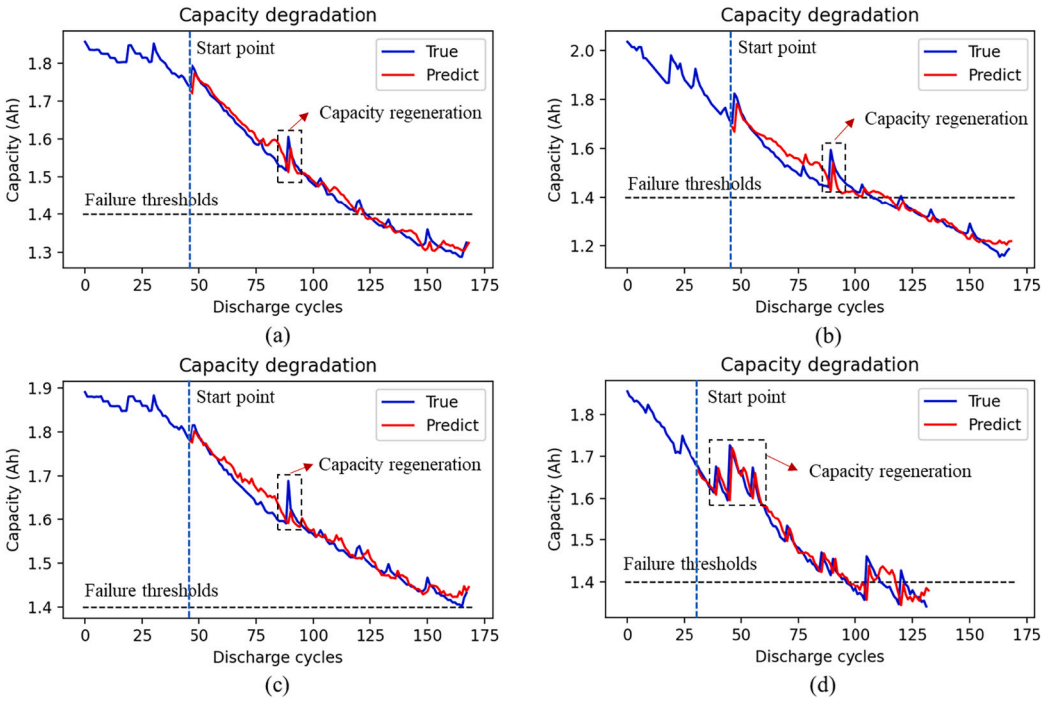

Fig. 10. Early prediction of RUL results for NASA battery dataset. (a): B0005 (b): B0006 (c): B0007 (d): B0018.

LSTM-based model for the AE and AER metrics. The average AER of CEEMDAN-Transformer-DNN is 0.0144, while the average AER of CEEMDAN-LSTM is 0.2656, indicating significant improvement. Accordingly, the proposed model has the smallest AE, reflecting its enhanced adaptability, robustness, and accuracy as compared to other models.

For the CALCE dataset, the average RMSE of CEEMDAN-Transformer-DNN is less than 3%, which is undoubtedly a very excellent performance. The average RMSE of the LSTM method is 0.0347, while the average RMSE of the CEEMDAN-Transformer-DNN method is 0.0244, with an average reduction of 29.6% for the RMSE metric. Meanwhile, the CEEMDAN-Transformer-DNN has the highest AE and AER metric accuracy. The average AER of CEEMDAN-Transformer-DNN is 0.0212, while the LSTM method is 0.0423, reducing about 49.8%. It demonstrates the effectiveness of replacing the LSTM with a Transformer-DNN model, achieving the best performance in prediction accuracy among the existing methods.

The above prediction results are based on the same starting point as the existing methods in the literature for comparison. To check the accuracy of the model's early prediction, we change the starting point to an earlier stage, as shown in Fig. 10 and Fig. 11. It is shown that the proposed method is able to accurately predict the capacity regeneration phenomenon using only 25%-30% of historical data, which is not available in other methods. The average RMSE of the proposed method using 25%-30% historical data is 0.0208 for the NASA dataset and 0.0337 for the CALCE dataset, which has a higher prediction accuracy than any other models that use even 50%-70% historical data. Hence, the proposed method has excellent early prediction precision.

The proposed method is found to achieve superior accuracy performances in several datasets when compared to other methods with CEEMDAN-Transf-ormer-DNN. Furthermore, even with a little capacity data, this model still offers better performance than other early prediction methods. As the transformer network can run in parallel, the calculation is faster than the mainstream recurrent neural networks and their variants, which can realize accurate prediction of the battery RUL in a shorter time.

5. Conclusion

Accurately predicting battery RUL is significant for monitoring the health state and enhancing operational safety through timely maintenance. Li-ion battery capacity regeneration problems during operation can seriously affect the accuracy of data-driven RUL prediction models. Additionally, using limited historical data, high-accurate early predictions of lithium-ion battery RUL are still challenging. In this research, we propose a CEEMDAN-Transformer-DNN hybrid model for RUL prediction that takes into account the capacity regeneration phenomenon on a global decreasing capacity. The hybrid model combines signal processing and deep learning to predict RUL.

According to test results using NASA and CALCE datasets, CEEMDAN can significantly address the mode mixing problem and improve the accuracy of decomposing battery capacity sequences. The proposed CEEMDAN-Transformer-DNN has better prediction accuracy and robustness. Furthermore, this method can predict RUL with high accuracy based only on 25%-30% of lifetime data.

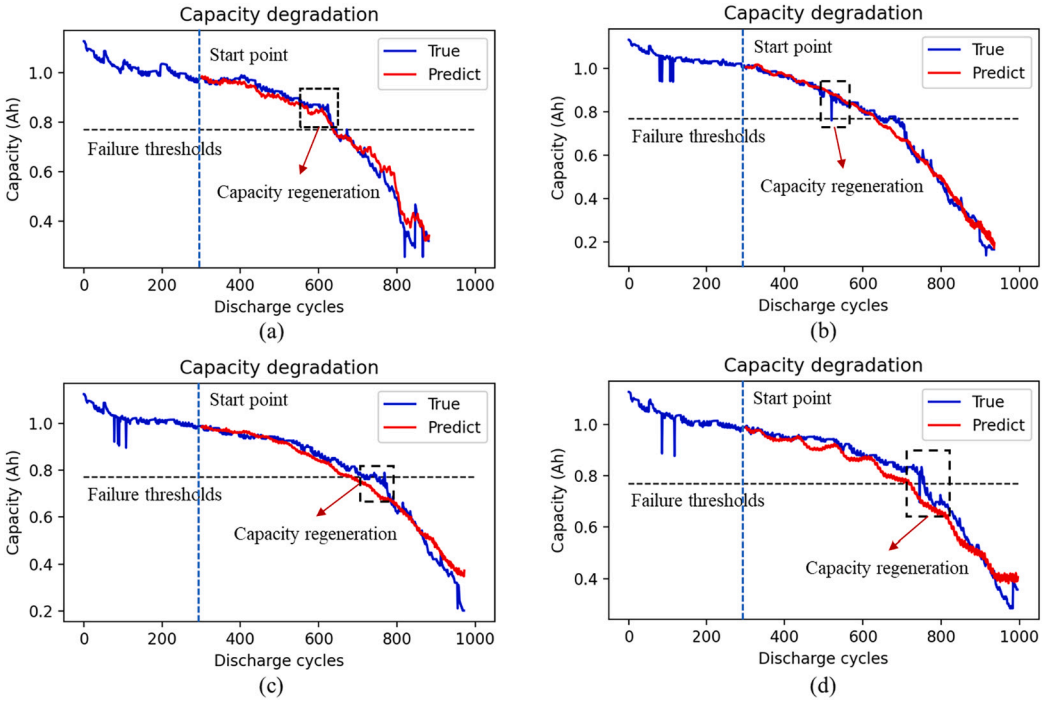

Fig. 11. Early prediction of RUL results for CALCE battery dataset. (a): CS235 (b): CS236 (c): CS237 (d): CS238.

Since capacity degradation behaviors are significantly affected by the actual operating conditions and working environment, multiple stress features like temperature, voltage, and current may be utilized as the input parameters to realize more advanced prediction models in future works.

CRediT authorship contribution statement

Yuxiang Cai: Conceived and designed the experiments; Performed the experiments; Contributed reagents, materials, analysis tools or data; Analyzed and interpreted the data; Wrote the paper. **Weimin Li; Qingguang Zhang:** Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper. **Taimoor Zahid; Chunhua Zheng:** Contributed reagents, materials, analysis tools or data; Wrote the paper. **Kun Xu:** Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationshipsthat could have appeared to influence the work reported in this paper.

Data availability statement

Data included in article/supp. material/referenced in article.

Acknowledgements

This work is partially supported by the National Natural Science Foundation of China (Grant No. 62073311), the Key Program of Natural Science Foundation of Shenzhen (Grant No. JCYJ20200109115403807, JCYJ20200109115414354) and the Sustainable Development Science and Technology Fund of Shenzhen (Grant No. KCXST20221021111210023).

References

[1] Zewang Chen, Na Shi, Yufan Ji, Mu Niu, Youren Wang, Lithium-ion batteries remaining useful life prediction based on bls-rvm, Energy 234 (2021) 121269.
[2] Qiuhui Ma, Ying Zheng, Weidong Yang, Yong Zhang, Hong Zhang, Remaining useful life prediction of lithium battery based on capacity regeneration point detection, Energy 234 (2021) 121233.
[3] Zan Li, Fengming Wang, Chengjie Wang, Qingpei Hu, Dan Yu, Reliability modeling and evaluation of lifetime delayed degradation process with nondestructive testing, Reliab. Eng. Syst. Saf. 208 (2021) 107358.

[4] Abdilbari Shifa Mussa, Liivat, Fast-charging effects on ageing for energy-optimized automotive lini1/3mn1/3co1/3o2 graphite prismatic lithium-ion cells, J. Power Sources 422 (2019) 175–184.
[5] Xiaosong Hu, Le Xu, Xianke Lin, Michael Pecht, Battery lifetime prognostics, Joule 4 (2) (2020) 310–346.
[6] Rui Xiong, Linlin Li, Jinpeng Tian, Towards a smarter battery management system: a critical review on battery state of health monitoring methods, J. Power Sources 405 (2018) 18–29.
[7] Jian Duan, Xuan Tang, Haifeng Dai, Ying Yang, Wangyan Wu, Xuezhe Wei, Yunhui Huang, Building safe lithium-ion batteries for electric vehicles: a review, Electrochem. Energy Rev. 3 (1) (2020) 1–42.
[8] Jin-zhen Kong, Fangfang Yang, Xi Zhang, Ershun Pan, Zhike Peng, Dong Wang, Voltage-temperature health feature extraction to improve prognostics and health management of lithium-ion batteries, Energy 223 (2021) 120114.
[9] Arijit Guha, Amit Patra, State of health estimation of lithium-ion batteries using capacity fade and internal resistance growth models, IEEE Trans. Transp. Electrif. 4 (1) (2017) 135–146.
[10] Jingwen Wei, Guangzhong Dong, Zonghai Chen, Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression, IEEE Trans. Ind. Electron. 65 (7) (2017) 5634–5643.
[11] Yuan Chen, Yigang He, Zhong Li, Liping Chen, Chaolong Zhang, Remaining useful life prediction and state of health diagnosis of lithium-ion battery based on second-order central difference particle filter, IEEE Access 8 (2020) 37305–37313.
[12] Xiaopeng Tang, Kailong Liu, Xin Wang, Furong Gao, James Macro, W. Dhammika Widanage, Model migration neural network for predicting battery aging trajectories, IEEE Trans. Transp. Electrif. 6 (2) (2020) 363–374.
[13] Jinsong Yang, Weiguang Fang, Jiayu Chen, Boqing Yao, A lithium-ion battery remaining useful life prediction method based on unscented particle filter and optimal combination strategy, J. Energy Storage 55 (2022) 105648.
[14] Ting Tang, Hui-Mei Yuan, Jun Zhu, Rul prediction of lithium batteries based on dlukf algorithm, in: 2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA), IEEE, 2020, pp. 1756–1761.
[15] Lijun Zhang, Zhongqiang Mu, Changyan Sun, Remaining useful life prediction for lithium-ion batteries based on exponential model and particle filter, IEEE Access 6 (2018) 17729–17740.
[16] Selina S.Y. Ng, Yinjiao Xing, Kwok L. Tsui, A naive Bayes model for robust remaining useful life prediction of lithium-ion battery, Appl. Energy 118 (2014) 114–123.
[17] Jianguo Wang, Shude Zhang, Chenyu Li, Lifeng Wu, Yingzhou Wang, A data-driven method with mode decomposition mechanism for remaining useful life prediction of lithium-ion batteries, IEEE Trans. Power Electron. 37 (11) (2022) 13684–13695.
[18] Lingling Li, Pengchong Wang, Kuei-Hsiang Chao, Yatong Zhou, Yang Xie, Remaining useful life prediction for lithium-ion batteries based on Gaussian processes mixture, PLoS ONE 11 (9) (2016) e0163004.
[19] Robert R. Richardson, Michael A. Osborne, David A. Howey, Gaussian process regression for forecasting battery state of health, J. Power Sources 357 (2017) 209–219.
[20] Jie Liu, Abhinav Saxena, Kai Goebel, Bhaskar Saha, Wilson Wang, An adaptive recurrent neural network for remaining useful life prediction of lithium-ion batteries, Technical report, National Aeronautics and Space Administration Moffett Field CA Ames Research . . . , 2010.
[21] Yongzhi Zhang, Rui Xiong, Hongwen He, Michael G. Pecht, Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries, IEEE Trans. Veh. Technol. 67 (7) (2018) 5695–5705.
[22] Zhanshe Yang, Yunhao Wang, Chenzai Kong, Remaining useful life prediction of lithium-ion batteries based on a mixture of ensemble empirical mode decomposition and gwo-svr model, IEEE Trans. Instrum. Meas. 70 (2021) 1–11.
[23] Zheming Tong, Jiazhi Miao, Shuiguang Tong, Yingying Lu, Early prediction of remaining useful life for lithium-ion batteries based on a hybrid machine learning method, J. Clean. Prod. 317 (2021) 128265.
[24] Guijun Ma, Zidong Wang, Weibo Liu, Jingzhong Fang, Yong Zhang, Han Ding, Ye Yuan, A two-stage integrated method for early prediction of remaining useful life of lithium-ion batteries, Knowl.-Based Syst. 259 (2023) 110012.
[25] Norden E. Huang, Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, Henry H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc. R. Soc. Lond., Ser. A, Math. Phys. Eng. Sci. 454 (1971) (1998) 903–995.
[26] Zhaohua Wu, Norden E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Adv. Adapt. Data Anal. 1 (01) (2009) 1–41.
[27] María E. Torres, Marcelo A. Colominas, Gaston Schlotthauer, Patrick Flandrin, A complete ensemble empirical mode decomposition with adaptive noise, in: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2011, pp. 4144–4147.
[28] Wojciech Zaremba, Ilya Sutskever, Oriol Vinyals, Recurrent neural network regularization, arXiv preprint, arXiv:1409.2329, 2014.
[29] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, Wang-chun Woo, Convolutional lstm network: a machine learning approach for precipitation nowcasting, Adv. Neural Inf. Process. Syst. 28 (2015).
[30] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, Illia Polosukhin, Attention is all you need, Adv. Neural Inf. Process. Syst. 30 (2017).
[31] Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio, Neural machine translation by jointly learning to align and translate, arXiv preprint, arXiv:1409.0473, 2014.
[32] Guorong Ding, Wenbo Wang, Ting Zhu, Remaining useful life prediction for lithium-ion batteries based on cs-vmd and gru, IEEE Access (2022).
[33] B. Saha, K. Goebel, Battery Data Set, NASA Ames Prognostics Data Repository, NASA Ames Research Center, Moffett Field, CA, 2007.
[34] M. Pecht, Battery data set. Center for advanced life cycle engineeringcalce, University of Maryland, 2013.