

Article

One-Time Prediction of Battery Capacity Fade Curve under Multiple Fast Charging Strategies

Xiaoming Han *, Zhentao Dai *, Mifeng Ren, Jing Cui and Yunfeng Shi

College of Electrical and Power Engineering, Taiyuan University of Technology, Taiyuan 030024, China; renmifeng@tyut.edu.cn (M.R.); cuijing0408@link.tyut.edu.cn (J.C.); shiyunfeng0473@link.tyut.edu.cn (Y.S.)
* Correspondence: hanxiaoming@tyut.edu.cn (X.H.); daizhentao0399@link.tyut.edu.cn (Z.D.)

Abstract: Using different fast charging strategies for lithium-ion batteries can affect the degradation rate of the batteries. In this case, predicting the capacity fade curve can facilitate the application of new batteries. Considering the impact of fast charging strategies on battery aging, a battery capacity degradation trajectory prediction method based on the TM-Seq2Seq (Trend Matching—Sequence-to-Sequence) model is proposed. This method uses data from the first 100 cycles to predict the future capacity fade curve and EOL (end of life) in one-time. First, features are extracted from the discharge voltage-capacity curve. Secondly, a sequence-to-sequence model based on CNN, SE-net, and GRU is designed. Finally, a trend matching loss function is designed based on the common characteristics of capacity fade curves to constrain the encoding features of the sequence-to-sequence model, facilitating the learning of the underlying relationship between inputs and outputs. TM-Seq2Seq model is verified on a public dataset with 132 battery cells and multiple fast charging strategies. The experimental results indicate that, compared to other popular models, the TM-Seq2Seq model has lower prediction errors.

Keywords: lithium-ion battery; multiple fast charging strategies; capacity fade curve; one-time prediction; trend matching



Citation: Han, X.; Dai, Z.; Ren, M.; Cui, J.; Shi, Y. One-Time Prediction of Battery Capacity Fade Curve under Multiple Fast Charging Strategies. *Batteries* **2024**, *10*, 74. <https://doi.org/10.3390/batteries10030074>

Academic Editor: Xin Sui

Received: 3 January 2024

Revised: 16 February 2024

Accepted: 20 February 2024

Published: 22 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the improvement of technology and the decrease of cost, lithium-ion batteries have gradually been promoted. However, their performance will accelerate to fade with the increase of usage times. This fade is the consequence of multiple coupled degradation mechanisms, including electrode–electrolyte interface side reactions, lithium deposition, electrode material structure damage and electrolyte decomposition [1]. Battery capacity is considered the leading indicator for evaluating battery performance. The battery capacity fade curve refers to the trend curve as the battery capacity gradually decreases with the increase of battery usage time. Generally, the battery capacity fade curve shows an exponential fade trend, that is, the longer the battery usage time, the faster the loss of battery capacity. Predicting the battery capacity fade curve can help people understand the battery life and performance, so that they can better manage, use, and maintain the battery. Therefore, it is necessary to develop models that can predict the trend of battery capacity fade using limited data.

Various models have been proposed by researchers to predict the fade trend of battery capacity. These models fall into two main categories: electrochemical models and data-driven models based on battery measurements such as capacity, voltage [2]. Electrochemical models [3] integrate pseudo-two-dimensional (P2D) models [4] or single-particle models [5] with battery aging mechanism models to predict battery capacity fade trends. This approach utilizes partial differential equations [6] to describe physical phenomena like mass transfer, charge transfer, thermodynamics, and chemical kinetics [7,8]. Although electrochemical models can provide highly accurate prediction results in some cases, there are many

uncertainties in the electrochemical process under different usage conditions, such as temperature, pressure, depth of charge–discharge, etc. These factors can affect the internal structure and chemical reactions of the battery that result in modeling failure.

Data-driven models have been widely applied. Some researchers used historical capacity data as input and fitted the capacity fade situation by using mathematical models such as exponential [9] and double exponential [10]. Other researchers consider more diverse inputs and employ deep learning models to capture capacity fade trends [11–14]. The aforementioned studies typically required sufficient charge–discharge data to accurately predict future capacities, which has limited the application of battery capacity prediction techniques in many scenarios.

Li et al. [15] implemented the earliest prediction of the entire capacity fade trend using a capacity-based method. This method relies solely on the capacity series data as input. They used data up to the first 100 cycles as input and used a sequence-to-sequence model based on a cyclic neural network. While this method predicts the future capacity fade curve in one-time, it fails to consider the impact of various charging conditions. Furthermore, certain high-performance batteries exhibit no capacity deterioration within the first 100 cycles [16]. Therefore, relying solely on capacity as a feature is insufficient for accurately estimating future capacity fade trajectories. It is necessary to extract more comprehensive aging features from data such as voltage, charge quantity, and other data.

Currently, considering the impact of fast charging strategies on batteries and the absence of capacity fade in the initial charge–discharge cycles, some methods have been successfully applied to predict end of life (EOL) [16–18], remaining useful life (RUL) [19], and rollover cycle (knee point) prediction [20]. These methods extract richer aging features from voltage, charge quantity, and other data, enabling the prediction of EOL, RUL and rollover cycle, and promoting the development of capacity degradation trajectory prediction methods. Building upon the various aging features validated by the aforementioned methods, researchers have demonstrated methods for predicting the capacity fade curve by using data from early charge–discharge cycles. Liu et al. [21] proposed an elastic net-based method to estimate parameters in an empirical power function that describes the capacity fade curve over more than 100 cycles. Similarly, Saurabh et al. [22] trained a 2D convolutional model to predict unknown parameters in the capacity fade curve parameter equation. Both of these methods can predict the capacity degradation trajectory. However, the parameter equation cannot fit all the original data without bias. This creates inherent errors in the constructed predicted labels, which can result in a low level of accuracy in the final outcomes.

Strange and dos Reis [23] used two different nine-layer 1D-CNN neural networks to predict the coordinates of multiple sampling points on the capacity degradation curve. Similarly, Herring et al. [24] used the first 100 cycles of data to build a multitask linear model that predicted the capacity fade curve. Training machine learning models using small amounts of battery data still presents a challenge, despite the absence of errors in label construction. Overfitting commonly plagues these approaches due to three primary reasons stemming from the nature of the data itself. Firstly, there are significant differences between samples, leading to complexities in modeling. Secondly, each sample exhibits high dimensions, with input features comprising multiple time series and the output label consisting of a single time series. Moreover, the resource-intensive nature of full-cycle experiments for batteries results in limited available training data. Consequently, fitting the training data necessitates larger models. However, without considering the distribution characteristics of the data, these models often lack adequate generalization ability. This predicament raises the unresolved issue of how to develop models with strong generalization capabilities in the face of minimal battery data.

In order to address the aforementioned issues, a battery capacity degradation trajectory prediction method based on the TM-Seq2Seq (Trend Matching—Sequence-to-Sequence) model is proposed. This method predicts the future capacity fade curve in one-time without iteration. Firstly, features are extracted during the first 100 charge–discharge

cycles. Secondly, the capacity fade curve is sampled at equal capacity intervals, and the sampled sequence is set as labels. Next, features are extracted using dilated convolution and SE-net, and the features are enhanced using a GRU decoder. Finally, a GRU decoder is used to predict the labels. Additionally, a trend feature matching loss function is designed based on the data characteristics of capacity fade curves. This loss function constrains the output from the intermediate layers of the model, ensuring that the intermediate features follow a certain pattern similar to the capacity fade curve. By controlling the learning space in this way, the model is able to enhance its robustness and improve its learning capabilities. Finally, a parameter equation for the capacity degradation curve is designed to fit the predicted sample points and obtain the complete capacity degradation curve. The proposed method is verified on a public dataset with 132 battery cells and multiple fast charging strategies.

2. The Battery Data for Multiple Fast Charging Strategies

Different fast charging strategies can impact the degradation rate of batteries. To reflect this scenario, this study utilizes the graphite/LiFePO₄ cylindrical batteries cycle dataset from Severson et al. [16] and Attia et al. [25]. The experimental dataset consists of 132 batteries of the same type with different fast charging strategies. Figure 1 shows six fast charging strategies' current-capacity curves and voltage-capacity curves. Firstly, these fast charging strategies involve constant current charging at different rates in different capacity ranges. Secondly, each battery follows the same fast charging strategy throughout different charge–discharge cycles. Lastly, all batteries have the same discharge conditions. Figure 2a displays the capacity fade curves of the 132 batteries. The horizontal axis represents the cycle number, the vertical axis represents the battery capacity, and each curve represents the capacity fade curve of a single battery. The color of the curve reflects the lifespan of the battery. The distribution of the total test cycles of 132 batteries is illustrated in Figure 2b. Under the influence of different charging strategies, the capacity degradation rate and trend of the batteries exhibit uncertainty, posing challenges to the prediction task of capacity fade curves.

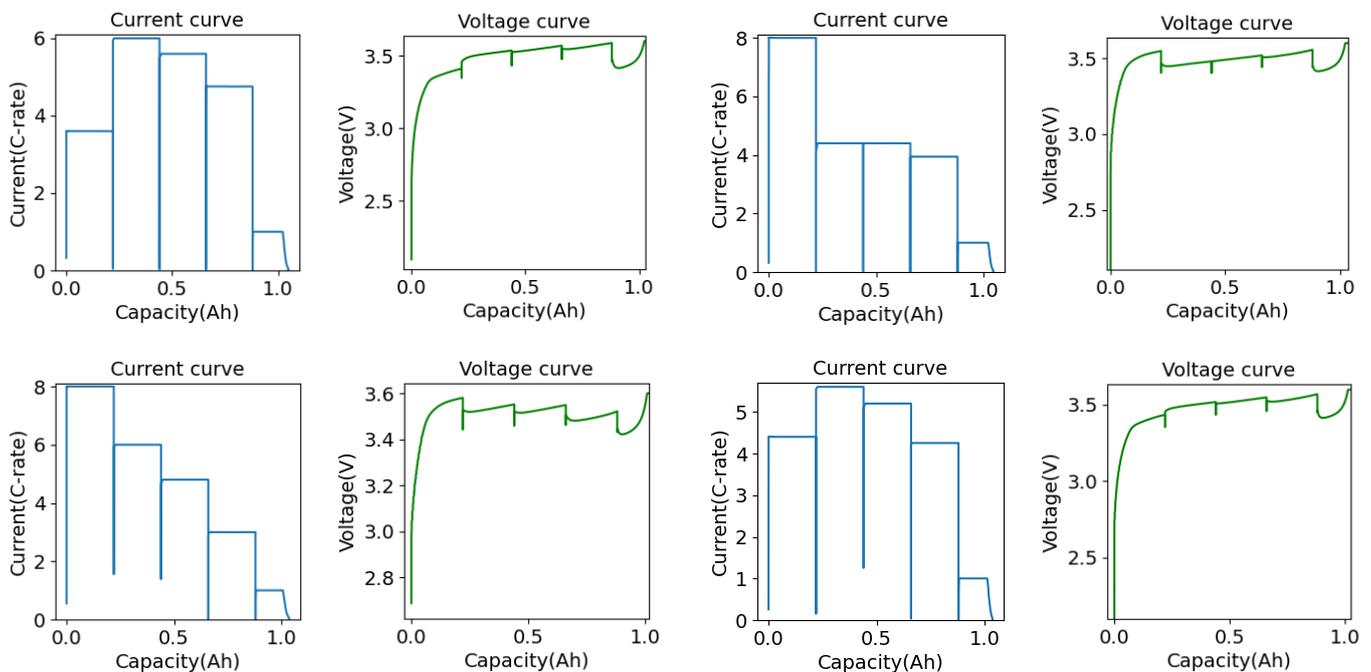


Figure 1. Cont.

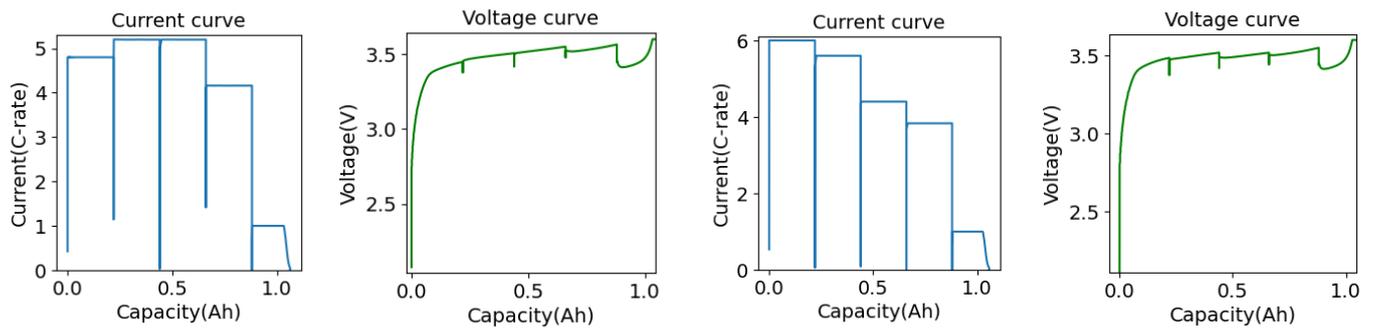


Figure 1. Current and voltage curves for six different fast charging strategies.

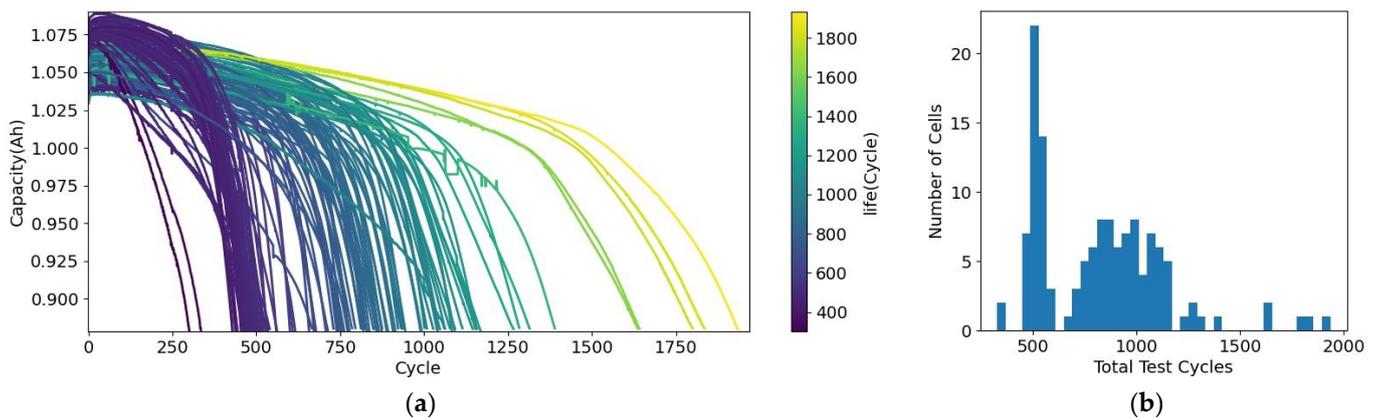


Figure 2. (a) Capacity fade curves of 132 experimental batteries. (b) The distribution of the total test cycles of 132 batteries.

Further information about the dataset can be found at the following web address:

1. <https://data.matr.io/1/projects/5c48dd2bc625d700019f3204>, accessed on 1 January 2024.
2. <https://data.matr.io/1/projects/5d80e633f405260001c0b60a>, accessed on 1 January 2024.

3. One-Time Prediction Method for Capacity Fade Curves

In this section, we explore a comprehensive approach to predict the capacity fade curve. This section comprises three subsections, each addressing a specific aspect of the methodology. Section 3.1 provides an overview of the feature extraction method as well as the method of quantifying the capacity fade curve into labels. Section 3.2 delves into the principles of the TM-Seq2Seq model, which serves as the core framework for our prediction methodology. Lastly, Section 3.3 presents the capacity fade curve parameter equations used to fit the labels into smooth and accurate representations of the actual fade curves.

3.1. Feature Structure and Label Structure

Although multiple fast charging strategies greatly affect the rate of battery capacity degradation, most batteries do not exhibit significant capacity fade within the first 100 charge–discharge cycles, as shown in Figure 3a. Relying solely on capacity data cannot reflect the battery’s aging rate. Therefore, it is necessary to identify features that reflect the battery’s aging rate and trend and map them onto future capacity fade curves. The voltage–discharge capacity curve is an important curve that reflects the discharge voltage strength of the battery at different states of charge. Many studies have extracted battery aging features through $Q(V)$ (discharge capacity–voltage) curves and conducted a series of prediction tasks. As the battery ages, the $Q(V)$ curve of the battery will undergo changes. Figure 3b illustrates the differences between the $Q(V)$ curves of the battery during the

first cycle and the 100th cycle. The horizontal axis of the graph represents the variance in capacity between the first cycle and the 100th cycle, while the vertical axis depicts the discharge voltage of the battery.

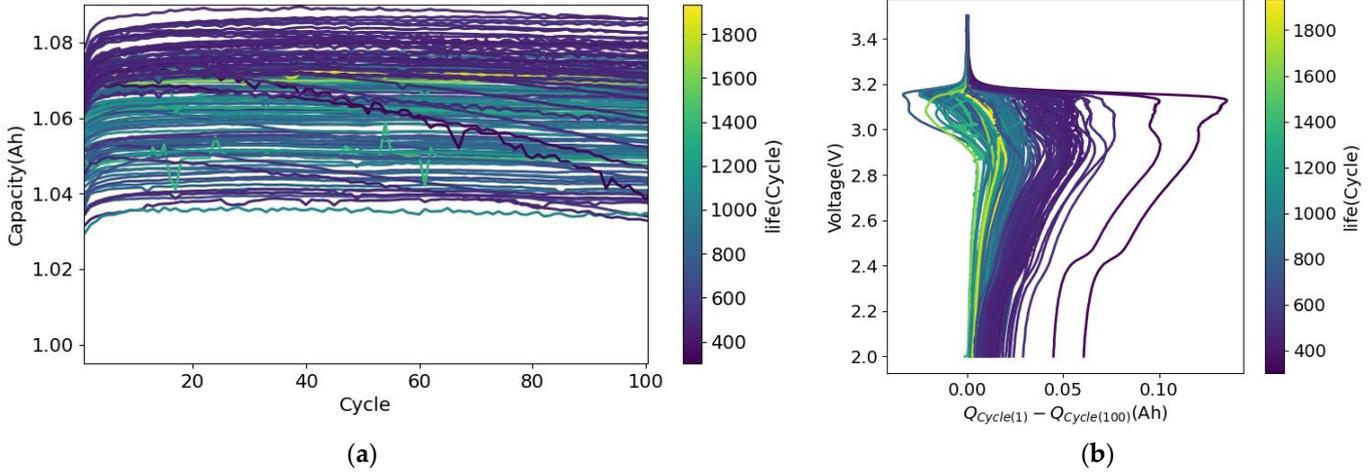


Figure 3. (a) Capacity curves within the first 100 charge–discharge cycles. (b) The subtraction of the Q(V) curve of the first cycle from the Q(V) curve of the 100th cycle of a battery.

The extracted features need to reflect the changes in the discharge Q(V) curve. Considering this aspect, two sets of temporal features are extracted. The specific feature extraction method is as follows.

Firstly, to ensure accurate feature extraction, it is important to establish a specific voltage range. The high-rate constant current discharge often leads to a reduction in the initial voltage, which can create difficulties in measuring voltages within the 3.5 V to 3.6 V range. Consequently, to address these challenges, we have opted to focus on extracting features within the voltage span of 2.0 V (discharge cut-off voltage) to 3.5 V. The discharge capacity values between 2.0 V and 3.5 V on the Q(V) curve are sampled at equal voltage intervals, generating a vector $\mathbf{Q}_i = \{Q_i(V_{\min}), \dots, Q_i(V_{\max})\}$ composed of the discharge capacity values. In this vector, the i is the cycle number and $Q_i(V_{\min})$ is the discharge capacity at a voltage of V_{\min} .

Next, the difference of vector \mathbf{Q}_i for each cycle relative to the 100th cycle is calculated using Equation (1). This difference serves to reflect the variations in the Q(V) curve.

$$\Delta \mathbf{Q}_{(100-i)} = \mathbf{Q}_{100} - \mathbf{Q}_i \quad (1)$$

Lastly, we use the mean and variance of the vector $\Delta \mathbf{Q}_{(100-i)}$ to construct two features. Specifically, the variance feature is calculated as shown in Equation (2), while the mean feature is computed following Equation (3).

$$\begin{cases} \mathbf{Var}_j = \{ \text{Var}(\Delta \mathbf{Q}_{(100-1)j}), \text{Var}(\Delta \mathbf{Q}_{(100-2)j}), \dots, \text{Var}(\Delta \mathbf{Q}_{(100-100)j}) \} \\ \text{Var}(\Delta \mathbf{Q}_{(100-i)}) = \frac{1}{n} \sum_{i=1}^n (\Delta \mathbf{Q}_{(100-i)}(V_i) - \overline{\Delta \mathbf{Q}_{(100-i)}})^2 \end{cases} \quad (2)$$

$$\mathbf{Mean}_j = \{ \overline{\Delta \mathbf{Q}_{(100-1)j}}, \overline{\Delta \mathbf{Q}_{(100-2)j}}, \dots, \overline{\Delta \mathbf{Q}_{(100-100)j}} \} \quad (3)$$

where \mathbf{Mean}_j and \mathbf{Var}_j is the features, j represents the j -th battery, which also corresponds to the j -th sample. These features reflect the speed and trend of battery aging within the first 100 cycles.

When the actual capacity of the battery reaches 80% of its rated capacity, the battery life ends. The end-of-life capacity of the battery can be calculated based on the rated capacity of the battery, as shown in Equation (4).

$$C_{EOL} = C_{Nominal} \times 80\% \quad (4)$$

where $C_{Nominal}$ is the nominal capacity of the battery, which is 1.1 Ah, and C_{EOL} is the end-of-life (EOL) capacity of the battery.

To achieve a reliable prediction of the capacity decay trajectory from the 100th cycle to end of life, we establish the target label that accurately reflects this progression. In order to maintain uniform label length across samples, we choose to evenly select 10 points along the capacity decay curve. This decision is driven by the intention to mitigate the risk of overfitting the model due to high label dimensions, particularly when dealing with limited sample data. Therefore, 10 points are uniformly sampled between C_{100} and C_{EOL} . These points can be obtained using Equation (5).

Next, to determine the coordinates of these 10 capacity values, we simply need to acquire their respective cycle numbers. The cycle numbers corresponding to these 10 capacity values are set as the labels for prediction, as shown in Equation (6).

$$\begin{cases} C_{ij} = C_{100j} - (C_{100j} - C_{EOL}) \times i \times 10\% \\ C_j = \{C_{1j}, C_{2j}, \dots, C_{10j}\} \end{cases} \quad (5)$$

$$y_j = \{Cycle_{1j}, Cycle_{2j}, \dots, Cycle_{10j}\} \quad (6)$$

where $Cycle_{ij}$ is the number of cycles to reach the capacity value of C_{ij} (specifically, $C_{10j} = C_{EOL}$, and $Cycle_{10j}$ is the EOL). Vector y_j is the label for prediction. J represents the j -th battery, which also corresponds to the j -th sample.

The initial measured capacity of different batteries deviates from the nominal capacity, resulting in variations in the numerical range of Vector C_j for different batteries. Additionally, the elements in Vector C_j and Vector y_j affect each other mutually. Given the above reasons, vector C_j is used to predict the label y_j . In summary, the features to be inputted into the model and the label to be predicted are shown in Equation (7).

$$D = [input|target] = \left[\begin{array}{ccc|c} Var_1, & C_1, & Mean_1 & y_1 \\ Var_2, & C_2, & Mean_2 & y_2 \\ & \dots & & \dots \\ Var_n, & C_n, & Mean_n & y_n \end{array} \right] \quad (7)$$

3.2. The Structure of TM-Seq2Seq Model

The original feature sequences used to predict the capacity fade curve represent a high-dimensional input for the sequence-to-sequence regression task. However, directly feeding these sequences into a Recurrent Neural Network (RNN) could hinder the GRU's ability to retain information over extended periods. Furthermore, the time series features being inputted into the model hold diverse physical interpretations, necessitating a careful evaluation of their relative importance for improving prediction accuracy. Therefore, to address these challenges, the TM-Seq2Seq model structure was specifically crafted with a focus on the sequence-to-sequence architecture [26] as shown in Figure 4. Firstly, the original features are learned and dimensionality-reduced through strided convolutions. Following this, an SE-net is added after the convolutional neural network to assign weights to features derived from various convolutional channels. Then, an encoder composed of GRU is used to encode the input sequence. Finally, a decoder composed of GRU is used to predict the labels y_j . In addition, the elements in Vector C_j and Vector y_j affect each other mutually. Given the above reasons, vector C_j is used as input for the decoder. Furthermore, considering that the capacity fade curve has characteristics approximating exponential

fade, the trend features of different sample encoding vectors are constrained. The specific roles and principles of SE-net, GRU, and trend matching are introduced below.

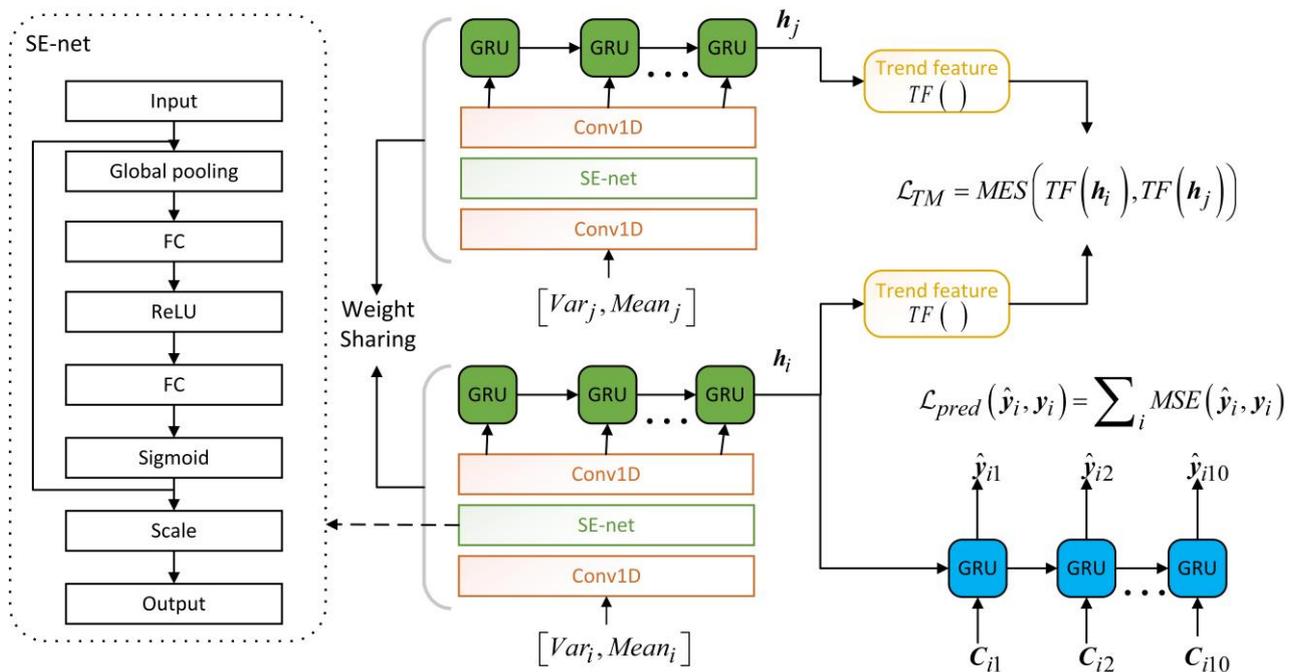


Figure 4. The structure of TM-Seq2Seq model.

3.2.1. SE-net

SE-Net (Squeeze-and-Excitation Network) [27] is an attention mechanism used to enhance the performance of convolutional neural networks (CNN). By introducing channel attention mechanism, SE-Net enables the network to dynamically learn the importance of each channel. First, for the input features, SE-Net performs global average pooling operation to transform the features of each channel into a single value. This value represents the global statistical information of that channel’s feature. Second, SE-Net introduces a fully connected two-layer network to learn the weight of each channel. Finally, by multiplying the learned channel attention weights with the original features, the effect of strengthening Important features and weakening useless features is achieved. The formula for the “squeeze” operation is shown in Equation (8).

$$\begin{cases} z_c = F_{sq}(u_c) = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H u_c(i, j) \\ \mathbf{U} = (u_1, u_2, \dots, u_c) \end{cases} \quad (8)$$

where \mathbf{U} is a set of input features, u_c is an input sample, and z_c is global statistical information.

The “excitation” operation is shown in Equation (9). The channel dependency is obtained through a simple gating mechanism.

$$s_c = F_{ex}(s_c, \mathbf{W}) = \sigma(g(z_c, \mathbf{W})) = \sigma(\mathbf{W}_2 \delta(z_c \mathbf{W}_1)) \quad (9)$$

where \mathbf{W}_1 and \mathbf{W}_2 are the weight matrix to be optimized. s_c represents the channel weight vector.

The output of the SE-net is shown in Equation (10).

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c \cdot u_c \quad (10)$$

where F_{scale} represents the multiplication operation channel-wise.

3.2.2. GRU

GRU [28] (Gated Recurrent Unit) comprises two gates: the update gate and the reset gate. The reset gate determines the influence of the old hidden state on the current state, while the update gate controls the influence of the new input on the current state. The updated equation is as follows:

$$z_t = \sigma(\mathbf{W}^z \cdot \mathbf{x}_t + \mathbf{U}^z \cdot \mathbf{h}_{t-1} + \mathbf{b}^z) \quad (11)$$

$$r_t = \sigma(\mathbf{W}^r \cdot \mathbf{x}_t + \mathbf{U}^r \cdot \mathbf{h}_{t-1} + \mathbf{b}^r) \quad (12)$$

$$\mathbf{h}'_t = \tanh(\mathbf{W}^{h'} \cdot \mathbf{x}_t + \mathbf{U}^{h'} \cdot (r_t \odot \mathbf{h}_{t-1}) + \mathbf{b}^{h'}) \quad (13)$$

$$\mathbf{h}_t = (1 - z_t) \odot \mathbf{h}_{t-1} + z_t \odot \mathbf{h}'_t \quad (14)$$

where x_t , h_t , and h'_t are input data, output hidden states, and candidate hidden states at time t . The parameters in the weight coefficient matrices, including \mathbf{W}^z , \mathbf{W}^r , and $\mathbf{W}^{h'}$ for the input data and \mathbf{U}^z , \mathbf{U}^r , and \mathbf{U}^h for the hidden states, are essential for the training of the network. The biases \mathbf{b}^z , \mathbf{b}^r , and \mathbf{b}^h are included for each corresponding component. Compared to LSTM, GRU utilizes fewer parameters, which results in faster training and testing speeds. Additionally, the simpler structure helps alleviate the problem of vanishing gradients, making GRU potentially perform better than LSTM when dealing with long-sequence data.

3.2.3. Trend Matching

To effectively fit complex and diverse input–output data, it is crucial to impose appropriate constraints on the model design, given the limited amount of training data available and the significant differences between samples. In designing a method to constrain the model, it is crucial to consider the factors that can impact its learning space. The constraints implemented should not alter the model's structure but rather enhance its generalization ability by limiting the range of parameters it can learn from the data. In many domains, special loss functions are used in generalization techniques to limit the output data from the intermediate layers of the network, thus restricting the learning space of the model [29,30]. The loss function plays a pivotal role in this process as it must be carefully designed to correspond to the data conditions and expected objectives. Given these factors, the method to constrain the model is designed as follows:

Firstly, based on the commonality that the capacity degradation curve roughly follows an exponential fade trend, Equation (15) is designed to quantify the trend feature of a vector.

$$\left\{ \begin{array}{l} \mathbf{y} = (\text{Cycle}_1, \text{Cycle}_2, \dots, \text{Cycle}_{10}) \\ TF(\mathbf{y}) = \left(\frac{|\text{Cycle}_1 - \bar{y}|}{\frac{1}{n} \sum_{i=1}^n |\text{Cycle}_i - \bar{y}|}, \frac{|\text{Cycle}_2 - \bar{y}|}{\frac{1}{n} \sum_{i=1}^n |\text{Cycle}_i - \bar{y}|}, \dots, \frac{|\text{Cycle}_n - \bar{y}|}{\frac{1}{n} \sum_{i=1}^n |\text{Cycle}_i - \bar{y}|} \right) \end{array} \right. \quad (15)$$

where \mathbf{y} represents the label and $TF(\mathbf{y})$ represents the trend feature of the label.

Secondly, the labels of different batteries were compared and the trend features of these labels were analyzed, as shown in Figure 5. Despite significant variations in the labels of different batteries, they exhibit similar trends. By imposing constraints on different feature encodings, we are able to follow certain patterns with relatively low levels of disorder in the different labels. This approach can help constrain the confusion level associated with various feature encodings, thus making the relationship between feature encodings and labels clearer and more effective. As a result, by enhancing the clarity and effectiveness of this relationship, we can significantly improve the performance and learning capability of the model.

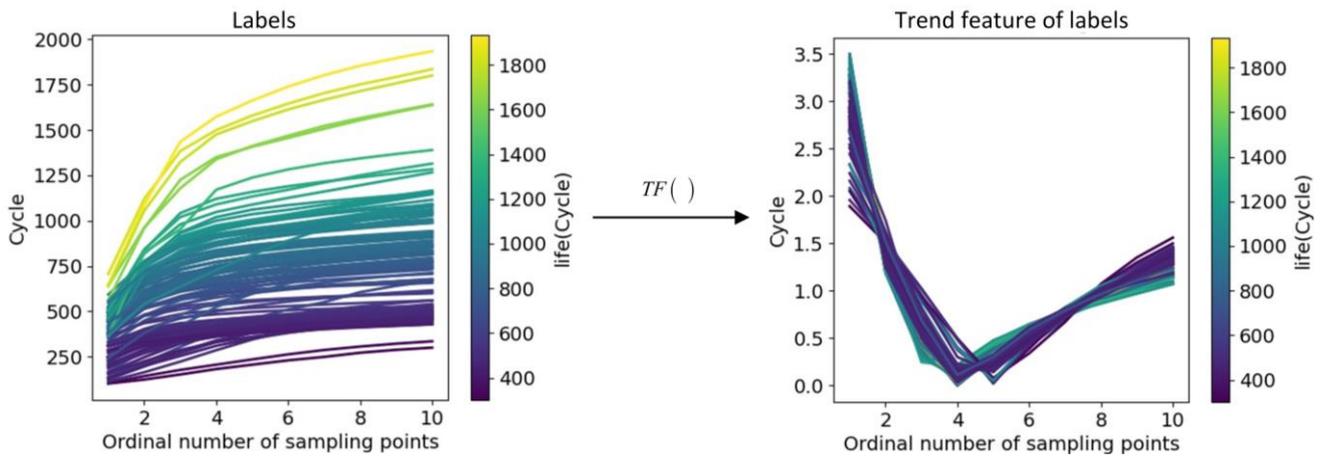


Figure 5. The distribution of the labels and the trend features of these labels.

Finally, Equation (16) is used as the loss function during model training. The purpose of this loss function is to minimize the differences in trend features of feature encodings across different samples. In addition, Equation (19) is also used as a loss function to fit the labels.

$$\mathcal{L}_{TM}(\mathbf{h}_i, \mathbf{h}_j) = \sum_{i,j} mse(TF(\mathbf{h}_i), TF(\mathbf{h}_j)) \tag{16}$$

$$\mathbf{h}_i = F_{encoder}(\mathbf{Mean}_i, \mathbf{Var}_i) \tag{17}$$

$$\hat{\mathbf{y}}_i = F_{decoder}(\mathbf{h}_i, \mathbf{C}_i) \tag{18}$$

$$\mathcal{L}_{pred}(\hat{\mathbf{y}}_i, \mathbf{y}_i) = \sum_i mse(\hat{\mathbf{y}}_i, \mathbf{y}_i) \tag{19}$$

where $F_{encoder}$ is the feature extractor consisting of CNN, SE-net, and GRU, $F_{decoder}$ is the GRU decoder, \mathbf{h}_i is the encoding vector of the i -th battery.

3.3. Parameter Equation of Capacity Fade Curve

When parameterizing the capacity fade curve, a reasonable parameter equation should be selected according to the actual situation to ensure accuracy and reliability. The exponential model is a common mathematical model that is used by many studies to fit battery capacity curve. In the dataset of this paper, some batteries experience capacity fade in the first 100 cycles, and full utilization of the known information from these cycles is necessary, such as the initial capacity fade rate. Therefore, we use the sum of the initial capacity fade rate m_0 and the exponential model to represent the change in battery capacity fade rate, where m_0 is the capacity fade rate of the battery in the 100th cycle, and the exponential model can describe the change in battery capacity fade rate. Specifically, the change in capacity fade rate is shown in Equation (20).

$$\frac{dC}{dn} = -m_0 - k \times e^{(n-N_k) \times \delta} \tag{20}$$

where n represents the number of cycles that are experienced by the battery after 100 cycles, m_0 is the capacity fade rate of the battery in the 100th cycle, dC/dn represents the capacity fade rate of the battery in the $100 + n$ th cycle, and k , N_k , and δ are three parameters to be optimized. Integrating Equation (20) yields the relationship between C and n as shown in Equation (21).

$$C = -m_0 \times n - k \times \frac{e^{(n-N_k) \times \delta}}{\delta} + C_1 \tag{21}$$

When $n = 0$, $C = C_{100}$ (C_{100} is the capacity in the 100th cycle). Substituting into Equation (21) can determine the constant C_1 :

$$C_{100} = -m_0 \times 0 - k \times \frac{e^{(0-N_k) \times \delta}}{\delta} + C_1 \tag{22}$$

$$C_1 = k \times \frac{e^{-N_k \times \delta}}{\delta} + C_{100} \tag{23}$$

$$C = -m_0 \times n - k \times \frac{e^{(n-N_k) \times \delta} - e^{-N_k \times \delta}}{\delta} + C_{100} \tag{24}$$

The capacity of the first 100 cycles is used to determine the initial capacity fade rate. Afterwards, the unknown parameters in Equation (24) are optimized using the nonlinear least squares method, so that Equation (24) can better fit the predicted 10 future capacity fade curve sampling points from the TM-Seq2Seq model.

4. Results and Discussion

In this section, the predicting results of our model and the fitting results of the capacity fade curve are presented. The prediction errors of different models are then compared.

To comprehensively evaluate the prediction performance of capacity fade curves, root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) are utilized as error quantification metrics to assess the accuracy of the estimation results. Additionally, the predicted label $y_i = (y_{i1}, y_{i2}, \dots, y_{i10})$ reflects the rate and trend of capacity fade, and y_{i10} represents the battery's end of life (EOL). The prediction result for EOL is crucial for formulating maintenance and recycling plans for the battery. Therefore, the prediction errors of the sampling sequence and EOL are separately calculated as shown below:

$$RMSE = \sqrt{\frac{1}{10 \times n} \sum_{i=1}^n \sum_{j=1}^{10} (\hat{y}_{ij} - y_{ij})^2} \tag{25}$$

$$RMSE_{EOL} = \sqrt{\frac{1}{n} \sum_{j=1}^n (\hat{y}_{i10} - y_{i10})^2} \tag{26}$$

$$MAE = \frac{1}{10 \times n} \sum_{i=1}^n \sum_{j=1}^{10} |\hat{y}_{ij} - y_{ij}| \tag{27}$$

$$MAE_{EOL} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_{i10} - y_{i10}| \tag{28}$$

$$MAPE = \frac{100\%}{10 \times n} \sum_{i=1}^n \sum_{j=1}^{10} \left| \frac{\hat{y}_{ij} - y_{ij}}{y_{ij}} \right| \tag{29}$$

$$MAPE_{EOL} = \frac{100\%}{n} \sum_{j=1}^{10} \left| \frac{\hat{y}_{i10} - y_{i10}}{y_{i10}} \right| \tag{30}$$

where n represents the total number of test samples. RMSE considers the square of estimation errors, thus penalizing larger errors more heavily and better reflecting the differences between model estimations and actual observations. MAE calculates the absolute value of estimation errors, directly measuring the average difference between estimated values and observed values. MAPE converts estimation errors into percentage form, intuitively reflecting the magnitude of estimation errors relative to observed values, facilitating the evaluation of model performance on different datasets.

To demonstrate the effectiveness of the model, comparative experiments were conducted against other popular machine learning architectures. We used the programming software 'Python 3.9' to build, train, and test all models. The participating models for comparison are as follows:

1. Seq2Seq: The Seq2Seq model consists of an encoder and a decoder. The encoder transforms the input sequence into a fixed-length vector, which is then used by the decoder to generate the output sequence. This model has demonstrated excellent performance in tasks such as predicting capacity fade curves [15].
2. CNN: The CNN model captures local features through convolution operations and reduces the number of parameters and computational complexity through pooling operations. It is also proficient in learning time series data. This model has been widely applied to tasks such as battery EOL prediction [18] and capacity fade curve prediction [22,23].
3. CNN-BI-LSTM: BI-LSTM is a type of recurrent neural network model used for natural language processing tasks such as named entity recognition and sentiment analysis. The model comprises both forward and backward LSTM layers, allowing it to utilize contextual information at each time step and better capture long-term dependencies. This model has been applied to tasks of battery SOH prediction [31].
4. SE-CNN-LSTM: SE-CNN-LSTM is a model that combines Squeeze-and-Excitation (SE) module, Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM). It leverages the SE module to enhance the feature representation of CNN, which allows the model to capture more informative features.

Each model is trained and tested under similar conditions to ensure a fair comparison. We conducted 10 experiments. In each of the 10 experiments, we randomly divided 132 sets of batteries into three groups: a training set (~60%), a validation set (~20%), and a test set (~20%). The model parameters were trained using the training set, and the parameters with the highest accuracy on the validation set were saved to ensure the model's performance on unseen data (generalization). Following this, the test data were input into the saved model to obtain prediction results. Finally, we calculated the average error from the 10 experiments for model comparison. The comparative results are presented in Table 1.

Table 1. Prediction errors of different models on the testing set.

Model	MAE	MAPE	RMSE
Seq2Seq	83.326	11.768%	109.921
CNN	99.730	15.619%	126.192
CNN-BI-LSTM	85.099	13.749%	105.192
SE-CNN-LSTM	90.418	13.898%	109.803
TM-Seq2Seq	77.043	12.246%	101.336
	MAE (EOL)	MAPE (EOL)	RMSE (EOL)
Seq2Seq	99.988	11.511%	127.745
CNN	113.772	14.853%	137.954
CNN-BI-LSTM	103.319	12.642%	131.097
SE-CNN-LSTM	103.478	12.207%	125.899
TM-Seq2Seq	87.838	10.409%	114.501

To justify the rationale behind incorporating the trend matching method and SE-Net, ablation experiments were performed on SE-Net and the trend matching method. The experimental results are shown in Table 2.

In addition, to visually represent the error distribution of different batteries, a distribution chart of the MAPE of all batteries is plotted, as shown in Figure 6. The width of each bar represents the corresponding error interval, and the height represents the number of batteries in that error interval. Figure 7 shows the predicted results of the EOL.

Table 2. Results comparison for ablation experiment.

SE-net	x	✓	x	✓
Trend Matching	x	x	✓	✓
MAE	88.502	81.440	85.089	77.043
MAPE	13.171%	13.088%	13.330%	12.246%
RMSE	114.178	108.281	122.588	101.336
MAE (EOL)	107.513	91.506	97.728	87.838
MAPE (EOL)	12.757%	11.384%	12.084%	10.409%
RMSE (EOL)	135.346	112.921	137.406	114.501

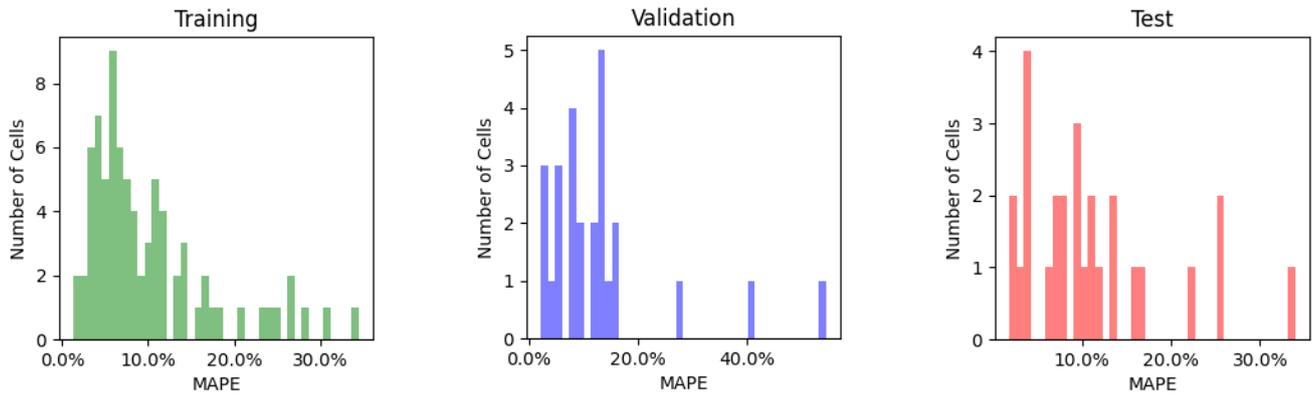


Figure 6. Histogram of error distribution for all batteries.

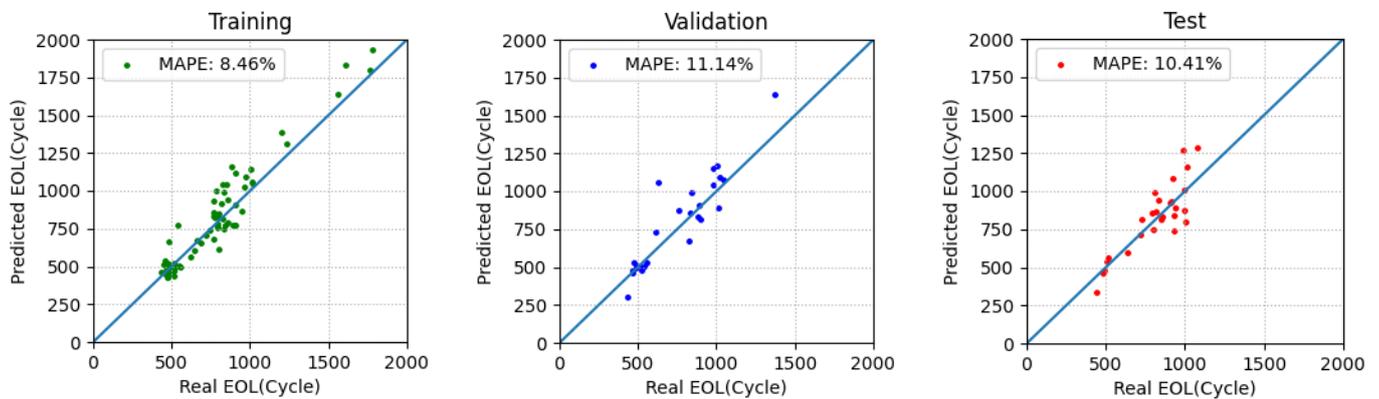


Figure 7. The predicted results of the EOL.

The researchers ranked the 27 batteries from the test set, with the first-ranked prediction being the most accurate and the predictions with lower rankings having larger errors. To showcase the fitting results of the capacity fade curve, nine batteries were selected for demonstration purposes, which are the 1th, 4th, 7th, 11th, 14th, 17th, 20th, 24th, and 27th batteries, as shown in Figure 8.

The TM-Seq2Seq model achieves the best overall performance in predicting both the capacity degradation trajectory and EOL. The ultimately predicted capacity degradation curve accurately reflects the true rate and trend of capacity fade. Additionally, incorporating both SE-net and trend matching methods improves the overall performance of the model. However, adding SE-net alone increases prediction errors. SE-net introduces channel importance weights, which enhance the model’s learning capability but also increase the risk of overfitting.

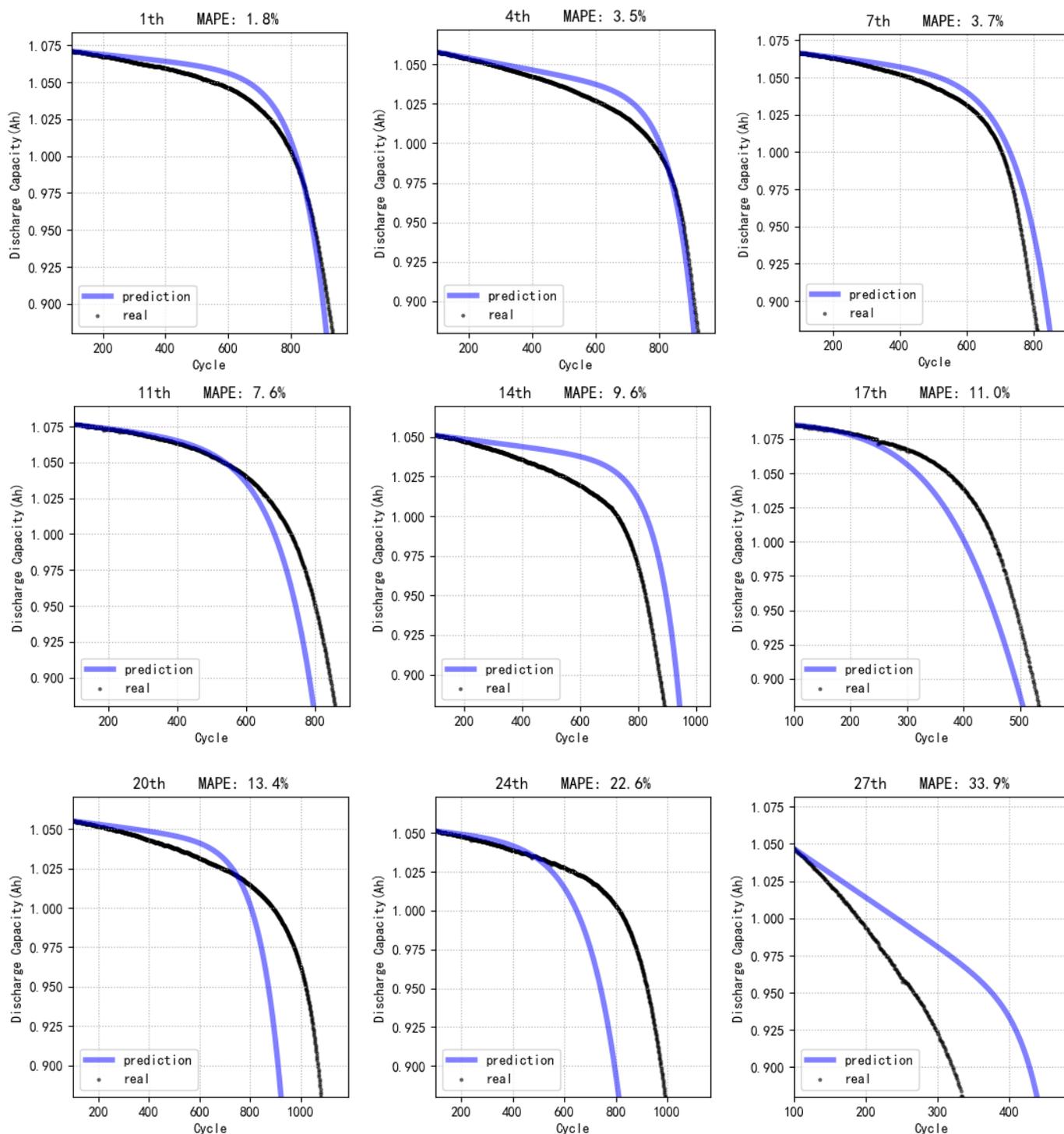


Figure 8. TM-Seq2Seq model predicted capacity fade curves for test data.

5. Conclusions

In conclusion, our study focuses on the prediction of capacity fade curves in lithium-ion batteries under different fast charging strategies. We proposed a battery capacity degradation trajectory prediction method. This method uses data from the first 100 cycles to predict the future capacity fade curve and EOL (end of life) in one-time. First, features are extracted from the discharge voltage–capacity curve. Secondly, the TM-Seq2Seq model based on CNN, SE-net, and GRU is designed. Finally, a trend matching loss function is designed based on the common characteristics of capacity fade curves. Through experi-

mentation, the following conclusions were discovered: The SE-net enhances the model's fitting capability, allowing it to capture more intricate patterns in the data. Furthermore, the trend matching is suitable for the data conditions in this study, and it can enhance the model's performance on unknown test data. Ultimately, the TM-Seq2Seq model improves the accuracy of one-time predictions for battery capacity fade curves under various fast charging strategies.

In the future, we will further optimize this method to improve prediction performance and generalization ability. We will also extend the idea of trend feature distribution matching to other machine learning and deep learning methods to further improve the accuracy and stability of capacity fade curve prediction. The method proposed in this paper is not only applicable to battery data processing but can also be applied to similar problems in other fields, with strong versatility and scalability. This method can improve the practicality and reliability of models and is of great significance for solving practical problems.

Author Contributions: Conceptualization, X.H. and Z.D.; methodology, X.H. and Z.D.; software, Z.D.; validation, Z.D., X.H. and J.C.; formal analysis, M.R.; investigation, Z.D.; resources, X.H.; data curation, Y.S.; writing—original draft preparation, Z.D.; writing—review and editing, X.H.; visualization, Z.D.; supervision, X.H.; project administration, X.H.; funding acquisition, X.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Research Project Supported by Shanxi Scholarship Council of China, grant number 2020-041.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: 1. <https://data.mtr.io/1/projects/5c48dd2bc625d700019f3204>, accessed on 1 January 2024. 2. <https://data.mtr.io/1/projects/5d80e633f405260001c0b60a>, accessed on 1 January 2024.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chen, M.; Ma, G.; Liu, W.; Zeng, N.; Luo, X. An overview of data-driven battery health estimation technology for battery management system. *Neurocomputing* **2023**, *532*, 152–169. [\[CrossRef\]](#)
2. Hu, X.; Xu, L.; Lin, X.; Pecht, M. Battery Lifetime Prognostics. *Joule* **2020**, *4*, 310–346. [\[CrossRef\]](#)
3. Fuller, T.F.; Doyle, M.; Newman, J. Simulation and optimization of the dual lithium ion insertion cell. *J. Electrochem. Soc.* **1994**, *141*, 1. [\[CrossRef\]](#)
4. Doyle, M.; Fuller, T.F.; Newman, J. Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell. *J. Electrochem. Soc.* **1993**, *140*, 1526–1533. [\[CrossRef\]](#)
5. Santhanagopalan, S.; Guo, Q.; Ramadass, P.; White, R.E. Review of models for predicting the cycling performance of lithium ion batteries. *J. Power Sources* **2006**, *156*, 620–628. [\[CrossRef\]](#)
6. Kindermann, F.M.; Keil, J.; Frank, A.; Jossen, A. A SEI modeling approach distinguishing between capacity and power fade. *J. Electrochem. Soc.* **2017**, *164*, E287–E294. [\[CrossRef\]](#)
7. Deshpande, R.; Verbrugge, M.; Cheng, Y.-T.; Wang, J.; Liu, P. Battery cycle life prediction with coupled chemical fade and fatigue mechanics. *J. Electrochem. Soc.* **2012**, *159*, A1730–A1738. [\[CrossRef\]](#)
8. Reniers, J.M.; Mulder, G.; Howey, D.A. Review and performance comparison of mechanical-chemical fade models for lithium-ion batteries. *J. Electrochem. Soc.* **2019**, *166*, A3189–A3200. [\[CrossRef\]](#)
9. Saha, B.; Goebel, K.; Poll, S.; Christophersen, J. Prognostics methods for battery health monitoring using a Bayesian framework. *IEEE Trans. Instrum. Meas.* **2009**, *58*, 291–296. [\[CrossRef\]](#)
10. He, W.; Williard, N.; Osterman, M.; Pecht, M. Prognostics of lithium-ion batteries based on Dempster-Shafer theory and the Bayesian Monte Carlo method. *J. Power Sources* **2011**, *196*, 10314–10321. [\[CrossRef\]](#)
11. Li, P.; Zhang, Z.; Xiong, Q.; Ding, B.; Hou, J.; Luo, D.; Rong, Y.; Li, S. State-of-health estimation and remaining useful life prediction for the lithium-ion battery based on a variant long short term memory neural network. *J. Power Sources* **2020**, *459*, 228069. [\[CrossRef\]](#)
12. 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. *J. Energy Storage* **2019**, *21*, 510–518. [\[CrossRef\]](#)
13. Wang, F.; Zhao, Z.; Ren, J.; Zhai, Z.; Wang, S.; Chen, X. A transferable lithium-ion battery remaining useful life prediction method from cycle-consistency of fade trend. *J. Power Sources* **2022**, *525*, 231027.
14. Zhang, Y.; Xiong, R.; He, H.; Pecht, M.G. Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries. *IEEE Trans. Veh. Technol.* **2018**, *67*, 5695–5705. [\[CrossRef\]](#)

15. Li, W.; Sengupta, N.; Dechent, P.; Howe, D.; Annaswamy, A.; Sauer, D.U. One-shot battery fade trajectory prediction with deep learning. *J. Power Sources* **2021**, *506*, 230024. [[CrossRef](#)]
16. 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 fade. *Nat. Energy* **2019**, *4*, 383–391. [[CrossRef](#)]
17. Ma, Y.; Wu, L.; Guan, Y.; Peng, Z. The capacity estimation and cycle life prediction of lithium-ion batteries using a new broad extreme learning machine approach. *J. Power Sources* **2020**, *476*, 228581. [[CrossRef](#)]
18. Shen, S.; Nemani, V.; Liu, J.; Hu, C.; Wang, Z. A hybrid machine learning model for battery cycle life prediction with early cycle data. In Proceedings of the 2020 IEEE Transportation Electrification Conference and Expo (ITEC), Chicago, IL, USA, 23–26 June 2020; pp. 181–184.
19. Hong, J.; Lee, D.; Jeong, E.R.; Yi, Y. Towards the swift prediction of the remaining useful life of lithium-ion batteries with end-to-end deep learning. *Appl. Energy* **2020**, *278*, 115646. [[CrossRef](#)]
20. Fermín-Cueto, P.; McTurk, E.; Allerhand, M.; Medina-Lopez, E.; Anjos, M.F.; Sylvester, J.; Reis, G.D. Identification and machine learning prediction of knee point and knee-onset in capacity fade curve of lithium-ion cells. *Energy AI* **2020**, *1*, 100006. [[CrossRef](#)]
21. Liu, J.; Thelen, A.; Hu, C.; Yang, X.G. An end-to-end learning framework for early prediction of battery capacity trajectory. In Proceedings of the Annual Conference of the PHM Society, Nashville, TN, USA, 29 November–20 December 2021; Volume 13.
22. Saxena, S.; Ward, L.; Kubal, J.; Lu, W.; Babinec, S.; Paulson, N. A convolutional neural network model for battery capacity fade curve prediction using early life data. *J. Power Sources* **2022**, *542*, 231736. [[CrossRef](#)]
23. Strange, C.; Reis, G.D. Prediction of future capacity and internal resistance of Li-ion cells from one cycle of input data. *Energy AI* **2021**, *5*, 100097. [[CrossRef](#)]
24. Herring, P.; Gopal, C.B.; Aykol, M.; Montoya, J.H.; Anapolsky, A.; Attia, P.M.; Gent, W.; Hummelshoj, J.S.; Hung, L.; Kwon, H.; et al. A Python library for battery evaluation and early prediction. *Software* **2020**, *11*, 100506. [[CrossRef](#)]
25. Attia, P.M.; Grover, A.; Jin, N.; Severson, K.A.; Markov, T.M.; Liao, Y.; 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*, 397–402. [[CrossRef](#)] [[PubMed](#)]
26. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to Sequence Learning with Neural Networks. In Proceedings of the 2014 Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS), Montreal, QC, Canada, 8–13 December 2014; pp. 3104–3112.
27. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In Proceedings of the 2018 IEEE CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018.
28. Song, Y.; Li, L.; Peng, Y.; Liu, D. Lithium-Ion Battery Remaining Useful Life Prediction Based on GRU-RNN. In Proceedings of the 2018 12th International Conference on Reliability, Maintainability, and Safety (ICRMS), Shanghai, China, 17–19 October 2018; pp. 317–322.
29. Zhou, K.; Liu, Z.; Qiao, Y.; Xiang, T.; Loy, C.C. Generalizing to Unseen Domains: A Survey on Domain Generalization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *35*, 4396–4415.
30. Du, Y.; Wang, J.; Feng, W.; Pan, S.; Qin, T.; Xu, R.; Wang, C. AdaRNN: Adaptive Learning and Forecasting for Time Series. In Proceedings of the 2021 Proceedings of the 30th ACM International Conference on Information & Knowledge Management (CIKM), Gold Coast, QD, Australia, 1–5 November 2021; pp. 402–414.
31. Yan, B.; Zheng, W.J.; Tang, D.Y.; Laili, Y.J.; Xing, Y. A knowledge-constrained CNN-BiLSTM model for lithium-ion batteries state-of-charge estimation. *Microelectron. Reliab.* **2023**, *150*, 115112. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.