

Article

Method for Evaluating Degradation of Battery Capacity Based on Partial Charging Segments for Multi-Type Batteries

Yujuan Sun ¹, Hao Tian ¹ , Fangfang Hu ² and Jiuyu Du ^{3,*}

¹ School of Mechanical and Energy Engineering, Beijing University of Technology, Beijing 100124, China; yjsun@bjut.edu.cn (Y.S.); haotian@emails.bjut.edu.cn (H.T.)

² Beijing Products Quality Supervision and Inspection Research Institute, National Automotive Quality Inspection and Testing Center, Beijing 101300, China; ff8345@126.com

³ School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China

* Correspondence: dujiuyu@tsinghua.edu.cn

Abstract: Accurately estimating the capacity degradation of lithium-ion batteries (LIBs) is crucial for evaluating the status of battery health. However, existing data-driven battery state estimation methods suffer from fixed input structures, high dependence on data quality, and limitations in scenarios where only early charge–discharge cycle data are available. To address these challenges, we propose a capacity degradation estimation method that utilizes shorter charging segments for multiple battery types. A learning-based model called GateCNN-BiLSTM is developed. To improve the accuracy of the basic model in small-sample scenarios, we integrate a single-source domain feature transfer learning framework based on maximum mean difference (MMD) and a multi-source domain framework using the meta-learning MAML algorithm. We validate the proposed algorithm using various LIB cell and battery pack datasets. Comparing the results with other models, we find that the GateCNN-BiLSTM algorithm achieves the lowest root mean square error (RMSE) and mean absolute error (MAE) for cell charging capacity estimation, and can accurately estimate battery capacity degradation based on actual charging data from electric vehicles. Moreover, the proposed method exhibits low dependence on the size of the dataset, improving the accuracy of capacity degradation estimation for multi-type batteries with limited data.

Keywords: lithium-ion battery; capacity degradation; convolutional neural network; data-driven; transfer learning; partial charging segments



Citation: Sun, Y.; Tian, H.; Hu, F.; Du, J. Method for Evaluating Degradation of Battery Capacity Based on Partial Charging Segments for Multi-Type Batteries. *Batteries* **2024**, *10*, 187.

<https://doi.org/10.3390/batteries10060187>

Received: 30 April 2024

Revised: 24 May 2024

Accepted: 27 May 2024

Published: 30 May 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

Lithium-ion batteries (LIBs) are widely used in electric vehicles and energy storage systems due to their high energy density, low self-discharge rate, long service life, high reliability, and high safety. LIBs are easily affected by factors such as the charging and discharging cycle, temperature changes, fluctuations, and stress in different operating conditions, resulting in continuous battery degradation [1]. Therefore, it is necessary to accurately and real-time estimate the capacity of lithium batteries to provide safer and more reliable energy for electric vehicles (EVs), smart grids, and other applications. This will enhance the stability and manageability of the energy system while maximizing the utilization of battery capacity. State of health (SOH) is a widely accepted indicator of battery life degradation, providing insight into the extent of degradation in a battery's current maximum discharging capacity compared to its factory-rated capacity [2]. SOH directly impacts battery performance and usable lifespan. Currently, three common methods are employed for SOH estimation: electrochemical mechanism-based estimation, equivalent circuit model (ECM)-based estimation, and data-driven methods. The electrochemical model (EM) is a mechanistic approach that simulates battery chemical reactions to characterize battery performance [3,4]. EM offers strong interpretability in describing external battery characteristics. However, EM relies on certain parameters that are challenging to

measure or identify, such as electrode materials and electrolyte chemistry. Consequently, the model may not accurately capture battery aging behavior [5]. In contrast to the EM, the ECM method offers an alternative approach that is not reliant on the internal chemistry of a battery. By using essential electrical components, such as resistance, capacitance, and inductance, the ECM method constructs a circuit model capable of describing the battery's external characteristics. Through the amalgamation of this model with substantial state data and the application of circuit analysis, the ECM method proves to be a powerful tool for accurately identifying the health state of a battery [6]. ECM is computationally efficient, making it the primary battery model employed for online capacity estimation in battery management systems (BMSs) for EVs. However, its accuracy is frequently hindered by the performance of model parameterization.

The data-driven approach offers a solution to the limitations of the aforementioned methods, as it does not heavily rely on parameters. This approach does not require an understanding of the internal principles of the battery or the construction of a physical battery model [7]. This method relies on collected battery aging data to extract and analyze information related to battery capacity degradation, enabling the estimation of the SOH for LIBs [8,9]. Since pure data-driven methods for capacity estimation lack transparency and interpretability [10], embedding physical or domain knowledge into the development of machine learning models has received more attention [11,12]. The use of a physics-informed neural network (PINN) is a classic method in this type of solution. It can fully integrate the physical characteristics of the problem into the loss function, thereby improving the generalization performance of data-driven models by utilizing data and mathematical operators based on underlying physical principles. Wen et al. [13] proposed a semi-physical and semi-empirical dynamic model to capture the capacity degradation of batteries, and based on this, they introduced a PINN to fuse the prior information of the dynamic model with the information extracted from monitoring data. Hofmann et al. combined simulated data generated by the pseudo-2-dimensional (P2D) model with real-world data for data-driven model training to reduce the computational cost of traditional physical and chemical models [14].

Data-driven methods can be broadly categorized into feature-based models and learning-based models. The accuracy of SOH estimation in feature-based models heavily relies on the extraction of relevant features. By employing correlation analysis, more features can be incorporated, leading to more accurate capacity estimation. One prominent feature-based method is incremental capacity curve analysis (ICA), which effectively quantifies battery capacity loss [15,16]. The ICA method is derived from the voltage differential capacity and extracts the incremental capacity (IC) curve from the discharging and charging phases. Zhang et al. [17] extracted peak values from the discharge IC curve for SOH estimation, while Li et al. [18] extracted the peak and its position from the charging IC curve as the health characteristics. Although IC analysis can explain the aging behavior inside the battery, it needs to be differentiated to filter the data, which increases the difficulty of data processing [19]. In addition, the noise in the data point is easily enlarged, resulting in a decrease in estimated accuracy. For the feature selection level, Li et al. [20] extracted the features of the IC curve while taking the proportion of the constant current (CC) mode, the equal voltage drop, and the sample entropy of the discharge voltage as health features. Wang et al. [21] extracted three health features under CC-CV charging and CC discharge modes: the time corresponding to the same charging voltage interval, the time corresponding to the same discharge voltage interval, and the time corresponding to the same discharge temperature change.

The learning-based model aims to simplify or eliminate the feature extraction process by directly inputting a certain length of charging segment data as model input [22], mainly using a convolutional neural network (CNN) for feature extraction. Yang et al. [23] used the CNN to extract the SOH index and the change in capacity between two consecutive charge–discharge cycles (Δ SOH) and used the random forest algorithm to obtain the final capacity degradation estimation result. The correlation of features has a great influence on

the accuracy of the model. Sun et al. [24] introduced an attention mechanism, by giving important feature weights, so that the key features in the input sequence that have a greater impact on the output results.

However, the existing data-driven methods are mostly based on feature extraction under relatively ideal conditions in the laboratory (the peak value of the IC curve, the voltage interval corresponding to the peak value, etc.), or the complete external characteristic curve (voltage, current, etc.) is used as the input of the data-driven model. Although this can achieve good results, it has certain limitations in real vehicle applications. First of all, the EVs are usually not fully charged and discharged, and the charging behaviors are complex and random. Secondly, in actual use and battery experiments, there are problems such as long experimental time, high cost, and difficulty in obtaining labels [25,26]. Furthermore, the effectiveness of the data-driven method depends on a substantial volume of historical data, which can be challenging to obtain in BMSs of EVs in practical conditions. Therefore, there is significant research value in developing a capacity estimation model that is suitable for real-world scenarios and can accurately predict capacity degradation using a limited amount of data.

For this issue, research on capacity estimation based on incomplete fragment inputs has shown great research value. For example, Tian et al. [27] used a small portion of the charging curve as input to estimate the complete charging curve and then used the predicted charging curve as an input feature for capacity estimation. In addition, Zhang et al. [28] introduced shapelets to further partition the charging segments based on the extraction of short charging segments and used the MRMR algorithm to find the optimal shapelet as the model input, in order to greatly reduce the input volume of the model. For relatively fixed fragment inputs, the model can learn the mapping relationship between features and capacity degradation well.

Considering that the state of charge (SOC) at the beginning (end) of charging is not fixed in actual vehicle operation scenarios, this article uses fixed charging segments and random charging segments as inputs for research. In addition, this paper introduces a flexible and robust data-driven method capable of achieving accurate estimation of battery capacity degradation across diverse battery datasets, even with limited samples available. The overall framework of this method is illustrated in Figure 1.

In this method, a learning-based capacity estimation model is first created using a gated convolutional and bi-directional long short-term memory hybrid neural network (GateCNN-BiLSTM). The CNN is used to capture the advanced spatial features of the battery sequence [29], and the bi-directional long short-term memory (BiLSTM) neural network is used to extract the positive and negative context information of the battery sequence [30]. The gating mechanism can suppress noise information and enhance the sensitivity of the network to important features, thereby improving the generalization ability of the network. The model can use easily measured data (such as voltage and current) recorded in any short charging time to estimate the current battery capacity state. Secondly, based on the learning-based model, combined with the capacity estimation framework of single-source domain and multi-source domain migration, the model also shows excellent performance under a very small amount of data.

In summary, the main contributions of this paper include the following three points:

- The proposed model has a flexible input and output structure and can accurately estimate the current capacity state based on short charging data segments. This method does not depend on the laboratory aging test environment of the complete charge and discharge cycle and avoids the problem that the existing data-driven methods need to perform artificial feature extraction. Therefore, this method has wide applicability in actual vehicle charging scenarios.
- Through the verification of different LIB aging datasets, it is proved that the learning-based model is robust to datasets of various battery types and charge–discharge conditions. In the case of random charging segment input, the minimum root mean

square error (RMSE) and mean absolute error (MAE) of the capacity estimation results are 0.384% and 0.339%, respectively.

- The transfer learning framework proposed in this study shows good results in the case of a very small amount of data, which significantly improves the training speed and stability. It overcomes the problem of the capacity label being difficult to obtain and the labeling cost being high in practical applications, and it further enhances the robustness and versatility of the method.

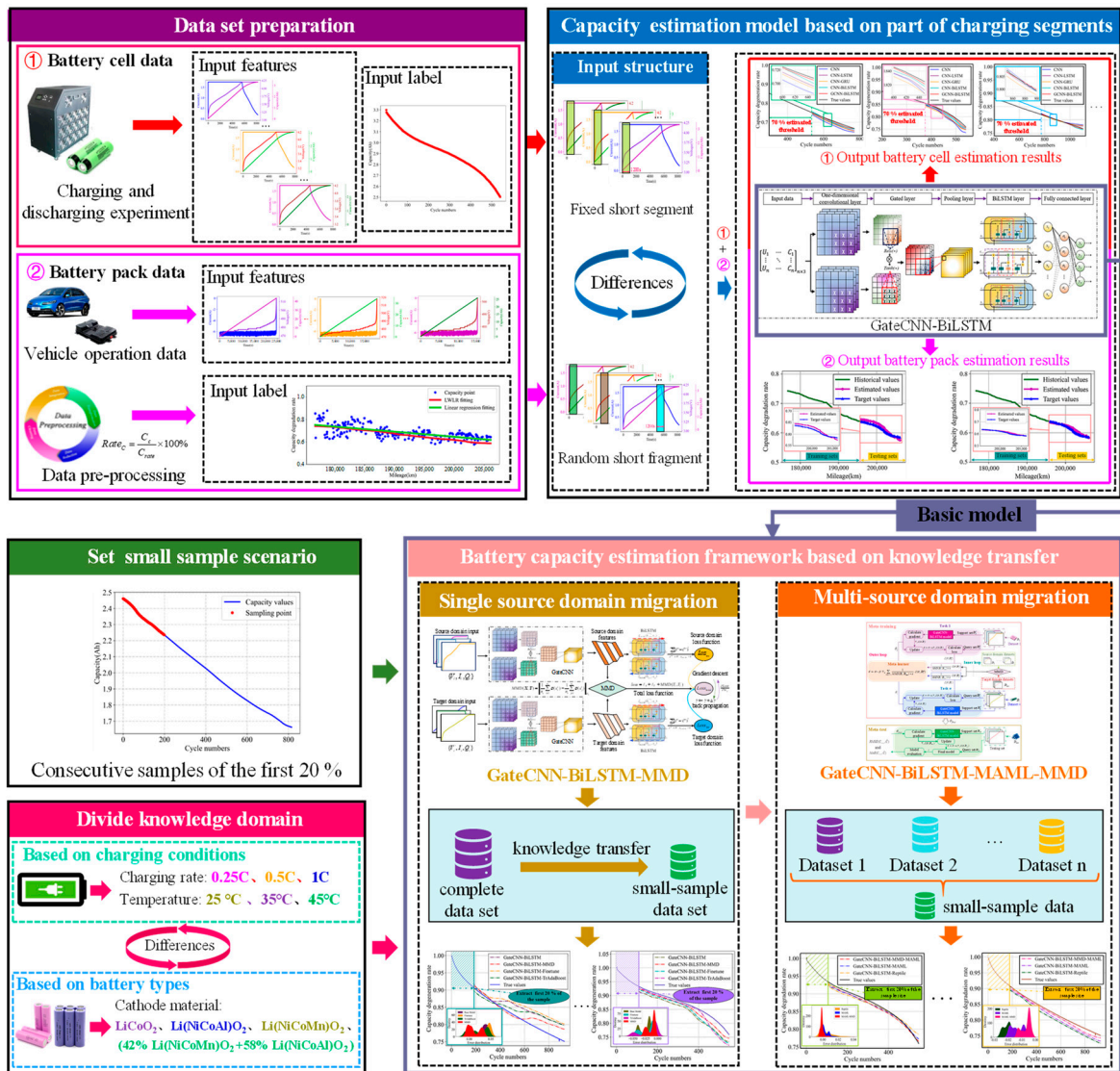


Figure 1. Overall framework of battery capacity estimation method.

The remaining sections of this paper are organized as follows: Section 2 presents the lithium battery dataset used in this study. Section 3 introduces the basic estimation model for lithium battery capacity degradation estimation, focusing on the utilization of partial charging segments as input and discussing different output structures. In Section 4, we propose a transfer learning framework that combines the basic capacity degradation estimation model with small-sample learning to enhance the model’s generalization capability in scenarios with limited samples. Section 5 presents the experimental results and analysis. Finally, Section 6 concludes the paper.

2. Lithium-Ion Battery Dataset

2.1. Battery Cell Dataset

The datasets involved in this study include battery cell data obtained in the laboratory and battery pack data collected for EVs. The battery cell datasets were obtained from the aging experiments conducted by Professor Dai Haifeng's team at Tongji University [31] and the University of Maryland [32]. Table 1 displays the information regarding battery cell data. The Tongji University dataset is divided into three subsets, namely NCA battery, NCM battery, and NCA+NCM battery, based on the distinct positive electrode materials used in the batteries. The University of Maryland dataset comprises CS2 and CX2 batteries, both of which utilize LiCoO_2 as the cathode material.

Table 1. Overview of battery cell data.

Battery Type	Rated Capacity (Ah)	Temperature ($^{\circ}\text{C}$)	Charging Protocol	Cut-Off Voltage (V)	Charging/Discharging Rate (C)
NCA	3.5	25/35/45	CC-CV	2.65–4.2	0.25/1, 0.5/1, 1/1
NCM	3.5	25/35/45	CC-CV	2.5–4.2	0.5/1
NCA+NCM	2.5	25	CC-CV	2.5–4.2	0.5/1, 0.5/2, 0.5/4
CS2_	1.1	20/25	CC-CV	2.7–4.2	1
CX2_	1.35	25/35/45/55	CC-CV	2.7–4.2	0.5

In the dataset of Tongji University, the battery cells are labeled as CYX-Y/Z, with X representing the temperature, Y representing the charging current, and Z representing the discharge current based on the charge–discharge cycle conditions. In this study, the input for the battery capacity estimation model consists of current, voltage, and charging capacity. These three data features are derived from measured data, enabling end-to-end capacity estimation encompassing data acquisition, processing, model establishment, and final health state estimation. To address variations in data sampling caused by experimental design and test environment, linear interpolation is applied. This ensures that all charge and discharge cycle data are uniformly converted into time series data with a consistent 1 s time interval. Additionally, the capacity label's maximum value in each dataset is normalized to achieve data uniformity.

The model training utilized datasets from NCA, NCM, and NCA + NCM batteries, which were further evaluated using transfer learning to enhance the generalizability of the proposed method. Figure 2 illustrates the evolution of voltage, current, and capacity during charging across different cycle periods. It is evident that distinct variations exist in charging voltage, current, and capacity for different aging states at the same charging duration. Moreover, throughout the entire charging cycle, these three variables display noticeable differences in their respective time series. This observation signifies the significance of utilizing the time series of charging voltage, current, and capacity as crucial features for accurately estimating battery capacity degradation.

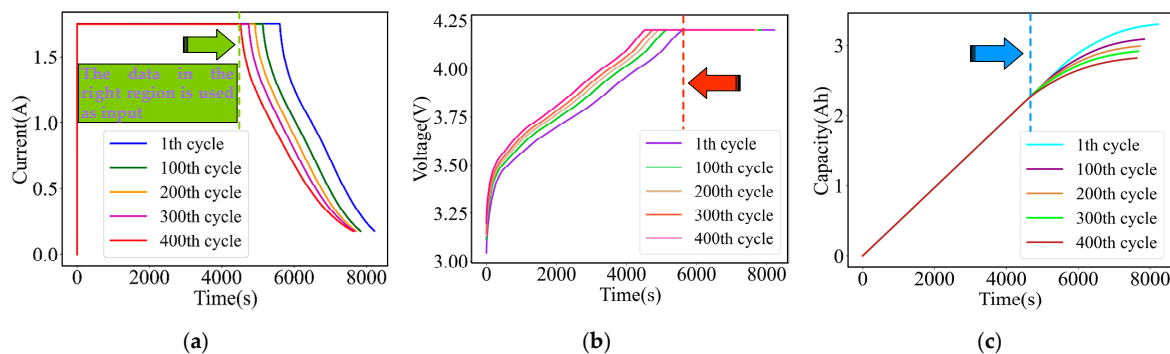


Figure 2. Evolution of battery voltage, current, and capacity during charging under different charge–discharge cycle durations: (a) current; (b) voltage; (c) capacity.

2.2. Battery Pack Dataset

Based on the previous research [33], we extracted the charging segments from January to November 2020 from the EV operation data. Figure 3 illustrates that SOC of the battery pack varies at the beginning and end of charging, contrasting with controlled charging and discharging experiments conducted in laboratory settings. The charging behavior of the user shows great randomness. Therefore, it is of great significance to extract the random segments of the charging data as the input of the battery charging capacity estimation model.

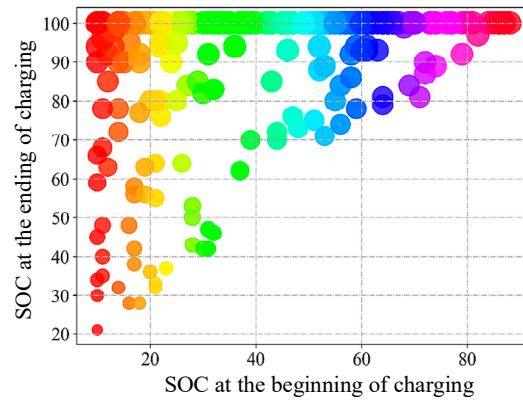


Figure 3. A distribution map of SOC at the beginning (end) of charging for an electric vehicle.

The health status of the battery is greatly affected by the environment and load conditions, making battery degradation a highly dynamic process. Generally, capacity degradation and resistance increase are two indicators used to assess the health status of LIBs. In this study, the capacity degradation trend of the battery is quantified through the maximum charging capacity observed in each charging cycle. Additionally, a capacity degradation rate is defined to further characterize the degradation process:

$$Rate_C = \frac{C_c}{C_{rate}} \times 100\% \tag{1}$$

where C_c is the current maximum rechargeable capacity of the battery; C_{rate} is the rated charge capacity of the battery.

In actual EV operation conditions, it is not always possible to meet the full charging, so the ampere-hour integral method is used to update the battery SOC in real time by recording the amount of electricity accumulated during the charging process [34]. If the initial SOC of a charging segment is SOC_s , the terminated SOC of the segment SOC_e is

$$SOC_e = SOC_s + \frac{1}{C_c} \int_s^e \eta I(t) dt \tag{2}$$

In the formula, η is the coulombic efficiency, $\eta = 1$ in the charging process; $I(t)$ is the current value at time t . An inverse transformation of the above formula is performed to obtain the maximum available capacity C_c :

$$C_c = \frac{\int_s^e \eta I(t) dt}{SOC_e - SOC_s} \tag{3}$$

The data collected by the EV are not continuous, so the integral of the Equation (3) needs to be transformed into a discrete form:

$$C_c = \frac{\sum_s^e \eta I \Delta t}{SOC_e - SOC_s} = \frac{\sum_s^e \eta I \Delta t}{\Delta SOC} \tag{4}$$

Currently, SOC update algorithms are typically internally set by the BMS. The accuracy of SOC collection can significantly impact capacity calculation results. To minimize this impact, this study follows the methodology presented in references [34,35] and increases the charging duration to enable the battery capacity estimation model to fully learn the information. As shown in Figure 3, the random charging behavior of electric vehicles results in unequal changes in SOC (ΔSOC) during each charging cycle. Equation (4) demonstrates that the accuracy of C_c is heavily influenced by ΔSOC . The smaller the ΔSOC , the greater the impact of SOC accuracy on C_c . Therefore, this paper introduces ΔSOC as a constraint:

$$\Delta\text{SOC} = \text{SOC}_e - \text{SOC}_s \geq 10\% \quad (5)$$

In order to mitigate battery instability at a low SOC and minimize the impact of small current at a high SOC, it is recommended to set the SOC of the charging segment between 25% and 95%.

In addition, temperature is one of the main factors affecting the current available capacity of the battery. Specifically, in the low-temperature environment, the lithium-ion activity in the battery is low, while in the high-temperature environment, the lithium-ions react faster between the electrodes [36]. Therefore, based on the capacity value at 25 °C, the corresponding temperature influence coefficient is obtained by referring to the initial capacity values of the battery cell at different temperatures (0 °C, 10 °C, 25 °C, 45 °C). The relationship between temperature and influence coefficient is as follows:

$$\delta(i) = -6.667 \times 10^{-7}T(i)^3 - 2.600 \times 10^{-5}T(i)^2 + 1.047 \times 10^{-3}T(i) + 0.968 \quad (6)$$

$$C_T = C_{\text{initial}} \times \delta \quad (7)$$

In Equations (6) and (7), $T(i)$ is the average temperature of the battery pack in the i -th charge–discharge cycle; $\delta(i)$ is the temperature influence coefficient of the i -th cycle, C_{initial} is the calculated initial capacity, and C_T is the capacity after temperature correction. The corrected capacity value of the real vehicle battery pack is shown in Figure 4.

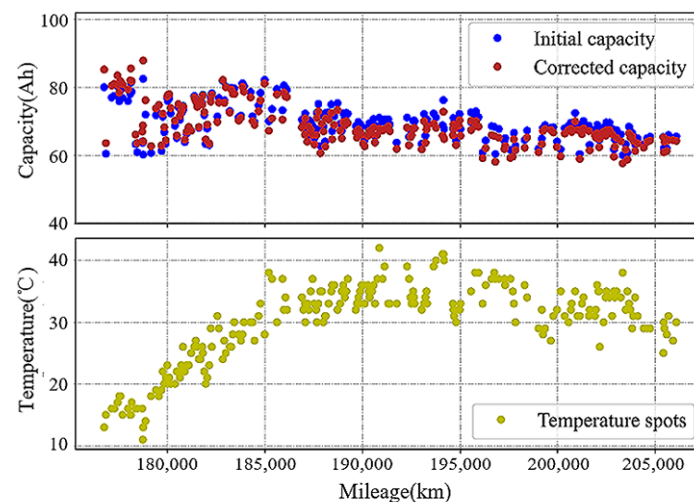


Figure 4. Temperature correction capacity of the real vehicle battery pack.

The linear regression method is one of the common methods for fitting measurement parameters [37]. However, the degradation process of the battery is nonlinear, so it is obviously difficult to ensure accuracy and robustness by using only linear regression methods. The locally weighted linear regression (LWLR) algorithm not only retains the simplicity of linear least squares regression, but also has the wide applicability of nonlinear regression. It adapts to the nonlinear relationship of data by fitting a local linear regression model around each data point. Therefore, in this paper, we use the LWLR algorithm to

perform autoregressive processing on the battery capacity, and we use the goodness of fit R^2 to evaluate the results. The fitting results are shown in Figure 5. Through calculation, it can be found that the R^2 of the LWLR method is 0.847, while the R^2 of linear regression is only 0.440. Therefore, the LWLR method can better cope with the nonlinear characteristics of battery capacity degradation and improve the accuracy of capacity evaluation.

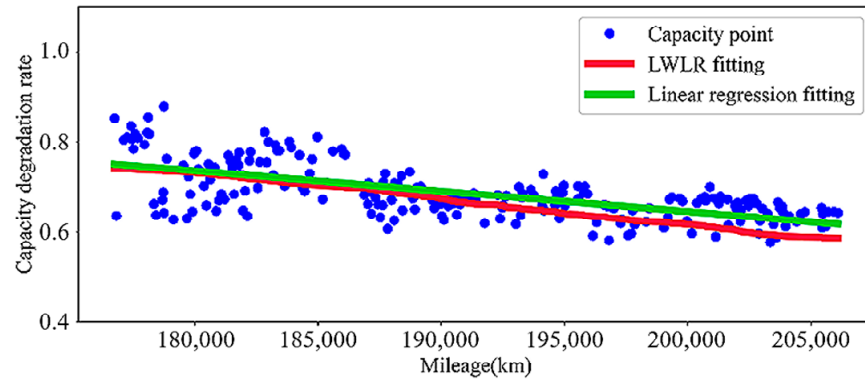


Figure 5. Battery capacity degradation fitting curve.

3. Capacity Degradation Estimation Model Based on Partial Short Segments

3.1. Input and Output Structure

For the estimation of capacity degradation, under the premise of the complete charge and discharge hypothesis, feature extraction is more convenient, and the method of battery capacity estimation based on charge and discharge characteristics is also more mature [38,39]. However, the charging and discharging behavior of the battery during the vehicle operation conditions has a certain complexity and randomness. Considering that the battery is usually only partially charged in the application and/or charged at a higher current, it does not meet the conditions of complete charging and discharging in many cases [40].

Therefore, this paper designs two different input structures for comparison, as shown in Figure 6. In structure 1 of Figure 6a, the fixed segment of the charging data is selected as the input, and the period is the initial 1200 s of the charging data (about 14% of the whole charging segment). In structure 2 of Figure 6b, the random segment of the charging data is selected as the input, which makes the input more flexible and allows end-to-end capacity degradation estimation.

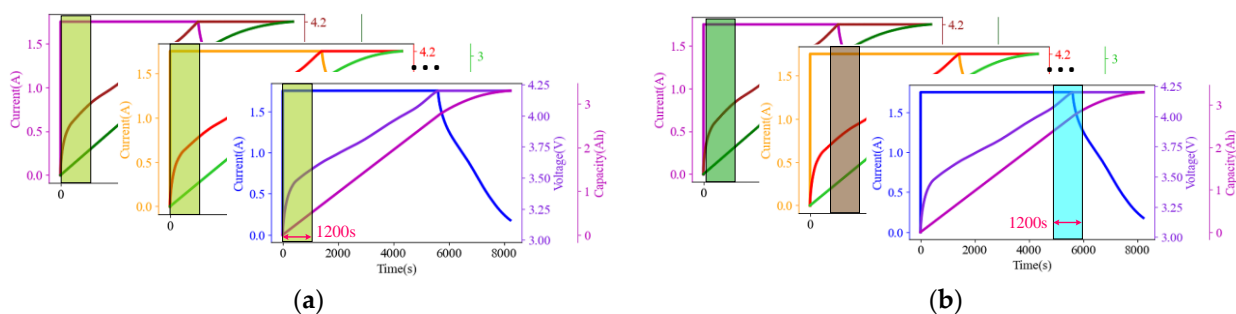


Figure 6. Input structure of capacity estimation model: (a) fixed segment input; (b) random segment input.

For the selection of random segments, in the battery experimental data in Section 2.1, it is observed that during the battery aging process, there is a significant mapping relationship between the current maximum available capacity of the battery and the voltage, current, and charging capacity during the charging process. Therefore, 1200 s of data are randomly selected from the voltage, current, and charging capacity of the charging section as the

input variables of the capacity estimation model. These input variables are organized into a 1200×3 matrix, where each column corresponds to one of the three variables (voltage, current, charging capacity). Therefore, the input of the model can be represented by the following matrix:

$$M_{input} = \begin{bmatrix} V_1 & I_1 & Q_1^c \\ V_2 & I_2 & Q_2^c \\ \vdots & \vdots & \vdots \\ V_i & I_i & Q_i^c \\ \vdots & \vdots & \vdots \\ V_L & I_L & Q_L^c \end{bmatrix} \quad (8)$$

where V_i , I_i , and Q_i^c are the voltage, current, and charging capacity of the input data at the i -th second, respectively.

The current output can be expressed as follows:

$$Rate_C = \frac{f_{GateCNN}(M_{input})}{C_{rate}} \quad (9)$$

where C_{rate} is the rated charging capacity of the battery; $f_{GateCNN}$ is the parameterized function of the GateCNN model.

3.2. GateCNN-BiLSTM Model

To improve the applicability of the battery capacity estimation model in practical applications, this paper introduces a deep learning model GateCNN-BiLSTM based on a gated convolutional and bi-directional long short-term neural network. The model structure is shown in Figure 7.

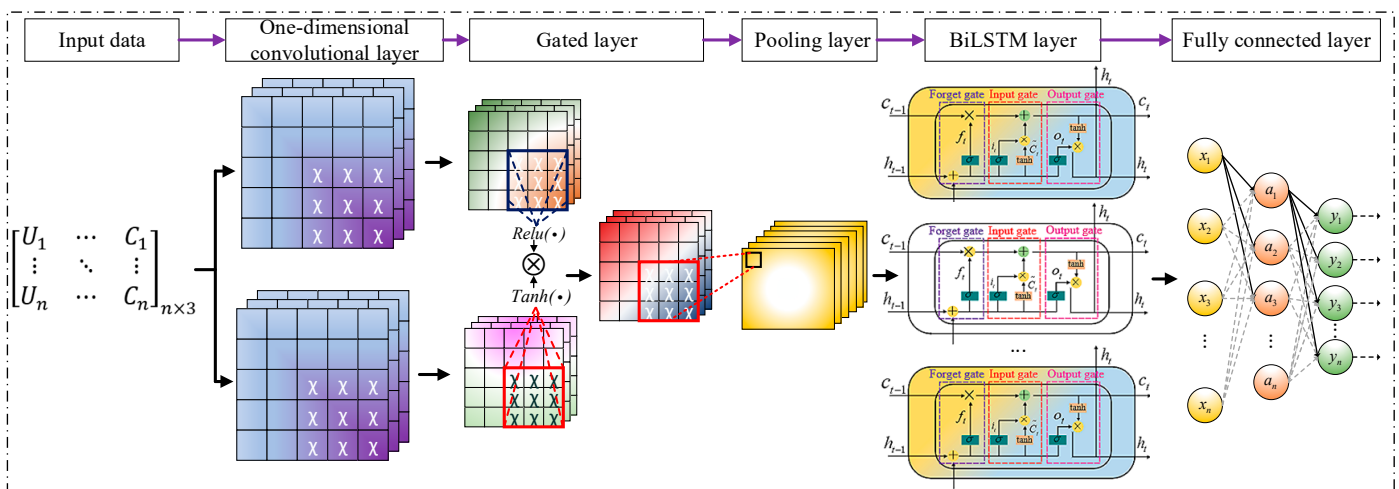


Figure 7. GateCNN-BiLSTM model framework.

Based on the feature extraction of traditional convolutional networks, this paper draws on the gated recurrent unit (GRU) [41–43], and introduces a gating mechanism. This gating mechanism can dynamically adjust the activation state of the network according to the current input data, suppress noise information, and enhance the sensitivity of the network to important features, thereby improving the generalization ability of the network. We combine it with CNN so that it can automatically extract information-rich features from large-scale charging data.

Traditional recurrent neural network (RNNs) and long short-term memory (LSTM) networks are unidirectional, which means that they can only capture the historical information of time series while ignoring the future information in time series. To obtain

more complete time series information, we introduce a BiLSTM network. The network is improved based on the LSTM network.

As shown in Figure 8, the hidden layer of the BiLSTM neural network consists of two independent LSTM hidden layers. The forward layer is used to capture the historical information of the time series data, and the reverse layer is used to capture the future information of the time series data. Therefore, compared with the one-way LSTM neural network, the BiLSTM network can simultaneously mine the past and future information of time series data, thereby improving the utilization of data.

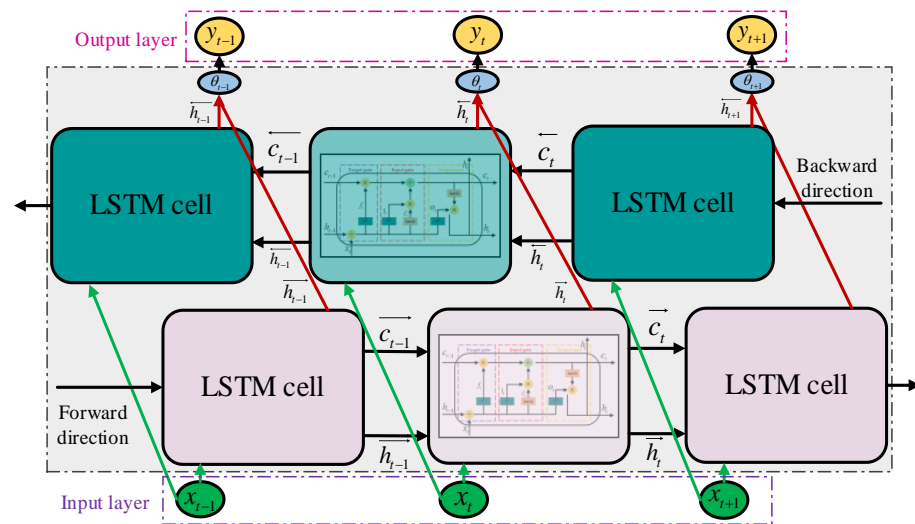


Figure 8. BiLSTM structure diagram.

The input data are first initialized by two convolutional layers to extract local features. The output of the convolutional layer passes through a gated recurrent layer, and the output of the recurrent layer is processed by a fully connected layer.

The convolutional layer is used to convolve the input sequence and extract features from the sequence. The one-dimensional convolutional layer consists of N parallel W -dimensional convolutional kernels, where N and W are hyperparameters. Firstly, the number of types of input features is F , and the length of input segments is L . We fill $(w - 1)/2$ zero vectors before and after the input matrix M_{input} to obtain $M_{padding}$, and the sequence is shown in Equation (10) as follows:

$$M_{padding} = [x_1, x_2, \dots, x_{w+F}] \tag{10}$$

where $M_{padding} \in R^{L \times (w+F)}$. It is known that the n -th convolutional kernel is $K^{(n)} = [k_1, k_2, \dots, k_w]$, and the sequence $y^{(n)}$ is output after convolution with the n -th convolutional kernel, as shown in Equation (11):

$$y^{(n)} = K^{(n)} * M_{padding} + b^{(n)} \tag{11}$$

where $b^{(n)}$ is the bias vector, $*$ is the convolution operation, $y^{(n)} \in R^5$, and $y_i^{(n)}$ is

$$y_i^{(n)} = \sum_{j=1}^w k_i^{(n)} \cdot x_{(i+j-1)} + b^{(n)} \tag{12}$$

The final output sequence is

$$y_{conv} = [y^{(1)}, y^{(2)}, \dots, y^{(n)}]^T \tag{13}$$

The activation function is a key element in the gated CNN because it is responsible for introducing nonlinearity, thereby increasing the capacity of the model. The model uses two commonly used neural network activation functions, ReLU and Tanh.

The gating layer is used to filter redundant feature information and reduce model complexity. The output results are as follows:

$$y_{gate} = \text{Tanh}(y_{conv}^{(1)}) \otimes \text{ReLU}(y_{conv}^{(2)}) \tag{14}$$

where $y_{conv}^{(1)}$ and $y_{conv}^{(2)}$ are two independent convolutional layers; \otimes represents the element-by-element product of the matrix; y_{gate} is the output sequence of the gating layer. After maximum pooling of the gating layer, $y_{maxpool}$ is exported and imported into BiLSTM. For BiLSTM, $x_{l,t}$ is the input at time t of the l -th layer, and $h_{l,t-1}$ is the output at the time $t-1$. The BiLSTM module consists of a forward layer and a reverse layer. The outputs of the forward layer $\overrightarrow{h_{l,t}}$ and the reverse layer $\overleftarrow{h_{l,t}}$ are

$$\overrightarrow{h_{l,t}} = f_{BiLSTM}(h_{l,t-1}, x_{l,t}) \tag{15}$$

$$\overleftarrow{h_{l,t}} = f_{BiLSTM}(h_{l,t-1}, x_{l,t}) \tag{16}$$

In Equations (15) and (16), f_{BiLSTM} is the BiLSTM algorithm.

According to the output of the forward layer and the reverse layer, the output of the BiLSTM network layer at the moment can be obtained as follows:

$$h_{l,t} = \delta(\overrightarrow{h_{l,t}}, \overleftarrow{h_{l,t}}) \tag{17}$$

In Equation (17), $\delta(\cdot)$ is a linear function for combining the forward layer and the reverse layer.

$$y_{BiLSTM,t} = [h_{1,t}, h_{2,t}, \dots, h_{L,t}]^T \tag{18}$$

The linear layer performs linear transformation on the input data, which can be expressed as follows:

$$y_{out,t} = W \cdot y_{BiLSTM,t} + b \tag{19}$$

In Equation (19), $W \in R$ is the weight vector of the fully connected layer, and $b \in R$ is the bias vector.

4. Capacity Estimation Method Based on Small-Sample Learning

4.1. Small-Sample Scenario Division

The blue curve in Figure 9 represents the capacity degradation curve of the target domain. To verify the knowledge transfer performance of LIBs' health status in the case of small samples, the corresponding small-sample scenarios in the target knowledge domain data are set. As shown in the red dots in Figure 9, the sample data of the target domain only contain limited information and fail to fully carry the knowledge of the entire information space. This means that the GateCNN-BiLSTM model cannot accurately infer the knowledge that is not included in the complete information space. At this time, the model faces the challenge of incomplete knowledge and needs to be solved by transfer learning. Given this situation, the first 5%, 10%, 15%, and 20% charging cycle samples in the target domain are selected as the training set of the target domain to study the influence of different sample ratios on the capacity estimation results.

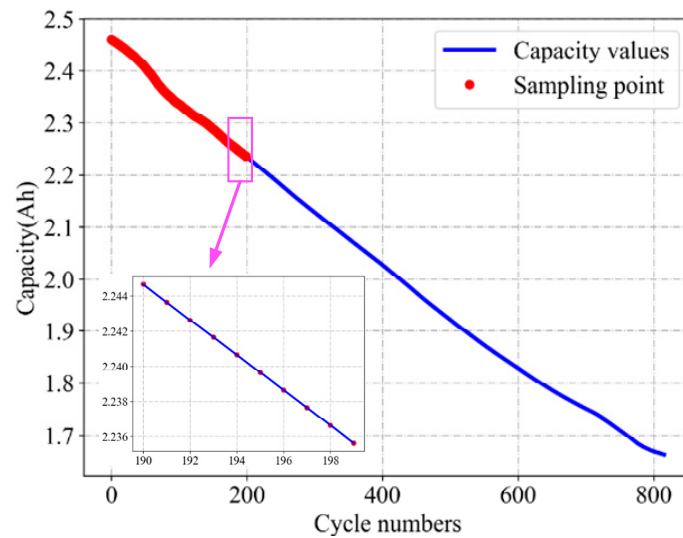


Figure 9. A small-sample scenario with only partial charging cycle capacity data.

4.2. Capacity Estimation Method of Single-Source Domain Migration Based on Domain Adaptation

To improve the performance of the basic capacity estimation model when the battery training samples are scarce, a domain adaptive transfer learning algorithm based on multi-core maximum mean difference (MMD) is proposed. By transferring the relevant feature knowledge of other battery datasets to the capacity estimation model, the information deficiency caused by insufficient training data is compensated for. In addition, this method also breaks the limitation that the training set and the test set of the neural network must be identically distributed, and it further expands the application range of the algorithm.

The MMD-based transfer learning algorithm integrates the MMD metric into the basic model GateCNN-BiLSTM of transfer learning to form the GateCNN-BiLSTM-MMD model. When the basic model extracts the features of the source domain and the target domain, the MMD is used to reduce the distribution difference between the source domain and the target domain features, so that the basic model can learn the features of the source domain and the target domain at the same time and realize the transfer of the source domain feature knowledge to the target domain. The transfer learning framework based on MMD is shown in Figure 10.

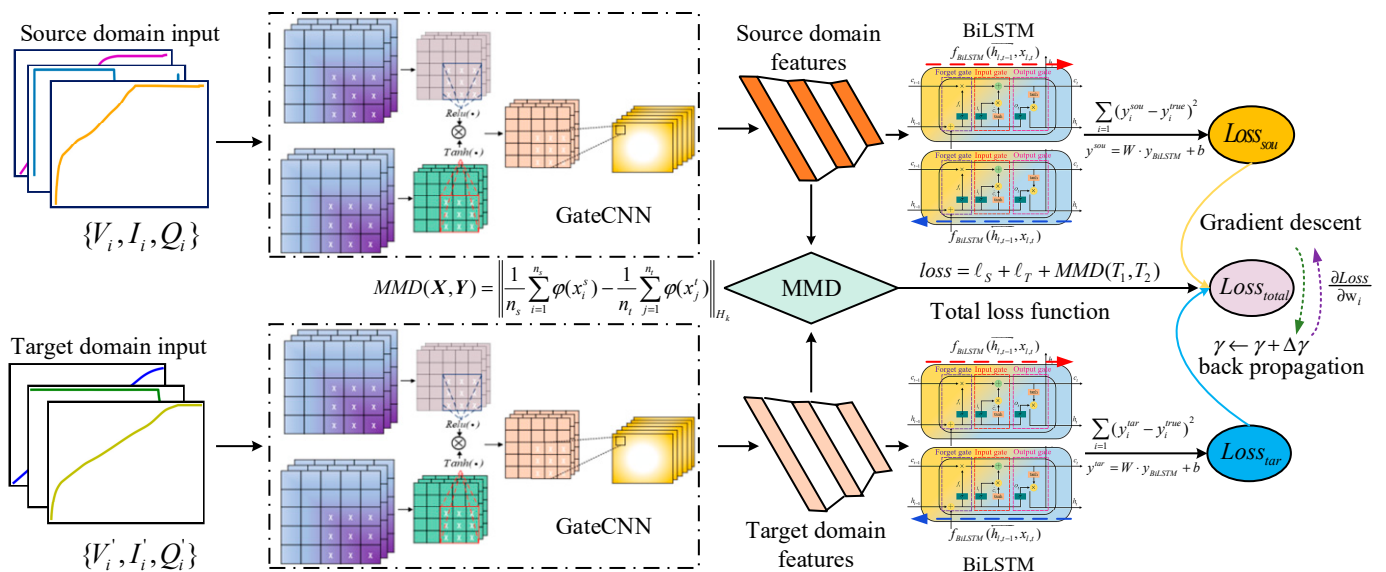


Figure 10. Transfer learning framework based on multi-core MMD.

The main idea of the MMD algorithm is to map the samples of the two datasets to a high-dimensional space through a specific mapping function, that is, the reproducing kernel Hilbert space (RKHS) [44]. In RKHS, the samples of the two datasets are separated, and the MMD value is the Euclidean distance between the average data points in the two datasets [45].

MMD is widely used in transfer learning algorithms to quantify the distribution difference between the source domain and the target domain [46,47]. When $MMD = 0$, it means that the distribution of the two domains is the same. On the contrary, the larger the MMD value, the greater the distribution difference between the two domains. According to the main idea of MMD, its definition is as follows:

$$MMD(X, Y) = \left\| \frac{1}{n_s} \sum_{i=1}^{n_s} \varphi(x_i^s) - \frac{1}{n_t} \sum_{j=1}^{n_t} \varphi(x_j^t) \right\|_{H_k} \tag{20}$$

where X and Y represent the sample space of two different distributions; x_i^s and x_j^t represent the samples extracted from different sample spaces; H_k represents RKHS; n_s and n_t represent the number of samples; $\varphi(\cdot)$ represents the nonlinear function of data mapping to RKHS. However, it is difficult to find a suitable mapping function, so the kernel trick is used to deal with it. The common kernel trick is as follows:

$$k(x^s, x^t) = \langle \varphi(x^s), \varphi(x^t) \rangle \tag{21}$$

Therefore, Equation (20) can be converted to:

$$MMD(X, Y) = \frac{1}{n_s^2} \sum_{i=1}^{n_s} \sum_{j=1}^{n_s} k(x_i^s, x_j^s) - \frac{2}{n_s n_t} \sum_{i=1}^{n_s} \sum_{j=1}^{n_t} k(x_i^s, x_j^t) + \frac{1}{n_t^2} \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} k(x_i^t, x_j^t) \tag{22}$$

The battery capacity estimation algorithm based on MMD transfer learning is shown in Algorithm 1.

Algorithm 1 Domain adaptation capacity estimation based on MMD transfer learning

$$\text{Input: } X^S = \begin{bmatrix} V_1^S & I_1^S & Q_1^S \\ V_2^S & I_2^S & Q_2^S \\ \vdots & \vdots & \vdots \\ V_m^S & I_m^S & Q_m^S \end{bmatrix}_{m \times 3}^T \quad X^T = \begin{bmatrix} V_1^T & I_1^T & Q_1^T \\ V_2^T & I_2^T & Q_2^T \\ \vdots & \vdots & \vdots \\ V_n^T & I_n^T & Q_n^T \end{bmatrix}_{n \times 3}^T$$

Output: Capacity degradation rate $Rate_c$

1. Normalize X^S and X^T
2. Divide X^T into small-sample scenarios. That is, the first 20% of the samples are extracted as the target domain training set *train*, and the last 80% of the samples are used as the test set *test*.
3. X^S is used to form a source domain dataset for knowledge transfer.
4. **while** $loss > 0.0001$ **do**
5. Taking X^S and *train* as inputs, the target domain feature T_1 and the source domain feature T_2 are obtained by GateCNN-BiLSTM.
6. Calculate the maximum mean difference $MMD(T_1, T_2)$ of the two features.
7. The final loss value $loss = \ell_S + \ell_T + MMD(T_1, T_2)$ is calculated by superimposing the loss value obtained by each input variable.
8. **end while**
9. The predicted value $result = \text{model.predict}(\text{train})$ is obtained and *result* is normalized.
10. The root mean square error (RMSE) and mean absolute error (MAE) are calculated.
11. **Return** RMSE, MAE, model.

In this model, the battery task dataset $\{V_i^T, I_i^T, Q_i^T\}_{i=1}^{n_T}$ with a limited number of samples is used as the target domain dataset, and the complete battery dataset $\{V_i^S, I_i^S, Q_i^S\}_{i=1}^{n_S}$ with different distributions from the target domain is used as the source domain dataset. Based on the input data of the source domain and the target domain, the data characteristics

of the source domain and the target domain are obtained through the basic GateCNN-BiLSTM model. To measure the distribution difference between the source domain and the target domain features, the MMD values of the source domain and the target domain features are calculated. In addition, while calculating the MMD value, the source domain and target domain features are used as the input of the BiLSTM model to obtain the output values of the source domain and the target domain. Then, based on the output values of the source domain and the target domain, the loss functions of the source domain and the target domain are calculated by using the source domain label $\{y_i^S\}_{i=1}^{n_S}$ and the target domain label $\{y_i^T\}_{i=1}^{n_T}$. Finally, combining these two loss functions with the MMD value, the loss function of the GateCNN-BiLSTM-MMD model is obtained. During the model training process, by adjusting the weight of the CNN model, the feature distribution of the source domain will continue to approach the feature distribution of the target domain until the two distributions are the same. Finally, the source domain features that are distributed with the target domain are applied to the target domain model, and the migration of the source domain features to the target domain is realized.

Due to the differences in charging conditions and battery types, the capacity degradation mechanism of LIBs is not consistent, and the battery aging data do not meet the characteristics of an independent and identical distribution. In this paper, different charging conditions and different types of battery data are divided into different source domains, and the target domain is established for small-sample data of specific charging conditions and specific battery types. Finally, knowledge transfer is completed across domains based on transfer learning modeling. The knowledge domain division methods based on different charging conditions and battery types are shown in Tables 2 and 3, respectively. The numbers after the symbol # represent different datasets.

Table 2. Knowledge domain division based on different charging conditions.

	Battery Type	Charging Protocol	Temperature	Charging Current	Charging Voltage	Capacity
Source domain	NCA _{CY45-0.5/1#28}	CC-CV	45 °C	1.75 A	4.2 V	3.5 Ah
Target Domain	NCA _{CY35-0.5/1#1}	CC-CV	35 °C	1.75 A	4.2 V	3.5 Ah

Table 3. Knowledge domain division based on different battery types.

	Battery Type	Cathode Material	Temperature	Charging Current	Charging Voltage	Capacity
Source domain	NCA _{CY45-0.5/1#1}	Li(NiCoAl)O ₂	45 °C	1.75 A	4.2 V	3.5 Ah
Target Domain	NCM _{CY45-0.5/1#27}	Li(NiCoMn)O ₂	45 °C	1.75 A	4.2 V	3.5 Ah

As an index for measuring the difference in feature distribution, MMD gradually reduces the difference in feature distribution between the source domain and target domain by adjusting parameters in the iterative process of the model to achieve the goal of transfer learning. This process is realized by an iterative optimization model, which makes the model have better generalization performance and adaptability in the target domain.

4.3. Capacity Estimation Method of Multi-Source Domain Migration Based on Meta-Learning

For LIBs, a single-source domain usually can only provide limited battery capacity degradation information, and when the distribution difference between the source domain data and the target domain data is too large, the mobility will be low, and even negative migration will occur in extreme cases. This has great limitations for solving the lack of domain knowledge in the case of small samples. Therefore, this section will expand on the differences between charging conditions and battery types. Based on the theory of meta-learning, combined with the basic learner of GateCNN-BiLSTM for training, the capacity estimation modeling of LIBs is completed by fusing multi-source domain data and the knowledge transfer from multi-distributed sources.

4.3.1. Model-Agnostic Meta-Learning

Both meta-learning and transfer learning can be applied to small-sample learning. The difference between them is that transfer learning needs to transfer the corresponding knowledge from the source domain to the target domain, which is highly dependent on the source domain. If the data of the source domain and the target domain are completely different, the effectiveness of knowledge transfer will be affected [48,49]. The essence of meta-learning is to increase the generalization ability of the learner in multiple tasks. Through continuous learning and adaptation of each specific task, the model network has a certain learning ability. Even if the source task and the target task are very different, this learning ability can also be applied to new tasks [50,51].

The model-agnostic meta-learning (MAML) algorithm is adopted to learn the general initialization of model parameters in a set of related tasks and then fine-tune the initialization according to the data of specific tasks, thus allowing the model to quickly adapt to new tasks with only a small amount of additional data. The dataset in each task can be divided into a support set and a query set. The support set is used to update the parameters of a specific model. After the update, the model obtains the corresponding loss through the query set. Finally, the loss sum of all tasks is counted to provide gradient information for the update of the initial parameters of the model [52].

4.3.2. An Improved Meta-Learning Method Based on Weight Guidance

The ultimate goal of the MAML algorithm is to make the initial parameters of the model balance the learning performance of each task. It does not seek the best learning performance of a single task, but it makes the performance of each task better and seeks the minimum joint training loss of all tasks. This is out of generalization considerations. However, at the level of knowledge transfer, considering that there are differences between tasks, treating the source tasks equally will inevitably lead to the problem of a low knowledge transfer rate.

To improve the efficiency of knowledge transfer, this section improves the learning process of MAML from the correlation between the source task and the target task and proposes a meta-learning improvement scheme based on weight guidance. The MMD value is used to measure the data distribution similarity between the target task and the source task dataset, and it is multiplied by the loss as a key reference for the final parameter update to increase the proportion of tasks with greater correlation in the meta-training process, thereby improving the task training efficiency of MAML. Unlike the transfer learning method in Section 4.2, the MMD here is not embedded in the loss function of the model for iterative optimization, but as a definition of a similar correlation. It can be seen as a measure that assigns corresponding weights to different datasets to ensure that in the meta-learning framework, tasks that are more relevant to the test set can be considered more pertinently when determining the initial parameters of the final model. The new meta-training loss function is as follows:

$$\mathcal{L}(\theta) = \sum_{T_i \sim p(T)} \frac{1}{MMD(T_S, T_T) + \alpha} \mathcal{L}_{T_i}(f_{\theta'}) \quad (23)$$

where $\frac{1}{MMD(T_S, T_T) + \alpha}$ is the weight of each task, and α is a specific value; in this section, $\alpha = 1.0 \times 10^5$ to prevent the denominator from being 0. The smaller $MMD(T_S, T_T)$ is, the larger $\frac{1}{MMD(T_S, T_T) + \alpha}$ is.

In this paper, the improved meta-learning framework is named MMD-MAML. As shown in Figure 11, the meta-learning framework contains two loops. In the inner loop, each task is trained to obtain the training loss of each task. In the outer loop, the loss of all tasks is summarized, and then the gradient update is performed to obtain the optimal parameters of the model.

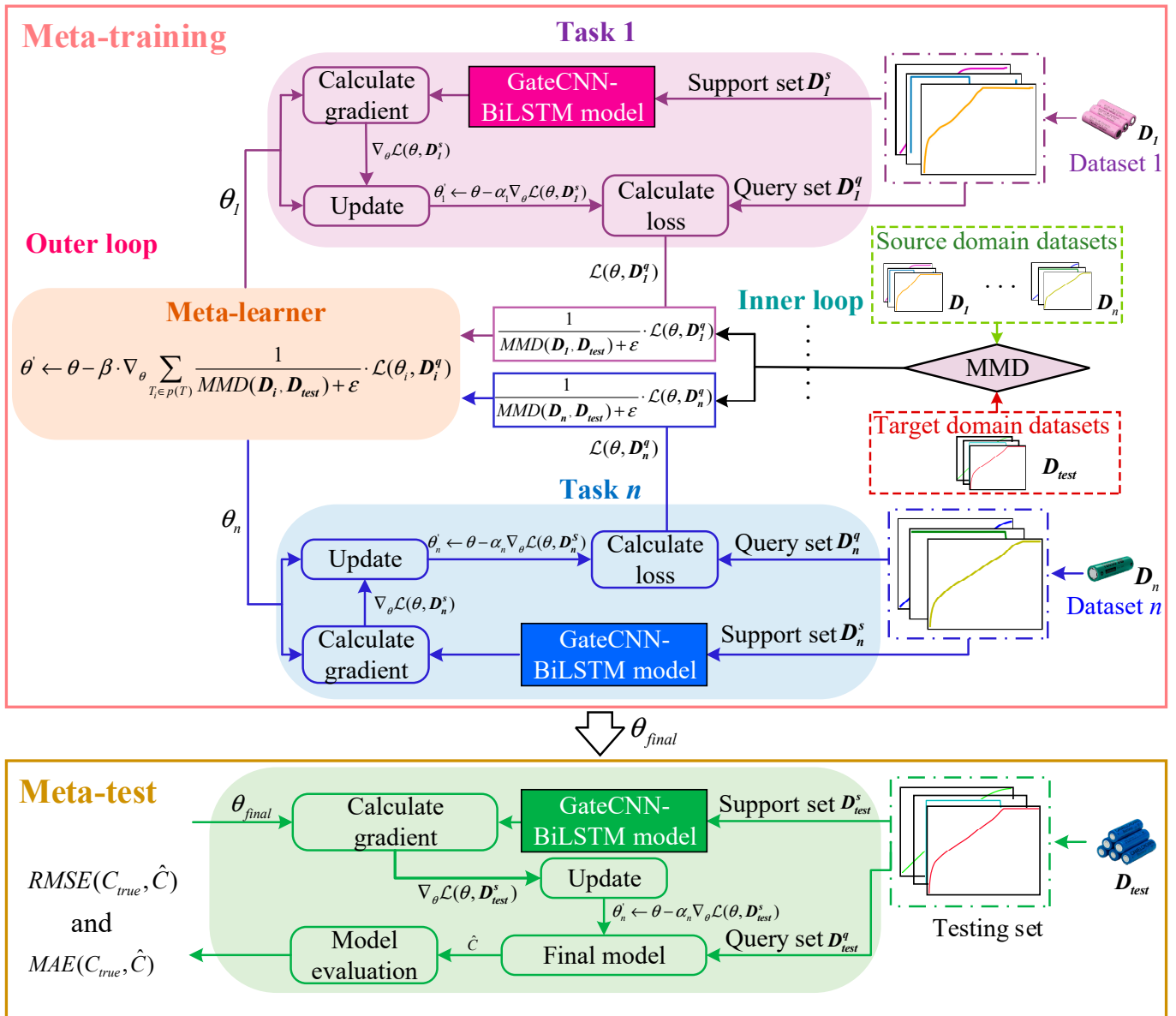


Figure 11. A multi-source domain knowledge transfer framework based on weight guidance.

Through the gradient update of the inner loop and the outer loop, when the model is learning in different tasks, the initial parameters have prior knowledge of the historical task. When there is a new task, this prior knowledge can make the model quickly adapt to the new task.

According to Algorithm 2, this section measures the domain of each source task and the target task, obtains the correlation between each auxiliary task and the target task, and converts the correlation into the weight of the task training loss to control the training effect of each task on the target task. By changing the loss function of MAML, the direction of updating the initial model parameters in the parameter space is controlled, that is, the loss weight of the target task is greater than the loss weight of the source task, so that the parameters of the initial model are updated as far as possible towards the optimal parameter position of the target task.

Algorithm 2 LIB capacity estimation algorithm based on MMD-MAML.

Input: Task distribution $p(T)$; inner and outer loop steps α and β ; multi-source domain migration framework f_{M-M}

Output: Capacity degradation rate $Rate_c$

1. **While not done do:**
2. **for** T_i **do:**
3. Extract specific task T_i from task distribution $p(T)$, and divide the dataset into support set and query set.
4. A task is iterated n times, and the support set is input for model training to obtain the corresponding loss $\mathcal{L}_{T_i}(f_{M-M}(\theta))$.
5. The model parameters are updated by gradient descent:

$$\theta'_i = \theta - \alpha \nabla_{\theta} \mathcal{L}_{T_i}(f_{M-M}(\theta))$$
6. Use the query set in T_i to test the accuracy of the updated model $f_{M-M}(\theta')$ and calculate the loss $\mathcal{L}_{T_i}(f_{M-M}(\theta'))$.
7. **end for**
8. Based on all θ'_i , the initial model parameters are updated:

$$\theta \leftarrow \theta - \beta \nabla_{\theta} \sum_{T_i \sim p(T)} \mathcal{L}_{T_i}(f_{M-M}(\theta'))$$
9. **end while**
10. **Return** Capacity degradation estimation model $Model_{meta}$.

4.3.3. Meta-Learning Task Division

Similar to the single-source domain small-sample migration scheme, in the multi-source domain migration scheme, the learning task is still set based on the difference in charging conditions and the difference in battery type. The specific division methods are shown in Tables 4 and 5, respectively. In the tables, # 1 represents the first dataset in a large class of batteries (NCA batteries), while * 1 and * 2 represent a small-sample scenario in a meta-testing dataset and the corresponding remaining samples, respectively.

Table 4. Task division based on charging condition difference.

	No.	Charging Rate	Discharge Rate	Temperature	Support Set	Query Set
Meta-training	Task ₁	0.5 C	1 C	45	NCA _{CY45-0.5/1#13}	NCA _{CY45-0.5/1#23}
	Task ₂	0.25 C	1 C	25	NCA _{CY25-0.25/1#3}	NCA _{CY25-0.25/1#4}
	Task ₃	1 C	1 C	25	NCA _{CY25-1/1#1}	NCA _{CY25-1/1#9}
	Task ₄	0.5 C	1 C	25	NCA _{CY25-0.5/1#10}	NCA _{CY25-0.5/1#18}
Meta-testing	Task _{tar}	0.5 C	1 C	35	NCA _{CY35-0.5/1#1} *1	NCA _{CY35-0.5/1#1} *2

Table 5. Task division based on battery type difference.

	No.	Cathode Material	Capacity (Ah)	Charging Rate	Support Set	Query Set
Meta-training	Task ₁ '	LiCoO ₂	1.1	0.5 C	CS2_33	CS2_34
	Task ₂ '	42% Li(NiCoMn)O ₂ +58% Li(NiCoAl)O ₂	2.5	0.5 C	(NCM + NCA) _{CY25-0.5/1#2}	(NCM + NCA) _{CY25-0.5/1#3}
	Task ₃ '	Li(NiCoAl)O ₂	3.5	0.5 C	NCA _{CY45-0.55/1#2}	NCA _{CY45-0.55/1#14}
	Task ₄ '	LiCoO ₂	1.35	0.5 C	CX2_35	CX2_37
Meta-testing	Task _{tar} '	Li(NiCoMn)O ₂	3.5	0.5 C	NCM _{CY45-05/1-#27} *1	NCM _{CY45-05/1#27} *2

Figure 12 shows the capacity degradation curve corresponding to each task dataset. It is worth emphasizing that the similarity of capacity degradation does not directly reflect the MMD value between knowledge domains, because MMD mainly focuses on the distribution of sample features without considering label information (i.e., capacity degradation). The features of different datasets are normalized to eliminate the influence of dimension. Then, the training weight of each task is calculated according to Equation (23) by calculating the MMD value of each dataset and the meta-test dataset. The smaller the MMD value is, the closer the feature distribution of the dataset is, and the larger the corresponding training weight is.

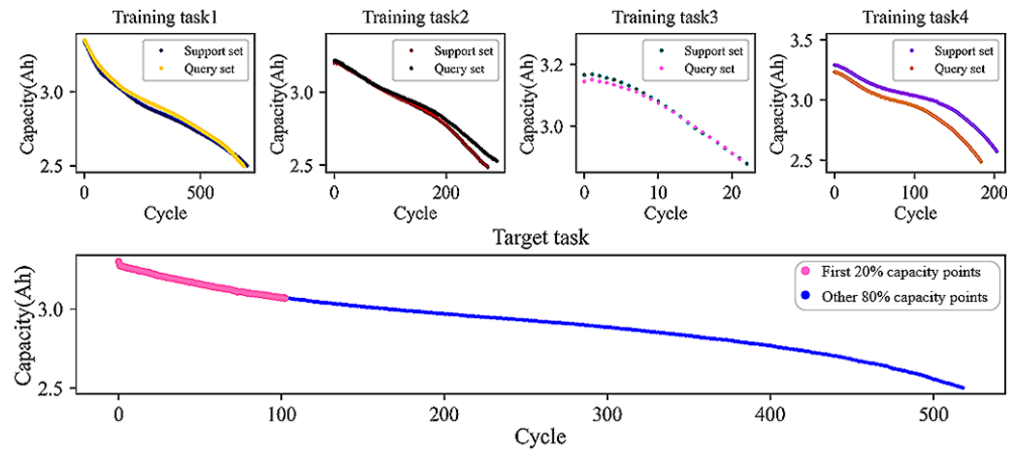


Figure 12. Capacity degradation curves of the corresponding datasets of meta-training and meta-testing based on the difference in charging conditions.

Figure 13 shows the weight ratio of different tasks in the case of fixed segment input and random segment input. It can be observed that there is a large correlation between the dataset of training task 2 and the dataset in the target task. Therefore, when performing the meta-training outer loop, training task 2 is given a greater weight to promote the parameter update more inclined to adapt to the dataset of task 2. This weight-guided meta-learning scheme helps to enhance the ability to learn the relevance of different tasks, thereby improving the performance of the meta-learning model.

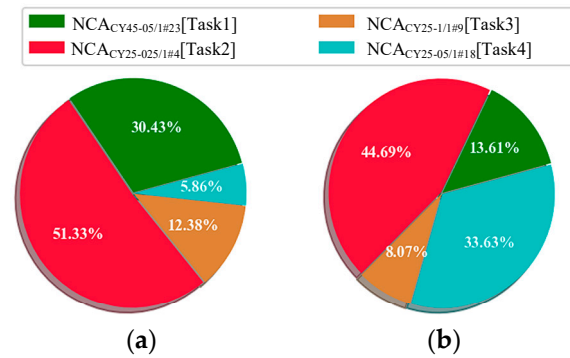


Figure 13. The correlation between meta-training task and meta-testing task dataset based on the difference of charging conditions: (a) fixed segment input; (b) random segment input.

In Section 4.2, knowledge transfer between two different battery types under the single-source domain framework is carried out. As shown in Table 5, according to the difference in battery type, four meta-training task datasets and one meta-testing target task dataset are divided. The battery capacity and cathode material between each task have a certain degree of difference, but the charging rate is uniformly selected at 0.5 C. Among them, the support set corresponds to the first $n\%$ small-sample scenario, and the query set corresponds to the remaining samples.

The capacity degradation curve presented by the dataset based on the difference in battery type is shown in Figure 14. It can be seen that the battery degradation curves corresponding to training tasks 2 and 3 show a certain similarity with the target dataset, which also increases the possibility of the model learning common features and patterns to a certain extent, thus improving the generalization performance on unknown tasks.

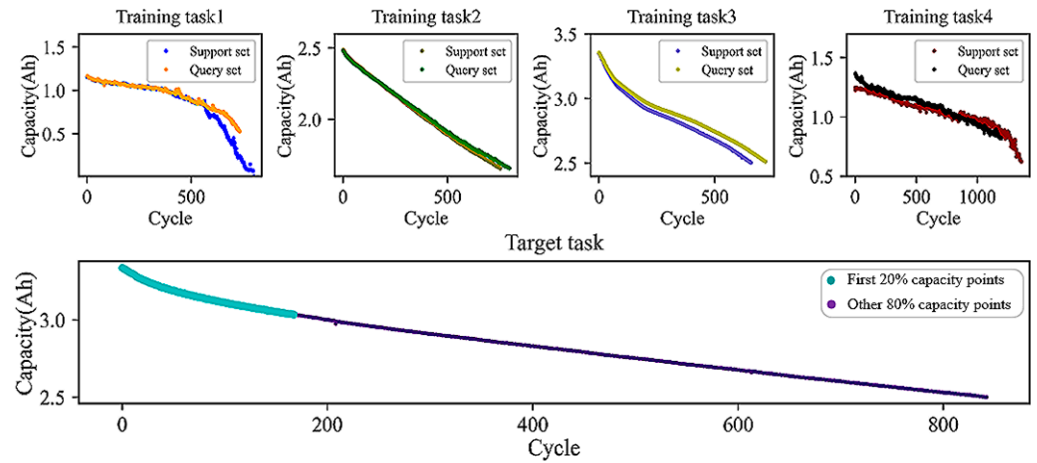


Figure 14. Capacity degradation curves of the corresponding datasets of meta-training and meta-testing based on battery type differences.

In addition, as can be seen from Figure 15, compared with the situation based on different charging conditions, the weight difference between training tasks based on different battery types is not significant, and the weight of task 3 and task 1 is slightly higher than that of other tasks, which is slightly different from the traditional MAML method’s assumption that all task weights are consistent.

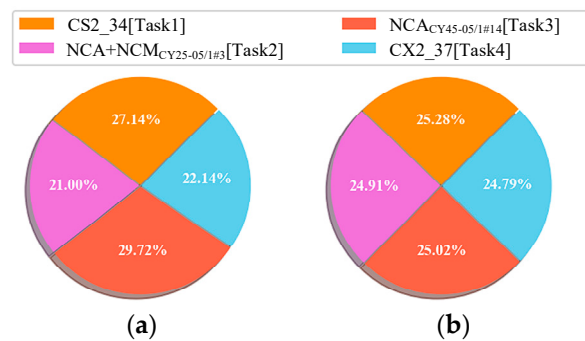


Figure 15. Correlation between meta-training task and meta-testing task dataset based on battery type difference: (a) fixed segment input; (b) random segment input.

5. Experimental Results and Analysis

To eliminate the influence of dimension, a data normalization method is used. To prevent the model from overfitting, a regularization term and dropout layer are introduced. At the same time, RMSE and MAE are used to evaluate the generalization performance of the capacity estimation model.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \tag{24}$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \tag{25}$$

In Equations (24) and (25), \hat{y}_i is the estimated value of capacity, y_i is the true value of capacity, and N is the number of true values.

Due to the uncertainty caused by random segments, unexpected or unwanted fluctuations or changes occur in the data. Therefore, the wavelet threshold denoising method is used to remove the interference noise in the capacity data to extract the real signal in the data. As shown in Equations (26)–(28), the core idea of wavelet denoising is to decompose

the signal into sub-bands of different frequencies by wavelet transform and to reduce noise by threshold processing of these sub-bands.

$$x(t) = \sum_j \sum_k (x, \psi_{j,k}) \cdot \psi_{j,k}(t) \tag{26}$$

$$D_{thresh} = sign(Wx) \cdot (|Wx| - \lambda) \tag{27}$$

$$\hat{x}(t) = \sum_j \sum_k (D_{thresh}, \psi_{j,k}) \cdot \psi_{j,k}(t) \tag{28}$$

where $\psi_{j,k}(t)$ is the wavelet basis function, and in this paper, Daubechies wavelet is selected; j and k represent the scale and translation parameters, respectively; Wx is the wavelet coefficient, and $Wx = \int x(t)\psi_{a,b}(t)dt$; D_{thresh} is the detail coefficient after soft threshold processing.

5.1. Capacity Estimation Results and Discussion Based on Partial Segment Input

From the Tongji University dataset, 60% was selected as the training set for model training, 10% as the validation set for model parameter tuning, and the remaining 30% was used as a test set to test the estimation performance of the model. Each partition of the training set, validation set, and test set is random, and two different input structures designed in Section 3.1 are used for comparison. The battery capacity degradation estimation results based on partial segment input are shown in Figure 16.

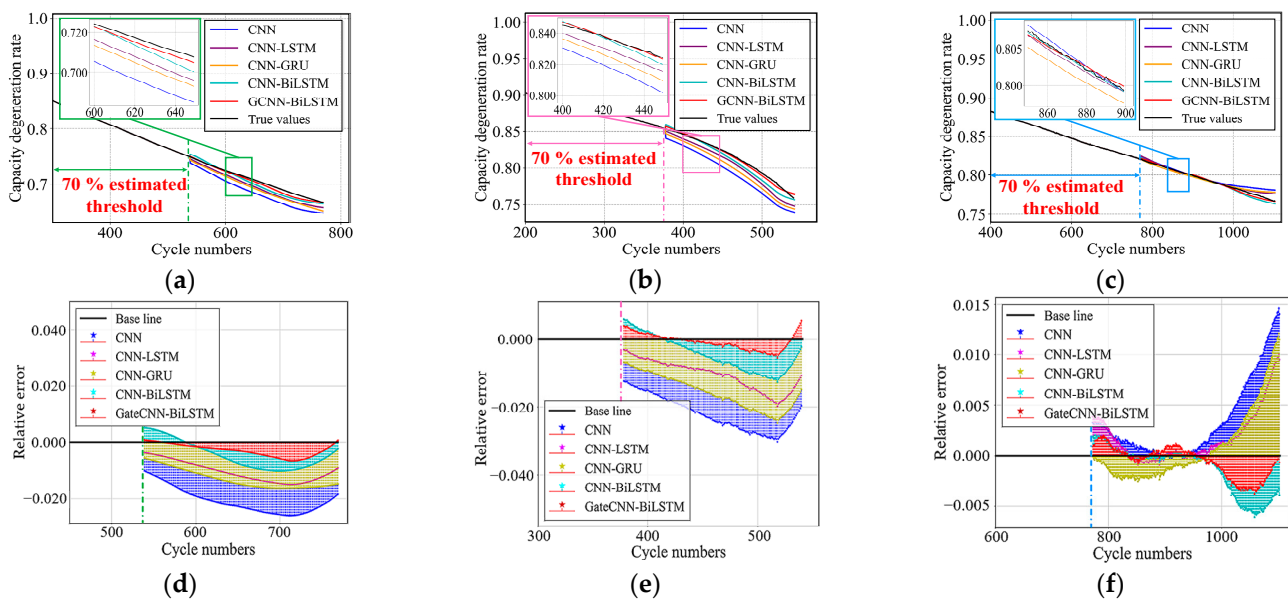


Figure 16. Battery capacity degradation estimation results based on partial segment input (fixed segment input): (a) NCA + NCM battery estimation results; (b) NCA battery estimation results; (c) NCM battery estimation results; (d) NCA + NCM battery estimation error; (e) NCA battery estimation error; (f) NCM battery estimation error.

Compared with the CNN model, the hybrid neural network model (CNN-LSTM, CNN-GRU) can better fit the test data. The CNN-BiLSTM network has achieved better estimation results than the CNN-LSTM/GRU network because it can extract the time series characteristics of the battery from the forward and backward directions. At the same time, GateCNN combines the neural network structure of the convolutional layer and the gated recurrent layer, which can suppress the noise information in the features and improve the generalization ability of the network. Therefore, the basic model proposed in this paper obtains the optimal estimation results. At the same time, to make the model training results more stable, the instantaneous capacity appreciation caused by the experimental operation is eliminated.

Table 6 gives the RMSE and MAE values of the capacity estimation using different methods when the initial fixed segment is used as the input. It can be seen that the RMSE and MAE values of the GateCNN-BiLSTM model in the three batteries are the lowest. The corresponding RMSE values are 0.245%, 0.161%, and 0.391%, respectively. The MAE values are 0.201%, 0.121%, and 0.308%, respectively. This shows that the proposed model has good robustness.

Table 6. RMSE and MAE values of different methods for capacity degradation estimation on Tongji University datasets (Fixed segment input).

Model	NCA		NCM		NCA+NCM	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
CNN	2.273%	2.215%	0.506%	0.332%	2.138%	2.085%
CNN-LSTM	1.136%	1.049%	0.304%	0.197%	1.118%	1.048%
CNN-GRU	1.647%	1.568%	0.365%	0.252%	1.362%	1.311%
CNN-BiLSTM	0.669%	0.569%	0.264%	0.174%	0.690%	0.598%
GateCNN-BiLSTM	0.245%	0.201%	0.161%	0.121%	0.391%	0.308%

When random segment charging data are used as input for Tongji University data, the variation in input features is large, which makes the data distribution more complicated, as shown in Figure 17. Compared with the fixed segment input, the capacity estimation results obtained by random segment input fluctuate slightly, but the relative error is within -0.05% . According to Table 7, the maximum RMSE and MAE in the three batteries are 0.623% and 0.594%, respectively.

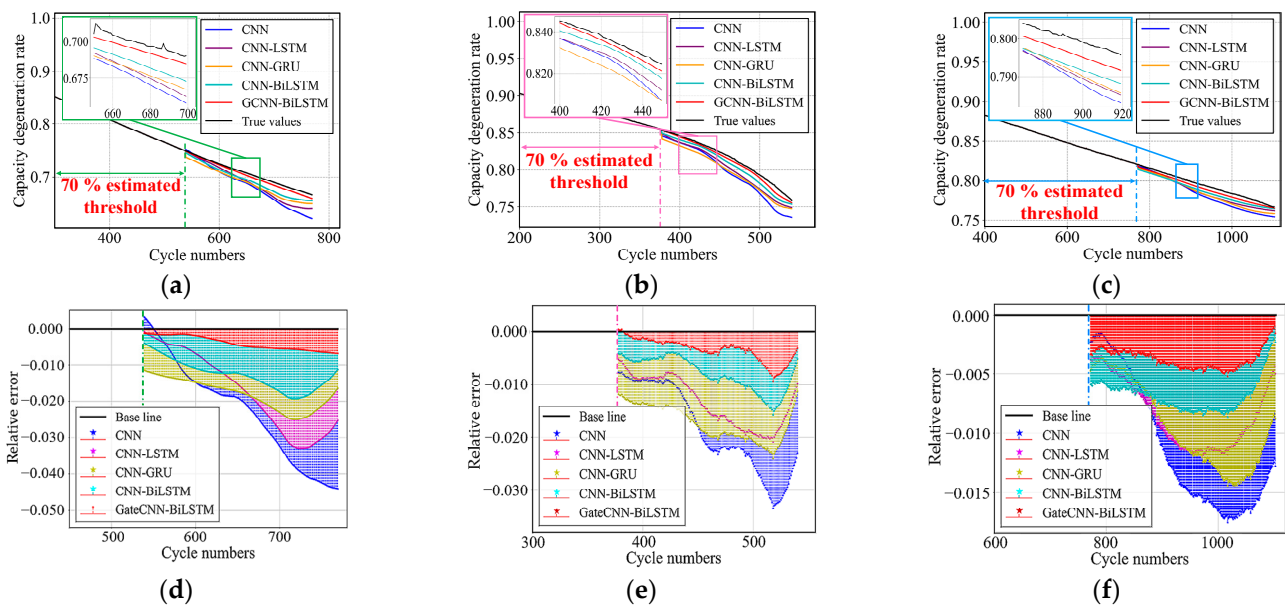


Figure 17. Battery capacity degradation estimation results based on partial segment input (random segment input): (a) NCA + NCM battery estimation results; (b) NCA battery estimation results; (c) NCM battery estimation results; (d) NCA + NCM battery estimation error; (e) NCA battery estimation error; (f) NCM battery estimation error.

The above results show that the output results of the GateCNN-BiLSTM model proposed in this paper show great stability on the datasets of these three different batteries, whether there is a fixed segment or random segment input. It is further proved that the GateCNN-BiLSTM model has strong robustness and generalization performance and can flexibly and accurately estimate the capacity degradation of battery cells.

Table 7. RMSE and MAE values of different methods for capacity degradation estimation on Tongji University datasets (Random segment input).

Model	NCA		NCM		NCA+NCM	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
CNN	2.030%	1.860%	1.245%	1.119%	2.837%	2.488%
CNN-LSTM	1.409%	1.329%	0.893%	0.849%	2.209%	1.920%
CNN-GRU	1.761%	1.729%	1.011%	0.945%	2.045%	2.005%
CNN-BiLSTM	0.857%	0.794%	0.696%	0.682%	1.535%	1.482%
GateCNN-BiLSTM	0.398%	0.339%	0.384%	0.374%	0.623%	0.594%

To validate the proposed model on capacity degradation in other public datasets, the University of Maryland datasets were used, specifically the CS2_35, CS2_36, and CS2_37 datasets. Figure 18 displays the results of the comparison of the proposed model to other contrastive models, under the condition of fixed initial short input segments. The proposed model achieved the best estimation results across all three battery datasets.

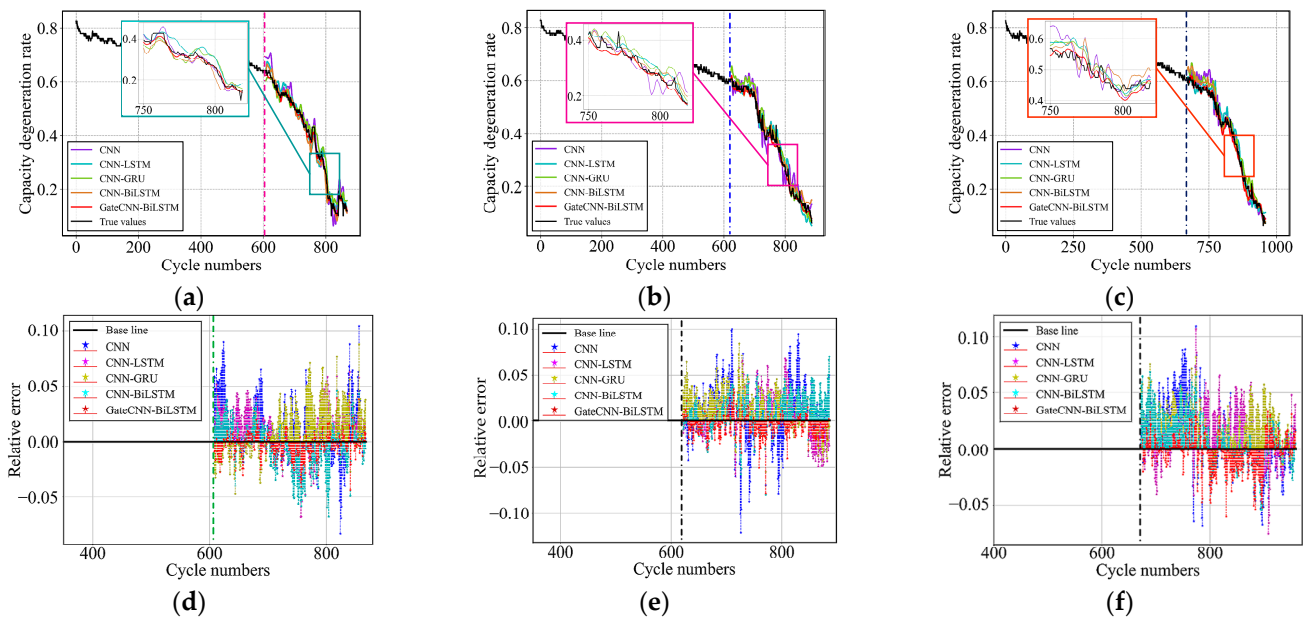


Figure 18. Battery capacity degradation estimation results based on partial segment input (fixed segment input): (a) CS2_35 battery estimation results; (b) CS2_36 battery estimation results; (c) CS2_37 battery estimation results; (d) CS2_35 battery estimation error; (e) CS2_36 battery estimation error; (f) CS2_37 battery estimation error.

Table 8 presents the capacity degradation estimation errors of different data-driven models that rely on fixed segment input obtained from the University of Maryland datasets. The GateCNN-BiLSTM model exhibits the lowest RMSE and MAE for estimated capacity, which are 1.292% and 0.996%, respectively. The average RMSE and MAE are 1.551% and 1.196%. The proposed model outperforms all other models on all datasets, indicating its superior estimation accuracy and robustness.

Figure 19 illustrates the capacity estimation results of different models for University of Maryland datasets when subjected to random fragment input scenarios. By employing gating mechanisms to select feature information from random fragment input, the proposed model exhibits greater stability and closely approximates the actual capacity curve.

Table 8. RMSE and MAE values of different methods for capacity degradation estimation on University of Maryland datasets (Fixed segment input).

Model	CS2_35		CS2_36		CS2_37	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
CNN	3.271%	2.485%	3.746%	2.894%	3.482%	2.700%
CNN-LSTM	2.598%	2.219%	2.678%	2.159%	2.867%	2.302%
CNN-GRU	2.763%	2.182%	2.828%	2.393%	3.070%	2.567%
CNN-BiLSTM	2.337%	1.929%	2.466%	1.978%	2.658%	2.118%
GateCNN-BiLSTM	1.292%	0.996%	1.492%	1.150%	1.869%	1.443%

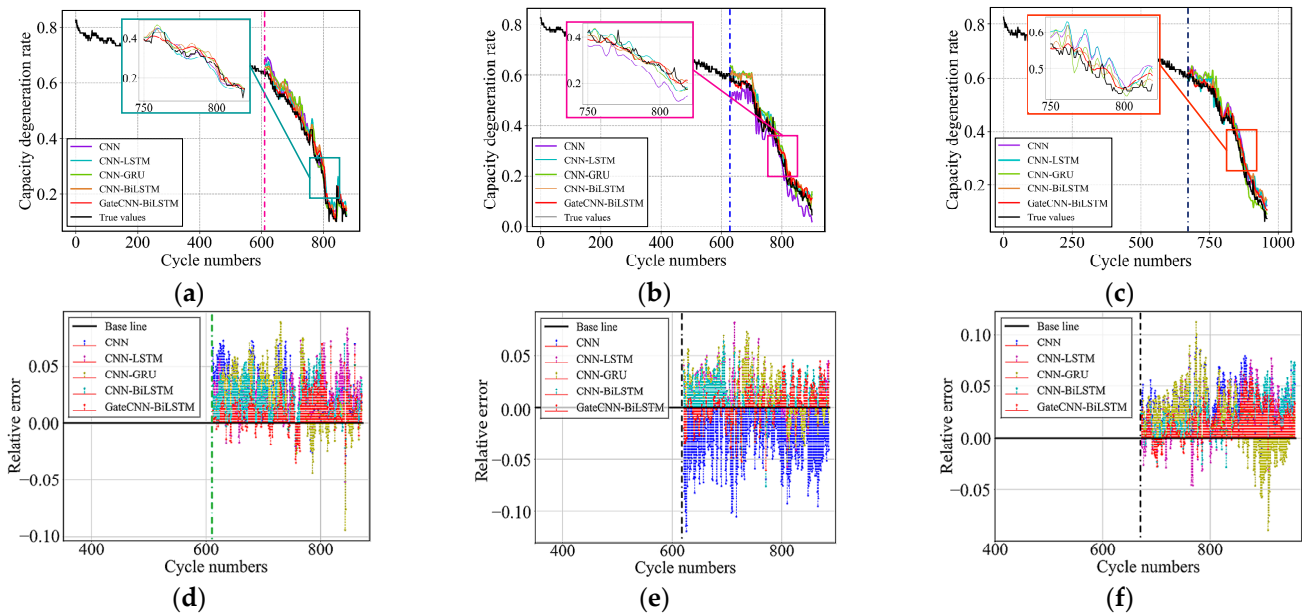


Figure 19. Battery capacity degradation estimation results based on partial segment input (random segment input): (a) CS2_35 battery estimation results; (b) CS2_36 battery estimation results; (c) CS2_37 battery estimation results; (d) CS2_35 battery estimation error; (e) CS2_36 battery estimation error; (f) CS2_37 battery estimation error.

Additionally, Table 9 indicates that the proposed model achieves lower capacity estimation errors when subjected to random fragment input scenarios. Furthermore, it can be observed that the model performs marginally better on the CS2_35 battery dataset, compared to other datasets, under both input structures. This superior performance may be attributed to the better mapping relationship between the health characteristics and capacity of this dataset.

Table 9. RMSE and MAE values of different methods for capacity degradation estimation on University of Maryland datasets (Random segment input).

Model	CS2_35		CS2_36		CS2_37	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
CNN	3.925%	3.402%	4.396%	3.674%	4.312%	3.919%
CNN-LSTM	3.420%	2.976%	3.676%	3.065%	3.424%	2.859%
CNN-GRU	3.691%	2.980%	3.822%	3.257%	3.778%	3.156%
CNN-BiLSTM	3.078%	2.745%	3.241%	2.875%	2.917%	2.316%
GateCNN-BiLSTM	2.275%	1.826%	2.449%	2.088%	2.632%	2.202%

Next, the data of the actual vehicle battery pack are used to verify the applicability of the battery capacity degradation evaluation model in practical applications. The results are shown in Figure 20. The estimated value is very close to the target value. The maximum relative error based on the fixed short segment input is 0.944%, and the RMSE is 0.540%,

respectively. The maximum relative error based on random short segment input is 0.998%, and RMSE is 0.761%, respectively. Therefore, the model proposed in this paper shows excellent accuracy and practical application in solving the problem of battery pack capacity degradation estimation.

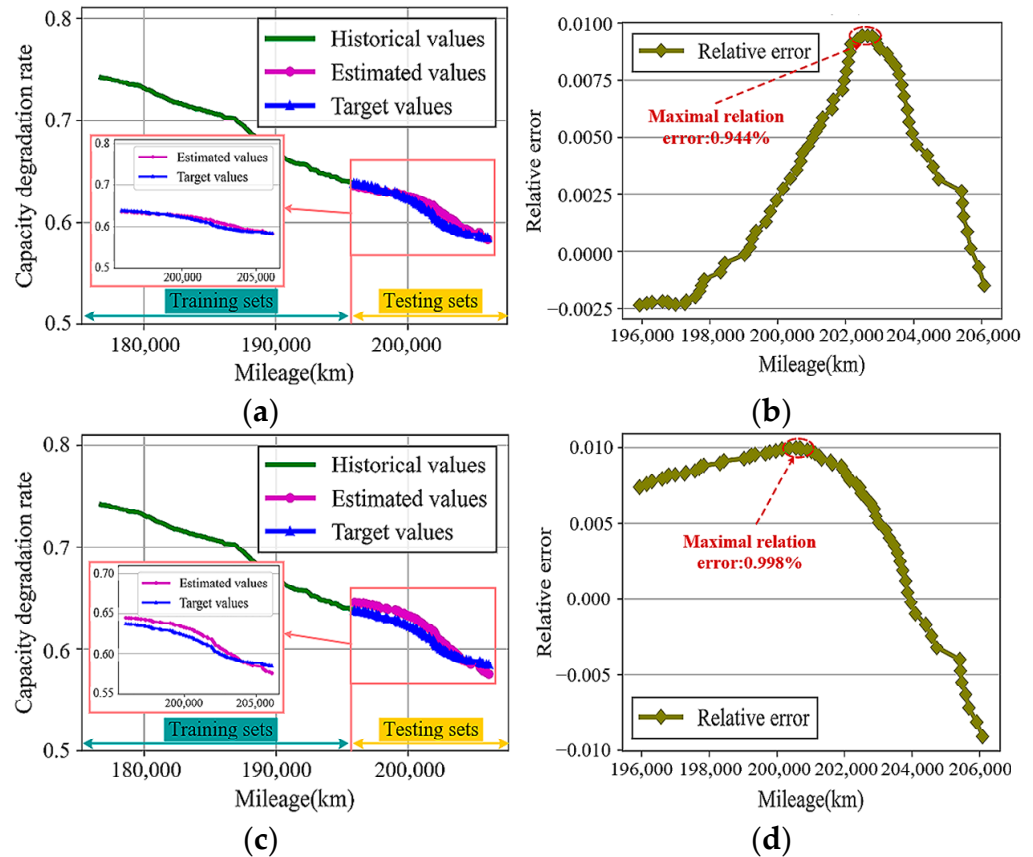


Figure 20. Estimation results of battery pack capacity degradation based on partial short segment input: (a) estimation results based on fixed segment input; (b) estimation error based on fixed segment input; (c) estimation results based on random segment input; (d) estimation error based on random segment input.

5.2. Analysis and Discussion of Capacity Estimation Results Based on Small-Sample Learning

5.2.1. Single-Source Domain Migration Method Based on Domain Adaptation

In this section, the transfer learning model based on domain adaptation is used to estimate the capacity of different small-sample scenarios, and the influence of different sampling ratios is discussed. Figures 21 and 22 show the battery capacity estimation results and errors for Tongji University data in the case of using the first $n\%$ continuous samples, and n is taken as 5, 10, 15, and 20. It can be seen that in the case of only extracting the first $n\%$ of continuous samples, the model may rely too much on this part of the data, resulting in a weak generalization ability of the entire dataset. As the sample size increases, the generalization ability of the model on subsequent data may be affected due to the lack of follow-up information learning, resulting in a slower decline in error.

For the two migration schemes, with the increase in the sample size of the target dataset, there are also some differences in the model error based on the difference in charging conditions and the difference in battery types. For the migration scheme based on the difference in charging conditions, by performing migration learning on the same battery datasets of two different charging conditions, the model may more easily learn common patterns and features, because the datasets between the two charging conditions come from the same battery type. Therefore, as the sample size of the target data collection increases, the

model may converge faster, and the accuracy will increase faster. For the migration schemes of different battery types, although the datasets under the same charging conditions have certain similarities, the model needs to adapt to the differences between different battery types, which may lead to more challenges in the learning and generalization process of the model. Therefore, as the sample size of the target data collection increases, the accuracy of the model may grow more slowly, because the model requires more data to better adapt to the differences between different batteries.

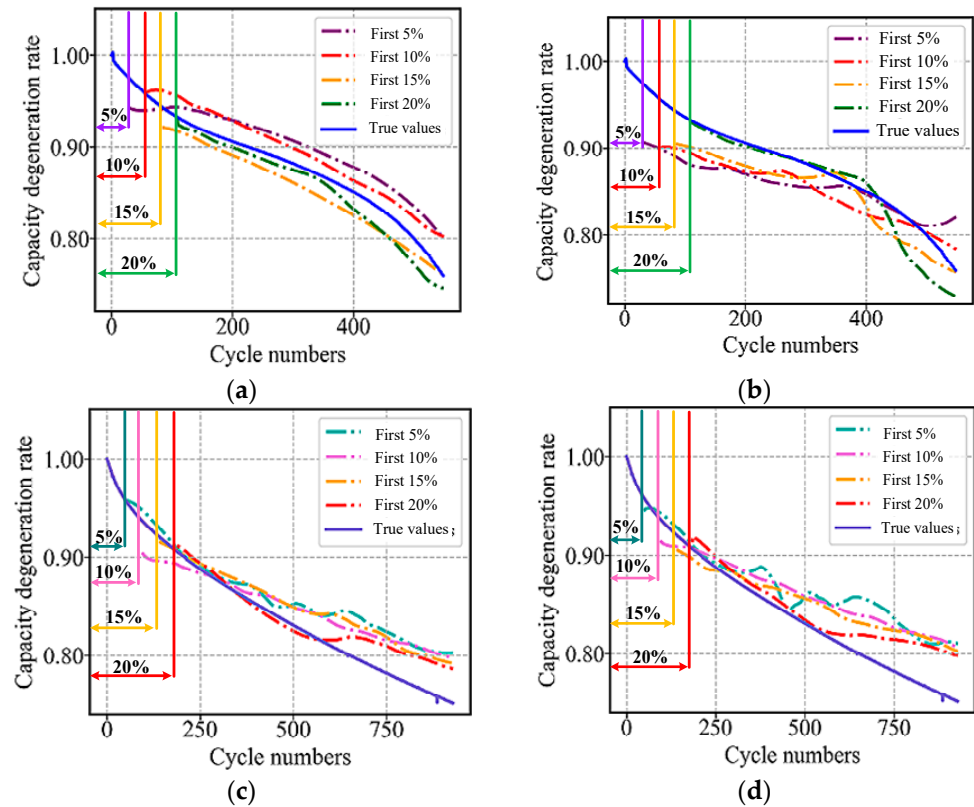


Figure 21. Estimation results of battery capacity under the first $n\%$ continuous sample: (a) condition difference + fixed segment input; (b) condition difference + random segment input; (c) battery type difference + fixed segment input; (d) battery type difference + random segment input.

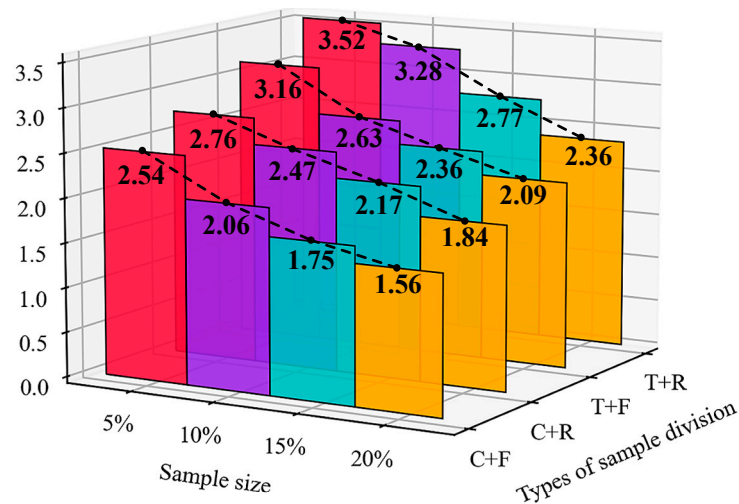


Figure 22. Capacity estimation error (RMSE) under different first ratios of the sample: C(T) + F(R): charging condition (battery type) difference + fixed (random) segment input.

The MMD transfer learning algorithm proposed in this paper belongs to the feature-based transfer learning algorithm. In addition, there are also instance-based and parameter-based transfer learning algorithms. In this section, the MMD-based feature transfer learning algorithm, the instance-based transfer learning algorithm, and the parameter-based transfer learning algorithm are compared on the same GateCNN-BiLSTM model to verify the effectiveness of the MMD feature transfer learning algorithm. For the instance-based transfer learning method, the classic TrAdaBoost algorithm [53] is selected. The main idea of this method is to select samples similar to the target domain from the source domain, re-weight them, and then add the weighted samples to the target domain training set to assist in the training of the target domain model. For the parameter-based transfer learning method, the Finetune algorithm is selected, which is also a classical algorithm [54]. The core idea of this method is as follows: firstly, the neural network is trained by using the source domain data with sufficient samples, then some parameters of the network are frozen, and finally the network is retrained by using the target domain training set with limited samples. In this way, the migration of source domain model parameters to the target domain model is realized.

Figure 21 shows the estimation results of capacity degradation based on the difference in charging conditions when extracting the first 20% of the samples. To better evaluate the performance of each transfer learning algorithm, Table 10 gives the RMSE and MAE values of each transfer learning algorithm for capacity degradation estimation. Among the three transfer learning algorithms, the MMD-based method has better accuracy, and the minimum RMSE and MAE are 1.558% and 1.318%, respectively.

Table 10. Comparison of evaluation indices of different transfer learning methods.

Small-Sample Scenario	Input Features	Estimated Error							
		GateCNN-BiLSTM							
		Base Model		TrAdaBoost		Finetune		MMD	
Charging condition difference	Fix	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
	Random	2.772%	2.549%	2.378%	1.567%	2.550%	2.108%	1.558%	1.318%
Battery type difference	Fix	3.179%	2.265%	3.214%	2.209%	4.226%	2.978%	2.094%	1.328%
	Random	3.137%	2.743%	2.861%	2.378%	3.662%	3.225%	1.841%	1.395%
		3.095%	2.419%	3.171%	2.639%	3.526%	2.866%	2.356%	1.838%

In the case of capacity degradation estimation with battery type differences, as depicted in Figure 23c,d, the knowledge transfer performance in the later stages of battery aging is relatively poor. This can be attributed to significant changes in the degradation mechanism of different batteries during the later stages of their lifespan. Consequently, the samples in the source domain struggle to effectively assist in training during the training process. According to the results presented in Table 10, knowledge transfer based on battery type differences does not yield better estimation performance improvements compared to those based on differences in operating conditions. This suggests that the sample disparities caused by battery type outweigh the differences between charging conditions.

In the case of using the first 20% of the charging cycle as the target domain dataset, the neural network model lacks complete knowledge expression ability due to the absence of late aging information about the battery in the training data. This indicates that a data-driven model without the inclusion of a transfer learning algorithm cannot construct an accurate capacity degradation estimator solely based on a small amount of early cycle data from the battery charging condition.

It is important to highlight that the TrAdaBoost and Finetune methods did not yield satisfactory results. This can be attributed to significant distribution differences between the source domain data and the target domain data. TrAdaBoost attempts to address this issue by re-weighting certain source domain data and incorporating them into model training alongside the target domain data. On the other hand, the Finetune algorithm fine-tunes a model trained on source domain data using a small amount of target domain training data to better adapt to the target domain's characteristics. Both of these transfer learning algorithms

heavily rely on the similarity between the source and target domain data. However, in the capacity estimation study conducted here, there is a transfer between different datasets, resulting in substantial distribution differences between the source and target domain data. These significant distribution disparities greatly impact the performance of these two transfer learning algorithms, making it challenging to achieve the desired outcomes.

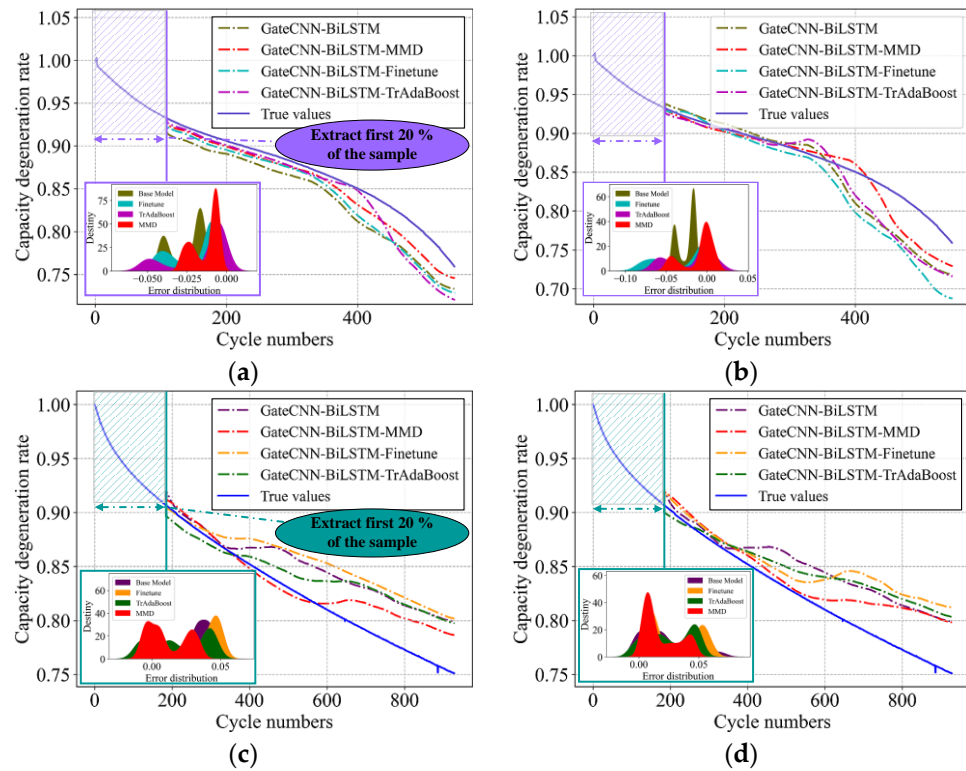


Figure 23. Capacity degradation estimation results based on domain adaptation transfer learning under different knowledge domain differences: (a) condition difference + fixed segment input; (b) condition difference + random segment input; (c) battery type difference + fixed segment input; (d) battery type difference + random segment input.

5.2.2. Multi-Source Domain Migration Method Based on Meta-Learning

To verify the effectiveness of the improved MMD-MAML framework based on weight guidance proposed in this paper, the Reptile [55] and MAML methods are selected for comparison. The learning mechanism of Reptile is similar to that of MAML, as it focuses on improving the direction of gradient descent for model parameters. However, as mentioned earlier, the single-source domain transfer learning method tends to exhibit deviations in the later stages of the cycle due to differences in charging conditions and battery types.

From Figure 24, it is evident that the capacity degradation estimation results achieved through multi-source domain transfer learning are significantly superior to those of single-source domain transfer. This improved performance effectively compensates for the errors in the later stages of the cycle. The reason behind this improvement is that MAML leverages its ability to comprehensively learn the knowledge expression of lithium-ion battery capacity degradation across multiple tasks. By dividing various charging conditions or battery types into source tasks and conducting meta-training, MAML can acquire the domain knowledge necessary for the target task. Additionally, the initial model parameters are tested using new small-scale data. Based on this, we introduce MMD to enhance MAML, considering the relevance of tasks, with the aim of further improving the accuracy of the model.

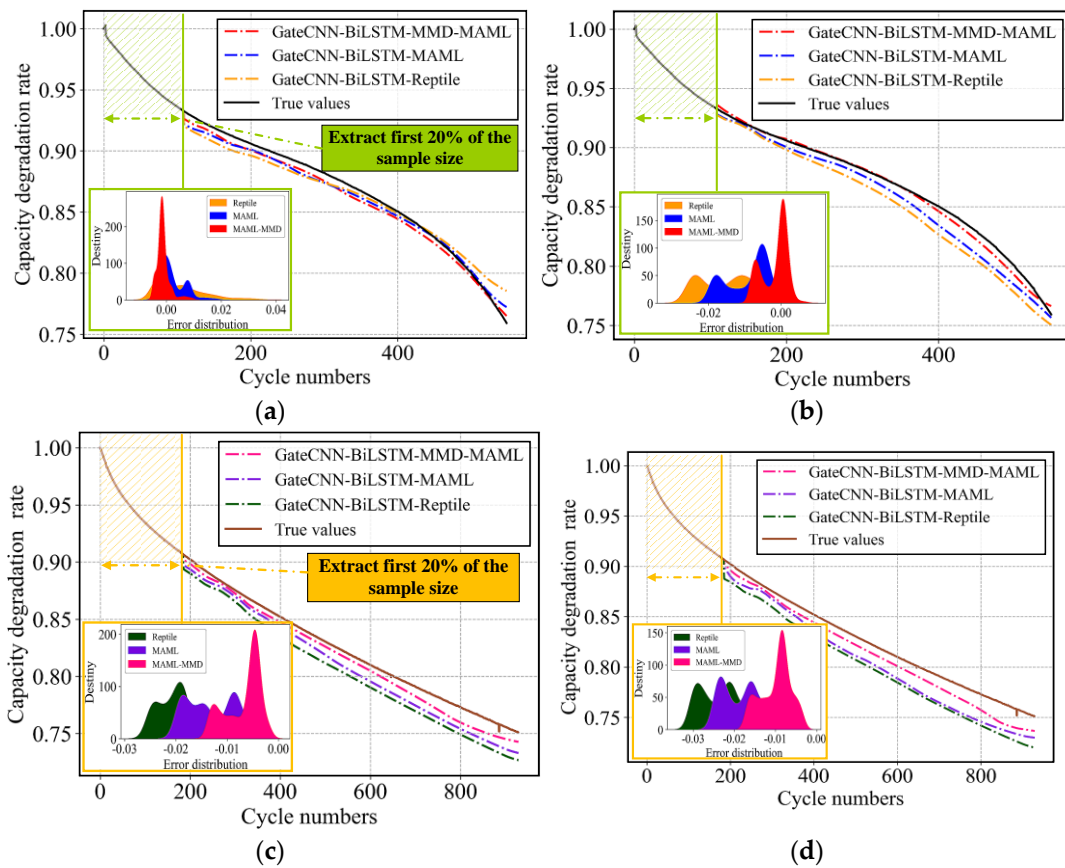


Figure 24. Capacity degradation estimation results of multi-source domain transfer learning under different knowledge domain differences: (a) condition difference + fixed segment input; (b) condition difference + random segment input; (c) battery type difference + fixed segment input; (d) battery type difference + random segment input.

Table 11 presents the capacity estimation results for different charging conditions and battery types using the first 20% sample condition. In a comparison with the MAML and Reptile frameworks, it can be concluded that the capacity estimation results obtained by the MMD-MAML framework under different charging conditions closely align with the actual curve. Additionally, the RMSE and MAE values for MMD-MAML reach the minimum values, which are 0.254% and 0.157%, respectively. Compared to the single-source domain migration scheme, the estimation curve of the MMD-MAML framework better approximates the actual value in the later stages of battery capacity degradation. Despite the lack of post-cycle data in the input sample of the target task, the model can still learn the post-capacity-degradation pattern of the target battery through training on the source task, resulting in improved capacity estimation performance.

Table 11. Comparison of different multi-source domain transfer learning methods.

Small-Sample Scenario	Input Features	Estimated Error					
		GateCNN-BiLSTM-Reptile		GateCNN-BiLSTM-MAML		GateCNN-BiLSTM-MMD-MAML	
		RMSE	MAE	RMSE	MAE	RMSE	MAE
Charging condition difference	Fix	1.223%	0.875%	0.678%	0.515%	0.254%	0.157%
	Random	1.635%	1.466%	1.086%	0.928%	0.386%	0.275%
Battery type difference	Fix	1.928%	1.871%	1.381%	1.303%	0.758%	0.678%
	Random	2.133%	2.066%	1.579%	1.522%	0.987%	0.941%

Regarding battery type differences, when considering the input features, the utilization of fixed segments as input yields the lowest RMSE and MAE for the estimated capacity, which are 0.758% and 0.678%, respectively. The variation in degradation mechanisms

among different cathode materials is more significant compared to that of charging conditions. Consequently, the RMSE and MAE obtained using corresponding fixed short segment input increased by 198.42% and 221.85%, respectively. Similarly, the RMSE and MAE obtained using corresponding random short segment input increased by 155.70% and 242.18%, respectively. The meta-learning-based knowledge transfer method serves as an active training mechanism. Hence, even if the meta-testing task lacks sample data for a specific interval, the meta-learning method can actively learn the capacity degradation patterns across various batteries and aging intervals during the learning process of multiple source tasks. Although the capacity degradation estimator's estimation error may increase in the later stages of the cycle, it is significantly reduced compared to the single-source domain method. Overall, the estimator provides a better fit for the battery's capacity degradation.

6. Conclusions

This paper presents a novel learning-based approach for estimating the capacity degradation of lithium-ion batteries (LIBs). We introduce a hybrid neural network, GateCNN-BiLSTM, which combines gated convolutional and bi-directional long short-term memory, to establish a data-driven basic model. To address scenarios with limited previous charge and discharge cycle data, we develop a learning framework that combines single-source domain and multi-source domain transfer learning. This framework is applied to estimate the capacity degradation of LIBs and is validated using various aging datasets of different LIB types. The research findings are summarized as follows:

The GateCNN-BiLSTM basic model demonstrates the ability to capture advanced spatial features and forward–reverse time features of battery sequence data, enabling comprehensive feature extraction. By utilizing voltage and current data recorded during a short charging period, the model accurately estimates the current battery capacity state. Compared to traditional data-driven models, it achieves higher accuracy in battery capacity estimation across multiple battery datasets, even with random fragment inputs, the average root mean square error (RMSE) and mean absolute error (MAE) are 0.468% and 0.436% respectively. For vehicle battery pack capacity degradation estimation, the maximum relative error and RMSE for battery charging capacity estimation based on fixed short segment inputs are 0.944% and 0.540%, respectively. For charging capacity estimation based on random short segment inputs, the maximum relative error and RMSE are 0.998% and 0.761%, respectively. These results highlight the model's wide applicability in capacity estimation for different battery types, without relying on a fixed input structure. Furthermore, considering the uncertainties in charging and discharging behavior in EVs, the model accurately estimates battery capacity degradation using short charging segment data.

In addressing the challenge of limited data in practical applications, we compare our transfer learning method with other approaches and find that the single-source domain feature transfer learning method based on maximum mean difference (MMD) effectively transfers battery health state information from a source domain with sufficient data to a target domain. This solves the problem of limited generalization in small-sample scenarios. By dividing the knowledge domain into charging condition difference and battery type difference, we analyze the variation in battery capacity estimation results based on different proportions of continuous charging samples. As the sample size increases, the capacity estimation error decreases, albeit at a slower rate. For the first 20% of the samples, the lowest RMSE and MAE are 1.558% and 1.318%, respectively. This indicates that our method significantly reduces the dependence of data-driven models on historical data quantity, improving the model's performance even with limited data. Multi-task learning is implemented using the multi-source domain meta-learning MAML algorithm. By incorporating MMD to measure the correlation between the source task and the target task, the efficiency of meta-training is enhanced, compensating for any deviation in single-source domain knowledge transfer during the later stages of capacity degradation. This leads to more accurate and efficient capacity degradation estimation. In the future, with further

optimization, we anticipate that the proposed battery capacity estimation method can be applied in cloud-based or vehicle-mounted battery management systems (BMSs).

Author Contributions: Conceptualization, Y.S.; supervision, J.D.; data curation, F.H.; writing—original draft, H.T.; writing—review and editing, Y.S., H.T. and J.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

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

References

1. Kim, S.; Jung, H.; Lee, M.; Choi, Y.Y.; Choi, J.-I. Model-Free Reconstruction of Capacity Degradation Trajectory of Lithium-Ion Batteries Using Early Cycle Data. *eTransportation* **2023**, *17*, 100243. [[CrossRef](#)]
2. Hu, X.; Feng, F.; Liu, K.; Zhang, L.; Xie, J.; Liu, B. State Estimation for Advanced Battery Management: Key Challenges and Future Trends. *Renew. Sustain. Energy Rev.* **2019**, *114*, 109334. [[CrossRef](#)]
3. Xu, L.; Lin, X.; Xie, Y.; Hu, X. Enabling High-Fidelity Electrochemical P2D Modeling of Lithium-Ion Batteries via Fast and Non-Destructive Parameter Identification. *Energy Storage Mater.* **2022**, *45*, 952–968. [[CrossRef](#)]
4. Li, J.; Adewuyi, K.; Lotfi, N.; Landers, R.G.; Park, J. A Single Particle Model with Chemical/Mechanical Degradation Physics for Lithium Ion Battery State of Health (SOH) Estimation. *Appl. Energy* **2018**, *212*, 1178–1190. [[CrossRef](#)]
5. Yao, L.; Xu, S.; Tang, A.; Zhou, F.; Hou, J.; Xiao, Y.; Fu, Z. A Review of Lithium-Ion Battery State of Health Estimation and Prediction Methods. *World Electr. Veh. J.* **2021**, *12*, 113. [[CrossRef](#)]
6. Tian, H.; Qin, P.; Li, K.; Zhao, Z. A Review of the State of Health for Lithium-Ion Batteries: Research Status and Suggestions. *J. Clean. Prod.* **2020**, *261*, 120813. [[CrossRef](#)]
7. Wang, Y.; Tian, J.; Sun, Z.; Wang, L.; Xu, R.; Li, M.; Chen, Z. A Comprehensive Review of Battery Modeling and State Estimation Approaches for Advanced Battery Management Systems. *Renew. Sustain. Energy Rev.* **2020**, *131*, 110015. [[CrossRef](#)]
8. Yang, S.; Zhang, C.; Jiang, J.; Zhang, W.; Zhang, L.; Wang, Y. Review on State-of-Health of Lithium-Ion Batteries: Characterizations, Estimations and Applications. *J. Clean. Prod.* **2021**, *314*, 128015. [[CrossRef](#)]
9. Zhao, J.; Ling, H.; Liu, J.; Wang, J.; Burke, A.F.; Lian, Y. Machine Learning for Predicting Battery Capacity for Electric Vehicles. *eTransportation* **2023**, *15*, 100214. [[CrossRef](#)]
10. Li, H.; Zhang, Z.; Li, T.; Si, X. A Review on Physics-Informed Data-Driven Remaining Useful Life Prediction: Challenges and Opportunities. *Mech. Syst. Signal Process.* **2024**, *209*, 111120. [[CrossRef](#)]
11. Navidi, S.; Thelen, A.; Li, T.; Hu, C. Physics-Informed Machine Learning for Battery Degradation Diagnostics: A Comparison of State-of-the-Art Methods. *Energy Storage Mater.* **2024**, *68*, 103343. [[CrossRef](#)]
12. Weddle, P.J.; Kim, S.; Chen, B.-R.; Yi, Z.; Gasper, P.; Colclasure, A.M.; Smith, K.; Gering, K.L.; Tanim, T.R.; Dufek, E.J. Battery State-of-Health Diagnostics during Fast Cycling Using Physics-Informed Deep-Learning. *J. Power Sources* **2023**, *585*, 233582. [[CrossRef](#)]
13. Wen, P.; Ye, Z.-S.; Li, Y.; Chen, S.; Xie, P.; Zhao, S. Physics-Informed Neural Networks for Prognostics and Health Management of Lithium-Ion Batteries. *IEEE Trans. Intell. Veh.* **2024**, *9*, 2276–2289. [[CrossRef](#)]
14. Hofmann, T.; Hamar, J.; Rogge, M.; Zoerr, C.; Erhard, S.; Schmidt, J.P. Physics-Informed Neural Networks for State of Health Estimation in Lithium-Ion Batteries. *J. Electrochem. Soc.* **2023**, *170*, 090524. [[CrossRef](#)]
15. Chen, J.; Marlow, M.N.; Jiang, Q.; Wu, B. Peak-Tracking Method to Quantify Degradation Modes in Lithium-Ion Batteries via Differential Voltage and Incremental Capacity. *J. Energy Storage* **2022**, *45*, 103669. [[CrossRef](#)]
16. Jones, C.; Sudarshan, M.; Serov, A.; Tomar, V. Investigation of Physical Effects on Prismatic Lithium-Ion Cell Electrodes after Partial Nail Puncture Using Raman Spectroscopy and Incremental Capacity Analysis. *eTransportation* **2022**, *12*, 100174. [[CrossRef](#)]
17. Zhang, S.; Zhai, B.; Guo, X.; Wang, K.; Peng, N.; Zhang, X. Synchronous Estimation of State of Health and Remaining Useful Lifetime for Lithium-Ion Battery Using the Incremental Capacity and Artificial Neural Networks. *J. Energy Storage* **2019**, *26*, 100951. [[CrossRef](#)]
18. Li, X.; Yuan, C.; Wang, Z. State of Health Estimation for Li-Ion Battery via Partial Incremental Capacity Analysis Based on Support Vector Regression. *Energy* **2020**, *203*, 117852. [[CrossRef](#)]
19. Pan, W.; Luo, X.; Zhu, M.; Ye, J.; Gong, L.; Qu, H. A Health Indicator Extraction and Optimization for Capacity Estimation of Li-Ion Battery Using Incremental Capacity Curves. *J. Energy Storage* **2021**, *42*, 103072. [[CrossRef](#)]
20. Li, Y.; Zhong, S.; Zhong, Q.; Shi, K. Lithium-Ion Battery State of Health Monitoring Based on Ensemble Learning. *IEEE Access* **2019**, *7*, 8754–8762. [[CrossRef](#)]
21. Wang, J.; Deng, Z.; Yu, T.; Yoshida, A.; Xu, L.; Guan, G.; Abudula, A. State of Health Estimation Based on Modified Gaussian Process Regression for Lithium-Ion Batteries. *J. Energy Storage* **2022**, *51*, 104512. [[CrossRef](#)]

22. Zheng, Y.; Hu, J.; Chen, J.; Deng, H.; Hu, W. State of Health Estimation for Lithium Battery Random Charging Process Based on CNN-GRU Method. *Energy Rep.* **2023**, *9*, 1–10. [[CrossRef](#)]
23. Yang, N.; Song, Z.; Hofmann, H.; Sun, J. Robust State of Health Estimation of Lithium-Ion Batteries Using Convolutional Neural Network and Random Forest. *J. Energy Storage* **2022**, *48*, 103857. [[CrossRef](#)]
24. Sun, S.; Sun, J.; Wang, Z.; Zhou, Z.; Cai, W. Prediction of Battery SOH by CNN-BiLSTM Network Fused with Attention Mechanism. *Energies* **2022**, *15*, 4428. [[CrossRef](#)]
25. Shen, S.; Sadoughi, M.; Li, M.; Wang, Z.; Hu, C. Deep Convolutional Neural Networks with Ensemble Learning and Transfer Learning for Capacity Estimation of Lithium-Ion Batteries. *Appl. Energy* **2020**, *260*, 114296. [[CrossRef](#)]
26. Zhang, S.; Liu, Z.; Su, H. State of Health Estimation for Lithium-Ion Batteries on Few-Shot Learning. *Energy* **2023**, *268*, 126726. [[CrossRef](#)]
27. Tian, J.; Xiong, R.; Shen, W.; Lu, J.; Yang, X.-G. Deep Neural Network Battery Charging Curve Prediction Using 30 Points Collected in 10 Min. *Joule* **2021**, *5*, 1521–1534. [[CrossRef](#)]
28. Zhang, X.; Fan, J.; Zou, Y.; Sun, W. Realizing Accurate Battery Capacity Estimation Using 4 Min 1C Discharging Data. *Energy* **2023**, *282*, 128744. [[CrossRef](#)]
29. Mei, P.; Karimi, H.R.; Chen, F.; Yang, S.; Huang, C.; Qiu, S. A Learning-Based Vehicle-Cloud Collaboration Approach for Joint Estimation of State-of-Energy and State-of-Health. *Sensors* **2022**, *22*, 9474. [[CrossRef](#)]
30. Tian, Y.; Wen, J.; Yang, Y.; Shi, Y.; Zeng, J. State-of-Health Prediction of Lithium-Ion Batteries Based on CNN-BiLSTM-AM. *Batteries* **2022**, *8*, 155. [[CrossRef](#)]
31. Zhu, J.; Wang, Y.; Huang, Y.; Gopaluni, R.B.; Cao, Y.; Heere, M.; Mühlbauer, M.J.; Mereacre, L.; Dai, H.; Liu, X.; et al. Data-driven capacity estimation of commercial lithium-ion batteries from voltage relaxation. *Nat. Commun.* **2022**, *13*, 2261. [[CrossRef](#)] [[PubMed](#)]
32. 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](#)]
33. Tian, H.; Sun, Y.; Hu, F.; Du, J. Charging Behavior Analysis Based on Operation Data of Private BEV Customers in Beijing. *Electronics* **2023**, *12*, 373. [[CrossRef](#)]
34. Zhao, X.; Hu, J.; Hu, G.; Qiu, H. A State of Health Estimation Framework Based on Real-World Electric Vehicles Operating Data. *J. Energy Storage* **2023**, *63*, 107031. [[CrossRef](#)]
35. Song, L.; Zhang, K.; Liang, T.; Han, X.; Zhang, Y. Intelligent State of Health Estimation for Lithium-Ion Battery Pack Based on Big Data Analysis. *J. Energy Storage* **2020**, *32*, 101836. [[CrossRef](#)]
36. Zhu, S.; He, C.; Zhao, N.; Sha, J. Data-Driven Analysis on Thermal Effects and Temperature Changes of Lithium-Ion Battery. *J. Power Sources* **2021**, *482*, 228983. [[CrossRef](#)]
37. Hong, J.; Wang, Z.; Chen, W.; Wang, L.; Lin, P.; Qu, C. Online Accurate State of Health Estimation for Battery Systems on Real-World Electric Vehicles with Variable Driving Conditions Considered. *J. Clean. Prod.* **2021**, *294*, 125814. [[CrossRef](#)]
38. Driscoll, L.; de la Torre, S.; Gomez-Ruiz, J.A. Feature-Based Lithium-Ion Battery State of Health Estimation with Artificial Neural Networks. *J. Energy Storage* **2022**, *50*, 104584. [[CrossRef](#)]
39. Gou, B.; Xu, Y.; Feng, X. State-of-Health Estimation and Remaining-Useful-Life Prediction for Lithium-Ion Battery Using a Hybrid Data-Driven Method. *IEEE Trans. Veh. Technol.* **2020**, *69*, 10854–10867. [[CrossRef](#)]
40. Liu, H.; Deng, Z.; Yang, Y.; Lu, C.; Li, B.; Liu, C.; Cheng, D. Capacity Evaluation and Degradation Analysis of Lithium-Ion Battery Packs for on-Road Electric Vehicles. *J. Energy Storage* **2023**, *65*, 107270. [[CrossRef](#)]
41. Ungurean, L.; Micea, M.V.; Cârstoiu, G. Online State of Health Prediction Method for Lithium-Ion Batteries, Based on Gated Recurrent Unit Neural Networks. *Int. J. Energy Res.* **2020**, *44*, 6767–6777. [[CrossRef](#)]
42. Zhu, W.; Guo, B.; Li, Y.; Yang, Y.; Xie, C.; Jin, J.; Gooi, H.B. Uncertainty Quantification of Proton-Exchange-Membrane Fuel Cells Degradation Prediction Based on Bayesian-Gated Recurrent Unit. *eTransportation* **2023**, *16*, 100230. [[CrossRef](#)]
43. Du, Y.; Cui, N.; Cui, W.; Chen, Z.; Zhang, C. Receding Horizon Control Based Energy Management Strategy for PHEB Using GRU Deep Learning Predictive Model. *eTransportation* **2022**, *13*, 100179. [[CrossRef](#)]
44. Shen, L.; Li, J.; Liu, J.; Zhu, L.; Shen, H.T. Temperature Adaptive Transfer Network for Cross-Domain State-of-Charge Estimation of Li-Ion Batteries. *IEEE Trans. Power Electron.* **2023**, *38*, 3857–3869. [[CrossRef](#)]
45. Li, Z.; Li, A.; Bai, F.; Zuo, H.; Zhang, Y. Remaining Useful Life Prediction of Lithium Battery Based on ACNN-Mogriifier LSTM-MMD. *Meas. Sci. Technol.* **2023**, *35*, 016101. [[CrossRef](#)]
46. Ma, G.; Xu, S.; Yang, T.; Du, Z.; Zhu, L.; Ding, H.; Yuan, Y. A Transfer Learning-Based Method for Personalized State of Health Estimation of Lithium-Ion Batteries. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *35*, 759–769. [[CrossRef](#)]
47. Zhu, Y.; Li, X.; Zhang, Y.; Zhang, W. Cross-Domain Prognostic Method of Lithium-Ion Battery in New Energy Electric Aircraft With Domain Adaptation. *IEEE Sens. J.* **2023**, *23*, 14487–14498. [[CrossRef](#)]
48. Bhattacharjee, A.; Verma, A.; Mishra, S.; Saha, T.K. Estimating State of Charge for xEV Batteries Using 1D Convolutional Neural Networks and Transfer Learning. *IEEE Trans. Veh. Technol.* **2021**, *70*, 3123–3135. [[CrossRef](#)]
49. Zhang, Y.; Liu, K.; Chuang, Y.; Zhang, J. Estimation of State of Charge Integrating Spatial and Temporal Characteristics with Transfer Learning Optimization. *Meas. Sci. Technol.* **2023**, *34*, 045112. [[CrossRef](#)]
50. Schmitt, J.; Horstkötter, I.; Bäker, B. Effective Estimation of Battery State-of-Health by Virtual Experiments via Transfer- and Meta-Learning. *J. Energy Storage* **2023**, *63*, 106969. [[CrossRef](#)]

51. Jeong, D.; Bae, S. Estimating Battery State-of-Charge with a Few Target Training Data by Meta-Learning. *J. Power Sources* **2023**, *553*, 232238. [[CrossRef](#)]
52. Satrya, W.F.; Yun, J.-H. Combining Model-Agnostic Meta-Learning and Transfer Learning for Regression. *Sensors* **2023**, *23*, 583. [[CrossRef](#)]
53. Qin, H.; Fan, X.; Fan, Y.; Wang, R.; Shang, Q.; Zhang, D. A Transferable Prediction Approach for the Remaining Useful Life of Lithium-Ion Batteries Based on Small Samples. *Appl. Sci.* **2023**, *13*, 8498. [[CrossRef](#)]
54. Chen, Z.; Chen, L.; Shen, W.; Xu, K. Remaining Useful Life Prediction of Lithium-Ion Battery via a Sequence Decomposition and Deep Learning Integrated Approach. *IEEE Trans. Veh. Technol.* **2022**, *71*, 1466–1479. [[CrossRef](#)]
55. Ortiz, J.P.; Ayabaca, G.P.; Cardenas, A.R.; Cabrera, D.; Valladolid, J.D. Continual Reinforcement Learning Using Real-World Data for Intelligent Prediction of SOC Consumption in Electric Vehicles. *IEEE Lat. Am. Trans.* **2022**, *20*, 624–633. [[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.