MDPI

Article

Health State Assessment of Lithium-Ion Batteries Based on Multi-Health Feature Fusion and Improved Informer Modeling

Jun He, Xinyu Liu *, Wentao Huang, Bohan Zhang, Zuoming Zhang, Zirui Shao and Zimu Mao

School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan 430070, China; apm874@163.com (J.H.); 20181094@hbut.edu.cn (W.H.); 102210310@hbut.edu.cn (B.Z.); zuomingzhang361@163.com (Z.Z.); 102210320@hbut.edu.cn (Z.S.); m1254722103@163.com (Z.M.)
* Correspondence: 102210357@hbut.edu.cn

Abstract: Accurately assessing the state of health (SOH) of lithium batteries is of great significance for improving battery safety performance. However, the current assessment for SOH suffers from the difficulty of selecting health features and the lack of uncertainty using data-driven methods. To this end, this paper proposes a health state assessment method for lithium-ion batteries based on health feature extraction and an improved Informer model. First, multiple features that can reflect the SOH of lithium-ion batteries were extracted from the charging and discharging time, the peak value of incremental capacity curve (ICC), and the inflection point value of differential voltage curve, etc., and the correlation between multiple health features and the health state was evaluated by gray correlation analysis. Then, an improved Informer model is proposed to establish a health state estimation method for lithium-ion batteries. Finally, the proposed algorithm is tested and validated using publicly available battery charge/discharge datasets and compared with other algorithms. The results show that the method in this paper can realize high-precision SOH prediction with a root-mean-square error (RMSE) of 0.011, and the model fit reaches more than 98%.

Keywords: lithium-ion batteries; health evaluation; feature extraction; deep learning

1. Introduction

With the increasing scale of new energy on-grid, the importance of energy storage in the new power system generation, transmission, distribution, and use of the link is becoming more and more prominent. Among them, lithium-ion batteries have become one of the mainstream energy storage systems in new power systems due to their high energy density, fast conversion rate, and easy deployment [1–3]. However, with the charging and discharging cycles of lithium-ion batteries, it will lead to the aging of lithium batteries, causing fire, explosion, and other accidents. The health state of lithium-ion batteries can effectively reflect the degree of battery aging, and in actual operation, replacing low SOH batteries based on a set threshold is one of the effective measures to ensure the safe and stable operation of battery energy storage systems.

Currently, Li-ion battery SOH assessment methods are divided into two main categories: model-driven-based methods and data-driven-based methods. Among them, the model-driven approach is mainly based on electrochemical models and equivalent circuit models. Electrochemical modeling usually employs partial differential equations to describe the electrochemical processes inside the battery, such as the growth mechanism of the solid electrolyte interfacial layer film, the loss process of lithium ions, and positive and negative active materials, etc., which leads to the derivation of the battery SOH [4,5].

Equivalent circuit modeling, on the other hand, abstracts the battery into components such as resistance and capacitance and reassembles them into circuits, thereby obtaining internal and external characteristics consistent with those of the battery. Equivalent circuit model parameter optimization is the key to improve the efficiency of battery SOH evaluation. Ref. [5] used a first-order equivalent circuit model to simulate the battery

and estimated the key parameters of the model by injecting different frequency currents sequentially using the frequency scale separation method. Ref. [6] used a stochastic variational ant colony optimization algorithm to identify the key parameters of the PNGV (the partnership for a new generation of vehicles) battery model. In addition, recursive least squares [7], Kalman filtering [8], and other methods are widely used for battery model parameter identification. In general, the model-based battery SOH assessment method is more accurate, but the calculation process consumes a lot of arithmetic power, which makes it more difficult to be deployed in the BMS.

In order to solve the difficulties of algorithm deployment in practical application scenarios, data-driven battery SOH evaluation methods based on data have attracted extensive attention from researchers. Compared with the model-driven method, the data-driven method does not need to study the internal operation mechanism of the battery in depth, and only learns the change rule of input data through loop iteration, and then can give accurate evaluation results, which is more flexible [9]. Data-driven battery SOH-based assessments are usually performed based on timing monitoring data collected by the BMS, and thus the core of such approaches is the analysis and processing of timing data. Ref. [10] used a modified gated recurrent unit convolutional neural network to exploit the potential correlation between battery timing parameters and SOH by taking the voltage, current, and temperature profiles in the charging and discharging curves as network inputs. Ref. [11] proposed an intelligent battery health prediction method based on integrated deep transfer learning (EDTL) for efficient robust aging trajectory matching. It is demonstrated experimentally that the method can achieve accurate and reliable battery health prediction under different aging stages or incomplete data. Ref. [12] used a deep confidence network based on partial incremental capacity curves to evaluate the health of lithium-ion batteries, and the evaluation error can be reduced to less than 5%. The above data-driven SOH assessment method realizes the accurate assessment of SOH by mining the potential trends of battery monitoring data and constructing mapping relationships. However, most of the data-driven methods are based on gradient iterative optimization to determine the internal parameters, and the lack of interpretability hinders the engineering application of the methods to a certain extent.

In order to accurately estimate the SOH, in addition to improving the accuracy and efficiency of the SOH estimation algorithm, it is also important to accurately extract the Health Features (HFs) that characterize the aging of the battery. For this reason, some scholars have conducted battery SOH assessment based on extracting battery health characteristics from BMS timing monitoring data. Ref. [13] used a genetic algorithm to search for the optimal charging voltage range and parameters, and the obtained parameters were directly used to build a machine-learning-based SOH model for Li-ion batteries. Ref. [14] proposed a battery health state estimation method based on long- and short-term memory neural networks, in order to accurately describe the aging mechanism of the battery, features such as time and energy were extracted from the battery data as features for health state estimation. Ref. [15] proposed a weighted least squares support vector machine (WLS-SVM)-based early prediction method for lithium-ion battery cycle life using health indicators as input.

With the research in recent years, SOH estimation methods for Li-ion batteries based on the Informer model have begun to emerge. Ref. [14] extracted the actual currents, voltages, and temperatures by combining the Gaussian Process Regression (GPR) model with the Informer model using the Informer customized convolution kernel. Although the Informer model can realize accurate SOH estimation, there are still several problems: firstly, the traditional Informer model has some limitations and local information extraction ability. Second, the Informer model, as a deep learning model, has too high of a computational cost.

Based on the above conclusions, this paper proposes an SOH prediction method based on multi-feature extraction and a CCN-Informer model. By introducing a causal convolutional neural network in the self-attention mechanism, the local information extraction capability is enhanced to capture the surrounding anomalous fluctuation information that has a large impact on the time point and learn it, while retaining the important positional

order and temporal information between the sequence data, and encoding the positional and temporal information of each sequence of the input data. The probabilistic sparse self-attention layer and distillation layer are utilized to control the sample computational cost of the encoder.

The main innovations of this paper include the following:

1. In order to accurately characterize the battery aging process, eight features were extracted from the battery data and classified into two categories: direct measurement features and second-order processing features.
2. An SOH estimation model for lithium-ion batteries based on causal convolutional neural network and the Informer model is established, which enhances the local information extraction capability of the Informer model, and is able to well capture the nonlinear relationship reflected by different features in the aging process of lithium-ion batteries.

The remaining chapters of this paper are organized as follows: Section 2 describes the process of extracting health features, Section 3 describes the health state assessment model based on CCN-Informer, Section 4 validates the possibility and accuracy of the methodology, and Section 5 summarizes the whole paper.

2. Health Feature Extraction Based on Battery Charging Profile

2.1. Health Feature Extraction Framework

Battery SOH: The maximum usable capacity of a usable battery is measured as the ratio of the nominal rated capacity of the battery [16], which is:

$$V_{SOH} = \frac{C_{now}}{C_{rated}} \times 100\% \tag{1}$$

where V_{SOH} is the SOH of the battery, C_{rated} is the rated capacity of the battery, and C_{now} is the current maximum usable capacity of the battery. The rated capacity is measured and calibrated before the battery leaves the factory, while the current maximum usable capacity is measured after the battery charge/discharge cycle is completed. Most of the existing battery BMSs can only monitor parameters such as battery charging voltage, current, and temperature, and it is difficult to directly measure the current maximum capacity of the battery. Therefore, existing data-driven methods usually take the battery monitoring parameters as inputs and construct mapping models to realize the estimation of battery SOH.

2.2. Experimental Dataset

CALCE: In this experiment, three lithium batteries were selected from the CALCE dataset, CS2-34, CS2-36, and CS-37, respectively, for experimental study [17,18]. The data were obtained from the lithium battery experimental system of ArbinBT2000, and the rated capacity of the battery was 1.1 Ah, which was charged and discharged at room temperature. Firstly, constant-current charging was carried out at a constant current of 0.45 A until the voltage reached 4.2 V, keeping the terminal voltage of the battery at 4.2 V, and continued charging until the charging current decreased to 0.05 A, and then discharged at a constant current of 0.45 A until the voltage decreased to 2.7 V. The discharge rate was kept at a constant 1 C during the experimental process.

NASA: The NASA battery degradation dataset is based on 18650 model lithium-ion batteries [18,19]. There are three batteries cited in this dataset, namely B0005, B0006, and B0007. The three battery datasets were obtained by charging and discharging the batteries in operation mode at room temperature of 24 degrees. The rated capacity of all three battery sets is 2 Ah.

The battery degradation curves for the two datasets are shown in Figure 1a,b.
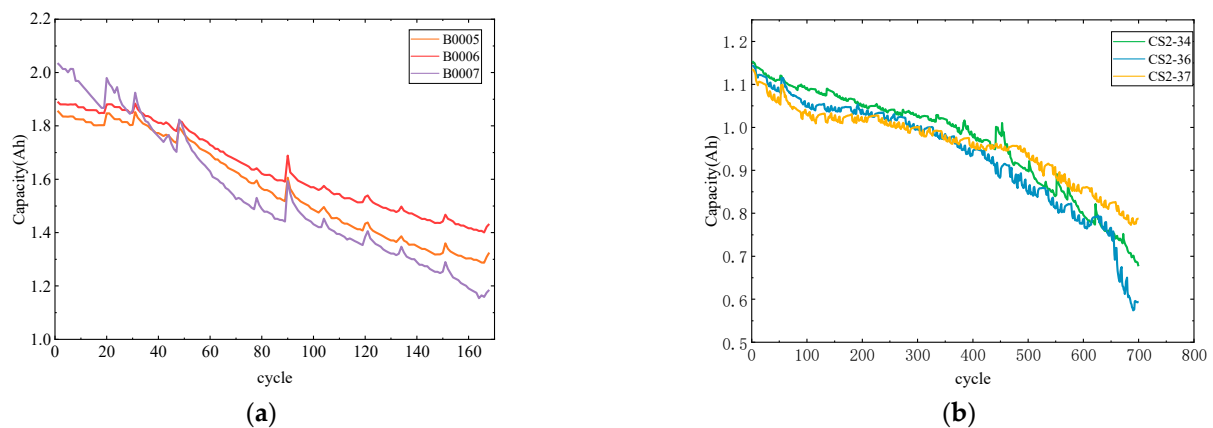
Figure 1. (**a**) Battery Capacity Degradation Curve for NASA Dataset; (**b**) Battery capacity degradation curves for the CALCE dataset.

2.3. Direct Measurement Feature Extraction

A large number of studies have shown that the charging voltage curve has a certain rule of change over time [20]. Taking the data of the battery numbered CS2-36 as an example, Figure 2a shows the charging voltage curves under different cycle times, with the increase in cycle times, the constant-current charging time is decreasing, while the constant-voltage charging time is increasing. For example, when cycling 30 times, the constant-voltage charging time is about 1500 s, while when cycling 150 times, the constant-voltage charging time increases to 3000 s. This phenomenon is caused by the decay of Li-ion battery materials over time. The charging time in the constant voltage stage tends to increase, and the slope of the curve in the constant-current charging stage gradually becomes steeper. In addition, the charging process of the battery is more stable than the discharging process, and the features extracted from it can be better adapted to different working conditions. Therefore, in this paper, four typical measurable parameters are extracted from the charging voltage curve of Li-ion batteries as HFs: constant-current charging time, constant-voltage charging time, optimal interval isobaric charging time, optimal interval isobaric charging voltage rise, and are named as HF1 to HF4. The specific meanings corresponding to each characteristic parameter in the charging voltage curve are shown in Figure 2b.
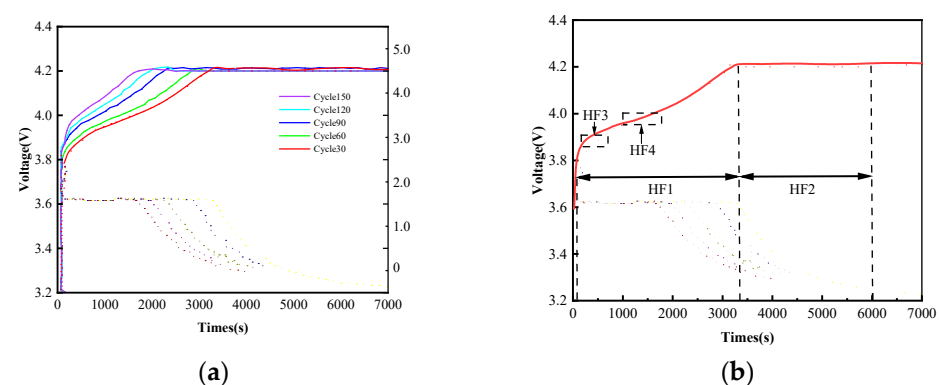

Figure 2. (**a**) Variation in charging voltage profile with different number of cycles; (**b**) Schematic diagram of directly extracted health features at 30 cycles.

In addition, in order to extract the segment features HF3 and HF4 in the voltage profile, this paper uses the traversal method, which traverses the charging voltage profile by adjusting the time–voltage interval in order to obtain the relevant charging intervals of the features HF3 and HF4 [21,22]. And using the correlation calculation method in Section 2.5, the feature parameters corresponding to the time intervals and voltage intervals with the largest correlation coefficients are taken as the target features HF3 and HF4, which

finally result in the charging intervals of [300, 450] s for HF3, and [3.85, 4.00] V for HF4. This traversal process is shown in Figure 3.
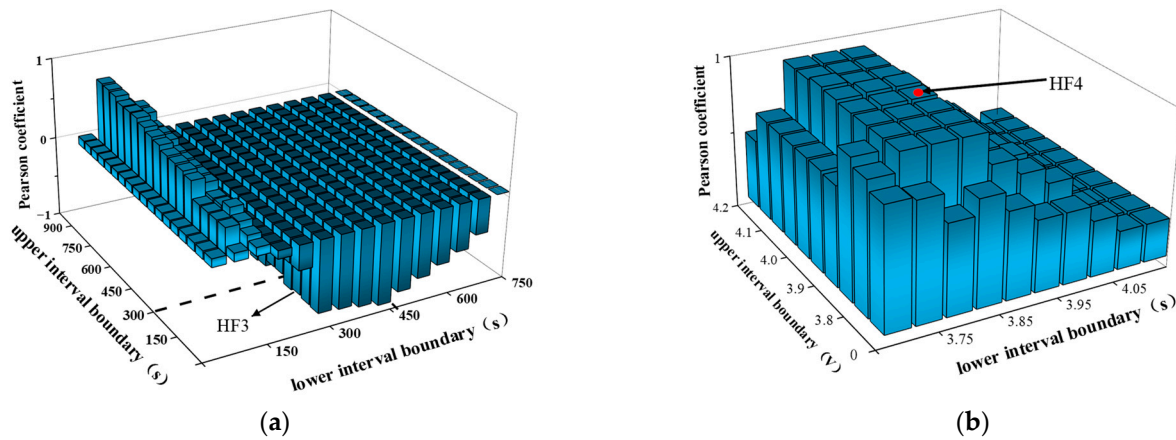

(**a**)　　　　　　　　　　　　　　　　(**b**)

Figure 3. (**a**) Charging time interval Pearson coefficient calculation results; (**b**) Charging voltage range Pearson coefficient calculation results.

2.4. Second-Order Processing Feature Extraction

Battery second-order processing curves are obtained by derivation calculations on the battery monitoring data collected by the BMS, such as battery capacity increment curves, differential voltage curves, and so on.

The first-order derivative of the terminal voltage–capacity (*V–Q*) curve under constant-current charging or discharging conditions is used to obtain the terminal voltage–capacity rate of change (*V–dQ/dV*) curve, based on which the incremental capacity analysis (ICA) method is developed [23]. This type of method is based on the battery charge/discharge voltage curve, and several characteristic parameters (including peak height, peak position, etc.) are extracted from the IC curve to be analyzed as health characteristics. The IC curve can illustrate the phase change characteristics of the battery during the insertion of lithium-ion active materials, including information about lithium ions participating in electrochemical reactions under different aging states of the battery. During the CC discharge of the Li-ion battery, the IC calculation formula can be expressed as:

$$\frac{dQ}{dV} = \frac{dQ}{dT} \cdot \frac{dT}{dV} = I \cdot \frac{dT}{dV} \tag{2}$$

where *Q*, *V*, *I*, and *T* are the battery capacity, charge/discharge voltage, charge/discharge current, and charge/discharge time, respectively.

Figure 4 shows the IC curves of Li-ion batteries under different cycle times. It can be found that the peak value of IC gradually decreases and moves to the right as the number of cycles increases [24]. The reason for this phenomenon is that the electron transfer ability of the internal active material embedded in lithium gradually decreases during the charging and discharging process of the battery. In this paper, the peaks of the IC curves and the voltages at their corresponding positions are extracted as IC HFs, which are denoted as HF5 and HF6.

Differential Voltage (DV) value combines the external characteristics of the battery with the internal electrochemical process to reflect the degradation state of the battery capacity through the rate of change in voltage under a specific capacity increment [25]. According to the external characteristics of the battery, the capacity of the battery is measured by coulometer method, and its calculation formula is shown in Equation (3); then, the differential voltage curve is expressed in Equation (4). According to the electrochemical process inside the battery, the differential voltage of the battery is calculated as Equation (5). The DV curve of battery charging is shown in Figure 5.

$$Q = \int_{t=1}^{r} I \mathrm{d}t \tag{3}$$

$$\frac{\mathrm{d}V}{\mathrm{d}Q} = \frac{\mathrm{d}V}{I \times \mathrm{d}T} = \frac{1}{I} \times \frac{\mathrm{d}V}{\mathrm{d}T} \tag{4}$$

$$\frac{\mathrm{d}V_{cell}(Q)}{\mathrm{d}Q} = \frac{\mathrm{d}V_p(Q)}{\mathrm{d}Q} - \frac{\mathrm{d}V_n(Q)}{\mathrm{d}Q} = \frac{1}{m_p} \cdot \frac{\mathrm{d}V_p(q_p)}{\mathrm{d}q_p} - \frac{1}{m_n} \cdot \frac{\mathrm{d}V_n(q_n)}{\mathrm{d}q_n} \tag{5}$$
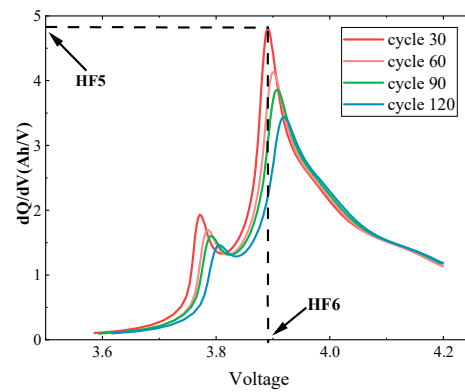

Figure 4. Incremental capacity curves.

From Equations (3)–(5) [26], it can be seen that the differential voltage inflection point value and its location are related to the battery capacity degradation [26]; therefore, the initial inflection point value of the charging differential voltage curve and its corresponding amount of time, and the peak value of the discharging differential voltage curve and its corresponding voltage value, are used as the indirect health factors for the SOH prediction of the batteries, which are denoted as HF7-HF8.
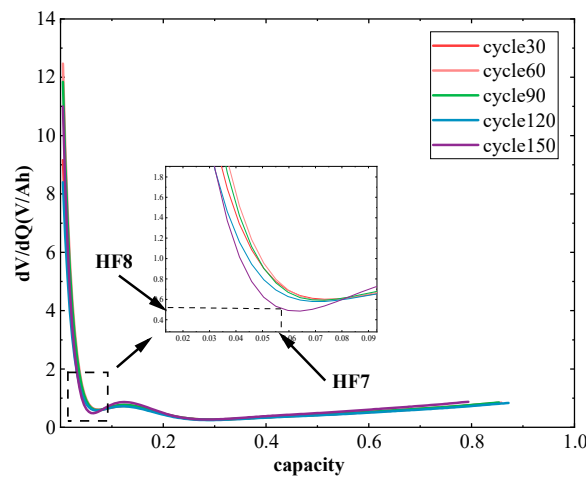

Figure 5. Differential Voltage Curve.

2.5. Feature Correlation Analysis

From the above analysis, it can be seen that the pattern of change of each extracted HF with battery aging is not consistent, and it is difficult to directly determine its correlation with capacity. Therefore, this paper adopts the Pearson correlation coefficient to further quantify the characteristic expression ability of each HF [27]. The Pearson correlation coefficient can effectively analyze the linear correlation between two variables, the larger the absolute value indicates that the correlation between the variables is higher, and if

the absolute value is 1, it means that the two are completely correlated. Its calculation expression is:

$$Pearson = \frac{E(xy) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}} \tag{6}$$

where X and Y are the extracted feature samples and battery capacity samples, respectively. Table 1 shows the Pearson correlation coefficients between each HF and battery capacity for four different batteries, which confirms the correlation between the eight extracted feature parameters and the battery capacity.

Table 1. Pearson correlation coefficient.

		HF1	HF2	HF3	HF4	HF5	HF6	HF7	HF8
	CS2_34	0.9238	0.8630	0.9789	0.9686	0.9984	0.9265	0.8677	0.9739
CALCE	CS2_36	0.6109	0.5398	0.8283	0.9791	0.9884	0.7177	0.6233	0.8001
	CS2_37	0.9245	0.8215	0.9730	0.9287	0.9927	0.8243	0.9281	0.9656
	B0005	0.9303	0.8715	0.9731	0.9516	0.9947	0.8725	0.9321	0.9668
NASA	B0006	0.9282	0.9799	0.9724	0.9818	0.9455	0.8487	0.9439	0.9754
	B0007	0.9686	0.9491	0.9809	0.9756	0.9211	0.8977	0.9502	0.9234

3. Causal Convolutional Self-Attention-Based Health State Prediction Model for CCN-Informer Batteries

3.1. Causal Convolutional Neural Network

Causal convolutional neural network (CCN) consists of input, hidden, and output layers, respectively. Each layer uses the same type of neurons, and each layer is realized by masking with non-full connectivity between them [28]. Its structure is shown in Figure 6.
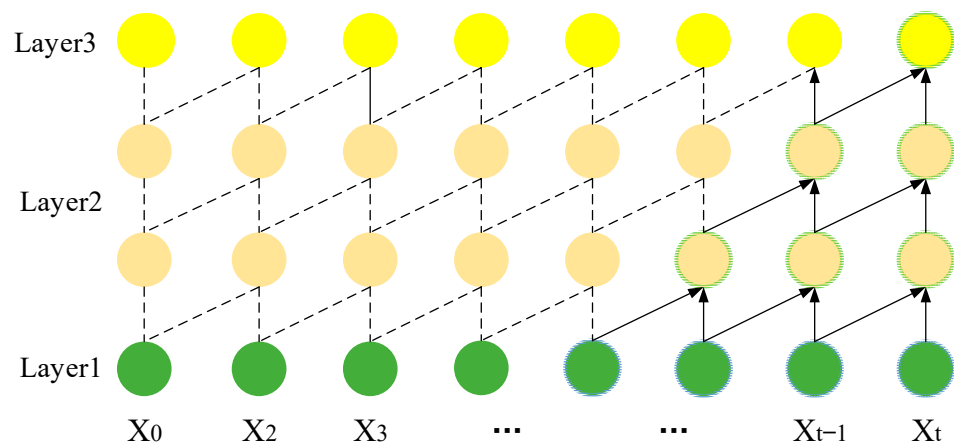

Figure 6. Schematic of a causal convolutional neural network.

Layer1 denotes the input layer for inputting feature vectors, Layer2 denotes the hidden layer for extracting information from the original data, and Layer3 denotes the output layer for outputting results. A convolution kernel of size k is used, and the expression of the causal convolution function F for input X is:

$$F(s) = (X * f)(s) = \sum_{i}^{k-1} f(i) \cdot x_{s-i} \tag{7}$$

where s denotes the input time series information, and $s - i$ denotes the localization of a certain location information in history. The convolution kernel k is used to set the sampling frequency of the network, and the edge positions are appropriately filled to ensure that the output and input dimensions are consistent. By deepening the layers of the network, the

next layer of the causal convolutional neural network will have a wider field of view of the historical information of the neurons in the previous layer. The causal convolutional neural network in Figure 6 mainly stacks three one-dimensional convolutional layers with convolution kernel 2, and after three convolutions, the output has four fields of view of historical information.

3.2. Self-Attention Mechanisms Incorporating Causal Convolutions

The self-attention mechanism focuses on the correlation between different parts of the whole input, enabling the model to pay more attention to the critical information of the current task and less attention to other non-critical information in a large amount of information. The formula for the original self-attention internal operation model is:

$$A(Q, K, V) = \text{Softmax}(QK^T \cdot d^*)V \tag{8}$$

In the conventional self-attention mechanism, the correlation between query q and index k is calculated based on the point-by-point dot product of the two, and the corresponding weighted aggregation is carried out, which does not incorporate the local information before and after the current node, and it is easy to ignore the influence of the change in the surrounding environmental factors at the current point in time.

The causal convolutional self-attention in this paper incorporates causal convolutional neural network characteristics in the canonical self-attention, fully considering that the time data have a strong dependence on the characteristics of the temporal dimension as well as the abnormal fluctuation problem that may occur; it deals with the time series problem by incorporating the input sequences into a one-dimensional causal convolutional module, using the convolution to extract the information around each node, and then utilizing the self-attention mechanism to learn the node relationships between them. At this point, the query q and index k in self-attention rely on the convolved information for computation, which enhances the similarity expression ability of the two in combination with local information, and the value vector maintains the characteristics of the sequence itself. By choosing different sizes of convolution kernel k, the causal convolution self-attention is able to focus on the influence of different scopes of local temporal information around the time point, instead of only considering the long-term dependency of a single time point of the regular pattern, which helps to make accurate predictions. The internal structure of the self-attention mechanism incorporating causal convolution is shown in Figure 7 in the following steps:

(1) Using the feature that causal convolution can capture local features, the input time series is subjected to a one-dimensional convolution operation, the one-dimensional convolution kernel parameter is set to k, which is used to set the sampling frequency of the network, and the number of convolution kernels is g_n, which yields the query vector $q[k, g_n]$ and the key value $k[k, g_n]$.

(2) Q and K for similarity calculation, which is Softmax-normalized to obtain the weights on the time series as shown in Equation (8).

(3) After the fully connected layer is converted into a value vector V with the same shape as the result of the weight calculation, it is subjected to matrix dot product to obtain the data with attention features.

3.3. Informer Model Structure

Informer is a supervised learning model based on self-attention mechanism, which is mainly composed of four parts: input module, encoder, decoder, and output module. The encoder can obtain robust remote dependence of long sequence inputs, and the decoder realizes one-step generation of predicted output results, which effectively solves the problems of Transformer in time series prediction and computational complexity. The

interior of its encoder consists of a stack of multi-head probabilistic sparse self-attention and distillation layers [29]. The encoder input vector expression is:

$$x_{en} = x^t \tag{9}$$

where the encoder input x_{en} comes from the historical data of the training model. In addition, the decoder input vector expression is:

$$x_{ed} = x^t_{2nd-half} + x_0 \tag{10}$$

where the decoder input x_{ed} consists of a partial sequence from the encoder and a zero-filled prediction sequence. The Informer model uses multi-head probabilistic sparse self-attention in the encoding and decoding layers to optimize the point-by-point accumulation operation of the original Transformer self-attention, which reduces the cost of the original self-attention quadratic dot product computation. Probabilistic sparsity self-attention uses a small number of dot products contributing to the main attention to pick out values whose probability distribution of attention for the corresponding query is far from the uniform distribution, converting the self-attention formula expression into a probabilistic formula:

$$A(q_i, k, v) = \sum_j^{L_k} \frac{k(q_i, k_j)}{\sum_i k(q_i, k_i)} = E_\rho(k_j \mid q_i)[v_i] \tag{11}$$

$$p(k_j, q_i) = \frac{k(q_i, k_j)}{\sum_l k(q_i, k_l)} \tag{12}$$

$$q(k_j, q_i) = \frac{1}{L_k} \tag{13}$$

where let q_i, k_l, and v_i denote the i-th row in Q, K, and V, respectively. $p(k_j, q_i)$ denotes the probability distribution of the attention of the ith query for all keys. $q(k_j, q_i)$ denotes the uniform distribution of query.
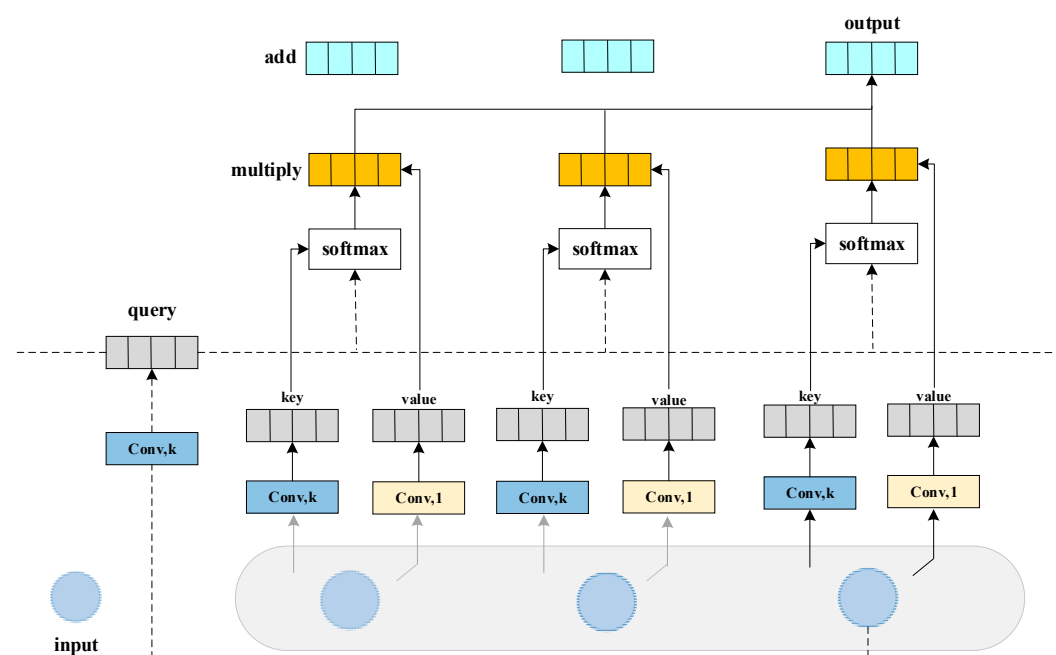

Figure 7. Causal Convolution-Self-Attention running mode.

Using KL scatter formula, the distance between *P* and *Q* distributions is calculated as a measure of the sparsity of the query, and the discrete KL scatter is defined as follows:

$$D(P \parallel Q) = \sum_{i \in X} P(i) * [\log(\frac{P(i)}{Q(i)})] \tag{14}$$

$$D(P \parallel Q) = \int_x P(x) * [\log(\frac{P(i)}{Q(i)})]dx \tag{15}$$

The attention probability distribution and uniform distribution of the query are added to the KL dispersion, and the approximate expression for the sparsity of the query is finally defined as:

$$KL(q \parallel \mathbf{p}) = \max_j\{q, k_j^T \cdot d^*\} - \frac{1}{L_k}\sum_{j=1}^{L_k}(q_i k_j^T \cdot d^*) \tag{16}$$

In the formula, the first term calculates the inner product of the *i*th query and all keys and selects the maximum value of it. Comparing with the arithmetic mean of the second term, the larger the difference of the result, the larger the variability of *p* and *q* is. In the set selection interval, the query with the higher difference ranking is selected, at this time, the dot product operation formula of sparsity probability self-attention in the Informer model is:

$$A(Q, K, V) = \text{Softmax}(\overline{Q}K^T \cdot d^*) \tag{17}$$

where *Q* denotes the first *n* query selected by sparsity computation. At this point, the value corresponding to the unselected query will be updated using the mean value.

Convolutional pooling is added between adjacent self-attention layers to perform a down-sampling operation on the output features of the self-attention, which effectively reduces the network parameters and output dimensions. The distillation operation from self-attention *m* to *m* + 1 layer is calculated as:

$$X_{m+1}^t = MAXPool(ELU(Convld[X_{m+1}^t]_{AB}) \tag{18}$$

where $[X_{m+1}^t]_{AB}$ denotes the output information of the sparse self-attention with multiple probabilities, Convld denotes the maximum pooling one-dimensional convolution operation on the time series, and ELU denotes the activation function. After the pooling layer, the length of the output sequence of the self-attention layer will be reduced by half.

The mask self-attention mechanism is generally used in decoders and serves to ensure that the prediction of the current moment is not affected by unknown future information. After the self-attention computation in this layer, the output query q and the index *k* and value of the encoder output in the last layer are fed into the encoder–decoder multi-attention layer, where the features of the encoder are incorporated into the decoder and recomputed [30,31]. Finally, after the fully connected layer, the long time series prediction output is obtained.

The prediction model based on causal convolutional self-attention and Informer is shown in Figure 8.
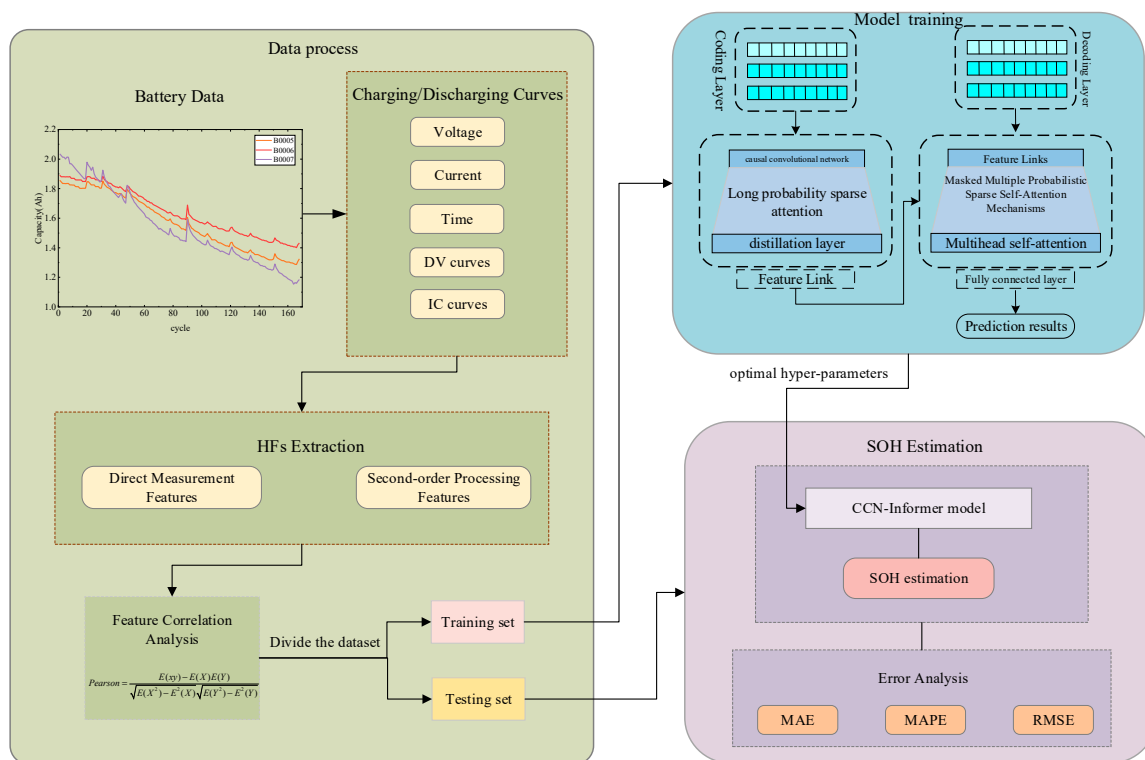
Figure 8. A Framework for Predictive Modeling Based on Causal Convolutional Self-Attention and Informer.

4. Calculation Validation and Analysis

4.1. Evaluation Indicators

In order to better quantify the SOH assessment accuracy of the proposed method, this paper introduces the mean absolute error (MAE), root-mean-square error (RMSE), mean absolute percentage error (MAPE), and decision coefficient (R^2) to evaluate the error between the assessed value of the algorithm and the true value, which is calculated as shown in the following formula [32,33]:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| y_i - y_i^* \right| \tag{19}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(y_i - y_i^* \right)^2} \tag{20}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \frac{\left| y_i - y_i^* \right|}{y_i} \tag{21}$$

$$R^2 = 1 - \frac{\sum\limits_{i=1}^{N} \left(y_i - y_i^* \right)^2}{\sum\limits_{i=1}^{N} \left(y_i - \overline{y} \right)^2} \tag{22}$$

where y_i is the true SOH of the battery under the ith charging cycle; y_i^* is the SOH of the battery evaluated by the model under the ith charging cycle; and N is the total number of charging cycles.

4.2. Analyzing and Validating the Prediction Results of Different Datasets

After constructing the structure of the CCN-Informer model, the datasets are analyzed and validated. In this paper, the leave-one-out validation method is used to produce the training set and validation set, i.e., the data of one battery are selected each time to produce the validation set, and the data of the remaining batteries are used as the training set, and it is repeated for four times until the data of each battery are validated as completing the validation process once [30]. The battery charging data used for validation are extracted as health features and then input into the trained CCN-Informer model to obtain the health state assessment results of this battery, and a 95% confidence interval is introduced to judge the reliability of the prediction results, i.e., the maximum allowable error range is [−0.05, 0.05].

Figure 9 shows the SOH estimates and their relative error distributions for the six cell monomers under the CCN-Informer model. It can be found that the SOH curves predicted by the two types of datasets with a total of six battery monomers are in general agreement with the actual situation. The bandwidth in the figure indicates the uncertainty of the estimated model. In general, the smaller the 95% area, the more reliable the fixed value. The estimation result plots show that the bandwidths of all six tested battery cells are small. By analyzing the relative errors of the different cells, it can be seen that the relative errors on each cell are small, with most points having deviations in the range of [−0.02, 0.02]. The results of the monomer SOH prediction for the six cells are given in Table 2. From the table, it can be seen that the best estimated single cell in the CALCE dataset is CS2_36, with MAE, MAPE, RMSE, and R^2 of 0.40%, 1.50%, 0.65%, and 99.87%, respectively. The best SOH-estimated cell in the NASA dataset is B0007, with MAE, MAPE, RMSE, and R^2 of 0.27%, 1.27%, 0.23%, and 99.89%, respectively. The above results show that the method proposed in this paper can accurately identify the SOH of Li-ion batteries.
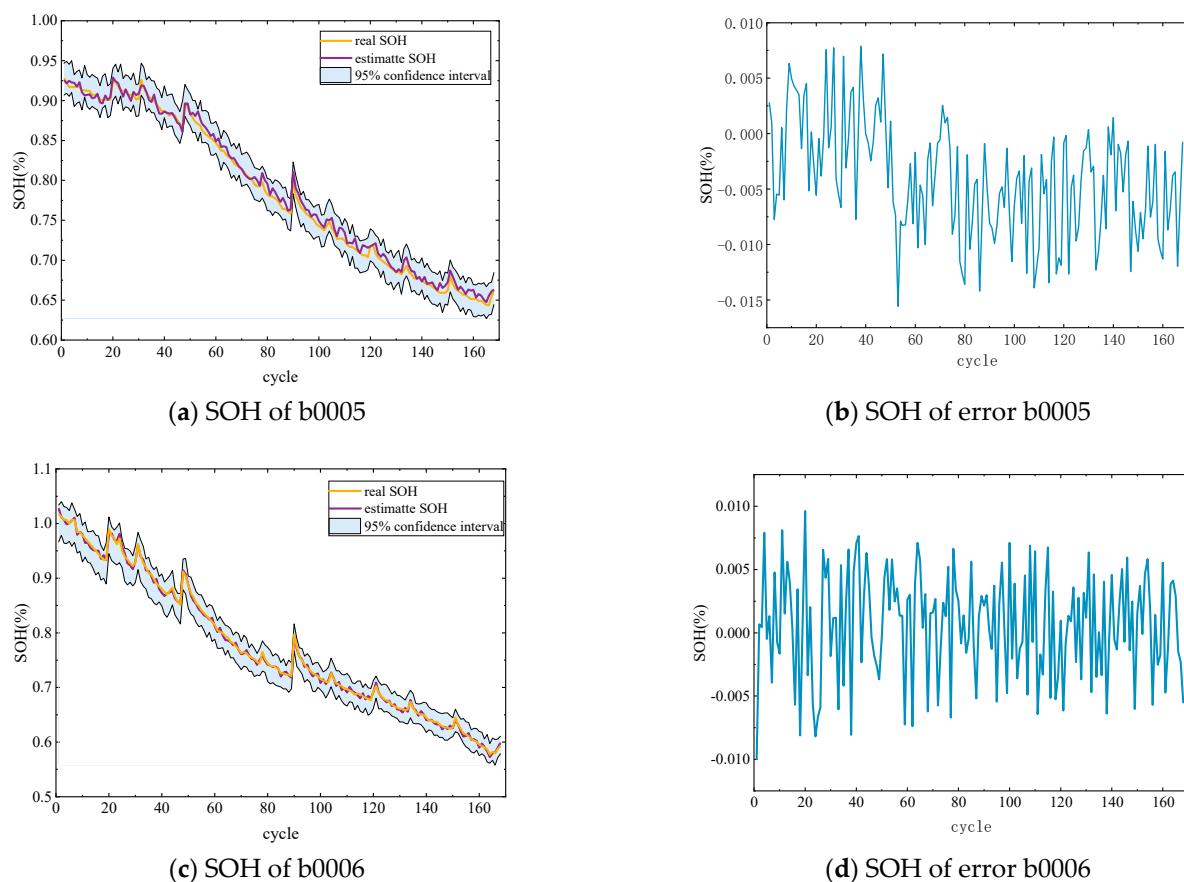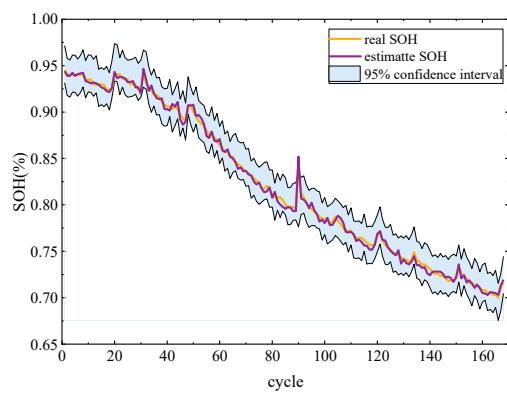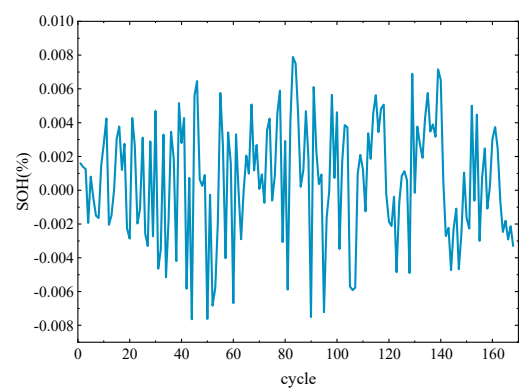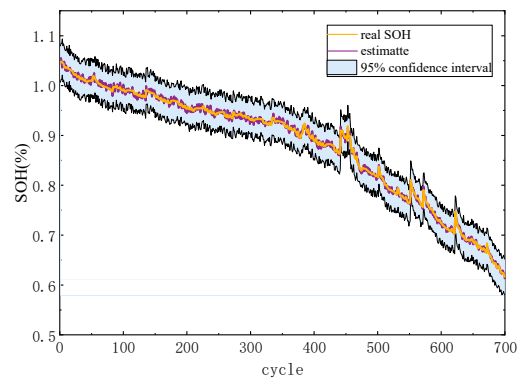

(**a**) SOH of b0005

(**b**) SOH of error b0005

(**c**) SOH of b0006

(**d**) SOH of error b0006

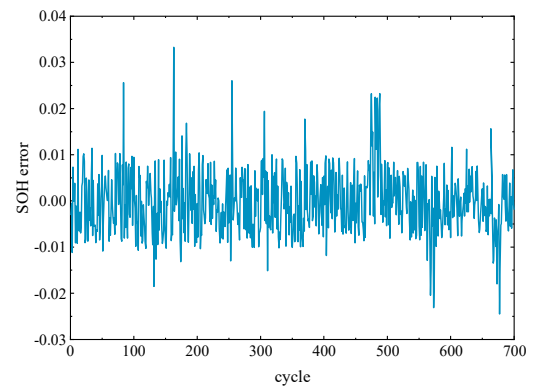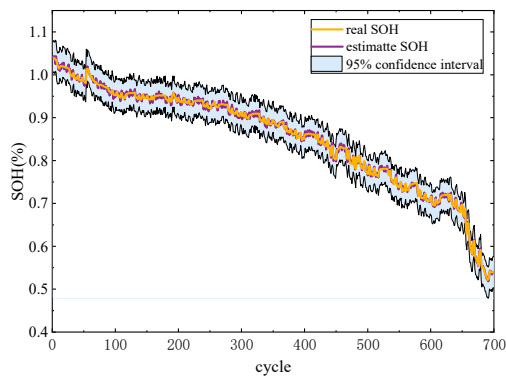**Figure 9.** *Cont.*

(**e**) SOH of b0007

(**f**) SOH of error b0007
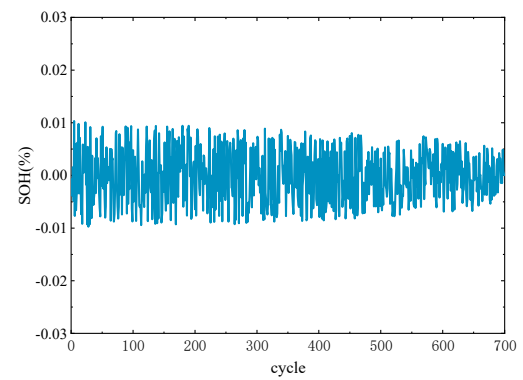
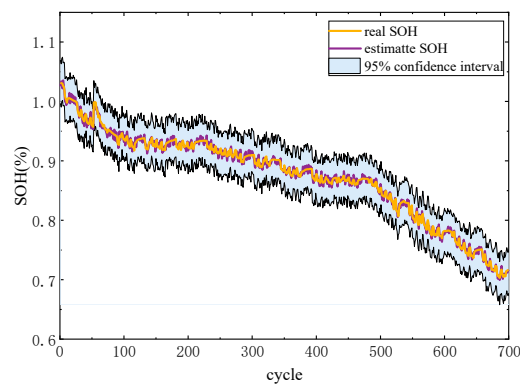(**g**) SOH of CS2_34

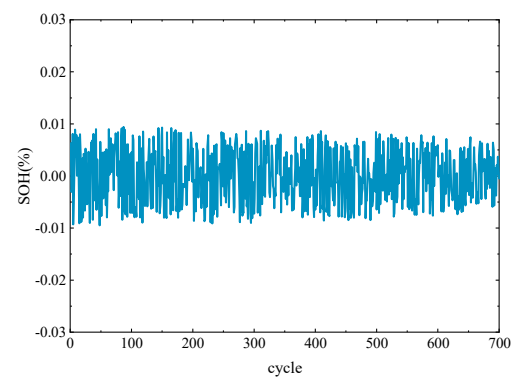(**h**) SOH of error CS2_34

(**i**) SOH of CS2_36

(**j**) SOH of error CS2_36

(**k**) SOH of CS2_37

(**l**) SOH of error CS2_37

Figure 9. Results of health state estimation under different datasets.

Table 2. SOH evaluation results of lithium-ion batteries based on CCN-Informer modeling.

		MAE	MAPE	RSME	R²
	CS2_34	0.0051	0.0153	0.0069	0.9977
CALCE	CS2_36	0.0040	0.0150	0.0065	0.9987
	CS2_37	0.0043	0.0167	0.0066	0.9986
	B0005	0.0027	0.0131	0.0032	0.9988
NASA	B0006	0.0044	0.0161	0.0011	0.9854
	B0007	0.0027	0.0127	0.0023	0.9989

4.3. Comparative Analysis of Different SOH Prediction Models

In order to further validate the effectiveness and superiority of the proposed method, a comparative study with the Informer model and LSTM model, which are widely used in the existing Ref. [14], is conducted in this section. The SOH estimation results and errors of the three methods for different batteries are shown in Figure 10.
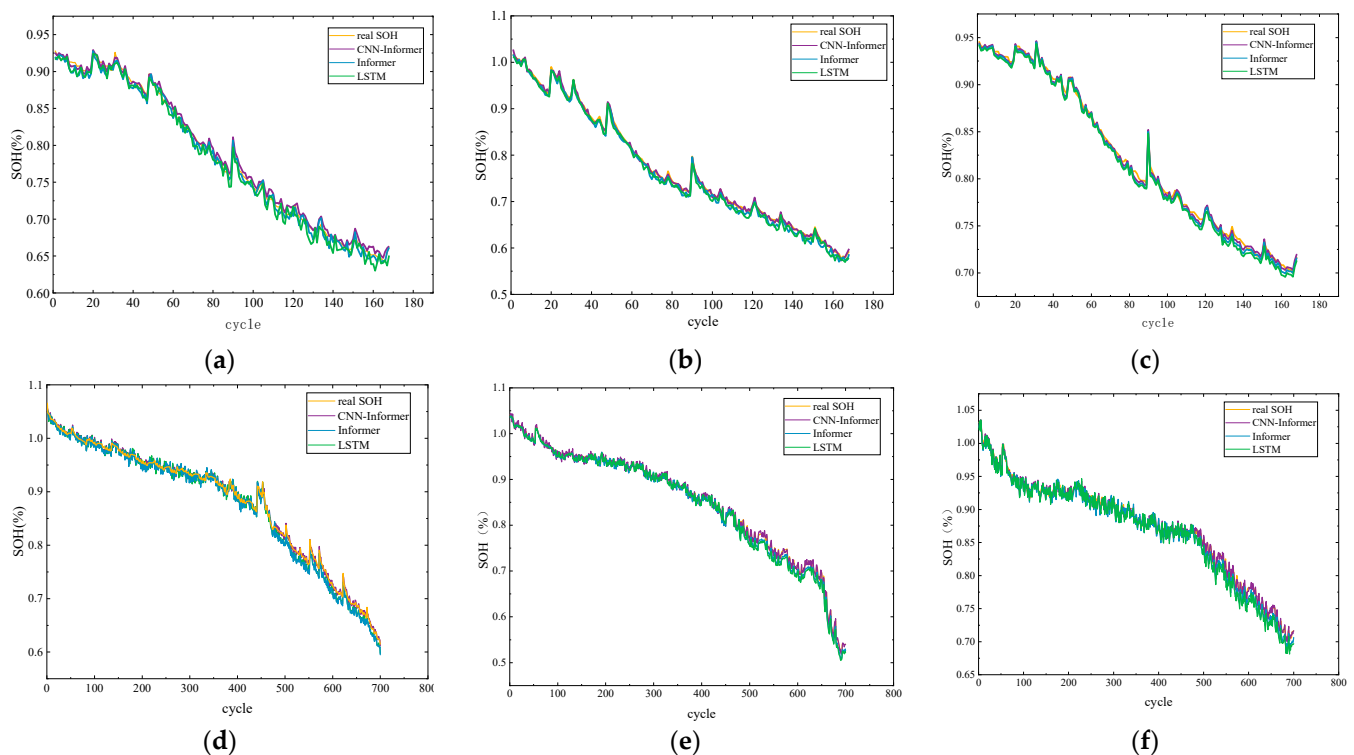

Figure 10. Results of health state estimation under different datasets. (**a**) b0005; (**b**) b0006; (**c**) b0007; (**d**) CS2_34; (**e**) CS2_36; (**f**) CS2_37.

From the figure, it can be seen that the prediction results of each model are good in the pre-prediction period, but the prediction errors of the Informer model and the LSTM model increase significantly with the increase in the number of prediction steps. This is because during the aging process of Li-ion batteries, the structural changes and complex electrochemical processes within the battery may not be immediately apparent in the pre-capacity degradation stage but have a cumulative effect in the long-term cycling. In the SOH prediction, the four lithium batteries constituting the training set and the test set follow the general aging pattern of the battery type, i.e., the aging trends of the batteries are basically the same. In the early-to-middle stage of prediction, the aging pattern of each battery is relatively single, and each model can basically fulfill the prediction task. However, in fact, it is difficult for the operating state and environment of each battery to be completely consistent, and the uncertainty factors in the prediction will affect the aging changes of different batteries and become more and more significant with the increase in

the number of cycles, so the SOH prediction of the sub-models becomes more difficult at the later stage of the prediction. In this paper, the uncertainty problem in the prediction process is solved by introducing the CCN-Informer method, and the CCN-Informer prediction model, compared with LSTM and other predictions, is able to effectively avoid the gradient explosion and other situations that occur in this type of deep neural network, and has higher stability and prediction accuracy, which reflects the advantages of the prediction model in this paper in the problem of SOH prediction of batteries.

Table 3 lists the MAE, MAPE, RMSE, and R^2 of the four prediction methods for estimating SOH, and it can be seen that the MAE, MAPE, RMSE, and R^2 of the method used in this paper are higher than that of the other three methods. Taking the CS2-34 battery as an example, the MAE, MAPE, RMSE, and R^2 of this paper's method are 0.51%, 1.53%, 0.69%, and 99.77%, respectively; those of the Informer model are 0.77%, 3.84%, 0.79%, and 99.02%, respectively; the MAE, MAPE, RMSE, and R^2 are 1.02%, 5.10%, 1.19%, and 97.99%, respectively, for LSTM. From the above analysis, it can be seen that the accuracies of the methods used in this paper are higher than the other two methods, and there is a significant improvement in the accuracy, which also reflects the importance of developing accurate battery state estimation methods.

Table 3. Evaluation results of different evaluation models for SOH of Li-ion batteries on CALCE and NASA datasets.

Battery	Method	MAE	MAPE	RSME	R^2
B0005	CCN-Informer	0.0027	0.0131	0.0032	0.9988
	Informer	0.0054	0.0283	0.0069	0.9971
	LSTM	0.0090	0.0492	0.0079	0.9967
B0006	CCN-Informer	0.0044	0.0161	0.0011	0.9854
	Informer	0.0063	0.0237	0.0040	0.9854
	LSTM	0.0108	0.0442	0.0045	0.9845
B0007	CCN-Informer	0.0027	0.0127	0.0023	0.9989
	Informer	0.0056	0.0323	0.0097	0.9821
	LSTM	0.0088	0.0486	0.0133	0.9788
CS2_34	CCN-Informer	0.0051	0.0153	0.0065	0.9986
	Informer	0.0077	0.0384	0.0079	0.9902
	LSTM	0.0102	0.0510	0.0119	0.9799
CS2_36	CCN-Informer	0.0040	0.0150	0.0069	0.9977
	Informer	0.0066	0.0412	0.0089	0.9921
	LSTM	0.0093	0.0633	0.0132	0.9798
CS2_37	CCN-Informer	0.0043	0.0167	0.0066	0.9986
	Informer	0.0066	0.0348	0.0141	0.9801
	LSTM	0.0087	0.0433	0.0135	0.9659

5. Conclusions

Battery health state assessment is a prerequisite for the safe operation of battery energy storage systems. In this paper, indicators that can reflect the health state of lithium batteries, such as IC, DV, and CC level charge/discharge curves, are fused to obtain the characteristics of lithium-ion battery SOH. The CCN-Informer-based lithium-ion battery SOH assessment and analysis framework is constructed to realize high-precision and stable and reliable lithium battery health state detection and safety assessment, and the main conclusions of this paper are as follows:

1. Eight features that can reflect the health state of the battery are extracted from the battery charge/discharge data, which are categorized into direct measurement features and second-order processing features to effectively and comprehensively describe the degradation of the battery.
2. Improved the Informer model. Add causal convolutional neural network to the self-attention mechanism to enhance the ability of local information extraction, so that the CCN-Informer model has the function of capturing both global and local information, avoiding gradient explosion, and optimizing the model prediction performance.

In summary, the lithium battery health state assessment method based on multiple health features and an improved Informer model proposed in this paper can effectively improve the accuracy and reliability of the prediction results, provide more valuable reference information for battery management and maintenance, and thus guarantee the safe and stable operation of energy storage power stations and other installations with energy storage needs.

References

1. Abdul, D.; Wenqi, J.; Tanveer, A. Prioritization of renewable energy source for electricity generation through AHP-VIKOR integrated methodology. *Renew. Energy* **2022**, *184*, 1018–1032. [CrossRef]
2. Shah, S.A.A.; Solangi, Y.A. A sustainable solution for electricity crisis in Pakistan: Opportunities, barriers, and policy implications for 100% renewable energy. *Environ. Sci. Pollut. Res.* **2019**, *26*, 29687–29703. [CrossRef] [PubMed]
3. Wen, J.; Zhao, D.; Zhang, C. An overview of electricity powered vehicles: Lithium-ion battery energy storage density and energy conversion efficiency. *Renew. Energy* **2020**, *162*, 1629–1648. [CrossRef]
4. Wang, D.; Zhang, Q.; Huang, H.; Yang, B.; Dong, H.; Zhang, J. An electrochemical–thermal model of lithium-ion battery and state of health estimation. *J. Energy Storage* **2022**, *47*, 103528. [CrossRef]
5. Song, Z.; Hou, J.; Li, X.; Wu, X.; Hu, X.; Hofmann, H.; Sun, J. The sequential algorithm for combined state of charge and state of health estimation of lithium-ion battery based on active current injection. *Energy* **2020**, *193*, 116732. [CrossRef]
6. Jiang, Z.; Li, J.; Li, L.; Gu, J. Fractional modeling and parameter identification of lithium-ion battery. *Ionics* **2022**, *28*, 4135–4148. [CrossRef]
7. Guo, X.; Kang, L.; Yao, Y.; Huang, Z.; Li, W. Joint estimation of the electric vehicle power battery state of charge based on the least squares method and the Kalman filter algorithm. *Energies* **2016**, *9*, 100. [CrossRef]
8. Hongwen, H.; Rui, X.; Hongqiang, G. Online estimation of model parameters and state-of-charge of LiFePO4 batteries in electric vehicles. *Appl. Energy* **2012**, *89*, 413–420.
9. Wei, Z.; Ruan, H.; Li, Y.; Li, J.; Zhang, C.; He, H. Multistage state of health estimation of lithium-ion battery with high tolerance to heavily partial charging. *IEEE Trans. Power Electron.* **2022**, *37*, 7432–7442. [CrossRef]
10. Fan, Y.; Xiao, F.; Li, C.; Yang, G.; Tang, X. A novel deep learning framework for state of health estimation of lithium-ion battery. *J. Energy Storage* **2020**, *32*, 10–41. [CrossRef]
11. Zhao, G.; Kang, Y.; Huang, P.; Duan, B.; Zhang, C. Chenghui Battery health prognostic using efficient and robust aging trajectory matching with ensemble deep transfer learning. *Energy* **2023**, *282*, 128228. [CrossRef]
12. Wu, J.; Fang, L.; Dong, G.; Lin, M. State of health estimation of lithium-ion battery with improved radial basis function neural network. *Energy* **2023**, *262*, 125380. [CrossRef]
13. Wei, J.; Dong, G.; Chen, Z. Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression. *IEEE Trans. Ind. Electron.* **2017**, *65*, 5634–5643. [CrossRef]
14. Liu, Q.; Kang, Y.; Qu, S.; Duan, B.; Wen, F.; Zhang, C. An online SOH estimation method based on the fusion of improved ICA and LSTM. In Proceedings of the 2020 IEEE/IAS Industrial and Commercial Power System Asia (I&CPS Asia), Weihai, China, 13–15 July 2020; pp. 1163–1167.
15. Xiong, W.; Xu, G.; Li, Y.; Zhang, F.; Ye, P.; Li, B. Early prediction of lithium-ion battery cycle life based on voltage-capacity discharge curves. *J. Energy Storage* **2023**, *62*, 106790. [CrossRef]
16. Choi, S.; Wang, G. Advanced lithium-ion batteries for practical applications: Technology, development, and future perspectives. *Adv. Mater. Technol.* **2018**, *3*, 1700376. [CrossRef]
17. Sui, X.; He, S.; Vilsen, S.B.; Meng, J.; Teodorescu, R.; Stroe, D.-I. A review of nonprobabilistic machine learning-based state of health estimation techniques for lithium-ion battery. *Appl. Energy* **2021**, *300*, 117346. [CrossRef]
18. Bian, X.; Wei, Z.; Li, W.; Pou, J.; Sauer, D.U.; Liu, L. State-of-health estimation of lithium-ion batteries by fusing an open circuit voltage model and incremental capacity analysis. *IEEE Trans. Power Electron.* **2022**, *37*, 2226–2236. [CrossRef]
19. Dong, J.X.; Yu, Z.S.; Zhang, X.K.; Luo, J.J.; Zou, Q.H.; Feng, C.; Ma, X.Q. Data-driven predictive prognostic model for power batteries based on machine learning. *Process Saf. Environ. Prot.* **2023**, *172*, 894–907. [CrossRef]
20. Zheng, L.F.; Zhu, J.G.; Lu, D.D.C.; Wang, G.X.; He, T.T. Incremental capacity analysis and differential voltage analysis based state of charge and capacity estimation for lithium-ion batteries. *Energy* **2018**, *150*, 759–769. [CrossRef]

21. Agudelo, B.O.; Zamboni, W.; Monmasson, E. Application domain extension of incremental capacity-based battery SoH indicators. *Energy* **2021**, *234*, 121224. [CrossRef]

22. Li, X.; Wang, Z.; Zhang, L.; Zou, C.; Dorrell, D.D. State-ofhealth estimation for Li-ion batteries by combing the incremental capacity analysis method with grey relational analysis. *J. Power Sources* **2019**, *410*, 106–114. [CrossRef]

23. Chen, D.H.; Zhang, W.G.; Zhang, C.P.; Sun, B.X.; Zhang, L.J.; Cong, X.W. Data-driven rapid lifetime prediction method for lithium-ion batteries under diverse fast charging protocols. *J. Energy Storage* **2023**, *74*, 109285. [CrossRef]

24. Su, X.; Sun, B.; Wang, J.; Zhang, W.; Ma, S.; He, X.; Ruan, H. Fast capacity estimation for lithium-ion battery based on online identification of low-frequency electrochemical impedance spectroscopy and Gaussian process regression. *Appl. Energy* **2022**, *322*, 119516. [CrossRef]

25. Wang, Z.; Yuan, C.; Li, X. Lithium battery state-of-health estimation via differential thermal voltammetry with Gaussian process regression. *IEEE Trans. Transp. Electrif.* **2021**, *7*, 16–25. [CrossRef]

26. Richardson, R.R.; Birkl, C.R.; Osborne, M.A.; Howey, D.A. Gaussian Process Regression for In Situ Capacity Estimation of Lithium-Ion Batteries. *IEEE Trans. Ind. Inform.* **2019**, *15*, 127–138. [CrossRef]

27. Montaru, M.; Fiette, S.; Kon'e, J.-L.; Bultel, Y. Calendar ageing model of Li-ion battery combining physics-based and empirical approaches. *J. Energy Storage* **2022**, *51*, 104544. [CrossRef]

28. Gong, D.; Gao, Y.; Kou, Y.; Wang, Y. State of health estimation for lithium-ion battery based on energy features. *Energy* **2022**, *257*, 124812. [CrossRef]

29. Li, Y.; Stroe, D.-I.; Cheng, Y.; Sheng, H.; Sui, X.; Teodorescu, R. On the feature selection for battery state of health estimation based on charging–discharging profiles. *Energy Storage* **2021**, *33*, 102–122. [CrossRef]

30. Liu, W.; Xu, Y. Data-driven online health estimation of Li-ion batteries using a novel energy-based health indicator. *IEEE Trans. Energy Convers.* **2020**, *35*, 1715–1718. [CrossRef]

31. Pang, X.; Liu, X.; Jia, J.; Wen, J.; Shi, Y.; Zeng, J.; Zhao, Z. A lithium-ion battery remaining useful life prediction method based on the incremental capacity analysis and Gaussian process regression. *Microelectron. Reliab.* **2021**, *127*, 114405. [CrossRef]

32. Damianou, A.; Lawrence, N.D. Deep gaussian processes. In *Artificial Intelligence and Statistics*; PMLR: New York City, NY, USA, 2013; pp. 207–215.

33. Xing, Y.; Ma, E.W.M.; Tsui, K.-L.; Pecht, M. An ensemble model for predicting the remaining useful performance of lithium-ion batteries. *Microelectron. Reliab.* **2013**, *53*, 811–820. [CrossRef]