



Article

# A Method for Estimating the SOH of Lithium-Ion Batteries Based on Graph Perceptual Neural Network

Kang Chen \* Dandan Wang and Wenwen Guo

School of Information Engineering, Zhengzhou College of Finance and Economics, Zhengzhou 450054, China  
\* Correspondence: chenkang@zzife.edu.cn

**Abstract:** The accurate estimation of battery state of health (SOH) is critical for ensuring the safety and reliability of devices. Considering the variation in health degradation across different types of lithium-ion battery materials, this paper proposes an SOH estimation method based on a graph perceptual neural network, designed to adapt to multiple battery materials. This method adapts to various battery materials by extracting crucial features from current, voltage, voltage–capacity, and temperature data, and it constructs a graph structure to encapsulate these features. This approach effectively captures the complex interactions and dependencies among different battery types. The novel technique of randomly removing features addresses feature redundancy. Initially, a mutual information graph structure is defined to illustrate the interdependencies among battery features. Moreover, a graph perceptual self-attention mechanism is implemented, integrating the adjacency matrix and edge features into the self-attention calculations. This enhancement aids the model's understanding of battery behaviors, thereby improving the transparency and interpretability of predictions. The experimental results demonstrate that this method outperforms traditional models in both accuracy and generalizability across various battery types, particularly those with significant chemical and degradation discrepancies. The model achieves a minimum mean absolute error of 0.357, a root mean square error of 0.560, and a maximum error of 0.941.

**Keywords:** lithium-ion battery; state of health; graph neural networks; self-attention mechanism; transformer



**Citation:** Chen, K.; Wang, D.; Guo, W. A Method for Estimating the SOH of Lithium-Ion Batteries Based on Graph Perceptual Neural Network. *Batteries* **2024**, *10*, 326. <https://doi.org/10.3390/batteries10090326>

Academic Editor: Vaclav Knap

Received: 7 August 2024

Revised: 6 September 2024

Accepted: 12 September 2024

Published: 13 September 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

In electric vehicles and renewable energy storage systems, batteries are core components, and their performance and reliability are of paramount importance. The state of health (SOH) of a battery is a crucial indicator for assessing its performance degradation and remaining lifespan [1,2]. Accurate estimation of the SOH is essential for ensuring safe operation, optimizing battery management systems (BMS), and extending the battery lifespan [3,4].

With advances in battery technology, various battery materials, including lithium-ion and nickel-metal hydride batteries, have emerged on the market. Each material has unique chemical characteristics and degradation patterns [5], leading to different requirements for SOH estimation methods. Traditional model-based methods [6,7] are typically grounded in well-defined physical principles, allowing for the explanation of battery degradation behavior by understanding internal chemical and physical processes. These methods also offer high adaptability, as model parameters can be adjusted to accommodate different battery types. However, they often require precise parameter tuning, have high computational complexity, and may struggle to maintain stable performance in complex environments. On the other hand, empirical methods [8], which rely on large datasets and leverage techniques such as pattern recognition and statistical analysis, enable rapid predictions. These methods are characterized by simplicity and flexibility, focusing on uncovering patterns from data rather than on physical or chemical principles. This often makes them more

efficient for handling complex systems. Nonetheless, they have notable limitations when data are scarce and lack the ability to explain internal battery mechanisms. As such, both traditional model-based and empirical methods often fail to fully meet the needs of all battery types, exhibiting constraints in terms of accuracy and generalizability. For example, Sang et al. [9] proposed a dual adaptive central difference H-infinity filter algorithm to address issues of low accuracy, slow convergence, and insufficient robustness in existing state of charge (SOC) and SOH joint estimation algorithms for practical applications. This algorithm improves measurement updates by incorporating a robust discrete H-infinity filter equation based on the central difference Kalman filter with high accuracy and integrates the Sage–Husa adaptive filter for joint SOC and SOH estimation. However, compared with data-driven methods, this approach is generally more complex, requires precise model adjustments, and may offer lower adaptability and flexibility to new environments. Additionally, Li et al. [10] developed a lithium-ion battery state estimation method based on an improved fractional-order model, employing an endogenous immune algorithm to identify model parameters, and incorporating aging effects through an adaptive dual square root Kalman filter for precise SOC and SOH estimation. The test results indicate that this method provides high-precision and robust battery state monitoring. However, model-based approaches may overly rely on specific settings and assumptions, limiting their applicability to new or different types of battery technologies.

In recent years, deep learning technology has shown significant potential in battery SOH estimation because of its exceptional feature extraction and pattern recognition capabilities [11]. Xu et al. [12] proposed a method that combines feature selection with CNN-LSTM model optimization (including skip connections), which effectively enhances the accuracy and efficiency of lithium-ion battery SOH estimation. A validation on two major battery datasets demonstrated the model's efficiency and robustness in handling complex battery data, indicating superior performance compared with traditional neural network models. Qian et al. [13] proposed an innovative attention mechanism-based multi-source sequence-to-sequence model (AM-seq2seq) to more accurately predict the SOH of lithium-ion batteries under dynamic load conditions. This model uniquely integrates the battery's historical state information with expected future load data, effectively capturing dependencies through advanced attention layers, thus enhancing prediction accuracy and model robustness. Peng et al. [14] developed a method to estimate the health status of lithium-ion batteries by combining multiple health feature extractions with an improved long short-term memory (LSTM) network. By comprehensively analyzing various health indicators and using gray relational analysis, along with optimizing neural network hyperparameters via an improved quantum particle swarm optimization algorithm, the accuracy and robustness of health status estimation were effectively enhanced. Miao et al. [15] proposed a dual-layer analysis framework that combines multiscale features with a long short-term memory (LSTM) network and random forest regression (multi-LSTM-RFR2) to increase the accuracy and interpretability of lithium-ion battery SOH estimation. This method achieves a deep understanding and high-accuracy prediction of battery health status by analyzing performance data across different scales. He et al. [16] proposed a new method for estimating the health status of lithium-ion batteries, which is particularly suitable for fast charging strategies. This method combines graph generation techniques with deep learning networks (graph convolutional networks and long short-term memory networks) to effectively handle the data challenges posed by fast charging. By analyzing voltage sequences and current changes during battery charging and using graphical data to analyze battery degradation characteristics in detail, this method can efficiently and accurately estimate battery health status. By introducing two types of undirected graphs and conditional graph convolutional networks, Wei et al. [17] improved the application of traditional graph convolutional networks (GCNs) in battery SOH and RUL prediction. This approach considers both feature correlations and temporal relationships through dual spectral graph convolution and dilated convolution operations, offering a more comprehensive data analysis. The experimental results show that this method outperforms existing

machine learning methods in terms of prediction accuracy, confirming its potential and effectiveness in battery health monitoring.

In particular, graph neural networks (GNN) provide a novel perspective for precise SOH prediction by modeling the intricate structural relationships and interactions within batteries [18]. However, current research often focuses on a single battery type and tends to overlook the detailed representation of complex interactions between different battery materials.

High-quality data are essential for training deep learning models, but data acquisition and quality pose significant challenges for lithium-ion batteries made from different materials. Variations in manufacturing processes, usage environments, and operating conditions can lead to data inconsistencies, affecting the model's ability to generalize. Second, training deep learning models requires diverse data that encompass various materials, usage conditions, and environmental factors to ensure robust generalizability. However, real-world data often lack sufficient diversity and representativeness, leading to sample bias issues. Finally, batteries composed of different materials display significant differences in electrochemical performance, aging mechanisms, and environmental responses, making it difficult for a single model to adapt to all types of batteries. The model may overfit data from specific materials, leading to poor performance on other materials. Data-driven methods can learn battery behavior patterns and degradation characteristics from large volumes of historical operational data without requiring an in-depth understanding of the battery's physical and chemical processes. By collecting operational data from different types of batteries under various conditions, machine learning models can be trained to identify unique degradation markers and influencing factors for different battery materials. Additionally, data-driven methods facilitate the implementation of feature engineering techniques. Effective feature extraction and engineering can enhance the model's adaptability to different battery materials, mitigate sample bias issues, and improve generalization with diverse data [19].

To address these issues, this paper proposes an SOH estimation method for different types of battery materials via graph-aware neural networks. First, 20 features were extracted from the following four aspects: current, voltage, voltage–capacity, and temperature data. A graph structure based on mutual information was constructed and trained via graph-aware neural networks to estimate the SOH of three different battery materials in the SNL dataset. The contributions of this work can be summarized as follows:

1. This study conducted a comprehensive analysis of key parameters, including current, voltage, voltage capacity, and temperature, from which 20 significant features were extracted from the battery's charge–discharge data. These features cover diverse aspects of battery operation, offering robust data support for an in-depth understanding of battery performance and health.

2. To increase the accuracy of capturing complex interactions and dependencies among battery components, this study introduces an innovative graph-aware self-attention mechanism. By integrating adjacency matrices and edge features into self-attention calculations, this approach not only improves the model's comprehension of battery behaviors but also greatly enhances the transparency and interpretability of the decision-making process through the direct inclusion of graph structural data.

3. By utilizing a graph structure informed by mutual information, this study successfully implemented SOH prediction for three different battery materials via graph-aware neural networks.

## 2. Data Preparation and Processing

### 2.1. Battery Dataset

Dataset A from Sandia National Laboratories [20] includes data from commercial 18650 NCA, NMC, and LFP batteries cycled to 80% capacity and then continuously. This study investigated the impacts of temperature, depth of discharge, and discharge current on the long-term degradation of batteries. Cyclic aging tests were performed via Arbin

SCTS and Arbin LBT21084 multichannel battery testing systems. Batteries were secured in 18650 battery holders connected to the Arbin system with an 18-gauge wire, with cable lengths maintained below 8 feet to minimize voltage drops. Cycling was carried out in an SPX Tenney T10C-1.5 environmental chamber, with surface temperatures monitored via K-type thermocouples.

At the beginning of the study, the batteries were brought to the required temperature in an environmental chamber for equilibration. The cycling protocol comprised initial capacity checks, defined cycling conditions, and follow-up capacity assessments. These checks involved three full 0–100% SOC charge/discharge cycles using a 0.5 C charging current. A full SOC was achieved through a 0.5 C charge and a discharge of 0.05 A at the maximum voltage. The number of cycles varied from 125 to –1000, with adjustments for any capacity loss exceeding 5%. The batteries underwent charging at 0.5 C and cycled under different SOC ranges as follows: 40–60% SOC (CC protocol), 20–80% SOC (fresh cell voltage limits), and 0–100% SOC (CCCV protocol). Specifically, LFP batteries cycled between 2 and 3.6 V, NCA batteries between 2.5 and 4.2 V, and NMC batteries between 2 and 4.2 V under 100% DOD conditions. This cycling protocol was consistently applied throughout the study.

Dataset B: The Toyota Research Institute dataset (in collaboration with Stanford University and the University of Maryland) [21] contains data from 124 commercial high-power LFP/Graphite A123 APR 18650 M1A batteries cycled in a temperature-controlled environment. The cycling process was conducted on a 48-channel Arbin LBT potentiostat with horizontal cylindrical clamps in a forced convection temperature chamber set to 30 °C. After a small portion of the plastic insulation layer was removed, T-type thermocouples were attached to the exposed battery casing via thermal epoxy (OMEGATHERM 201) and Kapton tape for temperature measurement.

All batteries in this dataset employed either a one-step or two-step fast charging approach. Initially, they were charged at a constant current ranging from 5.5 C to 6.1 C until a voltage of 3.6 V was reached. In the second phase, the charging current was maintained between 2 C and 4 C until the voltage again reached 3.6 V, reaching 80% SOC. Subsequently, the batteries were charged in 1 C CC–CV mode. The voltage limits were set at 3.6 V and 2.0 V, respectively, in line with the manufacturer's guidelines. These cutoff voltages remained consistent across all charging phases, including during rapid charging cycles. In some instances, batteries might reach the upper voltage limit during fast charging, leading to a notable period of constant voltage charging. The batteries were consistently discharged at a 4 C rate.

In this study, the ratio of the current available capacity of the battery to the rated capacity is used to express the SOH of the battery according to the following expression:

$$SOH = \frac{C_c}{C_i} \quad (1)$$

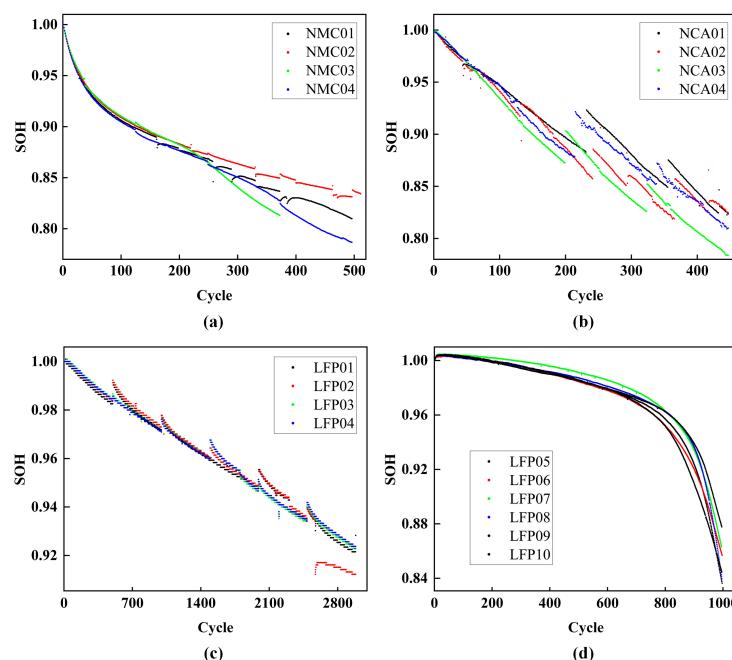
where  $C_c$  denotes the current maximum discharge capacity (remaining capacity) of the battery and  $C_i$  denotes the initial capacity.

Table 1 details the dataset specifications. Figure 1 displays a scatter plot of the SOH decay curves for three types of battery materials, illustrating distinct decay trajectories. LFP batteries exhibit a relatively linear decay pattern over numerous cycles, which is indicative of their superior structural stability and lower rate of capacity decay. Analysis from LFP01 to LFP10 indicates that these batteries maintain more than 90% SOH after 2500 and 1000 cycles, respectively, significantly outperforming the other materials. This remarkable stability is attributed to the robust structure of the LFP chemical system and its low risk of thermal runaway [22]. The gentle slope of the SOH curve further suggests that LFP batteries are well suited for applications demanding long cycle lives and consistent performance. In contrast, NCA batteries deteriorate more swiftly, with most falling below 85% SOH within approximately 400 cycles. The variability observed from NCA01 to NCA04 could be influenced by operational factors such as temperature and discharge rate. The

steep slope of their SOH curve highlights that, despite its high energy density, the NCA system suffers significant capacity loss, associated with its high operational voltage and sensitivity to thermal and mechanical stress [23]. The decay paths of the NMC batteries are between those of the LFP and NCA batteries. The SOH decreases to less than 85% after approximately 400 cycles, indicating a decay rate faster than that of LFPs but slower than that of NCAs. NMC01 to NMC04 show similar patterns with moderate variability, reflecting the balance between energy density and cycle life in the NMC system. The curve shape suggests that NMC batteries undergo both calendar and cyclic aging, with potential degradation mechanisms including electrolyte decomposition and active material loss.

**Table 1.** Commercial 18650-format lithium-ion battery manufacturer-specified operating bounds.

Battery	Dataset A		Dataset B	
	LFP(1–4)	NCA	NMC	LFP(5–10)
Nominal Capacity (Ah)	1.1	3.2	3	1.1
Nominal Voltage (V)	3.3	3.6	3.6	3.3
Voltage Range (V)	2 to 3.6	2.5 to 4.2	2 to 4.2	2 to 3.6
Max Discharge Current (A)	30	6	20	40
Nominal Mass (g)	39	6	20	-

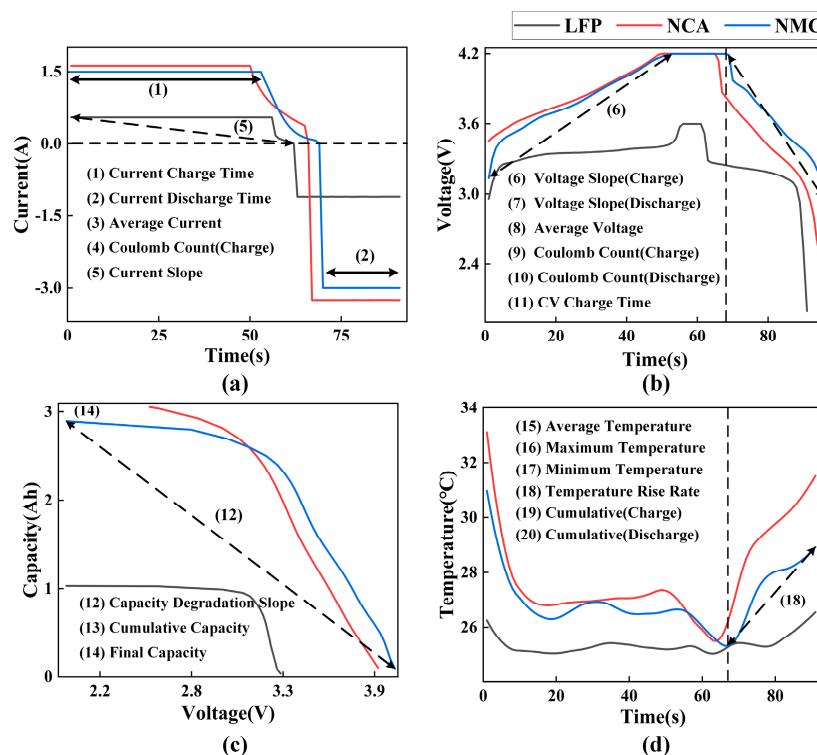


**Figure 1.** SOH decay curves for three battery material datasets: (a–c) for NMC, NCA, and LFP cells in Dataset A; (d) for LFP cells in Dataset B.

## 2.2. Feature Extraction

Feature extraction is essential for achieving accurate SOH estimation. This process simplifies data handling by converting large volumes of time series data from the charge-discharge cycles into key feature parameters, improving computational efficiency and reducing noise. It not only captures important physical and chemical changes during the battery's operational cycles, thereby enhancing the model's accuracy and robustness, but also provides a thorough assessment of the battery's overall health [24]. Multidimensional feature data are crucial in developing a detailed SOH estimation model that supports online monitoring and real-time diagnosis, improves model interpretability, and deepens the understanding of battery degradation mechanisms and performance change patterns [25]. Using Dataset A as an example, all the data were collected under consistent charge-discharge conditions, leading to similar trends in current, voltage, voltage–capacity

(IC), and temperature. To evaluate battery performance effectively, we extracted 20 features from these four categories. Figure 2 depicts the feature extraction process for the three types of battery materials. Given their different rated capacities, LFP batteries show distinct data trends compared with NCA and NMC batteries, which exhibit larger capacity fluctuations, whereas LFP batteries demonstrate smaller variations. Overall, the data curves for all three materials reveal similar trends. For example, in Figure 2d, the temperature curves drop to their lowest points after charge completion. The slope features mentioned below were all calculated as the ratio of the difference in the corresponding y-values to the difference in the x-values within the specified range.



**Figure 2.** Feature extraction process: (a) current profile, (b) voltage profile, (c) voltage–capacity profile, (d) temperature profile.

Group 1: From the current–time curve, five key features were extracted: charging time, discharging time, average current, coulomb count (during charging), and current slope. The charging and discharging times are indicative of the battery’s performance alterations across different cycling stages, assisting in assessing its charge–discharge efficiency [26]. The average current measures the battery’s stability and uniformity throughout standard cycles. The Coulomb count quantifies the total charged and discharged amounts by integrating the current over time, thus facilitating the assessment of the battery’s genuine capacity and its degradation. The current slope, which reflects the rate of current change, helps identify variations in the internal impedance and chemical reaction rates of the battery, offering essential baseline data for accurate SOH estimation. Together, these features provide comprehensive insights into the dynamic behaviors critical for SOH estimation of the battery.

Group 2: From the voltage–time curve, six features were extracted: charging voltage slope, discharging voltage slope, average voltage, coulomb count (during charging and discharging), and constant voltage charging time. The voltage slopes, both during charging and discharging, reflect the battery’s voltage response to the charge–discharge cycles and help identify voltage lag effects associated with aging. The average voltage assesses stability across different SOC levels, providing a consistent measure of battery behavior. By integrating voltage over time, the coulomb count can be used to evaluate the energy output

across varying voltage ranges, aiding in a comprehensive assessment of energy efficiency. Constant voltage charging time offers insights into the battery's performance and internal impedance alterations at elevated SOC stages. These voltage-related features are pivotal for SOH estimation, facilitating precise predictions of changes in internal resistance and electrochemical performance degradation [27].

Group 3: From the capacity–voltage curve, we extracted three features: capacity decay slope, cumulative capacity, and final capacity. The capacity decay slope reflects the rate of capacity change with voltage, revealing the actual capacity loss of the battery during use. The cumulative capacity is used to evaluate the working capacity of the battery in different voltage ranges. The final capacity directly reflects the battery's health status and remaining lifespan. Capacity features are core indicators for SOH estimation. By analyzing these features, the remaining useful life of the battery can be accurately estimated, identifying the battery's performance under different working conditions and providing data support for optimizing battery management systems [28].

Group 4: Six temperature-related features were extracted from the temperature–time curve: average temperature, maximum temperature, minimum temperature, temperature rise rate, and cumulative temperature during charging and discharging. The average, maximum, and minimum temperatures assess the battery's thermal management performance across various conditions. The temperature increase rate offers insights into the thermal response of a battery during charge–discharge cycles. The cumulative temperature quantifies the total thermal changes experienced during these cycles. These temperature metrics are crucial for SOH estimation. Monitoring these parameters helps identify potential thermal runaway risks under different conditions and facilitates the implementation of suitable thermal management strategies [29]. Cumulative temperature measurements are instrumental in evaluating long-term thermal impacts on battery life, enhancing the precision of SOH estimations and optimizing battery management.

### 2.3. Feature Selection

The number of features significantly influences model performance. The incorporation of more features increases model complexity, which can prolong the training time and increase computational demands. More features might also necessitate additional data to effectively learn and mitigate overfitting risks. Theoretically, an increased feature set can enrich the information base, potentially enhancing model performance for specific tasks, particularly if these features are pertinent to the prediction target. However, this approach can simultaneously diminish model interpretability and complicate the understanding of decision-making processes. Conversely, fewer features can improve the model's training efficiency and make it more interpretable, simplifying monitoring and maintenance. However, this might restrict the model's capacity to encapsulate essential complexities, thus impacting prediction accuracy. Consequently, choosing an optimal number of features is essential to balance model performance, efficiency, and interpretability [30].

#### 2.3.1. Correlation Analysis

The Pearson correlation coefficient is a statistical tool that quantifies the strength and direction of a linear relationship between two variables and is extensively utilized in feature selection and data analysis. It is calculated by dividing the covariance of the variables by the product of their standard deviations. The formula for this calculation is as follows:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (2)$$

Here,  $X_i$  and  $Y_i$  are the observed values of the two variables,  $\bar{X}$  and  $\bar{Y}$  are their means, and  $n$  is the total number of data points. The Pearson correlation coefficient values range

from  $-1$  to  $+1$ , where  $+1$  represents a perfect positive correlation,  $-1$  represents a perfect negative correlation, and  $0$  represents no linear correlation at all. This correlation measure is a crucial tool in data analysis and is particularly valuable in predictive analytics and model development. Identifying and discarding features with low correlation or focusing on those with high correlation can significantly enhance model performance and clarity. Table 2 presents the correlation analysis of 20 features. The correlation matrix helps identify features with high correlation. Notably, F1, F4, F9, F10, F13, F14, F16, F18, and F20 exhibit strong correlations with SOH. These features should be prioritized in future model development to improve prediction accuracy and model interpretability.

**Table 2.** Results of the correlation analysis between characteristics and SOH.

Features	Pearson Correlation Coefficient			Features	Pearson Correlation Coefficient		
	LFP01	NCA01	NMC01		LFP01	NCA01	NMC01
F1	0.9743	-0.9034	0.9784	F11	0.8912	0.5678	-0.6543
F2	0.6543	-0.5432	0.2345	F12	-0.9123	-0.6789	0.7654
F3	-0.7834	0.4321	-0.6456	F13	0.9345	0.9023	-0.9657
F4	0.9245	0.8654	-0.9345	F14	-0.9567	-0.9812	0.9734
F5	-0.2345	-0.6543	0.5432	F15	0.4567	0.5012	-0.0987
F6	0.3456	0.1234	-0.8765	F16	-0.9789	-0.9321	0.9013
F7	0.4567	0.7891	0.6543	F17	0.6789	0.1234	-0.3210
F8	0.5678	-0.2345	-0.4321	F18	-0.9945	-0.9456	0.9123
F9	0.9534	0.9412	0.9220	F19	0.8901	0.3456	-0.5432
F10	-0.9765	-0.9674	0.9345	F20	-0.9781	-0.9672	0.9454

### 2.3.2. Stochastic Removal of Features

While correlation analysis effectively identifies key features for SOH prediction, it often results in feature redundancy, with several features displaying high correlations or overlapping information. Such redundancy not only complicates the model but also risks overfitting, thereby diminishing its ability to generalize. For example, numerous similar features might arise from analyzing the battery's voltage and current data. If these features are derived in a closely similar manner, they end up supplying redundant information, which contributes repeatedly to the model. In these scenarios, judicious feature selection becomes imperative.

To enhance feature selection, we implemented random feature elimination. This straightforward technique assesses the model's reliance on specific features by randomly removing a subset of features and observing the resultant impact on model performance, thus establishing the importance of each feature. This method involves directly removing certain features from the dataset and subsequently retraining the model with the remaining features to discern any changes in performance.

After the model with the reduced feature set is retrained, its performance is re-evaluated. By comparing the performance of the model using the full feature set and the reduced feature set, the impact of removing specific features on model performance can be assessed. If the model's performance significantly decreases, it indicates that the removed feature set contains information crucial to the model's prediction; if the performance changes little or even improves, it suggests that the original feature set may include some redundant or irrelevant features. Specifically, randomly selecting and removing  $k$  features from feature set  $F$ , where  $k$  is the predetermined number of features to be removed. The remaining feature set is  $F' \subset F$ , where  $|F'|=|F|-k$ . The model is retrained using the remaining feature set  $F'$ .

#### 2.4. Constructing Graph-Structured Data

Battery performance and aging are influenced by a variety of factors, with interactions that are often nonlinear and complex. Graph structures are particularly adept at representing these intricate relationships. They are also flexible, allowing for easy expansion and updates, such as adding new nodes or adjusting edge weights to accommodate new data or changes in environmental conditions.

**Node Representation:** In our graph structure, each node represents a feature extracted from a single cycle of the battery. These features include variables such as current, voltage, capacity, and temperature. Each node  $v_i$  corresponds to a feature  $x_i$  extracted from a single cycle of the battery, i.e.,  $v_i = x_i$ , where  $x_i$  represents the  $i$ th feature extracted in that cycle.

**Edge Representation:** In the graph structure, edges depict the relationships among features. To encapsulate the global characteristics of the battery within a single cycle, we connect all feature nodes, forming a complete graph. This method allows the model to fully utilize the interdependencies among features for accurate SOH estimation.

Formally, the graph is defined  $G = (V, E)$ , where  $V$  is the set of nodes, and each node  $v_i$  represents a feature.  $E$  is the set of edges, and each edge  $e_{ij}$  represents the relationship between features  $x_i$  and  $x_j$ . In this graph representation method, the nodes and edges of the graph can be defined as follows:  $V = \{v_1, v_2, \dots, v_m\}$ ,  $E = \{(v_i, v_j) | 1 \leq i < j \leq m\}$ .

#### Graph Structure Based on Mutual Trustworthiness

To capture the interdependencies between battery features, we employ mutual information to construct the graph structure. Mutual information is a statistical measure of the interdependence between two variables, reflecting the correlation between battery features. We determine the connection relationships and edge weights between feature nodes by calculating the mutual information between these feature pairs. For two features  $x_i$  and  $x_j$ , their mutual information  $I(x_i; x_j)$  is calculated via the following formula:

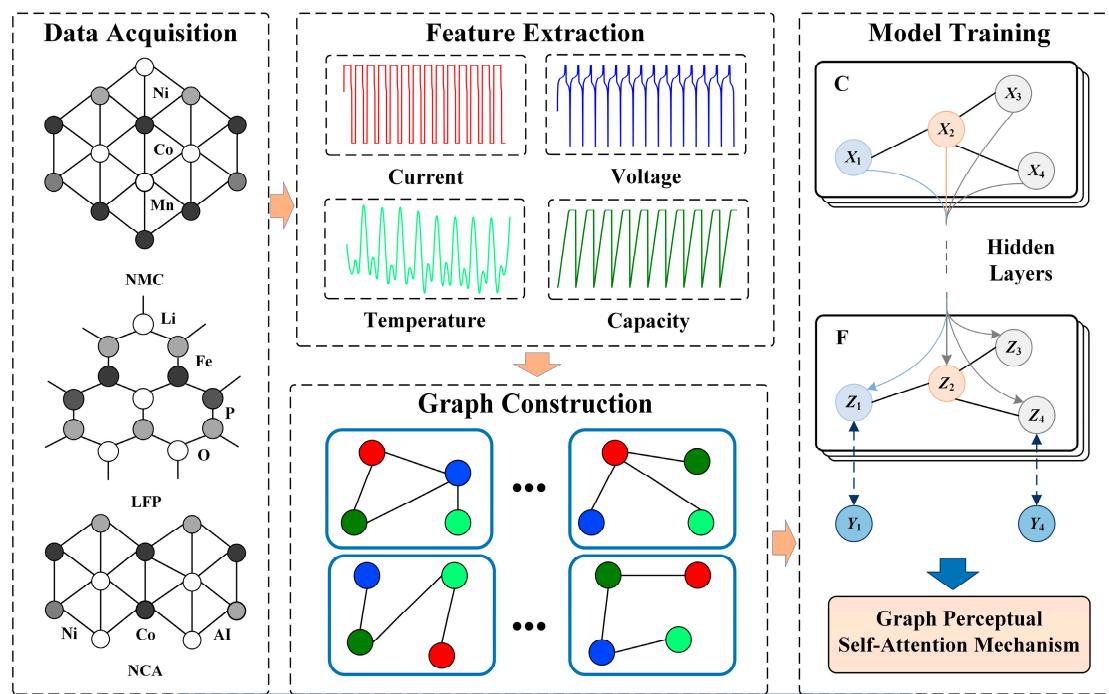
$$I(x_i; x_j) = \sum_{x_i} \sum_{x_j} p(x_i, x_j) \log \left( \frac{p(x_i, x_j)}{p(x_i)p(x_j)} \right) \quad (3)$$

where  $p(x_i, x_j)$  is the joint probability distribution of features  $x_i$  and  $x_j$ , whereas  $p(x_i)$  and  $p(x_j)$  are the marginal probability distributions of features  $x_i$  and  $x_j$ , respectively. Next, on the basis of the magnitude of mutual information, we determine whether to establish edges between feature nodes and set the edge weights. Specifically, for the mutual information matrix  $C$ , retain edges where  $C_{ij} > \tau$ , where  $\tau$  is a preset mutual information threshold, which is usually adjusted according to the specific data conditions.

On the basis of the mutual information matrix, the adjacency matrix  $A$  is constructed, where  $A_{ij} = I(X_i; X_j)\delta y$ . Using the adjacency matrix  $A$ , construct the graph  $G$ , where each node corresponds to a feature.

### 3. Model Architecture

The entire framework process is illustrated in Figure 3. First, data from three different types of batteries are obtained, followed by feature extraction of key parameters such as current, voltage, temperature, and capacity. Then, a graph structure is constructed on the basis of these features. Finally, model training is conducted.



**Figure 3.** Overall architecture of battery SOH estimation.

### 3.1. Graph Convolutional Networks

The graph convolutional network (GCN) is a deep learning model tailored for graph-structured data. It adapts the convolutional operations from traditional convolutional neural networks to suit graph data, efficiently capturing the relationships between nodes and their neighbors. GCNs work by aggregating the features of a node with those of its neighboring nodes via a message-passing mechanism, thus creating updated feature representations for each node. The fundamental formulation of the GCN model is described as follows:

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \quad (4)$$

where  $H^{(l)}$  is the node feature matrix of layer  $l$ ,  $\tilde{A} = A + I$  is the adjacency matrix with added self-connections,  $\tilde{D}$  is the degree matrix of  $\tilde{A}$ ,  $W^{(l)}$  is the weight matrix of layer  $l$ , and  $\sigma$  is the nonlinear activation function.

The GCN, through its convolution operations, aggregates the information of each feature node with that of its neighboring nodes, thus capturing local dependencies between features. For example, when a node represents the current feature of a battery, the GCN considers the neighboring voltage, capacity, and temperature features, integrating this information to generate a richer feature representation.

### 3.2. Self-Attention Mechanism of the Transformer

Moving away from the traditional recurrent neural network (RNN) framework, the transformer model relies solely on an attention mechanism to delineate the relationships among sequence elements. At the heart of this model are self-attention and multi-head attention mechanisms. The self-attention mechanism identifies global dependencies within a sequence by calculating relevance weights between each element and every other element in the input sequence. Specifically, for an input sequence, this mechanism first transforms it into three sets of vectors—query, key, and value—and then computes the attention weights via the formula below:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (5)$$

where  $Q$ ,  $K$ , and  $V$  represent the query, key, and value matrices, respectively, and  $d_k$  is the dimension of the key vector. This formula calculates the similarity between the query and key through an inner product operation, normalizes it via the Softmax function, and finally weights and sums the value matrix to obtain a new representation for each element. The self-attention mechanism can establish direct connections between any two features in the input sequence, regardless of their distance in the sequence. This is particularly useful for capturing the global dependencies of battery features.

**Multi-head attention mechanism:** The transformer enhances its expressive capabilities through the use of a multi-head attention mechanism. This mechanism operates by executing multiple self-attention processes in parallel, each independently focusing on different subspaces. The results from these processes are then concatenated and subjected to a linear transformation:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h)W^O \quad (6)$$

where  $\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$ ,  $QW_i^Q$ ,  $KW_i^K$ ,  $VW_i^V$ , and  $W^O$  are trainable parameter matrices. Through the multi-head attention mechanism, the transformer can compute attention in multiple subspaces in parallel, enriching and diversifying feature representations. This aids in capturing the complex relationships between battery features more comprehensively, thereby improving the accuracy of SOH estimation.

In the estimation of the battery SOH, battery features display both local and long-distance dependencies within the time series. Variations in certain battery features at different time points can exhibit significant correlations, which are essential for precise SOH predictions. The transformer model, which uses self-attention and multi-head attention mechanisms, effectively captures these long-distance dependencies.

### 3.3. Graph Perceptual Self-Attention Mechanism

In SOH estimation, interactions among battery components are characterized by intricate physical and chemical connections. Traditional self-attention mechanisms allocate attention weights primarily on the basis of feature similarity, potentially overlooking the complex structural relationships between components. Integrating graph awareness into the self-attention mechanism enables the model to consider both feature similarity and actual connections between components, leading to a more holistic understanding of the battery's overall state.

The adjacency matrix plays a crucial role in the graph-aware self-attention mechanism. It represents the connection relationships between the nodes in the graph, where each element  $A_{ij}$  of the matrix indicates whether nodes  $i$  and  $j$  are directly connected. Introducing the adjacency matrix into the self-attention mechanism allows the adjustment of attention weights to depend not only on feature similarity but also on the connectivity between nodes.

To incorporate the graph structure information, we adjust the dot product results as follows: Edge feature integration: For each pair of nodes  $i$  and  $j$ , calculate their edge feature vector  $E_{ij}$ . Edge Weight Calculation: Use a trainable parameter  $W$  to linearly transform the edge feature vector  $E_{ij}$  to obtain the edge weight adjustment value  $W \cdot E_{ij}$ . Adjacency Matrix Adjustment: The obtained edge weight adjustment value is multiplied with the adjacency matrix to ensure that only actually connected nodes have their weights adjusted. Finally, the adjusted attention weight formula is:

$$\text{Adjusted Attention}(Q, K, V, A) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}} + A \cdot (W \cdot E)\right)V \quad (7)$$

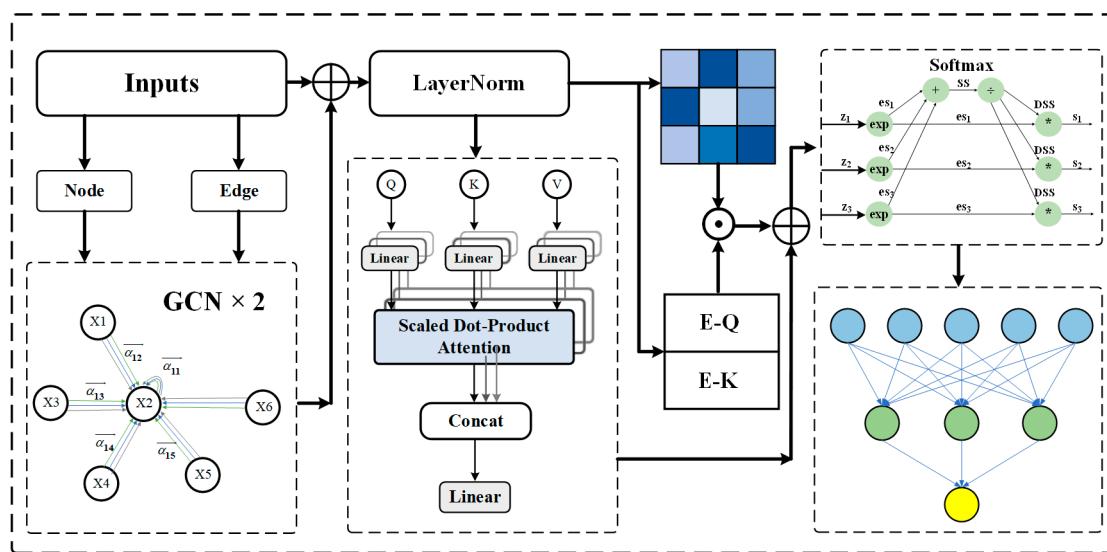
where  $A \cdot (W \cdot E)$  represents the attention scores adjusted by the graph structure.

### 3.4. Fusion of the GCN and Transformer

The transformer's strength lies in capturing long-distance dependencies and handling sequential data, whereas the GCN excels in analyzing the complex topological structure

between battery components. The goal of this fusion strategy is to effectively integrate the deep features of graph-structured data with long-distance dependencies.

Figure 4 depicts the overall framework of the graph perceptual neural network (GPNN). In the initial phase of the model's forward propagation, the input node features and edge indices are processed by the GCN layer. This layer applies graph convolution operations to ensure that each node's updated features encapsulate not only its own information but also that of its immediate neighbors, thus providing enriched feature representations that incorporate local graph structural details. These enhanced features from the GCN are subsequently merged with the original input features via residual connections, with the goal of preserving valuable information from the original data while integrating contextual details provided by the graph structure.



**Figure 4.** Graph perceptual neural network.

The fused features are input into a series of graph-aware multi-head self-attention layers, each of which is enhanced beyond the traditional self-attention mechanism. Initially, basic attention scores are calculated within these layers. Subsequently, edge features and the adjacency matrix are incorporated. By merging the query and key vectors of each node pair and processing them through a linear layer, edge features are created. These features, along with the adjacency matrix, modify the initial attention scores, shifting the focus of the attention distribution toward the actual connections within the graph, thus leveraging the graph structure information more effectively.

In the deeper layers of the model, graph-aware self-attention layers and layer normalization modules are applied iteratively. Each iteration integrates and refines the output from the previous layer, enabling the model to harness deeper feature sets and relationships. Layer normalization within each layer ensures numerical stability and mitigates the risk of overfitting during training. After extensive processing and feature fusion through multiple layers, the model consolidates the final node features, typically by averaging them to create a global feature representation. This global feature is subsequently processed through a fully connected layer for final output prediction, which is generally used to assess the battery's health status. With this architecture, the model adeptly processes traditional sequential data while thoroughly extracting and utilizing the complex relationships and dynamic variations found in graph-structured data, thus enhancing prediction accuracy and practical applicability.

## 4. Experimental Results and Discussion

### 4.1. Experimental Preparation

To comprehensively evaluate the performance of our developed predictive model, we selected three statistical indicators: the root mean square error (RMSE), maximum error (MAXE), and mean absolute percentage error (MAPE). These metrics help us analyze the prediction accuracy and reliability of the model from different perspectives.

The RMSE is a very important performance metric and is especially suitable for situations where larger prediction errors need to be emphasized. In battery SOH estimation, larger deviations can lead to safety risks or poor performance. Therefore, the RMSE allows us to capture these extreme error cases effectively and optimize them accordingly. The RMSE calculation formula is as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (8)$$

MAXE directly reflects the model's performance under the most adverse conditions, which is crucial for battery management systems. The stability and reliability of battery systems require the predictive model to maintain high accuracy under all possible operating conditions. MAXE provides a direct indicator to assess this capability. The MAXE calculation formula is as follows:

$$MAXE = \max(|f(x)_t - y_t|) \quad (9)$$

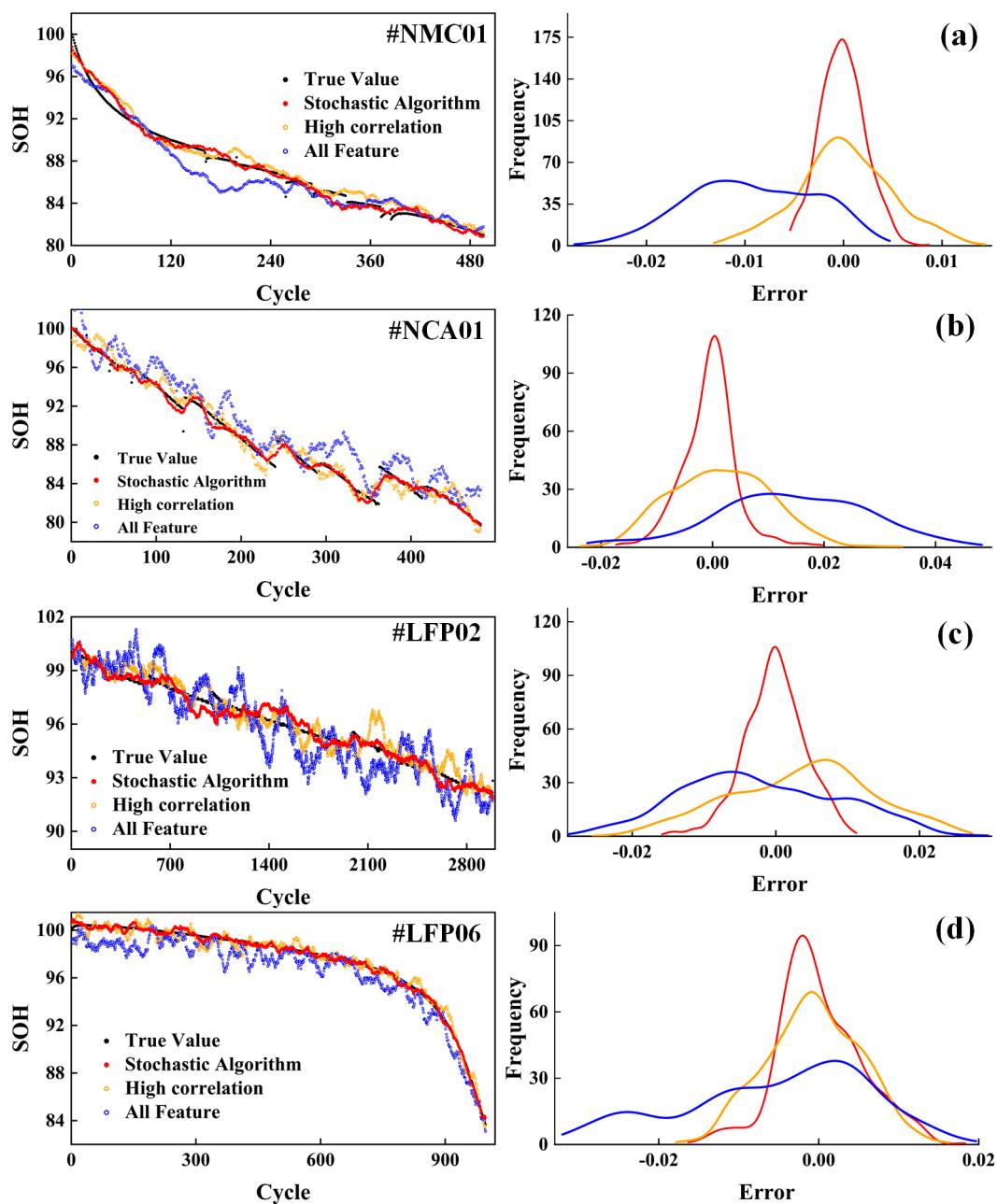
The MAPE provides a perspective for evaluating the relative size of model errors, which is particularly useful for comparing battery performance degradation models under different technologies or operating conditions. Since the MAPE is percentage-based, it allows us to compare different types and capacities of batteries, making it an ideal metric for evaluating battery SOH. The formula is as follows:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{f(x)_t - y_i}{\bar{y}} \right| \quad (10)$$

The symbols mentioned above are as follows:  $y_t$  and  $y_i$  are the actual values,  $f(x)_t$  and  $\hat{y}_i$  are the predicted values,  $\bar{y}$  is the average of the actual values, and  $n$  is the number of samples.

### 4.2. Characterization Experiments

To address potential feature redundancy, additional feature analysis experiments were conducted. In these experiments, a series of refined steps were taken to optimize the model's input features. Initially, we conducted an extensive correlation analysis using the Pearson correlation coefficient to quantify the linear relationships between features. By applying a predetermined correlation threshold, we identified and eliminated highly correlated features, reducing redundant data through a process known as feature reduction. This reduction not only simplifies the model's complexity but also mitigates stability issues from multicollinearity among features. To further refine the feature set, we implemented random feature elimination, removing certain features entirely to evaluate the impact on the model's performance. The removal proportions were set at 10%, 20%, and complete, with each proportion tested through multiple experimental iterations and feature combinations as part of a hyperparameter search. Figure 5 depicts the SOH estimation results and corresponding error probability density curves for different feature selections, whereas Table 3 presents the evaluation metric results.



**Figure 5.** SOH estimation result curves (left) and error probability density curves (right) for different feature selections: (a) NMC01; (b) NCA01; (c) LFP02; (d) LFP06.

In the feature selection experiments, significant differences emerged among the random algorithm, high-correlation feature selection, and all-feature methods. The random algorithm consistently showed a lower mean absolute error (MAE) across all batteries, notably for LFP02 and LFP06, with values of 0.384 and 0.366, respectively. The root mean square error (RMSE) was relatively stable, ranging from 0.609 to −0.692, and the maximum error values remained low, particularly for LFP06 at only 0.968, demonstrating the algorithm's effectiveness in reducing errors across battery types. In contrast, the high-correlation feature selection method typically resulted in a higher MAE, especially for the NMC01 and NCA01 batteries, at 0.628 and 0.709, respectively, and less favorable RMSE values between 0.774 and 0.853. The maximum error values were considerably higher, especially for NCA01, which reached 3.437, indicating poor performance in managing extreme errors. This method also suffers from feature redundancy, where multiple features

mirror the same information. The all-feature method showed a high MAE for all batteries, particularly reaching 1.84 for NCA01, with the RMSE also performing poorly, ranging from 0.821 to –2.01. The maximum error values were highest, particularly for NCA01 at 4.829, suggesting that using all features introduces excessive noise, increasing model complexity and compromising stability. Error probability density curve analysis revealed that the random algorithm (red curve) displayed the most concentrated error distribution across all battery types, indicating greater stability and effective reduction of errors from feature redundancy. The high-correlation feature selection (orange curve) results in a wider error distribution, especially for NCA01, which fails to effectively address feature redundancy. The all-feature method (blue curve) exhibited the broadest error distribution, particularly in the NMC01 and NCA01 batteries, highlighting the adverse effects of feature redundancy on model stability. By randomly eliminating some features, the random algorithm mitigates feature redundancy issues, avoiding the impacts seen with high-correlation and all-feature methods and thus enhancing model performance.

**Table 3.** Estimation results for different feature selection methods.

Feature Select	Evaluation Indicators	Battery Number			
		NMC01	NCA01	LFP02	NMC01
Stochastic algorithm	MAE	0.455	0.573	0.384	0.366
	RMSE	0.609	0.692	0.621	0.592
	MAXE	1.829	1.967	1.120	0.968
High correlation	MAE	0.628	0.709	0.674	0.479
	RMSE	0.774	0.853	0.821	0.692
	MAXE	2.211	3.437	2.249	1.431
All Feature	MAE	0.853	1.84	0.774	0.879
	RMSE	0.824	2.01	0.821	0.992
	MAXE	2.893	4.829	2.826	2.728

Different battery materials have unique characteristics and stability profiles. In feature selection experiments, the NMC01 battery showed heightened sensitivity to feature redundancy, with relatively high MAE and RMSE values under all feature selection methods. This suggests that NMC batteries are particularly prone to issues with redundant features and multicollinearity during model training, potentially leading to greater model complexity and instability in performance. Conversely, the LFP02 and LFP06 batteries consistently exhibited stable performance across the feature selection experiments. Notably, the LFP02 battery maintained low error values across all feature selection methods.

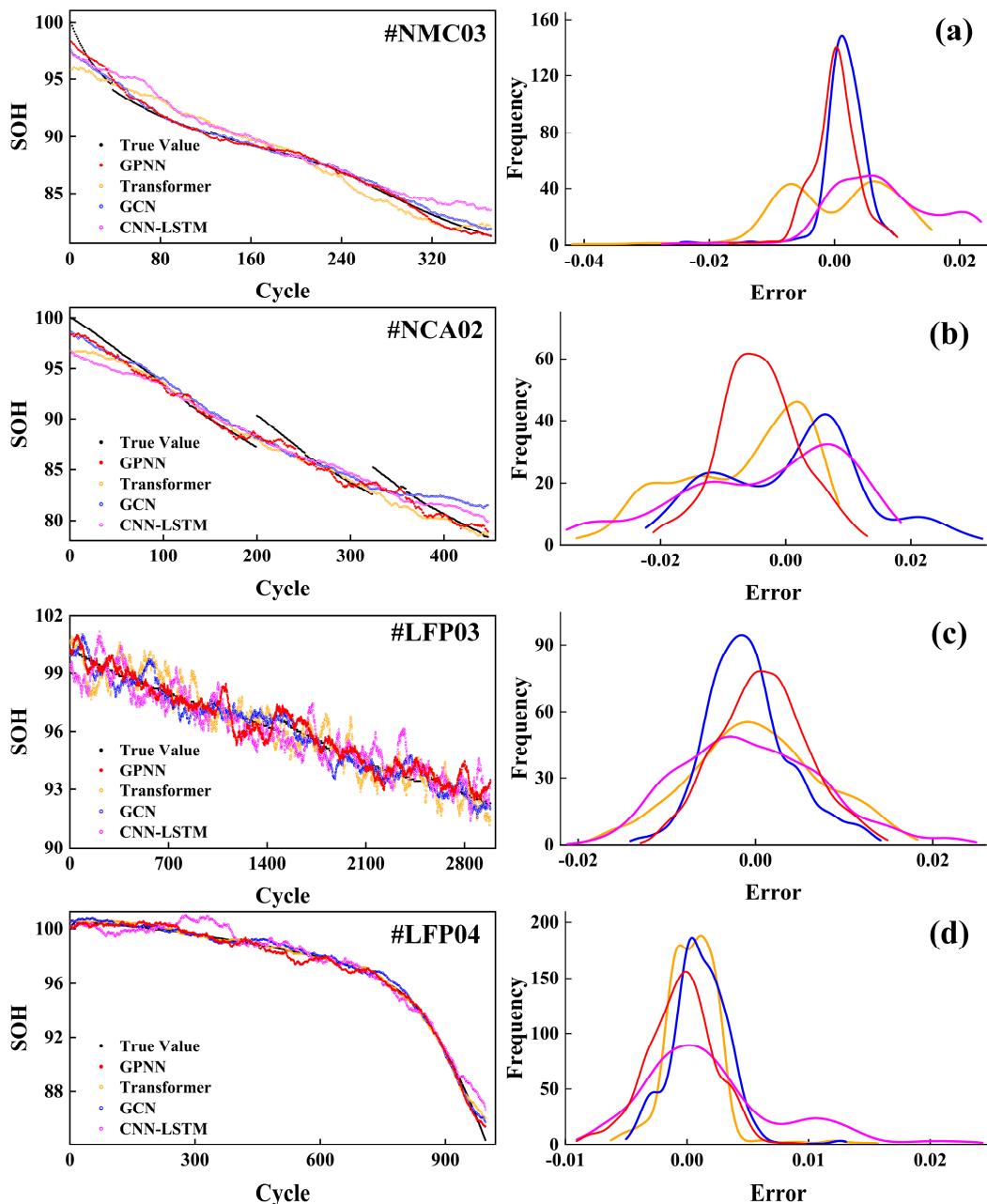
In summary, different types of battery materials demonstrate varied levels of stability in feature selection experiments. NMC batteries are particularly sensitive to redundant features, which can destabilize the model. However, using a random algorithm can effectively enhance model stability. NCA batteries show the highest sensitivity to feature selection, with the all-feature method introducing significant redundant information that adversely affects model performance. Conversely, the random algorithm considerably reduces errors and enhances stability. LFP batteries maintain relatively stable performance during feature selection and are less affected by the process. Notably, the random algorithm yields the best performance for these batteries. Therefore, in practical applications, it is crucial to employ suitable feature selection methods tailored to different battery materials to improve model prediction stability and overall performance.

#### 4.3. Comparative Experiments of Different Models

In recent studies on lithium battery SOH estimation, several popular and effective modeling approaches have been explored. These include the transformer model discussed in reference [31], the temporal convolutional network (TCN) mentioned in reference [26], as

well as other models such as combined convolutional neural network and long short-term memory network (CNN-LSTM) and graph convolutional networks (GCN).

In this study, we evaluated the performance of four distinct models for estimating SOH of batteries across various types. The models assessed include the graph perceptual neural network (GPNN), transformer, GCN, and CNN-LSTM. Each of these models was trained via the optimal feature sets derived from our feature selection experiments. Figure 6 displays the SOH estimation results and the corresponding error probability density curves for each model. Table 4 presents the outcomes of the evaluation metrics.



**Figure 6.** SOH estimation result curves (**left**) and error probability density curves (**right**) for different models: (a) NMC03; (b) NCA02; (c) LFP03; (d) LFP04.

**Table 4.** Estimation results for different models.

Model	Evaluation Indicators	Battery Number			
		NMC03	NCA02	LFP03	LFP08
GPNN	MAE	0.473	0.570	0.406	0.357
	RMSE	0.594	0.609	0.595	0.560
	MAXE	1.785	2.017	1.374	0.941
Transformer	MAE	3.135	2.340	0.640	0.565
	RMSE	2.679	1.633	0.728	0.640
	MAXE	4.192	3.340	1.972	1.716
GCN	MAE	0.960	2.374	0.538	0.463
	RMSE	0.862	1.672	0.573	0.484
	MAXE	2.386	3.141	1.411	1.384
CNN-LSTM	MAE	2.430	1.587	1.930	0.682
	RMSE	1.558	1.259	1.385	0.782
	MAXE	2.865	3.500	2.489	2.429

GPNN uses graph structure information to capture relationships between nodes through graph convolution layers, effectively handling data with complex dependencies. Across various battery types, the GPNN maintained stable MAE and RMSE values, with outstanding performance on the LFP08 battery, which registered an MAE of 0.357 and an RMSE of 0.560. This demonstrates the ability of the GPNN to capture feature interactions effectively and process graph-structured data. Overall, the GPNN ensures high prediction accuracy and stability across diverse battery types.

The transformer model, through its self-attention mechanism, can effectively capture global dependencies, making it suitable for handling sequential data and complex dependencies. However, in battery SOH prediction, transformers may face challenges in adequately capturing local feature relationships. The transformer performed poorly on the NMC03 and NCA02 batteries, with MAEs of 3.135 and 2.340, respectively, indicating difficulties in handling these types of battery features. The performance improved on the LFP03 and LFP08 batteries, but the overall results were still inferior to those of the GPNN.

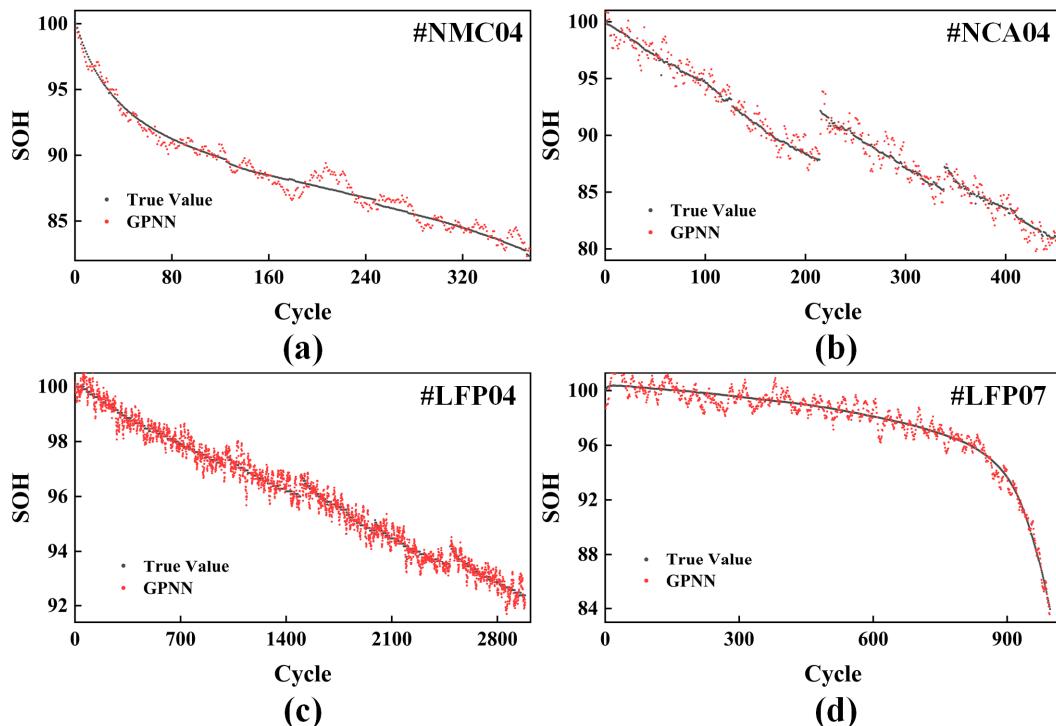
The GCN processes graph-structured data through graph convolution layers, effectively capturing local and global information between nodes. However, the performance of the GCN in battery SOH prediction is limited by its number of layers and the over-smoothing problem. The GCN performed well on the LFP03 and LFP08 batteries, with MAEs of 0.538 and 0.463, respectively, but performed poorly on the NMC03 and NCA02 batteries, especially on the NCA02 battery, with an MAE as high as 2.374. This finding indicates that the GCN has limited adaptability to certain battery types, with an overall performance inferior to that of the GPNN. CNN-LSTM combines the feature extraction capability of convolutional neural networks with the sequence data processing ability of LSTM, making it suitable for handling battery data with significant temporal features. However, its complex structure is prone to overfitting, and it is less effective than the GPNN and GCN in capturing certain feature relationships. CNN-LSTM performed well on the LFP08 battery, with an MAE of 0.682, but performed poorly on the other battery types, particularly NMC03 and LFP03, with MAEs of 2.430 and 1.930, respectively. This indicates limitations in handling these types of battery features, with an overall performance inferior to that of the GPNN.

#### 4.4. Noise Experiments

To further validate the robustness of the models in battery SOH prediction, we added 100 dB Gaussian noise to the feature data of four batteries (NMC03, NCA02, LFP03, LFP08) to simulate potential noise interference encountered in real-world scenarios. The noise formula is as follows:

$$X_{\text{noisy}} = X + \mathcal{N}(0, \sigma^2) \quad (11)$$

where  $\sigma$  is determined by a noise level of 100 dB. Figure 7 displays the curves of the SOH estimation results for different types of batteries. Table 5 presents the results of the evaluation metrics.



**Figure 7.** SOH estimation curves for different types of batteries: (a) NMC04; (b) NCA04; (c) LFP04; (d) LFP07.

**Table 5.** Estimation results for different noises.

Evaluation Indicators	Battery Number			
	NMC03	NCA02	NMC03	LFP08
MAE	0.532	0.635	0.652	0.410
RMSE	0.630	0.643	0.730	0.590
MAXE	1.996	2.193	1.905	1.600

The experimental results show that despite the addition of high noise, the prediction error of the GPNN model increased only slightly across all battery types, with the best performance for the LFP08 battery, where the MAE and RMSE were 0.390 and 0.590, respectively. This indicates that the GPNN can effectively capture complex relationships between battery features and has good noise resistance. Overall, the performance of the GPNN model in high-noise environments further demonstrates its broad applicability and reliability in battery SOH prediction tasks. For both the NMC03, NCA02, LFP03, and LFP08 batteries, the GPNN can maintain high prediction accuracy and stability, showing its superior performance in handling noise interference. These results provide strong support for the practical application of the GPNN in battery SOH prediction, ensuring accurate and reliable predictions even in complex and noisy real-world environments.

## 5. Conclusions

This study proposes a battery state of health (SOH) estimation method based on a graph perception neural network (GPNN), aiming to address the issues of accuracy and robustness in current SOH estimation models when dealing with different types of batteries. By extracting 20 features from current, voltage, voltage–capacity, and temperature data and

constructing a graph structure based on mutual information, the GPNN model can more comprehensively capture the complex relationships between battery features. The graph-aware self-attention mechanism that combines the adjacency matrix and edge features into the self-attention calculation enhances the model's understanding of battery behavior.

Experiments show that the GPNN model mitigates the issue of feature redundancy through random feature removal, significantly improving prediction accuracy and stability. In SOH estimation for batteries of different material types, the GPNN model has demonstrated its wide applicability and high performance, confirming its significant value in practical applications. For NMC, NCA, and LFP batteries, the GPNN can provide stable and accurate SOH predictions. This is crucial for battery management in electric vehicles and renewable energy storage systems, helping to optimize battery management systems (BMS), extend battery life, and ensure safe operation of equipment.

**Author Contributions:** K.C.: conceptualization, methodology, validation, writing—review and editing, supervision, project administration, funding acquisition. D.W.: methodology, validation, writing—original draft, investigation, data curation. W.G.: investigation, visualization, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The relevant experiments of this study are still in progress. If you need our data for relevant research, you can contact the corresponding author of this article.

**Conflicts of Interest:** The authors declare that they have no known competing financial interests or personal relationships that may appear to influence the work reported in this paper.

## References

- Oji, T.; Zhou, Y.; Ci, S.; Kang, F.; Chen, X.; Liu, X. Data-driven methods for battery SOH estimation: Survey and a critical analysis. *IEEE Access* **2021**, *9*, 126903–126916. [[CrossRef](#)]
- Vignesh, S.; Che, H.S.; Selvaraj, J.; Tey, K.S.; Lee, J.W.; Shareef, H.; Errouissi, R. State of Health (SoH) estimation methods for second life lithium-ion battery—Review and challenges. *Appl. Energy* **2024**, *369*, 123542.
- Das, K.; Kumar, R.; Krishna, A. Analyzing electric vehicle battery health performance using supervised machine learning. *Renew. Sustain. Energy Rev.* **2024**, *189*, 113967. [[CrossRef](#)]
- Huang, H.; Bian, C.; Wu, M.; An, D.; Yang, S. A novel integrated SOC–SOH estimation framework for whole-life-cycle lithium-ion batteries. *Energy* **2024**, *288*, 129801. [[CrossRef](#)]
- Li, X.; Chen, X.; Bai, Q.; Mo, Y.; Zhu, Y. From atomistic modeling to materials design: Computation-driven material development in lithium-ion batteries. *Sci. China Chem.* **2024**, *67*, 276–290. [[CrossRef](#)]
- Fahmy, H.M.; Hasanien, H.M.; Alsaleh, I.; Ji, H.; Alassaf, A. State of health estimation of lithium-ion battery using dual adaptive unscented Kalman filter and Coulomb counting approach. *J. Energy Storage* **2024**, *88*, 111557. [[CrossRef](#)]
- Peng, S.; Miao, Y.; Xiong, R.; Bai, J.; Cheng, M.; Pecht, M. State of charge estimation for a parallel battery pack jointly by fuzzy-PI model regulator and adaptive unscented Kalman filter. *Appl. Energy* **2024**, *360*, 122807. [[CrossRef](#)]
- Chen, M.; Ma, G.; Liu, W.; Zeng, N.; Luo, X. An overview of data-driven battery health estimation technology for battery management system. *Neurocomputing* **2023**, *532*, 152–169. [[CrossRef](#)]
- Sang, B.; Wu, Z.; Yang, B.; Wei, J.; Wan, Y. Joint Estimation of SOC and SOH for Lithium-Ion Batteries Based on Dual Adaptive Central Difference H-Infinity Filter. *Energies* **2024**, *17*, 1640. [[CrossRef](#)]
- Li, Z.; Shen, S.; Zhou, Z.; Cai, Z.; Gu, W.; Zhang, F. Novel method for modelling and adaptive estimation for SOC and SOH of lithium-ion batteries. *J. Energy Storage* **2023**, *62*, 106927. [[CrossRef](#)]
- Zhao, J.; Feng, X.; Pang, Q.; Wang, J.; Lian, Y.; Ouyang, M.; Burke, A.F. Battery prognostics and health management from a machine learning perspective. *J. Power Sources* **2023**, *581*, 233474. [[CrossRef](#)]
- Xu, H.; Wu, L.; Xiong, S.; Li, W.; Garg, A.; Gao, L. An improved CNN-LSTM model-based state-of-health estimation approach for lithium-ion batteries. *Energy* **2023**, *276*, 127585. [[CrossRef](#)]
- Qian, C.; Xu, B.; Xia, Q.; Ren, Y.; Sun, B.; Wang, Z. SOH prediction for Lithium-Ion batteries by using historical state and future load information with an AM-seq2seq model. *Appl. Energy* **2023**, *336*, 120793. [[CrossRef](#)]
- Peng, S.; Sun, Y.; Liu, D.; Yu, Q.; Kan, J.; Pecht, M. State of health estimation of lithium-ion batteries based on multi-health features extraction and improved long short-term memory neural network. *Energy* **2023**, *282*, 128956. [[CrossRef](#)]
- Lyu, G.; Zhang, H.; Miao, Q. An Adaptive and Interpretable SOH Estimation Method for Lithium-ion Batteries based-on Relaxation Voltage Cross-scale Features and Multi-LSTM-RFR2. *Energy* **2024**, *304*, 132167. [[CrossRef](#)]

16. He, Y.; Deng, Z.; Chen, J.; Li, W.; Zhou, J.; Xiang, F.; Hu, X. State-of-health estimation for fast-charging lithium-ion batteries based on a short charge curve using graph convolutional and long short-term memory networks. *J. Energy Chem.* **2024**, *98*, 1–11. [[CrossRef](#)]
17. Wei, Y.; Wu, D. State of health and remaining useful life prediction of lithium-ion batteries with conditional graph convolutional network. *Expert Syst. Appl.* **2024**, *238*, 122041. [[CrossRef](#)]
18. Vatter, J.; Mayer, R.; Jacobsen, H. The evolution of distributed systems for graph neural networks and their origin in graph processing and deep learning: A survey. *ACM Comput. Surv.* **2023**, *56*, 6. [[CrossRef](#)]
19. Xiong, R.; Sun, Y.; Wang, C.; Tian, J.; Chen, X.; Li, H.; Zhang, Q. A data-driven method for extracting aging features to accurately predict the battery health. *Energy Storage Mater.* **2023**, *57*, 460–470. [[CrossRef](#)]
20. Preger, Y.; Barkholtz, H.M.; Fresquez, A.; Campbell, D.L.; Juba, B.W.; Romà-Kustas, J.; Ferreira, S.R.; Chalamala, B.R. Degradation of commercial lithium-ion cells as a function of chemistry and cycling conditions. *J. Electrochem. Soc.* **2020**, *167*, 120532. [[CrossRef](#)]
21. Severson, K.A.; Attia, P.M.; Jin, N.; Perkins, N.; Jiang, B.; Yang, Z.; Chen, M.H.; Aykol, M.; Herring, P.K.; Fragedakis, D.; et al. Data-driven prediction of battery cycle life before capacity degradation. *Nat. Energy* **2019**, *4*, 383–391. [[CrossRef](#)]
22. Nájera, J.; Arribas, J.R.; de Castro, R.M.; Núñez, C. Semi-empirical ageing model for LFP and NMC Li-ion battery chemistries. *J. Energy Storage* **2023**, *72*, 108016. [[CrossRef](#)]
23. Ubaldi, S.; Conti, M.; Marra, F.; Russo, P. Identification of key events and emissions during thermal abuse testing on NCA 18650 cells. *Energies* **2023**, *16*, 3250. [[CrossRef](#)]
24. Demirci, O.; Taskin, S.; Schaltz, E.; Demirci, B.A. Review of battery state estimation methods for electric vehicles-Part II: SOH estimation. *J. Energy Storage* **2024**, *96*, 112703. [[CrossRef](#)]
25. Dai, H.; Wang, J.; Huang, Y.; Lai, Y.; Zhu, L. Lightweight state-of-health estimation of lithium-ion batteries based on statistical feature optimization. *Renew. Energy* **2024**, *222*, 119907. [[CrossRef](#)]
26. Liu, S.; Chen, Z.; Yuan, L.; Xu, Z.; Jin, L.; Zhang, C. State of health estimation of lithium-ion batteries based on multi-feature extraction and temporal convolutional network. *J. Energy Storage* **2024**, *75*, 109658. [[CrossRef](#)]
27. Sun, J.; Fan, C.; Yan, H. SOH estimation of lithium-ion batteries based on multi-feature deep fusion and XGBoost. *Energy* **2024**, *306*, 132429. [[CrossRef](#)]
28. Wang, Z.; Zhao, X.; Zhen, D.; Pombo, J.; Yang, W.; Gu, F.; Ball, A. Adaptable capacity estimation of lithium-ion battery based on short-duration random constant-current charging voltages and convolutional neural networks. *Energy* **2024**, *306*, 132541. [[CrossRef](#)]
29. Zhao, H.; Chen, Z.; Shu, X.; Xiao, R.; Shen, J.; Liu, Y.; Liu, Y. Online surface temperature prediction and abnormal diagnosis of lithium-ion batteries based on hybrid neural network and fault threshold optimization. *Reliab. Eng. Syst. Saf.* **2024**, *243*, 109798. [[CrossRef](#)]
30. Feng, X.; Zhang, Y.; Xiong, R.; Wang, C. Comprehensive performance comparison among different types of features in data-driven battery state of health estimation. *Appl. Energy* **2024**, *369*, 123555. [[CrossRef](#)]
31. Madani, S.S.; Ziebert, C.; Vahdatkhah, P.; Sadrnezhaad, S.K. Recent Progress of Deep Learning Methods for Health Monitoring of Lithium-Ion Batteries. *Batteries* **2024**, *10*, 204. [[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.