



State-of-Health Estimation for Sustainable Electric Vehicle Batteries Using Temporal-Enhanced Self-Attention Graph Neural Networks

Yixin Zhao

Mem. ASME
Environmental Engineering Sciences,
University of Florida,
Gainesville, FL 32611
e-mail: yixin.zhao@ufl.edu

Sara Behdad¹

Associate Professor
Environmental Engineering Sciences,
University of Florida,
Gainesville, FL 32611
e-mail: sarabehdad@ufl.edu

Electric vehicles (EVs) have emerged as an environmentally friendly alternative to conventional fuel vehicles. Lithium-ion batteries are the major energy source for EVs, but they degrade under dynamic operating conditions. Accurate estimation of battery state of health is important for sustainability as it quantifies battery condition, influences reuse possibilities, and helps alleviate capacity degradation, which finally impacts battery lifespan and energy efficiency. In this paper, a self-attention graph neural network combined with long short-term memory (LSTM) is introduced by focusing on using temporal and spatial dependencies in battery data. The LSTM layer utilizes a sliding window to extract temporal dependencies in the battery health factors. Two different approaches to the graph construction layer are subsequently developed: health factor-based and window-based graphs. Each approach emphasizes the interconnections between individual health factors and exploits temporal features in a deeper way, respectively. The self-attention mechanism is used to compute the adjacent weight matrix, which measures the strength of interactions between nodes in the graph. The impact of the two graph structures on the model performance is discussed. The model accuracy and computational cost of the proposed model are compared with the individual LSTM and gated recurrent unit (GRU) models.

[DOI: 10.1115/1.4065146]

Keywords: battery degradation, product lifecycle modeling, state of health (SOH), energy storage systems

1 Introduction

The development of electric vehicles (EVs) provides an environmentally friendly alternative to conventional fuel vehicles. While the shift to electrification in the automotive industry contributes to sustainable transportation, it presents a new set of challenges. Specifically, the market demand for more efficient, durable, and faster-charging batteries is increasing. Challenges related to battery lifespan, energy density, and safe operation under various conditions are currently hot topics in this area [1].

Lithium-ion batteries have become the energy source for many EVs due to their high energy density, long cycle life, and relatively low self-discharge [2]. However, these batteries are susceptible to degradation due to factors such as driving patterns, temperature fluctuations, and rates of charging and discharging [3]. As the battery goes through repeated charge–discharge cycles, its capacity

gradually decreases which leads to reduced performance, shorter lifespan, and increased energy consumption [4]. Each component of a lithium-ion battery contributes differently to the complicated degradation mechanism. The involved processes include solid electrolyte interphase formation, lithium plating, and elevated impedance [5]. In addition, operations such as overcharging, deep discharge, and exposure to extreme temperatures accelerate the aging process.

The related research on improving the performance, safety, and sustainability of lithium-ion batteries can be categorized into three groups based on the lifecycle stages: design, in-use, and end-of-life. In the design phase, optimized materials and structure could be developed to improve the energy density, durability, and fast charging capability of batteries [6,7]. On the other hand, modifications to battery chemistry can minimize dependence on rare and expensive materials, such as cobalt, thus lowering manufacturing costs and mitigating the overexploitation of critical resources [8].

During the in-use phase, a battery management system with effective condition monitoring, thermal control, and failure detection ensures the safe and high-performance operation of the EV battery [9,10]. As an illustration, predictive analysis of thermal data could assist in avoiding thermal runaway incident as well as

¹Corresponding author.

Contributed by the Advanced Energy Systems Division of ASME for publication in the JOURNAL OF ENERGY RESOURCES TECHNOLOGY. Manuscript received November 4, 2023; final manuscript received March 14, 2024; published online April 3, 2024. Assoc. Editor: Tatiana Morozuk.

optimizing energy efficiency [11]. Battery usage life can be extended by optimizing battery operation and avoiding detrimental practices such as overcharging and over-aging [12]. Furthermore, these measures facilitate the preservation of a substantial portion of the active material, making it prequalify for secondary use scenarios after the EV battery is retired.

Lithium-ion batteries have two environmentally friendly end-of-use phases: secondary use and recycling. Moving batteries to a second-life option can extend their lifecycle, thus reducing the environmental impact of manufacturing new batteries and even reducing the cost of batteries. If batteries are recycled, most of the valuable components, such as rare metals, will be recovered and reused to make new batteries, thus alleviating the impact of increasing demand for batteries on resource shortages [13]. However, the majority of retired EV batteries still have up to 80% of their original capacity and have the potential to be repurposed as energy storage modules for less demanding applications [14]. In fact, these batteries are expected to continue functioning for more than ten years when begin their second life at 70–80% of their initial capacity [15].

State of health (SOH) is a key parameter for quantifying the condition of a battery in use and evaluating the reuse value of the second-life stages [16]. Accurate estimation of SOH can help ensure that the battery is repaired or upgraded in case of failure, thus prolonging its first-use life and contributing to a circular economy [17]. However, SOH can only be inferred from indirect parameters such as current and voltage [5]. In addition, the dynamic operating conditions and interacting side reactions introduce uncertainty into the degradation trajectory [18]. The existing methods for SOH estimation are categorized into model-based and data-driven methods. Model-based methods can be further divided into equivalent circuit models [19,20], electrochemical models [21,22], and empirical models [23]. The principle of model-based methods is to describe battery behavior using simplified physico-chemical reactions, often requiring calibration and validation against experimental data. Therefore, it is difficult to have models that can comprehensively characterize aging under different operating conditions. The computational complexity of the model and the estimation accuracy are typically positively correlated and need to be balanced [24].

To overcome the limitations of model-based methods, there has been significant development in data-driven battery models. These techniques utilize large datasets that capture a variety of operating parameters and battery responses to train predictive models [25]. They do not require knowledge of the physical and chemical behavior inside the battery. For instance, Beganovic and Söfker [26] used features from acoustic emission measurements to directly estimate battery aging metrics, without taking into account underlying physical mechanisms. In particular, machine learning-based methods can be trained quickly on large dataset, which allows them to be adapted to various battery chemistries and configurations [27].

This paper proposes the temporal enhanced self-attention-based graph neural networks. The model cooperatively utilizes the temporal insight of long short-term memory (LSTM), the capture of spatial relationships by graph neural network (GNN), and the focus on important features by the attention mechanism to provide an accurate and robust method for battery SOH estimation. The LSTM layer is set before the graph construction layer and is performed as a complex feature engineering step. It captures the temporal dependency of each health factor (HF) from the battery charging data. Two unique principles of graph construction are introduced: individual health factor-based and window-based graphs. Health factor-based graph focuses on the interactions between individual health factors to capture spatial dependencies. In contrast, window-based graph investigates temporal dependencies more deeply. The self-attention mechanism is used to compute the adjacency weight matrix of the graph structure. This matrix quantifies the strength of interactions between nodes, which leads to more efficient weighted information aggregation.

The proposed model starts from the LSTM layer for temporal information extraction to the graph neural network layer for further exploitation of temporal and spatial dependencies, and finally completes the estimation of SOH. The impacts of different graph construction strategies on the model performance are discussed.

The remainder of this paper is organized as follows. Section 2 is the related works on traditional data-driven models, self-attention mechanisms, and graph neural networks for battery state prediction. Section 3 includes the dataset description and feature extraction process. Section 4 describes the proposed SOH prediction framework, including the architecture of the model. Section 5 gives details of the experimental implementation and the SOH prediction results of the models. Finally, Sec. 6 concludes the work in this study and the perspectives for future research.

2 Related Work

2.1 Data-Driven Models for Battery Aging Estimation. Traditional machine learning models such as support vector machine and Gaussian process regression are examples of data-driven machine learning models. To name several studies, Patil et al. [28] employed a multi-stage approach that integrates the classification and regression properties of support vector machines to achieve efficient remaining useful life prediction of EV batteries. Yang et al. [29] introduced a Gaussian process regression model that utilizes specific features of the charging curve as inputs and incorporates it with gray relational analysis to achieve high SOH estimation accuracy.

Besides machine learning models, deep learning models have been used to more closely approximate complicated nonlinear battery systems by training multilayer neural networks. Given the temporal nature of battery aging data, recurrent neural networks, which specialize in handling sequential data, are proposed solutions for battery modeling. Among them, the LSTM and gated recurrent unit (GRU) networks, were designed to address the vanishing gradient problem in traditional recurrent neural networks by introducing gating mechanisms that regulate the flow of information [30]. To name several examples, Kaur et al. [31] implemented feed-forward neural network, convolutional neural network, and LSTM for battery capacity estimation and proved that the LSTM has the highest accuracy among them. Venugopal and Vigneswaran [32] proposed an independent recurrent neural network and compared it with the performance of LSTM and GRU in state-of-health estimation. Their results further demonstrated that the recurrent neural networks that eliminate the gradient problem can learn the long-term dependence between battery capacity degradation well.

2.2 Self-Attention Mechanism: Emphasize the Highly Important Parts of the Data. A considerable number of studies have focused on introducing new trainable parameters and operations into a single model to improve accuracy and generalizability [30,33]. The self-attention mechanism, which is widely recognized for its ability to process sequential data, is an effective improvement for recurrent neural networks [34]. It can help the network allocate resources and extract more vital information during training by assigning different weights to the input features [35]. Qu et al. [36] introduced particle swarm optimization in LSTM to optimize the key parameters and solved the distraction problem by using the attention mechanism to obtain higher accuracy than baseline models. Jiang et al. [37] utilized a convolutional autoencoder to autonomously extract features from battery data and combined it with a self-attention mechanism to achieve accurate SOH estimation. Ge et al. [38] decomposed the battery data with variational mode decomposition to reduce the effect of instability. Then particle filter and the LSTM with self-attention were applied to the decomposed components respectively to improve the accuracy and robustness of the final prediction.

2.3 Graph Neural Network: Extract Spatial Dependencies Between Features. Besides utilizing self-attention to emphasize the relevant part of the data, another direction to improve model performance is the enrichment and augmentation of features by uncovering hidden patterns. Ren et al. [39] used multiple layers of convolutional neural networks before the LSTM to deeply mine the hidden information in the battery data. In addition, an autoencoder was used to expand the dimensionality of the data to match the demand of the convolutional layer implementation. Tian et al. [40] employed an equivalent circuit model to decompose the raw current and voltage data into open-circuit voltage, ohmic response, and polarization voltage. This approach provided the neural network with information about the battery's internal state to improve the performance prediction.

GNN is a technique that effectively extracts and utilizes potential spatial dependencies between features [41]. This method represents the data as a graph structure and allows nodes to pass information to their neighbors through edges to achieve feature fusion. This helps each feature contain more information without increasing the dimensionality of the data. Moreover, if a feature is missing or abnormal, performing information fusion can make the model less perturbed by noise. Yao et al. [42] proposed a GNN using manually extracted battery features as nodes and linear correlation coefficients between features and SOH as edges. In addition, the model combined convolutional neural network and LSTM layers to deepmine the information and achieve accurate SOH prediction. Wei and Wu [43] aggregated the information on battery features using graph convolutional layer, after which the prediction of the target is accomplished using LSTM. Their model outperforms gradient boosting decision tree, single LSTM, and Gaussian process. Wang et al. [44] performed graph construction based on raw battery data, overlaying three GNN layers selected by neural architecture search for feature fusion prior to prediction. Although these works demonstrate the potential of GNNs in battery modeling, the field is still in the beginning stage. The robustness of GNNs to anomalous data is well suited for avoiding interference from noise and uncertainty in the field data, and facilitates real-time estimation of the battery state. However, the limited perspectives of current research tend to neglect the discussion of different methods for constructing graphical data and strategies for aggregating features. In addition, the smoothing problem inherent to GNN and its computational complexity for its application in the battery management are also lacking exploration. Based on the time dependence of battery aging and the need for robustness and computational efficiency, it is necessary to explore GNN architectures that are more applicable to battery modeling.

3 Dataset and Feature Extraction

3.1 Data Description. The data used in this study are sourced from the NASA lithium-ion battery aging dataset [45], with commercial lithium-ion batteries as experimental targets. Specifically, Battery 0005 is analyzed in this study, with its charge–discharge cycles executed under ambient conditions. An initial constant

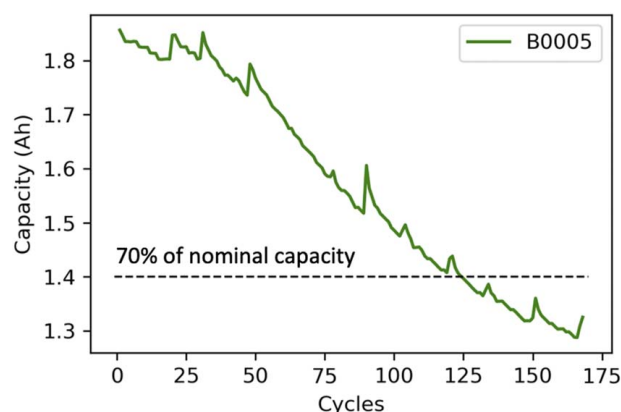


Fig. 1 Capacity degradation of Battery 0005 [47]

current (CC) of 1.5 A was set for battery charging, which was transitioned to a constant voltage (CV) mode once 4.2 V was achieved. This was sustained until a decline to 20 mA in the current was observed [46]. On the other hand, a constant current of 2 A was maintained during discharging until a voltage of 2.7 V was reached. Throughout these procedures, various signals, including current (I), voltage (V), and temperature (T) were continuously monitored. When the capacity of the battery has deteriorated to 70% of its initial state due to repeated charging and discharging, the service life of the battery as well as the experiment is considered to be finished. Figure 1 illustrates the aging trajectory of Battery 0005, which also was used in a previous study by the authors [47].

3.2 Health Factors Extraction. Feature extraction is performed on I, V, and T signals acquired during the charging phase [48]. Figure 2 visualizes the variations in charging profiles as the battery undergoes aging. The SOH is defined by the ratio between the nominal and releasable capacities in this study. Therefore, the key to predicting SOH is to extrapolate the available capacity of the battery. This capacity represents the reversible electric charge accumulation during the charging cycle and subsequent release during discharge [49]. The time required for the constant current phase is related to the charging amount, which is a good response to the aging of the battery.

The constant voltage phase, on the other hand, is a process in which the charging voltage of the battery is constant at its maximum value while the current gradually decreases to a stop [50]. However, the migration of lithium ions is hindered by aging phenomena such as increased impedance and solid electrolyte inter-phase development within the battery [48]. This leads to a growth in the time required for the constant voltage phase. Therefore, analyzing I, V, and T-related features in the constant current and constant voltage phases, respectively, can capture the patterns associated with battery degradation.

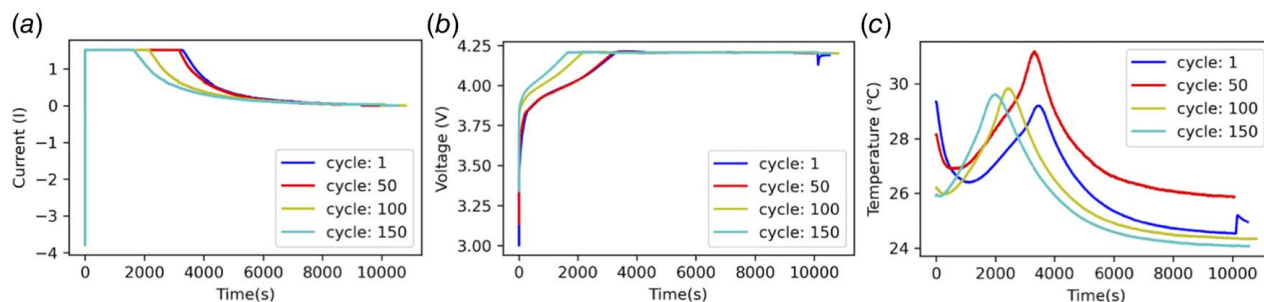


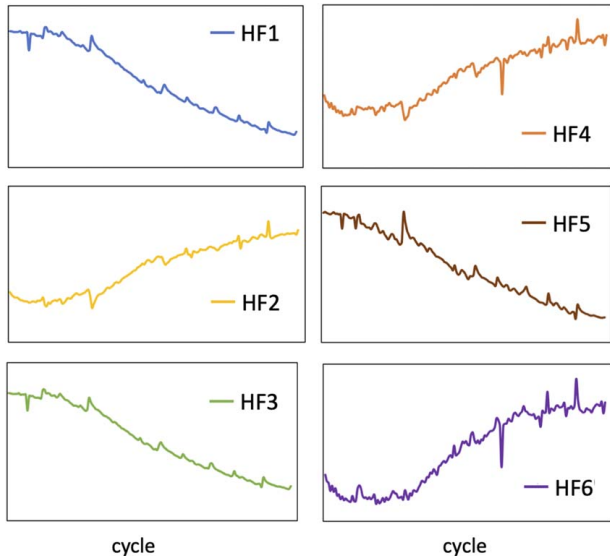
Fig. 2 (a) Current, (b) voltage, and (c) temperature curves during charging of Battery 0005 after different number of cycles [47]

Table 1 Health factors extraction [42,51]

Health factor	Description	Type
HF1	Area covered by current curves of the CC charging	Current-related
HF2	Area covered by current curves of CV charging	
HF3	Area covered by voltage curves of CC charging	Voltage-related
HF4	Area covered by voltage curves of CV charging	
HF5	Area covered by temperature curves of the CC charging	Temperature-related
HF6	Area covered by temperature curves of the CV charging	

Based on the above analysis and references from other similar battery feature extraction and selection methods [42,48,51], a total of six charging health factors are extracted and summarized in Table 1. Examples of the profiles of the extracted features are displayed in Fig. 3. Although the variation of the discharge curve can also indicate the current capacity of the battery, data from the discharge process were not used in this study. The reason for this is that the experimental condition of Battery 0005 is a continuous release at a constant current during the discharge process until the cutoff voltage is reached. According to the calculation method of battery capacity, when the battery is discharged at a constant current, its capacity is given by the discharge current multiplied by the discharge duration [49]. Therefore, for the dataset adopted in this study, the features related to the discharge duration would be nearly linearly correlated with the battery SOH. It is not reasonable to use these discharge features to examine the effectiveness of the developed time-series prediction model. In addition, it is not common in practice to allow an EV battery to complete a full discharge process.

After the feature extraction, the system starts with six time-series health factors. The x -axis of the time-series data represents the battery cycle, while the SOH provides as ground truth for each cycle. The health factors and target SOH are split into training and test sets in a time-series aware manner, with the first 70% of the sequences used for training, and the remaining 30% for testing.

**Fig. 3 Profiles of health factors [47]**

4 Methodology

In this study, a temporal enhanced self-attention GNN is proposed for predicting the SOH of lithium-ion batteries (Fig. 4). The methodology leverages the power of conventional GNN and LSTM network to handle complex time-series health factors extracted from the battery charging data. The self-attention mechanism is specifically employed to learn latent correlations between multiple features that serve as nodes in our graph.

4.1 LSTM Layer for Temporal Dependencies Extraction.

The first component of the proposed model is an LSTM layer. The LSTM layer is set before the graph construction layer as a complex feature engineering step. This layer processes the input time-series health factors data using a sliding window, as shown in Fig. 5, which lets the LSTM capture the temporal dependencies within each HF. The extraction process is executed as follows:

The feature vector X at time-step t can be denoted as Eq. (1)

$$X_t = \{HF_1(t), HF_2(t), HF_3(t), HF_4(t), HF_5(t), HF_6(t)\} \quad (1)$$

The time-step t corresponds to the battery cycle in the time-series data. The sliding window of LSTM layer is defined as SW with a window size of 10. Each input sequence for the LSTM is then a window SW_t at each time-step t containing the battery features over the previous nine cycles and the current cycle, where

$$SW_t = \{X_{t-9}, X_{t-8}, X_{t-7}, \dots, X_t\} \quad (2)$$

The LSTM layer takes the window SW_t as input and processes it to produce temporal features. The LSTM has hidden states h_t and cell states c_t that change over time as the network processes each input time-step. The LSTM recurrence equations at each time-step t are given in Eqs. (3)–(8) [52,53]

$$i_t = \sigma(W_i \cdot [h_{t-1}, SW_t] + b_i) \quad (3)$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, SW_t] + b_f) \quad (4)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, SW_t] + b_o) \quad (5)$$

$$g_t = \tanh(W_g \cdot [h_{t-1}, SW_t] + b_g) \quad (6)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot g_t \quad (7)$$

$$h_t = o_t \cdot \tanh(h(c_t)) \quad (8)$$

where i_t , f_t , o_t , g_t are input gate, forget gate, output gate, and candidate cell state, respectively; σ is the sigmoid activation function; $\tanh$ is the hyperbolic tangent activation function; W_i , W_f , W_o , W_g are the weight matrices; and b_i , b_f , b_o , b_g are the bias vectors associated with the respective gates.

The LSTM's hidden state h_t acts as a form of memory, and captures information from the earlier data points in the window to aid in the prediction of the target. The LSTM updates its hidden state h_t for each time-step t in the input sequence sequentially. After processing the entire sequence, the final hidden state is then used as the extracted temporal feature vector for the following analysis.

4.2 Graph Neural Network Layer for Temporal and Spatial Dependencies Extraction

4.2.1 Graph Construction. In order to implement GNN, the inputs of this layer need to be converted into a representation of the graph structure. Graph-structured data are a mathematical representation consisting of nodes and edges, where nodes represent entities and edges describe connections or relationships between those entities [42]. A fully connected graph is assumed for this dataset, which means each node is connected to every other feature node except itself. The set of nodes V and edges E is defined as follows:

$$V = \{v_1, v_2, \dots, v_n\} \quad (9)$$

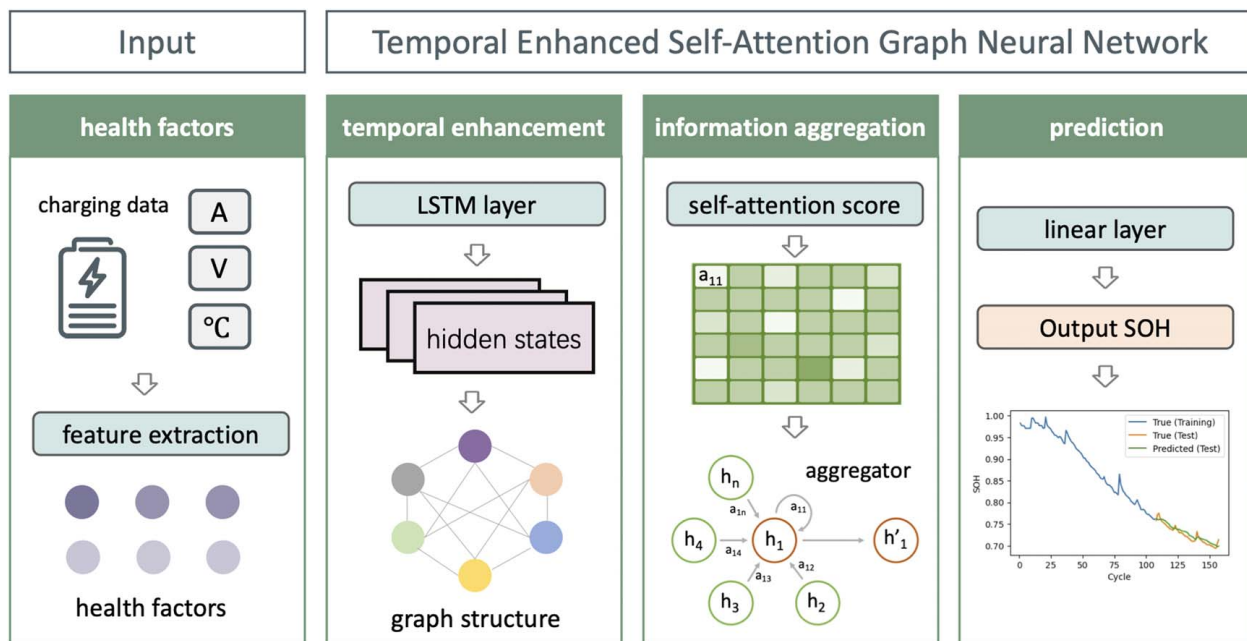


Fig. 4 Framework of the proposed SOH estimation model

$$E = \{(i, j) \mid i, j \in V, i \neq j\} \quad (10)$$

where n is the number of nodes in the graph, and each pair (i, j) represents an edge connecting node i and node j .

Two different approaches are taken to construct graph structures based on the output of the LSTM layer: health factor-based graph and window-based graph (Table 2). In the first approach, the six health factors are processed separately using LSTM layer to contain a deeper level of temporal and spatial information. This produces six final hidden states, each represents information derived from a specific health factor, which is defined as temporally enhanced factors in this paper. These six temporally enhanced factors are then used as individual nodes in the constructed graph. In the second approach, the six health factors are input together into the LSTM for processing, producing an output with a shape of: (window size, number of health factors). To construct the window-based graph, each row of the LSTM output is treated as a node's representation. The graph is built with the number of nodes equal to the window size of the LSTM layer. The nodes contain information about the time-step t and its surrounding battery cycles, respectively.

These two graphs provide different perspectives for discovering relationships within the data. Individual feature-based graphs emphasize the connections between health factors. Thereby, the interactions between factors and their contribution to the overall system behavior can be explored. With this approach, the model has the ability to capture dependencies between series rather than just within individual time series. In contrast, window-based graphs focus on further exploiting the temporal features present in the data. By utilizing the LSTM output in window format, each node has information from a specific temporal context. This approach allows the model to go further than LSTM to capture

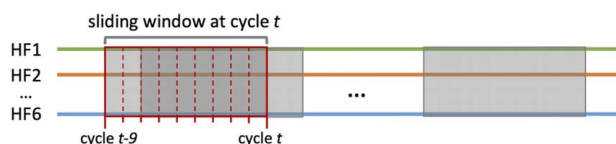


Fig. 5 Sliding window of LSTM layer

temporal dependencies and the trends that evolve with successive time-steps. The performance of models based on these two graph structures is compared in Sec. 3. In the statements that follow, the systems that utilize these two graph construction methods will be referred to as health factor-based GNN and window-based GNN, respectively.

4.2.2 Self-Attention Mechanism and Adjacency Matrix Calculation. The self-attention mechanism operates on the two constructed graphs. It computes an adjacency weight matrix for the graph, where each element represents the weight of the edge connecting nodes. This weight denotes the strength of interaction or correlation between the two nodes as learned from the data.

The process begins by transforming the LSTM output for each node into query, key, and value vectors [38]. They are different representations of the input to self-attention, designed to perform different roles in the process. The query vector usually represents the node that is currently being concerned. It is used to probe other nodes for their relevance to the current one. The keys represent other nodes, which are paired with the query to determine the attention score of each component with respect to the query node. Once an attention score is obtained from query and key pairings, these scores are used to perform weighted message passing on

Table 2 Graph construction comparison

	Health factor-based graph	Window-based graph
Approach	Process each health factor separately with LSTM	Process all health factors together with LSTM
Node definition	Temporal information about each health factor	Temporal information about each cycle in a sliding window
Number of nodes	Six (number of health factors)	Ten (size of sliding window)
Focus	Interactions between health factor	Trends in successive time-steps
Perspective	Emphasizes spatial features between health factors	Exploits deeper temporal features within health factors
Edge	Fully connected	Fully connected

Table 3 Details of the performance comparison

	GNN (health factor-based)	GNN (window-based)	LSTM	GRU
RMSE	0.00809 ± 0.00164	0.01457 ± 0.00895	0.01343 ± 0.00176	0.01531 ± 0.00128
MAE	0.00628 ± 0.00116	0.01128 ± 0.00792	0.01076 ± 0.00137	0.01177 ± 0.00094
Time cost (s)	21.2	13.6	6.5	9.5

Note: Standard deviation is calculated based on five runs of the model.

value vectors. For node i , the transformations are as follows:

$$q_i = \text{Linear}_q(h_i) \quad (11)$$

$$k_i = \text{Linear}_k(h_i) \quad (12)$$

$$v_i = \text{Linear}_v(h_i) \quad (13)$$

where h_i is the LSTM output for node i , and Linear_q , Linear_k , and Linear_v represent the linear transformations for generating the query, key, and value vectors, respectively.

The self-attention mechanism then computes the attention score a_{ij} between two nodes i and j using their query and key vectors [54]

$$a_{ij} = \text{softmax} \left(\frac{q_i \cdot k_j^T}{\sqrt{d_k}} \right) \quad (14)$$

where d_k is the dimension of key vector.

The softmax function in Eq. (14) is used to convert the raw attention scores into probabilities so that the attention scores sum to one. This basically reflects the amount of attention each element in the sequence should receive relative to the other elements. It enables nodes to assign different importance to information from neighboring nodes based on the learned attention weights [55]. The attention scores a_{ij} form the adjacency weight matrix A of the graph. This matrix represents the correlation structure between the nodes and is used in the subsequent aggregation step to weigh the contributions of the different characteristics.

The adjacency weight matrix is used to propagate messages in a graph structure, preventing uniform smoothing and preserving node individuality. The multiply aggregation strategy is employed for health factor-based graph, where the aggregated information from neighbors is obtained by multiplying the messages weighted by attention scores. Multiply aggregation allows for complex interactions between different factors. This will further increase the sensitivity of the model to specific nodes. The influence of nodes that are assigned higher weights in the self-attention mechanism may be amplified exponentially.

However, in window-based graph, additive aggregation is adopted by weighted summation of information from different time windows, which will better maintain temporal continuity. The updated feature vector for each node after the message passing is

$$h'_i = \prod_{j \in \text{Neigh}(i)} a_{ij} \cdot v_j \text{ (health factor based graph)} \quad (15)$$

$$h'_i = \sum_{j \in \text{Neigh}(i)} a_{ij} \cdot v_j \text{ (window based graph)} \quad (16)$$

where $\text{Neigh}(i)$ denotes the set of neighbors of node i , a_{ij} is the attention score, and v_j is the value vector.

4.3 Linear Layer for Prediction. Finally, the updated feature vectors are predicted through two linear layers separated by a ReLU activation function. The output of the last layer is the predicted SOH for each battery cycle. In summary, the proposed system constructs a fully connected graph from the LSTM output, learns

the weights of the edges in the graph using a self-attention mechanism, and forms a neighbor matrix. This graph structure is then used to realize the passing of messages between the nodes and update the features for prediction. This process effectively captures intra-sequence dependencies and inter-sequence correlations in time-series data and provides a more accurate prediction of SOH.

5 Results and Discussion

5.1 Experiments Implementation Details. The graph-based network employed for SOH estimation is implemented in PYTHON utilizing the PyTorch framework. The models are run on a single CPU and utilize the mean squared error loss function during the training process. The developed GNN models have three key parameters that can be fine-tuned: the window size of LSTM layer, hidden unit of LSTM layer, and the initial learning rate. The early stopping mechanism and learning rate scheduler are adopted during model training to prevent overfitting as well as to speed up model convergence.

The validity of the designed SOH estimation model was assessed using two key metrics: the root mean square error (RMSE) and the mean absolute error (MAE). The RMSE serves as an indicator of the stability of the model and quantifies the difference between the actual and predicted values. Meanwhile, the MAE is a linear indicator of the prediction error and provides the accuracy of the model.

5.2 Model Performance Evaluation. The prediction performance of the two temporal enhanced self-attention GNNs and two baseline models is summarized in Table 3 and Fig. 6. The health factor-based GNN gives the lowest prediction error, followed by the LSTM, GRU, and finally the window-based GNN. The high accuracy of health factor-based GNN can be attributed to its ability to capture the complex relationships between temporally enhanced factors. The use of a network structure that facilitates interactions between nodes, can discover subtle patterns and dependencies that are important for more accurate SOH prediction. The model has the lowest error variance and is the most stable of all the models. The lower variance indicates that the method performs consistently across runs, which may be due to the information passing and aggregation brought about by the graph structure, which makes

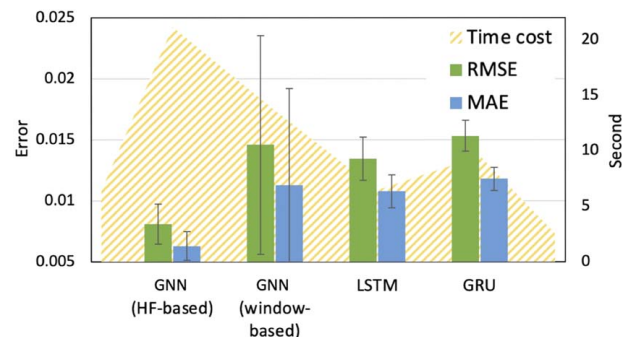


Fig. 6 Performance comparison of the models

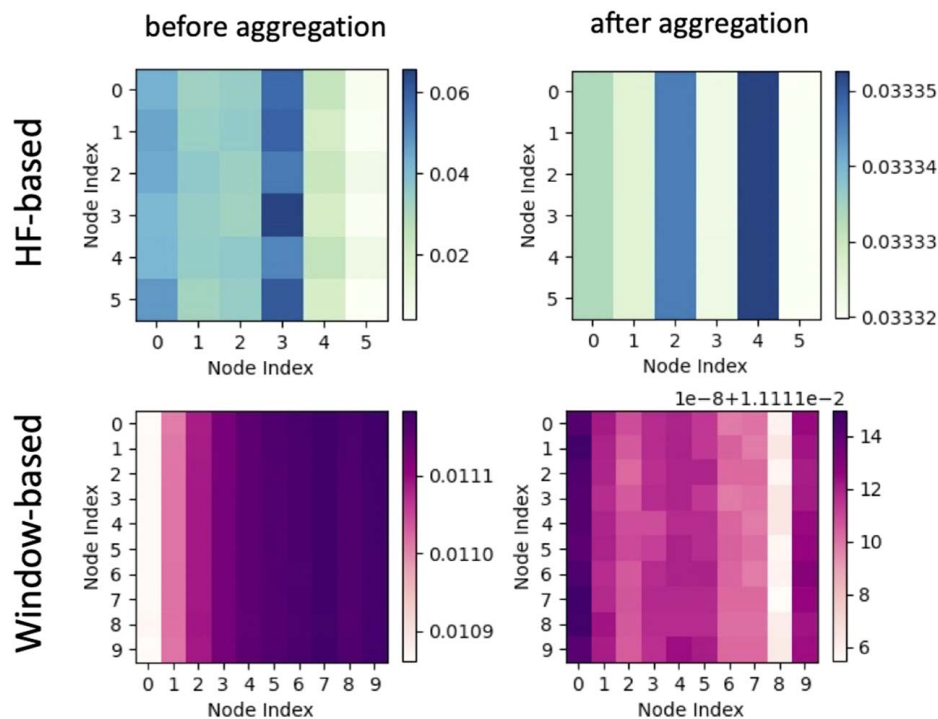


Fig. 7 Visualization of attention score

it robust enough to capture the feature interactions and temporal dependencies.

However, despite its high predictive power, health factor-based GNN requires the longest computational time. The reason is that the model requires separate LSTM layers for each health factor to preserve their distinct temporal patterns. However, the model

using a single layer of LSTM has a faster running speed, such as window-based GNN and LSTM, and a potential loss of information.

On the other hand, window-based GNN has the highest prediction error as well as the most unstable output (based on five runs). In window-based GNN, the graph structure is built based

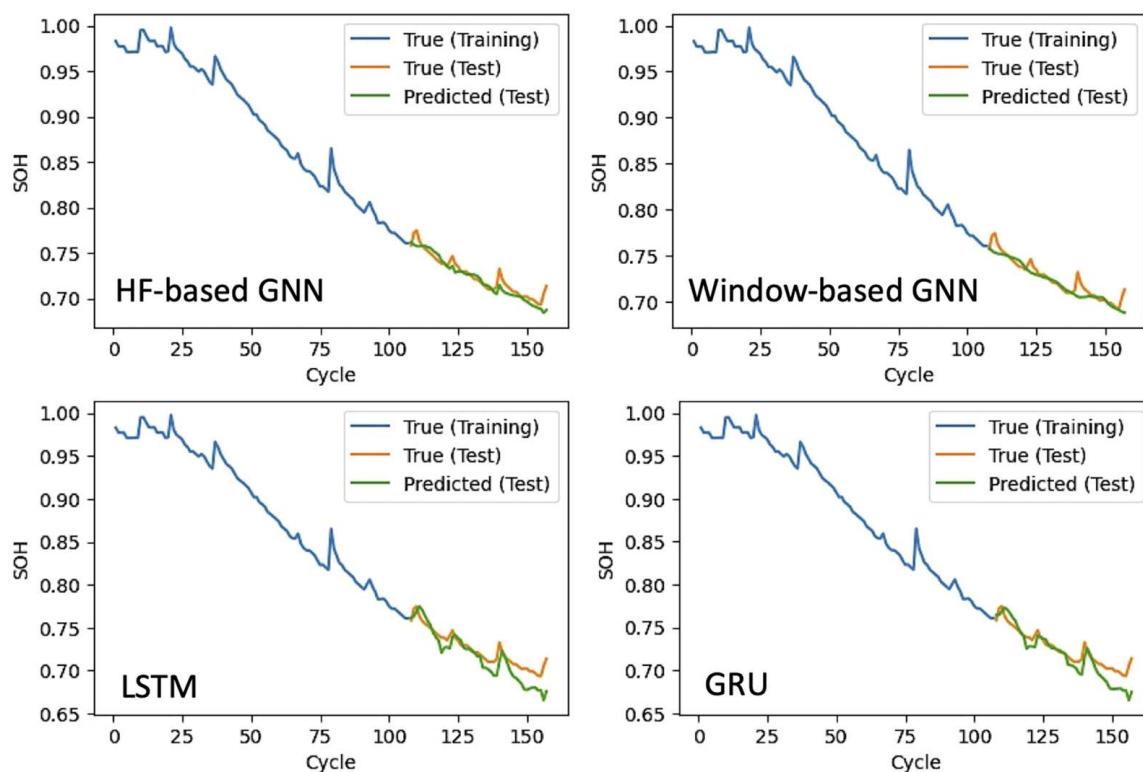


Fig. 8 Visualization of model predictions

on the battery information generated by LSTM at different time-steps. Each of these nodes represents a specific time window. After performing the self-attention based information aggregation, one observed trend is that the attention weights converge, which suggests an over-smoothing problem. It essentially masks the discriminative features of node representations in the aggregation. The potential reason is that the output from the LSTM may not have enough distinctiveness between different time-steps, which makes the nodes become too similar after information fusion. As illustrated in Fig. 7, in the health factor-based graph, the disparity in attention weights amongst nodes before information aggregation reaches the order of 10^{-2} . Following aggregation, this difference shrinks to 10^{-5} . In comparison, the window-based graph presents a more pronounced reduction: initially showcasing a variation at 10^{-4} , which further narrows down to an order of 10^{-8} after the aggregation process.

5.3 Smoothness of Graph Neural Network Predictions.

Figure 8 shows that GNNs generate smoother predictions than LSTM and GRU, which is both a strength and a limitation. The information aggregation mechanism in GNN tends to average the extremes, which results in a smoother curve. This mitigates noise interference and potential overfitting to the training data. However, it also results in the model failing to capture rapid fluctuations in the test data. The challenge here is to find a tradeoff between the model generalization ability and remaining sensitive to minor changes in the data.

5.4 Implications and Future Directions. The health factor-based GNN model has the best accuracy and robustness, but has the highest computational complexity. In contrast, the window-based GNN has the largest prediction error although it runs faster. Also, it shows instability in operation, which could be attributed to the over-smoothing problem. The performance of the LSTM and GRU single models falls between the two GNNs, while the GNNs produce relatively smooth predictions that do not capture random fluctuations well.

It has been shown that the running time of health factor-based GNN models is high, especially when multiple LSTM layers are employed to process the six series separately. A key consideration for future research is the effectiveness of deploying a single LSTM layer rather than multiple. This approach, while computationally more compact, requires a well-developed strategy to preserve the unique dynamics of each factor. Future research may explore ways to circumvent the potential loss of information. Potential approaches are to use feature engineering to emphasize the characteristics of individual health factors before the data are fed into the LSTM, or to optimize the attentional mechanism.

Another important observation is that window-based GNNs have a tendency to experience over-smoothing. Moreover, there is a need to fine-tune health factor-based GNNs to detect more subtle changes in the data without compromising prediction accuracy. Both of these limitations are related to the smoothing properties brought about by information aggregation. Investigating ways to highlight the differences in the output of the LSTM at different health factors or time-steps, as well as adjusting the correlation between the nodes, are both viable ways to address this issue.

6 Conclusion

This study explores the application of graph-based neural networks incorporating LSTM and self-attention mechanism for estimating battery SOH. The framework first employs LSTM to efficiently extract temporal-based information from the battery health factors. Two approaches to graph construction, health factor-based and window-based, are applied to hidden states from the LSTM layer. While the first focuses on understanding the spatial interactions between health factor data, the second delves deeper

into the temporal dependencies of each cycle. Self-attention mechanism is used to quantify the interactions between graph nodes and to weigh the aggregation of information from neighbors to enrich features. The performance of two proposed GNNs is compared with vanilla LSTM and GRU, which are widely used in the field of battery SOH estimation.

Among the various models examined, the health factor-based GNN outperforms other models including LSTM, GRU, and window-based GNN. Its effectiveness is attributed to the ability to capture the complex interactions between temporally enhanced health factors. While window-based GNN is an innovative approach, it suffers from the problem of over-smoothing. This may be caused by the minor difference in the output of the LSTM at different time-steps. It leads to high similarity of each node after the integration of information. Furthermore, the information integration behavior of graph-structured data enhances the robustness of the model, but at the cost of difficulty in capturing stochastic fluctuations in the data. Practical application of the model requires a tradeoff between these two aspects.

For future research, the computational speed of GNN models can be improved, especially when multiple LSTM layers are used. In addition, advanced strategies can be employed to prevent over-smoothing in GNN. Potential solutions include strengthening the distinctness of each feature or restructuring the correlation between nodes. Finally, the model can be applied to more diverse and robust datasets.

Acknowledgment

This material is based upon work supported by the National Science Foundation—USA under Grant No. #2324950. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

References

- [1] Deng, J., Bae, C., Denlinger, A., and Miller, T., 2020, "Electric Vehicles Batteries: Requirements and Challenges," *Joule*, **4**(3), pp. 511–515.
- [2] Diouf, B., and Pöde, R., 2015, "Potential of Lithium-Ion Batteries in Renewable Energy," *Renewable Energy*, **76**, pp. 375–380.
- [3] Wang, D., Coignard, J., Zeng, T., Zhang, C., and Saxena, S., 2016, "Quantifying Electric Vehicle Battery Degradation From Driving Vs. Vehicle-to-Grid Services," *J. Power Sources*, **332**, pp. 193–203.
- [4] Yang, F., Xie, Y., Deng, Y., and Yuan, C., 2021, "Temporal Environmental and Economic Performance of Electric Vehicle and Conventional Vehicle: A Comparative Study on Their US Operations," *Resour. Conserv. Recycl.*, **169**, p. 105311.
- [5] Khodadadi Sadabadi, K., Jin, X., and Rizzoni, G., 2021, "Prediction of Remaining Useful Life for a Composite Electrode Lithium Ion Battery Cell Using an Electrochemical Model to Estimate the State of Health," *J. Power Sources*, **481**, p. 228861.
- [6] Patel, P., and Nelson, G. J., 2020, "The Influence of Structure on the Electrochemical and Thermal Response of Li-Ion Battery Electrodes," *ASME J. Energy Resour. Technol.*, **142**(5), p. 050906.
- [7] Li, J., Du, Z., Ruth, R. E., An, S. J., David, L. A., Hays, K., Wood, M., et al., 2017, "Toward Low-Cost, High-Energy Density, and High-Power Density Lithium-Ion Batteries," *JOM*, **69**(9), pp. 1484–1496.
- [8] Rajaeifar, M. A., Ghadimi, P., Raugei, M., Wu, Y., and Heidrich, O., 2022, "Challenges and Recent Developments in Supply and Value Chains of Electric Vehicle Batteries: A Sustainability Perspective," *Resour. Conserv. Recycl.*, **180**, p. 106144.

- [9] Omariba, Z. B., Zhang, L., and Sun, D., 2018, "Review on Health Management System for Lithium-Ion Batteries of Electric Vehicles," *Electronics*, **7**(5), p. 72.
- [10] Hamut, H. S., Dincer, I., and Naterer, G. F., 2014, "Experimental and Theoretical Efficiency Investigation of Hybrid Electric Vehicle Battery Thermal Management Systems," *ASME J. Energy Resour. Technol.*, **136**(1), p. 011202.
- [11] Nazari, A., Kaviani, S., and Nazari, A., 2020, "Lithium-Ion Batteries' Energy Efficiency Prediction Using Physics-Based and State-of-the-Art Artificial Neural Network-Based Models," *ASME J. Energy Resour. Technol.*, **142**(10), p. 102001.
- [12] Hassanzadeh, M., and Rahmani, Z., 2021, "An Intelligent Predictive Controller for Power and Battery Management in Plug-In Hybrid Electric Vehicles," *ASME J. Energy Resour. Technol.*, **143**(11), p. 112105.
- [13] Maisel, F., Neef, C., Marscheider-Weidemann, F., and Nissen, N. F., 2023, "A Forecast on Future Raw Material Demand and Recycling Potential of Lithium-Ion Batteries in Electric Vehicles," *Resour. Conserv. Recycl.*, **192**, p. 106920.
- [14] Yang, J., Gu, F., and Guo, J., 2020, "Environmental Feasibility of Secondary Use of Electric Vehicle Lithium-Ion Batteries in Communication Base Stations," *Resour. Conserv. Recycl.*, **156**, p. 104713.
- [15] Zhu, J., Mathews, I., Ren, D., Li, W., Cogswell, D., Xing, B., Sedlatschek, T., et al., 2021, "End-of-Life or Second-Life Options for Retired Electric Vehicle Batteries," *Cell Rep. Phys. Sci.*, **2**(8), p. 100537.
- [16] Hua, Y., Liu, X., Zhou, S., Huang, Y., Ling, H., and Yang, S., 2021, "Toward Sustainable Reuse of Retired Lithium-Ion Batteries From Electric Vehicles," *Resour. Conserv. Recycl.*, **168**, p. 105249.
- [17] Albertsen, L., Richter, J. L., Peck, P., Dalhammar, C., and Plepys, A., 2021, "Circular Business Models for Electric Vehicle Lithium-Ion Batteries: An Analysis of Current Practices of Vehicle Manufacturers and Policies in the EU," *Resour. Conserv. Recycl.*, **172**, p. 105658.
- [18] Ma, B., Yang, S., Zhang, L., Wang, W., Chen, S., Yang, X., Xie, H., Yu, H., Wang, H., and Liu, X., 2022, "Remaining Useful Life and State of Health Prediction for Lithium Batteries Based on Differential Thermal Voltammetry and a Deep-Learning Model," *J. Power Sources*, **548**, p. 232030.
- [19] Allafi, W., Uddin, K., Zhang, C., Mazuir Raja Ahsan Sha, R., and Marco, J., 2017, "On-Line Scheme for Parameter Estimation of Nonlinear Lithium Ion Battery Equivalent Circuit Models Using the Simplified Refined Instrumental Variable Method for a Modified Wiener Continuous-Time Model," *Appl. Energy*, **204**, pp. 497–508.
- [20] Hu, X., Li, S., and Peng, H., 2012, "A Comparative Study of Equivalent Circuit Models for Li-Ion Batteries," *J. Power Sources*, **198**, pp. 359–367.
- [21] Doyle, M., Fuller, T. F., and Newman, J., 1993, "Modeling of Galvanostatic Charge and Discharge of the Lithium/Polymer/Insertion Cell," *J. Electrochem. Soc.*, **140**(6), pp. 1526–1533.
- [22] Li, J., Adewuyi, K., Lotfi, N., Landers, R. G., and Park, J., 2018, "A Single Particle Model With Chemical/Mechanical Degradation Physics for Lithium Ion Battery State of Health (SOH) Estimation," *Appl. Energy*, **212**, pp. 1178–1190.
- [23] Zhang, S., Guo, X., Dou, X., and Zhang, X., 2020, "A Rapid Online Calculation Method for State of Health of Lithium-Ion Battery Based on Coulomb Counting Method and Differential Voltage Analysis," *J. Power Sources*, **479**, p. 228740.
- [24] Meng, J., Luo, G., Ricco, M., Swierczynski, M., Stroe, D.-I., and Teodorescu, R., 2018, "Overview of Lithium-Ion Battery Modeling Methods for State-of-Charge Estimation in Electrical Vehicles," *Appl. Sci.*, **8**(5), p. 659.
- [25] Chen, Z., Zhao, H., Zhang, Y., Shen, S., Shen, J., and Liu, Y., 2022, "State of Health Estimation for Lithium-Ion Batteries Based on Temperature Prediction and Gated Recurrent Unit Neural Network," *J. Power Sources*, **521**, p. 230892.
- [26] Beganovic, N., and Söffker, D., 2019, "Estimation of Remaining Useful Lifetime of Lithium-Ion Battery Based on Acoustic Emission Measurements," *ASME J. Energy Resour. Technol.*, **141**(4), p. 041901.
- [27] Tian, J., Xiong, R., Shen, W., Lu, J., and Yang, X.-G., 2021, "Deep Neural Network Battery Charging Curve Prediction Using 30 Points Collected in 10 Min," *Joule*, **5**(6), pp. 1521–1534.
- [28] Patil, M. A., Tagade, P., Hariharan, K. S., Kolake, S. M., Song, T., Yeo, T., and Doo, S., 2015, "A Novel Multistage Support Vector Machine Based Approach for Li Ion Battery Remaining Useful Life Estimation," *Appl. Energy*, **159**, pp. 285–297.
- [29] Yang, D., Zhang, X., Pan, R., Wang, Y., and Chen, Z., 2018, "A Novel Gaussian Process Regression Model for State-of-Health Estimation of Lithium-Ion Battery Using Charging Curve," *J. Power Sources*, **384**, pp. 387–395.
- [30] Li, X., Zhang, L., Wang, Z., and Dong, P., 2019, "Remaining Useful Life Prediction for Lithium-Ion Batteries Based on a Hybrid Model Combining the Long Short-Term Memory and Elman Neural Networks," *J. Energy Storage*, **21**, pp. 510–518.
- [31] Kaur, K., Garg, A., Cui, X., Singh, S., and Panigrahi, B. K., 2021, "Deep Learning Networks for Capacity Estimation for Monitoring SOH of Li-Ion Batteries for Electric Vehicles," *Int. J. Energy Res.*, **45**(2), pp. 3113–3128.
- [32] Venugopal, P., and Vigneswaran, T., 2019, "State-of-Health Estimation of Li-Ion Batteries in Electric Vehicle Using IndRNN Under Variable Load Condition," *Energies*, **12**(22), p. 4338.
- [33] Li, P., Zhang, Z., Xiong, Q., Ding, B., Hou, J., Luo, D., Rong, Y., and Li, S., 2020, "State-of-Health Estimation and Remaining Useful Life Prediction for the Lithium-Ion Battery Based on a Variant Long Short Term Memory Neural Network," *J. Power Sources*, **459**, p. 228069.
- [34] Liu, G., and Guo, J., 2019, "Bidirectional LSTM With Attention Mechanism and Convolutional Layer for Text Classification," *Neurocomputing*, **337**, pp. 325–338.
- [35] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I., 2017, "Attention Is All You Need," *Neural Inf. Process. Syst.*, **30**, pp. 5998–6008.
- [36] Qu, J., Liu, F., Ma, Y., and Fan, J., 2019, "A Neural-Network-Based Method for RUL Prediction and SOH Monitoring of Lithium-Ion Battery," *IEEE Access*, **7**, pp. 87178–87191.
- [37] Jiang, Y., Chen, Y., Yang, F., and Peng, W., 2023, "State of Health Estimation of Lithium-Ion Battery With Automatic Feature Extraction and Self-Attention Learning Mechanism," *J. Power Sources*, **556**, p. 232466.
- [38] Ge, Y., Zhang, F., and Ren, Y., 2022, "Lithium Ion Battery Health Prediction Via Variable Model Decomposition and Deep Learning Network With Self-Attention Mechanism," *Front. Energy Res.*, **10**, p. 810490.
- [39] Ren, L., Dong, J., Wang, X., Meng, Z., Zhao, L., and Deen, M. J., 2021, "A Data-Driven Auto-CNN-LSTM Prediction Model for Lithium-Ion Battery Remaining Useful Life," *IEEE Trans. Ind. Inf.*, **17**(5), pp. 3478–3487.
- [40] Tian, J., Xiong, R., Lu, J., Chen, C., and Shen, W., 2022, "Battery State-of-Charge Estimation Amid Dynamic Usage With Physics-Informed Deep Learning," *Energy Storage Mater.*, **50**, pp. 718–729.
- [41] Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., and Sun, M., 2020, "Graph Neural Networks: A Review of Methods and Applications," *AI Open*, **1**, pp. 57–81.
- [42] Yao, X.-Y., Chen, G., Pecht, M., and Chen, B., 2023, "A Novel Graph-Based Framework for State of Health Prediction of Lithium-Ion Battery," *J. Energy Storage*, **58**, p. 106437.
- [43] Wei, Y., and Wu, D., 2023, "Prediction of State of Health and Remaining Useful Life of Lithium-Ion Battery Using Graph Convolutional Network With Dual Attention Mechanisms," *Reliab. Eng. Syst. Saf.*, **230**, p. 108947.
- [44] Wang, Z., Yang, F., Xu, Q., Wang, Y., Yan, H., and Xie, M., 2023, "Capacity Estimation of Lithium-Ion Batteries Based on Data Aggregation and Feature Fusion Via Graph Neural Network," *Appl. Energy*, **336**, p. 120808.
- [45] "Li-Ion Battery Aging Datasets | NASA Open Data Portal," <https://data.nasa.gov/dataset/Li-ion-Battery-Aging-Datasets/uj5r-zjdb>. Accessed March 11, 2023.
- [46] Khumprom, P., and Yodo, N., 2019, "A Data-Driven Predictive Prognostic Model for Lithium-Ion Batteries Based on a Deep Learning Algorithm," *Energies*, **12**(4), p. 660.
- [47] Zhao, Y., and Behdad, S., 2023, "State of Health Estimation of Electric Vehicle Batteries Using Transformer-Based Neural Network," *Proceedings of the ASME IDETC/CIE*, Boston, MA, Aug. 20–23, American Society of Mechanical Engineers Digital Collection, p. V005T05A017.
- [48] Li, Y., Stroe, D.-I., Cheng, Y., Sheng, H., Sui, X., and Teodorescu, R., 2021, "On the Feature Selection for Battery State of Health Estimation Based on Charging–Discharging Profiles," *J. Energy Storage*, **33**, p. 102122.
- [49] Kirchev, A., 2015, "Chapter 20—Battery Management and Battery Diagnostics," *Electrochemical Energy Storage for Renewable Sources and Grid Balancing*, P. T. Moseley, and J. Garche, eds., Elsevier, Amsterdam, pp. 411–435.
- [50] Wang, H., Frisco, S., Gottlieb, E., Yuan, R., and Whitacre, J. F., 2019, "Capacity Degradation in Commercial Li-Ion Cells: The Effects of Charge Protocol and Temperature," *J. Power Sources*, **426**, pp. 67–73.
- [51] Ma, Y., Shan, C., Gao, J., and Chen, H., 2022, "A Novel Method for State of Health Estimation of Lithium-Ion Batteries Based on Improved LSTM and Health Indicators Extraction," *Energy*, **251**, p. 123973.
- [52] Yang, H., Wang, P., An, Y., Shi, C., Sun, X., Wang, K., Zhang, X., Wei, T., and Ma, Y., 2020, "Remaining Useful Life Prediction Based on Denoising Technique and Deep Neural Network for Lithium-Ion Capacitors," *eTransportation*, **5**, p. 100078.
- [53] Zhang, X., Sun, J., Shang, Y., Ren, S., Liu, Y., and Wang, D., 2022, "A Novel State-of-Health Prediction Method Based on Long Short-Term Memory Network With Attention Mechanism for Lithium-Ion Battery," *Front. Energy Res.*, **10**, p. 972486.
- [54] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., and Bengio, Y., 2018, "Graph Attention Networks," *Proceedings of ICLR*, Vancouver, BC, Canada, Apr. 30–May 3, arxiv.org/abs/1710.10903.
- [55] Brody, S., Alon, U., and Yahav, E., 2022, "How Attentive Are Graph Attention Networks?," *Proceedings of ICLR*, Virtual, Apr. 25–29, arxiv.org/abs/2105.14491.