

# An Attention-Based Deep Learning Approach for Lithium-ion Battery Lifespan Prediction: Analysis and Experimental Validation

Ahmed Darwish<sup>1,\*</sup> 

<sup>1</sup> Department of Computer Science, Faculty of Computer and Informatics, Zagazig University, Zagazig, 44519, Egypt; [adarwish@fci.zu.edu.eg](mailto:adarwish@fci.zu.edu.eg)

\* Correspondence: [adarwish@fci.zu.edu.eg](mailto:adarwish@fci.zu.edu.eg)

**Abstract:** The potential for lithium-ion batteries to become unstable can lead to operational malfunctions within the system and result in safety incidents. Therefore, accurately forecasting the remaining useful life (RUL) is beneficial in mitigating the likelihood of battery failure and prolonging its operational lifespan. Hence, precise estimation of RUL can help prevent numerous safety incidents and minimize resource wastage, presenting a significant and complex issue. This paper introduces a Deep Learning (DL) model that utilizes Long Short-Term Memory (LSTM) and attention mechanism to improve the accuracy of predicting the RUL of lithium-ion batteries. Initially, the battery capacity regeneration phenomenon is captured by applying four LSTM layers, followed by implementing an attention mechanism to align input and output sequences based on the content or semantics of the input sequence. Finally, the final prediction outcomes are generated via a Fully Connected (FC) layer. The efficacy of the proposed model is assessed through the utilization of the NASA dataset, and its performance is contrasted with various deep learning models to highlight its efficacy. Results from the experiments demonstrate that the suggested At-LSTM presents a robust option for forecasting the RUL of lithium-ion batteries, as it delivers superior results compared to all other models examined.

**Keywords:** Lithium-ion Batteries Prognosis; Remaining Useful Life; Long Short-term Memory Network; Attention Mechanism.

---

## 1. Introduction

Two major issues facing the world community are environmental degradation and the energy crisis [1]. Consequently, numerous nations have proactively initiated the growth of the electric vehicle sector in recent times [2]. In recent years, there has been a significant increase in the utilization of conventional coal-based fossil fuels, leading to energy shortages and significant environmental harm [3]. Lithium-ion batteries (LIBs) have gained significant popularity as a substitute for conventional energy sources in the realm of new sustainable energy. Their utilization spans a range of sectors including Electric Vehicles (EVs), Automated Guided Vehicles (AGVs), and aerospace due to their exceptional attributes, such as high energy density, extended lifespan, and minimal self-discharge rate [4-8]. Nevertheless, the efficiency of lithium-ion batteries tends to decline over time due to usage, resulting in a range of issues such as reduced capacity and shorter driving distances. Once the maximum discharging capacity of the battery drops to 70%–80% of its original rating, it is considered to have reached the threshold for capacity failure and is deemed unsuitable for electric vehicle (EV) use, necessitating retirement[9]. Over time, lithium batteries will deteriorate in capacity due to prolonged use, potentially resulting in equipment malfunctions and serious incidents [10]. Conversely, premature battery replacement may result in the wastage of valuable battery resources

[11]. Hence, accurately forecasting the Remaining Useful Life (RUL) of lithium-ion batteries is crucial. Precise predictions can help avoid accidents resulting from battery degradation and reduce unnecessary waste from premature battery replacements [12]. Several research methodologies have been created by academics in recent years to study battery capacity and predict RUL. These approaches can typically be categorized into three groups: model-based methods [13], and data-driven approaches [14].

In model-based approaches involve creating mathematical or physical representations based on data collected through measurements to comprehend the degradation process of batteries [9]. These techniques frequently demand existing information to delineate the inner workings of the lithium-ion battery, such as the Electrochemical model (EM), equivalent circuit model (ECM), and empirical model, which are commonly employed for forecasting the capacity or RUL of lithium-ion batteries within model-based methodologies [15]. Sadabadi et al. [16] formulated an advanced single-particle model incorporating improved parameters to forecast the RUL, which has the potential to be applied utilizing data from Electric Vehicle (EV) charging. This model offers a foundational explanation for the degradation of batteries resulting from electrochemical processes occurring within them. Kim et al. [17] formulated a method for predicting RUL by integrating anomaly detection techniques with particle filtering (PF) based forecasting using a basic empirical degradation model. Chen et al. [18] introduced an enhanced method for predicting RUL using a particle filter. This method was developed by merging the linear optimization resampling particle filter (LORPF) with the sliding-window gray model (SGM). Jiao et al. [19] introduced an innovative PF framework utilizing a conditional variational autoencoder (CVAE) and a reweighting technique for the estimation of the RUL of batteries. Despite being able to predict the RUL, these methods still have certain drawbacks. Firstly, it can be challenging to fine-tune the model parameters. Secondly, observer techniques like the particle filter are prone to issues such as particle impoverishment, resulting in inaccurate RUL predictions. Thirdly, the computational requirements of model-based approaches are significant, making it challenging to develop a model that accurately captures the aging properties of a battery at all stages of its lifespan due to the intricate chemical reactions involved. On the other hand, data-driven methods do not necessitate considering chemical reactions and aging dynamics within batteries. These methods involve extracting aging characteristics from extensive data collected at various stages of aging and establishing their correlation with RUL, making them widely utilized for RUL estimation.

The second approach involves data-driven techniques. In contrast to model-based approaches, data-driven methods do not require the construction of a sophisticated electrochemical model for lithium-ion batteries. Instead, they primarily extract implicit information from capacity degradation data of lithium-ion batteries to enable the prediction of RUL. Data-driven techniques encompass Machine Learning (ML) and Deep Learning (DL) methodologies. Machine learning techniques have become a powerful tool in various aspects of our lives, providing computers with the ability to learn from data without explicit programming, continuously improving their performance. The ability of machine learning to extract knowledge from data and perform tasks automatically is transforming how we live, work, and interact with technology. With further advancements in this field, we can expect even more profound impacts on our global society. Machine learning techniques can leverage large sets of sensor data, operating parameters, and past maintenance records. This data-driven approach allows machine learning models to understand complex relationships among different factors that influence the state and deterioration of machinery. Patil et al. [20] identified crucial characteristics from voltage and temperature data, combining support vector machine techniques to address both classification and regression objectives in order to estimate the gross value and predict the RUL. Zhang et al. [21] utilized the relevance vector machine (RVM) enhanced by the differential evolution (DE) algorithm to predict the RUL of batteries using denoised data generated by the wavelet denoising algorithm. Liu et al. [22] conducted prognostics of RUL by employing an RVM and

several Gaussian Process Regression (GPR) models, with the health indicators (HIs) being utilized as input for the models. Machine learning algorithms have advanced in the estimation of RUL for Lithium-ion batteries; however, many of these approaches necessitate extensive feature engineering in order to identify key features. Furthermore, the modest complexity of these models constrains their ability to adequately capture information and demonstrate strong generalization performance.

Deep learning, a prominent subset of machine learning, has led to substantial changes in various aspects of our daily lives by utilizing artificial neural networks with multiple layers to process information in a way that mimics the cognitive functions of the human brain. The advancement of deep learning techniques is primarily attributed to their versatility, as they eliminate the requirement for manual feature engineering by autonomously extracting feature representations. Predicting RUL involves addressing a time series regression issue, where deep learning models have proven adept at capturing temporal patterns from historical data. widely utilized deep learning structures like Convolutional Neural Networks (CNNs) [23], Recurrent Neural Networks (RNNs) [24], and Transformer [25] are commonly utilized in the prediction of RUL for Lithium-Ion Batteries. Catelani et al. [26] combined the RNN with a filtering-based method to predict the RUL. In this approach, the optimization of the RNN's performance is guaranteed through the utilization of Genetic Algorithms (GA). Zhang et al. [27] crafted a hybrid parallel residual CNN architecture, leveraging the network to capture characteristics from charging data by integrating voltage, current, and temperature profiles from various cycles. chinomona et al. [28] carried out a forward selection-LSTM method, which effectively identifies an optimal subset of features from the original signals by excluding unnecessary ones. Ren et al. [8] merged an enhanced CNN with LSTM, with the CNN focusing on capturing profound information and the LSTM specializing in extracting temporal information. Park et al. [29] introduced a many-to-one architecture of LSTM for precise prediction of the RUL for lithium-ion batteries. Particle swarm optimization was utilized to optimize the parameters of the model. A hybrid Elman-LSTM model was introduced by [30] for the prediction of RUL. The Elman neural network and LSTM were employed for the prediction of both high- and low-frequency components derived from the empirical mode decomposition algorithm. Ren et al. [8] utilized an auto-encoder to enhance the initial dataset, subsequently extracting in-depth insights through the utilization of CNN and Long Short-Term Memory Neural Network (LSTM NN).

Most methods for predicting RUL rely on having a sufficient amount of historical data. However, the availability of data is limited due to the constraints of charge and discharge cycles in Lithium-ion Batteries. Even though the cycle life of the battery is increasing gradually as Lithium-ion battery manufacturing technology advances, the existing historical data remains inadequate for developing accurate prediction models. Moreover, assessing the battery's lifespan based on long-term historical cycle data becomes irrelevant as the battery nears the end of its useful life after prolonged cycle sampling.

Driven by the aforementioned rationale, this study introduces a new methodology that integrates the LSTM network and attention mechanism (At-LSTM) to enhance the accuracy of RUL prediction for lithium-ion batteries. Initially, a variety of temporal characteristics are chosen and employed as inputs for the training and evaluation of the suggested model. Through experimentation, these features have been demonstrated to exhibit a significant correlation with capacity, which is widely recognized as a key indicator of battery health. An attention mechanism has been integrated into the LSTM architecture to allocate weight values to the extracted data, thereby emphasizing crucial information and enhancing the RUL prediction of the model. We assess the effectiveness of the AT-LSTM model on the commonly used NASA [31] dataset for predicting the RUL of Lithium-ion batteries. Our experimental findings demonstrate that the proposed approach reduces uncertainty in multi-step prediction tasks and surpasses the performance of alternative existing models in terms of accuracy. The major contributions of this paper are listed as follows:

- 1) A new approach for predicting the RUL of Lithium-ion batteries is proposed, involving the utilization of LSTM to capture intricate temporal characteristics from aging data, followed by the reconstruction of the output state using layers of attention mechanism.
- 2) An attention mechanism has been developed with the aim of enhancing the accuracy of predicting the RUL by selectively filtering input characteristics and assigning greater importance to crucial features.
- 3) The model's validation was conducted with the NASA dataset, demonstrating its superiority in predicting RUL compared to other models.

The subsequent sections of this paper are structured as follows. The materials and methods are outlined in Section 2. Subsequently, the proposed model will be discussed in Section 3. This will be followed by the presentation of experiments and results in Section 4. Followed by applications in section 5. Lastly, the conclusions will be presented in Section 6.

## 2. Materials and Methods

### 2.1. NASA Lithium-ion batteries dataset

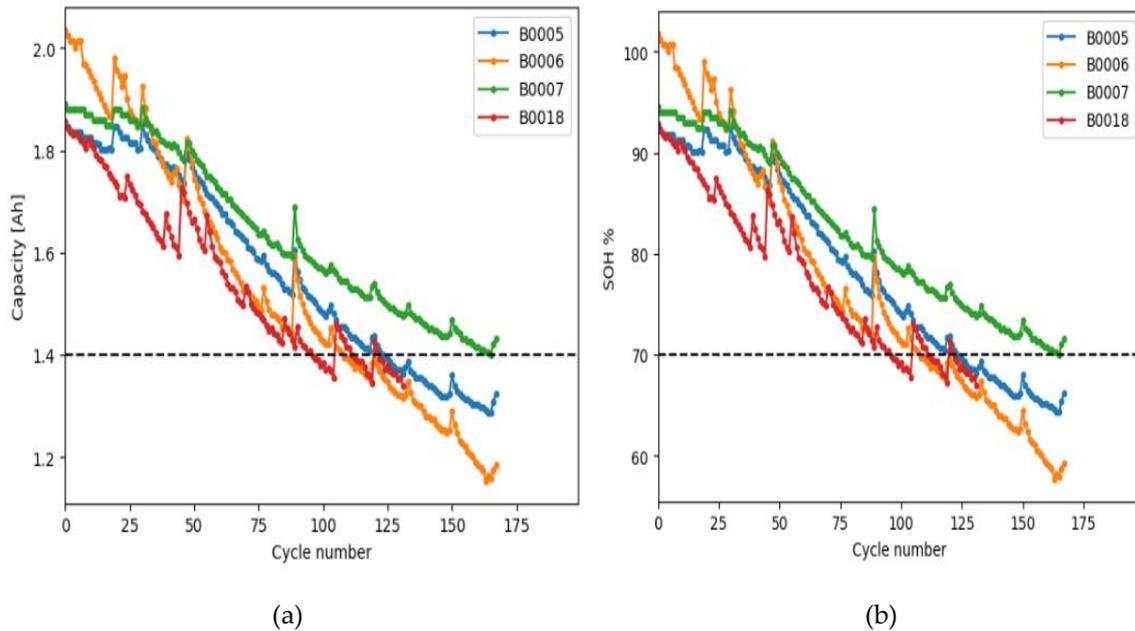
To assess the effectiveness of the suggested model in predicting the Remaining Useful Life (RUL) of Lithium-ion batteries, the NASA dataset is utilized. This dataset is made available by the NASA Ames Prognostics Center of Excellence [31] and comprises information from four 18650 lithium-ion batteries identified as B0005, B0006, B0007, and B0018. The operational cycles of these batteries, each possessing a rated capacity of 2Ah, involve both charging and discharging phases. During charging, a constant current is applied until the battery voltage reaches 4.2V, followed by a shift to constant voltage charging until the current decreases to 20mA. In the discharging phase, the batteries are discharged at a 2A current until specific voltage thresholds are reached. The battery capacity and state of health (SOH) decay curves of these four batteries are shown in Figure 1. In this paper, the Lithium-ion batteries are configured with a capacity threshold of 1.4Ah as shown in Figure 1. Given the variations in operational parameters among the four batteries, the failure threshold is 1.4Ah [32]. The precise specifications of the chosen lithium-ion batteries from NASA are presented in Table 1. The status of a battery is closely linked to its capacity, which is widely regarded as the primary indicator of battery health [33]. The capacity of a battery diminishes as it undergoes more charging and discharging cycles, and this decline in capacity exhibits a consistent pattern over time. As a result, the longevity of Lithium-ion batteries can be defined by their actual capacity. Relevant parameters were identified based on the timing of when each cycle of voltage, current, and temperature hits their respective cutoff points.

**Table 1.** The precise specifications of the chosen lithium-ion batteries from NASA.

Battery	B0005	B0006	B0007	B0018
Discharge current (A)	2	2	2	2
Rated capacity (Ah)	2	2	2	2
Charing/discharge cut-off voltage (V)	4.2/2.7	4.2/2.5	4.2/2.3	4.2/2.5
Minimal charge current (mA)	20	20	20	20
Temperature (C)	24	24	24	24
Failure threshold (Ah)	1.4	1.4	1.4	1.4
Capacity data (Cycles)	168	168	168	132

These parameters exhibit unique time-series properties, exhibit strong associations with capacity, and are closely linked to the overall health of the battery system. The evolving patterns of different parameters during charging for B0005, B0006, B0007, and B0018 are depicted in Figure 2, Figure 3, and Figure 4, while the evolving patterns during discharging for B0005, B0006, B0007, and B0018 are illustrated in Figure 5, Figure 6, and Figure 7. With an increasing number of cycles, the

battery's capacity will gradually diminish. In scenarios involving constant current charging and discharging, an earlier attainment of the cutoff voltage leads to a quicker approach to the cutoff current.



**Figure 1.** Decay curves. (a) present the capacity degradation curves of the NASA dataset, (b) present the SOH degradation curves of the NASA Lithium-ion batteries dataset.

## 2.2. Long Short-Term Memory (LSTM)

LSTM, introduces a novel iteration of the recurrent neural network (RNN) architecture, strategically designed to tackle the issue of vanishing gradients encountered in traditional RNNs, with a specific focus on addressing the challenges posed by long-term dependencies in predictive tasks. This innovative model showcases a sophisticated memory cell structure that sets it apart by its ability to effectively retain and utilize information over prolonged sequences, making it particularly well-suited for applications requiring the prediction of long-term dependencies, such as RUL estimation in Lithium-ion battery systems. The LSTM cell architecture is shown in Figure 8. Within the LSTM framework, the forget gate ( $f_t$ ), input gate ( $i_t$ ), and output gate ( $o_t$ ) constitute the three key components that collectively govern the flow of information and regulate the interactions within the network. The forget gate is responsible for determining which information should be excluded from the previous cell state. By considering the current input and the previous hidden state, the forget gate generates an output value ranging from 0 (indicating complete forgetfulness) to 1 (indicating full retention). The  $f_t$  is mathematically defined using equation 1. The input gate is responsible for determining the novel data to be retained in the cell state. By taking the current input and the preceding hidden state as inputs, it generates an output ranging from 0 to 1. Furthermore, it generates a fresh candidate value intended for incorporation into the cell state. It is mathematically defined using equation 2. The output gate regulates the selection of information to be transmitted as the hidden state of the current LSTM cell, based on both the current input and the previous hidden state. Its output is a numerical value ranging from 0 to 1. The  $o_t$  is mathematically defined using equation 3. The candidate value  $c'_t$  denotes fresh data that may be incorporated into the cell state at the present step ( $t$ ). This value is generated by the input gate, taking into account the current input and the previous hidden state. The  $c'_t$  mathematically defined using equation 4. Next, determine the  $ct$  value representing the unit state at time  $t$  that is mathematically defined using equation 5. Subsequently, ascertain the  $ht$  value representing the hidden state at time  $t$  through mathematical equation 6.

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{1}$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{2}$$

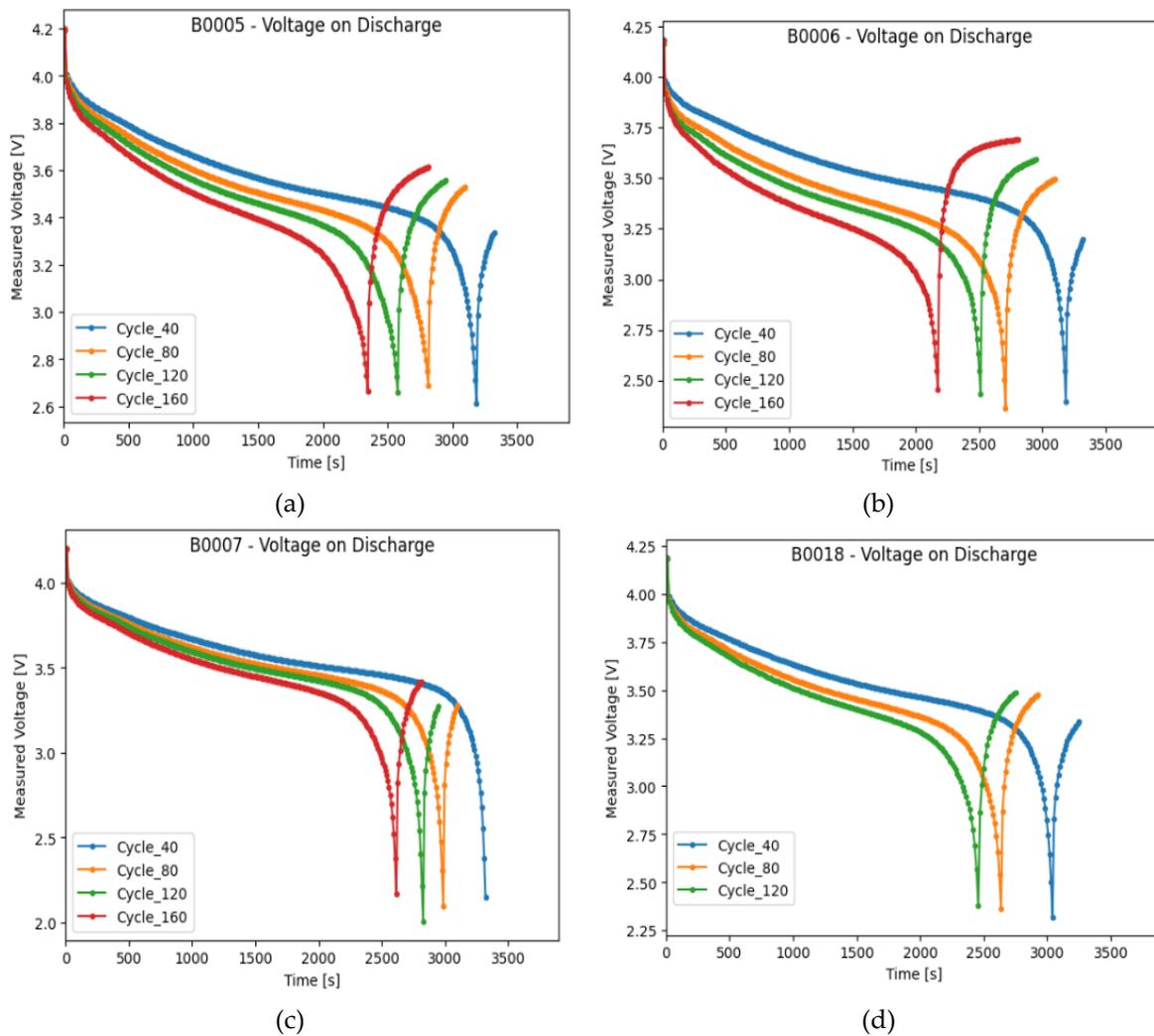
$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{3}$$

$$c'_t = \tanh(W_a x_t + U_a h_{t-1} + b_a) \tag{4}$$

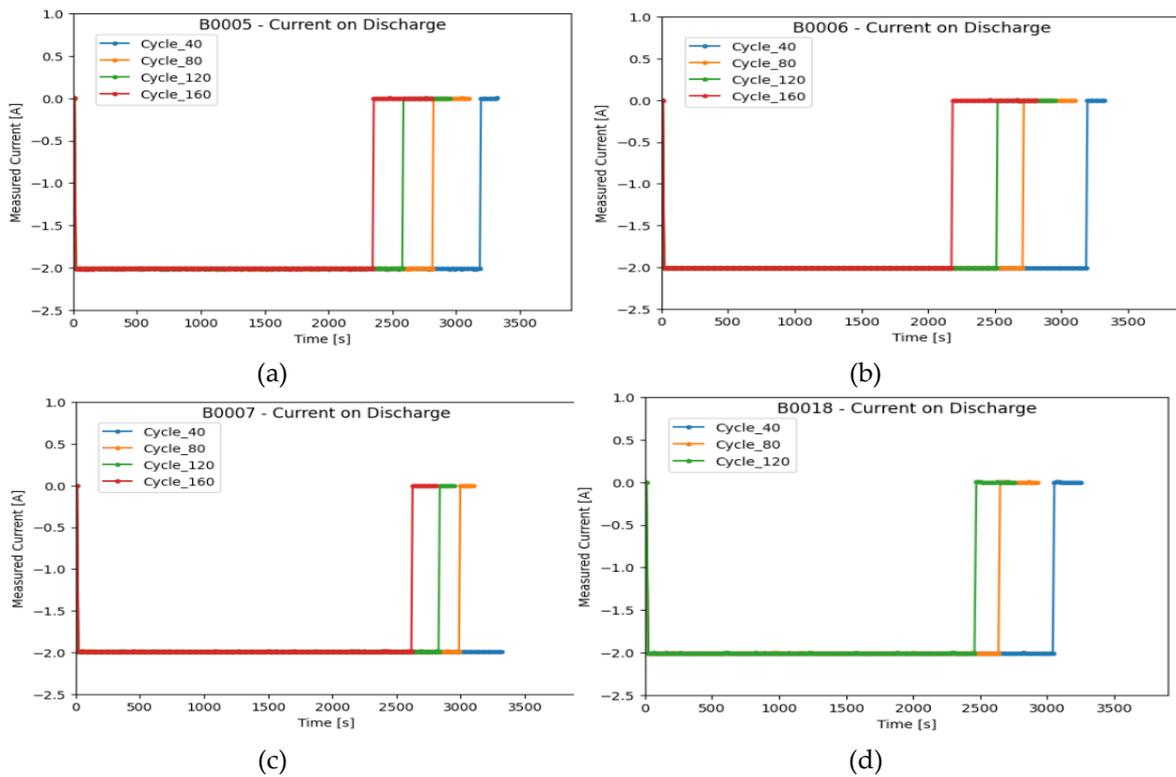
$$c_t = f_t \cdot c_{t-1} + i_t \cdot c'_t \tag{5}$$

$$h_t = o_t \cdot \tanh(c_t) \tag{6}$$

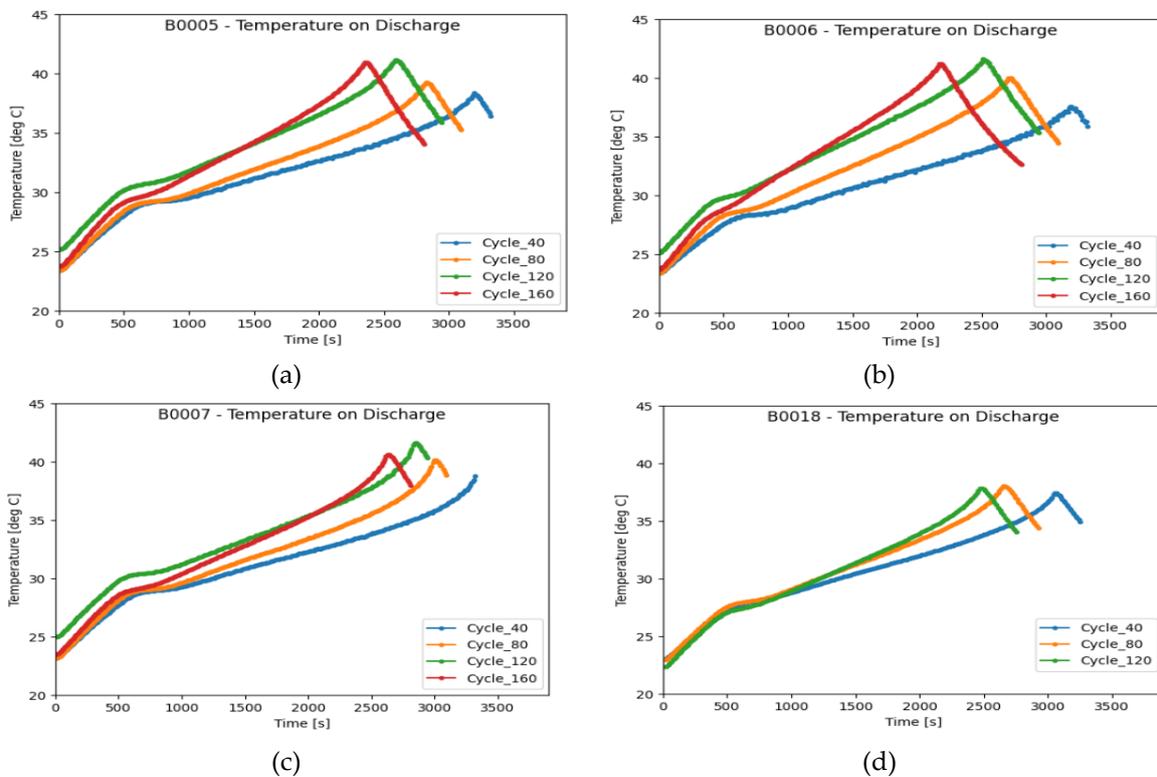
Where the symbol  $\sigma$  denotes the sigmoid function,  $t$  represents the time step,  $x_t$  signifies the input feature at time  $t$ ,  $h_{t-1}$  denotes the output hidden state from the previous time sample, the parameters  $W_f, W_i, W_o, W_a, U_f, U_i, U_o, U_a, b_f, b_i, b_o, b_a$  are optimized during the training process.



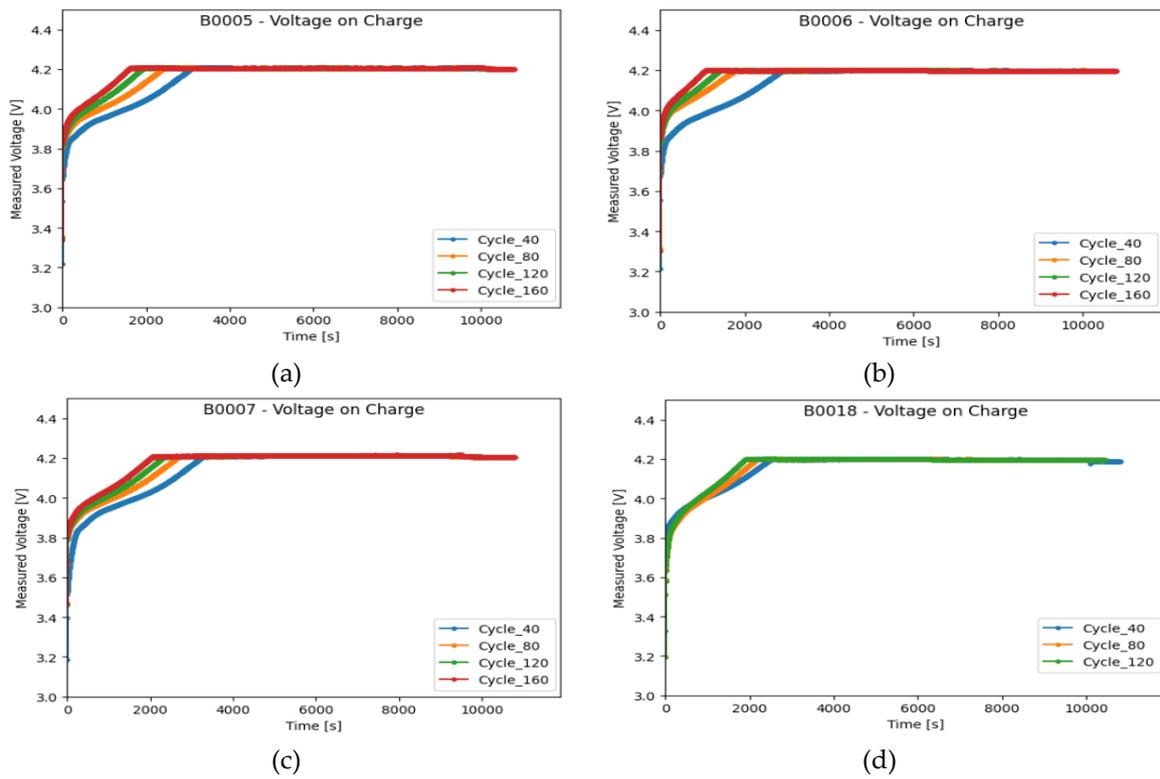
**Figure 2.** Commencing from the 40th cycle, the alterations in voltage during the discharging process are recorded at intervals of every 40 cycles. (a) B0005, (b) B0006, (c) B0007, and (d) B0018.



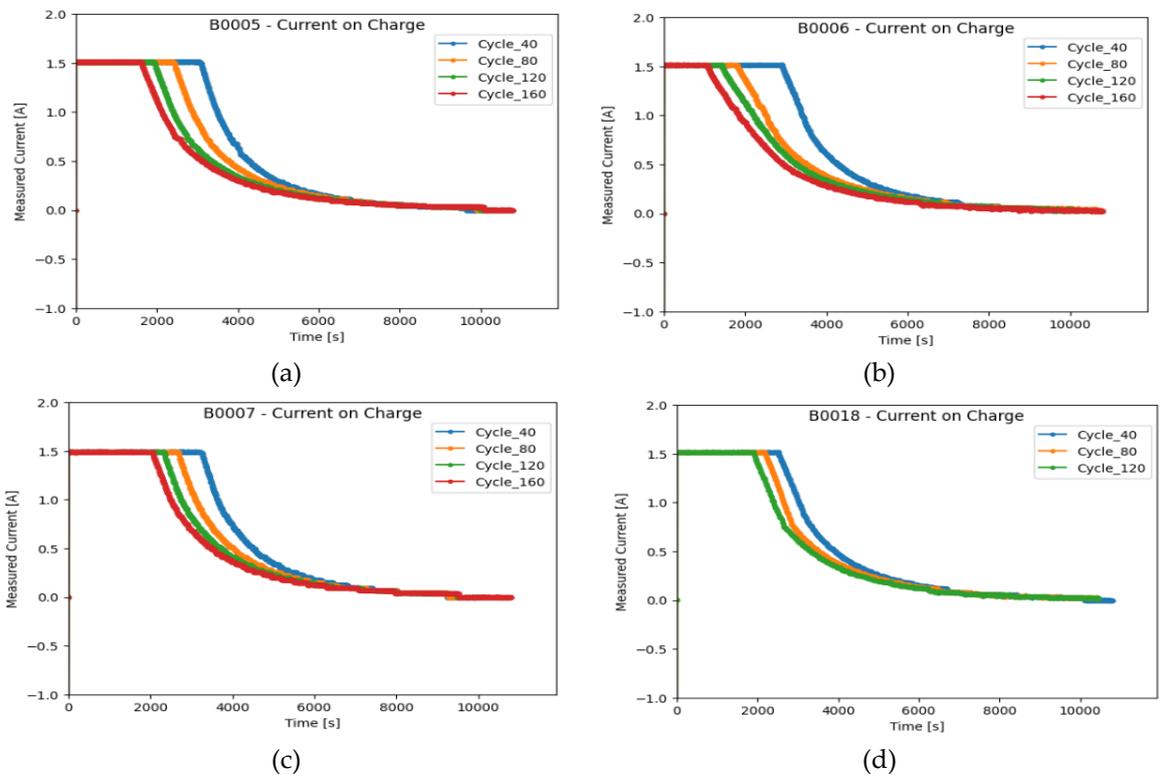
**Figure 3.** Commencing from the 40th cycle, the alterations in current during the discharging process are recorded at intervals of every 40 cycles. (a) B0005, (b) B0006, (c) B0007, and (d) B0018.



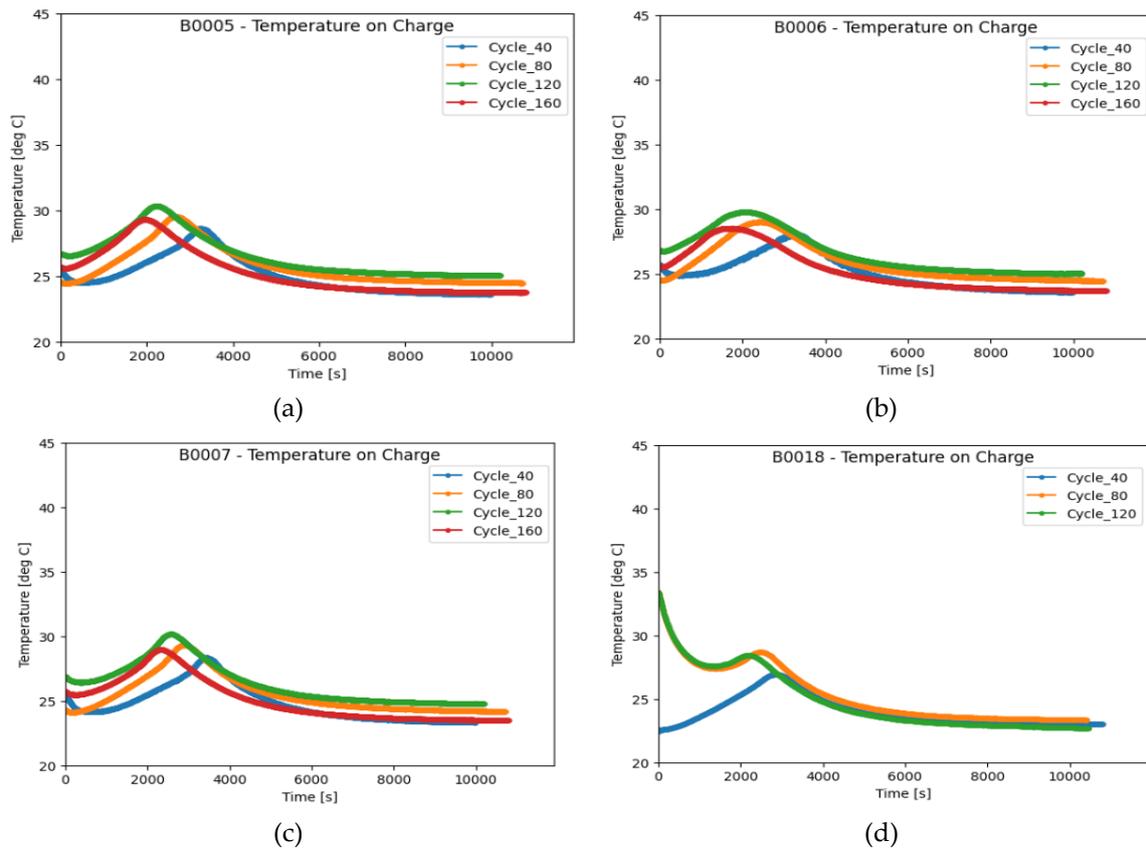
**Figure 4.** Commencing from the 40th cycle, the alterations in temperature during the discharging process are recorded at intervals of every 40 cycles. (a) B0005, (b) B0006, (c) B0007, and (d) B0018.



**Figure 5.** Commencing from the 40th cycle, the alterations in voltage during the charging process are recorded at intervals of every 40 cycles. (a) B0005, (b) B0006, (c) B0007, and (d) B0018.



**Figure 6.** Commencing from the 40th cycle, the alterations in current during the charging process are recorded at intervals of every 40 cycles. (a) B0005, (b) B0006, (c) B0007, and (d) B0018.



**Figure 7.** Commencing from the 40th cycle, the alterations in temperature during the charging process are recorded at intervals of every 40 cycles. (a) B0005, (b) B0006, (c) B0007, and (d) B0018.

### 2.3. Content-based Attention Mechanism

Content-based attention, also referred to as content-based addressing or general attention, is a form of attention mechanism that is frequently employed in sequence-to-sequence models and neural networks. This mechanism is designed to align input and output sequences by considering the content or meaning of the input sequence. In contrast to dot-product attention, which depends on a learned weight vector, content-based attention directly evaluates the content of each element in the input sequence against the current processing step. This process aids in the identification of elements within the sequence that exhibit semantic similarity or relevance to the current point of interest. The computation of attention weights involves a linear transformation utilizing a Linear layer on the attention input, followed by a non-linear activation function (tanh) and a softmax operation to yield normalized attention weights. These weights are then utilized to calculate a context vector through a weighted sum of the encoder outputs. The energy is calculated by applying a linear transformation to the attention inputs as shown in equation 7. Then the attention weights are obtained by applying the softmax function to the energy tensor along the sequence length dimension as shown in equation 8. The context vector is computed by element-wise multiplication of the attention weights with the attention inputs followed by summation along the sequence length dimensions shown in equation 9. Content-based attention enables the model to selectively concentrate on different segments of the input sequence, prioritizing relevant information while disregarding irrelevant or noisy input. Its application is prevalent in sequence-to-sequence models.

$$energy = \tanh(Linear(\text{attention\_inputs})) \tag{7}$$

$$\text{attention\_weights} = \text{softmax}(energy) \tag{8}$$

$$\text{context\_vector} = \sum_{i=1}^{seq\_length} \text{attention\_weights}_i \times \text{attention\_inputs}_i \tag{9}$$

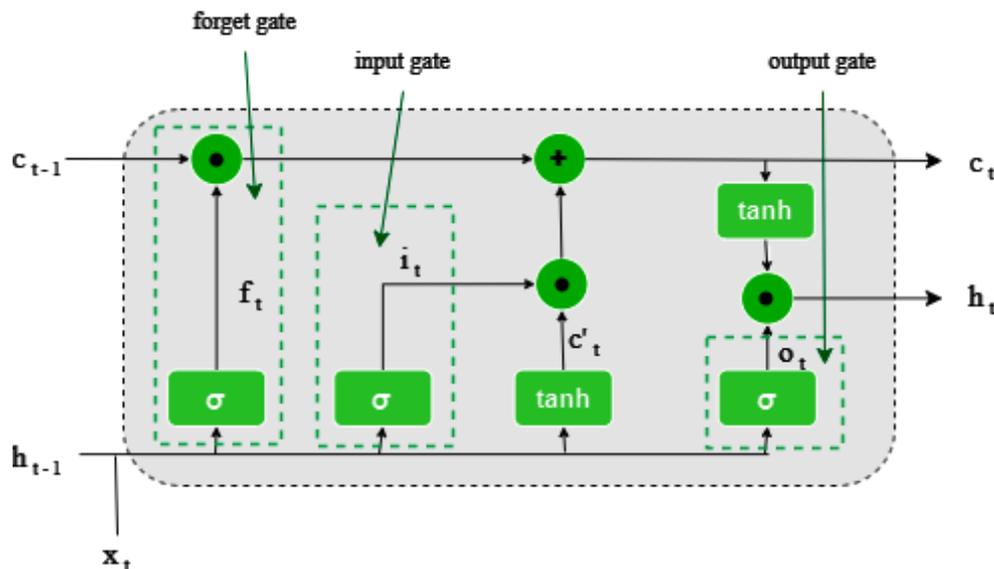


Figure 8. LSTM cell architecture.

### 3. The proposed model

The prediction of the RUL of Lithium-ion batteries is considered a supervised regression problem, involving the utilization of data gathered from different sensors for the purpose of training and evaluating a range of Deep Learning models. This paper introduces a novel DL model called At-LSTM, which integrates LSTM and attention mechanism, for the purpose of predicting the RUL of Lithium-ion batteries, as shown in Figure 9. The quantity of LSTM layers indicates the network's depth, which pertains to how many LSTM layers are arranged on top of each other to create the model. Each LSTM layer within the sequence processes the input sequentially, transmitting information through memory cells and gates. The output of a single LSTM layer functions as the input for the subsequent layer, enabling the model to acquire hierarchical representations of the input data. Incorporating multiple LSTM layers can empower the model to grasp intricate temporal relationships and develop more abstract representations of the data. In this paper, a network structure consisting of four LSTM layers is utilized to analyze the time sequence information extracted multiple times, allowing for a deep integration of the input sample data. Subsequent to the final LSTM layer, an attention mechanism layer is introduced to compute the neuron weights of the hidden state layer. This layer ultimately delivers the output through the LSTM, enabling the allocation of weight coefficients and data reconstruction to pinpoint the crucial aspects of the extracted features. The processed and hidden layer data are combined and forwarded to the fully connected (FC) layer for the ultimate RUL prediction. Lastly, Algorithm 1 outlines the pseudocode of the suggested model.

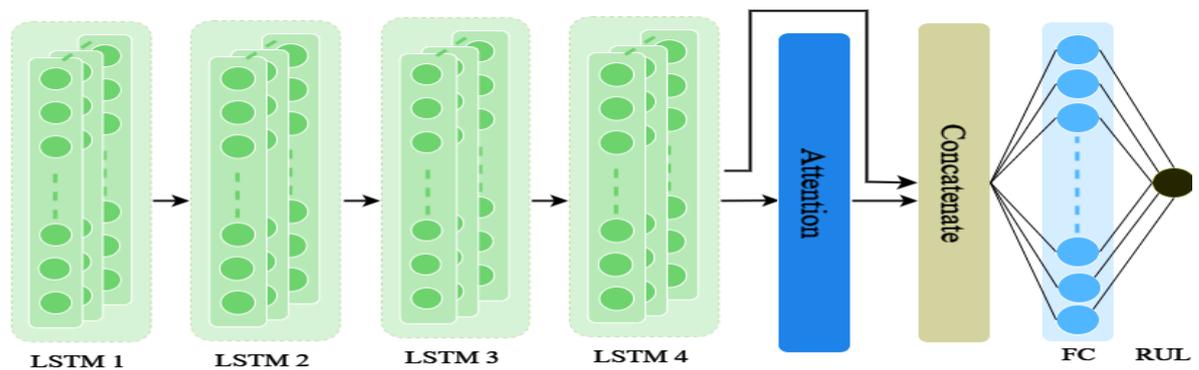


Figure 9. Flowchart of the proposed At-LSTM.

As outlined in Algorithm 1, the suggested framework processes the input data through a series of steps. Initially, the input data is preprocessed and then passed to the input layer, which in turn feeds it to a network architecture consisting of four LSTM layers with 512 neurons each, along with a Tanh activation function. This architecture is designed to analyze the temporal sequence information iteratively, enabling a comprehensive integration of the input data samples. Subsequently, the network's output is directed to the Content-based Attention Mechanism to calculate the weights of the neurons in the hidden state layer. The resulting attention output is then concatenated with the data from the hidden layer and input into a fully connected layer with a single neuron to forecast the remaining useful life of lithium-ion batteries.

---

#### Algorithm 1 Pseudo-code of At-LSTM

---

**Input:** Input data ( $D$ ), batch size ( $B_s$ ), maximum epoch ( $T$ ), and learning rate ( $lr$ )

**Output:**  $loss$  ( $MSE$ ),  $RMSE$

- 1: Conducting the **preprocessing step**
  - 2: **Input:** Construct an **input layer** to receive the input data  
/\* **Feature extraction based on the LSTM** \*/
  - 3:  $x$ : Add an **LSTM** layer with 512 units to  $x$ .
  - 4:  $x$ : Add an **LSTM** layer with 512 units to  $x$ .
  - 5:  $x$ : Add an **LSTM** layer with 512 units to  $x$ .
  - 6:  $x$ : Add an **LSTM** layer with 512 units to  $x$ .
  - 7: **attention\_output:** Add **Content-based Attention** to  $x$ .
  - 8:  $x$ : concatenate the attention\_output with  $x$ .  
/\* **Prediction Block** \*/
  - 9:  $x$ : Add a **Linear layer** with 1 node to  $x$ .  
/\* **Optimization process** \*/
  - 10:  $N = Size(D)/B_s$  /\* Estimate the number of batches \*/
  - 11:  $t = 0$ , Current epoch
  - 12: **while**  $t < T$
  - 13:      $i = 0$ , the current batch size.
  - 14:     **while**  $i < N$
  - 15:         Compute the **Score function** using the  $ith$  batch.  
           Update the weights based on the Adam to optimize the **score** function.
  - 16:          $i = i + 1$
  - 17:     **end while.**
  - 18:      $t = t + 1$
  - 19: **end while**
-

## 4. Experiments and Results

### 4.1. Data preprocessing

In data processing, the NASA dataset presents various features with disparate scales, causing detrimental impacts on the efficacy of Deep Learning (DL) models throughout the training phase. The dataset encompasses parameters such as voltage, current, temperature, and capacity obtained from batteries. Accordingly, normalization of these parameters is imperative to eliminate potential distortions and biases, thereby enhancing the accuracy of DL models. Min-max scaling is the technique employed to normalize the feature values. This scaling method holds significance in machine learning, especially when addressing features with diverse scales or units. It adjusts the data to a consistent range, typically from 0 to 1, maintaining the original distribution shape while standardizing the range [34]. Mathematically, min-max scaling is defined by equation 10. Where  $x_{i,j}$  donates the value of  $i$ th sample, and  $j$ th feature,  $x_{j \min}$ ,  $x_{j \max}$  donates the minimum, and maximum values in  $j$ th feature, respectively.

$$x'_{i,j} = \frac{x_{i,j} - x_{j \min}}{x_{j \max} - x_{j \min}} \quad (10)$$

The utilization of the sliding window technique is a foundational method employed in a variety of signal processing and machine learning scenarios, especially in the context of handling sequential data such as time series. This method entails dividing a continuous flow of data points into smaller, partially overlapping (or non-overlapping) segments for subsequent analysis. In this process, a sliding window is applied to partition the input data. if the  $d$  is the initial sliding window value, and the number of samples in the entire training data is  $n$ , then the vector  $x^{(i)}$  within the sliding window as follows:

$$x^{(i)} = [x_i, x_{i+1}, x_{i+2}, \dots, x_{i+d-1}], \quad i = 1 \rightarrow n - d \quad (11)$$

Upon determining the optimal embedding dimension  $d$ , the train set  $Data_{train}$  can be formulated accordingly. These sets are represented as:

$$Data_{train} = [(x^{(1)}, x_{d+1}), \dots, (x^{(i)}, x_{d+i}), \dots, (x^{(n-d)}, x_n)], i = 1 \rightarrow n - d \quad (12)$$

Where  $x_{d+i}$  is the label for input data, and the testing set follows the same technique.

### 4.2. Evaluation metrics

in this paper, The Adam [35] optimization algorithm and mean square error (MSE) loss computed using equation 13 are employed for the optimization of the network parameters. The introduction of root mean square error (RMSE) serves as a metric to evaluate the proposed model by comparing the RMSE values between actual and predicted labels of all instances in the dataset. The RMSE value is derived through mathematical computation as described in equation 14. where  $N$  represents the number of samples, and  $y_i$  and  $y'_i$  represents the true and predicted labels of the  $i$ th sample, respectively. Both MSE and RMSE should be minimized to enhance the accuracy of predicting the Remaining Useful Life (RUL) of Lithium-ion batteries.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2 \quad (13)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2} \quad (14)$$

### 4.3. Hyperparameter tuning

The At-LSTM model proposed in this paper involves several hyper-parameters, including learning rate and window size, no. of LSTM layers, which must be precisely determined to optimize its efficiency and reduce the RMSE. Consequently, a series of experiments are carried out in this research, exploring different settings for each parameter to identify the optimal values that lead to a considerable enhancement in the model's performance as shown in Table 2. For instance, the model's performance is impacted by the quantity of hidden units present in each LSTM layer. Consequently, numerous experiments were conducted to determine the optimal number of hidden dimensions for the LSTM layer, ranging from 64 to 512. The outcomes of these experiments indicate that the most suitable number of hidden dimensions for the LSTM layer is 512. Similarly, a series of experiments were carried out using window sizes of 8, 16, 32, and 64 to determine the optimal window size. The results of these experiments indicate that a window size of 64 is the most suitable for the given dataset.

**Table 2.** Experimental Analysis of the influence of parameters on prediction results.

		<b>B0005</b>	<b>B0006</b>	<b>B0007</b>	<b>B0018</b>
<b>Hidden dim of LSTM</b>	64	0.0163	0.0593	0.0165	0.0229
	128	0.0156	0.0314	0.0173	0.0235
	256	0.0188	0.0472	0.0721	0.0221
	<b>512</b>	<b>0.0146</b>	<b>0.0124</b>	<b>0.0121</b>	<b>0.0214</b>
<b>Window size</b>	8	0.0428	0.0995	0.0204	0.0220
	16	0.0402	0.0315	0.0146	0.0224
	32	0.0154	0.0157	0.0132	0.0219
	<b>64</b>	<b>0.0146</b>	<b>0.0124</b>	<b>0.0121</b>	<b>0.0214</b>
<b>No. of LSTM</b>	3	0.0199	0.0385	0.0160	0.0223
	<b>4</b>	<b>0.0146</b>	<b>0.0124</b>	<b>0.0121</b>	<b>0.0214</b>
	5	0.0201	0.0399	0.0187	0.0312
<b>Learning rate</b>	0.0001	0.0168	0.0480	0.0742	0.0448
	<b>0.001</b>	<b>0.0146</b>	<b>0.0124</b>	<b>0.0121</b>	<b>0.0214</b>
	0.002	0.0176	0.0315	0.0236	0.0215
	0.01	0.0212	0.0310	0.0354	0.0268

The deeper networks pose a risk of overfitting and increased computational complexity, necessitating several experiments to determine the optimal number of LSTM layers. experimental findings suggest that a network with 4 LSTM layers exhibits the highest performance as shown in Table 2. The learning rate is a crucial factor in the training phase and can influence the speed at which convergence occurs, the effectiveness of the model, and the overall stability of the training process. experiments indicate that a learning rate of 0.001 yields the most optimal performance. The final hyperparameters of the proposed model are detailed in Table 3.

### 4.4. Ablation Experiments

In the conducted ablation experiment, an analysis is carried out on the efficacy of a composite model structure that incorporates both the LSTM and Content-based Attention Mechanism models. The primary aim is to scrutinize the influence of each element on forecasting RUL of Lithium-ion

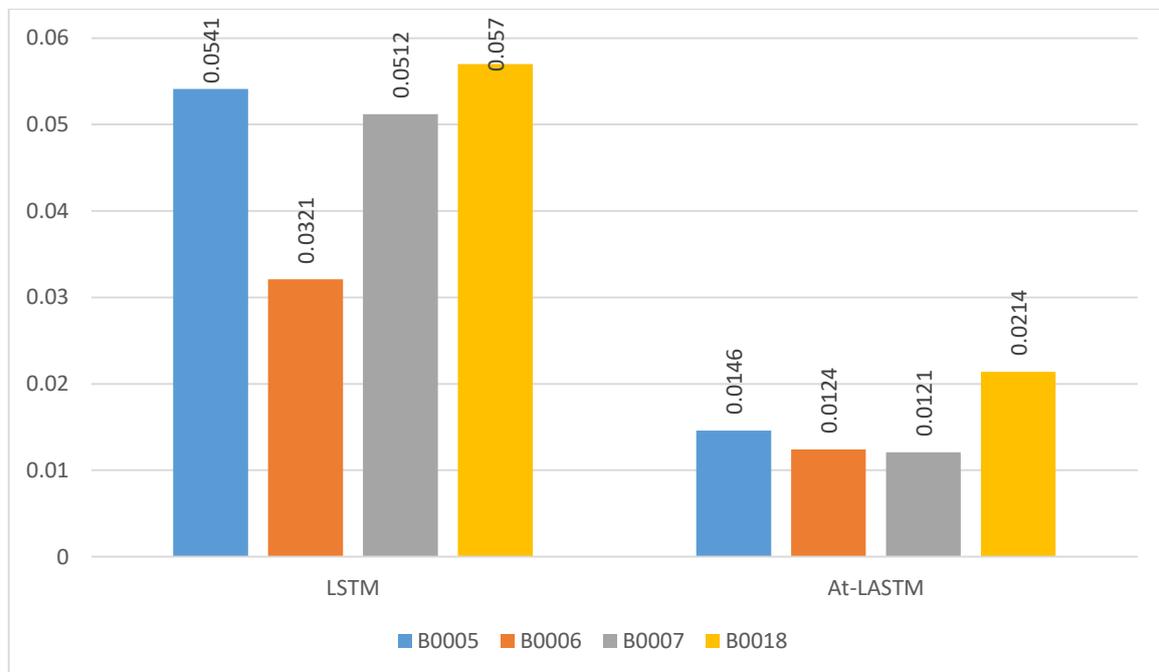
batteries and to gauge the collective enhancement in performance resulting from their amalgamation. The assessment will rely on the metrics of Root Mean Squared Error (RMSE). An analysis is conducted to determine how the attention mechanism affects the performance of the model. Results are presented in a table showing prediction errors for both LSTM and At-LSTM. Notably, the model performs worst in terms of prediction errors across all batteries when the attention mechanism is not utilized, with RMSEs of 0.0541, 0.0321, 0.0512, and 0.0570, respectively. All results are detailed in Table 4 and illustrated in Figure 10. The incorporation of the attention mechanism embedding leads to a significant reduction in errors. Experimental findings demonstrate that the inclusion of an attention mechanism substantially improves the model's predictive capability. Moreover, the performance of the method gradually improves as the attention mechanism is strengthened, providing further evidence of its effectiveness.

**Table 3.** The At-LSTM hyperparameters.

Parameter	value
No. of LSTM layer	4
Hidden dim of LSTM	512
Window size	64
Learning rate	0.001
Max no. of epoch	1000
Loss	MSE
Optimizer	Adam

**Table 4.** The results of the ablation study.

	B0005	B0006	B0007	B0018
<b>LSTM</b>	0.0541	0.0321	0.0512	0.0570
<b>LSTM+Attention (At-LASTM)</b>	<b>0.0146</b>	<b>0.0124</b>	<b>0.0121</b>	<b>0.0214</b>



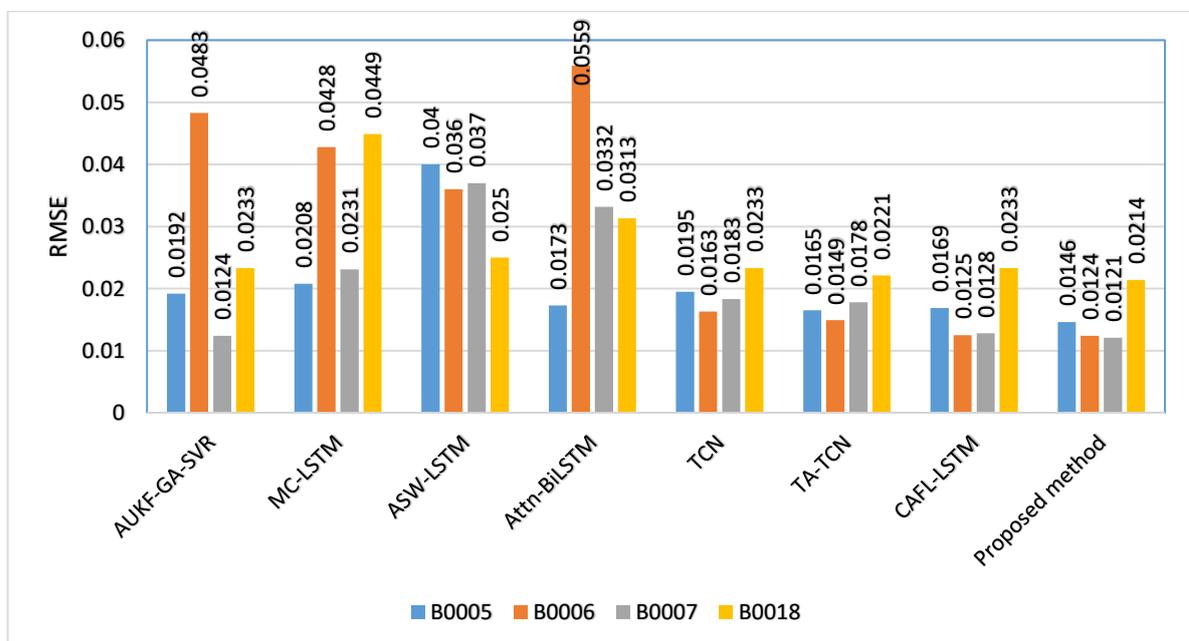
**Figure 10.** The representation of RMSE values acquired through ablation experiments.

## 5. Applications

This section presents the results achieved by the proposed At-LSTM model and various competing models for batteries B0005, B0006, B0007, and B0018 sourced from the NASA dataset. These results are evaluated using the RMSE metric to demonstrate the models' effectiveness in reducing the discrepancy between the predicted and target RUL. The outcomes of At-LSTM for batteries B0005, B0006, B0007, and B0018 sourced from the NASA dataset are compared with seven competing models to demonstrate their effectiveness and efficiency. A comprehensive comparison is conducted between the outcomes of At-LSTM and several rival models, including AUKF-GA-SVR [36], MC-LSTM [29], ASW-LSTM [37], Attn-BiLSTM [38], TCN [39], TA-TCN [39], and CAFL-LSTM [40], to showcase its superior performance. These results are displayed in the RMSE values as indicated in the Table 5.

**Table 5.** The precise specifications of the chosen lithium-ion batteries from NASA.

	B0005	B0006	B0007	B0018
AUKF-GA-SVR	0.0192	0.0483	0.0124	0.0233
MC-LSTM	0.0208	0.0428	0.0231	0.0449
ASW-LSTM	0.04	0.036	0.037	0.025
Attn-BiLSTM	0.0173	0.0559	0.0332	0.0313
TCN	0.0195	0.0163	0.0183	0.0233
TA-TCN	0.0165	0.0149	0.0178	0.0221
CAFL-LSTM	0.0169	0.0125	0.0128	0.0233
Proposed method	<b>0.0146</b>	<b>0.0124</b>	<b>0.0121</b>	<b>0.0214</b>



**Figure 11.** illustrates the representation of RMSE values acquired from different models.

The superior outcomes are highlighted in bold. At-LSTM surpasses the other approaches in terms of the RMSE metric. The table illustrates that At-LSTM may outperform all the models considered in relation to the RMSE for the selected batteries, achieving RMSE values of 0.0146, 0.0124,

0.0121, and 0.0214 for B0005, B0006, B0007, and B0018, respectively, which significantly outperforms the comparison methods under similar circumstances. Upon juxtaposition of our results with the leading outcomes attained by different mentioned models, our suggested model showcases a decrease in RMSE by 13.6%, 0.8%, 5.4%, and 3.17% for NASA lithium-ion batteries B0005, B0006, B0007, and B0018, respectively. This proposed model is viewed as a robust solution for addressing this issue due to its potential to excel in the RMSE metric, which assigns equal importance to both early and late predictions. To visually exhibit the superiority of the proposed model, Figure 11 is included to display the RMSE values obtained by different algorithms for B0005, B0006, B0007, and B0018 batteries.

## 6. Conclusions

Precise prediction of the remaining useful life (RUL) of lithium-ion batteries is essential for ensuring the safe and dependable functionality of batteries while also mitigating safety hazards. A novel approach called At-LSTM is suggested for predicting the RUL of lithium-ion batteries. The At-LSTM model is based on LSTM and content-based attention mechanism, LSTMs are utilized to capture and retain patterns across extended sequences, rendering them suitable for representing intricate temporal connections in time series datasets. After this is the integration of an attention mechanism designed to synchronize input and output sequences by considering the content or meaning of the input sequence. Ultimately, the final prediction results are produced through a Fully Connected (FC) layer. We opted to utilize the NASA lithium battery dataset for our experimental investigations. Upon comparison of our findings with the top results achieved by various models cited in the references, our proposed model demonstrates a reduction in RMSE of 13.6%, 0.8%, 5.4%, and 3.17% for NASA lithium-ion batteries B0005, B0006, B0007, and B0018, correspondingly.

## Funding

This research received no external funding.

## Ethical approval

This research does not contain any studies with human participants or animals performed by any of the authors.

## Conflicts of Interest

The authors declare that there is no conflict of interest in the research.

## Institutional Review Board Statement

Not applicable.

## Informed Consent Statement

Not applicable.

## Data Availability Statement

Not applicable.

## References

- [1] Xiong, J., & Xu, D. (2021). Relationship between energy consumption, economic growth and environmental pollution in China. *Environmental Research*, 194, 110718. <https://doi.org/10.1016/j.envres.2021.110718>
- [2] Wei, Z., Zhao, J., Xiong, R., Dong, G., Pou, J., & Tseng, K. J. (2018). Online estimation of power capacity with noise effect attenuation for lithium-ion battery. *IEEE Transactions on Industrial Electronics*, 66(7), 5724-5735. <https://doi.org/10.1109/TIE.2018.2878122>
- [3] Lyu, Z., Gao, R., & Li, X. (2021). A partial charging curve-based data-fusion-model method for capacity estimation of Li-Ion battery. *Journal of Power Sources*, 483, 229131. <https://doi.org/10.1016/j.jpowsour.2020.229131>
- [4] Zheng, S., Zhu, X., Xiang, Z., Xu, L., Zhang, L., & Lee, C. H. (2022). Technology trends, challenges, and opportunities of reduced-rare-earth PM motor for modern electric vehicles. *Green Energy and Intelligent Transportation*, 1(1), 100012. <https://doi.org/10.1016/j.geits.2022.100012>
- [5] Tang, X., Liu, K., Li, K., Widanage, W. D., Kendrick, E., & Gao, F. (2021). Recovering large-scale battery aging dataset with machine learning. *Patterns*, 2(8). <https://doi.org/10.1016/j.patter.2021.100302>
- [6] He, H., Sun, F., Wang, Z., Lin, C., Zhang, C., Xiong, R., . . . Zhang, S. (2022). China's battery electric vehicles lead the world: Achievements in technology system architecture and technological breakthroughs., 2022, 1. *Green Energy and Intelligent Transportation*, 100020. <https://doi.org/10.1016/j.geits.2022.100020>
- [7] Li, S., Zhang, C., Du, J., Cong, X., Zhang, L., Jiang, Y., & Wang, L. (2022). Fault diagnosis for lithium-ion batteries in electric vehicles based on signal decomposition and two-dimensional feature clustering. *Green Energy and Intelligent Transportation*, 1(1), 100009. <https://doi.org/10.1016/j.geits.2022.100009>
- [8] Ren, L., Dong, J., Wang, X., Meng, Z., Zhao, L., & Deen, M. J. (2020). A data-driven auto-CNN-LSTM prediction model for lithium-ion battery remaining useful life. *IEEE Transactions on Industrial Informatics*, 17(5), 3478-3487. <https://doi.org/10.1109/TII.2020.3008223>
- [9] Wang, Y., Tian, J., Sun, Z., Wang, L., Xu, R., Li, M., & Chen, Z. (2020). A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems. *Renewable and Sustainable Energy Reviews*, 131, 110015. <https://doi.org/10.1016/j.rser.2020.110015>
- [10] Chen, L., Zhang, Y., Zheng, Y., Li, X., & Zheng, X. (2020). Remaining useful life prediction of lithium-ion battery with optimal input sequence selection and error compensation. *Neurocomputing*, 414, 245-254. <https://doi.org/10.1016/j.neucom.2020.07.081>
- [11] Li, X., Ma, Y., & Zhu, J. (2021). An online dual filters RUL prediction method of lithium-ion battery based on unscented particle filter and least squares support vector machine. *Measurement*, 184, 109935. <https://doi.org/10.1016/j.measurement.2021.109935>
- [12] Yu, H., Meng, L., & Downey, A. (2021). Physics-based prognostics of implantable-grade lithium-ion battery for remaining useful life prediction [J]. *Journal of Power Sources*, 485(6273), 229327. <https://doi.org/10.1016/j.jpowsour.2020.229327>
- [13] Sun, T., Wu, R., Cui, Y., & Zheng, Y. (2021). Sequent extended Kalman filter capacity estimation method for lithium-ion batteries based on discrete battery aging model and support vector machine. *Journal of Energy Storage*, 39, 102594. <https://doi.org/10.1016/j.est.2021.102594>
- [14] Wei, Z., Zhao, J., Ji, D., & Tseng, K. J. (2017). A multi-timescale estimator for battery state of charge and capacity dual estimation based on an online identified model. *Applied energy*, 204, 1264-1274. <https://doi.org/10.1016/j.apenergy.2017.02.016>
- [15] Hu, X., Jiang, J., Cao, D., & Egardt, B. (2015). Battery health prognosis for electric vehicles using sample entropy and sparse Bayesian predictive modeling. *IEEE Transactions on Industrial Electronics*, 63(4), 2645-2656. <https://doi.org/10.1109/TIE.2015.2461523>
- [16] Sadabadi, K. K., Jin, X., & 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. *Journal of Power Sources*, 481, 228861. <https://doi.org/10.1016/j.jpowsour.2020.228861>
- [17] Kim, S., Park, H. J., Choi, J.-H., & Kwon, D. (2020). A novel prognostics approach using shifting kernel particle filter of Li-ion batteries under state changes. *IEEE Transactions on Industrial Electronics*, 68(4), 3485-3493. <https://doi.org/10.1109/TIE.2020.2978688>

- [18] Chen, L., An, J., Wang, H., Zhang, M., & Pan, H. (2020). Remaining useful life prediction for lithium-ion battery by combining an improved particle filter with sliding-window gray model. *Energy Reports*, 6, 2086-2093. <https://doi.org/10.1016/j.egy.2020.07.026>
- [19] Jiao, R., Peng, K., & Dong, J. (2020). Remaining useful life prediction of lithium-ion batteries based on conditional variational autoencoders-particle filter. *IEEE Transactions on Instrumentation and Measurement*, 69(11), 8831-8843. <https://doi.org/10.1109/TIM.2020.2996004>
- [20] Patil, M. A., Tagade, P., Hariharan, K. S., Kolake, S. M., Song, T., Yeo, T., & Doo, S. (2015). A novel multistage Support Vector Machine based approach for Li ion battery remaining useful life estimation. *Applied energy*, 159, 285-297. <https://doi.org/10.1016/j.apenergy.2015.08.119>
- [21] Zhang, C., He, Y., Yuan, L., Xiang, S., & Wang, J. (2015). Prognostics of lithium-ion batteries based on wavelet denoising and DE-RVM. *Computational intelligence and neuroscience*, 2015, 14-14. <https://doi.org/10.1155/2015/918305>
- [22] Liu, J., & Chen, Z. (2019). Remaining useful life prediction of lithium-ion batteries based on health indicator and Gaussian process regression model. *Ieee Access*, 7, 39474-39484. <https://doi.org/10.1109/ACCESS.2019.2905740>
- [23] Ding, P., Liu, X., Li, H., Huang, Z., Zhang, K., Shao, L., & Abedinia, O. (2021). Useful life prediction based on wavelet packet decomposition and two-dimensional convolutional neural network for lithium-ion batteries. *Renewable and Sustainable Energy Reviews*, 148, 111287. <https://doi.org/10.1016/j.rser.2021.111287>
- [24] Zhao, S., Zhang, C., & Wang, Y. (2022). Lithium-ion battery capacity and remaining useful life prediction using board learning system and long short-term memory neural network. *Journal of Energy Storage*, 52, 104901. <https://doi.org/10.1016/j.est.2022.104901>
- [25] Chen, D., Hong, W., & Zhou, X. (2022). Transformer network for remaining useful life prediction of lithium-ion batteries. *Ieee Access*, 10, 19621-19628. <https://doi.org/10.1109/ACCESS.2022.3151975>
- [26] Catelani, M., Ciani, L., Fantacci, R., Patrizi, G., & Picano, B. (2021). Remaining useful life estimation for prognostics of lithium-ion batteries based on recurrent neural network. *IEEE Transactions on Instrumentation and Measurement*, 70, 1-11. <https://doi.org/10.1109/TIM.2021.3111009>
- [27] Zhang, Q., Yang, L., Guo, W., Qiang, J., Peng, C., Li, Q., & Deng, Z. (2022). A deep learning method for lithium-ion battery remaining useful life prediction based on sparse segment data via cloud computing system. *Energy*, 241, 122716. <https://doi.org/10.1016/j.energy.2021.122716>
- [28] Chinomona, B., Chung, C., Chang, L.-K., Su, W.-C., & Tsai, M.-C. (2020). Long short-term memory approach to estimate battery remaining useful life using partial data. *Ieee Access*, 8, 165419-165431. <https://doi.org/10.1109/ACCESS.2020.3022505>
- [29] Park, K., Choi, Y., Choi, W. J., Ryu, H.-Y., & Kim, H. (2020). LSTM-based battery remaining useful life prediction with multi-channel charging profiles. *Ieee Access*, 8, 20786-20798. <https://doi.org/10.1109/ACCESS.2020.2968939>
- [30] Li, X., Zhang, L., Wang, Z., & 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. *Journal of Energy Storage*, 21, 510-518. <https://doi.org/10.1016/j.est.2018.12.011>
- [31] Saha, B., & Goebel, K. (2022). Nasa ames prognostics data repository: battery data set. Online], <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository>.
- [32] Zhao, L., Song, S., Wang, P., Wang, C., Wang, J., & Guo, M. (2024). A MLP-Mixer and mixture of expert model for remaining useful life prediction of lithium-ion batteries. *Frontiers of Computer Science*, 18(5), 185329. <https://doi.org/10.1007/s11704-023-3277-4>
- [33] Lin, C.-P., Cabrera, J., Yang, F., Ling, M.-H., Tsui, K.-L., & Bae, S.-J. (2020). Battery state of health modeling and remaining useful life prediction through time series model. *Applied energy*, 275, 115338. <https://doi.org/10.1016/j.apenergy.2020.115338>
- [34] Ali, P. J. M. (2022). Investigating the Impact of min-max data normalization on the regression performance of K-nearest neighbor with different similarity measurements. *ARO-The Scientific Journal of Koya University*, 10(1), 85-91. <https://doi.org/10.14500/aro.10955>

- [35] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980. <https://doi.org/10.48550/arXiv.1412.6980>
- [36] Xue, Z., Zhang, Y., Cheng, C., & Ma, G. (2020). Remaining useful life prediction of lithium-ion batteries with adaptive unscented kalman filter and optimized support vector regression. *Neurocomputing*, 376, 95-102. <https://doi.org/10.1016/j.neucom.2019.09.074>
- [37] Wang, Z., Liu, N., & Guo, Y. (2021). Adaptive sliding window LSTM NN based RUL prediction for lithium-ion batteries integrating LTSA feature reconstruction. *Neurocomputing*, 466, 178-189. <https://doi.org/10.1016/j.neucom.2021.09.025>
- [38] Zhang, Z., Zhang, W., Yang, K., & Zhang, S. (2022). Remaining useful life prediction of lithium-ion batteries based on attention mechanism and bidirectional long short-term memory network. *Measurement*, 204, 112093. <https://doi.org/10.1016/j.measurement.2022.112093>
- [39] Li, L., Li, Y., Mao, R., Li, L., Hua, W., & Zhang, J. (2023). Remaining useful life prediction for lithium-ion batteries with a hybrid model based on TCN-GRU-DNN and dual attention mechanism. *IEEE Transactions on Transportation Electrification*. <https://doi.org/10.1109/TTE.2023.3247614>
- [40] Wang, Y., Wang, S., Fan, Y., Xie, Y., Hao, X., & Guerrero, J. M. (2024). High-precision collaborative estimation of lithium-ion battery state of health and remaining useful life based on call activation function library-long short term memory neural network algorithm. *Journal of Energy Storage*, 83, 110749. <https://doi.org/10.1016/j.est.2024.110749>

**Received:** 04 Jan 2024, **Revised:** 15 Mar 2024,

**Accepted:** 14 Apr 2024, **Available online:** 16 Apr 2024.



© 2024 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

**Disclaimer/Publisher's Note:** The perspectives, opinions, and data shared in all publications are the sole responsibility of the individual authors and contributors, and do not necessarily reflect the views of Sciences Force or the editorial team. Sciences Force and the editorial team disclaim any liability for potential harm to individuals or property resulting from the ideas, methods, instructions, or products referenced in the content.