Proceedings of the 2nd International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/67/20240691

Predictive State-of-Health (SOH) of lithium-ion batteries throughout their lifecycle based on LSTM neural networks

Hanchao Lv

School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai, China

2135061319@st.usst.edu.cn

Abstract. With the rapid development of new energy sources, the demand for energy storage systems in the power industry has surged. While lithium-ion battery energy storage systems have gained widespread market acceptance due to their performance advantages, concerns regarding battery safety and durability have also escalated. Therefore, accurately predicting the health status of batteries has become a crucial aspect in assessing their safety and durability. State of Health (SOH), as a vital indicator of battery health status, has become a focal point for many researchers aiming to enhance the accuracy of its prediction. Today, owing to the rapid advancements in machine learning and neural networks, data-driven approaches have gradually emerged as one of the primary methods for SOH prediction. This paper proposes a method based on Long Short-Term Memory (LSTM) neural networks, which combines theoretical analysis with neural network models, utilizing authentic battery aging data from NASA to predict the health status of lithium-ion batteries throughout their lifecycle. Experimental results demonstrate that the proposed method exhibits high accuracy.

Keywords: Lithium-ion Battery, State of Health(SOH), Long short-term memory, Aging Analysis.

1. Introduction

With the rapid development in areas such as electric transportation and renewable energy, there is an increasing societal demand for clean energy and sustainable development, consequently driving the need for high-performance, reliable, and long-lasting battery systems. Over time, batteries undergo aging due to environmental factors such as charge-discharge cycles and temperature variations [1]. Consequently, precise assessment of battery health throughout its lifecycle is crucial to ensure the longevity and reliable operation of battery systems.

The State of Health (SOH), as one of the most critical indicators for evaluating battery health, requires precise prediction to enable researchers to better optimize battery management systems and address issues such as battery aging and degradation. The proposition of battery SOH prediction aims to achieve real-time assessment of battery condition through monitoring and analyzing various physical and chemical processes within the battery, thereby delving into the understanding of battery aging mechanisms. This prediction not only aids in early detection of battery aging and performance degradation trends but also guides battery maintenance and management [2]. By taking timely measures,

© 2024 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/67/20240691

battery lifespan can be maximized, reducing system maintenance costs while ensuring sustainable development in electric transportation and renewable energy sectors.

Direct measurement methods of battery SOH involve various technologies. Firstly, the Electrochemical Impedance Spectroscopy (EIS) method measures the impedance spectrum of the battery at different frequencies, obtaining the complex impedance Z. The impedance spectrum comprises internal elements of the battery such as resistance(R) and capacitance(C). By analyzing the shape and characteristic frequencies of the impedance spectrum using equivalent circuit models like the Randles circuit, parameters of internal electrochemical processes can be resolved. These parameters are closely related to the battery's state, such as electrolyte conductivity and electrode surface area. By substituting these parameters into appropriate mathematical models, estimates of SOH can be obtained. Its advantages lie in non-invasiveness and real-time capability, although it requires specialized equipment and complex data analysis [3].

Traditional empirical methods are then utilized to predict the SOH of batteries. Monitoring the temperature variations within the battery during operation is employed, as the aging process typically accompanies internal energy loss and exothermic battery reactions. The relationship between battery SOH and temperature is commonly expressed through empirical formulas, estimating SOH by recording temperature changes under various operational conditions. Alternatively, the open circuit voltage (OCV) method assesses SOH by monitoring the battery's OCV when idle. This approach typically employs empirical formulas to estimate SOH by comparing measured OCV with reference values, thereby determining the battery's SOH percentage. However, this method may fail to accurately detect localized aging or non-uniformity within the battery, as it primarily relies on the overall average performance of the battery [4]. Moreover, its applicability may be limited for high-rate charge-discharge applications.

Model-based approaches are among the commonly used methods for estimating battery SOH, manifested mainly in electrochemical model methods, equivalent circuit model methods, and fusion model methods. Traditional electrochemical models are primarily based on battery reaction mechanisms, including porous electrode theory and concentrated solution theory models. These models employ partial differential equations to describe charge conservation in the solid phase or electrolyte [5]. However, the complexity of the equations and high computational costs limit their widespread use in practical applications. To reduce computational complexity, researchers have proposed simplified models such as polynomialized porous electrode models and single particle models. These models provide new insights into battery SOH prediction at different time scales, but their complexity also makes them difficult to adapt to various battery types and operating conditions [6].

Equivalent circuit models utilize circuit elements such as resistors and capacitors to describe the dynamic characteristics of batteries. Typically, they include models for internal resistance, resistance-capacitance, and PNGV. While widely employed in battery management systems due to their parsimonious parameters, equivalent circuit models may exhibit reduced robustness in parameter estimation. To enhance accuracy, researchers integrate Kalman filters, utilizing adaptive algorithms and multi-order Kalman filters for state parameter estimation. Despite their extensive application in battery management systems, equivalent circuit models may demonstrate diminished robustness in parameter estimation. This stems from oversimplification of the batteries' internal complex dynamic behaviors, overlooking crucial nonlinear and time-varying properties. This oversimplification may lead to diminished performance of the model under complex operating conditions [7]. Model fusion integrates the strengths of different models to improve the estimation capability of battery state parameters. For instance, utilizing particle swarm optimization algorithms for parameter optimization significantly enhances model performance. However, integrating multiple models introduces sources of errors, making it challenging to determine criteria for model failure.

In recent years, machine learning techniques have seen significant development. Compared to direct measurement methods and model-based prediction approaches, researchers have recognized the convenience of data-driven nature of machine learning methods and their substantial improvement in model universality, flexibility, and feature learning ability. They have applied these methods to address issues in battery SOH prediction. Supervised learning methods such as linear regression, random forests,

DOI: 10.54254/2755-2721/67/20240691

and neural networks require labeled datasets for training. Machine learning methods offer greater flexibility, adapting to various types and scales of battery systems. Their model structures and parameters can be adjusted according to real-world scenarios, enhancing their applicability in different application scenarios [8]. Feature learning automatically extracts key features of battery dynamics without the need for manually defining complex physical models. This makes the models more versatile, capable of handling complex nonlinear and non-stationary characteristics, significantly reducing constraints imposed by complex equations and high computational costs in practical applications.

With the maturity of machine learning methods and deeper research, constructing neural networks with complex network structures has become the primary means of estimating battery health status. Neural networks, through multi-level learning structures, can better capture the complex dynamic characteristics within battery systems. In SOH prediction, neural network-based processes generally include the input layer for data, feature learning in hidden layers, and SOH prediction in the output layer. Neural networks can learn abstract features of battery systems from vast amounts of input data, making the prediction process more intelligent and comprehensive.

In terms of SOH prediction throughout the battery's entire lifecycle, batteries undergo multiple charge-discharge cycles, each containing long-term dependencies such as capacity degradation and internal resistance changes during the cycling process. Meanwhile, batteries in practical applications are influenced by external environments, charging/discharging strategies, among other factors, leading to dynamic changes in their characteristics. Combining the aforementioned, the Long Short-Term Memory (LSTM) network's units enable the network to effectively capture and retain long-term dependencies between these recurrent cycles, providing an advantage for a better understanding of the temporal evolution of battery lifecycles [9]. Its gating mechanism allows the network to selectively update and forget information based on the dynamic changes in input data, thereby adapting better to the dynamic characteristics of battery systems.

SOH prediction requires continuous tracking of important information in time series to capture changes in system state trends. Compared to traditional short-term memory networks (SMNs) or other neural networks, the primary advantage of LSTM networks lies in their handling of time series information. The memory units inside LSTM networks can effectively preserve and transmit this time series information, ensuring accurate prediction of SOH changes throughout the battery's lifespan. This stands in contrast to traditional short-term memory networks or other neural networks, which may suffer from information loss when dealing with long-term cycle data.

2. Aging Mechanism

During the aging process, batteries undergo various complex changes, accompanied by capacity degradation. Understanding the mechanisms of battery aging is of paramount importance for research purposes. This section will provide a detailed explanation of the State of Health (SOH) definition and the aging mechanisms during charge and discharge cycles, laying the theoretical groundwork for extracting input features in subsequent sections.

2.1. Definition of State of Health

The State of Health (SOH) of lithium-ion batteries is a crucial parameter for assessing the overall health status of batteries, reflecting their aging degree and predicting their lifespan. SOH is defined based on the battery's capacity, a widely adopted metric, with its predictive expression as follows:

$$SOH = \frac{C_{aged}}{C_{new}} * 100\% \tag{4}$$

Here, the "rated capacity" denotes the specified capacity of the lithium-ion battery, while the "actual capacity" of the battery is determined by recording the charge and discharge capacities during the process after undergoing a series of charge and discharge cycles. The SOH value for a new battery is 100%, which gradually decreases with battery usage. In this study, when the SOH declines to 70%, it indicates that the lithium-ion battery's operational lifespan is nearing its end.

DOI: 10.54254/2755-2721/67/20240691

2.2. Aging Process

2.2.1. Formation and Growth of Anode SEI layer During Charge-Discharge Cycle Aging Process

During the charge-discharge cycling aging process, the formation and growth of the anode Solid Electrolyte Interphase (SEI) layer primarily result from the changes in the electrode/electrolyte interface due to reactions between the anode and electrolyte. Many researchers consider this as the main cause of electrode aging [10]. Since the working voltage of lithium-ion battery anode exceeds the electrochemical stability window of electrolyte components, the electrode/electrolyte interface undergoes electrolyte reduction during charging (polarization to low potential), accompanied by irreversible consumption of lithium ions. At the beginning of cycling, especially during the first cycle, decomposition products form a "protective layer" covering the electrode surface. The SEI layer formed on graphite surface is primarily classified into two types, depending on their distinct functions. At some sites on the graphite surface (i.e., at edge and basal plane defects), lithium ions enter/leave the graphite structure through intercalation/deintercalation processes. At these sites, the SEI layer must act as a Solid Electrolyte Interphase (SEI) [11]. The SEI layer exhibits unique characteristics, being permeable to lithium cations to a first approximation but relatively impermeable to other electrolyte components and electrons. Therefore, the SEI layer is generally believed to serve two purposes: (i) protecting electrolyte compounds from further reduction and (ii) safeguarding charging electrodes from corrosion. The protective layer formed on graphite surface, where no process of lithium-ion entry/exit from the graphite structure occurs (i.e., at basal plane surface), lacks SEI functionality in the above sense.

Due to the SEI layer does not exhibit a transfer number of 1 (= 1) for lithium cations, thus lacking the characteristics of a true lithium cation solid electrolyte. Other charged species (such as anions, electrons, dissolved cations) and neutral species (such as solvents, impurities) appear to still be capable of diffusing/migrating through the SEI layer. Furthermore, migration of solvated lithium ions and other electrolyte components may still occur even in later stages, as irreversible charge capacity and self-discharge reactions can still be observed after extended cycling periods. Consequently, corrosion and electrolyte decomposition persist throughout the battery's lifespan, leading to electrolyte loss and further SEI formation, albeit to a lesser extent and at a slower rate compared to the initial cycles. On a longer time scale, the SEI layer permeates into the pores of the electrode and may also infiltrate the pores of the separator membrane, potentially reducing the accessible active surface area of the electrode. Many researchers have measured the increase in electrode impedance, directly linked to power attenuation of the battery [12]. The rise in electrode impedance is believed to stem from the growth of SEI and its changes in composition and morphology, although SEI formation predominantly occurs in the initial few charge-discharge cycles, its transformation, stabilization, and growth also continue throughout further cycles and storage processes.

Given that SEI comprises electrolyte decomposition products, it naturally follows that the characteristics of SEI are greatly dependent on the composition of the electrolyte employed. Formation of a beneficial SEI, characterized by low irreversible capacity, stable SEI, and low self-discharge rates, occurs when electrolyte components conducive to SEI formation are present, thus also retarding SEI aging [13]. In essence, electrolyte solvents and additives with membrane-forming capabilities should not only support the formation process of SEI (during initial cycles) but also bolster its aging process (during extended cycling and storage periods).

2.2.2. Influence of Positive Electrode SEI Layer on Battery Capacity

Throughout the charge-discharge cycles of a battery, electrode aging, increased internal resistance, power decay, and capacity attenuation arise from various complex mechanisms. Material parameters as well as storage and cycling conditions all affect the battery's lifespan and performance. During lithium insertion and extraction processes, the electrode experiences volume expansion and contraction, causing stress on spherical materials; during aging, some particles may fracture or become isolated, resulting in a reduction in lithium capacity. Additionally, side reactions occur during electrode aging, dissolving into the electrolyte, thereby reducing its own capacity. Ions dissolved due to side reactions further

DOI: 10.54254/2755-2721/67/20240691

deposit onto the negative electrode, catalyzing the growth of the negative SEI membrane and accelerating battery aging [14]. For instance, Li_yCoO_2 reversible phase transitions occur in the electrode during normal lithium insertion/extraction processes, but when y<0.5, the material undergoes irreversible crystalline transformation, losing its lithium insertion/extraction ability. All these mechanisms contribute to a decrease in electrode capacity.

Depending on the battery's chemical system, both high and low charge states can lead to performance degradation and reduced battery capacity. Decay accelerates at high temperatures, but low temperatures, especially during the charging process, can also have adverse effects. Among material parameters, surface chemistry plays a crucial role in both negative and positive electrode materials. On the positive electrode, phase transitions and structural changes of bulk materials strongly influence aging, whereas changes in negative electrode bulk materials are considered to have minor significance. Typically, the formation of the anode SEI is accompanied by the release of electrolyte decomposition products. The irreversible charge capacity consumed during SEI formation depends on the graphite's specific surface area and layer formation conditions [15].

2.2.3. Formation of Negative Electrode SEI Film and Its Impact on Battery Capacity

For the negative electrode, almost all negatives are carbon materials. During the initial few cycles of the battery, a SEI film forms on the surface of active particles, increasing the negative electrode resistance. This process benefits by forming a protective layer that can prevent further electrode alteration. However, due to about 10% volume change in the negative electrode during lithium insertion/extraction, the SEI film may rupture, and a new SEI film continues to grow. Similar to positive electrode materials, negative electrode materials also degrade, such as stress on the materials during lithium insertion/extraction, leading to possible particle fracture or local isolation.

When the resistance of the negative electrode SEI film to lithium-ion insertion/extraction is significant, the negative electrode resistance increases, and lithium ions in the electrolyte decrease. The generation of the SEI film leads to a decrease in the negative electrode's conductivity and an increase in the liquid-phase ohmic resistance. Moreover, aging resistance of positive electrodes, current collectors, and other components also increases. These factors result in an overall increase in the battery's polarization voltage, causing the battery to reach charge-discharge cutoff points earlier and thus reducing its charge-discharge capacity [16].

Simultaneously, the process of SEI film generation consumes a large number of lithium ions, and in the initial cycling stages, the growth rate of the SEI film is relatively fast, which then becomes relatively steady. Additionally, at high charging rates or low temperatures, lithium plating may occur on the negative electrode, forming dendrites, resulting in lithium-ion loss. Furthermore, some side reaction products may consume lithium ions and generate deposits. The decrease in lithium ions in the electrode leads to a reduction in the transfer of lithium ions between positive and negative electrodes, directly resulting in a decrease in battery capacity [17]. Moreover, the reduction of lithium ions in the solution reduces the electrolyte's conductivity, increasing the polarization voltage during charge-discharge processes, thereby reducing the available capacity of the battery.

3. Long Short-Term Memory (LSTM) Neural Networks

3.1. Basic Introduction to LSTM Neural Networks

In order to handle time-series data, a type of neural network with a recurrent structure called Recurrent Neural Network (RNN) has been proposed. RNNs are widely used for processing sequential data. However, traditional RNNs often encounter the vanishing or exploding gradient problem when dealing with long sequential data, making it difficult to effectively capture long-term dependencies. This limitation restricts the performance of RNNs in tasks involving time-series and natural language processing.

To address this issue, Sepp Hochreiter and Jürgen Schmidhuber introduced the Long Short-Term Memory (LSTM) neural network in 1997. As an improved form of RNN, LSTM overcomes the

DOI: 10.54254/2755-2721/67/20240691

shortcomings of traditional RNNs by introducing structures such as memory units, input gates, forget gates, and output gates, especially making significant progress in handling long-term dependencies. In current research, LSTM has been proven to be an effective tool for processing sequence data, time-series analysis, and natural language processing tasks.

Researchers have successfully applied LSTM networks in various fields. Taking natural language processing as an example, Hochreiter and Schmidhuber focused on discussing its application in character recognition tasks in the original paper introducing LSTM [18]. Additionally, in 2013, Alex Graves and others introduced successful applications of LSTM in handwritten recognition and text generation. In the field of speech recognition, Awni Hannun and colleagues used LSTM for speech recognition in 2014, further confirming its superior performance in handling sequential data [19].

The importance of LSTM in neural network deep learning lies in its ability to effectively capture and learn long-term dependencies. Compared to traditional RNNs, LSTM introduces structures such as memory units, input gates, forget gates, and output gates, enabling it to better handle and maintain information in sequences. This makes LSTM networks excel in handling long sequences, long-term memory, and avoiding gradient vanishing problems, making it one of the preferred models for processing time-series data and natural language processing tasks.

3.2. Principles Explanation

Long Short-Term Memory (LSTM) neural networks effectively address the vanishing and exploding gradient problems encountered in traditional RNNs when processing long sequential data by introducing memory units and gate structures. Its principles mainly include input gates, forget gates, output gates, and memory units.

The forward propagation process is outlined as follows:

(1) Input Gate:

The input gate determines which information will be incorporated into the memory cell update at the current time step. The output of the input gate is obtained by operating on the current input and the previous time step's hidden state, expressed as follows:

$$i_t = \sigma(W_{ii}{*}x_t + b_{ii} + W_{hi}{*}h_{t-1} + b_{hi}) \tag{5}$$

Here, σ represents the sigmoid activation function, W_{ii} and W_{hi} are weight matrices, b_{ii} and b_{hi} is the bias term.

(2) Forget Gate:

The forget gate decides which past memories need to be forgotten at the current time step. The calculation of the output f_t of the forget gate is as follows:

$$f_t = \sigma(W_{if}{*}x_t + b_{if} + W_{hf}{*}h_{t-1} + b_{hf}) \tag{6}$$

(3) Output Gate:

The output gate determines the output of the hidden state at the current time step. The computation of the output of the output gate is as follows:

$$o_t = \sigma(W_{io}{*}x_t + b_{io} + W_{ho}{*}h_{t-1} + b_{hf}) \tag{7}$$

Simultaneously, the hidden state at the current time step is obtained through the control of the memory cell state by the output gate:

$$h_t = o_t \otimes tanh(C_t) \tag{8}$$

(4) Cell State (Updating Cell State):

The core of LSTM is the memory cell, which enables the network to maintain and update information when processing sequences. The state of the memory cell represents the network's memory of past information, and through gate control, memories can be selectively updated or retained.

The updating rule of the memory cell can be represented by the following formula:

DOI: 10.54254/2755-2721/67/20240691

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \tilde{C}_t \tag{9}$$

Here, C_t is the current memory cell state, f_t is the output of the forget gate, i_t is the output of the input gate, and $\tilde{C}_t$ is the current candidate memory cell state.

Through the collaboration of these gates, the LSTM network can selectively update and retain information when processing sequence data, thereby better capturing long-term dependencies, avoiding the vanishing and exploding gradient problems, and achieving outstanding performance in various sequence modeling tasks. This paper also employs the Adam optimizer to update the weights of the LSTM neural network during the training process to minimize the loss function. Adam (Adaptive Moment Estimation) is an optimization algorithm with adaptive learning rates, which combines momentum and adaptive learning rate concepts and effectively handles sparse gradients and non-stationary objectives. Section 3.4 of this paper will explain the specific parameter settings in the Adam optimizer.

3.3. Feature Extraction

Building upon the mechanistic analysis presented in Section 2, this section elucidates the aging mechanisms and degradation characteristics generation process of lithium-ion batteries (LIBs) to investigate the aging and inconsistent characteristics of LIB packs. Robust features, reflecting the inherent inconsistencies of each cell, are extracted as input features for the neural network through mechanistic analysis.

As elucidated in Section 2.2, during the charge-discharge cycling aging process of batteries, the formation and growth of the Solid Electrolyte Interphase (SEI) layer at the positive electrode-electrolyte interface represent a pivotal mechanism that significantly influences battery performance and lifespan. Particularly in the initial stages of cycling, notably within the first charge-discharge cycle, interactions between the electrode and electrolyte induce changes at the electrode/electrolyte interface, leading to the formation of the SEI layer.

Primarily, it is imperative to comprehend that the formation of the SEI layer arises from reactions between electrode materials and electrolytes. In lithium-ion batteries, under charging conditions, reactions may occur that reduce the electrolyte, accompanied by irreversible lithium-ion consumption. This process is particularly pronounced during initial cycling, resulting in decomposition byproducts forming a protective SEI layer covering the electrode surface. The formation of the SEI layer not only occurs in the early stages of battery cycling but undergoes various changes throughout the battery's lifespan. Initially, its formation rate is rapid, whereas during prolonged cycling and storage, the SEI layer may undergo transformation, stabilization, and growth [20].

Variations in charge and discharge rates directly influence the formation and growth of the SEI layer. Higher charging rates may expedite SEI layer formation, as increased electrode-electrolyte reactions during battery charging prompt more dissolution products to form the SEI layer. This may lead to greater irreversible capacity loss, accelerating battery aging. Conversely, lower charging rates may decelerate SEI layer growth due to relatively fewer electrolyte decomposition reactions. While this may decelerate battery aging, a balance between charging rate and battery performance needs to be struck in practical applications.

Additionally, discharge rate may impact the stability of the SEI layer. At higher discharge rates, increased stress on the electrode surface and electrolyte decomposition may intensify, resulting in more pronounced SEI layer alterations.

Furthermore, the formation and growth of the SEI layer are associated with the increase in charge transfer resistance, an estimated measure of which can reflect internal electrochemical processes, performance, and dynamics within the battery. Higher charge transfer resistance may impede the transfer of electrons and ions within the electrode material, leading to diminished battery performance. This could be linked to SEI layer growth as the formation of the SEI layer may increase the resistance to electron and ion transfer. Additionally, fluctuations in charge transfer resistance may directly impact battery charge and discharge rates. Elevated charge transfer resistance leads to greater polarization,

DOI: 10.54254/2755-2721/67/20240691

slowing down charge-discharge reactions [21]. This is closely linked to concerns regarding charge and discharge rates, as increased charge transfer resistance may limit the battery's ability to charge and discharge rapidly.

On a longer timescale of battery life, the SEI layer may infiltrate the pores of the electrode and even those of the separator. This could reduce the accessible active surface area of the electrode, thereby affecting battery capacity. The increase in electrode impedance, related to SEI growth and compositional changes, is considered a direct indicator of power attenuation within the battery over its lifespan.

To provide a more specific focus on feature extraction, the following aspects can be considered:

1. Charge and discharge rates: Given that the rate of SEI layer formation and growth during battery cycling may be influenced by charge and discharge rates, this feature can reflect changes in battery performance under different rates.

2. Battery impedance: Changes in battery impedance are closely linked to the formation and growth of the SEI layer. The increase in impedance may result from SEI layer growth and can thus serve as an important feature reflecting the state of battery lifespan.

3. Estimated charge transfer resistance: Variations in charge transfer resistance can reflect the internal dynamics of the battery. By estimating charge transfer resistance, we can better understand battery performance and lifespan states.

Through monitoring and analyzing these features, a more accurate understanding of the battery's lifespan state can be achieved, facilitating effective prediction and management. These features can be extracted and utilized as inputs for LSTM neural networks, providing robust support for State of Health (SOH) prediction throughout the battery's lifecycle.

3.4. Construction and Training Process of LSTM Neural Network

As described above, this section employs the charging and discharging experimental data of NASA lithium-ion batteries, utilizing features such as battery charging rate, discharging rate, estimated charge transfer resistance, and battery impedance as input variables for neural network training. Subsequently, the model's generalization capability and reliability are assessed based on the test dataset.

To ensure data consistency and model applicability, preprocessing steps are undertaken. By identifying the longest sequence length within the data and padding all sequences accordingly, uniform input data length is ensured, thereby enhancing the model's robustness.

Regarding the parameterization of the LSTM network, a multi-input single-output model architecture is adopted. The input feature dimension (`inputSize`) is defined as 52, corresponding to the four features. A layer structure containing 100 LSTM hidden units (`numHiddenUnits`) is selected to capture temporal information within the data. Lastly, the output size (`outputSize`) is set to 1, aligning with the objective of predicting battery health status.

During the training process, the Adam optimizer is utilized, with training options configured, including a maximum of 50 training epochs, a batch size of 8, a gradient threshold of 1, an initial learning rate of 0.01, and a piecewise learning rate adjustment strategy. This aids in achieving better convergence during training, balancing the risks of overfitting and generalization ability, and fine-tuning the learning rate more precisely in the later stages of training. Additionally, parameters are set to decrease the learning rate every 10 epochs, allowing for finer adjustments near convergence to prevent oscillations near the optimal solution and better adapt to model variations.

Furthermore, data shuffling is performed every epoch to enhance the model's generalization ability, and `'training-progress'` is enabled as a real-time monitoring window for tracking training progress. Through the visualization of loss functions, users are prompted to adjust and improve neural network hyperparameters in real-time to monitor training progress effectively.

In the model evaluation phase, 30% of all data is used as test data to predict SOH using the trained LSTM neural network. The model's predictions are compared with actual values, providing a direct reflection of the model's generalization performance.

DOI: 10.54254/2755-2721/67/20240691

4. Results and Discussion

4.1. Data Source

The dataset utilized in this study was sourced from the National Aeronautics and Space Administration (NASA) Predictive Capability for Excellence (PCoE). Samples from batteries B5, B6, B7, and B18, each rated at 2Ah, were selected from the dataset provided by the NASA Research Center.

Experiments were conducted under constant room temperature of 24°C, involving four lithium-ion batteries subjected to three different operational modes (charging, discharging, and impedance). Charging was performed in constant current (CC) mode at 1.5A until the battery voltage reached the charge termination voltage of 4.2V, followed by constant voltage (CV) mode until the charging current dropped to 20mA. Discharging was conducted at a constant current (CC) level of 2A until the battery voltage reached 2.7V, 2.5V, 2.2V, and 2.5V (corresponding to the discharge termination voltages for batteries 5, 6, 7, and 18, respectively). Impedance data were measured through Electrochemical Impedance Spectroscopy (EIS) frequency scans ranging from 0.1Hz to 5kHz. Repetitive charge and discharge cycles accelerated battery aging, while impedance measurements provided insights into the evolving internal battery parameters with aging. The experiments ceased when the batteries reached the End-of-Life (EOL) criterion, defined as a 30% decrease in rated capacity (from 2Ahr to 1.4Ahr).

4.2. Overall Prediction Performance

4.2.1. Model Evaluation Criteria

To comprehensively evaluate the performance of the models proposed in this paper for State of Health (SOH) prediction, we employed Root Mean Square Error ($RMSE$), Mean Absolute Error (MAE), model coefficient of determination, and three coefficients as evaluation metrics. Herein, smaller values of MAE and $RMSE$ indicate better prediction accuracy and stronger average predictive performance of the model, respectively. Conversely, higher values (within [0,1]) of the coefficients, evaluated against true attributions, indicate closer resemblance of predicted results. The definitions of the three aforementioned criteria are formulated as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{10}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|(y_i - \hat{y}_i)| \tag{11}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y - \bar{y}_i)^2} \tag{12}$$

$n$represents the total number of samples, y_i and $\hat{y}_i$ denote the true and predicted values of battery health state SOH at the "i"th cycle, respectively, while $\bar{y}$ stands for the mean of actual values. These metrics provide a comprehensive assessment of the predictive performance of the models under consideration.

4.2.2. Overall prediction Performance

Initially, feature extraction is conducted in conjunction with aging mechanism analysis, and the extracted features are utilized as input for the model. In this study, the charging and discharging cycle data of battery samples B0005, B0006, and B0007 provided by NASA Research Center are employed as the training set for the model; battery sample B0018 is used as the testing set to train the Long Short-Term Memory (LSTM) neural network model. Subsequently, an optimal model diagram, as illustrated in Figure 2, has been obtained, and the estimation results and errors of SOH are computed and evaluated.

Generalization ability serves as a crucial metric for assessing machine learning models, reflecting their performance when faced with unseen data. To validate the generalization ability of our proposed LSTM model, we adopted cross-validation methodology.

DOI: 10.54254/2755-2721/67/20240691

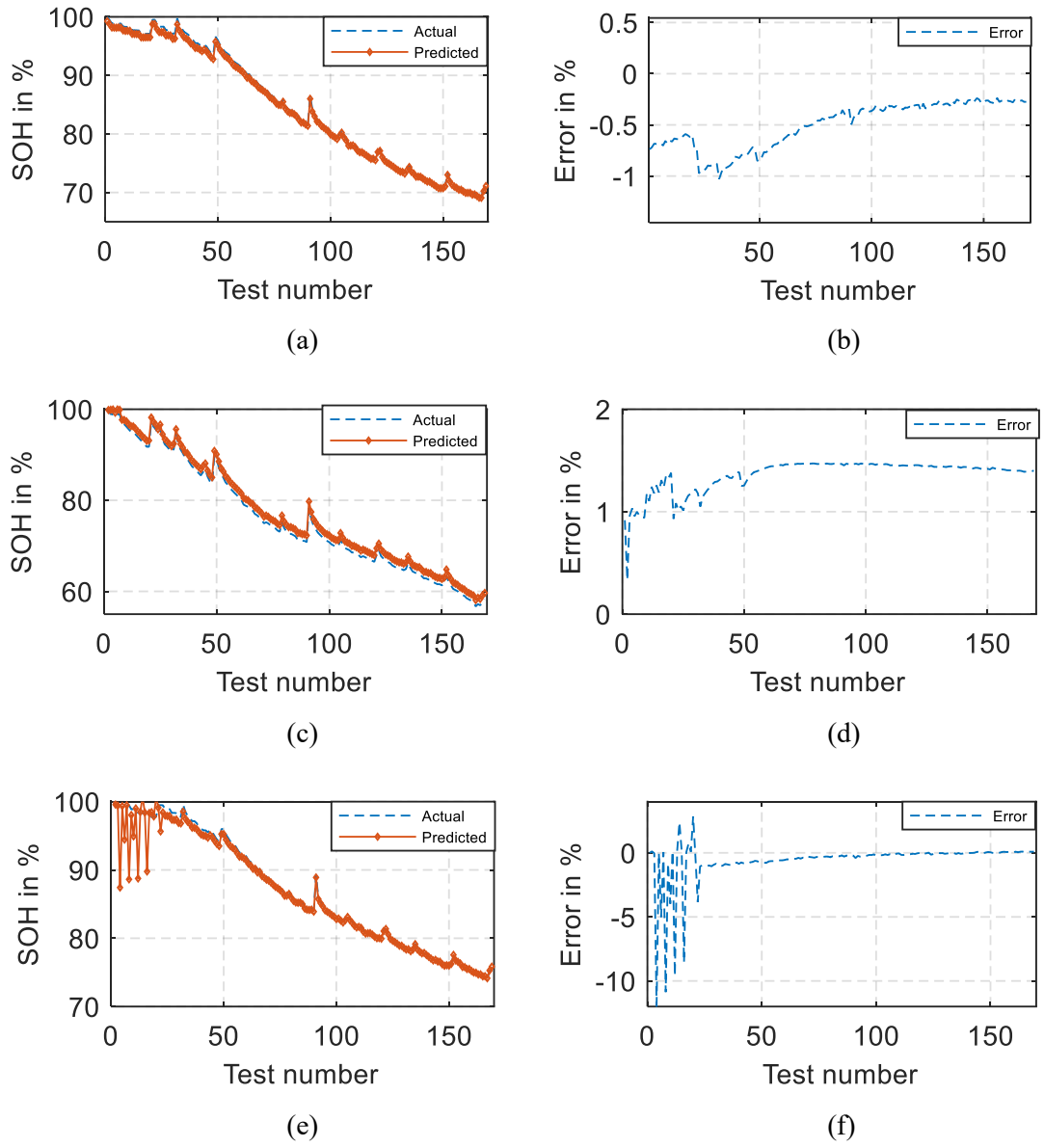


Figure 1. Evaluation of generalization performance using K-fold cross-validation. Figures a, c, and e represent SOH prediction plots, while figures b, d, and f depict prediction error plots.

In cross-validation, we employed the K-fold cross-validation approach, dividing the dataset into three equal parts. Each time, one part is utilized as the validation set, while the remaining portions serve as the training set. The model is then trained multiple times, and the average performance metrics are computed. As illustrated in the Figure 1, through this method, we can comprehensively evaluate the model's performance across different data subsets and mitigate dependence on specific data distributions.

$$CV_{(k)} = \frac{1}{k}\sum_{i=1}^{k} MSE_i \tag{13}$$

As illustrated in the table, during the validation process, we observed that MSE_1 and MSE_3 even under significant variations, our proposed LSTM model consistently achieves high-precision SOH predictions across different batteries, with the majority of estimation errors below 1%. By validating the

DOI: 10.54254/2755-2721/67/20240691

model's generalization ability, we obtain a more comprehensive assessment of its reliability in practical scenarios, further demonstrating its applicability.

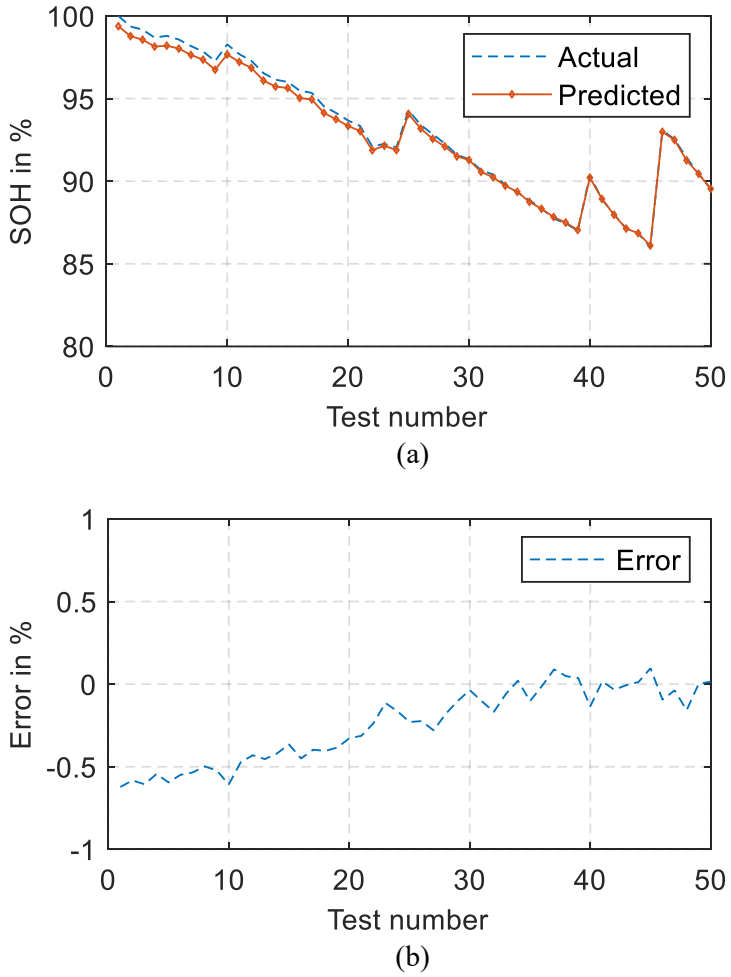


Figure 2. Overall Model Prediction Performance.

Table 1. Evaluation metrics for LSTM-based model approaches.

Method	Error Criterion		
	MAE	R^2	$RMSE$
LSTM	0.3681	0.9876	0.4414

4.3. Results Analysis and Discussion

4.3.1. SOH Prediction using Single Battery Data

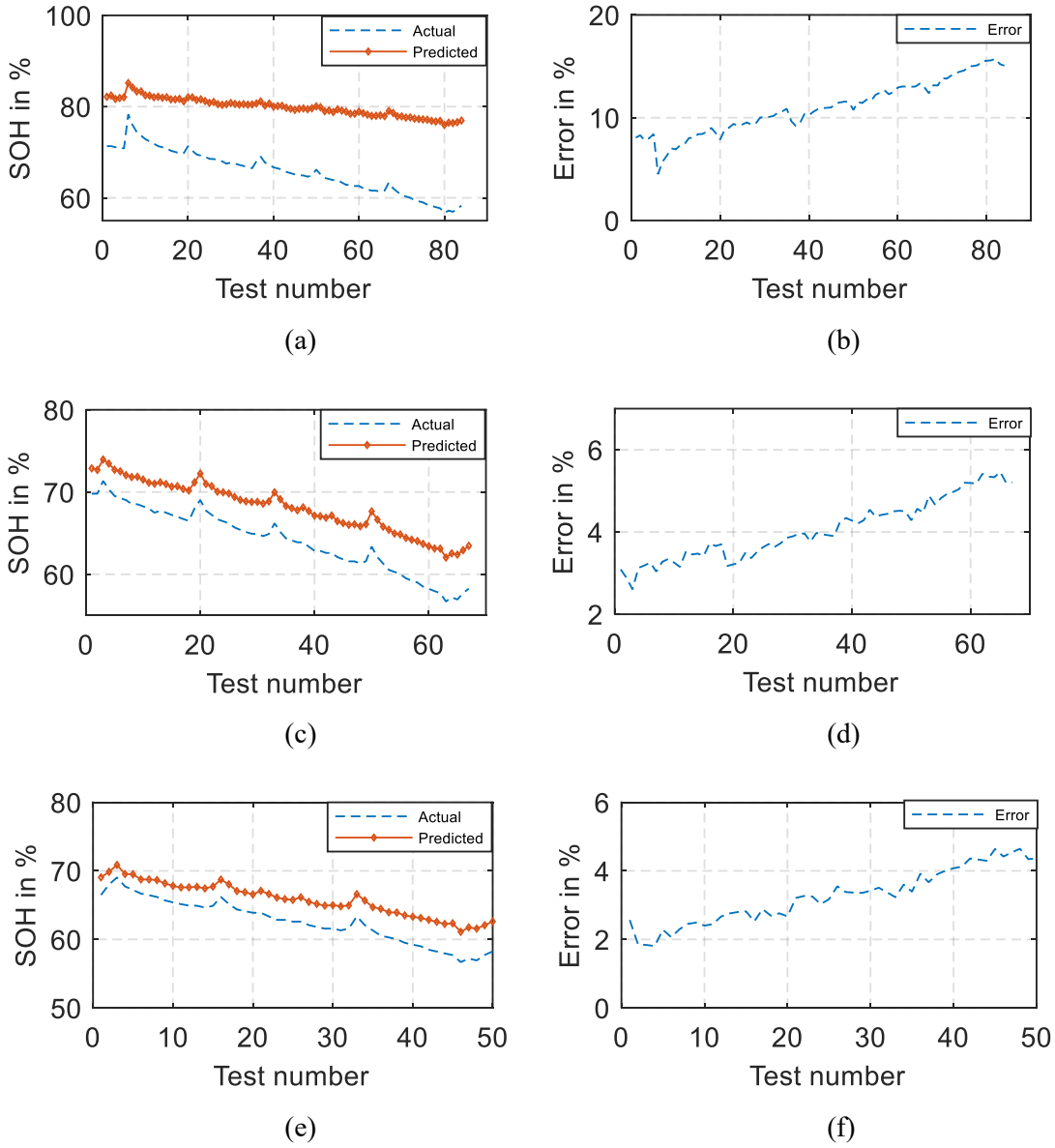
To investigate the impact of data volume on the Long Short-Term Memory (LSTM) network, we gradually increased the training data volume for battery B6, utilizing 50%, 60%, 70%, and 80% of the experimental data for model training, respectively. The results are presented in the table. With the increase in data volume, the overall prediction error gradually decreases.

As depicted in the Figure 3, when only 50% of the data is used for training, the maximum prediction error reaches 5.07%. However, when 60% and 70% of the cycle data are used to train the model, the maximum prediction errors decrease to 2.80% and 2.59%, respectively. It is evident that the LSTM

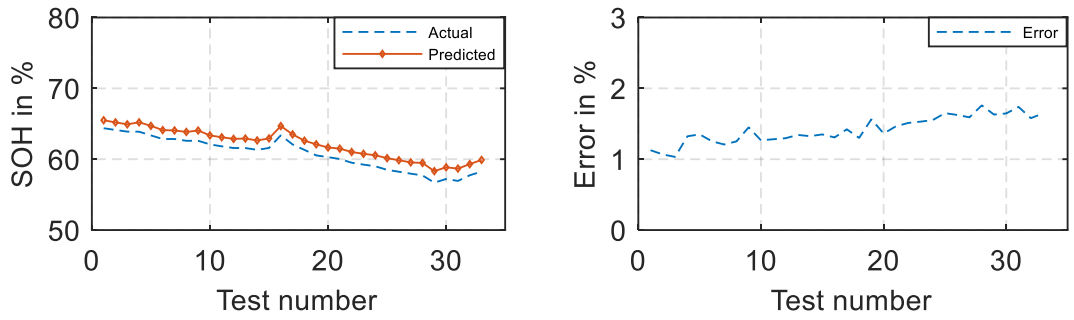
DOI: 10.54254/2755-2721/67/20240691

retains useful historical information when predicting the next state of the system, thereby causing a lag between the predicted and actual values.

From the aforementioned estimation error graph, it can be observed that during the prediction phase, as the training data decreases, the estimated SOH gradually deviates from the true SOH, with errors becoming significant towards the end of the cycle. When using 80% of the data for training, the majority of estimation errors are below 1.8%, with a maximum error of only 2.08%. In practical applications, SOH prediction errors below 2% are typically acceptable. Hence, we chose to train the model using 80% of the cycle data.



DOI: 10.54254/2755-2721/67/20240691



(g) (h)

Figure 3. Analyzing the Effect of Data Volume on Model Accuracy for Forecasting the SOH of Individual Cells.

Table 2. Evaluation metrics for different data volumes in a single battery.

Rate(%)	Error Criterion		
	MAE	R^2	$RMSE$
50	5.0717	-0.2092	5.3547
60	2.5915	0.5007	2.7462
70	2.8015	0.2122	2.8885
80	1.5077	0.5566	1.5377

4.3.2. Benefits of Utilizing Hybrid Battery Dataset

By analyzing the impact of varying training data volumes on model performance, we demonstrate the robustness of the proposed algorithm in tracking the State of Health (SOH) of batteries.

To further validate the influence of training data volume on model performance, we amalgamated three battery datasets, namely B5, B6, and B7, as the training set, while utilizing batteries B18 and B27 as the testing set. We observed prediction errors under different scenarios. By amalgamating data from multiple batteries, it becomes possible to better capture the features and patterns of the data, comprehensively consider the trend changes across different stages, and enable the predictive model to more accurately capture these patterns, thereby enhancing prediction accuracy. Moreover, increasing the training data volume can further mitigate the risk of model overfitting.

Results indicate that when training with data from batteries B5, B6, and B7, the prediction errors for batteries B18 and B27 are 0.3681% and 1.2081%, respectively. It is noteworthy that the slightly higher prediction error for battery B27 is attributed to minor variations in battery characteristics. However, upon evaluating various performance metrics, we found the proposed SOH prediction algorithm to perform satisfactorily, especially in scenarios where predicted trajectories closely resemble actual SOH changes.

5. Conclusion

This paper presents a method for estimating the State of Health (SOH) of lithium-ion batteries based on an LSTM model. Initially, we provide a detailed explanation of battery aging mechanisms, focusing on the variations in the thickness of the SEI layer on both the positive and negative electrodes during charge-discharge cycling. Specific parameters reflecting the battery aging process are extracted. By analyzing the correlation between aging-related features and battery SOH, we further extract input features for the neural network model: battery impedance, estimated charge transfer resistance, and charge-discharge rate. Utilizing these three features, SOH prediction is performed using LSTM, while Adam algorithm is employed to find optimal model parameters.

To demonstrate the model's generalization ability, we employ K-fold cross-validation for further training. Experimental results indicate that the majority of estimation errors can be maintained below

DOI: 10.54254/2755-2721/67/20240691

1%. For individual batteries, it is verified that when using 80% of cycle data for model training, SOH estimation error can be kept within 1.8%.

Furthermore, to expand and enhance the universality of the research, our next step will involve the introduction of other advanced algorithms to improve SOH prediction accuracy. We will focus on extracting relevant health features based on time-varying currents and partial charge-discharge cycle data to achieve more accurate SOH predictions.

References

[1] Guha A & Patra A. (2018). Online estimation of the electrochemical impedance spec-trum and remaining useful life of lithium-ion batteries. IEEE Trans Instrum Meas 2018, 67:1836-49. https://doi.org/xx.xxx/yyyy

[2] Alvarez-Monteserin I & Sanz-Bobi MA. (2022). An online fade capacity estimation of lithium-ion battery using a new health indicator based only on a short period of the charging voltage profile. IEEE Access 2022, 10:11138-11146.

[3] Yao L., Xu S., Tang A., Zhou F., Hou J., Xiao Y. & Fu Z. (2021). A Review of Lithium-ion Battery State of Health Estimation and Prediction Methods. World Electr Veh. J.2021, 12, 113.

[4] X. Hu., J. Jiang., D. Cao. & E. Bo. (2016). Battery Health Prognosis for Electric Vehicles Using Sample Entropy and Sparse Bayesian Predictive Modeling. IEEE T Ind Electron, 63 2645-2656.

[5] Zxa B., Jwa B. & Pdla C. (2021). Co-estimating the state of charge and health of lithium batteries through combining a minimalist electrochemical model and an equivalent circuit model.

[6] Li J., Loti N. & Landers R G. (2017). A single paticle model for lithium-ion batteries with electrolyte and stress-enhanced difusion physics. J Electrochem Soc, 164:A874-A883.

[7] Hu Xiao song, Li Sheng bo & Peng Huei. (2012). A comparative study of equivalent circuit models for Li-ion batteries. Journal of Power Sources, 198: 359-367.

[8] Khumprom P. & Yodo N. (2019). A Data-Driven Predictive Prognostic Model for Lithium-Ion Batteries Based on a Deep Learning Algorithm. Energies, 12, 660.

[9] Q. Liu, Y. Kang, S. Qu, B. Duan, F. Wen & C. Zhang (2020). An Online SOH Estimation Method Based on the Fusion of Improved ICA and LSTM IEEE/IAS Industrial and Commercial Power System Asia, pp. 1163-1167.

[10] Sumukh S., Rao V., Williamson S.S. (2021). Comprehensive Review on Smart Techniques for Estimation of State of Health for Battery Management System Application. Energies, 14, 4617.

[11] Lv Y., Shang M. & Chen X. (2019). Double-Net Enclosed Sulfur Composite as a New Cathode in Lithium Sulfur Batteries. The Journal of Physical Chemistry C, 123(29):17719-17727.

[12] Agubra V & Fergus J. (2013). Lithium-Ion Battery Anode Aging Mechanisms. Materials, 6(4): 1310-1325.

[13] J. Zeng & S. Liu (2023). Research on aging mechanism and state of health prediction in lithium batteries. Energy Storage, 72 108274.

[14] Qi Z. & White R E. (2008). Capacity fade analysis of a lithium ion cell. Journal of Power Sources, 179(2): 793-798.

[15] Petzl M, Kasper M & Danzer M A. (2015). Lithium plating in a commercial lithium-ion battery–A low-temperature aging study. Journal of Power Sources, 275: 799-807.

[16] Vetter J, Novak ´ P, Wagner MR, Veit C, Moller ¨ KC. & Besenhard JO. (2005). Ageing mechanisms in lithium-ion batteries. J Power Sources, 147(1–2):269–81.

[17] C. Pastor-Fernández, T. F. Yu, W. D. Widanage & J. Marco. (2019). Critical review of non-invasive diagnosis techniques for quantification of degradation modes in lithium-ion batteries. Renew. Sustain. Energy Rev., vol. 109, pp. 138–159.

[18] Hochreiter S. & Schmidhuber J. (1997). Long short-term memory. Neural Computation, 9(8).1735-1780.

DOI: 10.54254/2755-2721/67/20240691

[19] Awni Hannun, Carl Case, Jared Casper, Bryan Catanzaro, Greg Diamos, Erich Elsen, Ryan Prenger, Sanjeev Satheesh, Shubho Sengupta & Adam Coates. (2014). Deep speech: Scaling up end-to-end speech recognition. arXiv, preprint arXiv:1412.5567.

[20] Yamin H & Peled E. (1983). Electrochemistry of a nonaqueous lithium/sulfur cell. Journal of Power Sources, 9(3):281-287.

[21] Wright R. B., Motloch C. G. & Belt J R. (2002) Calendar- and cycle-life studies of advanced technology development program generation 1 lithium-ion batteries. Journal of Power Sources, 110(2): 445-470.