

Article

Artificial Neural Networks for Residual Capacity Estimation of Cycle-Aged Cylindric LFP Batteries

Pasquale Franzese ¹, Diego Iannuzzi ¹, Roberta Merolla ², Mattia Ribera ¹ and Ivan Spina ^{1,*}

¹ Department of Electrical Engineering and Information Technology (D.I.E.T.I.), University of Naples Federico II, 80125 Naples, Italy; pasquale.franzese@unina.it (P.F.); diego.iannuzzi@unina.it (D.I.); mattia.ribera@unina.it (M.R.)

² Dipartimento di Ingegneria dell'Informazione ed Elettrica e Matematica Applicata (D.I.E.M.), University of Salerno, 84084 Fisciano, Italy; rmerolla@unisa.it

* Correspondence: ivan.spina@unina.it

Abstract

This paper introduces a data-driven methodology for accurately estimating the residual capacity (RC) of lithium iron phosphate (LFP) batteries through a tailored artificial neural network (ANN) architecture. The proposed model integrates a long short-term memory (LSTM) layer with a fully connected layer, leveraging their combined strengths to achieve precise RC predictions. A distinguishing feature of this study is its ability to deliver highly accurate estimates using a limited dataset that was derived from a single cylindrical LFP battery with a 40 Ah capacity and collected during a controlled experimental campaign. Despite the constraints imposed by the dataset size, the ANN demonstrates remarkable performance, underscoring the model's capability to operate effectively with minimal data. The dataset is partitioned into the training and testing subsets to ensure a rigorous evaluation. Additionally, the robustness of the approach is validated by testing the trained ANN on data from a second battery cell subjected to a distinct aging process, which was entirely unseen during training. This critical aspect underscores the method's applicability in estimating RC for batteries with varying aging profiles, a key requirement for real-world deployment. The proposed LSTM-based architecture was also benchmarked against a GRU-based model, yielding significantly lower prediction errors. Furthermore, beyond LFP chemistry, the method was tested on a broader NMC dataset comprising seven cells aged under different C-rates and temperatures, where it maintained high accuracy, confirming its scalability and robustness across chemistries and usage conditions. These results advance battery management systems by offering a robust, efficient modeling framework that optimizes battery utilization across diverse applications, even under data-constrained conditions.



Academic Editors: Truong Minh Ngoc Bui, Truong Quang Dinh and Mona Faraji Niri

Received: 14 May 2025

Revised: 27 June 2025

Accepted: 4 July 2025

Published: 10 July 2025

Citation: Franzese, P.; Iannuzzi, D.;

Merolla, R.; Ribera, M.; Spina, I.

Artificial Neural Networks for Residual Capacity Estimation of Cycle-Aged Cylindric LFP Batteries.

Batteries **2025**, *11*, 260. <https://doi.org/10.3390/batteries11070260>

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license

(<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

At present, road, maritime, air, and railway transport account for approximately 64%, 18%, 16%, and 2% of the total transport-related final energy consumption, respectively [1]. An unprecedented revolution in electrified transport is unfolding worldwide. This revolution spans a wide range of applications [2,3], from electric vehicles (EVs) [4–8] and buses used for both public and private transport to small- and medium-sized maritime vessels, and even to the promising vertical take-off and landing (VTOL) aircraft in the

aeronautical sector, which are intended to serve urban areas. The increasing complexity of these applications has driven the development of advanced power conversion and control systems [9], as well as the improvement of energy storage technologies in terms of performance, safety, and integration. Currently, EVs rely on six well-established battery technologies: lithium nickel cobalt aluminum (NCA), lithium cobalt oxide (LCO), lithium nickel manganese cobalt (NMC), lithium manganese spinel (LMO), lithium titanate (LTO), and lithium iron phosphate (LFP), with no single technology clearly outperforming the others [10–12]. At present, the market share for LFP batteries has grown fivefold, from just 6% in 2020 to 30% in 2022. LFP batteries are gaining increasing popularity among car manufacturers and the research community. One of the major safety advantages of LFP batteries is their resistance to thermal runaway, a phenomenon that can lead to fires when batteries are damaged or defective. Another significant advantage of LFP batteries is their longer lifespan and resistance to degradation under fast-charging conditions. However, LFP technology does have some drawbacks, including relatively a low energy density, slower lithium diffusion, and poor electronic conductivity [13]. In this context, recent work has emphasized the need for a broader and more systematic research framework for storage batteries—one that considers not only materials and architectures but also chemical bonding mechanisms and cost-effectiveness as part of a multidimensional design space [14]. A critical challenge in the field of electrified transport is the accurate forecasting of battery state of health (SoH) and remaining useful life (RUL) in real-world operating conditions. Furthermore, the lack of statistically significant databases on battery usage across different electrified transport sectors throughout their life cycles complicates these predictions. For example, EVs used in urban areas typically operate within a limited range and rarely fall below 50% of their state of charge during normal use. As a result, the voltage–Ah characteristics of these vehicles tend to overlap, making it difficult to accurately estimate the SoH and RUL [15]. Only when batteries experience high depths of discharge (greater than 80%) can more precise estimations of the SoH and RUL be made. In addition, it is not enough to simply know the current SoH; understanding the rate of change over time is essential for predicting RUL. Therefore, precise predictions of the SoH and RUL [16] are crucial for the efficient use of batteries, particularly when making decisions related to battery reuse, recycling, and disposal [17]. Two primary approaches have been developed to predict the SoH and RUL of Lithium-ion batteries (LIBs): empirical models based on the electrochemical characteristics of batteries and data-driven methods. The accurate prediction of lithium-ion battery degradation is essential for ensuring long-term performance and safety, particularly in applications such as electric vehicles and grid storage. Traditional physics-based models often struggle to generalize across different usage patterns and cell chemistries. To overcome these limitations, data-driven approaches such as artificial neural networks (ANNs) have been increasingly adopted for modeling degradation trajectories, particularly given their ability to capture highly nonlinear relationships. Recent studies have emphasized the complex, nonlinear nature of degradation processes, which are influenced by multiple operational and intrinsic factors [18]. These models can be broadly categorized into traditional machine learning and deep learning approaches. Traditional machine learning techniques involve extracting relevant features from experimental data, which are then used as inputs to models such as support vector machines (SVMs [19]), relevance vector machines (RVMs), Gaussian process regression (GPR), k-nearest neighbors (k-NNs), ensemble learning, and artificial neural networks (ANNs) [20]. For example, Patil et al. [21] utilized SVMs and Support Vector Regression (SVR) for RUL estimation using features derived from battery voltage and temperature profiles, enabling faster computation. Guo et al. [22] employed a Bayesian RVM approach, optimizing health features from voltage, current, and temperature curves using grey relational analysis and principal

component analysis (PCA). The authors in [23] propose a two-step machine learning approach: a genetic algorithm-based fuzzy C-means clustering technique is used to partition the training data, which helps determine the model topology. Deep learning models, such as feed-forward multilayer perceptrons, convolutional neural networks (CNNs), and recurrent neural networks (RNNs), including long short-term memory (LSTM) networks, have been successfully applied to battery SoH and RUL estimation. For instance, Wu et al. [24] utilized feed-forward neural networks (FNNs) trained on Li-ion battery terminal voltage curves, incorporating importance sampling for input selection. LSTM-based RNNs, such as those proposed by [25,26], capture long-term dependencies in data, making them suitable for SoH estimation under various driving conditions. Zhang et al. [27] further explored LSTM networks for battery prognostics, highlighting their capability to process variable-sized input data [28]. The paper [29] combines offline model training and online weight generation to address the divergence problem caused by inconsistency between individual battery cells. The method uses incremental capacity analysis, focusing on the height of the incremental capacity peak in the high-voltage region as the feature for battery aging prediction. Ungurean et al. [30] and Fan et al. [31] implemented gated recurrent unit (GRU) networks for online SoH prediction, noting higher prediction errors but fewer required parameters compared to LSTM. Choi et al. [32] compared FNN, CNN, and LSTM models for battery capacity estimation using multi-channel charging profiles, but their study was limited by model complexity variations, and they did not evaluate the individual contributions of different measurable signals. In [33], an SG-GRU architecture is employed. The model uses a Savitzky–Golay (SG) filter for denoising and a GRU network to capture temporal dependencies for RUL prediction. Their results, showing an RMSE within 1%, confirm the effectiveness of GRU-based models in this context. Although Support Vector Regression (SVR) and multilayer perceptrons (MLPs) have been widely adopted in various regression tasks, their applicability to lithium-ion batteries is limited due to their inability to adequately capture the complex temporal dynamics and nonlinearity inherent in battery degradation processes. The recent literature consistently demonstrates the superior performance of LSMT- and GRU-based models in these tasks. For instance, Huo and Chen [34] showed that GRU networks outperform conventional methods in SoC estimation due to their ability to retain temporal dependencies over long sequences. Cui and Joe [35] further refined GRU-based estimation by integrating spatial–temporal attention mechanisms and handcrafted health indicators, achieving significant improvements in SoH prediction accuracy.

In contrast, models such as SVR and MLPs lack inherent mechanisms to model time-dependent degradation trajectories. They treat each input independently and thus fail to exploit the sequential structure of battery aging data. This limitation makes them less effective in capturing gradual capacity fade trends or sudden shifts in battery behavior. Alharbi et al. [36] compared various deep learning techniques and reported that recurrent models consistently outperformed feed-forward alternatives like MLPs, particularly in long-horizon forecasting scenarios. Likewise, Ding et al. [37] combined signal decomposition techniques with GRU networks to successfully predict the remaining useful life (RUL), again underlining the importance of temporal modeling.

Despite advancements in data-driven battery health estimation, previous studies have not thoroughly examined which training factors most significantly influence estimation precision. Additionally, other aspects, such as pre-processing operations, model complexity, and sub-sampling criteria, have not been fully addressed. More importantly, previous papers typically train neural networks using data from a specific battery cell and then evaluate the model's performance using a portion of the same battery's dataset. In these studies, the total dataset from a single battery cell is split into the 'training', 'validation,' and

'testing' sets. The 'training' and 'validation' sets are used for the learning process, while the 'testing' set is used to evaluate the trained ANN's performance. However, this method does not demonstrate the robustness of the trained ANN when applied to different battery cells, whose data was not used during training. This is a critical limitation, as creating a dataset requires the battery to undergo a significant reduction in its RUL (often nearing its end of life).

The paper [38] proposes a feed-forward migration neural network for predicting the aging trajectories of lithium-ion batteries. The proposed method first establishes a base model from existing battery aging data using a dual exponential function. In other studies, 'transfer learning' models are developed [39]. In these cases, knowledge learned from the dataset of a specific cell is transferred to other cells. In [40], a hierarchical extreme learning machine (HELM) is used to improve the estimation robustness and accuracy without the complex parameter model. The results show that the SoH estimation errors are no more than 1.5%, while the training and estimation datasets are from the same temperature; when the SoH estimation is conducted at different temperatures, the maximum error is only 3.36%.

The present research paper addresses existing gaps by evaluating the state of health (SoH) of Li-ion batteries using a custom artificial neural network (ANN) architecture inspired by [39]. The model integrates a long short-term memory (LSTM) layer and a fully connected (FC) layer. It explores the influence of model complexity, pre-processing strategies, and training configurations [41] on battery residual capacity (RC) prediction, adopting a purely data-driven approach. The proposed method is thoroughly tested, and its robustness is systematically assessed.

A key original contribution of this work lies in the detailed analysis of multiple training-related factors. Furthermore, two different battery cells, each subjected to distinct aging processes, are considered. Notably, the ANN is trained using data exclusively from the first battery cell, without exposure to any data from the second cell during training. After internal validation using the test set derived from the first battery, the trained model is evaluated on the second cell to verify its ability to generalize across different aging profiles. This testing setup, based on non-overlapping datasets, confirms the consistency and effectiveness of the proposed architecture and training methodology.

Additional and significant enhancements have been introduced to strengthen experimental validation. Specifically, for lithium iron phosphate (LFP) cells, the proposed method has been benchmarked against a gated recurrent unit (GRU)-based approach. The comparative analysis confirms that the LSTM-FC model achieves superior performance in terms of both prediction accuracy and generalization capability.

Moreover, in order to further validate the generalization capabilities of the model, a second dataset involving lithium nickel manganese cobalt oxide (NMC) cells has been introduced. This dataset includes cells aged under varying C-rates and ambient temperatures, thus offering a more challenging test scenario. In this case as well, the ANN is trained on data from only one cell and then tested on other cells whose data were never used during training. Despite the heterogeneity in aging conditions, the proposed method maintains high prediction accuracy and stable error metrics, underscoring its robustness and scalability across chemistries and usage profiles.

Therefore, another central contribution of this paper is the demonstration that high-quality results can be achieved even with a limited number of training samples. The ability to train the network on a single cell while successfully generalizing to different cells—operated under various stress conditions and chemistries—proves the practical applicability of the method and confirms its potential in real-world battery health diagnostics. This comprehensive framework contributes to a deeper understanding of battery degradation

behavior and advances the development of reliable and accurate data-driven prognostic tools for Li-ion batteries.

2. Battery Aging Setup

2.1. Batteries Under Test and Experimental Setup

The experiments were conducted in collaboration with FAAM FIB S.p.A., using cylindrical LiFePO₄ batteries with a rated capacity of 40 Ah. These batteries support a maximum discharge current of 40 A (1 C) and a maximum charge current of 20 A (0.5 C), with a voltage range from a minimum cut-off of 2.7 V to a maximum cut-off of 3.65 V. The detailed parameters of the tested LFP cells are provided in Table 1.

Table 1. Technical data of the FAAM LFP cell.

Quantities	Values
Technology	LiFePO ₄
Nominal capacity [Ah]	40
Nominal voltage [V]	3.2
Cut-off voltage [V]	2.7
Max voltage [V]	3.65
Nominal charging current [A]	20
Nominal discharging current [A]	40
Nominal working temperature [°C]	25 ± 5
Storage temperature [°C]	25 ± 5
Expected cycle life (DoD 100%)	2500

Cycle aging was carried out using an ACT0550 cycler integrated with an ACS DM600ESP climate chamber that is capable of maintaining ambient temperatures between -15°C and 80°C . Thermistors were placed on the battery surfaces to monitor the temperature, and data collection was automated with variable sampling rates tailored to different cycle stages. The experimental setup is depicted in Figure 1.

The cyclic testing procedure, managed by a host PC, continued until the batteries reached their end of life (EoL), defined as 80% of their initial capacity measured at the start of the aging process. All tests were conducted at a constant ambient temperature of 25°C . The cycler schedule, illustrated in Figure 2a, consisted of 25 aging cycles followed by a hybrid pulse power characterization (HPPC) cycle.

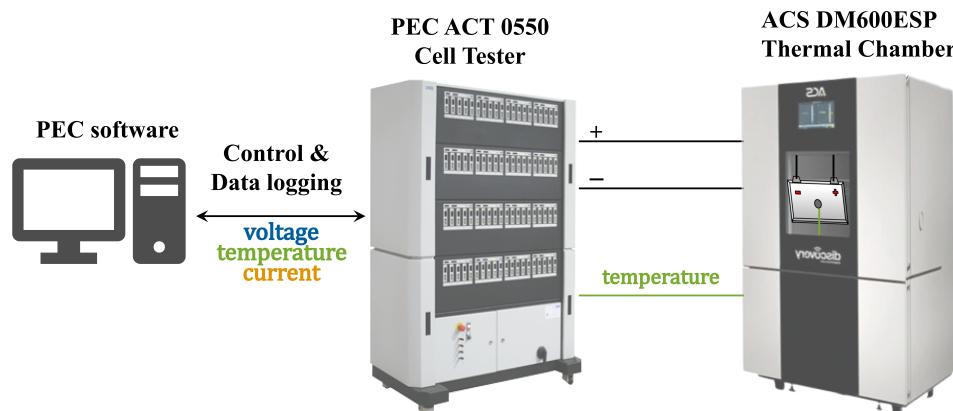


Figure 1. Experimental setup: PC control terminal cell tester, thermal chamber, and cell under test.

Aging cycle (Figure 2b and Table 2): This cycle begins with a CC-CV charge phase, where the battery is charged at 20 A until the terminal voltage reaches 3.65 V. This is

followed by a 30 min discharge phase at -40 A, a 1 min rest, and a continuation of discharge at -40 A until the voltage drops to the cut-off value of 2.7 V.

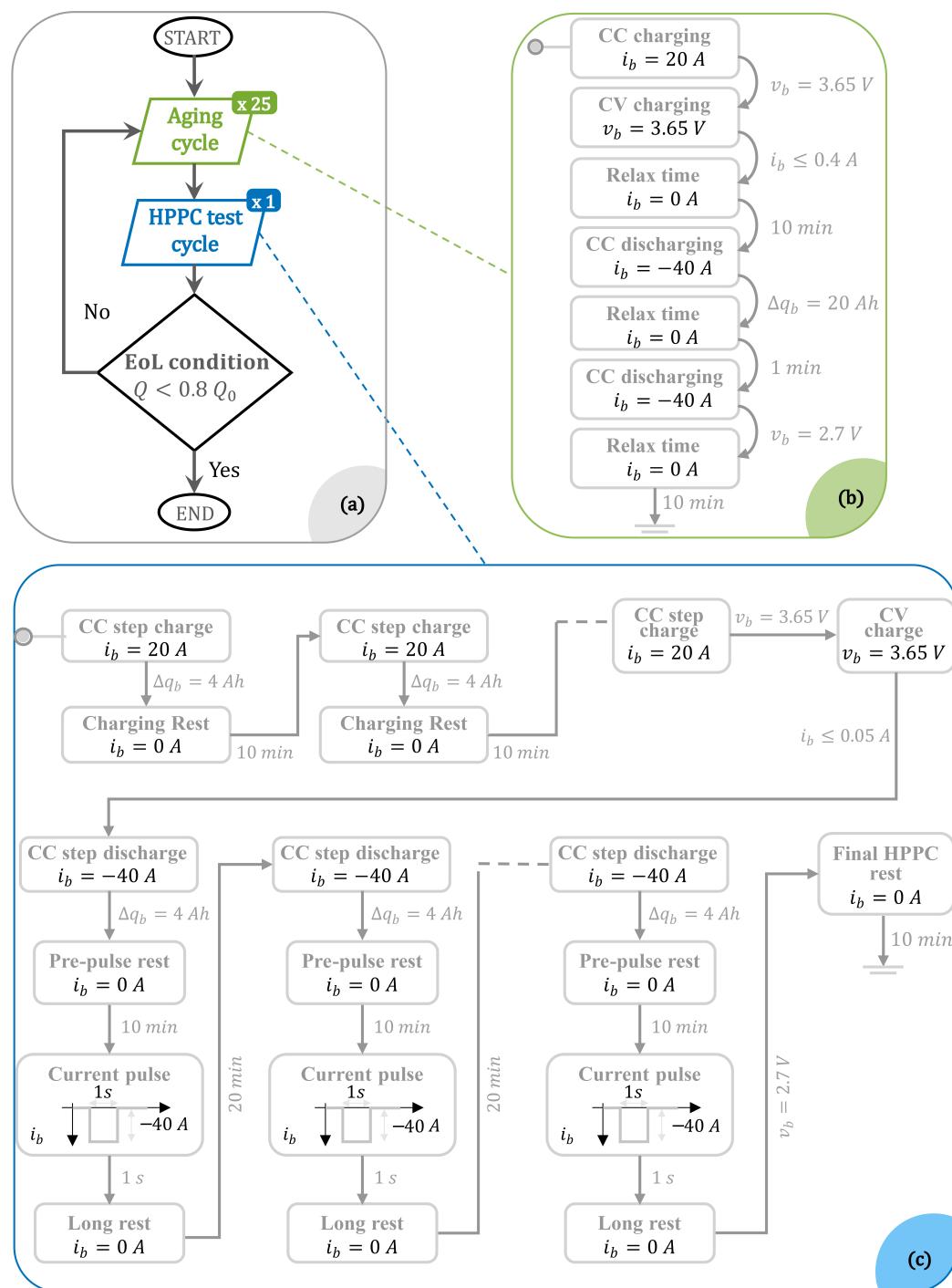


Figure 2. Flow chart of the implemented aging and data-gathering methods, with a focus on the aging cycle and HPPC procedure. Subfigure (a) illustrates the complete flow of the experimental protocol, alternating 25 Aging cycles and 1 HPPC test until the end-of-life condition is reached. Subfigure (b) details the structure of a single Aging cycle. Subfigure (c) shows the full sequence of the HPPC procedure used for data acquisition.

HPPC cycle (Figure 2c and Table 3): This cycle starts with partial charges, where the battery is charged by 4 Ah increments at a constant current of 20 A until the terminal voltage reaches 3.65 V. After each partial charge, a 10 min rest allows the battery to stabilize. During the discharge phase, the battery undergoes multiple 4 Ah partial discharges at a

current of 40 A, interspersed with 30 min rest periods. A 1 s discharge pulse of 40 A is applied midway through each rest period. The discharge phase ends when the voltage drops to 2.7 V.

Table 2. Charging and discharging cycles to age the cell.

Step	Description	Cycler Control Values	End Event	Next Step
1	CC charging phase	Constant current: 20 A	Battery voltage = 3.65 V	2
2	CV charging phase	Constant voltage: 3.65 V	Battery current \leq 0.4 A	3
3	Rest	Relax time: 10 min	-	4
4	CC discharging phase	Constant current: -40 A	Moved charge = 20 Ah	5
5	Rest	Relax time: 1 min	-	6
6	CC discharging phase	Constant current: -40 A	Battery voltage = 2.7 V	7
7	Rest	Relax time: 10 min	-	End

Table 3. HPPC procedure.

Step	Description	Cycler Control Values	End Event	Next Step
1	CC step charge	Constant current: 20 A	Moved charge = 4 Ah Battery voltage = 3.65 V	2 3
2	Charging rest	Relax time: 10 min	-	3
3	CV step charge	Constant voltage: 3.65 V	Battery current \leq 0.05 A	4
4	Rest Ch-Dis	Relax time: 10 min	-	5
5	Before pulse rest	Relax time: 10 min	-	6
6	Pulse	1-s current pulse: -40 A	-	7
7	Long rest	Relax time: 20 min	-	8
8	CC step discharge	Constant current: -40 A	Moved charge = 4 Ah Battery voltage = 2.7 V	5 9
9	Final HPPC rest	Relax time: 10 min	-	End

For a generic n-th aging cycle, electrical profiles for both HPPC and aging cycles are shown in Figures 3 and 4.

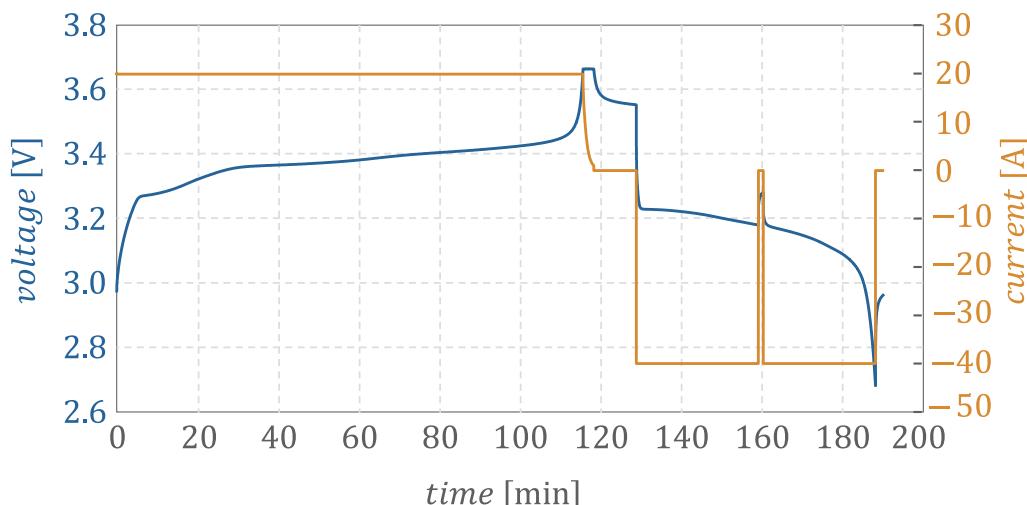


Figure 3. Voltage and current profile of the aging cycle.

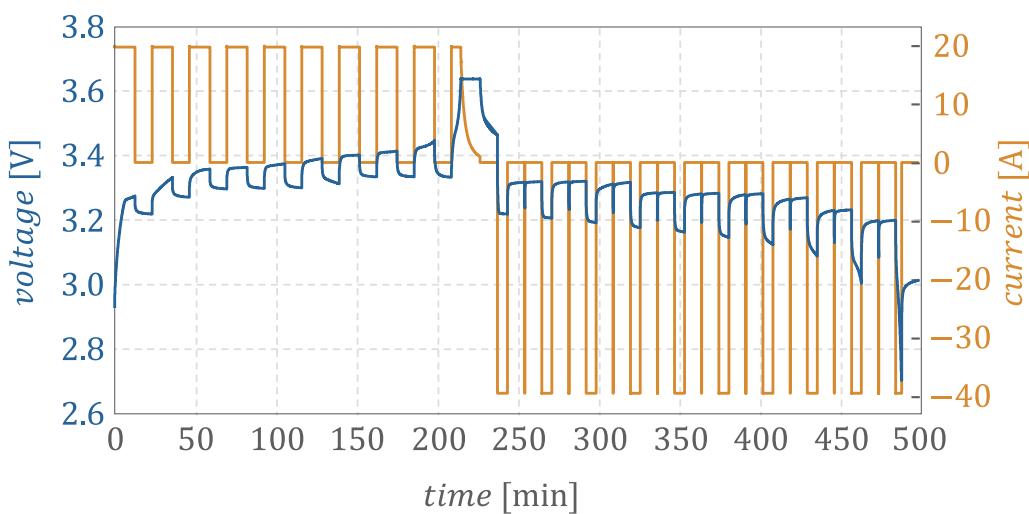


Figure 4. Voltage and current profile of the HPPC procedure.

2.2. Data Description

The battery under test exhibited an initial rated capacity (RC) of 38.73 Ah, which decreased to 30.00 Ah after 1543 aging cycles. This capacity fade phenomenon is illustrated in Figure 5.

Voltage (V), current (I), and temperature (T) measurements were recorded for all cycles from beginning of life (BoL) to end of life (EoL). These measurements are essential for constructing data-driven models to estimate battery capacity.

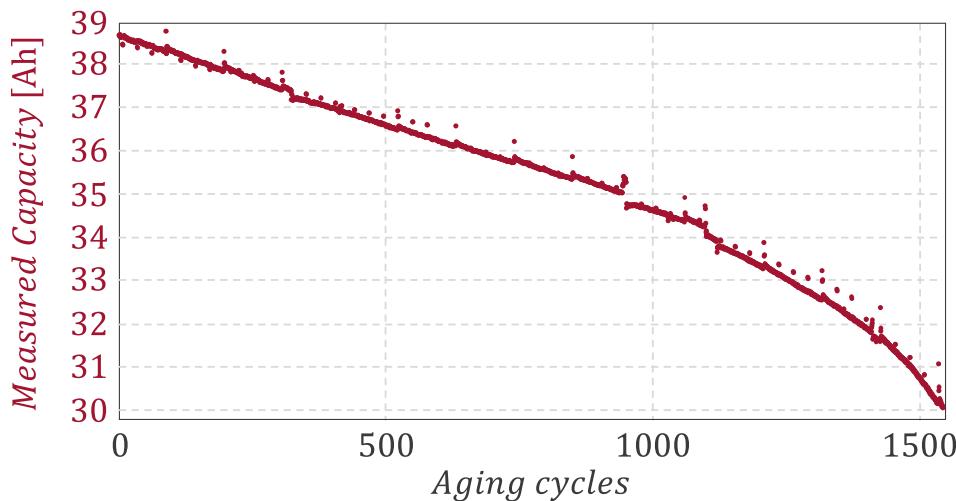


Figure 5. Cell capacity fade over cycle life.

Additionally, post-processed data on the moved charge (Q) trend for each cycle was included, providing four input signals (V , I , T , and Q) for training and testing. Figure 6 illustrates the discharge measurement behaviors of the aged LFP battery at both BoL and EoL.

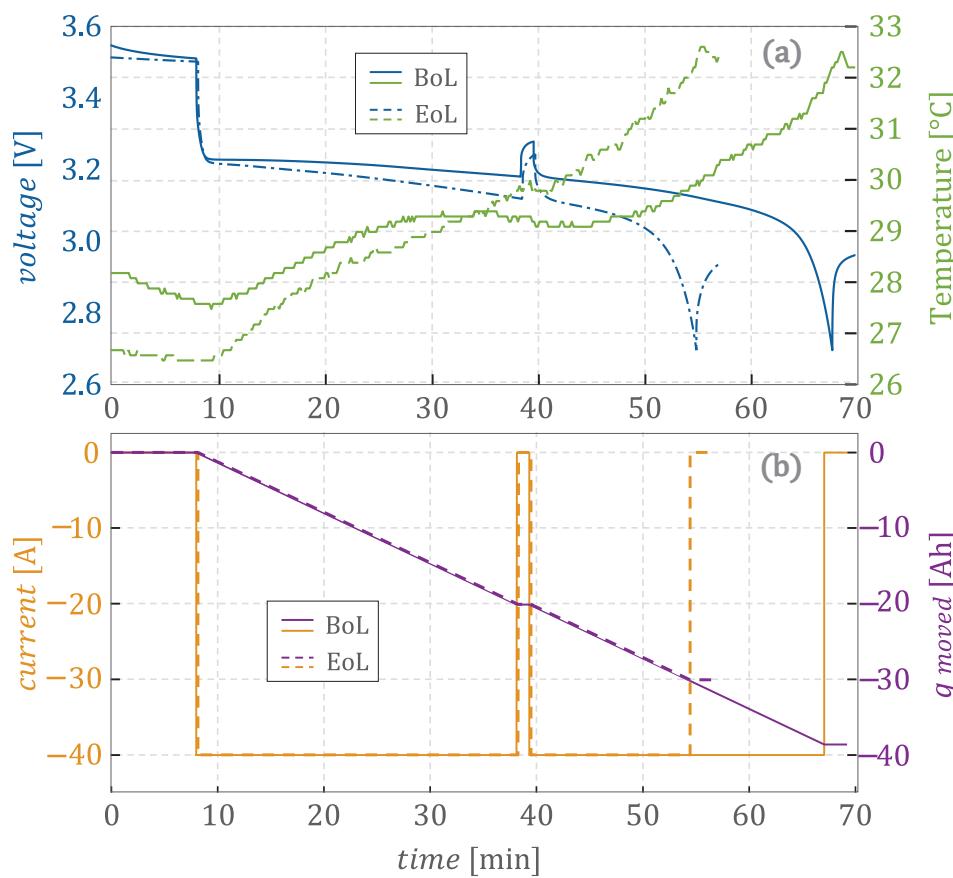


Figure 6. Measurement of voltage, current, temperature, and moved charge on the cell at BoL and EoL. (a) Profiles of voltage and temperature during the test at BoL and EoL. (b) Profiles of current and moved charge during the test at BoL and EoL.

3. Methodology of the Neural Network Approach

This study's methodology follows a structured approach, involving data pre-processing, validating data-driven models for optimal hyperparameter selection, and choosing the most effective artificial neural network (ANN) architecture based on the test results. The overall procedure is summarized in the flowchart in Figure 7, which outlines each step described in the following subsections.

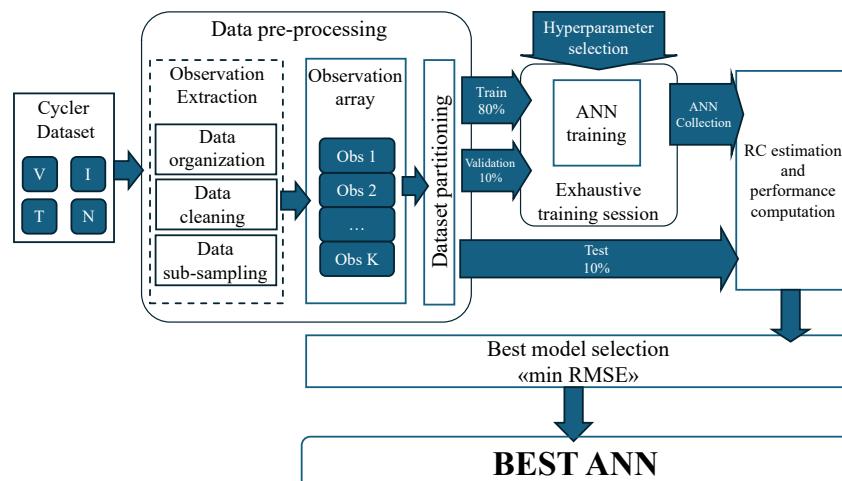


Figure 7. Workflow of data processing, ANN training, and model selection based on RMSE minimization.

3.1. Description of the Tested ANN Model

This study evaluates a hybrid ANN architecture that combines a long short-term memory (LSTM) layer with a fully connected (FC) layer. This architecture leverages LSTM's sequential processing capabilities, which capture complex temporal dependencies within data through specialized gating mechanisms. The LSTM layer selectively retains relevant historical information, addressing the vanishing gradient problem common in standard recurrent neural networks (RNNs). The FC layer then transforms the LSTM's temporal representations into final model outputs, enhancing sequence representation and output precision to meet the study's objectives.

LSTMs, which are a variant of RNNs, are designed to capture both short- and long-term dependencies in sequential data. While RNNs struggle with long-term dependencies due to the vanishing gradient problem, LSTMs mitigate this issue through gating mechanisms that regulate the information flow within memory cells, allowing for the selective retention of relevant information over time and avoiding gradient vanishing and explosion issues.

During the training phase, the ANN determines one value of the internal weights W per training interaction using several observations input: one observation is a complete charge (or discharge) phase, regardless of whether this belongs to an aging cycle or to an HPPC.

The core innovation of LSTMs lies in the cell state C_t , which is modulated by three gates: The input gate controls how much new information is added to the cell. The forget gate regulates past information retention, and the output gate determines how much of the cell state contributes to the current output. These gates use sigmoid and hyperbolic tangent activation functions to control information flow.

At each step, the LSTM computes the current cell state C_t by combining the forgetting gate's effect on the previous state C_{t-1} and the input gate's modulation of new candidate values, while the output gate generates the hidden state h_t for the current iteration. This process allows LSTMs to effectively model long-term dependencies in sequential data. The equations governing the gates and cell state are detailed in Equation (1).

This architecture, illustrated in Figure 8, enables LSTMs to overcome the limitations of standard RNNs, managing complex temporal patterns in data effectively.

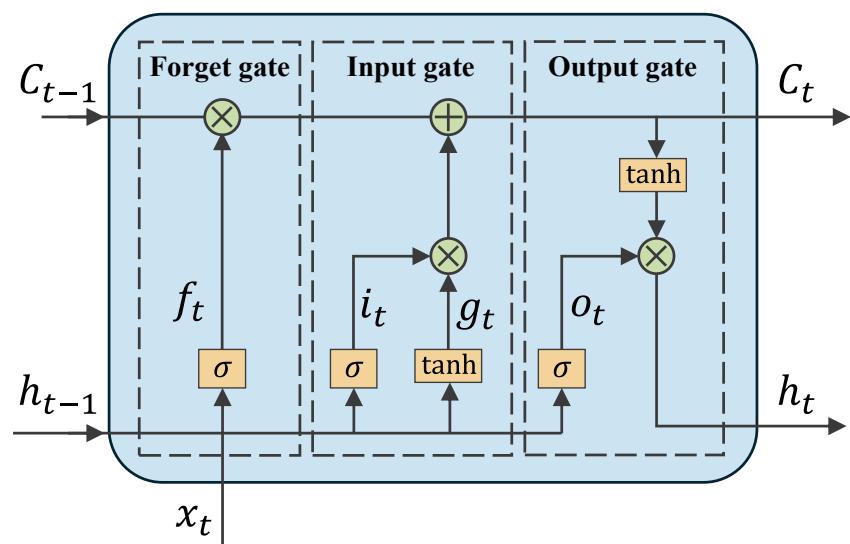


Figure 8. LSTM architecture.

$$\begin{aligned}
i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\
f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\
g_t &= \tanh(W_g \cdot [h_{t-1}, x_t] + b_g) \\
C_t &= f_t \odot C_{t-1} + i_t \odot g_t \\
o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\
h_t &= o_t \odot \tanh(C_t)
\end{aligned} \tag{1}$$

The FC layer consists of M elemental neurons (or nodes), where each neuron in the layer is connected to all outputs from the previous layer. Each neuron computes its output as the weighted sum of the input signals plus a bias term, with the result passed through an activation function. Common activation functions include the rectified linear unit (ReLU), the sigmoid function, the hyperbolic tangent, and the softmax function.

The process for each neuron j is mathematically described by Equation (2):

$$z_j = \phi \left(\sum_{i=1}^M \omega_i x_i + b_j \right) \tag{2}$$

where x_i and ω_i represent the input signals and their associated weights, b_j is the bias term for neuron j , and ϕ is the activation function applied uniformly across all neurons in the layer.

Figure 9 illustrates a schematic representation of a single neuron, showing its inputs, weights, bias, and activation function.

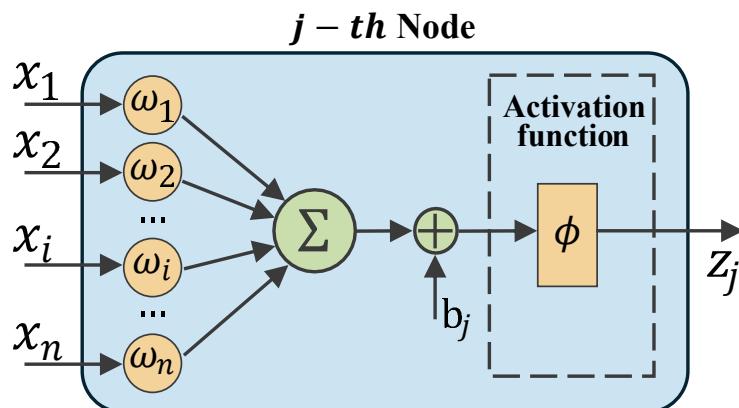


Figure 9. FC node architecture.

The outputs from all neurons in the current layer are then propagated to the next layer. All nodes within the layer operate simultaneously, with each performing the same type of computation on their respective inputs.

In this study, the LSTM layer is paired with an FC layer to output the predicted remaining capacity of the battery. This LSTM-FC architecture was chosen for its dual advantages: the LSTM layer effectively reduces model sensitivity to noise, ensuring robust performance, while the FC layer consolidates features learned by the LSTM, enhancing the model's flexibility and generalization capacity. This combination allows for accurate capacity predictions and adaptation to varied input configurations. The architecture used in this study is depicted in Figure 10.

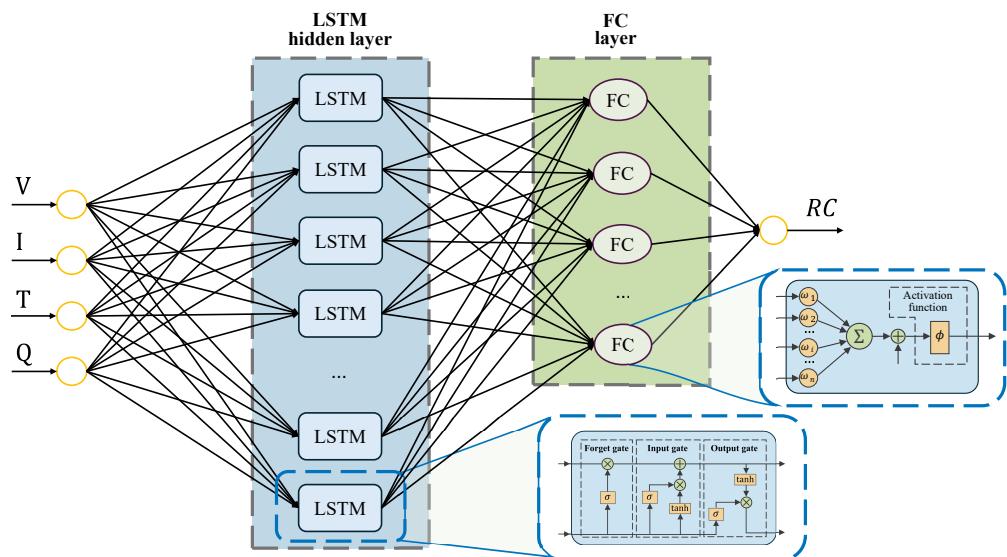


Figure 10. Architecture of the selected ANN. The LSTM layer is coupled with an FC layer.

3.2. Data Pre-Processing

Data pre-processing ensures the input data's quality and consistency for the ANN model. The original Cyler dataset consists of a continuous time series of current, voltage, and temperature measurements recorded from the beginning of life (BoL) to the end of life (EoL) of the lithium-ion cell. These raw signals are supplemented by metadata provided by the cycling equipment, including an incremental aging cycle counter.

The observation extraction begins with data organization, where the full dataset is segmented into individual observations, each corresponding to either a charge or discharge phase of a single aging cycle. For each observation, the transferred charge is computed through current-time integration (Coulomb counting), providing an estimate of the capacity associated with that segment.

Subsequently, a data cleaning process is applied, where corrupted observations are identified and removed. These anomalies, often stemming from instrumentation resets or shutdowns, result in invalid timestamps that compromise the accurate computation of the transferred charge. The data cleaning process relies on two criteria for identifying and discarding faulty observations. First, the time intervals between consecutive samples within each observation are analyzed. If any consecutive timestamps exceed the maximum expected sampling interval—indicating an abnormal timestamp gap—the entire observation is discarded. Second, the measured capacity of each observation is evaluated to detect incomplete or corrupted cycles, which may result from interruptions due to unexpected errors or maintenance operations. The mean difference in capacity between successive observations is computed, and any observation whose capacity difference exceeds five times this mean value is removed from the dataset. Such faulty segments are discarded to preserve the integrity of downstream processing and learning steps.

Given the long duration and variable sampling rates of recorded measurements, data sub-sampling is performed to reduce the data size while preserving signal fidelity. Different sampling rates are tested to analyze their effect on RC predictions by the ANN. Multiple sub-sampling strategies were applied, as the sampling rate was treated as a hyperparameter in the search for the optimal ANN configuration. The result of these steps is the observation cell array, a MATLAB-compatible structure where each cell element contains the segmented measurements of voltage (V), current (I), temperature (T), and the computed moved charge (q).

Finally, dataset partitioning is applied to split the data into training, validation, and testing subsets with an 80-10-10% ratio, ensuring each subset includes representative capacity fade behavior. Although formal k-fold cross-validation was not performed,

multiple data splits were tested during model tuning, with the chosen ratio yielding the most consistent and favorable performance. This structure supports robust model generalization across different battery aging conditions.

3.3. Exhaustive Training Session

After partitioning the dataset into the training, validation, and test subsets, an exhaustive training and evaluation process is carried out to identify the most suitable ANN model for capacity estimation. The training phase is performed on the designated training set using multiple ANN configurations that differed in architecture and hyperparameters. This section investigates the impact of multiple hyperparameters on the performance of the ANN-based estimator, focusing on minimizing the root mean square error (RMSE) of the test dataset. Key factors analyzed include the network's complexity, training parameters, and input signal selection.

The hyperparameters explored are as follows:

- **Batch Size:** It determines the number of observations per training iteration and affects the frequency of weight updates. Smaller batch sizes can improve convergence and reduce overfitting but may increase computational costs due to more frequent updates.
- **The Length of Sub-sampling Data:** It specifies the number of samples per observation during training, balancing computational efficiency and model accuracy. Excessive sub-sampling may lead to information loss, while minimal sub-sampling enhances accuracy at the cost of increased computational demands.
- **The Number of LSTM Nodes:** It influences the complexity of the LSTM layer, which captures temporal dependencies. Higher node counts enable learning of intricate patterns but incur greater computational costs.
- **The Number of Fully Connected Nodes:** It determines the complexity of the fully connected layer. While higher node counts can enhance the model's learning capacity, excessive complexity risks overfitting.
- **Input Signal Selection:** It examines the impact of different input signal combinations (e.g., V, I, T, and Q) on estimation performance. Including all signals yielded the highest accuracy and robustness, although results for specific combinations are omitted for brevity.

In this study, the impact of various hyperparameters on the performance of artificial neural networks was systematically evaluated. Batch sizes of 16, 32, 64, and 128 were tested, allowing us to assess how different sizes influence training efficiency and model accuracy. Additionally, sub-sampling lengths of 200, 400, 600, 800, and 1000 samples were employed, together with some models trained without sub-sampling to investigate its overall impact on model performance. Furthermore, LSTM node counts of 16, 32, and 64 were evaluated to determine the optimal architecture for our networks. Lastly, configurations with 5, 10, and 15 fully connected nodes were explored to identify an effective balance between model complexity and generalization capability. Regarding input signal selection, various combinations of input signals (i.e., V, I, T, and Q) were investigated. The training employed the Adam optimizer, a stochastic gradient descent method utilizing adaptive estimation of first- and second-order moments. Early stopping was implemented to prevent overfitting, halting training if no significant improvement was observed after a specified number of validation checks. Considering the values chosen for each hyperparameter (Batch Size: four values; Length of Sub-sampling Data: five values; Number of LSTM Nodes: three values; and Number of Fully Connected Nodes: three values), the total possible combinations resulted in $4 \cdot 5 \cdot 3 \cdot 3 = 180$.

An exhaustive grid search was performed across all possible combinations of hyperparameters; therefore, 180 ANNs were trained. A parallel hyperparameter selection process

guides the search toward optimal combinations. Each candidate model is trained and then validated on the dedicated validation set to monitor generalization performance and avoid overfitting. This process results in a collection of trained models, each associated with its respective validation error.

3.4. Performance Computation and Best Model Selection

Error measures are essential for quantifying a neural network model's accuracy and assessing deviations from expected values. The primary error metrics are listed below:

- **Root Mean Square Error (RMSE):** The RMSE calculates the square root of the mean of the squared differences between predicted and actual values. It gives more weight to larger errors, making it valuable in contexts where significant errors need heavier penalization (3)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

A lower RMSE indicates higher model accuracy with respect to the test data.

- **Max Absolute Error (MAE):** This metric represents the maximum absolute differences between predicted and actual values. It may be more sensitive to outliers than other metrics; it provides an exhaustive view of the maximum deviation of samples (4).

$$MAE = \max_i |y_i - \hat{y}_i| \quad (4)$$

where y_i and \hat{y}_i represent the actual and predicted values for the i -th observation.

- **Coefficient of Determination (R^2):** Known as *R-square*, this metric shows the proportion of variability in the data that the model explains, making it especially useful for evaluating the goodness of fit in regression problems. It ranges from 0 to 1, with a higher value indicating a better fit (5).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (5)$$

where \bar{y} is the average of the observed values of the dependent variable.

The MAE and RMSE metrics quantify model precision on the test dataset, while $|R^2 - 1|$ (rather than R^2) is used to evaluate performance from a different perspective, with values approaching zero indicating optimal performance, consistently with MAE and RMSE behaviors. Performance indicators (RMSE, MAE, and R^2) were calculated for each combination to identify the optimal configuration. RMSE is used as the primary metric; thus, the model achieving the lowest RMSE on the test set is selected as the best ANN.

4. Error Behavior of the Proposed ANN Versus Factor Variation

The neural networks were trained on a workstation equipped with MATLAB R2024b, running on an AMD FX-8320 CPU (8 cores at 3.5 GHz) and an Nvidia RTX 4060Ti GPU with 8GB of GDDR6 memory. MATLAB was chosen for its widespread application in deep learning, facilitating the sharing and deployment of trained networks. The GPU's parallel processing capabilities significantly accelerated the training process. Additionally, MATLAB's Experimental Manager Toolbox streamlined the exploration of training configurations, allowing the testing of various parameter combinations efficiently.

The relationship between performance indicators and hyperparameters is inherently multidimensional, with each indicator varying as a function of four independent parameters. To isolate the effect of a single hyperparameter, the other three were held constant at their optimal values. Figure 11 illustrates the impact of each hyperparameter on the RMSE,

evaluated on the test dataset of the LFP cell. Each graph shows the RMSE as a function of the hyperparameter under investigation, with the remaining hyperparameters fixed at the values corresponding to the best-performing configuration.

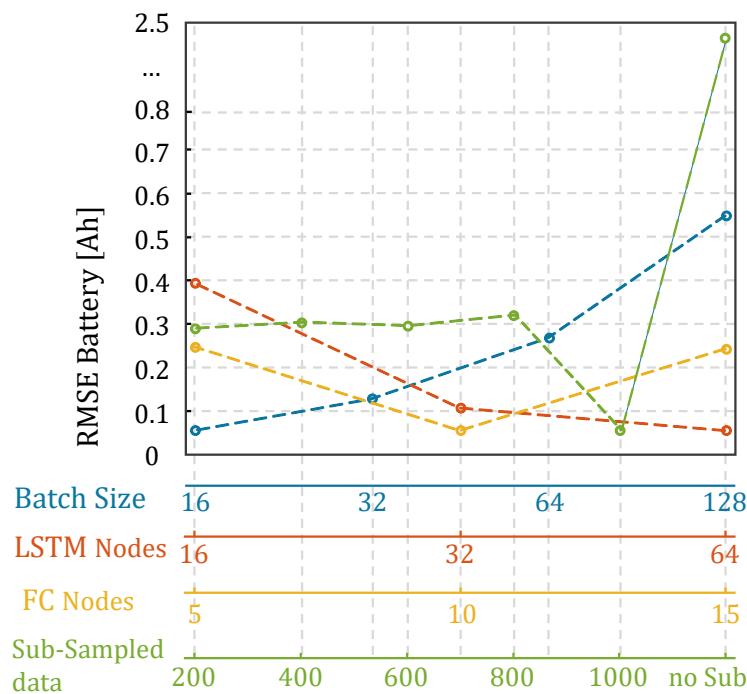


Figure 11. Effects of the hyperparameters' variation on the RMSE.

Figure 12 presents a spider plot comparing performance metrics (RMSE, MAE, and $|R^2 - 1|$) for three representative LSTM-FC ANN configurations:

- *Best ANN*: The configuration with the lowest RMSE.
- *Worst ANN*: The configuration with the highest RMSE.
- *Average ANN*: The configuration with the RMSE closest to the mean value of all trained models.

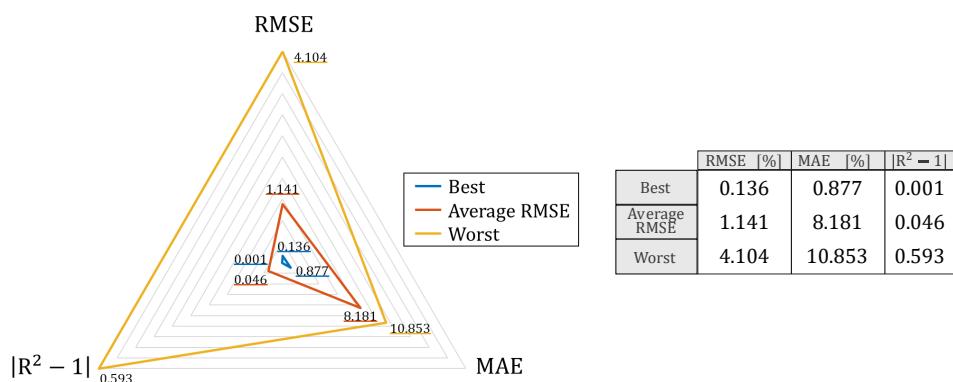


Figure 12. Spider plot of ANN performance. The three plots refer to the ANNs with the best, worst, and average RMSEs.

The spider plot uses $|R^2 - 1|$ instead of R^2 for direct performance comparison, and both the RMSE and MAE are expressed in % with reference to the rated capacity. The optimal ANN configuration is closest to the plot's origin, indicating superior performance across all metrics.

The vertices of the plot represent the maximum values recorded for the RMSE, MAE, and $|R^2 - 1|$ across all training sessions. As noted, the best ANN not only has the lowest

RMSE but also the lowest MAE and $|R^2 - 1|$, positioning it closer to the origin of the plot. Conversely, the worst ANN occupies the largest area, reflecting its high RMSE and $|R^2 - 1|$ values. The average ANN lies between these extremes.

The optimal hyperparameter values for the best-performing ANN are summarized in Table 4. The performance evaluation presented in the following section pertains to this neural network, which will remain fixed throughout the analysis.

Table 4. Factors of the assessed best-performing ANN.

Factors	Values
Sub-sampling length	1000
Batch size	16
Number of LSTM nodes	64
Number of fully connected nodes	10
Maximum number of epochs	100
Validation patience	10
Validation frequency	50
Initial learning rate	0.01
Learning rate drop factor	0.1 after 10 epochs
Training algorithm	Adam
Training computational time (s)	520 (8 min e 40 s)

5. Performance Evaluation of Proposed ANN

This section focuses on evaluating the performance of the proposed ANN, which was trained using fixed hyperparameter values determined through preliminary optimization. The finalized configuration of these hyperparameters enables a thorough assessment of the model's predictive capabilities and its ability to generalize to data from cells not included in the training process.

5.1. Traditional Evaluation Method

The neural network training employed a holdout approach, dividing the dataset into three subsets: 80% for training, 10% for validation, and 10% for testing. This strategy ensured that the model could learn from a substantial portion of the data while also being validated during training and tested on an independent subset. All subsets originated from data collected on a single cell. The RMSE values computed in this phase helped identify the optimal network configuration. The network accurately captured the RC trend over time (see Figure 13).

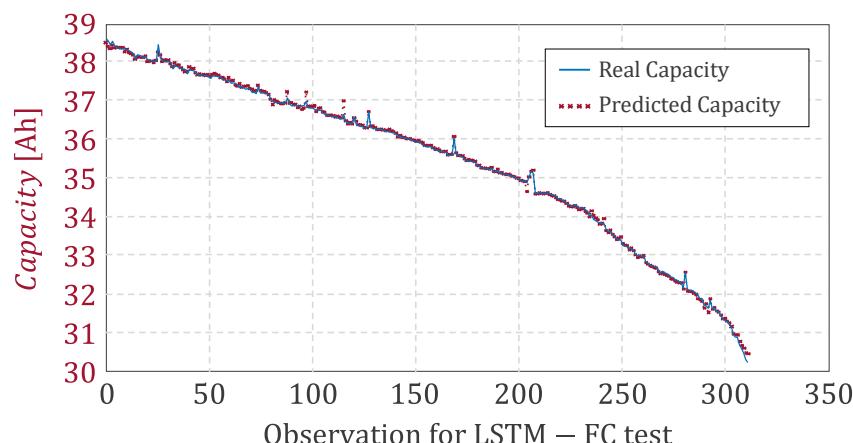


Figure 13. Proposed ANN predicted capacity.

The low percentage errors relative to the nominal capacity of the tested cell, reported in Figure 14, further confirm the model's accuracy.

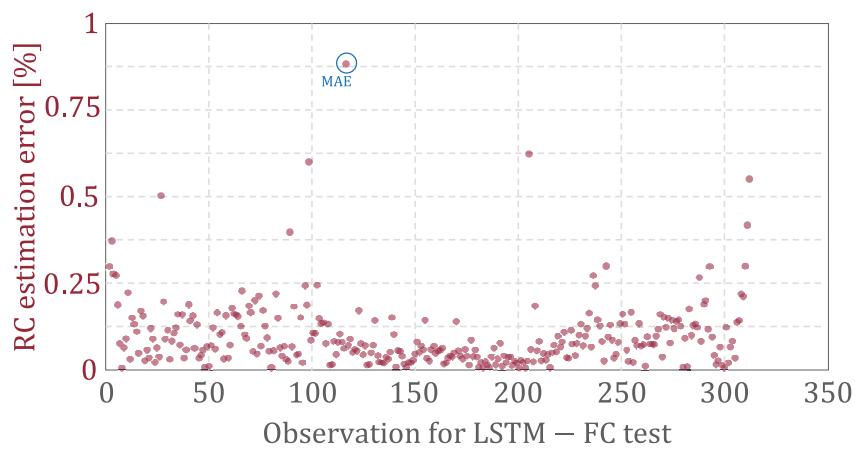


Figure 14. Relative accuracy estimation error.

The average relative error was below 0.1%, with a maximum of 0.875%. Moreover, the ANN achieved an RMSE of 0.136% and an R^2 value of 0.999 (see Figure 11), demonstrating its reliability in estimating RC.

5.2. Performance Evaluation on a Different Aged Cell

This section evaluates the performance of the proposed ANN on a second LFP cell with the same nominal specifications but a different aging history. These tests assess the model's robustness when faced with entirely unseen data. Notably, the ANN was trained without access to any data from the second cell.

The second cell's dataset was obtained through an aging campaign with a modified protocol, resulting in a slightly different aging profile. While the overall cycling procedure was similar to that of the first cell (see Figure 3), the discharge step from cycle 106 to cycle 208 differed. During this interval, after the 1 min pause for internal resistance measurement, the discharge current was reduced to 32 A instead of 40 A. Table 5 summarizes the applied protocol, and a sample acquisition is shown in Figure 15.

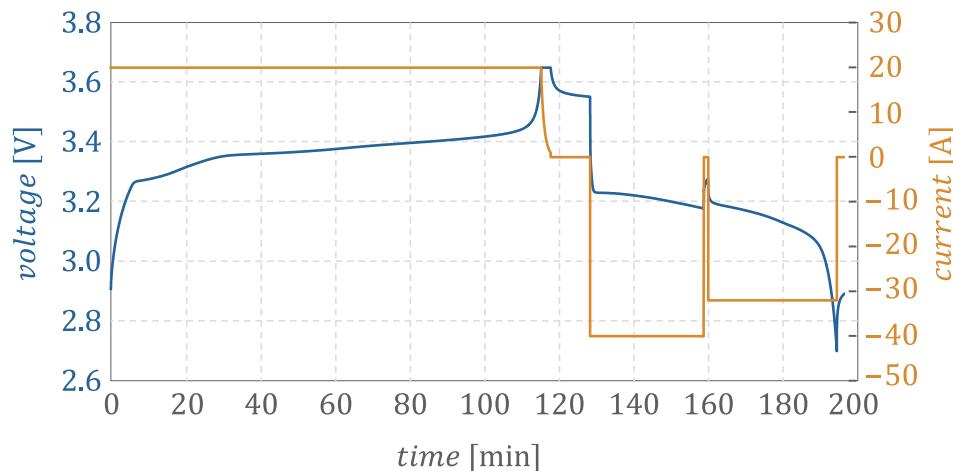


Figure 15. Voltage and current profile of the aging cycle of cell 2 (cycle number 106 to cycle number 208).

The HPPC protocol remained identical to that of the first battery (see Table 3). Figure 16 displays the discharge acquisition data for the second cell at BoL and EoL, while Figure 17

highlights the capacity fade differences between the two batteries. The second cell's capacity decreased from 38.72 Ah to 30.02 Ah after 1532 aging cycles.

Table 5. Charging and discharging cycles to age cell 2 (cycle number 106 to cycle number 208).

Step	Description	Cycler Control Values	End Event	Next Step
1	CC Charging phase	Constant current: 20 A	Battery voltage = 3.65 V	2
2	CV Charging phase	Constant voltage: 3.65 V	Battery current < 0.4 A	3
3	Rest	Relax time: 10 min	-	4
4	CC Discharging Phase 1 C	Constant current: -40 A	Duration 30 min	5
5	Short Rest	Relax time: 1 s	-	6
6	CC Discharging Phase 0.8 C	Constant current: -32 A	Battery voltage = 2.5 V	7
7	Rest	Relax time: 10 min	-	End

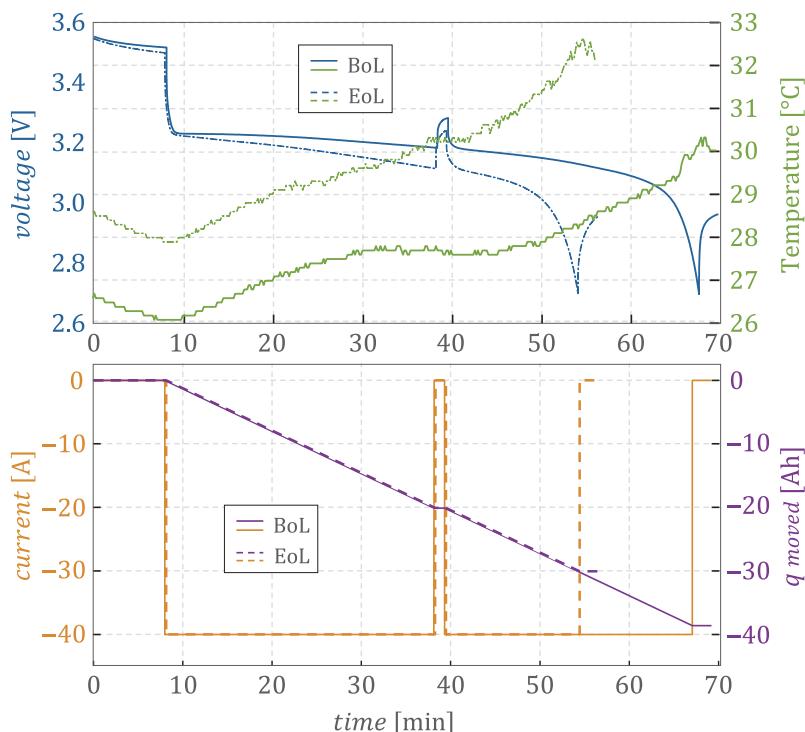


Figure 16. Measurement of voltage, current, temperature, and moved charge on cell 2 at BoL and EoL.

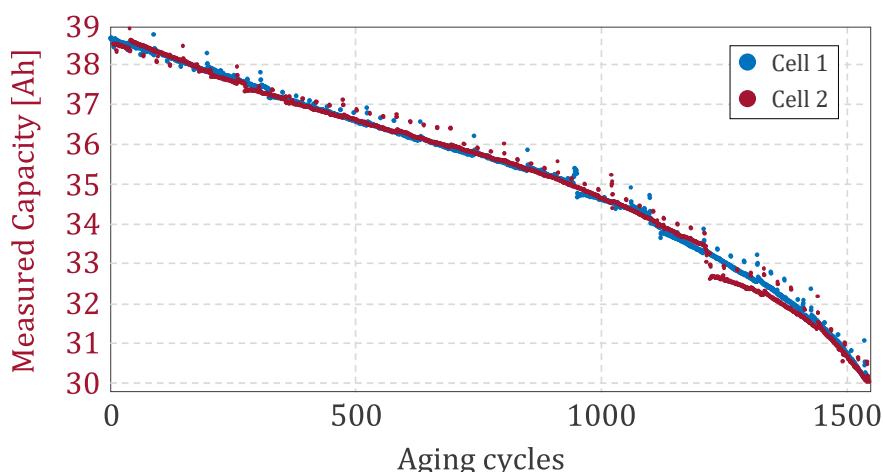


Figure 17. Comparison of capacity fade over the cycle life of cell 1 and cell 2.

Data processing, including filtering and charge calculation, was applied to this dataset. The same ANN, which was trained exclusively on the first cell's dataset, was used to estimate the RC of the second cell. Figure 18 compares the measured and predicted values, showing a non-perfect fit with occasional outliers.

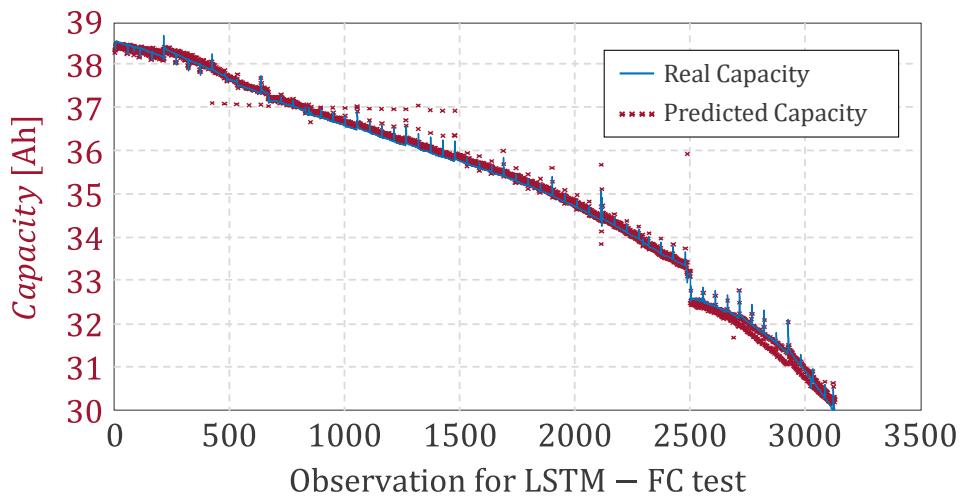


Figure 18. Trend of cell 2's predicted capacity by the LSTM-FC network compared to the actual values.

However, the analysis of error distribution estimation in Figure 19 reveals an average estimation error of 0.2%, with a few outliers below 2.5%. These results are outstandingly good and underscore the ANN's robustness, considering that the ANN never had access to data from the second cell.

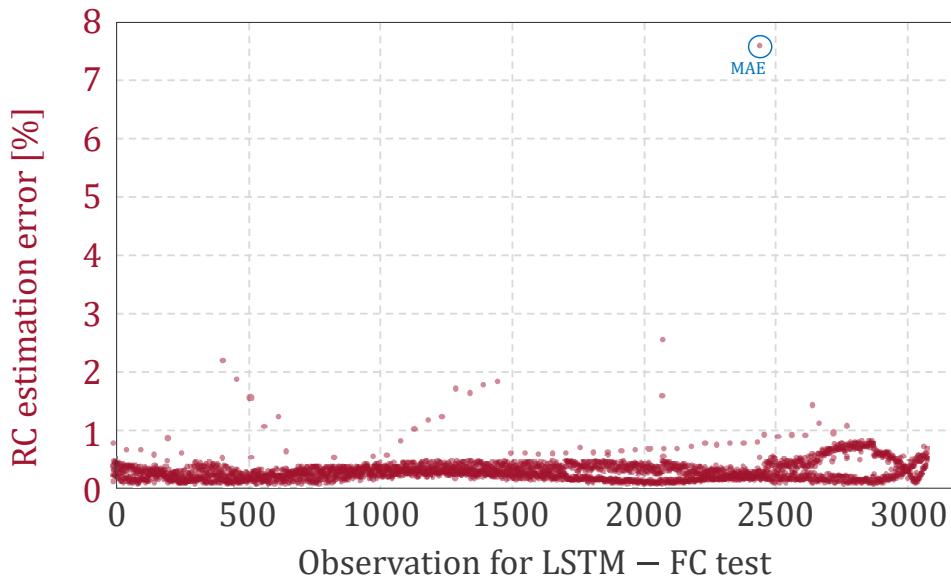


Figure 19. Relative accuracy estimation error of cell 2.

Once we established that the 'first-cell-best-case-ANN' (proposed ANN) also performs outstandingly good on the second cell, it is worth investigating if a different choice of hyperparameter values may even guarantee better results on the second cell. Counting on the availability of the remaining 179 trained ANNs (not the best cases for the first cell), it was possible to calculate the performance indicators for all possible combinations to look for a 'second-cell-best-case-ANN' candidate. All 180 trained ANNs were calculated for each configuration to identify the "second-cell-best-case-ANN". As an example, the results obtained for the first cell and shown in Figure 11 are now reported for the second cell.

Figure 20 illustrates the effect of each hyperparameter on the RMSE for the second cell while keeping the remaining hyperparameters fixed at their optimal values that were identified for the first cell. Notably, the hyperparameter values that minimized the RMSE for the first cell also achieved the lowest RMSE for the second cell.

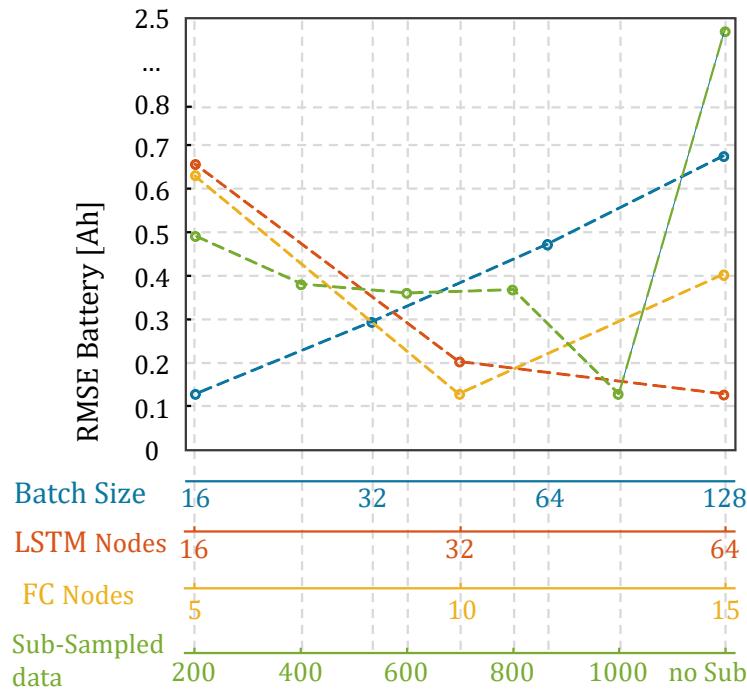


Figure 20. Effects of the hyperparameters' variation of the RMSE on cell 2.

Additionally, Figure 21 presents a spider plot summarizing the performance metrics for the second cell. Among the 180 trained ANNs, the proposed ANN achieved the lowest RMSE (0.3%) and the best $|R^2 - 1|$ value (0.003), although its MAE (7.59%) was not the minimum. Notably, this MAE value corresponds to a single outlier point; excluding this point, the MAE would fall below 3%. Another network among the 180 achieved the lowest MAE (3.7%), but with a significantly worse RMSE (0.96%) and $|R^2 - 1|$ (0.028).

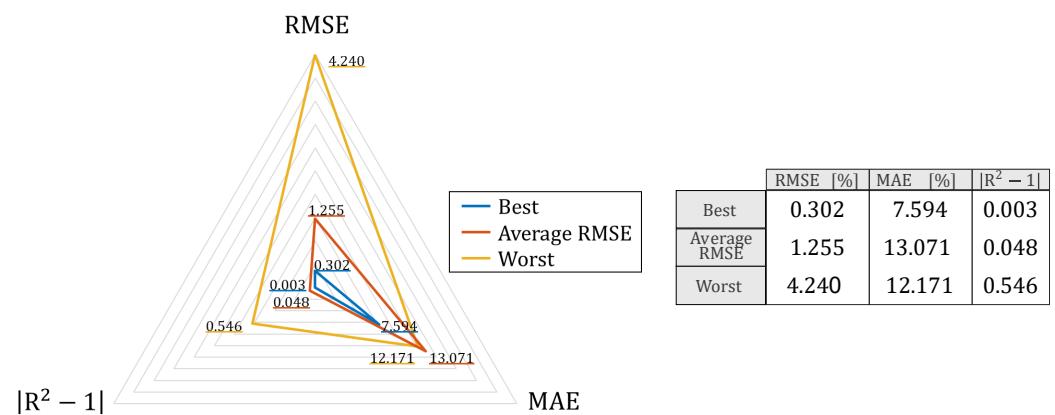


Figure 21. Spider plot of ANN performance based on the described indexes of cell 2. The three plots refer to the ANNs with the best, worst, and average RMSEs.

Thus, the “second-cell-best-case-ANN” coincides with the “first-cell-best-case-ANN”, confirming the consistency of the optimal hyperparameter values across different cells. However, this conclusion is valid strictly in terms of hyperparameter configuration, since the actual performance indices (RMSE, MAE, and $|R^2 - 1|$) for the second cell are slightly

worse than those obtained for the first cell. This is consistent with the fact that the ANN was trained exclusively on the first cell's dataset.

Finally, Figure 22 presents a spider plot comparing the performance metrics across both cells. The slight degradation in performance highlights the inherent challenges of cross-cell generalization. Nevertheless, the overall results remain outstanding, especially considering the training constraints.

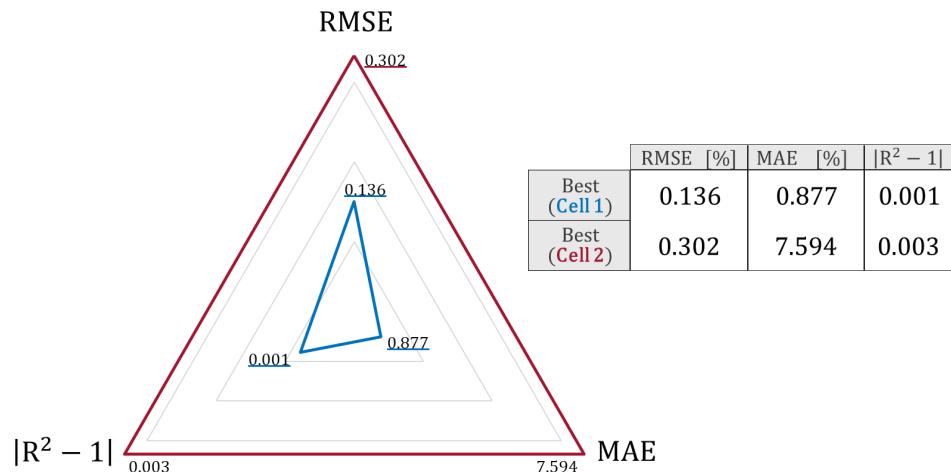


Figure 22. Spider plot of ANN performance based on the described indexes of the best cases of cell 1 and cell 2.

5.3. Comparison with Other Types of Regressive Nodes

In this section, the proposed network based on the combination of LSTM and FC nodes is compared with a similar network in which the LSTM nodes have been replaced with GRU nodes. The aim of the comparison is to prove that the LSTM nodes promise better long-term temporal dependency modeling, which has been considered beneficial for the nature of the charge/discharge sequences.

A GRU node is a simplified version of an LSTM node, which is designed to capture temporal dependencies in sequential data while using fewer parameters and a more compact architecture [33]. As mentioned, an LSTM node contains three gates: the input gate, the forget gate, and the output gate, along with a cell state that carries long-term memory across time steps. In contrast, a GRU node merges the cell and hidden states into a single state vector and uses only two gates:

- The *update gate*, which determines how much of the past information should be carried forward;
- The *reset gate*, which controls how much of the previous state should be forgotten when computing candidate activation.

The architecture of a GRU node is illustrated in Figure 23, and the model equations are carried out in (6)

$$\begin{aligned}
 z_t &= \sigma(W_z \cdot x_t + U_z \cdot h_{t-1} + b_z) \\
 r_t &= \sigma(W_r \cdot x_t + U_r \cdot h_{t-1} + b_r) \\
 \hat{h}_t &= \tanh(W_h \cdot x_t + U_h(r_t \odot h_{t-1}) + b_h) \\
 h_t &= (1 - z_t) \odot h_{t-1} + z_t \odot \hat{h}_t
 \end{aligned} \tag{6}$$

Because of this streamlined structure, GRUs are typically faster to train and require less computational power while still performing comparably to LSTMs in many sequence modeling tasks.

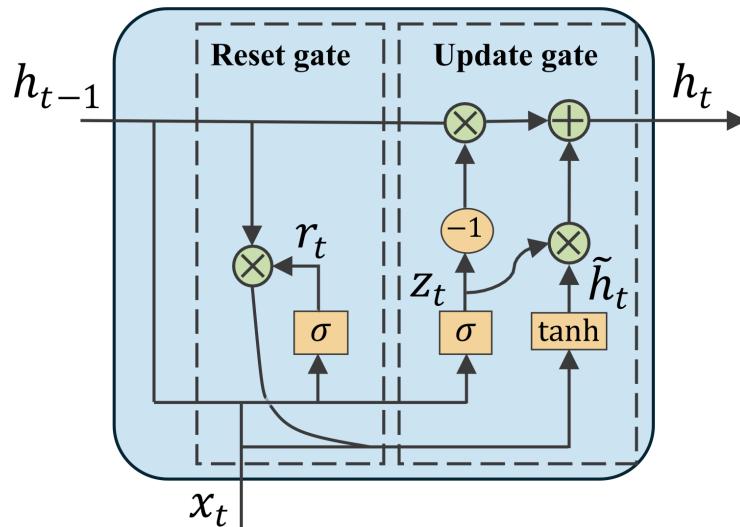


Figure 23. GRU architecture.

To provide a clearer assessment of the difference in performance between LSTM- and GRU-based architectures, a comparative experiment was conducted. Using the same LFP dataset, 180 GRU-based networks were trained by replacing the LSTM nodes in the original architecture with GRU nodes. The networks were trained using the same hyperparameter search strategy and within the same parameter range defined in this paper.

The performance of the GRU-based ANNs are presented in Table 6 in terms of the RMSE, MAE, and $|R^2 - 1|$. The hyperparameter of the best cases are derived from the training result of cell 1 (as explained in previous sections), and they are different for the two types of network. The RMSE and $|R^2 - 1|$ of the GRU-based networks are clearly higher than those of the LSTM-based networks for both cells in the best cases. This result can confirm the effectiveness of LSTM nodes versus the simpler structure of the GRU nodes due to the better long-term temporal dependency modeling of LSTM.

Table 6. LSTM and GRU nodes based on the best RMSE ANN performance of cell 1 and cell 2.

ANN with the Best RMSE	RMSE [%]	MAE [%]	$ R^2 - 1 $
Cell 1—LSTM	0.136	0.877	0.001
Cell 1—GRU	0.3961	2.9706	0.0051
Cell 2—LSTM	0.302	7.594	0.003
Cell 2—GRU	1.5342	9.5243	0.0714

To further investigate the relationship between hyperparameters and the RMSE of LSTM-based and GRU-based networks, Figure 24a (for LSTM-based networks) and Figure 24b (for GRU-based networks) are presented. Both figures consist of nine subplots that are organized into three groups according to the number of LSTM or GRU nodes. Within each group, three levels correspond to different values of FC nodes. The x - and y -axes represent the batch size and sub-sampling length, respectively. A color gradient from deep blue to light yellow is used to indicate RMSE performance: darker blue tones correspond to lower RMSE values, while light yellow indicates RMSE values exceeding 0.5.

In each figure, the red dots highlight the top 10 networks in terms of RMSE performance. From the comparison, several consistent patterns emerge for both LSTM and GRU models:

- Networks with a small number of recurrent (LSTM or GRU) nodes tend to perform worse in terms of RMSE;
- The most effective configurations typically use a batch size not exceeding 32;

- High-performing networks are distributed across all tested values of sub-sampling lengths and FC nodes, suggesting that these parameters are less critical to performance than the others.

This additional experiment supports the choice of LSTM over the GRU in the proposed design, as LSTM-based models consistently achieve lower RMSE values in the majority of configurations tested. The results in Figure 24a,b deal with the RMSE of cell 1 only, but the same results can be obtained for cell 2.

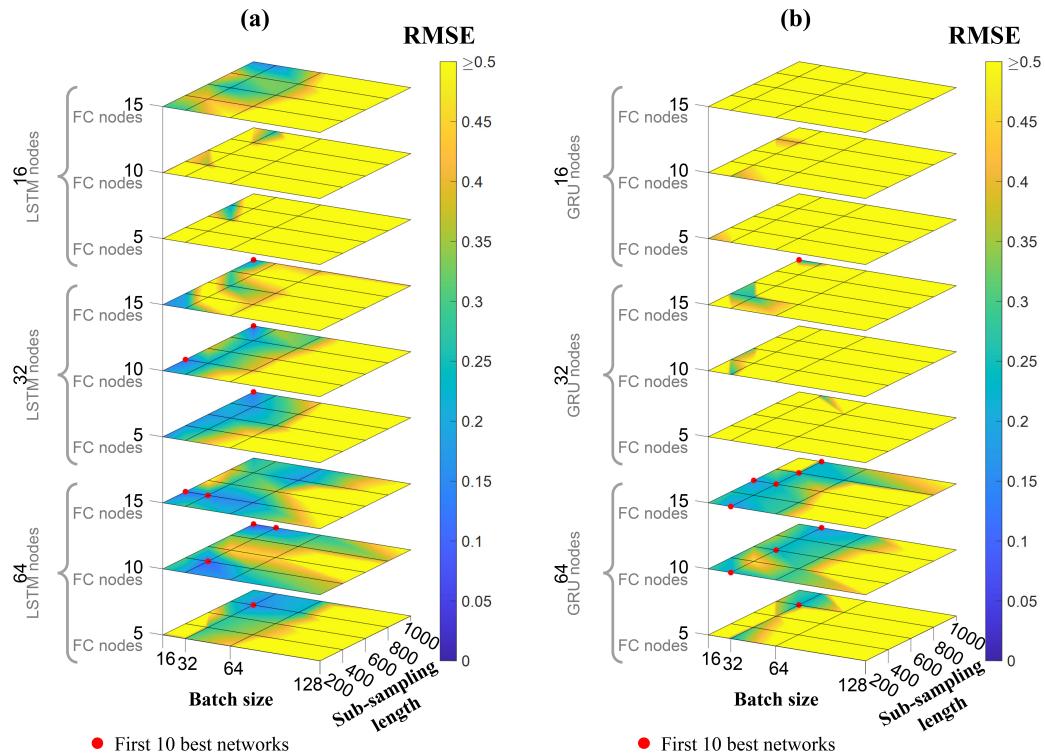


Figure 24. RMSE-based comparison of the exhaustive hyperparameter combinations for two ANN architectures: (a) LSTM-based and (b) GRU-based. The red dots indicate the ten best-performing networks in each configuration.

5.4. Cross-Chemistry Validation on Differently Aged NMC Cells

To further assess the robustness and generalization capability of the proposed method, this section extends the validation process by applying the same learning framework to a different lithium-ion chemistry. Due to the limited availability of LFP data, a more extensive dataset [42] is employed; it contains several NMC (nickel manganese cobalt) cells with a rated capacity of 5 Ah, each aged under distinct temperature and C-rate conditions. This cross-chemistry validation allows for a broader evaluation of the method's applicability to cells differing from those used for training.

Following the same protocol adopted for the LFP, the artificial neural network is trained exclusively on cell 1's dataset and then tested on six different cells aged in heterogeneous conditions in term of C-rates and ambient temperature. The left part of Table 7 summarizes the aging condition for all NMC cells.

The training procedure mirrors the one previously described, with hyperparameters optimized through a grid search involving various configurations of sub-sampling lengths (100 to 1000), batch sizes (32 to 512), and the number of fully connected nodes (1 to 20) and LSTM nodes (1 to 32). These ranges differ slightly from those adopted for the LFP cells, as they were empirically selected to better suit the NMC dataset characteristics. As for the LFP case, the ANN achieving the best RMSE on cell 1 is selected as the best ANN

and is tested on the other cells. The performance indexes are presented in the right side of Table 7. Notably, the RMSE on cell 1 is even lower than that observed on the LFP reference cell, confirming the model's effectiveness on the training data. While the prediction errors increase when applied to the other NMC cells, the results remain within acceptable bounds, consistently under 4%.

Table 7. Aging conditions and performance metrics (RMSE, MAE, and $|R^2 - 1|$) for NMC cells under different C-rates and temperatures.

Cell ID	Aging Conditions			Performance Results		
	C-Rate Charging [p.u.]	C-Rate Discharging [p.u.]	Temperature [°C]	RMSE [%]	MAE [%]	$ R^2 - 1 $ [p.u.]
1	0.2	0.2	25	0.0010	0.0030	0.0023
2	0.2	0.2	-5	1.9246	3.1990	0.5160
3	0.2	0.2	45	2.6928	13.6111	0.5076
4	1.5	1.5	25	1.1056	5.2750	0.0302
5	1.5	1.5	45	3.7931	9.8810	0.5324
6	2	2	25	1.8352	9.8996	0.0764
7	0.2	1.5	25	0.1159	0.9233	0.0017

Figures 25 and 26a,b illustrate the predicted capacity fade for the training cell and for two representative test cases.

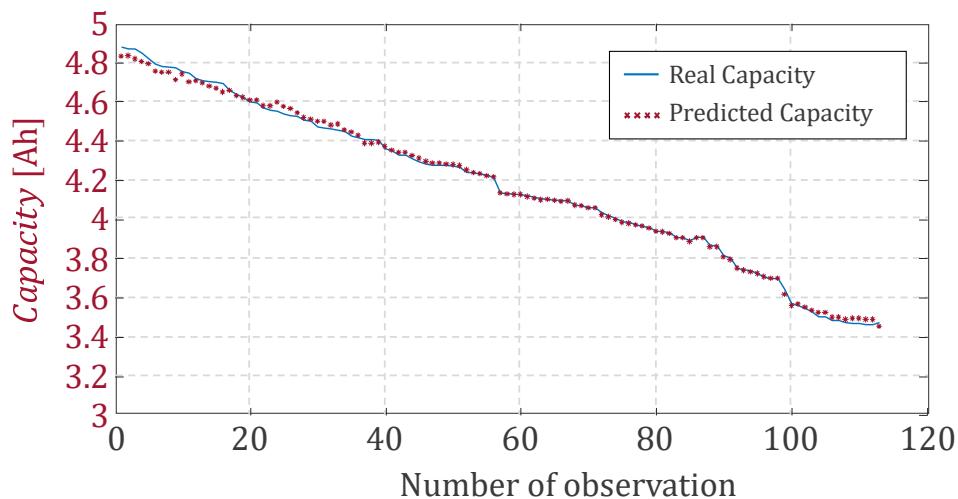


Figure 25. Capacity fade prediction for Cell 1.

Cell 7 shares the same temperature and charging C-rate as cell 1 but features a much higher discharge C-rate, over seven times greater. The ANN accurately tracks the degradation trend, with only minor deviations occurring after approximately 1200 observations. Conversely, cell 4, aged under a high C-rate for both charging and discharging, exhibits a divergence between the predicted and actual capacity after around 600 observations (equivalent to about 300 full cycles).

This behavior highlights a relevant consideration for practical applications. When the operating conditions of the test cell deviate significantly from those used in training, the model's accuracy naturally decreases beyond a certain usage horizon. However, the ability to maintain reliable predictions for up to 600 cycles under such conditions is noteworthy. In many real-world applications—such as electric vehicles—this corresponds

to approximately 1 to 2 years of operation. Providing accurate residual capacity estimates over such a period using a model trained on a single cell is a remarkable result, especially in the absence of complex handcrafted feature extraction or chemistry-specific tuning.

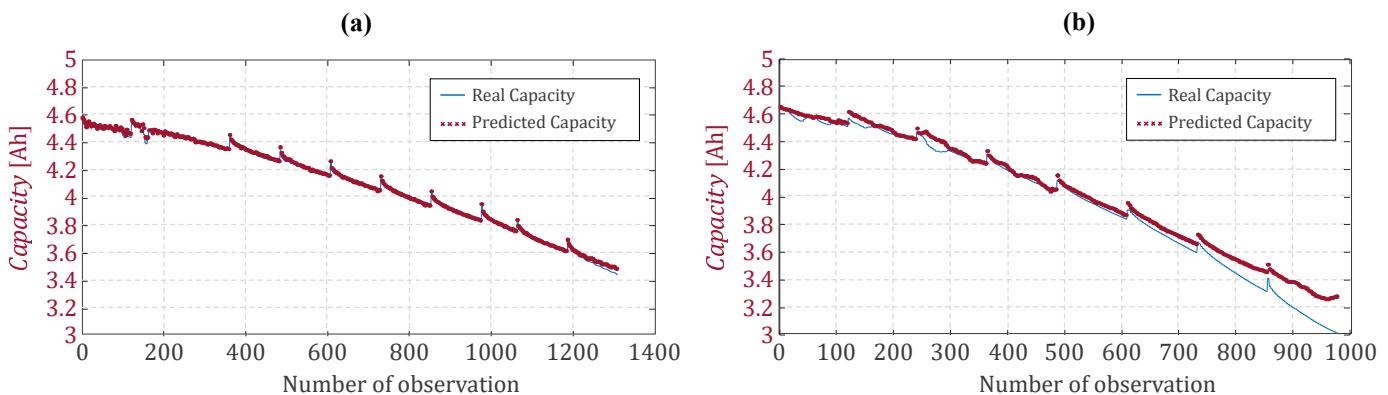


Figure 26. Capacity fade prediction for Cell 4 (a) and Cell 7 (b).

6. Conclusions

This study developed a custom artificial neural network (ANN) architecture that combines long short-term memory (LSTM) and fully connected (FC) layers for predicting the residual capacity (RC) of lithium iron phosphate (LFP) batteries. The primary objective was to design an effective training strategy and evaluate the estimation performance of this ANN architecture, leveraging a limited dataset derived from a single LFP battery cell.

The research methodology involved meticulous data preprocessing and comprehensive hyperparameter optimization. The dataset was partitioned into three subsets: a training set for model learning, a validation set for fine-tuning, and a testing set for performance evaluation. An exhaustive training campaign explored a wide range of hyperparameter configurations, identifying the optimal setup that minimized the root mean square error (RMSE) and max absolute error (MAE) while maximizing the coefficient of determination (R^2) on the testing set. The ANN demonstrated exceptional accuracy, with capacity estimations closely matching the measured values. The average estimation error was as low as 0.1%, with maximum errors remaining below 0.9% of the rated capacity. These results highlight the ANN's ability to deliver highly precise RC predictions, even when trained on a minimal dataset.

To validate the robustness of the proposed approach, the trained ANN was tested on a second LFP battery cell subjected to a distinct aging protocol. Importantly, the ANN was not exposed to this second cell's data during training or validation. On this new dataset, the ANN achieved an average estimation error of 0.2% with maximum errors under 2.5% of the rated capacity. These outcomes underscore the model's robustness and adaptability, demonstrating its capability to generalize effectively across cells with differing aging characteristics.

The proposed approach was also evaluated against a gated recurrent unit (GRU)-based model. In this comparison, the ANN architecture employing an LSTM layer achieved significantly lower estimation errors, confirming the benefits of including long-term memory structures in the network design.

Beyond the tests conducted on LFP chemistry, the method was also validated on a broader dataset comprising seven NMC cells aged under varying C-rates and ambient temperatures. In this more challenging scenario, the proposed ANN maintained excellent performance, highlighting its scalability across different chemistries and usage profiles.

Future research will aim to broaden validation efforts by incorporating data from cells aged under diverse depths of discharge (DoDs) and operational conditions, includ-

ing irregular discharge profiles derived from practical use cases. This expansion will support the development of more advanced neural network models, further enhancing the accuracy and reliability of capacity estimation for aged cells across a wider range of real-world applications.

Author Contributions: Conceptualization, D.I., R.M., M.R. and I.S.; Methodology, M.R. and R.M.; Software, M.R., R.M. and P.F.; Validation, M.R., R.M., P.F., D.I. and I.S.; Formal analysis, I.S., P.F., M.R. and R.M.; Investigation, M.R., R.M., P.F., D.I. and I.S.; Resources, D.I., M.R. and R.M.; Data curation, M.R. and R.M.; Writing—original draft preparation, M.R., R.M., P.F., D.I. and I.S.; Writing—review and editing, D.I., I.S., P.F. and M.R.; Visualization, M.R., P.F. and I.S.; Supervision, D.I. and I.S.; Project administration, D.I.; Funding acquisition, D.I. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by funding from the i-STENTORE (*Innovative Energy Storage Technologies TOWards increased Renewables integration and Efficient operation*), a part of the Horizon Europe Program, project ID 101096787, <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/projects-details/43108390/101096787> (accessed on 1 April 2025).

Data Availability Statement: Data are contained within the article.

Acknowledgments: This work originates from the Master’s thesis “Analisi e sviluppo di metodi di stima della capacità residua di batterie LiFePO₄ tramite reti neurali ricorrenti” by Roberta Merolla.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

<i>SoH</i>	State of Health
<i>ANNs</i>	Artificial Neural Networks
<i>RNNs</i>	Recurrent Neural Networks
<i>RC</i>	Residual Capacity
<i>RMSE</i>	Root Mean Square Error
<i>LIBs</i>	Lithium-ion batteries
<i>CV</i>	Constant Voltage
<i>FNNs</i>	Feed-forward Neural Networks
<i>EoL</i>	End of Life
<i>RUIL</i>	Remaining Useful Life
<i>CC</i>	Constant Current
<i>LSTM</i>	Long Short-Term Memory
<i>BoL</i>	Beginning of Life
<i>MAE</i>	Max Absolute Error
<i>GRU</i>	Gated Recurrent Unit

References

- IEA. Global EV Outlook 2018. 2018. Available online: <https://www.iea.org/reports/global-ev-outlook-2018> (accessed on 3 July 2025).
- Franzese, P.; Iannuzzi, D. Wireless Battery Charger Based on Sensorless Control for E-Bike Station. In Proceedings of the 2019 21st European Conference on Power Electronics and Applications (EPE’19 ECCE Europe), Genova, Italy, 2–6 September 2019; pp. P.1–P.10. [[CrossRef](#)]
- Iannuzzi, D.; Pagano, M.; Franzese, P.; Roscia, C. On-Board Energy Storage Systems Based on Lithium Ion Capacitors for LRT Energy Saving: Optimization Design Procedure. In Proceedings of the 2020 IEEE International Conference on Industrial Technology (ICIT), Buenos Aires, Argentina, 26–28 February 2020; pp. 717–722. [[CrossRef](#)]
- Franzese, P.; Iannuzzi, D.; Mottola, F.; Proto, D.; Pagano, M. Charging Strategies for Ultra-Fast Stations with Multiple Plug-In Electric Vehicle Parking Slots. In Proceedings of the 2020 AEIT International Annual Conference (AEIT), Catania, Italy, 23–25 September 2020; pp. 1–6. [[CrossRef](#)]

5. Franzese, P.; Di Pasquale, A.; Iannuzzi, D.; Pagano, M. Electric Ultra Fast Charging Stations: A Real Case Study. In Proceedings of the 2021 AEIT International Annual Conference (AEIT), Benevento, Italy, 15–17 September 2021; pp. 1–6. [CrossRef]
6. Dannier, A.; Brando, G.; Ribera, M.; Spina, I. Li-Ion Batteries for Electric Vehicle Applications: An Overview of Accurate State of Charge/State of Health Estimation Methods. *Energies* **2025**, *18*, 786. [CrossRef]
7. Spina, I.; Rogers, D.J.; Brando, G.; Chatzinikolaou, E.; Siwakoti, Y.P. Maximum Power per Ampere Modulation for Cascaded H-Bridge Converters. *IEEE J. Emerg. Sel. Top. Power Electron.* **2023**, *11*, 264–275. [CrossRef]
8. Spina, I.; Cervone, A. Energy Saving in Battery Electric Vehicles Equipped with Induction Machines and Modular Multilevel Converters. In Proceedings of the 2021 IEEE 15th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG 2021), Setúbal, Portugal, 14–16 July 2021. [CrossRef]
9. Brando, G.; Cervone, A.; Franzese, P.; Meo, S.; Toscano, L. Gain Scheduling Control with Minimum-Norm Pole-Placement Design of a Dual-Active-Bridge DC-DC Converter. In Proceedings of the 2020 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), Sorrento, Italy, 24–26 June 2020; pp. 846–851. [CrossRef]
10. The Boston Consulting Group. Batteries for Electric Cars: Challenges, Opportunities and the Outlook to 2020. Available online: <http://large.stanford.edu/courses/2016/ph240/enright1/docs/file36615.pdf> (accessed on 3 July 2025).
11. Reddy, T.B. *Linden's Handbook of Batteries*, 4th ed.; McGraw-Hill Education: New York, NY, USA, 2011.
12. Ohzuku, T.; Brodd, R.J. An overview of positive-electrode materials for advanced lithium-ion batteries. *J. Power Sources* **2007**, *174*, 449–456. [CrossRef]
13. Scrosati, B.; Garche, J. Lithium batteries: Status, prospects and future. *J. Power Sources* **2010**, *195*, 2419–2430. [CrossRef]
14. Ji, X. A paradigm of storage batteries. *Energy Environ. Sci.* **2019**, *12*, 3203–3224. [CrossRef]
15. Lee, S.; Mohtat, P.; Siegel, J.B.; Stefanopoulou, A.G.; Lee, J.W.; Lee, T.K. Estimation Error Bound of Battery Electrode Parameters With Limited Data Window. *IEEE Trans. Ind. Inform.* **2020**, *16*, 3376–3386. [CrossRef]
16. Mejdoubi, A.E.; Chaoui, H.; Gualous, H.; den Bossche, P.; Omar, N.; Mierlo, J.V. Lithium-Ion Batteries Health Prognosis Considering Aging Conditions. *IEEE Trans. Power Electron.* **2018**, *34*, 6834–6844. [CrossRef]
17. Brando, G.; Chatzinikolaou, E.; Rogers, D.; Spina, I. Electrochemical Cell Loss Minimization in Modular Multilevel Converters Based on Half-Bridge Modules. *Energies* **2021**, *14*, 1359. [CrossRef]
18. Tao, S.; Zhang, M.; Zhao, Z.; Li, H.; Ma, R.; Che, Y.; Sun, X.; Su, L.; Sun, C.; Chen, X.; et al. Non-destructive degradation pattern decoupling for early battery trajectory prediction via physics-informed learning. *Energy Environ. Sci.* **2025**, *18*, 1544–1559. [CrossRef]
19. Chen, Z.; Sun, M.; Shu, X.; Xiao, R.; Shen, J. Online State of Health Estimation for Lithium-Ion Batteries Based on Support Vector Machine. *Appl. Sci.* **2018**, *8*, 925. [CrossRef]
20. Dini, P.; Paolini, D. Exploiting Artificial Neural Networks for the State of Charge Estimation in EV/HV Battery Systems: A Review. *Batteries* **2025**, *11*, 107. [CrossRef]
21. Patil, M.; Tagade, P.; Hariharan, K.; Kolake, S.; Song, T.; Yeo, T.; Doo, S.G. A novel multistage Support Vector Machine based approach for Li ion battery remaining useful life estimation. *Appl. Energy* **2015**, *159*, 285–297. [CrossRef]
22. Guo, P.; Cheng, Z.; Yang, L. A data-driven remaining capacity estimation approach for lithium-ion batteries based on charging health feature extraction. *J. Power Sources* **2019**, *412*, 442–450. [CrossRef]
23. Hu, X.; Li, S.E.; Yang, Y. Advanced Machine Learning Approach for Lithium-Ion Battery State Estimation in Electric Vehicles. *IEEE Trans. Transp. Electrif.* **2016**, *2*, 140–149. [CrossRef]
24. Wu, J.; Zhang, C.; Chen, Z. An online method for lithium-ion battery remaining useful life estimation using importance sampling and neural networks. *Appl. Energy* **2016**, *173*, 134–140. [CrossRef]
25. You, G.-W.; Park, S.; Oh, D. Diagnosis of Electric Vehicle Batteries Using Recurrent Neural Networks. *IEEE Trans. Ind. Electron.* **2017**, *64*, 4885–4893. [CrossRef]
26. Wu, Y.; Xue, Q.; Shen, J.; Lei, Z.; Chen, Z.; Liu, Y. State of Health Estimation for Lithium-Ion Batteries Based on Healthy Features and Long Short-Term Memory. *IEEE Access* **2020**, *8*, 28533–28547. [CrossRef]
27. Zhang, W.; Li, X.; Li, X. Deep Learning-Based Prognostic Approach for Lithium-ion Batteries with Adaptive Time-Series Prediction and On-Line Validation. *Measurement* **2020**, *164*, 108052. [CrossRef]
28. de la Vega, J.; Riba, J.R.; Ortega-Redondo, J.A. Real-Time Lithium Battery Aging Prediction Based on Capacity Estimation and Deep Learning Methods. *Batteries* **2024**, *10*, 10. [CrossRef]
29. She, C.; Li, Y.; Zou, C.; Wik, T.; Wang, Z.; Sun, F. Offline and Online Blended Machine Learning for Lithium-Ion Battery Health State Estimation. *IEEE Trans. Transp. Electrif.* **2022**, *8*, 1604–1618. [CrossRef]
30. Ungurean, L.; Micea, M.; Carstoiu, G. Online state of health prediction method for lithium-ion batteries, based on gated recurrent unit neural networks. *Int. J. Energy Res.* **2020**, *44*, 6767–6777. [CrossRef]
31. Fan, Y.; Li, Y.; Zhao, J.; Wang, L.; Yan, C.; Wu, X.; Zhang, P.; Wang, J.; Gao, G.; Wei, L. Online State-of-Health Estimation for Fast-Charging Lithium-Ion Batteries Based on a Transformer–Long Short-Term Memory Neural Network. *Batteries* **2023**, *9*, 539. [CrossRef]

32. Choi, Y.; Ryu, S.; Park, K.; Kim, H. Machine Learning-Based Lithium-Ion Battery Capacity Estimation Exploiting Multi-Channel Charging Profiles. *IEEE Access* **2019**, *7*, 75143–75152. [[CrossRef](#)]
33. Zheng, Y.; Hu, J.; Chen, J.; Deng, H.; Hu, W. State of health estimation for lithium battery random charging process based on CNN-GRU method. *Energy Rep.* **2023**, *9*, 1–10. [[CrossRef](#)]
34. Huo, F.; Chen, C.-H. The State of Charge Estimation Based on GRU for Lithium-ion Batteries. In Proceedings of the 2022 IEEE 11th Global Conference on Consumer Electronics (GCCE), Osaka, Japan, 18–21 October 2022; pp. 220–221. [[CrossRef](#)]
35. Cui, S.; Joe, I. A Dynamic Spatial-Temporal Attention-Based GRU Model With Healthy Features for State-of-Health Estimation of Lithium-Ion Batteries. *IEEE Access* **2021**, *9*, 27374–27388. [[CrossRef](#)]
36. Alharbi, T.; Umair, M.; Alharbi, A. Lithium-Ion Battery State of Health Degradation Prediction Using Deep Learning Approaches. *IEEE Access* **2025**, *13*, 13464–13481. [[CrossRef](#)]
37. Ding, G.; Wang, W.; Zhu, T. Remaining Useful Life Prediction for Lithium-Ion Batteries Based on CS-VMD and GRU. *IEEE Access* **2022**, *10*, 89402–89413. [[CrossRef](#)]
38. Tang, X.; Liu, K.; Wang, X.; Gao, F.; Macro, J.; Widanage, W.D. Model Migration Neural Network for Predicting Battery Aging Trajectories. *IEEE Trans. Transp. Electrif.* **2020**, *6*, 363–374. [[CrossRef](#)]
39. Tan, Y.; Zhao, G. Transfer Learning With Long Short-Term Memory Network for State-of-Health Prediction of Lithium-Ion Batteries. *IEEE Trans. Ind. Electron.* **2020**, *67*, 8723–8731. [[CrossRef](#)]
40. Chen, L.; Ding, Y.; Wang, H.; Wang, Y.; Liu, B.; Wu, S.; Li, H.; Pan, H. Online Estimating State of Health of Lithium-Ion Batteries Using Hierarchical Extreme Learning Machine. *IEEE Trans. Transp. Electrif.* **2022**, *8*, 965–975. [[CrossRef](#)]
41. Kaur, K.; Garg, A.; Cui, X.; Singh, S.; Panigrahi, B. Deep learning networks for capacity estimation for monitoring SOH of Li-ion batteries for electric vehicles. *Int. J. Energy Res.* **2020**, *45*, 3113–3128. [[CrossRef](#)]
42. Mohtat, P.; Siegel, J.B.; Stefanopoulou, A.G.; Lee, S. UofM Pouch Cell Voltage and Expansion Cyclic Aging Dataset. *Dataset* **2021**. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.