

CHAPTER IV

INVESTIGATING THE IMPACT OF DATA RESAMPLING ON MODEL ACCURACY AND COMPLEXITY

4.1 Introduction

This chapter details the core experimental methodology and results of this study. The focus is on the application and evaluation of data resampling techniques for SOH estimation. The overall experimental framework for this chapter is illustrated in Fig. 4.1. The methodology will be applied to two distinct battery datasets: the Oxford Battery Degradation Dataset and the NASA Prognostics dataset. The performance of this approach will be systematically benchmarked across five different predictive models: LSTM, Bidirectional-LSTM (Bi-LSTM), AM-LSTM, GRU, and GPR.

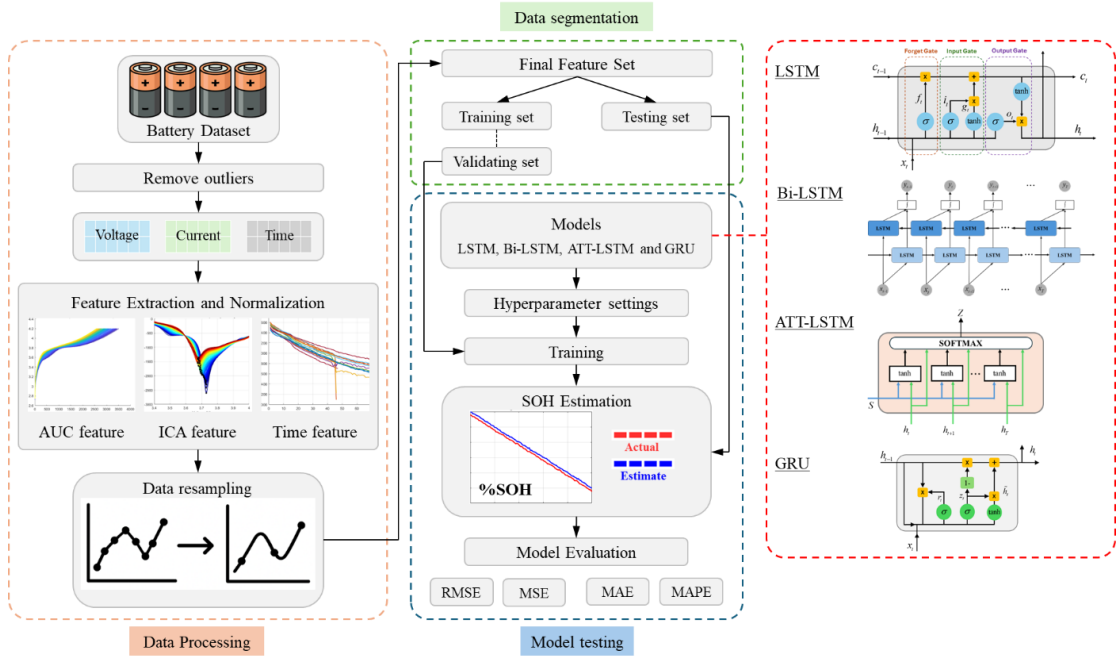


Figure 4.1 Experimental workflow for SOH estimation using data resampling.

4.2 Battery Datasets Utilized

In addition to the Oxford Battery Degradation Dataset, which was used and detailed in Chapter III, this chapter also incorporates the NASA Prognostics dataset for analysis and evaluation.

4.2.1 NASA Prognostics Center of Excellence Dataset

The NASA Prognostics Center of Excellence (PCoE) dataset comprises ageing data collected from four LIBs: B0005, B0006, B0007, and B0018. These cells were subjected to three operational modes: charging, discharging, and impedance measurements, all conducted under room temperature conditions. During charging, each battery was charged using a constant current (CC) mode at 1.5 A until the terminal voltage reached 4.2 V, after which charging continued in a constant voltage (CV) mode until the charging current decreased to 20 mA. The characteristic charging voltage profile is illustrated in Fig. 4.2, and the corresponding current profile is shown in Fig. 4.3. Discharge testing was performed using a constant current of 2 A, and the discharge process was terminated once the cell voltage dropped to predetermined cut-off levels, which were 2.7 V for cell B0005, 2.5 V for cells B0006 and B0018, and 2.2 V for cell B0007. Typical discharge voltage and current profiles are presented in Fig. 4.4 and Fig. 4.5, respectively.

Each battery's data was stored in a separate .mat file, and the recorded parameters included the measured voltage, current, and temperature, along with the corresponding charging current, charging voltage, and the timestamp in seconds. In this study, only the measured voltage and measured current were selected for both the charging and discharging cycles to ensure a consistent basis for analysis. It should be noted that this dataset contains certain limitations. Some cells exhibit inconsistencies in data logging, and a few segments show the presence of noise. For example, in some charging cycles, even after the cell voltage had reached 4.2 V and the current dropped to zero, the voltage occasionally decreased slightly, a phenomenon that is considered non-ideal and treated as noise. Similarly, during discharging, some records continued even after the battery had reached its designated cut-off voltage. To ensure data

integrity, all such anomalous segments were removed during a dedicated data-cleaning phase prior to feature extraction.

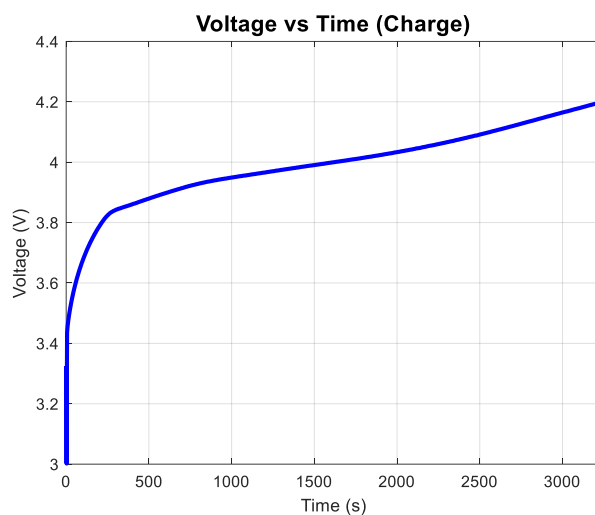


Figure 4.2 Voltage profiles of B0005 during charging cycle from the NASA Battery

Dataset

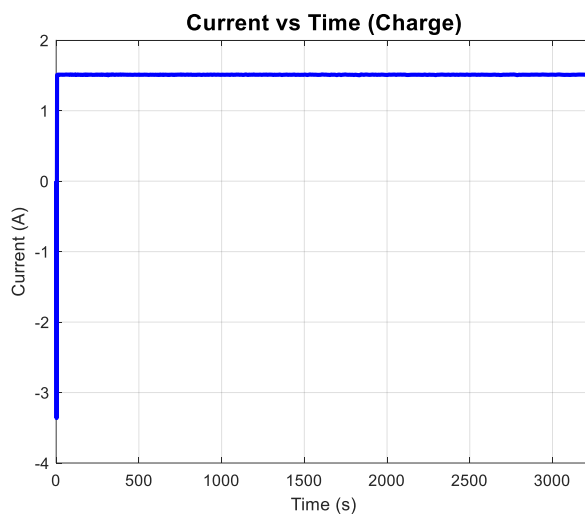


Figure 4.3 Current profiles of B0005 during charging cycle from the NASA Battery

Dataset

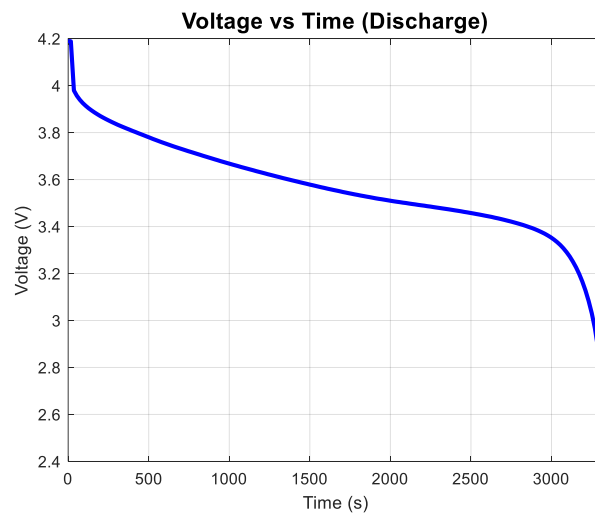


Figure 4.4 Voltage profiles of B0005 during discharging cycle from the NASA Battery

Dataset

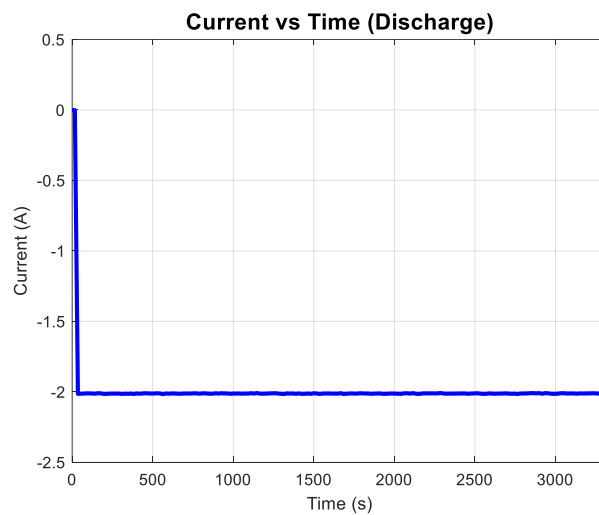


Figure 4.5 Current profiles of B0005 during discharging cycle from the NASA Battery

Dataset

The degradation behavior under different discharge cut-off voltages and operating profiles, the SOH trends of the four LIBs in the NASA dataset were plotted.

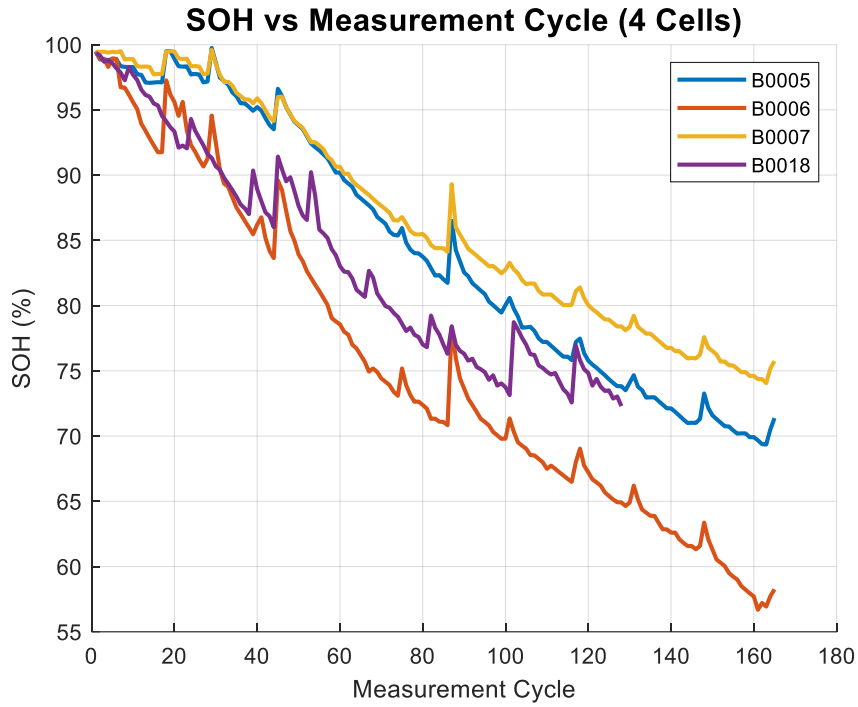


Figure 4.6 SOH degradation curves of all 8 cells from the Oxford Battery Degradation Dataset

4.2.2 Dataset Partitioning Strategy

Following the data preparation described in Chapter 3, the datasets were partitioned based on battery cell identity. This ensures that the testing set consists of a completely "unseen" battery. The specific allocation for both Oxford and NASA datasets is summarized in Table 4.1.

Table 4.1 Hyperparameter Settings for Training

Dataset	Training Set	Validation Set	Testing Set
Oxford	Cells 1, 2, 3, 4	Cells 5, 6, 7	Cell 8
NASA	B0006, B0007	B0005	B0018

4.3 Data Resampling

Data resampling is a process used to alter the temporal resolution of time-series data. This can involve increasing the data frequency (upsampling), often through interpolation, or decreasing it (downsampling), which involves selecting data at sparser

intervals rather than using every available data point. The primary goal of resampling is to adapt the dataset to improve the efficiency, speed, or analytical suitability of the model being applied. This study focuses exclusively on downsampling. The objective is to simplify the input data fed into the deep learning models, thereby reducing the computational load and training time, while critically aiming to preserve the accuracy of the battery's SOH estimation. To implement this, a fixed-step downsampling strategy was employed. Recognizing that the two datasets possess different characteristics and data logging frequencies, a distinct range of resampling intervals was applied to each. These specific intervals were determined based on preliminary exploratory experiments to establish a range that captures the spectrum from minor data reduction to highly sparse data. The selected intervals were:

- 1) For the Oxford Dataset: Intervals of 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 300, and 400 seconds were tested.
- 2) For the NASA Dataset: Intervals of 10, 15, 20, 25, 30, 40, and 60 seconds were tested.

For both datasets, the performance at these downsampled intervals was compared against the baseline (labeled 'None'), which utilizes the original, high-resolution (non-resampled) data. This systematic approach allows for a thorough evaluation of the threshold at which reduced data still maintains acceptable accuracy and the point at which information loss begins to significantly degrade model performance. Ultimately, this experiment serves as a key step in identifying the optimal trade-off between data volume and model performance.

The determination of resampling intervals in this research was based on criteria tailored to the specific temporal characteristics of each dataset, ensuring that the resampled data retained essential features for accurate estimation:

For the Oxford dataset, the interval selection was based on the Relative Temporal Resolution framework. Given the average charging duration (T_{avg}) of approximately 3,600 seconds, a baseline interval of 20 seconds was selected. This value represents approximately 0.5% of the total duration, serving as a high-fidelity baseline. Subsequent intervals were increased in arithmetic progression to investigate the impact of resolution degradation.

Conversely, for the NASA dataset, the interval selection was guided by the Signal-to-Noise Ratio Enhancement principle. Due to the inherent sampling irregularity of approximately 2.5 to 3 seconds in the raw data, a baseline interval of 10 seconds was established.

This value corresponds to a smoothing factor (k) of approximately 3 to 4 times the raw sampling rate, chosen based on Variance Reduction principles. Statistically, aggregating data within this range significantly mitigates random noise and provides a necessary safety margin against sampling jitter. Consequently, the resampling process effectively acts as a low-pass filter, standardizing time-series stability before model training.

4.4 Feature Extraction Methodology

Following the data cleaning and preprocessing for both the Oxford and NASA datasets, the feature extraction process was performed. This involved calculating key degradation-aware indicators such as the AUC, Charging and Discharging Time Duration, and features derived from ICA.

4.4.1 Area under curve

The detailed methodology for calculating the AUC as a health-aware feature was previously established in Section 3.3.1.

For the Oxford dataset, the identical set of AUC features presented in that section is utilized in this analysis. For the NASA dataset, the AUC features were calculated anew. To illustrate the raw data characteristics from which these features were extracted, representative cycle profiles for cell B0005 are shown in Figs. 4.7–4.10: the charge voltage profile (Fig. 4.7), charge current profile (Fig. 4.8), discharge voltage profile (Fig. 4.9), and discharge current profile (Fig. 4.10). The corresponding AUC data for the remaining NASA cells are presented in the Appendix B.

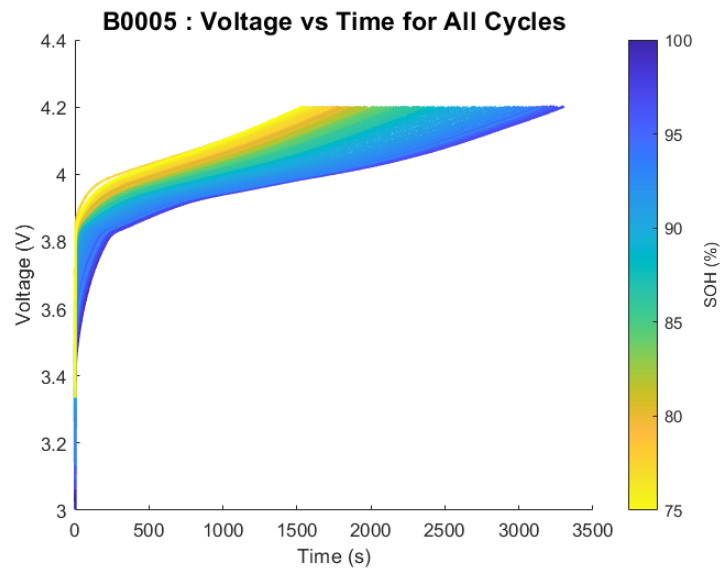


Figure 4.7 Voltage profiles B0005 of NASA are used to charge all the cycles

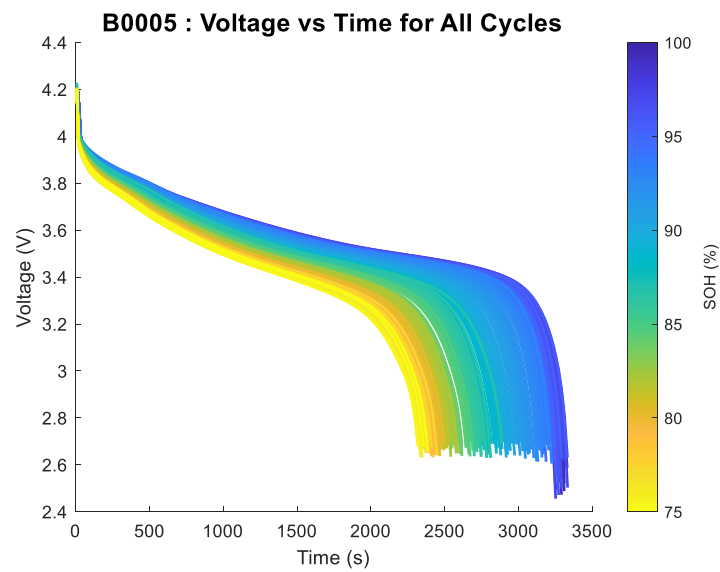


Figure 4.8 Voltage profiles B0005 of NASA are used to discharge all the cycles

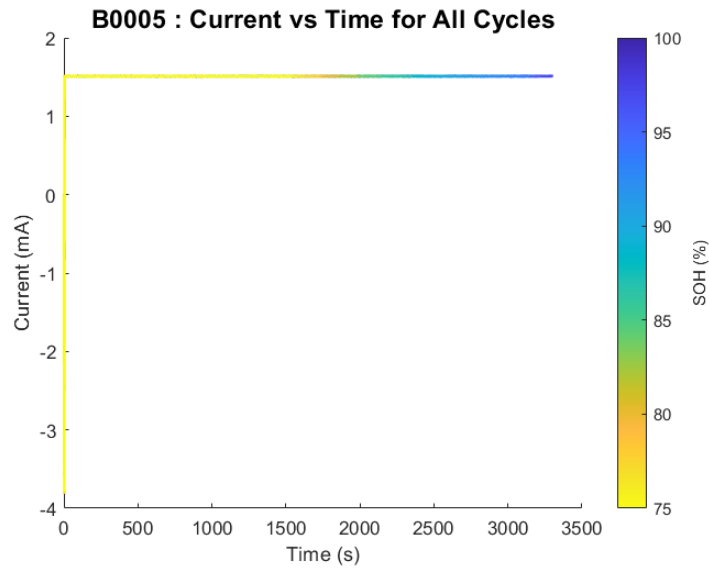


Figure 4.9 Current profiles B0005 of NASA are used to charge all the cycles

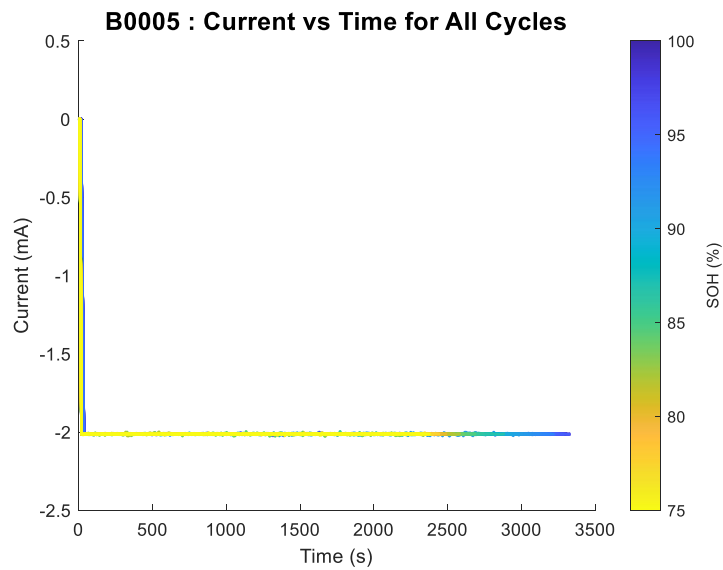


Figure 4.10 Current profiles B0005 of NASA are used to discharge all the cycles

After applying the AUC calculation to all cycles for all NASA cells, the resulting degradation features were normalized using Min-Max Normalization for visualization. The complete normalized AUC degradation trends for all NASA cells (B0005, B0006, B0007, and B0018) are presented in Figs. 4.11–4.14, showing the charge voltage AUC (Fig. 4.11), discharge voltage AUC (Fig. 4.12), charge current AUC (Fig. 4.13), and discharge current AUC (Fig. 4.14), respectively.

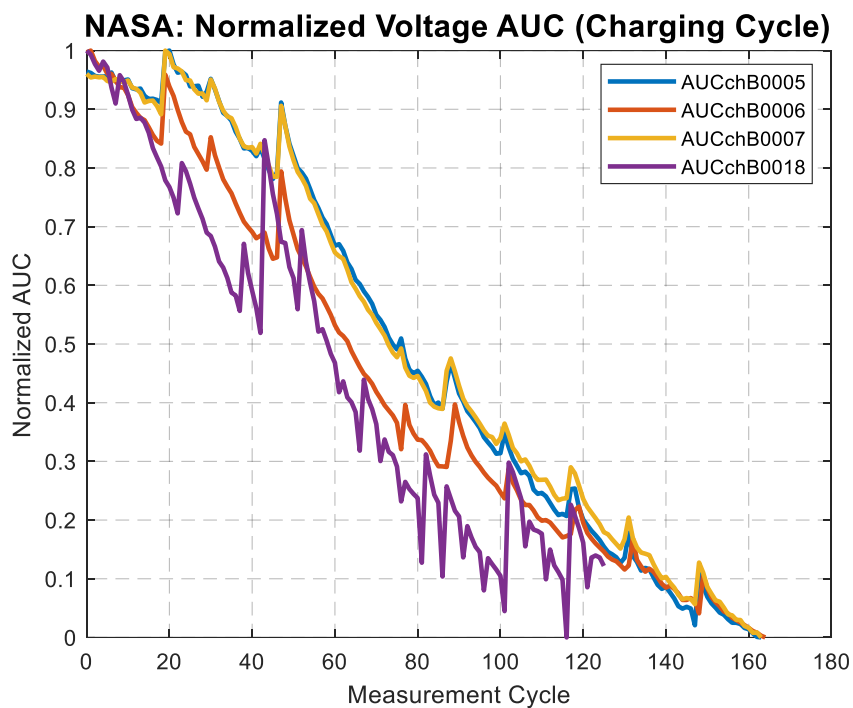


Figure 4.11 Normalized AUC of voltage during the charging cycle from the NASA dataset

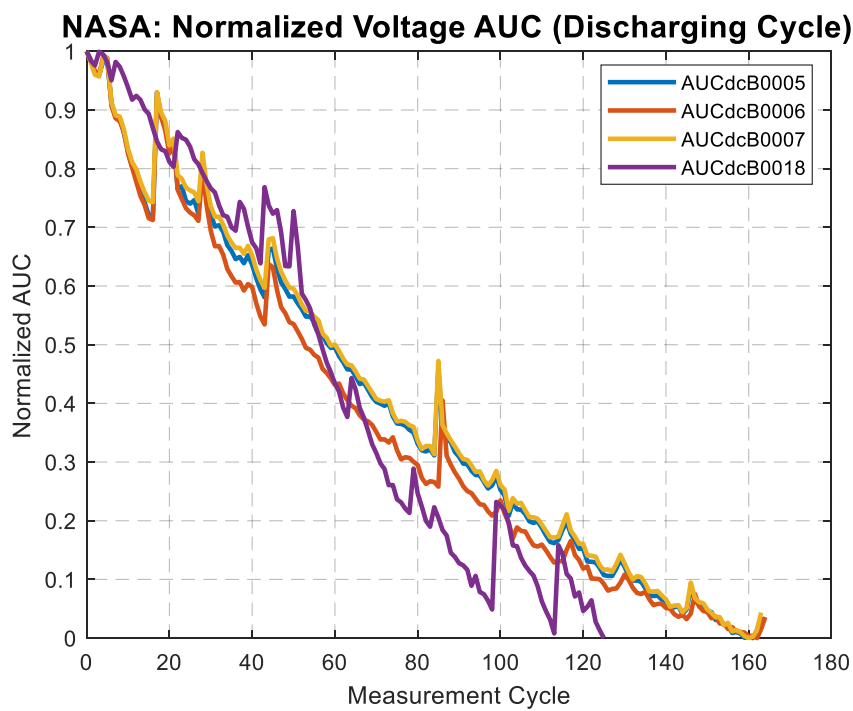


Figure 4.12 Normalized AUC of voltage during the discharging cycle from the NASA dataset

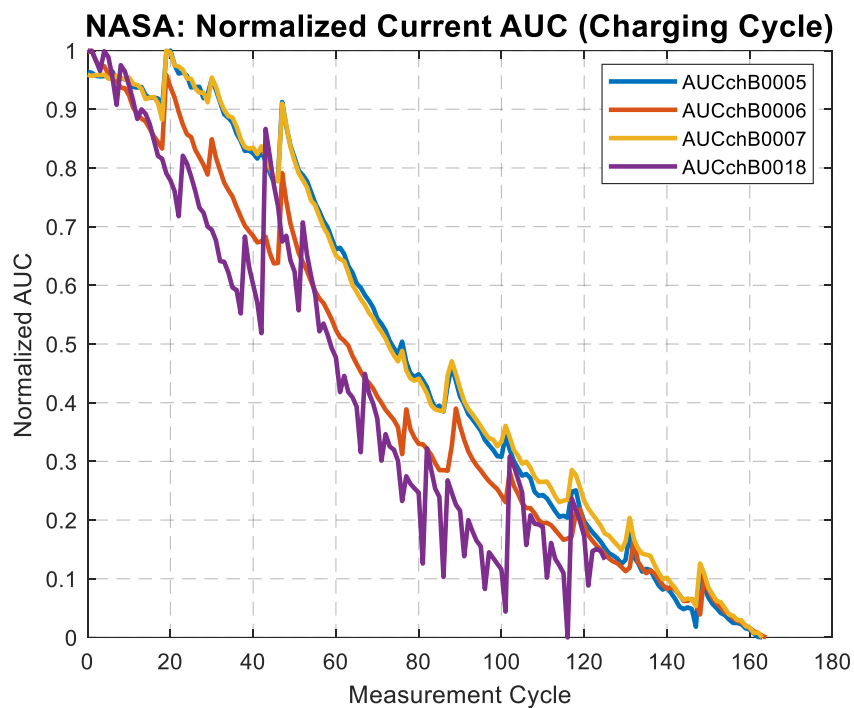


Figure 4.13 Normalized AUC of current during the charging cycle from the NASA dataset

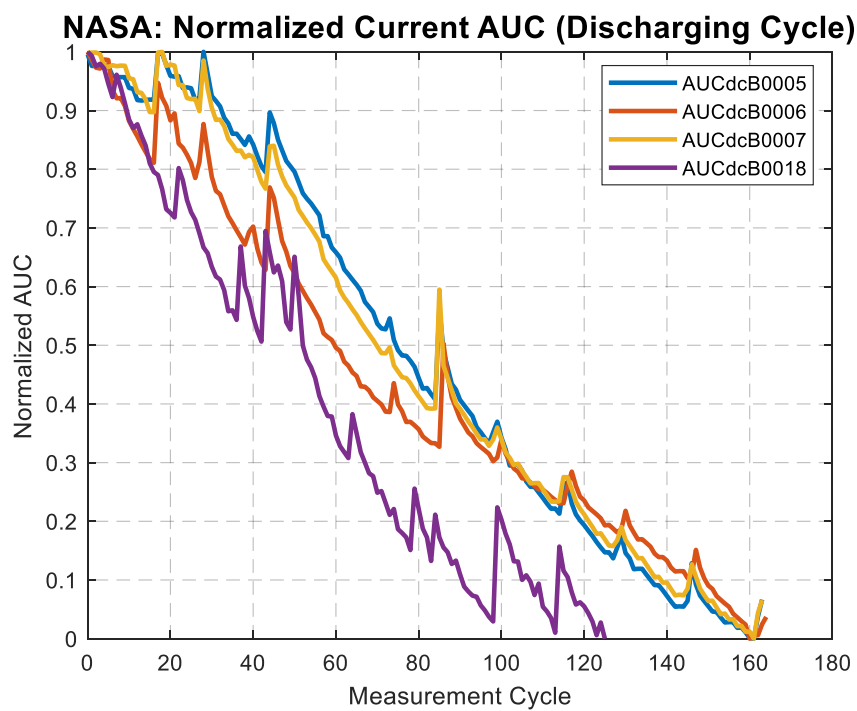


Figure 4.14 Normalized AUC of current during the discharging cycle from the NASA Dataset

The overall trends of the normalized AUC values derived from both voltage and current data show a similar decreasing pattern to the battery's SOH, indicating a strong correlation between AUC behavior and battery degradation.

4.4.2 The charging and discharging durations

For both the Oxford and NASA datasets, the charging and discharging processes were recorded with corresponding timestamps. Therefore, the final recorded time value for each cycle can be extracted and used as a feature representing the charging and discharging duration.

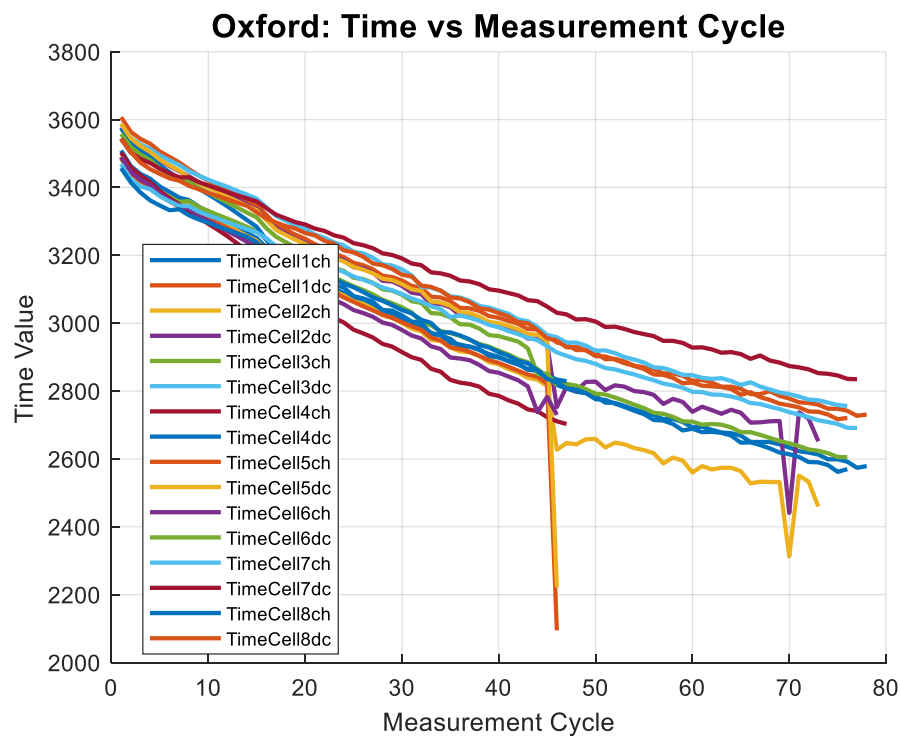


Figure 4.15 Charging and discharging time duration across cycles from the Oxford Dataset

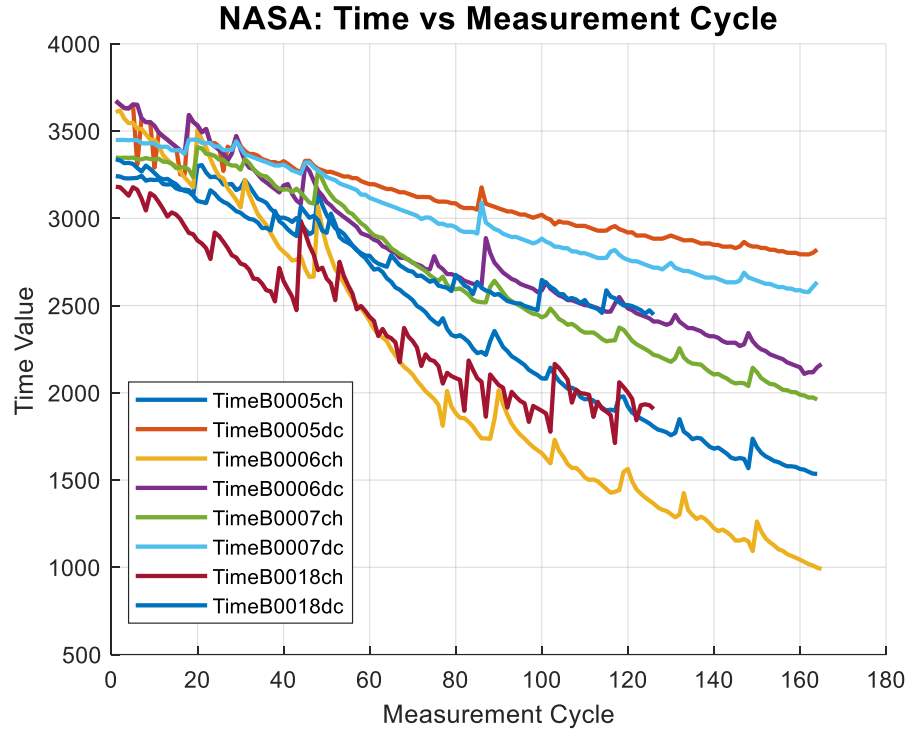


Figure 4.16 Charging and discharging time duration across cycles from the NASA dataset

As shown in Fig. 4.15, which presents the charging and discharging duration from the Oxford dataset, and Fig. 4.16 from the NASA dataset, the overall trend of time duration tends to decrease with increasing cycle count, which corresponds to the decline in the battery's SOH.

4.4.3 Incremental Capacity Analysis

ICA is an in-depth analytical technique used to investigate the behavior and diagnose the health of batteries during charging or discharging processes. It relies on calculating the derivative of electrical capacity (Q) with respect to voltage (V), known as shown in the equation (4.1)

$$\frac{dQ}{dV} \quad (4.1)$$

The hallmark of the ICA technique is its ability to transform the relatively flat voltage plateaus, which normally occur in a voltage-capacity (V - Q) graph, into distinct peaks or valleys on the dQ/dV plot. These peaks correspond directly to

the phase transitions of the active materials within the electrodes. The characteristics of these peaks, such as their height, area, width, and voltage position, are highly sensitive to the electrochemical changes within the battery cell. As the battery degrades, major degradation mechanisms such as Loss of Lithium Inventory (LLI), Loss of Active Material (LAM), and Increased Internal Resistance will cause these peak characteristics to change for example, the peaks may shift, shrink (in height or area), or broaden. Consequently, tracking the changes in key features on the dQ/dV curve such as the minimum point or peaks can serve as an effective indicator for evaluating the degradation of the battery's ability to store or deliver charge. This reflects internal structural changes, such as the loss of active material surface area or the aforementioned lithium loss. For this reason, ICA has been widely adopted in battery aging studies and is considered an effective indicator for evaluating the battery's SOH.

To improve the accuracy of the analysis, the raw dQ/dV data must first undergo a noise filtering process. In this study, a Moving Average technique was applied to smooth the dQ/dV curves. For the Oxford dataset, a window size of 400 was used, while for the NASA dataset, which has different characteristics, a smaller window size of 25 was chosen. These values were selected to match the characteristics and resolution of each dataset, enabling clearer and more reliable identification of battery degradation trends throughout its life cycle. Examples of the calculation results for Cell 1 of the Oxford dataset are shown in Fig. 4.17, and for B0005 of the NASA dataset are shown in Fig. 4.18. The ICA profiles for the remaining batteries from the Oxford and NASA datasets are provided in Appendix C and Appendix D, respectively.

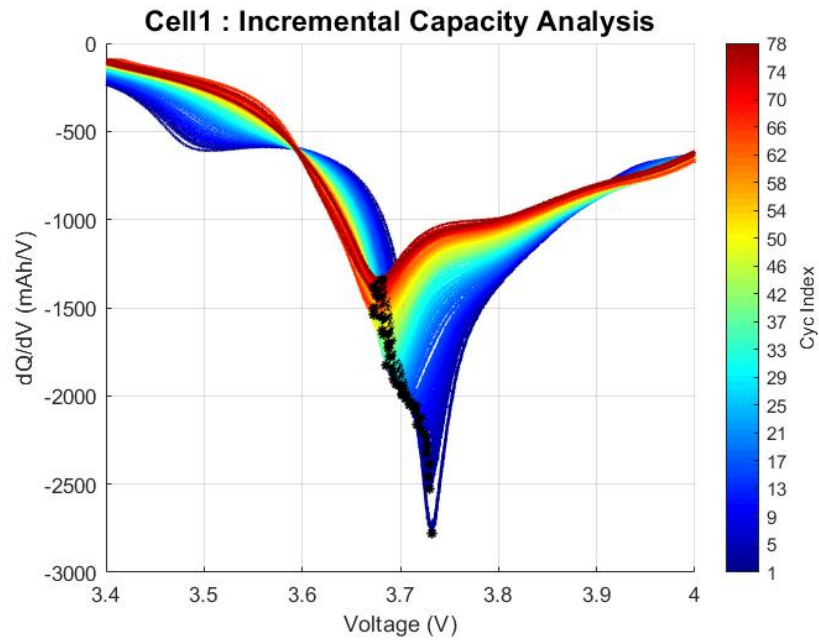


Figure 4.17 ICA curves for Oxford Cell 1

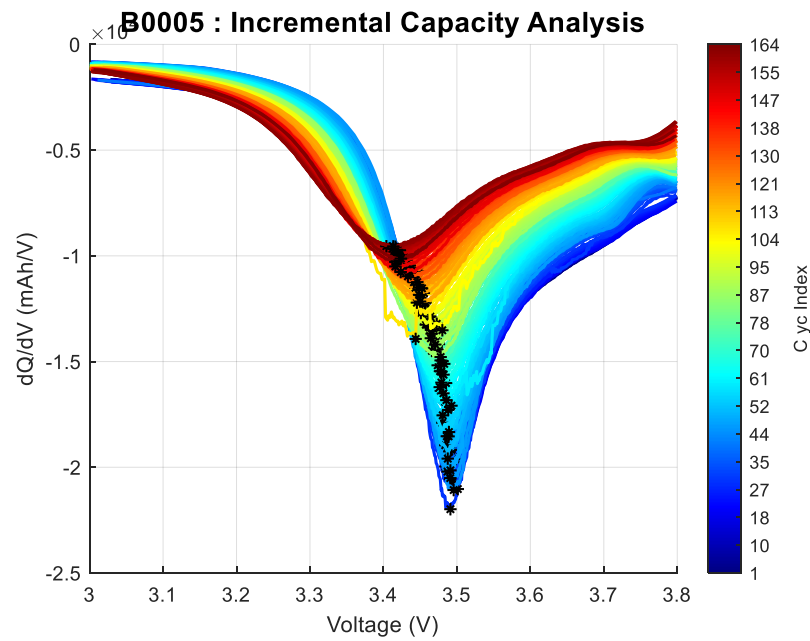


Figure 4.18 ICA curves for NASA B0005

As illustrated in the figures, the ICA curves reveal noticeable changes as the number of cycles increases. In this study, the minimum ICA value (the value at a key dQ/dV valley) from each cycle is selected as a feature for extraction. This feature is chosen because it is directly related to the electrochemical changes within the battery cell and can effectively reflect the battery's degradation condition.

All extracted features, including AUC, time-based features, and ICA features, are normalized using Min-Max Normalization, scaling the values into a 0 to 1 range. This process allows features with different units and scales to be fairly compared and processed on the same footing. It also helps the learning algorithm to converge faster and prevents the model from being biased toward features with larger numerical ranges. As a result, the normalization process enhances the model's stability and prediction accuracy over time.

4.5 Feature Correlation Analysis

After all relevant features (AUC, time-based features, and ICA) were extracted, a correlation analysis was performed to investigate the relationships between these features themselves (inter-feature correlation) and, more importantly, their correlation with the battery's SOH. For this analysis, a Correlation Matrix was constructed, following the same approach mentioned in Chapter 3, to evaluate the linear relationships between variables. The results for the Oxford dataset are presented in Fig. 4.19, and the results for the NASA dataset are shown in Fig. 4.20.

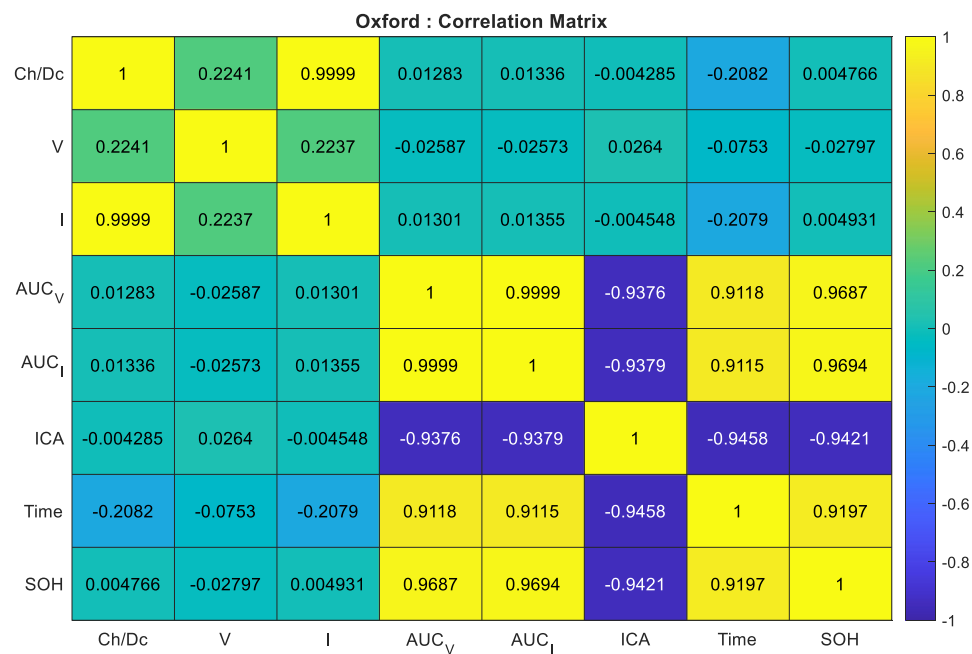


Figure 4.19 The correlation matrix from the Oxford dataset

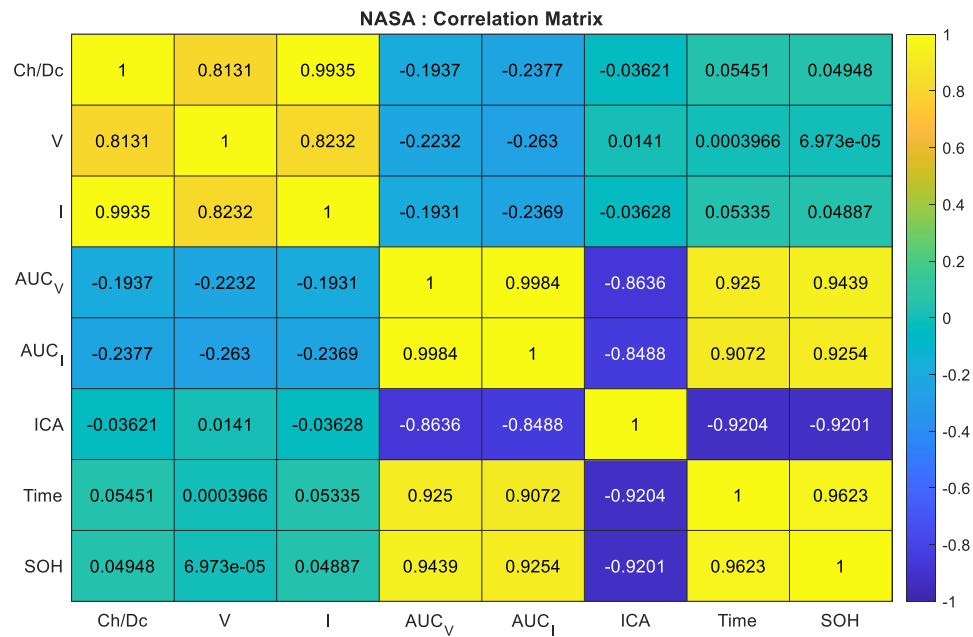


Figure 4.20 The correlation matrix from the NASA dataset

An examination of both correlation matrices reveals that all extracted features show a very high degree of correlation with the battery's SOH. The features AUC_v , AUC_i , and Time all exhibit an extremely strong positive correlation. In the Oxford dataset (Fig. 4.19), their coefficients are 0.9687, 0.9694, and 0.9197, respectively. The NASA dataset (Fig. 4.20) shows clearly consistent results, with values of 0.9439, 0.9254, and 0.9623, respectively. This strong positive correlation confirms that as the battery degrades (SOH decreases), the values of these features (such as the AUC or the time duration) also decrease, which aligns with the expected physical behavior of a degrading battery.

Meanwhile, the ICA feature, which was extracted from the minimum value of the dQ/dV curve, shows a very strong negative correlation (a coefficient close to -1) with SOH, posting a value of -0.9421 in the Oxford dataset and -0.9201 in the NASA dataset. This strong negative coefficient is highly significant; it does not imply it is a poor feature but rather indicates a clear inverse relationship. In other words, as the battery's SOH decreases (the battery degrades), the extracted ICA feature value (the value at the minimum point/valley) tends to increase. The fact that the coefficient is nearly -1 confirms that the change in this dQ/dV valley is also a highly reliable

indicator of degradation. In summary, this analysis confirms that all selected features (AUC_v , AUC_i , ICA, and Time) are excellent predictors of SOH, supporting their suitability for use in the subsequent modeling phase.

4.6 Hyperparameter Settings

The performance and accuracy of Deep Learning models depend heavily on proper hyperparameter settings. In this study, to ensure a fair and standardized comparison between models, and specifically to clearly evaluate the effectiveness of the Resampling technique without the influence of differing model settings, most basic parameters were kept identical for the LSTM, Bi-LSTM, and GRU models.

Key training parameters, including the number of hidden units, maximum epochs, mini-batch size, as well as the learning rate schedule and optimizer details, are summarized in Table 4.2.

However, the Attention-LSTM model required slightly different settings. The number of Hidden Units was set to 256, with an additional Attention Dimension (attnDim) of 64. This adjustment was necessary to satisfy the structural constraints of the Multi-Head Attention mechanism, which require the input dimension (Hidden Units) to be exactly divisible by the number of attention heads. Using 200 units, as in the other models, caused calculation errors due to this mismatch. In contrast, 256 (a power of two) ensures full compatibility and efficient operation within the attention layer's architecture.

Table 4.2 Hyperparameter Settings for Training

Parameter	Value
Number of Hidden Units	200
Fully Connected Layer	500
Max Epochs	100
Mini-Batch Size	1024
Initial Learn Rate	0.01
Learn Rate Drop Factor	0.5
Learn Rate Drop Period	40 epochs
Optimizer	Adam

The parameters listed in Table 4.2 were selected to balance accuracy with computational efficiency. Certain key values were adopted from the optimization methodology in (Jafari & Byun, 2025) as a robust baseline, while others were empirically fined-tuned. A more in-depth analysis regarding hyperparameter tuning and its impact on model performance will be presented and discussed in the following section.

4.7 Experimental Results and Discussion

This section presents the experimental results, organized into two main parts. First, the Hyperparameter Tuning Analysis details the process of identifying optimal model settings. Second, the Comparative SOH Estimation Results evaluates and compares the accuracy of the proposed models using these finalized parameters.

4.7.1 Hyperparameter Tuning Analysis

In this study, we performed hyperparameter tuning primarily on the Oxford dataset (specifically the version resampled at 100 cycles). Our goal was to find the settings that allow the model to learn degradation patterns most effectively.

Once the best parameters were identified, we applied these exact same settings to the NASA dataset (resampled at 20 cycles) without any further adjustments. We did this to test the model's robustness and ability to generalize. This approach

proves that the proposed model architecture works well even on datasets with different characteristics.

To determine the optimal network architecture, we used a Grid Search strategy. We simultaneously varied two key parameters: the number of hidden units (numHiddenUnits) and the size of the fully connected layer (hiddenLayerSize), using the RMSE as our main guide. The numerical results are listed in Table 4.2, and the trends are visualized in the bar chart in Fig. 4.21.

Table 4.2 RMSE results from the Grid Search analysis on the Oxford dataset

Hidden Layer Size (Nodes)	Num of Hidden Units		
	100	200	300
200	0.4224	0.3357	0.4736
300	0.3975	0.3215	0.4127
400	0.358	0.3286	0.3436
500	0.3353	0.3045	0.3239
600	0.3887	0.4376	0.3684
700	0.3872	0.3378	0.4783

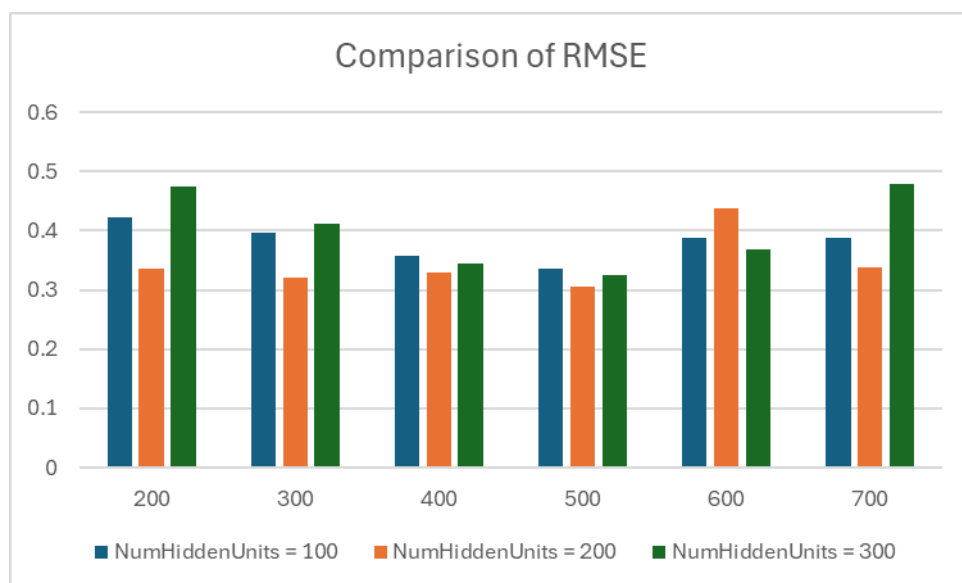


Figure 4.21 Comparison of RMSE values across different model architecture configurations

Impact of Hidden Layer Size: As clearly shown by the orange bars (where numHiddenUnits is fixed at 200), the RMSE tends to decrease as the hiddenLayerSize increases from 200 up to 500. The error drops from 0.3357 down to 0.3045 at size 500, which is the lowest point (minimum error) in this experiment. However, when we increased the size further to 600, the RMSE spiked significantly to 0.4376. This suggests the model became too complex (over-parameterized), leading to overfitting, where it started memorizing noise instead of learning the actual patterns.

Impact of Hidden Units: Similarly, if we compare the height of the bars at a hiddenLayerSize of 500, we see that using 100 units (blue bar) or 300 units (green bar) results in higher errors compared to 200 units (orange bar). Their values are 0.3353 and 0.3239, respectively. This indicates that 200 units offer the best balance between learning capability and model complexity.

Based on this grid search analysis, we confirmed that the combination of (numHiddenUnits = 200, hiddenLayerSize = 500) is the optimal choice for this dataset. Therefore, this setting was selected as the standard for the comparative experiments in the next section.

4.7.2 Comparative SOH Estimation Results

The performance evaluation of the four deep learning models LSTM, Bi-LSTM, GRU, AM-LSTM and the statistical model (GPR) begins with an analysis of the Oxford dataset, which is characterized by its high-resolution data. The quantitative results presented in Table 4.3 and Table 4.4 highlight significant differences in both estimation accuracy and computational time between using raw data and applying the proposed resampling technique.

When considering the baseline case Non-resampling, most models struggled to predict SOH accurately. This is illustrated in Fig. 4.21 Baseline Comparison, which compares the prediction results of all models against the actual values using raw data. The prediction curves for the standard models, particularly LSTM and GRU, exhibited considerable fluctuations throughout the battery's life cycle. This visual observation aligns with the high RMSE values of 0.922 and 1.548, respectively. The

volatility seen in Fig. 4.22 reflects the presence of significant noise in the raw data, preventing the models from effectively capturing the global degradation trend, resulting in jagged curves that fail to track the actual SOH closely.

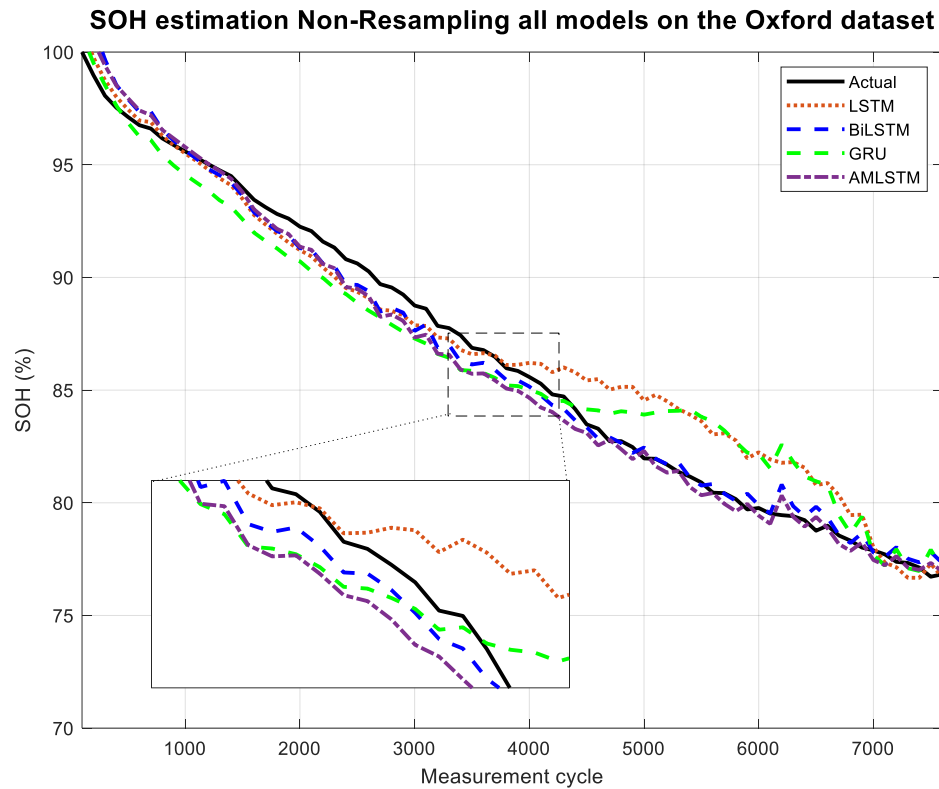


Figure 4.22 Baseline comparison of SOH estimation results using raw data
(Non-resampling) for all models on the Oxford dataset

However, when the data underwent resampling at the optimal interval of 100 seconds, the results improved significantly, as shown in Fig. 4.23 Proposed Method Comparison. The prediction curves for all models became much smoother and less volatile. Comparing the models, the proposed AM-LSTM (purple line) demonstrated the best ability to track the actual value curve (black line), especially during the final stages of the battery life where degradation is most severe. This is supported by the data in Table 4.3, where AM-LSTM achieved the lowest RMSE of 0.292, compared to LSTM (0.305) and Bi-LSTM (0.311).

It is also important to address the specific status of the GPR model in Table 4.3. Readers will notice that the entry for GPR under the raw data ('None')

category is marked with a hyphen (-). This symbol indicates that the model could not be successfully executed in this scenario due to its inherent computational limitations; GPR faces significant challenges with large-scale datasets as its computational complexity increases drastically with the number of data points. Consequently, it was unable to process the massive volume of the raw Oxford data directly. However, once the data was reduced via the resampling technique at the 100-second interval, the GPR model became operational. As shown in Fig. 4.23, although its resulting RMSE of 1.541 was higher than that of the deep learning models, the GPR model offers a distinct advantage by providing a 95% confidence interval. This feature offers valuable insight into the uncertainty of the predictions, even if the model lacks the capacity to capture the complex, non-linear relationships of the battery data as effectively as the neural networks.

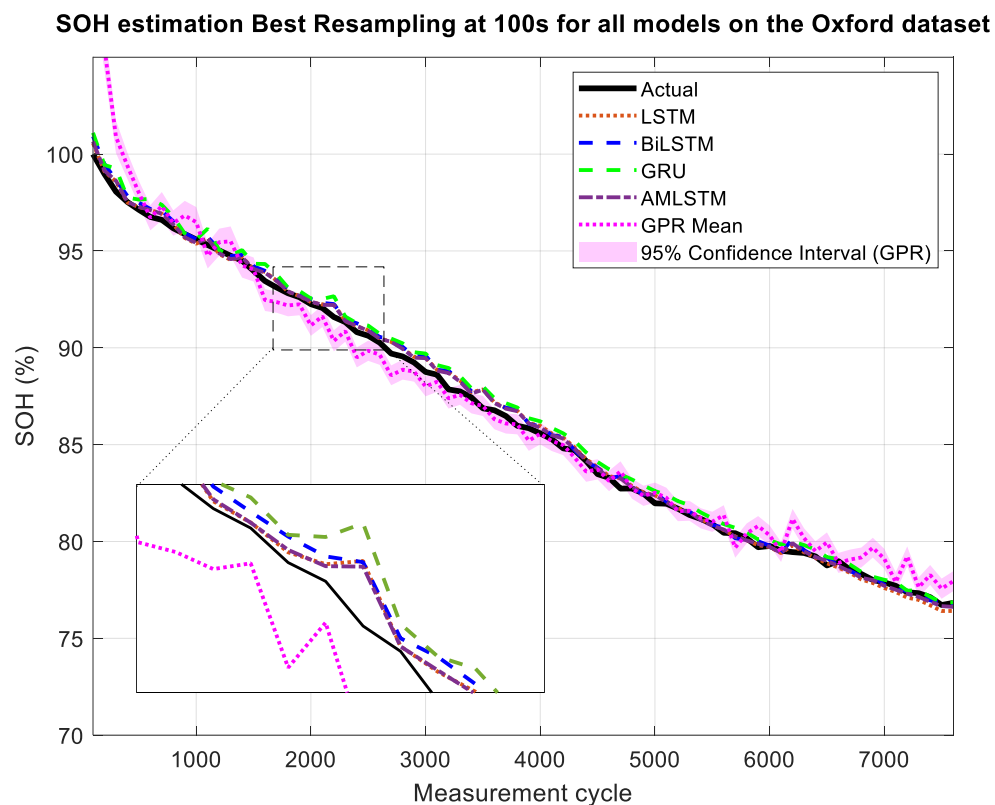


Figure 4.23 Comparison of SOH estimation results using the Best Resampling at 100s
For all models on the Oxford dataset.

To confirm the necessity and effectiveness of the proposed approach, the specific impact of the resampling technique on the AM-LSTM model is analyzed in Fig. 4.24 Impact of Resampling Analysis. This graph clearly contrasts the prediction using raw data (yellow line) versus the prediction using data resampled at 100 seconds (blue line). As seen in the figure, the yellow line is widely scattered with high error, whereas the blue line is tight, concise, and converges accurately to the actual SOH values. This visual evidence confirms that selecting an appropriate resampling interval effectively removes redundant noisy data, allowing the model to learn the underlying degradation structure more clearly.

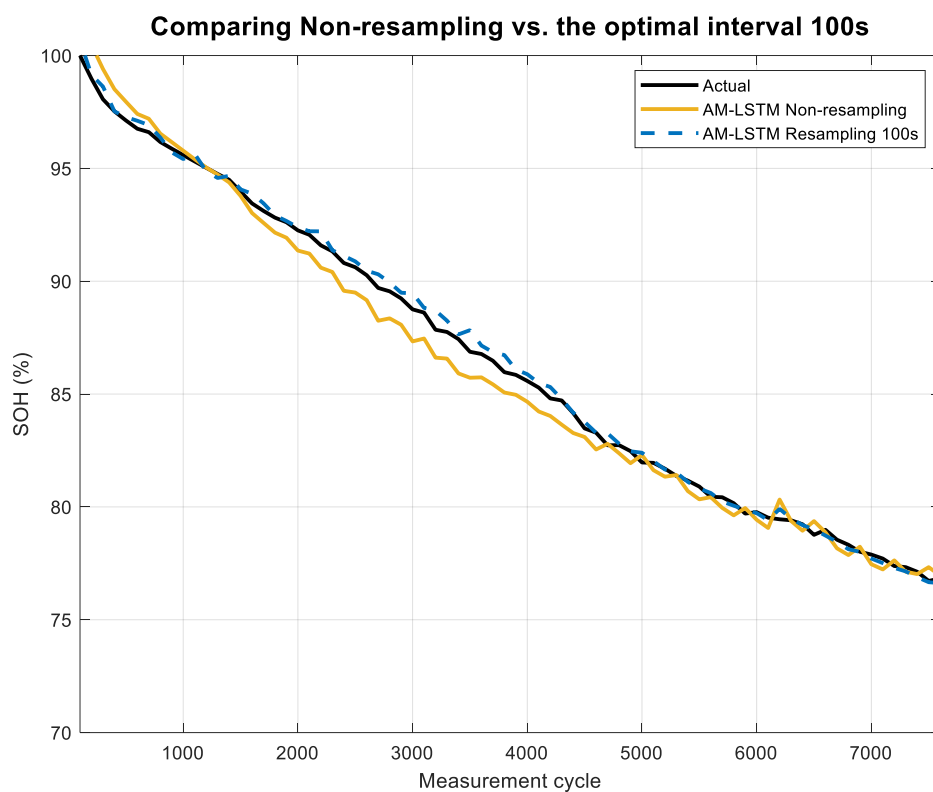


Figure 4.24 Analysis of the impact of resampling on the AM-LSTM model (Oxford dataset) comparing Non-resampling vs. the optimal interval (100s).

However, a key observation from the experiments is that a trade-off exists: expanding the resampling interval too much can negatively affect accuracy. As indicated in Table 4.3, when the resampling interval extends beyond the optimal point (e.g., beyond 120 seconds), the RMSE of the model starts to rise again. This

phenomenon is caused by excessive information loss; subtle degradation patterns are removed along with the noise. Consequently, the prediction curve becomes too coarse and fails to capture the detailed behavior of the battery. Therefore, the 100-second mark represents the "sweet spot" for this study, balancing maximum noise reduction with the preservation of essential data required for model learning. Fig. 4.25 illustrates the RMSE comparison of all models for the Oxford dataset.

Beyond improvements in accuracy, computational efficiency is another major achievement of this study. Data from Table 4.4 reveals that training the models with raw data requires a substantial amount of time. For instance, the AM-LSTM model took approximately 1.59×10^4 seconds (about 4.4 hours). However, by applying resampling at 100 seconds, the processing time dropped significantly to just 169.1 seconds (under 3 minutes). This dramatic reduction in training time is not merely due to data reduction, but because the model can learn faster when the input is compact and noise-free. This allows the optimization process to converge more easily, making the proposed method highly suitable for real-world applications where both accuracy and speed are critical.

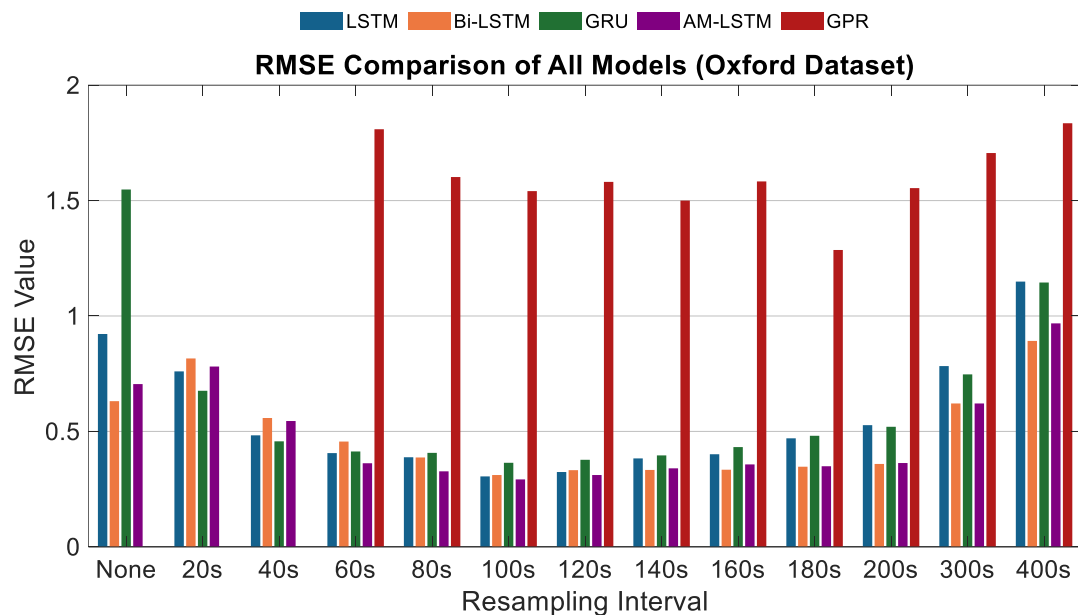


Figure 4.25 RMSE comparison of all models on the Oxford dataset

Table 4.3 Comparison of SOH Estimation Results on Oxford Dataset

Dataset	Resampling Interval	LSTM				Bi-LSTM				GRU				AM-LSTM				GPR			
		RMSE	MSE	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE	MAE	MAPE
Oxford	None	0.922	0.850	0.722	0.810	0.631	0.398	0.523	0.596	1.548	2.397	1.331	1.564	0.705	0.497	0.585	0.669	-	-	-	-
	20 s	0.760	0.577	0.609	0.684	0.816	0.666	0.667	0.747	0.676	0.456	0.400	0.445	0.781	0.610	0.592	0.658	-	-	-	-
	40 s	0.483	0.233	0.364	0.420	0.558	0.312	0.427	0.478	0.457	0.209	0.362	0.433	0.545	0.297	0.452	0.534	-	-	-	-
	60 s	0.406	0.165	0.332	0.385	0.456	0.208	0.382	0.450	0.413	0.170	0.328	0.373	0.362	0.131	0.286	0.339	1.809	3.272	0.869	0.972
	80 s	0.388	0.150	0.340	0.299	0.387	0.150	0.311	0.367	0.407	0.166	0.317	0.363	0.327	0.107	0.256	0.294	1.602	2.566	0.867	0.986
	100 s	0.305	0.093	0.242	0.279	0.311	0.097	0.249	0.290	0.364	0.132	0.294	0.336	0.292	0.085	0.232	0.271	1.541	2.375	0.828	0.929
	120 s	0.324	0.105	0.257	0.296	0.332	0.110	0.254	0.289	0.377	0.142	0.308	0.350	0.311	0.097	0.239	0.271	1.581	2.500	0.907	1.029
	140 s	0.383	0.147	0.314	0.360	0.333	0.111	0.263	0.299	0.396	0.157	0.325	0.369	0.340	0.115	0.270	0.305	1.500	2.250	0.892	1.019
	160 s	0.401	0.161	0.325	0.368	0.334	0.112	0.258	0.292	0.432	0.186	0.349	0.396	0.357	0.128	0.278	0.313	1.583	2.505	1.048	1.203
	180 s	0.470	0.221	0.389	0.440	0.347	0.120	0.272	0.308	0.481	0.232	0.394	0.447	0.349	0.122	0.278	0.316	1.286	1.654	0.817	0.929
Oxford	200 s	0.527	0.278	0.446	0.503	0.359	0.129	0.287	0.326	0.520	0.270	0.448	0.508	0.363	0.132	0.279	0.313	1.554	2.416	0.950	1.083
	300 s	0.783	0.613	0.694	0.795	0.621	0.385	0.527	0.591	0.747	0.557	0.669	0.760	0.621	0.385	0.527	0.593	1.706	2.909	0.985	1.124
	400 s	1.149	1.321	1.006	1.156	0.892	0.795	0.790	0.893	1.145	1.311	0.999	1.152	0.968	0.937	0.849	0.971	1.835	3.366	1.040	1.200

Table 4.4 Execution Time of Each Model for SOH Estimation on Oxford Dataset

Dataset	Resampling	LSTM	Bi-LSTM	GRU	AM-LSTM	GPR
	Interval	(s)	(s)	(s)	(s)	(s)
Oxford	None	14,056	17,071	13,403	15,891	-
	20 s	447.4362	556.6291	403.2034	848.2249	-
	40 s	217.598	261.4257	205.7885	382.1823	-
	60 s	160.9488	184.5931	139.2921	264.0181	1187.7686
	80 s	105.9322	139.2008	103.1165	228.6973	282.9889
	100 s	84.0947	115.3445	85.6528	169.1163	148.4209
	120 s	70.8816	97.7156	72.7832	131.5215	100.683
	140 s	58.675	82.5298	59.7725	106.8197	73.4618
	160 s	51.8306	66.6126	51.5024	93.0716	56.3858
	180 s	48.366	64.9802	51.0482	96.7402	48.5944
	200 s	44.8542	57.1521	45.14	84.3684	43.021
	300 s	30.0948	40.1414	30.8055	57.0947	27.4732
	400 s	23.5741	28.96	23.4261	41.0468	26.0582

Following the analysis of the Oxford dataset, this study extended the evaluation to the NASA dataset to assess the model's generalization capability and robustness under different data conditions. The quantitative results regarding estimation errors and computational time are summarized in Table 4.5 and Table 4.6, respectively.

Upon examining the results, the behavior of the NASA dataset differs interestingly from the Oxford one. In the baseline case using raw data (None), most models performed reasonably well, as shown in Fig. 4.26 Baseline Comparison. Notably, the AM-LSTM model (purple line) achieved the lowest error with an RMSE of 0.2351, representing the best performance in this experiment. This indicates that the Attention-LSTM architecture is highly effective at handling the raw NASA data directly without requiring complex preprocessing.

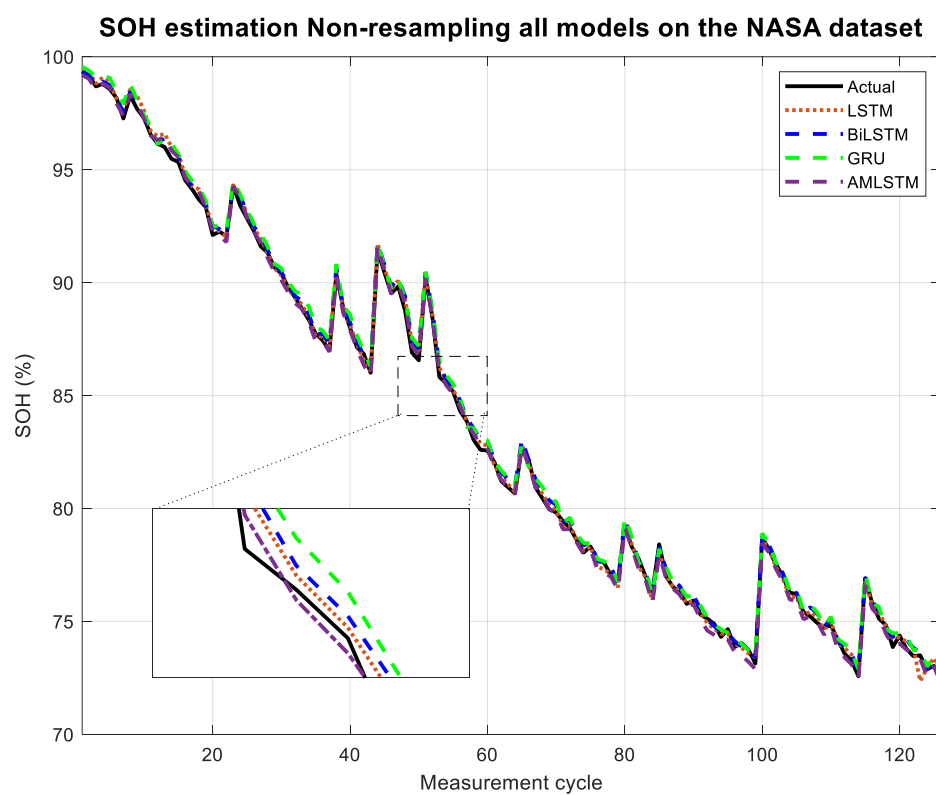


Figure 4.26 Baseline comparison of SOH estimation results using raw data (Non-resampling) for all models on the NASA dataset

However, a key insight emerges when applying the resampling technique at a 20-second interval (the optimal setting for this dataset). The results demonstrate the robustness of the model. Although the RMSE of the AM-LSTM increased slightly to 0.311, it remained highly accurate compared to other baseline models. This is visible in Fig. 4.27 Robustness Check, where the prediction curve

continues to track the actual degradation trend satisfactorily. This proves the model is flexible enough to handle data reduction without a critical loss of performance.

SOH estimation using data resampled at 20s for all models on the NASA dataset

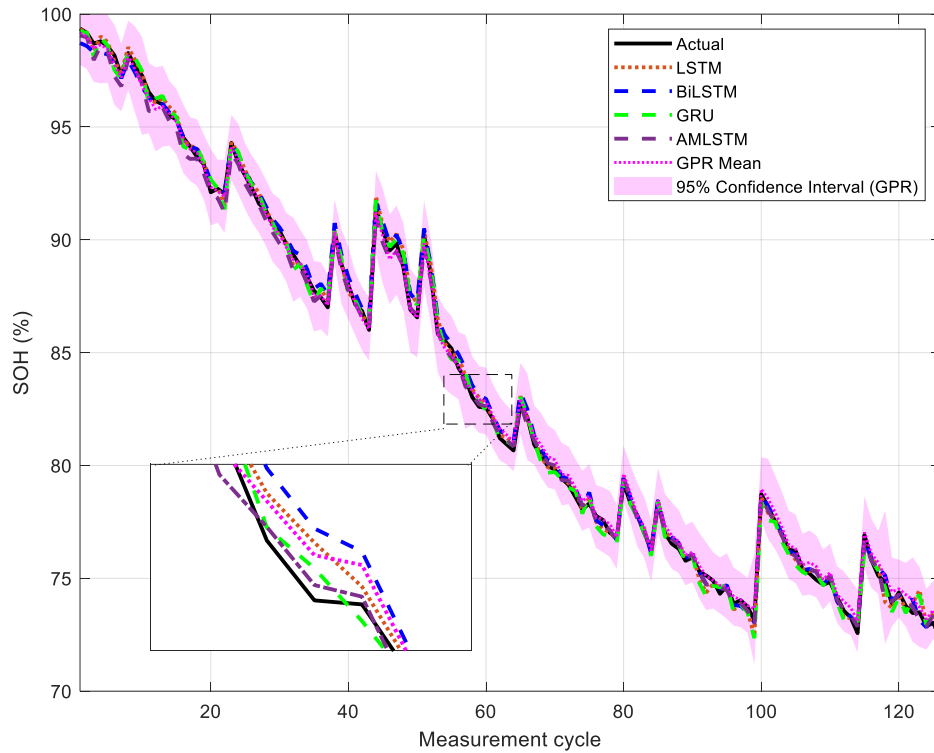


Figure 4.27 Comparison of SOH estimation results using the Best Resampling at 20s for all models on the NASA dataset

To further illustrate this robustness and the choice of the optimal interval, Fig. 4.28 Impact of Resampling Analysis compares the AM-LSTM model using raw data (yellow line) versus the 20-second resampling (blue line). It can be observed that the two lines almost overlap and follow the same trajectory smoothly, confirming that the 20-second interval is a "Safe Zone" for data reduction. Conversely, if the resampling interval is extended too far (e.g., 40 or 60 seconds), the error spikes rapidly RMSE exceeding 1.45, as seen in Table 4.5 due to the loss of fine-grained details which are much smaller in the NASA dataset compared to Oxford.

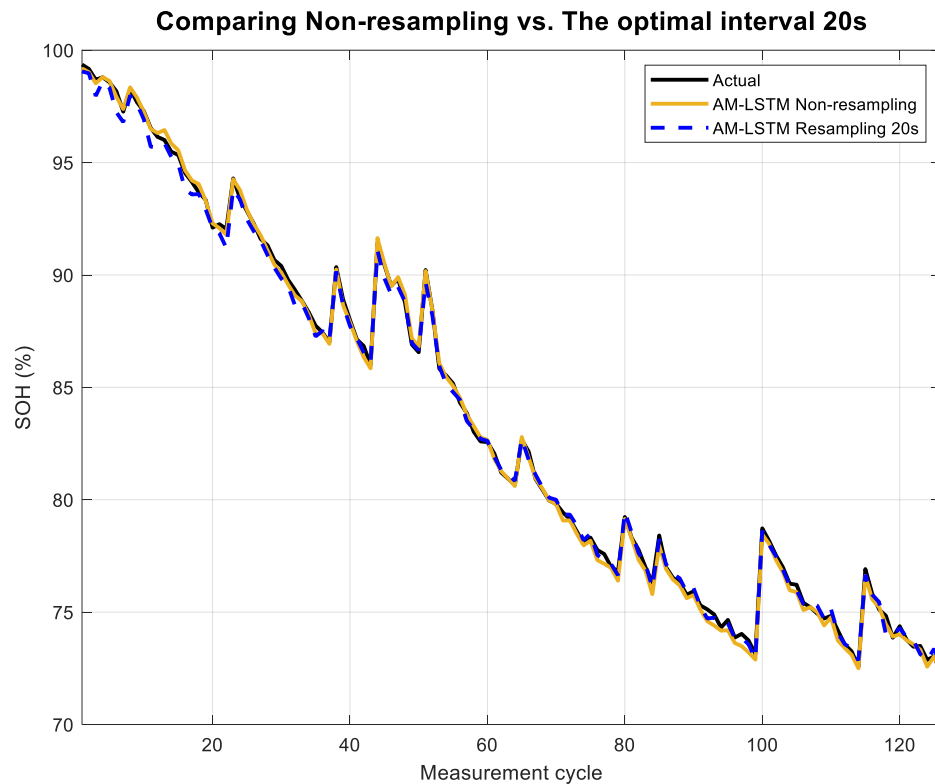


Figure 4.28 Analysis of the impact of resampling on the AM-LSTM model (NASA dataset) comparing Non-resampling vs. 20s interval

Regarding the GPR model, the same limitation persists: it could not process the large raw dataset (indicated by the hyphen symbol '-' in the table). However, it became operational once the data size was reduced. While its accuracy is not as high as the deep learning models, it remains useful for providing confidence intervals.

The deciding factor for using resampling with the NASA dataset is computational efficiency. Data from Table 4.6 clearly shows that training the AM-LSTM model with raw data took approximately 2,759 seconds. However, using resampling at 20 seconds reduced this time to just 121.8 seconds, a decrease of over 95%. Therefore, although resampling involves a slight trade-off in accuracy compared to raw data, the massive gain in processing speed makes it a worthwhile option for real-world applications with limited computational resources.

Comparing the SOH prediction results between the two datasets reveals a distinct relationship between the physical characteristics of the data and the optimal data processing strategy. For the Oxford, the degradation pattern exhibits a relatively smooth, monotonic decay but is superimposed with high-frequency noise throughout its lifecycle. This characteristic explains why the 100-second resampling interval yielded the best results; in this context, resampling acts effectively as a denoising filter. By averaging out the fluctuations, the model can more clearly discern the underlying global degradation trend, leading to improved accuracy.

In contrast, the NASA dataset displays a fundamentally different behavior. The SOH curve does not follow a simple linear decline but is marked by significant Capacity Regeneration Phenomena or capacity recovery. These are observed as sudden upward spikes in SOH during certain cycles, typically caused by the battery's relaxation effect during rest periods, allowing active ions to redistribute.

This unique characteristic of the NASA dataset is the key reason why using raw data (None) or a very short resampling interval (20s) resulted in the highest accuracy. These upward "bounces" are not noise, but significant features reflecting the battery's actual health status. If a wider resampling interval were applied here, the data reduction process would smooth out these critical peaks, causing the model to lose vital information required to predict the regeneration phases accurately.

However, the ultimate validation of the proposed AM-LSTM model lies in its adaptability. The experimental results demonstrate that the model is robust enough to handle these opposing data extremes. It successfully generated a smooth, trend-following curve for the noisy Oxford data, while simultaneously being responsive enough to track the sharp fluctuations and regeneration peaks of the NASA data without significant delay. This confirms that the proposed architecture coupled with the attention mechanism is capable of capturing complex degradation patterns regardless of the specific data characteristics. Fig. 4.29 presents the RMSE comparison of all models for the NASA dataset.

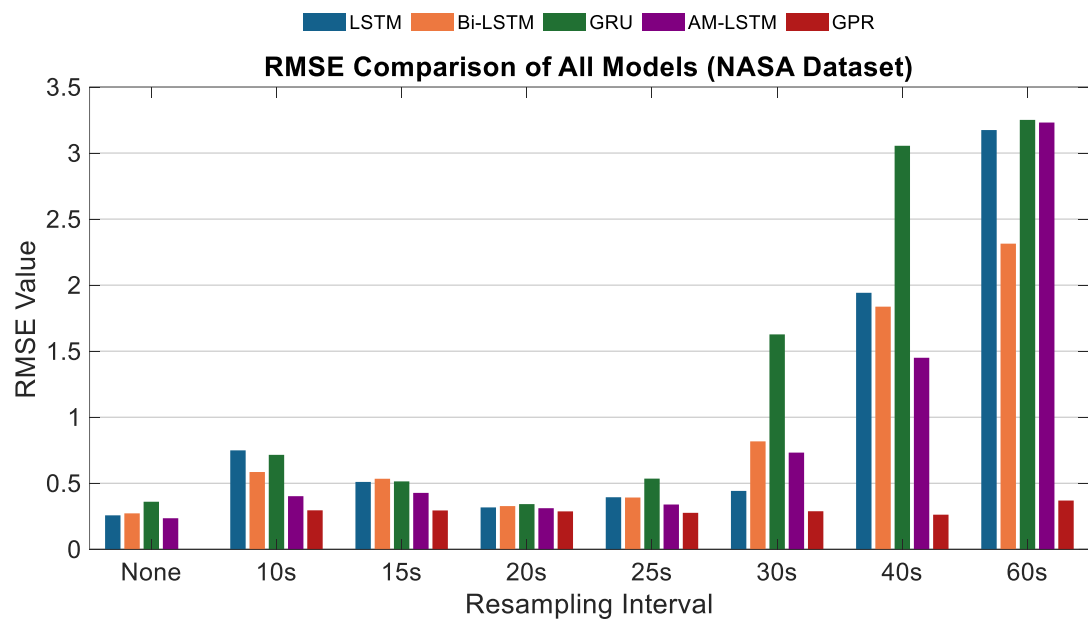


Figure 4.29 RMSE comparison of all models on the NASA dataset

Table 4.5 Comparison of SOH Estimation Results on NASA Dataset

Dataset	Resampling	LSTM				Bi-LSTM				GRU				AM-LSTM				GPR			
	Interval	RMSE	MSE	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE	MAE	MAPE
NASA	None	0.257	0.066	0.201	0.242	0.272	0.074	0.228	0.271	0.360	0.130	0.317	0.377	0.235	0.055	0.187	0.229	-	-	-	-
	10 s	0.749	0.562	0.636	0.737	0.585	0.342	0.498	0.617	0.715	0.512	0.648	0.795	0.402	0.161	0.336	0.407	0.295	0.087	0.240	0.298
	15 s	0.510	0.260	0.437	0.525	0.534	0.286	0.420	0.485	0.514	0.264	0.415	0.485	0.427	0.182	0.344	0.413	0.294	0.086	0.244	0.301
	20 s	0.317	0.100	0.252	0.306	0.327	0.107	0.273	0.325	0.342	0.117	0.276	0.336	0.311	0.097	0.247	0.290	0.287	0.082	0.237	0.293
	25 s	0.394	0.156	0.313	0.380	0.392	0.153	0.310	0.369	0.535	0.286	0.451	0.554	0.339	0.115	0.269	0.325	0.276	0.076	0.223	0.278
	30 s	0.442	0.195	0.350	0.424	0.817	0.667	0.643	0.807	1.627	2.645	1.568	1.905	0.732	0.535	0.582	0.725	0.288	0.083	0.238	0.294
	40 s	1.942	3.769	1.890	2.277	1.837	3.373	1.748	2.087	3.055	9.329	2.752	3.344	1.450	2.102	1.344	1.601	0.262	0.069	0.212	0.262
	60 s	3.174	10.073	2.762	3.358	2.314	5.356	2.259	2.719	3.251	10.570	2.820	3.437	3.231	10.439	2.813	3.419	0.369	0.137	0.236	0.295

Table 4.6 Execution Time of Each Model for SOH Estimation on NASA Dataset

Dataset	Resampling	LSTM	Bi-LSTM	GRU	AM-LSTM	GPR
	Interval	(s)	(s)	(s)	(s)	(s)
NASA	None	1,749.56	2,175.9	1,602.4	2,759.3	-
	10 s	134.4283	231.0435	134.9282	254.1212	1,249.0148
	15 s	108.6076	122.1542	91.2231	179.9841	224.0689
	20 s	68.7088	102.988	66.7452	119.223	98.7963
	25 s	62.6125	75.981	57.9775	107.0138	47.8137
	30 s	45.5109	58.8247	44.2024	83.9645	33.577
	40 s	34.8527	43.4884	33.3974	62.2606	17.8264
	60 s	28.7726	34.8007	26.8101	44.6983	11.0508

3) Summary

This chapter presented an analysis of the proposed methodology for SOH estimation, focusing on the interaction between dataset characteristics and processing strategies. The experimental results highlighted fundamental differences between the Oxford and NASA datasets: Oxford is characterized by high-density data with significant noise, while NASA exhibits complex electrochemical behaviors, specifically the capacity regeneration phenomena.

Empirical evidence confirmed that the AM-LSTM model achieved superior performance compared to baseline models. For the Oxford dataset, the model attained its best accuracy with an RMSE of 0.2920 when paired with a resampling interval of 100 seconds, which effectively filtered out high-frequency noise. Conversely, for the NASA dataset, the model achieved its optimal performance with an RMSE of

0.2351 by utilizing raw data or short-interval resampling, thereby preserving the fine-grained details required to track capacity recovery.

Beyond accuracy, the most significant advantage of applying the resampling technique proved to be computational efficiency. Specifically, for the Oxford dataset, the results demonstrated a drastic reduction in model training time by over 98%, accompanied by a significant 58.58% reduction in estimation error compared to the raw data baseline. In the case of the NASA dataset, although resampling led to a 32.34% increase in RMSE, this is considered an acceptable trade-off given the massive reduction in training time of over 95%. These findings confirm that the proposed approach successfully strikes a balance between high accuracy and computational feasibility, making it highly suitable for implementation in real-world battery management systems with limited resources.