Proceedings of the ASME 2025 44th International Conference on Ocean, Offshore and Arctic Engineering
OMAE2025
June 22-27, 2025, Vancouver, British Columbia, Canada

OMAE2025-156104

LITHIUM-ION BATTERY DEGRADATION FORECASTING USING DATA-DRIVEN TIME SERIES MODELS

Kishan Patel[1,*], Athira Pulickakudy Salin[2], Merten Stender[3], Moritz Braun[1], Sören Ehlers[1]

[1]German Aerospace Center (DLR) – Institute of Maritime Energy Systems, Geesthacht, Germany
[2]Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[3]Technische Universität Berlin, Berlin, Germany

ABSTRACT

The maritime industry faces significant challenges as it adapts from a major carbon emitter to a low-emission sector, intending to eventually achieve zero emissions. This transition requires innovative solutions for both new and old vessels, lithium-ion batteries show promise in achieving these goals. Battery management systems improve reliability and safety by monitoring voltage, current, and temperature through sensors. These parameters enable the prediction of remaining usable life, allowing for prompt maintenance and replacement before failure occurs. Publicly accessible lithium battery datasets provide a useful starting point for predictive degradation model development. This study investigates time series modeling methodologies for lithium-ion battery degradation, utilizing NASA's battery degradation dataset. Three models viz. Autoregressive, Autoregressive Integrated Moving Average, and its extension using seasonality parameters were developed. They were tested with four train/test ratios to predict the remaining useful life values and assess the accuracy of the predicted degradation curve against experimental results. From the results, it was observed that the Autoregressive Integrated Moving Average model had the least combined average Root Mean Square Error values, resulting in a good overall degradation curve fitting, whereas the Seasonal Autoregressive Integrated Moving Average model was able to predict the End of Life values more accurately.

Keywords: Lithium-ion batteries, Data Driven Time Series Model, AR, ARIMA, SARIMA, Remaining Useful Life

NOMENCLATURE

LIB	Lithium-ion Battery
EOL	End of Life
RUL	Remaining Useful Life
BMS	Battery Management System
SOC	State of Charge
PF	Particle Filter
NN	Neural Networks
ARMA	Autoregressive Moving Average
SOH	State of Health
AR	Autoregressive
ARIMA	Autoregressive Integrated Moving Average
SARIMA	Seasonal Autoregressive Integrated Moving Average
RMSE	Root Mean Square Error

1. INTRODUCTION AND OBJECTIVES

This section provides a background on the necessity of implementing new technologies to reduce carbon emissions, followed by the objectives of this research.

1.1 Background

Maritime transportation has long been a key driver of economic development in economies, serving as a major agent of international trade. More specifically, marine transportation accounts for 80-90% of global trade, which is predicted to go only upwards and play a significant role in the coming years for the trading business worldwide [1].

*Corresponding author
*Email Id: kishan.patel@dlr.de

With the growth of worldwide trade, international shipping has become one of the major fuel consumers in the transportation industry. According to the International Maritime Organization (IMO) which is the industry's rule maker, ships utilize around 300 million tons of fossil fuels per year [2]. The high consumption of fossil fuels results in a high level of emissions. Since fuel consumption is closely tied to the emission of greenhouse gases and pollutants, it is anticipated that the shipping industry's carbon dioxide (CO_2) emissions will account for 12-18% of worldwide anthropogenic CO_2 emissions by 2050 if no changes to the current technologies and existing infrastructures are made [2]. The usage of fossil fuels is mostly responsible for hazardous emissions from ships. Heavy fuel oil (HFO) accounts for approximately 72% of fossil fuel usage, followed by marine diesel oil (MDO) at 26% and liquefied natural gas (LNG) at 2% [3]. Despite the fact that LNG outperforms HFO and MDO in terms of energy efficiency and greenhouse gas emissions, problems such as methane leaks and difficulty with onboard storage have undermined its market position. Mitigating dangerous exhaust gas emissions from shipping remains a concern for the business.

As a result, more efficient and alternative solutions to internal combustion engines are required for the maritime industry's future. At this stage, the maritime industry has three potential options to remove or minimize airborne emissions: adopting new emissions abatement technology, marine alternative fuels, and finally, electrification of propulsion with the probable usage of hybrid power systems (HPS) [4]. Based on DNV's preliminary studies, vessel routing, increased electrification using energy storage (ESS) technologies, renewable integration, and the use of alternative fuels are identified as potential fuel and energy-efficient technologies that could significantly reduce the overall emissions of an all-electric ship (AES). The use of hybrid-electric power for ship propulsion and other electrical loads can reduce harmful emissions by lowering fuel consumption, among other benefits. In this sense, power generating and storage choices must be thoroughly evaluated to identify more effective solutions.

Lithium-ion batteries (LIBs) are one of the most widely utilized power sources in new energy systems due to their high energy density, power density, and low self-discharge rate [5]. The increasing market expansion for LIBs is driving down prices, but making LIBs endure longer is also vital. This improves lifetime economics, allows for longer warranties, and reduces the environmental consequences associated with raw material extraction and manufacturing [6]. The battery aging mechanism and degradation model are equally critical at the battery system level. The impact of electrical, mechanical, and thermal factors on battery health must be examined for the aging mechanism using predictive degradation models [7].

Battery degradation is an important topic to investigate since it has a considerable impact on battery performance, lifespan, and safety. Understanding battery degradation is crucial for cost-effective de-carbonization of both transportation and energy systems as well as critical in increasing the operational life of batteries [8]. The ability to precisely anticipate battery end-of-life (EOL) or remaining useful life (RUL) allows the dangers of battery runaway to be reduced [9]. Based on the aging mechanism and the corresponding battery degradation model, the Battery Management System (BMS) could anticipate the harm caused by various operating and environmental situations. The prediction of future performance refers to RUL. Given the battery's nonlinear fading characteristics, the usual extrapolation method cannot reliably anticipate its remaining life. It is required to produce a reliable prediction based on the primary aging process and associated battery life model under varied operating situations and performance fading stages. The life of the same batteries would be drastically different if they were operated under different conditions. The primary parameters influencing battery life are as follows [10]:

- High temperature, which accelerates internal side reactions;
- Low temperature, causing metal ions to reduce quickly and lithium deposition making the crystal structure of the active material easily destroyed;
- High or low State of Charge (SOC); and
- Overload or over-discharge.

Moreover, investigating degradation helps develop strategies to reduce environmental concerns related to battery disposal and recycling.

Recent research on LIBs RUL prediction highlights the importance of hybrid prognostic methods in achieving highly accurate predictions. Single-method approaches often struggle to capture the complex non-linear degradation patterns of LIBs, prompting the development of more sophisticated hybrid models. Various predictive techniques, including Particle Filter (PF), Gaussian Process Regression (GPR), and Neural Networks (NN), have been extensively studied and implemented. Saha et al. [11] explored the use of traditional PF algorithms for RUL prediction, framing the battery prognostic problem within a Bayesian learning context. Their findings suggested that PF models excel in scenarios where expert knowledge provides high fidelity for parts of the state space, while uncertainties dominate elsewhere. However, PF models often fail to fully capture the intricate trends in battery degradation data. To address this limitation, Hu et al. [12] integrated a kernel smoothing method with the PF model for RUL prediction of degrading components. Their approach yielded more accurate estimates of component RUL than conventional particle filtering techniques. Similarly, Miao et al. [13] demonstrated that prediction accuracy improves with the availability of more measurement data. Their comparative analysis of unscented PF and standard PF models revealed that the unscented PF model delivered a 2% improvement in accuracy. Kozlowski [14] introduced a data-driven RUL prediction model that combined Autoregressive (AR) Moving Average (MA) models, NNs, and fuzzy logic. Likewise, Liu et al. [15] proposed an integrated approach that merged PF with a non-linear degradation AR model for State of Health (SOH) estimations. These hybrid methodologies represent a significant leap forward in LIB RUL prediction, delivering improved accuracy and robustness compared to single-method approaches. As research continues, the

integration of complementary techniques remains a key focus in developing reliable and precise predictive models for LIB health management.

Given the time-varying nature of battery degradation data, time series analysis and forecasting methods can serve as effective alternatives to standard PF models for predicting battery health and RUL [16]. Time series models are particularly suited for analyzing cyclic data by leveraging historical values and error terms to forecast future states. Incorporating external factors, such as temperature, voltage, and current, can further enhance predictive accuracy, addressing limitations inherent in standard PF models [17]. This approach provides a comprehensive understanding of the degradation trajectory and RUL by analyzing the effects of cyclic charging and discharging processes. An accurate time series model thus offers valuable insights into the factors driving battery degradation and improves prognostic outcomes. In maritime contexts, some studies have applied data-driven methods to monitor battery health. For example, Liang et al. [18] presented a case study on 'Battery AI', which is a monitoring platform for the battery data from ships experiencing actual operating conditions. They also compared the prediction results of this platform with the annual SOH test results of four ships and found the largest deviation to be around 3.21% compared to actual experimental tests. Their study shows great potential of Artificial Intelligence (AI) in estimating battery health. Another research by Vanem et al. [19] followed a data-driven approach, they presented results using snap-shot method for modeling and monitoring the SOH of the battery systems of ships. They explored statistical models for predicting battery capacity values and concluded that models based on a limited set of features derived from sensor data could produce reliable and comparable results. However, they acknowledged that achieving accurate forecasts remains a significant challenge, especially when considering datasets from all cells in the battery, as this has a direct impact on the size and complexity of the training dataset.

Few studies have utilized LIB aging datasets from open sources, NASA aging datasets being one of the most frequently used in development of LIB degradation time series models. Xu et al. [20] utilized the NASA dataset for estimating SOH values using the Regression Tree model and Random Forest model and claimed that both models perform extremely well with the mean square error of 0.0006 and 0.0002 respectively. Nagulapati et al. [21] performed a study on the influence of training dataset size and how it affects the model outcomes. They implemented GPR and Artificial NN to the NASA dataset and compared the RMSE values to other Machine Learning (ML) algorithms. Wang et al. [22] and Liu et al. [23] used Autoregressive Integrated Moving Average (ARIMA) models in combination with Long Short-Term Memory (LSTM) and Empirical Mode Decomposition (EMD) models respectively to predict SOH and RUL LIBs. The common objectives of these studies were to use either standalone ML algorithms or time series models (AR or ARIMA) combined with PF or ML models for the NASA dataset. However, studies focusing on solely combining different models within the same time series family are limited. To the best knowledge of the authors, no research based on an extension of the ARIMA model to the Seasonal–ARIMA (SARIMA) model for the NASA dataset has been explored yet. This research, therefore, aims to bridge this gap by developing models from the same time series family, making this research unique.

1.2 Objectives

Understanding and predicting LIB degradation, especially in critical applications is important for improving the reliability and efficiency of the system. Time series predictive models, such as ARIMA and SARIMA, are well-suited for capturing trends, seasonality, and patterns in capacity degradation data. However, before implementing complex models, starting with a simpler and less complex model such as the AR model is a sensible approach. AR model serves as a foundational block for ARIMA and SARIMA models, as it captures simple linear dependencies and helps by establishing a baseline understanding of the data. This initial model will help to assess if additional components like trend, seasonality, etc. from the NASA dataset need to be considered or not.

The primary objective of this research is to develop data-driven models, specifically three time series models – AR, ARIMA, and SARIMA to predict battery degradation behavior. The initial step involves a thorough exploration and pre-processing of the raw dataset, followed by the development and implementation of these models. Three datasets (B0005, B0006, and B0007) from the NASA library will be used to develop these models, and four train/test split ratios will be examined to understand model behavior and its impact on prediction accuracy. Using these ratios, the goal is to predict capacity, RUL, and EOL values for each scenario and plot the predicted capacity values against the experimental degradation curve from NASA's dataset. Since the SARIMA model is an extension of the ARIMA model, establishing a good foundation with the ARIMA model is important. To achieve this, the ARIMA model developed in this research will be validated against findings from existing literature.

Finally, the three models will be compared and evaluated based on mean predicted EOL, standard deviation, and RMSE values. The mean predicted EOL, along with standard deviation, will provide insights into model performance in terms of predicting the RUL of the batteries. Meanwhile, RMSE values will help assess the accuracy of curve fitting for each model.

2. DATASET USED IN THIS RESEARCH

Obtaining real-world degradation data for LIBs in marine applications is difficult due to proprietary constraints, new topics of research, complex and versatile circumstances onboard ships, the absence of particular sensors in old vessels, and harsh offshore environmental conditions. This intricacy makes it challenging to get consistent and complete sensor data required for reliable modeling. To overcome this, many academics and research institutes have turned to open-source datasets, which provide reliable, easily available, and standardized data for creating prediction models. For example, NASA's (National Aeronautics and Space Administration) Prognostics Center of

Excellence supplies LIB aging datasets [24], which are frequently utilized in the development of prognostic algorithms.

By leveraging open-source data like NASA's, researchers can develop robust models to predict and mitigate battery degradation, enhancing the reliability and efficiency of LIBs in maritime applications. Even though NASA's battery data is not related to maritime applications, it still sets up a strong foundation for understanding battery degradation behavior. It would provide a solid base and a benchmark for developing and testing predictive models, allowing us to understand and identify trends and patterns of capacity degradation, cycle life, and other important processes. Once these models are validated, they can be adapted and tuned to different sensor data and parameters helping to address the battery load profiles and operating conditions of ships.

2.1 NASA Battery Degradation Data

This research uses the NASA LIB degradation dataset [24]. These batteries underwent testing across three distinct operational profiles: charge, discharge, and impedance at ambient temperature [11, 25]. The dataset contains charge and discharge cycles of 18,650 LIB cells, with varying operational profiles.

This research is focused on batteries B0005, B0006, and B0007, which had 168-170 complete charge-discharge cycles and covered a sufficient discharge capacity range for RUL estimation. These considered batteries cover the discharge capacity range from a nominal capacity of 2Ah to at least 1.4Ah which is 70% of the nominal capacity and also the failure threshold value, which is essential for the RUL estimation [26].

Other reasons behind using B0005, B0006, and B0007 are due to complete sensor data, no missing values and uninterrupted charge and discharge cycles, providing a good dataset to be used in predictive models.

2.2 Charge Profile

The charging profile starting from the initial up to the final cycles for battery B0005 was analyzed by visualizing the voltage, current, and temperature trends over time. After comparing the charging profiles of new and aged batteries, it showed that the battery behavior changes as it ages due to degradation, with similar results for batteries B0006 and B0007 depicted in Figure 1. In this dataset, the battery charging cycle followed the constant current-constant voltage (CC-CV) principle, where the current is initially held constant, and then the voltage is kept constant.

Assessing the voltage during the charging cycle, it was observed that the current remained constant at the start of the initial and final cycles after which the voltage increased to 4.2 V. Due to internal resistance and electrochemical reactions, the temperature rises slightly during charging, especially in the constant current phase. The temperature rise is more distinct in the final cycles due to increased internal resistance from degradation mechanisms like Solid Electrolyte Interphase (SEI) layer growth and active material loss. The last cycle displayed reduced capacity compared to the new battery with a slower voltage increase and extended constant voltage phase. These changes signify the impact of degradation mechanisms over the battery's lifetime.

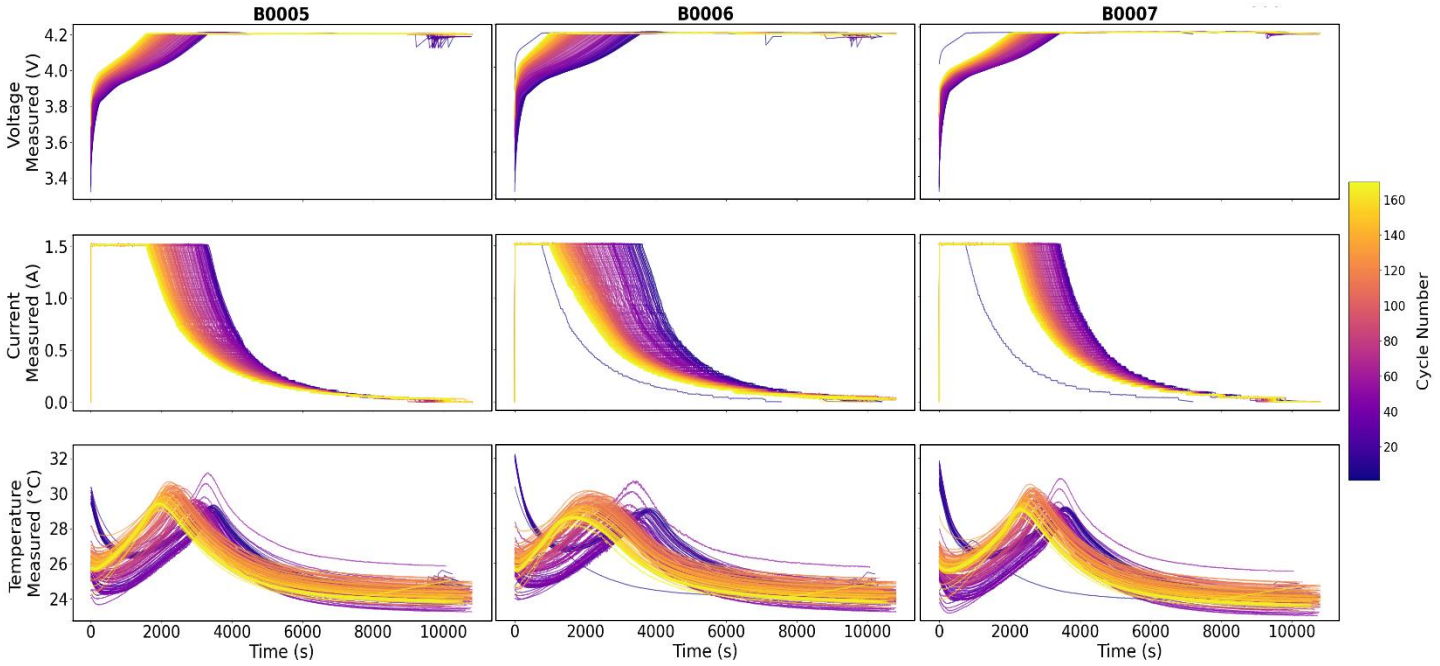

FIGURE 1: VOLTAGE, CURRENT AND TEMPERATURE PLOTS FOR B0005-0007 FOR ALL CYCLES

As the number of cycles increases, the duration of the charge cycle gets prolonged due to the increased internal resistance and chemical reactions inside the battery.

2.3 Discharge Profile

The batteries discharge at a constant current of 2.0A until the cell voltage reaches 2.7V, 2.2V, and 2.2V for B0005, B0006, and B0007, respectively. The test experiments continue until the battery's rated capacity decreases by 30%, i.e., 1.4 Ah [26]. Figure 2 shows that the discharge capacity degrades as the number of cycles increases. It follows a non-linear trend due to the degradation caused by charging, discharging, and rest time between charge-discharge cycles. The failure threshold marked (red color) in Figure 2 is 1.4Ah (70% of the initial capacity), indicating that the battery is considered failed when the capacity value drops below this level.

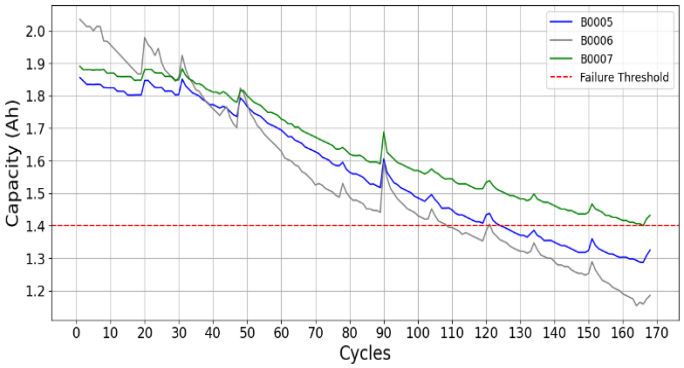

FIGURE 2: CAPACITY DEGRADATION OF BATTERIES WITH FAILURE THRESHOLD LINE

2.4 Data Preprocessing

The following three steps were used in this study to preprocess the raw NASA data:

a) Duplicate entries: Duplicate entries add no value to the analysis and can slow down processing. Therefore, removing them from the dataset is vital. The *'drop duplicates'* function from the 'Pandas' library was implemented to check for duplicate entries and remove them.

b) Missing values: The approach for dealing with missing data in columns and rows of sensor data varies based on the problem statement, as well as the type and size of the dataset. Some ML algorithms perform even in the absence of missing values, but in cases where the available dataset is relatively small, they should be managed by either entering the missing data or eliminating the incomplete entries that correspond to the missing value.

After performing the missing value check, it was observed that all 34 cells were complete, and therefore no further action was required to deal with missing values. Even though it is already claimed that this particular dataset (B0005-0007) is complete and has no missing values, there is a possibility of losing data in the initial data import and preparation phase. Hence missing values were checked from the raw data to make sure there is no data loss after importing into the main data frame.

c) Outliers detection: The Interquartile Range (IQR) Proximity Rule and the Z-Score method were implemented to detect any outliers if any. According to the IQR method, any values that were more than 1.5 times the interquartile range from the first or third quartile were considered outliers. Eliminating or changing them may increase accuracy in some situations but may cause the model to become skewed in others. In this study, no outliers were found.

As seen from the data-preprocessing step, this particular dataset proves to be a good starting point due to fewer duplicate cells, and no missing values and outliers, making the training and developing algorithm part more efficient.

3. METHODOLOGY

3.1 Time Series Analysis of Battery Data

Time series analysis is essential to understand the behavior of battery degradation and, hence to develop appropriate predictive models. This section carries out a time series analysis of the battery degradation data for B0005, B0006, and B0007 by decomposing its key components, followed by a stationarity check and the seasonal decomposition charts. The trend in the battery degradation data is non-stationary since it does not have a constant mean level over time. The paper by Yang et al. [17] introduces various time series forecasting models suitable for this kind of data and develops an integrated approach. For the seasonal decomposition of the battery capacity data, the trend, seasonality, and residual components are separated. Time series data usually consist of four components: level, trend, seasonality, and residuals [16]. The time series data contains all levels that provide the average value, enhanced by a noise component representing irregular or random fluctuations. While seasonality and trend are optional components, they provide a very detailed look into the intrinsic character of the data series. These critical components help build up an abstract picture of both the data and the problem, especially in the process of time series analysis and forecasting. The trend component represents the long-term variability of data and reflects the general direction taken by the data over time. In the case of battery aging, it would be a degrading or declining curve. In contrast, seasonality refers to a regular, repetitive pattern in data, and cycles refer to patterns repeating over a certain period.

3.1.1 Stationarity Check

The Augmented Dickey-Fuller (ADF) test was performed on data from three battery cells (0005, 0006, and 0007) to assess the stationarity of the degrading battery dataset.

The ADF statistics with corresponding p-values are shown in Table 1 below.

TABLE 1: ADF AND *P* VALUES FOR BATTERIES

Battery	ADF stats	*p* value
B0005	-0.5257	0.8869
B0006	-1.3704	0.5964
B0007	-0.6566	0.8577

As the *p*-values are greater than the significance level (0.05) and the ADF statistics do not exceed critical thresholds, the results suggest insufficient evidence to reject the null hypothesis of a unit room indicating that the data for all three cells may be non-stationary [27].

To address non-stationarity, first-order differencing was applied to the data. This transformation removed the trends and stabilized the mean making the data more useful for further time-series modeling and extracting relevant information more accurately. Stationarity was achieved after first-order differencing, as confirmed by a subsequent ADF test, hence the differencing parameter $d = 1$ is selected for the models which is explained later. If the data remains non-stationary, additional differencing steps may be necessary, incrementing d accordingly. However, in our case, performing first-order differencing gave us satisfying results and p-values less than 0.05 and hence there was no need to perform second-order differencing to the time series data.

3.1.2 Seasonal Trend Decomposition Plots

Seasonal-trend decomposition was performed on the capacity degradation data of the B0005 battery cell.

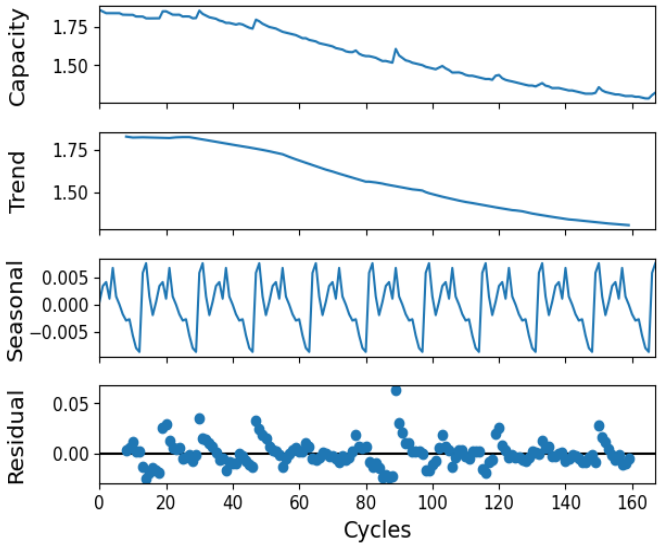

FIGURE 3: SEASONAL-TREND DECOMPOSITION PLOT OF B0005 FOR RAW DATASET

In Figure 3, the Seasonal-Trend Decomposition plot was plotted using the *'statsmodel'* library from Python which uses the Locally Estimated Scatter-plot Smoothing (LOESS) method for the original data. Based on this, the original capacity data was transformed into three major components: trend, seasonality, and residual.

The top panel in Figure 3 depicts the original capacity time series, which follows a downward trend, a sign of progressive battery degradation with time. The second panel in the same figure plots the original trend component, which is a smooth curve representing the long-term decline in the capacity of the battery. This is justified due to aging and degradation mechanisms inside the battery over time. The third panel highlights periodic fluctuations, or the seasonal component, arising from operating conditions or measurement cycles, for example, variations in temperature during charge and discharge cycles. The bottom panel plots the residual component, which is essentially the noise or randomness in the data.

After performing the step, it can be observed through Figure 4, that the trend component is absent compared to the original series of data.

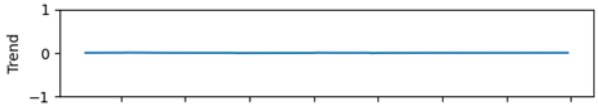

FIGURE 4: TREND COMPONENT AFTER PERFORMING SEASONAL-TREND DECOMPOSITION

Similarly, there was no pattern in the residual component and it was centered around zero. Hence, the decomposition step was successful in filtering out the systematic components of the time series. Similarly, seasonal decomposition has been performed on the remaining two battery datasets (B0006 and B0007).

3.2 Time Series Prediction Models

In this section, we look into the detailed model implementation algorithm of time series models. Figure 5 is inspired by the work of Moseley et al. [28] Shows the algorithm for battery capacity modeling using the AR and SARIMA models. The main steps include sampling, model training, prediction using a test set, and evaluation.

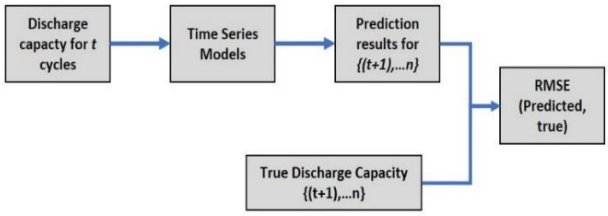

FIGURE 5: BATTERY CAPACITY ALGORITHM USING TIME SERIES MODELS

In this context, 'n' represents the total number of cycles for the battery, and 't' denotes the number of training cycles. So, the first step is sampling, which splits the data into training and test sets. The AR and SARIMA models require discharge capacity and the corresponding cycles to train the model with appropriate training parameters.

3.2.1 Choosing Algorithms Based on Battery Data

The time series models in the ARMA family can generally be classified based on their ability to study the trend, seasonality, and influence of external parameters [29]. AR, ARIMA, and SARIMA models from the time series family were utilized to develop the battery capacity degradation. Such models can detect temporal correlations and patterns. An AR model works well with time-series data when the current state is linearly dependent on prior values. This makes AR an easy and useful first step, especially in understanding the fundamental autoregressive nature of degradation, because the capacity decrease in batteries frequently follows predictable dynamics over time.

Battery capacity statistics are frequently non-stationary and seasonal, as indicated by both the Augmented Dickey-Fuller test and Seasonal-Trend-Decomposition plots. These indicate a more effective model capable of handling the series' features, particularly non-stationarity, and seasonality, which are handled by differencing and seasonal components, respectively, in the SARIMA. SARIMA allows for very accurate capture of both long-term trends in battery capacity and short-term seasonal fluctuations. This adaptability and ability to integrate a wide range of seasonality elements, such as periodic cycle patterns and autoregressive/moving average (MA) terms, helps the model suit the observed data attributes.

Zhou et al. [30] implemented time series models like ARIMA and Empirical Mode Decomposition ARIMA to predict the SOH values for the NASA battery dataset. Their work presented a comparison of actual and predicted RUL values showing comparable results. These models were chosen for their good interpretability and computational efficiency, which have proven to be quite productive in time-series forecasting applications.

In this scenario, AR and SARIMA provide a more streamlined approach by directly addressing both non-stationarity and seasonality. They also offer a structured data-driven strategy for modeling and predicting the dynamics of battery degradation, producing informative results relevant to health monitoring, RUL estimates, and battery management system enhancements.

3.2.2 AR Model

This AR model was developed initially, it forecasts the current value by combining the historical values of a time series linearly.

The intrinsic temporal dependencies in capacity loss, where the current state of change is heavily reliant on previous states, are captured by AR models quite well in the case of battery degradation. This allows for preliminary forecasts and creates a baseline understanding of the degradation dynamics.

The formula for an AR model of order *p* is given as:

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \epsilon_t \qquad (1)$$

Where, X_t is the value of time series at time *t*, ϕ_i are the parameter coefficient of the model, $X_{\{t-i\}}$ are the lagged values of the series, ϵ_t is the error term and lastly, p is the order of the AR model.

Before running the models, it is important to find the order of the simulation based on the data. To determine the order of the necessary model, at first, we use the Partial Autocorrelation Function (PACF) plot from the data of B0005 followed by the Information Criterion Based Analysis using Akaike & Bayesian Information Criterions (AIC & BIC values):

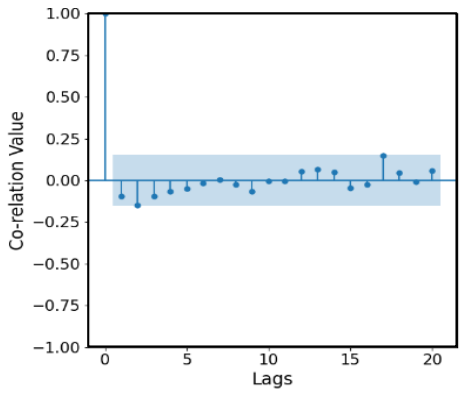

FIGURE 6: PACF PLOT FOR B0005 AR MODEL

The PACF plot in Figure 6, shows a significant spike at lag 0 (self-correlation), followed by the remaining lags falling in the shaded (blue) confidence interval, except for lag 2 near the start has a correlation value very close to the value of confidence interval zone. Since the confidence region only covers 95% statistically, there is a possibility that AR(2) would be the best choice, but also leaves some uncertainties between opting for AR(1) and AR(3) models. To mitigate this problem, a more objective approach, Akaike and Bayesian Information Criterion values also referred to as AIC & BIC values were calculated for AR(0,1,2 and 3) and presented in Table 2 for B0005.

TABLE 2: AIC AND BIC VALUES FOR B0005

Order	AIC	BIC
0	-206.710	-202.526
1	-336.721	-330.489
2	-328.179	-319.937
3	-322.083	-311.868

Based on the general rule of model selection using information criteria also known as the minimum AIC/BIC approach, the order with the lowest AIC and BIC values gives the most accurate results for the data and hence AR(1) model was chosen for this research. An AR(1) model, indicates that the current value of the series is mainly determined by the immediately preceding value. Similarly, based on AIC and BIC values for B0006 and B0007, AR(1) was determined to be the best-fit model.

For the first model for B0005, 35% of the available dataset was used, meaning the training data length was 60 cycles and the testing data length of 108 cycles. The standard safety threshold

for the battery is when its rated capacity drops below 70% of the nominal capacity ($C_{nominal}$), in this research 1.4Ah is considered as the failure threshold value [31].

It can be noted from Figure 8, that the prediction value based on 35% training length, does not provide good prediction results and it might indicate that the training data length provided was not sufficient for the AR model and it must be increased to have satisfactory and comparable results.

Hence, the training length of 40% which is 68 cycles was tested next leaving 100 cycles as testing data length. Figure 8 shows the graph of the AR model with a 40% training set and it can be observed that even though the prediction curve doesn't follow the original values, it still gives a reasonable result at the failure threshold line.

Similarly, two other scenarios were simulated for B0005, increasing the training data set to 45% and 50%. Four models were implemented for each battery data based on different training data lengths, resulting in a total of 12 cases for the AR prediction model. Subsequent results for all 12 cases are explained and validated in the next chapter of 'Results and Interpretation'.

Even though AR models don't account for seasonal patterns or non-stationarity, their interpretability and computational efficiency make it a perfect starting point for more intricate models like ARIMA and SARIMA.

3.2.3 ARIMA and SARIMA Model

The ARIMA model is defined by three components viz. AR, Integrated (I), and the MA. It is represented by ARIMA(p,d,q), where p indicates the order of the AR models, d refers to the differencing of the series and q represents the MA terms. The AR part is responsible for capturing the relationship between the current value with the past value. The Integrated part is responsible for differencing the series to make it stationary and finally, the MA component models the dependencies between current value and the forecast errors. ARIMA equation is given by:

$$\Delta^d x_t = c + \Sigma(\phi_i \cdot x_{t-i}) + \Sigma\left(\theta_j \cdot \epsilon_{t-j}\right) + \epsilon_t \qquad (2)$$

Where, $\Delta^d x_t$ is the differenced series of x_t (the integrated part) applied d times to make the series stationary, c is the intercept term, $\Sigma(\phi_i \cdot x_{t-i})$ represent the summation of past values (p) weighted by the coefficients ϕ_i which is the AR part and $\Sigma\left(\theta_j \cdot \epsilon_{t-j}\right)$ is the summation of q past error terms, weighted by the coefficients θ_j (the MA part).

As the ARIMA model excludes external variables, it is difficult to assess the impact of external characteristics and their link to dependent features, such as seasonality in capacity degradation.

The SARIMA model accounts for both non-seasonal and seasonal components hence making it more accurate for time series data that exhibits seasonal patterns. The general form of the SARIMA equation incorporates both non-seasonal and seasonal AR, differencing, and MA terms, along with a seasonal error term. It is commonly used in forecasting applications where data shows trends and seasonal patterns. This model contains seasonal factors in its formulation and is denoted as SARIMA(p,d,q)(P,D,Q)s [29]. SARIMA parameters include non-seasonal parameters such as *p, d,* and *q,* whereas the seasonal parameters *P, Q, and D.* Since SARIMA is an extension of the ARIMA model, two new terms are added to Equation (2). The first additional term is $\Sigma(\Phi_m \cdot x_{t-m})$which represents the Seasonal Autoregressive (SAR) term with *m* as the seasonal lag part. The second term is $\Sigma(\Theta_m \cdot \epsilon_{t-m})$ which represents the Seasonal Moving Average (SMA) terms with ϵ_{t-m} being the seasonal error terms at lag *m.* Therefore, equation for SARIMA model can be written as:

$$\Delta^d x_t = c + \Sigma(\phi_i \cdot x_{t-i}) + \Sigma\left(\theta_j \cdot \epsilon_{t-j}\right) + \Sigma(\Phi_m \cdot x_{t-m}) + \Sigma(\Theta_m \cdot \epsilon_{t-m}) + \epsilon_t \qquad (3)$$

Determining the p,d,q values, P,D and Q are quite important for both models. In this research, Correlation Function Analysis as well as Criterion-Based Analysis are implemented. Unlike the AR model, Autocorrelation Function (ACF) and PACF plots for ARIMA and SARIMA are built on the residual series after performing the differentiation to the original series. Figure 7 shows the ACF and PACF plots.

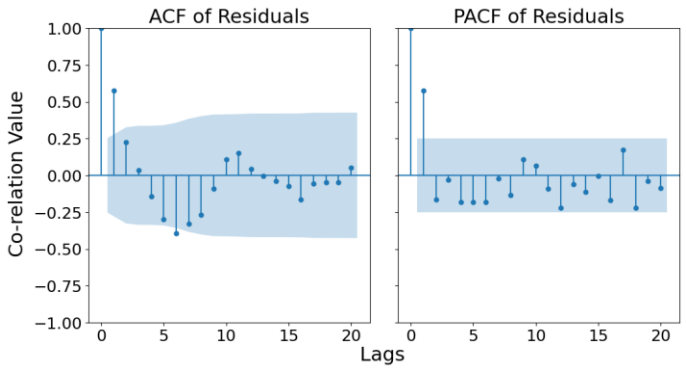

FIGURE 7: ACF AND PACF RESIDUAL PLOTS FOR B0005

Based on Figure 7, PACF shows a significant spike at lag 1, and multiple spikes close to the boundary of the confidence zone in blue, this suggests that opting for p = 1 might not be an ideal value for the order and relying just on PACF plots may not provide accurate results as seen from AR models, similarly in the case of ACF plot, spike 1 and 6 are out of the confidence zone but also spikes at lag 5 and 7 are close to the lower limits of the confidence zones, indicating that single or even double MA term might not be sufficient in producing optimum predictions. Based on the plots, and conservative approach, the ideal starting order for the SARIMA model would be (2,1,3) and thereafter reducing the p and q values until a desired curve fitting is obtained.

As observed from the AR models, an objective and a better approach would be selecting the order for the SARIMA model based on Criterion based information analysis. The lowest AIC and BIC values for B0005-0007 were calculated and are shown in Table 3.

TABLE 3: AIC AND BIC VALUES FOR B0005

p	*d*	*q*	AIC	BIC
1	1	1	-228.086	-219.397
2	1	1	-226.109	-215.683
3	1	3	-225.274	-209.635
1	1	2	-224.997	-214.571
1	1	3	-224.738	-212.574
2	1	2	-219.650	-207.487
3	1	1	-218.454	-206.291
2	1	3	-216.938	-203.037
3	1	2	-216.437	-202.535

Similarly, both values for B0006 and B0007 were also calculated. Based on the lowest values, the order of (1,1,1) for B0005, (1,1,3) for B0006, and (2,1,1) for B0007 was chosen. The use of first differencing was sufficient to make the series stationary for addressing the trends, hence *d=1* was chosen. Using other combinations of (p,d,q) values from Table 3 is also justified here since AIC and BIC values are in a very close range of each other. (2,1,3) combination from the ACF and PACF plots was found to be the 4th best possible combination of the order considering B0006 and B0007. The drawback of not choosing the lowest AIC and BIC values is that it affects the curve fitting parameters, especially the RMSE values in this research.

Similar to the AR model, a total of 12 simulations were performed for ARIMA and SARIMA each. Figure 8 below, below shows AR and SARIMA model predictions with training data lengths of 35% and 40% for B0005. By initial looks, compared to the same boundary conditions from the AR model, the SARIMA model gives a good prediction result for it.

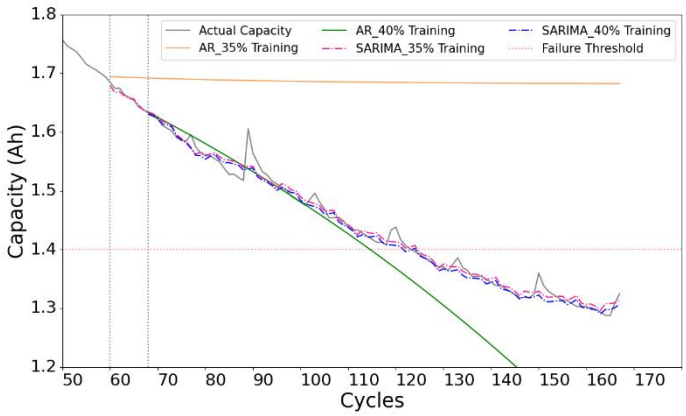

FIGURE 8: AR AND SARIMA MODEL PREDICTIONS WITH 35% AND 40% TRAINING DATA

Similarly, simulations for B0006 and B0007 data were run and the subsequent results for all 12 cases are explained and validated in the next chapter of 'Results and Interpretation'.

4. RESULTS AND INTERPRETATIONS

In this section, the results of the AR, ARIMA and SARIMA time series models in predicting battery capacity degradation are discussed in detail. Both models are evaluated based on different performance metrics such as the impact of data splitting of training/test ratios on prediction correctness, the difference between predicted and actual EOL cycles, the nature of actual and predicted curves, and how well the predicted RUL aligns with the actual battery performance. Finally, RMSE values obtained from this research are compared with different data splitting cases to better understand the trends.

The following tables in sub-chapters provide a detailed overview of the results, highlighting the strengths and limitations of the AR and SARIMA models.

4.1 AR Model Results

Table 4 and Figure 9 give us the comparison of values of actual data from NASA and the AR model for all three batteries along with the RMSE values. B0005, B0006, and B0007 displayed different trends in EOL forecasts, RUL predictions, and RMSE values. For B0005 and B0006 at 35% training data length, the AR model failed to predict the next states suggesting the need for more training data. Contrary, for B0007, the model was able to use the training data of 60 cycles and make a prediction curve, although not an accurate one proving that model training is more dependent on the values, trend, and pattern of the dataset.

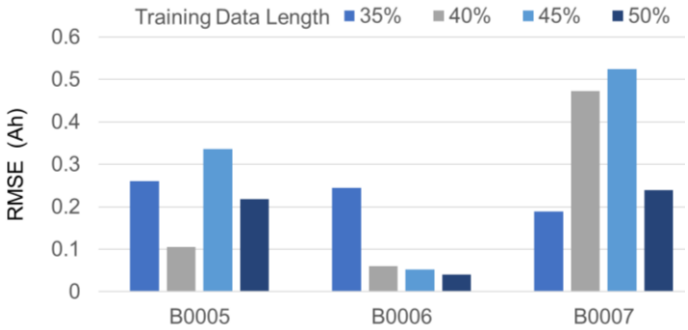

FIGURE 9: RMSE FOR AR MODELS

On the other hand, data from B0006 was accurately predicted compared to B0005 and B0007, giving more reasonable EOL forecasts and having the least RMSE values. Another interesting observation from Figure 9 is that the RMSE values for B0006 consistently decreased as training data length increased, which supports the general expectation that more training data improves the prediction model's accuracy. However, a mixed pattern for B0005 and B0007 was observed. These results emphasize that simply increasing training data does not guarantee better predictions. Instead, this highlights the necessity of tailoring optimization strategies to the unique characteristics of each cell dataset to achieve optimal results in battery health management and predictive modeling.

TABLE 4: AR MODEL RESULTS FOR B0005-0007 BASED ON VARIOUS TRAINING DATA LENGTH AND COMPARISON OF EOL CYCLES WITH EXPERIMENTAL DATA

	Training Data Length		EOL	EOL	EOL	RUL	RUL
	in Percent	in Cycle	Actual	Calculated	Difference	Actual	Calculated
B0005	35	60	124	N/A	N/A	64	N/A
	40	68		115	9	56	47
	45	76		102	22	48	26
	50	84		107	17	40	23
B0006	35	60	108	N/A	N/A	48	N/A
	40	68		114	-6	40	46
	45	76		96	12	32	20
	50	84		102	6	24	18
B0007	35	60	163	120	43	103	60
	40	68		106	57	95	38
	45	76		105	58	87	29
	50	84		117	46	79	33

For the 35% training length of B0005 and B0006, the models did not consist of values lower than 1.4Ah, and hence N/A (not available) was added to the result table.

4.2 ARIMA and SARIMA Results

Before discussing the SARIMA results, the ARIMA model was also developed to compare the results from already published literature by Zhou et al. [30]. ARIMA models act as a good base model for developing SARIMA as the only difference between them is the seasonality parameter. Table 3 compares the results of experimental data of NASA, published literature by Zhou et al. [30] and the work done in this research.

Table 5 presents a comparison of RUL predictions by the ARIMA model for B0005 and B0006 at various starting cycle numbers. The predicted RUL from this research demonstrated greater accuracy with RUL values consistently closer to the experimental results. This improved performance is further validated by the Root Mean Square Error (RMSE) values, where the proposed model consistently shows lower or comparable RMSE values.

The ARIMA model performed significantly better compared to the AR model in terms of capacity prediction values for B0005 and B0006. However, predictions for B0007 in the last phase could not be completed properly due to the limited number of values below the 1.4Ah threshold in the original dataset of B0007.

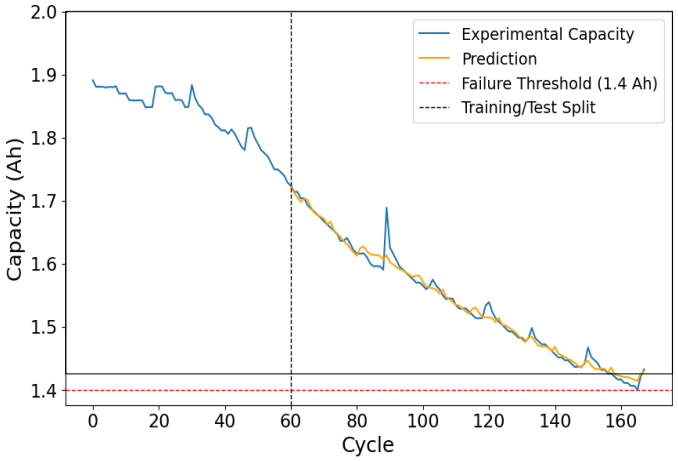

FIGURE 10: ARIMA MODEL FOR B0007 WITH 35% TRAINING DATA

TABLE 5: COMPARISON OF RUL VALUES FROM PUBLISHED LITERATURE [30] AND THIS RESEARCH USING ARIMA MODEL

	Starting Cycle	RUL Actual	RUL Zhou et al. [30]	RUL Calculated	RMSE Zhou et al. [30]	RMSE Calculated
B0005	60	64	53	64	0.0127	0.0119
	70	54	35	57	0.0235	0.0186
	80	44	29	34	0.0214	0.0259
	90	34	64	34	0.0376	0.0101
B0006	50	58	75	59	0.0491	0.0205
	60	48	34	49	0.0244	0.0209
	70	38	18	39	0.0168	0.0222

TABLE 6: SARIMA MODEL RESULTS FOR B0005-0007 BASED ON VARIOUS TRAINING DATA LENGTH AND COMPARISON OF EOL CYCLES WITH EXPERIMENTAL DATA

	Training Data Length		EOL	EOL	EOL	RUL	RUL
	in percent	in cycle	Actual	Calculated	Difference	Actual	Calculated
B0005	35%	60	124	125	-1	64	65
	40%	68		122	2	56	54
	45%	76		124	0	48	48
	50%	84		124	0	40	40
B0006	35%	60	108	109	-1	48	49
	40%	68		109	-1	40	41
	45%	76		108	0	32	32
	50%	84		108	0	24	24
B0007	35%	60	163	162	1	103	102
	40%	68		N/A	N/A	95	N/A
	45%	76		163	0	87	87
	50%	84		N/A	N/A	79	N/A

As shown in Figure 10, the prediction curve struggles to reach the failure threshold region for the ARIMA model of B0007, due to a low number of data points. This suggests that including the seasonality parameter might help the model to incorporate this shortcoming and predict values efficiently in the lower limit zone of 1.4Ah. Referring to Table 6 and Figure 8, the SARIMA model at 40% and 45% training data length had the smallest prediction errors which indicates that these configurations provide a good balance for accurate predictions for the datasets from B0005, B0006, and B0007. For Battery B0005 and B0006, the model provided reliable and consistent EOL predictions, with relatively small discrepancies between projected and actual cycles ranging from -1 to 2 cycles across several training splits (negative sign in the difference column indicates prediction value larger than the threshold value). Referring to the RMSE values from Figure 11 and comparing them against the results of the AR model, the SARIMA model has lower RMSE values across all configurations, which suggests the model can fit curves considerably better and capture the overall trends. On the whole, ARIMA and SARIMA models were reliable across the three datasets, especially at 45% and 50% training data length, with B0005 and B0006 showing the most stable and accurate predictions.

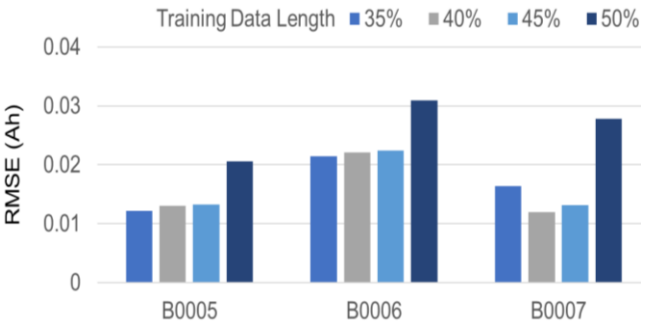

FIGURE 11: RMSE FOR SARIMA MODELS

These results show that the SARIMA model is able to capture both the long-term trend and the seasonal pattern in battery degradation, making it a good candidate for time-series modeling of battery capacity. The SARIMA model was also successful in giving us RUL estimates of B0007 whereas the AR and ARIMA models couldn't. However, there remains a gap between EOL prediction and experimental values in the current model. It can be reduced using an optimized SARIMA model by choosing different parameter values and observing how these changes affect the prediction results in each case.

4.3 EOL Comparison of All Models.

Figure 12 and Table 7 compares the performance of AR, ARIMA, and SARIMA models across three datasets (B0005, B0006, and B0007) with varying training data lengths in predicting the EOL cycles using mean values of the predicted EOL along with standard deviation and percent difference of actual and mean EOL. B0005 and B0006 datasets showed great adaptability with EOL values aligning closely with the actual experimental values for all three models. Since the B0007 dataset consisted of very few points in the lower limit of the failure threshold range, the prediction values for the ARIMA model and two instances in the SARIMA model did not cross the 1.4Ah mark. Figure 12 shows a comparison of predicted EOL values for all training dataset lengths as well as the performance of AR, ARIMA, and SARIMA models with the experimental value (red dotted line) at a glance.

ARIMA model for B0006 showed consistency in predicting the values for all training data lengths and yielded no EOL value at the threshold capacity for B0007 due to insufficient data at the defined failure zone value of 1.4Ah. AR model couldn't give any reliable output for 35% training data length due to low training data length and the predicted EOLs had a considerable difference.

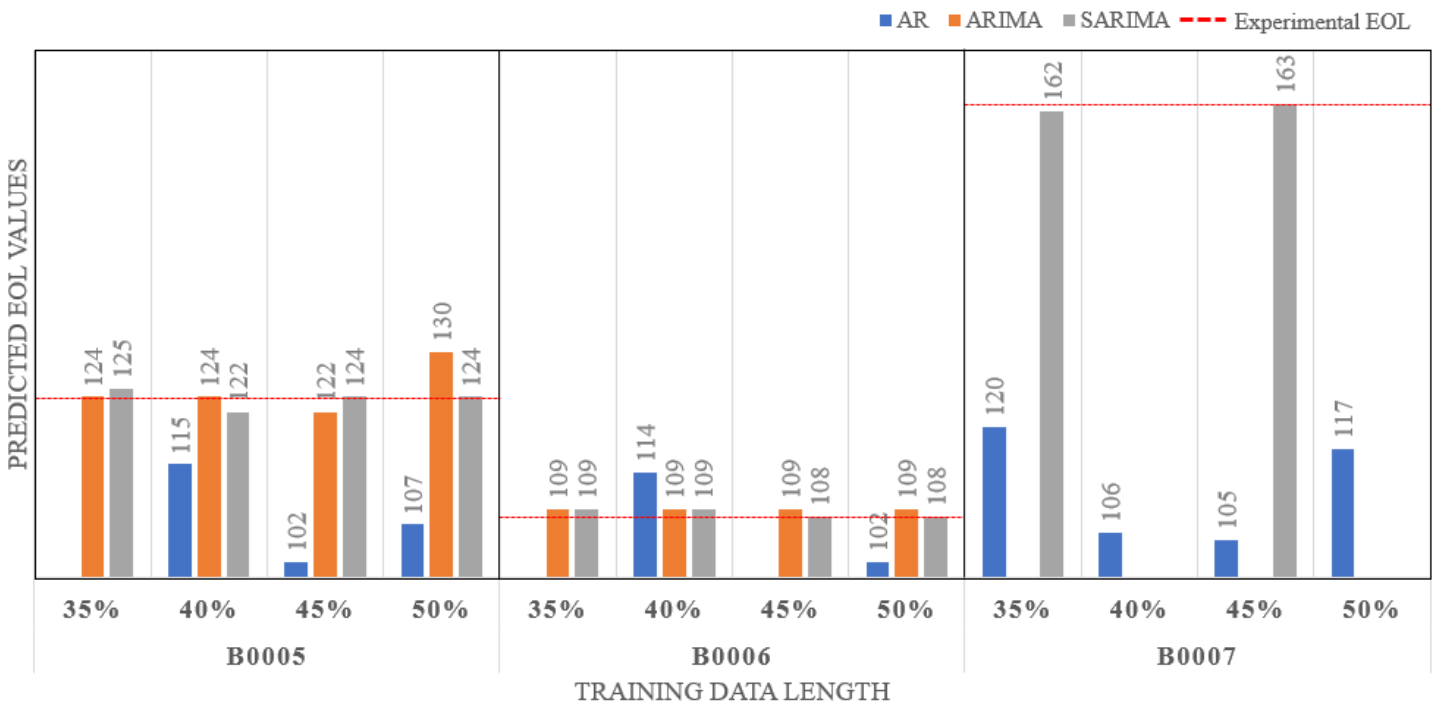

FIGURE 12: PREDICTET EOL VALUES FOR COMPARING MODEL PERFORMANCES

TABLE 7: MEAN EOL, STANDARD DEVIATION, AND PERCENT DIFFERENCE IN ACTUAL VERSUS MEAN EOL

	Mean Predicted EOL			Standard Deviation of Predicted EOL			Percent Difference of Actual and Mean EOL		
	AR	ARIMA	SARIMA	AR	ARIMA	SARIMA	AR	ARIMA	SARIMA
B0005	108	125	123.75	5.35	3	1.08	-12.9	0.80	-0.20
B0006	104	109	108.5	7.48	0	0.5	-3.7	0.92	0.46
B0007	112	N/A	162.5	6.59	N/A	0.5	-31.28	N/A	-0.30

Table 7 shows the mean predicted EOL values for B0005-0007 along with standard deviation and the percent difference in actual versus mean predicted EOL.

Lastly, the SARIMA model was able to train and give accurate results for two scenarios and no values at 40% and 50% training data length. It should be noted that even though SARIMA model results were not as accurate as ARIMA predictions in terms of RMSE values, it still showed significantly good curve fitting properties in terms of capturing the peaks and valleys of the original capacity curve, meaning a better understanding of the trends in the dataset caused by the charge-discharge cycles.

5. CONCLUSION

This research underlines the important role that predictive modeling plays in managing LIB degradation, especially if implemented in the maritime sector, where efficient and reliable energy storage systems are crucial for reducing fuel consumption and mitigating environmental impacts. A comparison study of the data predicted using the AR, ARIMA, and SARIMA models indicates relative strengths and limitations in capturing the pattern of battery degradation and, therefore, EOL and RUL prediction.

The following conclusions can be drawn from this research based on the defined objectives:

- From the results, it can be concluded that ARIMA and SARIMA models are well-suited to assess the degradation behavior of LIBs in terms of EOL prediction as well as RMSE. On the other hand, the AR model being the simplest in this time series family, helps set a base by developing easy and computationally efficient models. It sets up a foundation by providing a benchmark to access the added value of differencing (I term), seasonal, and moving average components.
- ARIMA model was validated using existing research by Zhou et al. [30] by comparing RUL and RMSE values. Referring to Table 5, ARIMA from this research was found to be more precise, therefore setting a good base to develop the SARIMA model.
- Overall RMSE values were calculated to be 0.22872, 0.01468 and 0.01878 for AR, ARIMA, and SARIMA

respectively for the same training data lengths. The high RMSE value of AR can be explained due to its linear prediction capability. Based on the RMSE, it can be concluded that the ARIMA model provides the best curve fitting for this research.

- Lastly, a model comparison was done using the mean predicted EOL, standard deviation and the percent difference between actual EOL and the mean predicted EOL values. Referring to Table 7, the AR model has the highest standard deviation for the predicted EOL for B0005-0007, and SARIMA had the lowest values. SARIMA has the closest EOL values to the experimental values and least standard deviation proving that the model is able to provide the most accurate EOL forecasts for this dataset. On the other hand, the ARIMA model also predicted EOL values with an error of less than 1% which also suggests it to be a useful degradation model. However, due to a lack of seasonal terms and a low number of data points for B0007, could not yield EOL values, proving the importance of the SARIMA model.

6. FUTURE WORK

This research focused on open-source battery degradation data and developed degradation prediction models without a specific target application. In future work, these degradation models will be extended to sector-specific applications, especially in maritime energy storage systems. Adapting and validating these models with relevant practical datasets is important for assessing the effectiveness in real-world maritime operations.

On the other hand, the range in predictions among individual battery cells might be reduced; it was observed that some cells had extremely accurate projections, while others exhibited significant differences, this similar observation was also noted in the work of Vanem et al. [19]. Cell-specific modeling or clustering strategies are likely to produce higher overall prediction accuracy. Uneven Depth of Discharge also known as DOD cycles results in a non-uniform pattern of charging and discharging, affecting model consistency during training due to increased trends and seasonality. Alternative time-series models, like Long Short-Term Memory, NNs, or even Physics inspired NNs, may provide superior predictive performance due to capturing nonlinear relationships and long-range dependencies present in degradation data. Addressing these challenges will enhance the robustness and reliability of battery health prediction models in the future.

7. REFERENCES

[1] J. Ling-Chin and A. P. Roskilly, "A comparative life cycle assessment of marine power systems," Energy Conversion and Management, vol. 127, pp. 477–493, 2016, DOI: 10.1016/j.enconman.2016.09.012.

[2] Ø. Buhaug et al., "Second IMO GHG study 2009," 2009.

[3] N. Gray, S. McDonagh, R. O'Shea, B. Smyth, and J. D. Murphy, "Decarbonising ships, planes and trucks: An analysis of suitable low-carbon fuels for the maritime, aviation and haulage sectors," Advances in Applied Energy, vol. 1, p. 100008, 2021, DOI: 10.1016/j.adapen.2021.100008.

[4] N. Bennabi, J. F. Charpentier, H. Menana, J. Y. Billard, and P. Genet, "Hybrid propulsion systems for small ships: Context and challenges," in 2016 XXII International Conference on Electrical Machines (ICEM), Lausanne, Switzerland, 2016, pp. 2948–2954, DOI: 10.1109/ICELMACH.2016.7732943.

[5] X. Han et al., "A review on the key issues of the lithium ion battery degradation among the whole life cycle," eTransportation, vol. 1, p. 100005, 2019, DOI: 10.1016/j.etran.2019.100005.

[6] C. P. Grey and J. M. Tarascon, "Sustainability and in situ monitoring in battery development," Nature materials, vol. 16, no. 1, pp. 45–56, 2016, DOI: 10.1038/nmat4777.

[7] L. Lu, X. Han, J. Li, J. Hua, and M. Ouyang, "A review on the key issues for lithium-ion battery management in electric vehicles," Journal of Power Sources, vol. 226, pp. 272–288, 2013, DOI: 10.1016/j.jpowsour.2012.10.060.

[8] G. E. Blomgren, "The Development and Future of Lithium Ion Batteries," J. Electrochem. Soc., vol. 164, no. 1, A5019-A5025, 2017, DOI: 10.1149/2.0251701jes.

[9] C. Hendricks, N. Williard, S. Mathew, and M. Pecht, "A failure modes, mechanisms, and effects analysis (FMMEA) of lithium-ion batteries," Journal of Power Sources, vol. 297, pp. 113–120, 2015, DOI: 10.1016/j.jpowsour.2015.07.100.

[10] R. Guo, L. Lu, M. Ouyang, and X. Feng, "Mechanism of the entire overdischarge process and overdischarge-induced internal short circuit in lithium-ion batteries," Scientific reports, DOI: 10.1038/srep30248.

[11] B. Saha, K. Goebel, S. Poll, and J. Christophersen, "Prognostics Methods for Battery Health Monitoring Using a Bayesian Framework," IEEE Trans. Instrum. Meas., vol. 58, no. 2, pp. 291–296, 2009, DOI: 10.1109/TIM.2008.2005965.

[12] Y. Hu, P. Baraldi, F. Di Maio, and E. Zio, "A particle filtering and kernel smoothing-based approach for new design component prognostics," Reliability Engineering & System Safety, vol. 134, pp. 19–31, 2015, DOI: 10.1016/j.ress.2014.10.003.

[13] Q. Miao, L. Xie, H. Cui, W. Liang, and M. Pecht, "Remaining useful life prediction of lithium-ion battery with unscented particle filter technique," Microelectronics Reliability, vol. 53, no. 6, pp. 805–810, 2013, DOI: 10.1016/j.microrel.2012.12.004.

[14] J. D. Kozlowski, "Electrochemical cell prognostics using online impedance measurements and model-

based data fusion techniques," in 2003 IEEE Aerospace Conference Proceedings (Cat. No.03TH8652), Big Sky, Montana, USA, 2003, pp. 3257–3270, DOI: 10.1109/AERO.2003.1234169.
[15] D. Liu, Y. Luo, J. Liu, Y. Peng, L. Guo, and M. Pecht, "Lithium-ion battery remaining useful life estimation based on fusion nonlinear degradation AR model and RPF algorithm," Neural Comput & Applic, vol. 25, 3-4, pp. 557–572, 2014, DOI: 10.1007/s00521-013-1520-x.
[16] G. T. Wilson, "Time Series Analysis: Forecasting and Control, 5th Edition, by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel and Greta M. Ljung, 2015. Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1," Journal Time Series Analysis, vol. 37, no. 5, pp. 709–711, 2016, DOI: 10.1111/jtsa.12194.
[17] Yun Yang et al., "A Battery Capacity and RUL Prognostic Approach Based on ARIMA and PF," in 2021. pp.220-232, DOI: 10.6911/WSRJ.202105_7(5).0023
[18] Q. Liang et al., "Data-driven state of health monitoring for maritime battery systems – a case study on sensor data from ships in operation," Ships and Offshore Structures, pp. 1–13, 2023, DOI: 10.1080/17445302.2023.2211241.
[19] E. Vanem, M. Bruch, Q. Liang, K. Thorbjørnsen, L. O. Valøen, and Ø. Å. Alnes, "Data-driven snapshot methods leveraging data fusion to estimate state of health for maritime battery systems," Energy Storage, vol. 5, no. 8, 2023, Art. no. e476, DOI: 10.1002/est2.476.
[20] H. Xu, Y. Peng, and L. Su, "Health State Estimation Method of Lithium Ion Battery Based on NASA Experimental Data Set," IOP Conf. Ser.: Mater. Sci. Eng., vol. 452, p. 32067, 2018, DOI: 10.1088/1757-899X/452/3/032067.
[21] V. M. Nagulapati, H. Lee, D. Jung, B. Brigljevic, Y. Choi, and H. Lim, "Capacity estimation of batteries: Influence of training dataset size and diversity on data driven prognostic models," Reliability Engineering & System Safety, vol. 216, p. 108048, 2021, DOI: 10.1016/j.ress.2021.108048.
[22] J. Wang et al., "Research on Lithium-Ion Battery State of Health Prediction Based on Autoregressive Integrated Moving Average Mode," in 2024 39th Youth Academic Annual Conference of Chinese Association of Automation (YAC), Dalian, China, 2024, pp. 1238–1242, DOI: 10.1109/YAC63405.2024.10598491.
[23] Z. Liu, J. Zhang, and F. Feng, "Research on the prediction method of lithium-ion battery remaining life based on EMD-Arima," in 2022 International Conference on Manufacturing, Industrial Automation and Electronics (ICMIAE), Rimini, Italy, 2022, pp. 357–360, DOI: 10.1109/ICMIAE57032.2022.00077.
[24] NASA. "Li-ion battery aging datasets." https://data.nasa.gov/dataset/Li-ion-Battery-Aging-Datasets/uj5r-zjdb/about_data
[25] M. J. Daigle and K. Goebel, "Model-Based Prognostics with Concurrent Damage Progression Processes," IEEE Trans. Syst. Man Cybern, Syst., vol. 43, no. 3, pp. 535–546, 2013, DOI: 10.1109/TSMCA.2012.2207109.
[26] S. S. Sheikh, F. A. Shah, S. O. Athar, and H. A. Khalid, "A Data-Driven Comparative Analysis of Lithium-Ion Battery State of Health and Capacity Estimation," Electric Power Components and Systems, vol. 51, no. 1, pp. 1–11, 2023, DOI: 10.1080/15325008.2022.2145389.
[27] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature," Geosci. Model Dev., vol. 7, no. 3, pp. 1247–1250, 2014, DOI: 10.5194/gmd-7-1247-2014.
[28] B. Moseley, "Physics-informed machine learning: from concepts to real-world applications," University of Oxford, 2022.
[29] S. Kim, P.-Y. Lee, M. Lee, J. Kim, and W. Na, "Improved State-of-health prediction based on auto-regressive integrated moving average with exogenous variables model in overcoming battery degradation-dependent internal parameter variation," Journal of Energy Storage, vol. 46, p. 103888, 2022, DOI: 10.1016/j.est.2021.103888.
[30] Y. Zhou and M. Huang, "Lithium-ion batteries remaining useful life prediction based on a mixture of empirical mode decomposition and ARIMA model," Microelectronics Reliability, vol. 65, pp. 265–273, 2016, DOI: 10.1016/j.microrel.2016.07.151.
[31] M. Catelani, L. Ciani, R. Fantacci, G. Patrizi, and B. Picano, "Remaining Useful Life Estimation for Prognostics of Lithium-Ion Batteries Based on Recurrent Neural Network," IEEE Trans. Instrum. Meas., vol. 70, pp. 1–11, 2021, DOI: 10.1109/TIM.2021.3111009.

