PREDICTION OF LITHIUM-ION BATTERY CAPACITY BY FUNCTIONAL MONITORING DATA USING FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS

A Thesis presented to
The Faculty of the Graduate School
At the University of Missouri-Columbia

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science, Industrial Engineering

By
MD SHORIAT ULLAH

Dr. Kangwon Seo, Thesis Supervisor

DECEMBER 2020

The undersigned appointed by the Dean of the Graduate School, have examined the thesis entitled:

PREDICTION OF LITHIUM-ION BATTERY CAPACITY BY FUNCTIONAL MONITORING DATA USING FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS

presented by **MD Shoriat Ullah**,

a candidate for the degree of Master of Science and hereby confirm that, in their opinion, is worthy of acceptance.

Dr. Kangwon Seo

Dr. James Noble

Dr. Xueju Wang

ACKNOWLEDGEMENTS

I am grateful to my adviser, Dr. Kangwon Seo, for his guidance, patience, and understanding throughout the duration of the research. His continuous advice and supervision have contributed a key role in determining my current position and providing me an encouraging experience at the University of Missouri. He encouraged me to remain focused and provided me the support I needed to overcome many situations throughout my Master's journey. It has been a matter of great pleasure to work under his supervision and I would like to express my sincere gratitude for making my academic life a memorable one. I would also like to express my deepest gratitude to my committee members for the time and effort. Dr. Noble and Dr. Wang helped me with their insightful comments and suggestions throughout the duration of my research and provided encouraging feedback on my oral defense. Justly, it's an honor and a pleasure for being your student and working under your supervision.

I would like to thank all the faculties, staffs, and my peers collectively for providing me the platform and ambiance to complete and excel in my academic as well as social life. I learned a lot from all of them which is very much precious to me. Lastly, I would like to thank my parents and relatives for their love and support throughout my life especially while staying alone that has allowed me to successfully complete my Master's and this journey of emotions.

Table of Contents

ACKNOWLEDGEMENTS .... **ii**

Table of Contents .... **iii**

List of Table .... **v**

List of Figure .... **vi**

Abstract .... **vii**

Chapter 1 .... **1**

Introduction 1

Chapter 2 .... **11**

Literature Review **11**

2.1. Physical Model. 12

2.2. Data-Driven Models. 14

2.3. Research scopes. 20

Chapter 3 .... **23**

Methodology 23

3.1. Overview of the problem and data analysis 23

3.2. Data collection 23

3.2. Model building and testing 25

3.3. Functional Principal Component Analysis: 28

3.4. Lasso Regression: 30

Chapter 4 .... **32**

Results and Analysis 32

4.1. Data visualization 32

4.2. Functional principal component analysis model 36

4.3. Linear regression model 44

4.4. Lasso regression model 46

Chapter 5 51

Conclusions and Future Work...51

5.1. Conclusions 51

5.2. Future Work..52

BIBLIOGRAPHY 54

List of Table

Table 2. 1 : Literature of data-driven models' classification 21
Table 2. 2: Input parameters of preceding studies involved with FPCA. 22
Table 4. 1: FPC scores for initial 20 cycles of voltage, current and temperature. 43
Table 4. 2: Prediction performance of battery No 7 for three separate cases. 50

List of Figure

Figure 1. 1: The battery comparison demonstrates the volumetric and specific densities (Miao et al. 2019).. 2

Figure 3. 1: Data structure of dataset demonstrates initial 25 sequential data of first cycle. 25

Figure 3. 2: Graphical representation of workflow.. 26

Figure 4. 1: The first 100 cycles of voltage rating against time steps. 33

Figure 4. 2: The first 100 cycles of current rating against time steps. 34

Figure 4. 3: The first 100 cycles of temperature against time steps. 35

Figure 4. 4: The degradation of capacity over the number of discharge cycle of lithium-ion battery. 36

Figure 4. 5: B-spline fitted for the first 50 cycles of voltage curves. 37

Figure 4. 6: Model fitting for actual voltage vs. voltage predicted................................... 38

Figure 4. 7: The mean and eigenfunctions of voltage cycles.. 38

Figure 4. 8: B-spline fitted for the first 50 cycles of current curves................................. 39

Figure 4. 9: The mean and eigenfunctions of current cycles. 40

Figure 4. 10: B-spline fitted for the first 50 cycles of temperature curves. 41

Figure 4. 11: Model fitting for actual temperature vs. temperature predicted.................. 41

Figure 4. 12: The mean and eigenfunctions of current cycles. 42

Figure 4. 13: The predicted 20 cycles of testing dataset for voltage. 44

Figure 4. 14: The predicted 20 cycles of testing dataset for current................................. 45

Figure 4. 15: The predicted 20 cycles of testing dataset for temperature. 45

Figure 4. 16: Actual versus predicted capacity (a) 100 cycles as training (b) 120 cycles as training for training and testing dataset 47

Figure 4. 17: Fitted curve for capacity prediction of initial 120 cycles............................ 48

Figure 4. 18: Fitted curve for capacity prediction of initial 140 cycles. 49

Abstract

Lithium-ion batteries have been a promising energy storage technology for applications such as electronics, automobiles, and smart grids over the years. Extensive research was conducted to improve the prediction of the remaining capacity of the lithium-ion battery. A robust prediction model would improve the battery performance and reliability for forthcoming usage. To develop a data-driven capacity prediction model of lithium-ion batteries most of past studies employed capacity degradation data, yet very few tried using other performance monitoring variables such as temperature, voltage, and current data to estimate and predict the battery capacity. In this thesis, we aim to develop a data-driven model for predicting the capacity of lithium-ion battery adopting functional principal component analysis applied to functional monitoring data of temperature, voltage, and current observations collected from NASA Ames Prognostics Center of Excellence repository. The result of capacity prediction has been substantiated with past studies and obtained root mean square error (RMSE) of 0.009. The proposed data-driven approach performs well to predict the capacity employing functional performance measures over the life span of a lithium-ion battery.

Chapter 1

Introduction

The concept of storing potential energy in the form of a battery has changed human lives to a great extent. The human race secured immense flexibility and adaptability availing such technology into their daily life. The applications of the battery have been dominating everyday life from a tiny headphone to the laptop, electrical vehicle to the airplane industry. Its wide-ranging operations not only assist human activities but also allowed us to make advancements in science and technology. Being too much dependent upon such crucial technology these days, the ramification of any kind of failure in performance is unbearable for us. Even in some cases, we would have faced functional impairment along with disastrous failures. While battery has become an indispensable part of our daily necessity, understanding battery performance degradation could aid in improving user satisfaction and would help to enhance the overall reliability of a system. There are several kinds of rechargeable batteries available such as Lead-acid, Nickel- cadmium and Lithium-Ion battery. Among these lithium-ion is the latest development and has been used for its compact in size and high energy density (Types of Battery Cells, 2019). Figure 1.1 shows the comparison of available materials and it’s volumetric and specific energy densities (Miao et al. 2019), and it shows lithium-ion batteries are preferable for most of the electrical systems. Recent studies are showing by 2030 the demand of rechargeable lithium batteries will be expected to reach approximately 350,000 metric tons (Statista,2019). Meanwhile, the lithium-ion battery is a vital constituent for lots of our known systems and became very crucial for the performance. As demand rises for such batteries, researchers are trying to keep more scrutiny to health monitoring and prognostics for lithium-ion batteries. Scientists are putting continuous effort to control the operating conditions, predicting replacement intervals for battery aiming to improve the system reliability and consistency.

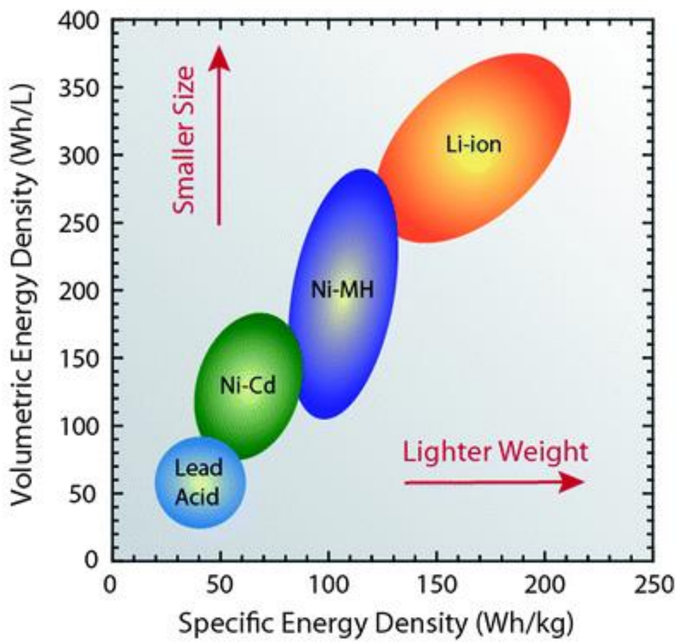


Figure 1. 1: The battery comparison demonstrates the volumetric and specific densities (Miao et al. 2019).

The study presented in this thesis focuses on these issues by developing an empirical model based on lithium-ion battery capacity degradation which is affected by the ambient environment and load condition. While developing the model, the study addresses the effect of current, voltage, and temperature on the capacity degradation applied to Functional Principal Component Analysis (FPCA) to predict the remaining capacity of a battery. An effective prognostic algorithm should be able to predict the remaining useful life or capacity using the past life cycle data of battery performance.

Most of the existing research on battery prognostics and health monitoring studies are based upon tracking the state-of-health (SOH) and state-of-charge (SOC) employing battery capacity data due to the lack of battery information (S. Lee et al. 2012). These algorithms frequently contribute measurement inaccuracies and misleading assumptions which are often root causes of failures. This unexpected behavior of battery also forcing people to avoid cost-effective and environment-friendly lithium-ion batteries over hazardous led-acid batteries (Geoff S. Fein,2003). However, lithium-ion batteries are significantly different at the rudimental level from other available conventional cells. First, the electrochemical reaction occurring inside the battery is a laborious task to measure and secondly, computable information that can be achieved from battery terminal is a form of continuous function such as current, voltage, impedance and temperature (Gao et al. 2002). Researchers are trying to imitate the dynamic behavior of lithium-ion batteries (Gao et al. 2002) with the help of advanced analytical modeling and data analysis (Zhang and Lee 2011). The SOC estimation is one of the most widespread research topics in recent days. An accurate SOC estimation helps to assess the reliability of products along with it delivers some crucial performance measures such as remaining useful energy and time. Guiheen et al. (2004) proposed SOC estimation process that needed to record relationships of the measurement of ramp-pick current and SOC. However, the overall test process requires extensive data collection and evaluation which is a cumbersome procedure. One major drawback was the idea of data tabulation which matched with the actual SOC estimation. This discrepancy also results in the ambiguity between estimated SOC and actual SOC (S. J. Lee et al. 2007). These uncertainties have compelled future research to develop more advanced SOC estimation methods such as simulation, physical cell degradation, and data-driven prediction instead of relying on SOC-OCV (open circuit voltage tables) (Zhang and Lee 2011). Most of the developed methods are model-based and data-driven process

(Goebel et al. 2008). Those models can be classified as electrochemical models, circuit-based models, performance-based models, and analytical models (Barré et al. 2013). The electrode theory and physical degradation of the cathode or other specific parts in a cell are the main focus of an electrochemical process. These models are largely described and discussed the Nano-mechanics of the battery (S. J. Lee et al. 2007) While circuit-based models are designed using a circuit and performance measures are being calculated with the assistance of a focused circuit. However, These models are not pertinent to separate types of batteries even in some cases the outcomes are not reproducible (Barré et al. 2013; Goebel et al. 2008; Peled, Golodnitsky, and Ardel 1997).Performance types model is primarily researched about the influence of ageing over the performance of a battery. Though the performance of a battery over its life-time may depend on several environmental and physical variables, the overall scenario can be imitate using advanced simulation techniques and plenty of accelerated life testing procedures under heavy workload (Eddahech et al. 2015). However, an equivalent circuit model was developed to investigate the performance measures where authors have emphasized at the influence of thermal ageing on the capacity fade phenomenon which is the foremost anticipated reason for degrading battery calendar life over the battery storage, standby even in the operation periods (Einhorn et al. 2010). Finally, numerous studies were conducted focusing on sudden capacity degradation, remaining useful life (RUL) estimation based upon data visualization and analysis and further development of mathematical algorithms (Zhang and Lee 2011). Saha et al. (2007) explored the way of addressing RUL for a complex system using historical data in anticipated operational conditions. They have used a Bayesian statistical approach as an analytical tool where relevance vector machines (RVMs) and particle filters (PFs) are examined for RUL prediction. Adaptive recurrent neural network (ARNN) is also proposed for prediction by J. Liu et al. (2010) where investigators developed an ANN architecture and weights were calculated and optimized using recursive Levenberg-Marquardt (RLM) method. Gao et al. (2002) developed a complete model for a lithium-ion battery which is designed for nonlinear equilibrium potential, rate and temperature dependencies, thermal effects using an extended Kalman filter (EKF). Moreover, other machine learning algorithms have also been used to predict the RUL of the lithium-ion battery. Goebel et al. (2008) combined state transitions, the ageing process and measurement fidelity in a framework and used Bayesian treatment of the support vector machine (SVM) to diagnose the process and estimate noise in the system. The analysis eventually helped them to measure the state of charge (SOC), SOH and SOL. Estimating the SOH and predicting the RUL now has become one of the most common practices where researchers are using a physical, chemical, and empirical model.

While most of the studies have been conducted in the past regarding RUL estimation and developed numerous algorithms such as machine learning, statistical and empirical models considering past capacity degradation data or SOH information, very few studies have been emphasized other crucial factors such as current, voltage, or temperature during discharge. However, the capacity data also depends upon other parameters such as current, voltage, temperature, and load (Birkl et al. 2017). F. K. Wang and Mamo (2018) investigated parameters especially current and voltage to establish a hybrid model using SVM and differential evolution to predict the RUL of a lithium-ion battery. They have calculated the mean value of each cycle for both the current and voltage and developed a model with corresponding cycle capacity which resultant improved prediction for RUL estimation.

There are very few studies has been conducted based on the complicated correlation and framework, to obtain the precise and reliable measures from these relations or analytical model seems very cumbersome approach. Lu et al. (2014) suggested a geometric approach for lithium-ion batteries for projecting the capacity degradation using current, voltage and temperature. They have extracted four separated features from lithium-ion battery data cycle which are correlated with capacity degradation for further assessment. However, inaccurate features might impact the accuracy of capacity estimation. Furthermore, this technique only provides the point estimates which has a tendency to ignore the confidence interval and does not provide the residual lifetime distribution (RLD). However, recently L. Li et al. (2016) established a residual lifetime estimation framework using the Gaussian process regression algorithm. The approach has solved the RLD and provide a confidence interval up to the failure threshold. The method also can be utilized for multimodality by fitting separate trajectories with different Gaussian regression models. Yet there is a drawback of the method since the model is a parametric model, it needs the parametric forms of the model component.

While the data seems in the shape of random curves rather than vectors and scaler values, the phenomenon of dimension reduction is a kind of compulsory procedure. Functional principal component analysis (FPCA) has become one of the common tools. Significant research has been done over the years emphasizing FPCA in different fields of the scientific world. In bioinformatics, the most challenging task is to deal with the dimensionality of measurements without losing the generality. The principal component analysis is one of the classic dimension reduction approaches while functional principal component analysis helps to analyze time-course expression data (Ma and Dai 2011). FPCA integrate the influence of predictor variables either through the mean response function or functional principal expansion which reduce random trajectory and variability (Yao et al. 2004). Furthermore, FPCA aids to classify the dominants modes of variation of a sample of random continuous functional data around an overall mean trajectory. The method was introduced while studying individual growth curves classified by litter, sex, and treatment (Rao 1958). While Rice and Silverman have studied the smoothing techniques of a random function with an unknown mean and covariance structure (Rice and Silverman 1991). They proposed smooth nonparametric estimates of the eigenfunctions and a method of cross-validation to evaluate the smoothing. There were other studies also conducted cluster the functional data. Basis spline (B-spline) has been proposed to model the individual curves with random coefficient for sparse, irregular spaced and separate spaced time points data set (James and Sugar 2003). A recent study Cheng et al. (2015) has been conducted to predict residual lifetime using FPCA and Bayesian approach where the authors developed a nonparametric degradation model. The authors only emphasized on the capacity degradation rather considering other parameter such as current, voltage and temperature. In this thesis, we have used the basis spline model of FPCA to fit the data along with generalized regression to select variables for further study.

FPCA is a statistical tool to analyze continuous functions and able to extract key features that epitomize dominant aspects of the original dataset (Kipp et al. 2011). These significant attributes are contributing factors to utilize the evolutional rules for useful lifetime prediction and extensively used in other research fields as well. Variation in sea surface temperature has been studied to predict entire smooth functions in the future since the cycle is considered smooth functions and well defined by FPCA (Besse et al. 2000). While dealing with multi-dimensional functional data of protein structure, researchers have introduced FPCA utilizing the Gaussian basis function (Kayano and Konishi 2009). Moreover, they have developed cross-product matrix performing regularized FPCA with B-spline. Financial time series data has been analyzed using FPCA and found that the first FPC score accounts for 89.4% of the whole variability (Ingrassia and Costanzo 2005). Studies also have been conducted for shrinkage estimation using FPCA for the population kinetics which contains an irregular grid of time points (Yao et al. 2003). The study introduced a scatterplot smoothing method to measure the mean function and covariance surface of the model. Y. Hu et al. (2009). also studied the children's growth data using FPCA and predicted weight growth in children. The authors investigate the probability of the impact on the transformation growth of children’s over their predicted model using FPCA. Furthermore, human movement waveforms have also been analyzed using FPCA (Epifanio et al. 2008). The statistical tool was applied to study the sit-to-stand movement of two separate groups such as osteoarthritis patients and healthy subjects. FPCA presented extensive discriminatory power relative to the classical multivariate approach. Coffey et al. (2011))also studied human movement using FPCA and found that the technique retains all the information intact for further analysis and the analysis is an extension of multivariate principal component analysis. The authors conducted a t-test of each group FPC score to classify significant differences among the experimental groups. These increasing numbers of research have reported that a more comprehensive conclusion can be depicted by FPCA than traditional discrete data analysis methods while data sets are continuous functions (Kipp et al. 2011). Thus, according to the aforementioned literature and research, it is visible that the FPCA technique is one of the leading alternatives to develop a capacity degradation model of a lithium-ion battery and predict the remaining useful life. Research using PCA applied to various datasets is widely recognized. However, analyzing PCA with regression creates one of the major difficulties which is multicollinearity.

Explaining and addressing multicollinearity require biased regression estimators which includes ridge regression, shrinkage estimator and lasso regression (I. T. Jolliffe 1986). These regression models are known as generalized regression. The generalized regression model (GRM) describes observed spatial variation in the current, voltage and temperature. Shamsudduha et al. (2015) have studied arsenic variation in the shallow water using a generalized regression model and found a significant concentration variation in shallow groundwater at both national and regional levels. Furthermore, the GRM demonstrated that the spatial variation explained by the statistical interactions. For functional data, the generalized linear regression model has also been estimated via penalized likelihood (Cardot and Sarda 2005). The authors initially developed the theoretical framework and defined the model. The functional coefficient has been estimated via penalized likelihood using spline approximation. Thus, based upon the generalized regression, the maximum likelihood of battery monitoring data (i.e., current, voltage, and temperature) can be calculated.

In this study, we aim to predict the capacity of lithium-ion batteries using a data-driven prognostics algorithm supporting uncertainty representation and management. While building the data-driven model, we have considered battery ageing stress factors such as voltage, current, and temperature. Specifically, we adopted FPCA to study functional data which is structured as cycle from battery terminal amid of discharge phase of a battery. As the outcome of FPCA is FPC scores of variables such as voltage, temperature and current treated as high dimensional data and needed to address model correlation, we have used lasso regression to fit the model and predict the remaining capacity.

The remaining thesis is ordered as follows. Chapter 2 provides in-depth literature reviews and past research studies regarding various data-driven techniques and models that have been employed to solve capacity prediction problems for lithium-ion battery. Chapter 3 demonstrates the methodology that we have applied for solving the problems along with proper mathematical explanations and models. The solutions obtained from the developed model are discussed and demonstrated in Chapter 4 as Results. Finally, Chapter 5 comprises of conclusion along with the scope of future research.

Chapter 2

Literature Review

The battery aging phenomenon and capacity degradation should be emphasized to make an unyielding optimum and reliable battery system, serving a longer life without compromising performance. The ageing mechanism and degradation model need to analyze the influence of critical parameters such as voltage, current, temperature, energy density, and power density at the cell level (Han et al. 2019). From the preview of the overall battery management system, the models are also significant to do the health estimation, which is often influenced by the history, optimization of working parameters, and prediction of desired performance.

The prognostics of lithium-ion batteries deal with the energy or power that is degraded over time of usage and predict how much time the battery will perform up to the desired level (Y. Li et al. 2019). The health monitor requires the preceding information of the performance measures and historical capacity fading signals, often can be achieved from a data-driven estimator to forecast systems forthcoming condition under specified operating loads. The current technological advancement in data science made it possible to make a breakthrough battery life and keep battery performance consistent for the lithium-ion battery lifetime. The battery will degrade over time due to ageing, environmental effects under dynamic loading conditions (Zhang and Lee 2011). Thus, it is necessary to perceive the underlying degradation process and address the issue with countermeasures to hinder the terrible failures from occurring midst of any system. Prognostic health care is another integral part of the management system. It deals with fault propagation and forecasts the remaining lifetime of a component. Health monitoring and prognostic of lithium-ion batteries have become a scope of mass attention in the researcher’s community and every year, hundreds of articles, studies are being published in conference proceedings and academic journals. The most widely cited models and technologies for lithium-ion batteries are discussed in the following sections. Most of the studies that we came across through detailed literature reviews either tried to build a data-driven model using capacity and ignoring other ageing stress factors (Chang, Fang, and Zhang 2017; Chen et al. 2016) or not addressing the issue of functional/sequential data of the ageing stress factors (F. K. Wang and Mamo 2018) which is why for the purpose of this research, we will consider these ageing stress factors such as current, voltage and temperature along with considering sequential data applied to the functional principal component analysis and further lasso regression.

2.1. Physical Model.

The first characteristic that would be presented here is the concept and studies related to the battery ageing phenomenon originated for electrochemical erosion. Based on these ageing erosions, on battery ageing factors, effects and attempted to measure the battery capacity. These studies are explained from several fields such as electrochemical models and analytical models. Notable studies are discussed in the following sections.

2.1.1. Electrochemical ageing.

Ageing of lithium-ion battery initiates by the chemical structure of the electrolyte. The capacity degradation proceeds from the positive and negative electrode are significantly different, and origin could be either chemical or mechanical, which sturdily depends upon battery composition (Vetter et al. 2005). The battery's lifetime is responsible and incites the component degradation, which can induce and modify structural characteristics, involves variation in chemical composition or erosion of active material such as electrolyte (Jannesari et al. 2011). Several types of deformation may be observed throughout the capacity degradation period, such as current collector tearing through a layer, kinking layer, and local melting (B. Liu et al. 2020). The most detrimental for capacity degradation happens due to the material composition, especially battery contained with LiPF6 in high wet ratio in the battery. These degradations or corrosion can be avoided by rudimental interphase or acid scavengers additives in solutions (Aurbach et al. 2007).

2.1.2. Ageing on cycling.

The ageing of a lithium-ion battery is typically defined and demonstrated as a result of the capacity degradation of active materials. Specifically, the positive material originated from phase transformation while lithium insertion. The fading rate mechanism decreases typically with cycle number progression, resulting in equilibrium with the time of utilization (Broussely et al. 2005). Occasionally, the fading rate may demonstrate an inflexion point, which is why increasing capacity loss instead of decreasing the phenomenon. Apparently, this effect is not common to this packaging design but a weak mechanically soft case may cause degradation (Kanamura et al. 2005). In terms of anode materials, Carbon is dominating among other alternatives for lithium-ion battery (Vetter et al. 2005b). Thus, most of the literature is based on the graphite carbons and it’s been noted that most of the cases presented in literature have an independent system. Changes are mostly influenced in the electrode and electrolyte interface over the cycle are noted by the researchers and detail studies are been conducted to investigate the implication of battery performance. The electrochemical stability depends on the voltage which operates at the lithium-ion battery anode. The study found that the shape of the carbon particle has a vital role as an active material for Lithium-ion batteries (Aurbach et al. 2002). The capacity degradation or fading also happens due to other associated mechanisms, associated with undesirable electrode and electrolyte reactions that are the product of the battery. These effects are most noticeable while overcharge or over-discharge, resulting in electrolyte decomposition, active material dissolution, and passive film formation (Arorat et al. 1998). Johnson and White (1998) have shown that the capacity degradation of lithium-ion battery can be fade away by 10-40% over the first 450 cycles, and it will then continue to degrade at a proportionate rate for the rest of the remaining lifetime.

2.2. Data-Driven Models.

Data-driven methods are often used for health monitoring and capacity prediction of a lithium-ion battery. These approaches are gaining interest in academia and industry because of its flexibility and novelty of work (Xiaosong Hu et al. 2016). This literature includes techniques requiring battery ageing history, effective on the dataset's data quality and quantity. Diverse algorithms are already developed through researchers which is a notable contribution for any field. Firstly, the dependency between monitoring variables such as SOH and electrical, thermal and mechanical effects the performance of the battery. Mathematical analysis proved as a useful tool that employs voltage data, terminal temperature and strain for battery loading states. Secondly, data collected under designed experimental conditions that fit a large number of datasets have become another widespread method due to its overwhelming computational efficiency and higher accuracy. Lastly, machine learning and deep learning algorithms are most favorable among data-driven techniques because of their flexibility and nonlinear relation problem-solving ability both for health estimation and prediction. Data-driven approaches are becoming a leading tactic for estimating battery health and predicting performance measures since they do not require any complex physical experimentation.

2.2.1. Analytical models with data fitting

Analytical model approaches typically employ a mathematical approach where battery ageing condition, service time, and cycle numbers are normally used as parameters. The empirical analytical model captures the correlation between the ageing stress factors (Y. Li et al. 2019) and the battery state of health factors to obtain the relationship of battery performance. These models are built by interpolation and fitting the dataset via a specified experimental dataset. For estimating a lithium-ion battery lifetime, an all-inclusive ageing investigation covers a wide range of battery operation conditions, including behavior under load. The existing studies mostly include battery cyclic and calendar ageing autonomously, combining the prediction model under dynamic load profile (Sarasketa-Zabala et al. 2016). The experiment regarding lithium-ion battery is designed under specific conditions that corroborate the exploration of influences such as temperature, SOC, charge and discharge phase terminal current, and voltage. The battery capacity loss is measured as a function of time, cycle number, or capacity throughput (Sarasketa-Zabala et al. 2016). Capacity throughput denotes the amount of energy delivered from the designed electrode to another during the cycling period. The choice of fitting the dataset into an equation depends upon the estimated capacity degradation. However, the model also helps to predict the battery life under several cycling conditions (Chonde et al. 2013). Sarasketa-Zabala et al. (2015) investigated cycling ageing performance upon constable operations conditions through an LFP-based lithium-ion cell performance model using DOD and C-rate stress factors. The prediction model under dynamic cycling conditions achieved RMSE 1.75% (P. Liu et al. 2010). Another study by Stroe et al. (2014) demonstrated a three-stage methodology for accelerated lifetime testing of the lithium-ion battery. The parameter obtained from the experiment were used for developing the performance degradation model, capable of predicting both the capacity and discharged power from the battery. Several other studies are also available that perform dynamic state estimation problems. The analytical variants of Bayesian filters, Kalman filter (KF) and particle filter and their variants, provide a general framework used to estimate lithium-ion battery performance measures under dynamic loading (Särkkä 2013). The Bayesian interface uses the observations to measure and update variables with the usage of the probability density function. The filter depends on the system's dynamic behavior and shape of data noise distribution (An et al. 2013). Linear system also incorporates Gaussian noise and KF provides better results relative to other models. Burgess developed a linear capacity degradation model associated with KF to measure the remaining lifetime. However, it is being observed the capacity degradation process of lithium-ion battery demonstrates non-linear correlation. For solving this issue, variants of KF have been proposed to resolve this issue (Goebel et al. 2008). In recent times in prognostic studies, the lasso technique has become one of the trusted approaches since the method improves the quality of forecasting by shrinking the coefficients and comparing the resultant forecasting result between the model fitting and unpenalized maximum likelihood. Musoro et al. (2014) proposed a discriminative model performance of shrinkage approach lasso model to predict chronic respiratory disease 6 months ahead.

2.2.2. Machine learning methods

Machine learning algorithms can be employed for health estimation and lifetime prediction. There is a significant dissimilarity between input features and the anticipated yield from the model. The features that used as input for health estimation are extracted from the battery performance measures while operating under different conditions at a given time. However, the machine learning algorithms require the estimated performance measures such as capacity values as inputs to the model and the model delivers the prediction of the remaining lifetime or cycle. Machine learning algorithms could be supervised and unsupervised. While supervised ML approaches can be either probabilistic model or non-probabilistic model. The results are defined through specified known relationships between different states underlying probability distributions. Some frequent models are the artificial neural network, SVN, elastic net, etc. (Y. Li et al. 2019). However, the sole aspect of predicting lifetime or cycle is diagnostics and the prediction's uncertainty level. Thus, recent time is seeing progressive development of Gaussian Process regression (GPR) and relevance vector machine (RVM), derived from the Bayesian model framework which are gaining increased attention apparently.

2.2.2.1. Autoregressive based models

Autoregressive (AR) models involve with time series data that typically utilizes a linear correlation from preceding time points as an input to forecast succeeding time points. The AR model comprises more straightforward computation and easier to extract the features among other models. Long et al. (2013) build an improved optimum model for lithium-ion battery capacity degradation prediction that a particle swarm approach was used to change the model order adaptively. However, the AR model demonstrates opposite behavior to the capacity degradation progress. The model is linear while the degrading process carries nonlinear patterns, the opposite behavior will result in model under-fitting while forecasting long-term performance measures (Xing et al. 2013). Solving the issue proposed an integrated moving average (ARIMA) framework combined with the AR and moving average models. The moving average model uses historical errors in the regression model. For example, Zhou et al. (2016) proposed combining ARIMA and decomposition model to improve the forecasting accuracy.

2.2.2.2. Neural network

The neural network comprises multiple nodes and layers, a straightforward mimic model of the human intelligence. It involves complex system modeling which adopts a black box, including variants of performance measure as a variable. However, neural network modeling has become the most widely adopted complex system design due to its simplicity and adaptability in handling statistical data structure and forms. Typically, the neural network compromises of 3 layers. As instances, an input layer, a hidden layer, and an output layer. The structural network connecting each node and pair is defined as weights, involving mapping functions from location to location between nodes and layers. JD Kozlowski (2003) developed an ANN model using variables that are involved in battery core such as charge transfer resistance and electrolyte resistance to measure the state of charge. Patillon et al. (1999) adopted two neural networks to predict the time or cycle while the specified variable levels designed for voltage and current would be reached at the battery terminal. The first NN model predicted remaining capacity while the second model was used to assign weights on the initial model to adjust for manufacturing design variation, ageing effects, and usage pattern under dynamic battery loading conditions. The input of the model includes initial voltage, cycle number and discharge period. Finally, the ANN model requires a large amount of input datasets for training and verification. The model accuracy depends upon the training method and data structure. Moreover, while the computational price is a hindrance of lifetime prediction, ANN playing an essential role in the current research world (X Hu et al. 2015.).

2.2.2.3. Functional principal component analysis

Functional principal component analysis (FPCA) is based on sequential data analysis that uses the mean function and covariance separately. The mean function could be measured using a smoothing function while assuming a specified mean function representing a degradation mechanism. Eigenvalues and eigenfunctions are employed to denote the covariance function based upon the Karhunen-Loéve decomposition (J.-L. Wang et al. 2016). The FPCA can extract the features that can explain the variability contained within the original dataset. It is one of the prime reasons behind using this statistical approach extensively among many other research fields where sequential data is concerned and plays a crucial role in seeking evolutional rules for forecasting modeling (Donà et al. 2009; Yao and Lee 2006). In the meantime, FPCA has been proven a useful tool while dealing with a dataset comprised of continuous functions in several number of studies in recent times (Kipp et al. 2011). In the case of a lithium-ion battery, several research studies are found where they have considered FPCA to develop the model for health estimation and battery prognostic. Cheng et al. (2015) studied and developed a novel prediction approach applied to the FPCA and Bayesian approach for predicting lithium-ion battery residual lifetime. Y. Hu et al. (2009) also proposed a method for prognostics of the lithium-ion battery applied to FPCA while using voltage and current profile for charge and discharge profile. The eigenfunctions they analyzed able to capture 99.9% data variation (Guo and Li 2017). These studies provided a rudimentary understanding to explore FPCA for developing a lithium-ion battery prognostic model using other ageing stress factors and performance measures since the dataset is a continuous function.

2.3. Research scopes.

A distinct research scope can be drawn for a data-driven prognostic model of lithium-ion battery based on literature reviews and past studies. The following Table 2.1 demonstrates the classification of the research scope

Table 2. 1 : Literature of data-driven models' classification

	Data-driven models										
Authors	Bayesian regression	Gaussian process	Kalman filter	Particle filter	Particle swarm optimization	Autoregressive based	Neural network	Support vector machine	Relevance vector machine	Functional Principal Component Analysis	Lasso Regression
X. Hu et al. (2015)	✓										
M.A. Patil et al. (2015)		✓						✓			
J. Zhou et al. (2012)			✓					✓			
S. Theodoridis (2012)	✓							✓			
B. Long et al. (2013)					✓	✓					
K. Liu et al (2018)			✓				✓				
M Kirk (2014)	✓						✓				
A. Nuhic et al. (2013)				✓				✓			
T. Qin et al. (2015)					✓			✓			
Q. Zhao et al. (2018)								✓	✓		
R.R. Richardson et al. (2017)	✓	✓									
Wang et al. (2013)				✓					✓		
Yujie Cheng et al. (2015)	✓									✓	
Jian Guo and Zhaojun Li (2017)										✓	

As it can be seen from the table as mentioned above that the majority of the work regarding li-ion battery have used SVM, Bayesian regression, and ML algorithms to develop an empirical model. In this research, we will take a novel approach, adopting functional principal component analysis to extract the feature from the dataset existing in the form of contentious function and use lasso regression, a penalized regression model to fit better models by shrinking the model coefficients. The combination of FPCA and lasso regression for analyzing functional monitoring data have never been used in any studies and remains untouched to date. The input parameter is also a crucial aspect of choosing FPCA over other models. The following Table 2.2 shows the input parameters of preceding studies of lithium-ion battery employing FPCA for monitoring and extracting features.

Table 2. 2: Input parameters of preceding studies involved with FPCA.

Authors	Input Parameters						
	Charge		Discharge				Capacity Degradation
	Current	Voltage	Current	Voltage	Temperature	Time	
Jian Guo and Zhaojun Li (2017)	✓	✓	✓	✓		✓	
Yujie Cheng et al. (2015)						✓	✓
Scope of this research			✓	✓	✓	✓	

As we are seeing from above mentioned Table that the influence of current, voltage and temperature for degrading capacity in preceding FPCA model is absent, we intended to study further and investigate the effects of those ageing stress factors applying FPCA to the capacity of lithium-ion battery. This thesis aims to explore the combination of current, voltage and temperature as input parameter and develop a data-driven lithium-ion battery model to predict the capacity applied to FPCA using lasso regression.

Chapter 3

Methodology

3.1. Overview of the problem and data analysis

Over the years, several studies have been conducted to develop prediction algorithms for the lithium-ion battery’s capacity. Mostly model-based and data-driven approaches are the leading field of studies. The lithium-ion battery has several performance measures and parameters as ageing stress factors that could be used to monitor data. These monitoring data includes voltage, current, and temperature. A standard lithium-ion battery may perform around 300-500 cycles before reaching its full discharge criteria (Geoff S. Fein,2003) . The monitoring data can be obtained as a continuous function of capacity throughout the cycle. The lithium-ion battery's capacity degrades after each cycle significantly and the battery needs to be replaced while it reaches to end-of-life criteria. In this study, we aim to predict the capacity of lithium-ion batteries using a data-driven prognostics algorithm supporting uncertainty representation and management. Specifically, we adopt FPCA applied to temperature, voltage, and current observations collected from NASA Ames Prognostics Center of Excellence repository.

3.2. Data collection

The model requires a real dataset of lithium-ion batteries to predict the remaining useful lifetime of a battery. The data used in this study to make a prognostic model is obtained from the public data repository of NASA Ames Prognostic Center of Excellence (PCoE). Three separate batteries especially 5, 6, and 7 were selected to validate the accuracy of the proposed model. The data set has been taken from a prognostic test experiment in NASA which includes commercial Lithium-ion 1850 sized battery, thermocouple sensors, power supply, load, etc. to record all the performance measures (B. Saha and K. Goebel 2007). The lithium-ion battery experiment was extended through three different operational profiles such as charge, discharge and impedance at ambient temperature. The charging procedure was follow-through in constant current (CC) mode at 1.5A. While the battery voltage touched at 4.2V mark, charging was continued with constant voltage (CV) until the charging current fell below 20mA. The discharge process was carried out with constant current (CC) up to 2A until the voltage level touched the specified end of discharge level 2.7V, 2.5V, and 2.2V for batteries 5,6 and 7 respectively. While the capacity degraded to the specified threshold called end of life (EOL) criteria which was a 30% fade in rated capacity, the experiment was stopped. Figure 3.1 demonstrates the data structure including performance monitoring variables.

	Cycle	Discharge Voltage	Discharge Current	Discharge Temperature	Charge Current	Charge Voltage	Time	Capacity
1	1	4.1994	-0.0018659	23.937	-0.0004	0	0	1.8911
2	1	4.1995	-0.0021394	23.924	-0.0004	4.215	16.781	1.8911
3	1	3.9856	-1.9888	24.004	-2	3.003	35.703	1.8911
4	1	3.9632	-1.9926	24.163	-2	2.987	53.781	1.8911
5	1	3.9466	-1.9885	24.346	-2	2.972	71.922	1.8911
6	1	3.9328	-1.9912	24.546	-2	2.963	90.094	1.8911
7	1	3.921	-1.9895	24.741	-2	2.957	108.28	1.8911
8	1	3.9109	-1.9892	24.937	-2	2.955	126.45	1.8911
9	1	3.9017	-1.9895	25.128	-2	2.951	144.64	1.8911
10	1	3.8932	-1.9896	25.313	-2	2.946	162.84	1.8911
11	1	3.8857	-1.9903	25.511	-2	2.94	181.02	1.8911
12	1	3.8784	-1.9888	25.695	-2	2.934	199.22	1.8911
13	1	3.8715	-1.9912	25.887	-2	2.929	217.39	1.8911
14	1	3.865	-1.9904	26.071	-2	2.923	235.59	1.8911
15	1	3.859	-1.9886	26.247	-2	2.919	253.75	1.8911
16	1	3.853	-1.9915	26.424	-2	2.913	271.98	1.8911
17	1	3.8471	-1.9892	26.596	-2	2.909	290.14	1.8911
18	1	3.8415	-1.9917	26.775	-2	2.903	308.36	1.8911
19	1	3.8361	-1.9913	26.932	-2	2.899	326.5	1.8911
20	1	3.831	-1.9896	27.098	-2	2.894	344.75	1.8911
21	1	3.8257	-1.9906	27.255	-2	2.889	362.91	1.8911
22	1	3.821	-1.9891	27.406	-2	2.885	381.05	1.8911
23	1	3.8161	-1.9925	27.557	-2	2.88	399.19	1.8911
24	1	3.8114	-1.9898	27.726	-2	2.875	417.28	1.8911
25	1	3.8068	-1.9889	27.875	-2	2.871	435.48	1.8911

Figure 3. 1: Data structure of dataset demonstrates initial 25 sequential data of first cycle.

3.2. Model building and testing

Suppose we have a battery that has been used for N charge-discharge cycles. We define the training data as $D_{train} = (X_{train}, Y_{train})$ where $X_{train} = \{v_{ij}, c_{ij}, t_{ij}, i = 1, \dots, N, j = 1, \dots, m_i\}$, which includes the monitoring measurements of voltage, current, and temperature for the first N cycles at m_i time points, and $Y_{train} = \{y_i, i = 1, \dots, N\}$, the battery capacity data for the corresponding cycles, as the target. Based on the training data, we aim to predict $D_{test} = (Y_{test}) = \{y_i, i = N + 1, \dots, N + P\}$ the battery capacity for forthcoming P cycles using simple linear regression. Note that the testing data does not include X_{test} as the monitoring observations are not available when the prediction is made in real applications. We instead attempt to predict $X_{test} = \{v_{ij}, c_{ij}, t_{ij}, i = N + 1, \dots, N +$

$P, j = 1, \ldots, m_i\}$ based on X_{train}, and use the predicted monitoring measurements as the input of the prediction model. Figure 3.2 depicts the graphical representation of the process for building and testing the capacity prediction model with the help of battery monitoring data.

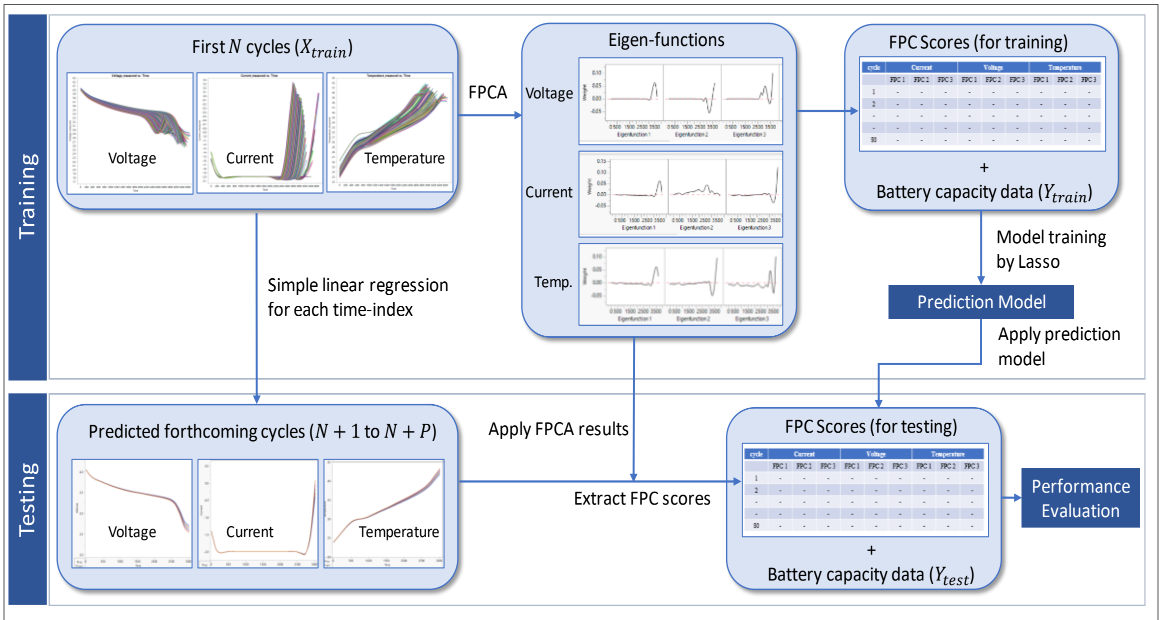


Figure 3. 2: Graphical representation of workflow

For the model training, the FPC scores are extracted from the first N cycles of voltage, current, and temperature measurements. These FPC scores are utilized to predict the battery capacity via the least absolute shrinkage and selection operator (lasso) which is a penalized regression procedure to fit the model by shrinking the coefficient into zero that results in a biased prediction outcome with low prediction variance and enhanced prediction accuracy. For the model testing, we use the simple linear regression for each time index to predict forthcoming P cycles of monitoring variables based upon preceding discharge cycles. Subsequently, FPCA is employed to extract the FPC scores from each predicted voltage, current, and temperature cycle. The trained regression model is applied to predict the lithium-ion battery's remaining capacity for cycles $\mathrm{N}+1$ to $\mathrm{N}+\mathrm{P}$. The result from the proposed model provides notable accuracy throughout different cycling conditions over the life span of a lithium-ion battery. The detailed model building and testing process in step by step as follows -

(a). The monitoring data of the lithium-ion battery comprises of 168 cycles of voltage, current, and temperature. To begin the model, we have considered the initial $N = 100$ cycles as training data to build a prediction model.

(b). The original discretized measurements are transformed to smooth curves by applying the B-spline basis expansion. The number of knots $K = 5$ is chosen, for each of the voltage, current, and temperature, such that the fitted model has the lowest BIC value.

(c). The FPCA is performed upon each monitoring variable, and the mean function μ(t), FPC scores $\xi_{ik}, k = 1, \dots, K$ and corresponding eigenfunctions are obtained for each predictor. These FPC scores characterize the status of the battery at the corresponding cycle.

(d). The FPCA provided $\mathrm{n} = 3, 3$ and 5 types of FPC scores for voltage, current and temperature respectively. These FPC scores are used as input feature for the lasso regression.

(e). The lasso regression model is trained based on FPC scores. The lasso complexity parameter is chosen by the K-fold cross-validation with 5 folds.

(f). To test the model, the monitoring variables' measurements for forthcoming $P = 20$ cycles are predicted. The simple linear regression was used for this task for each time point. For the voltage measurements for a time point of $j = 1$, for example, we fit the following model to the initial 100 cycles of data.

$$v_{i1} = \beta_0 + \beta_1 i + \varepsilon_{i1}, \qquad i = 1, \dots, 100 \tag{1}$$

Then, the fitted model is extrapolated to produce $\hat{v}_{i1}, i = 101, \dots, 120$. This process is repeated for each time point of $j = 1, \dots, m_i$. Merging those resulting responses, we got predicted 20 cycles which are almost indistinguishable from actual cycles. The B-spline with the number of knots learned by the training data is applied to these predicted measurements.

(g). Using the eigenfunctions obtained from the training phase, the FPC scores are extracted from each predicted curve.

(h). The prediction of $\hat{y}_i, i = 101, \dots, 120$ are obtained by lasso model obtained in (f), and compared with the true battery capacity values to evaluate the model performance.

3.3. Functional Principal Component Analysis:

The FPCA is an approach used to reduce the dimensionality of large datasets while helps to enhance interpretability, yet loses minimum information from the data (I. T. Jolliffe 1986; Ian T. Jolliffe and Cadima 2016). Herein, we briefly introduce the main idea of FPCA with an example of voltage measurement data v_{ij}. The FPCA can be thought of as an extended version of the traditional multivariate principal component analysis (PCA) with the infinite-dimensional vectors, i.e., the function. To apply the FPCA, the data has to be given as a form of a smooth function. As the voltage data is discrete measurements over time, we first attempt to represent v_{ij} as a function $v_i(t)$ by using a B-spline basis expansion with an appropriate number of knots.

The FPCA aims to find weight functions $\phi(t)$ which mostly explain the function-to-function variabilities. The first functional principal component $\phi_1(t)$ is chosen to maximize the mean square $\sum_{i=1}^{N} \xi_{i1}^2 / N$ where

$$\xi_{i1} = \int \phi_1(t) v_i(t) dt \tag{2}$$

with the constraint of the unit squared norm $\int \phi_1(t)^2 dt = \|\phi_1\|^2 = 1$. The second and subsequent FPCs can be also found by solving the same optimization problem with the additional orthogonality constraints of $\int \phi_k(t)\phi_m(t)dt = 0, k < m$. Let us define the covariance function $G(s,t) = \sum_{i=1}^{N} v_i(s)v_i(t) / (N-1)$, then it can be shown that the above optimization problem is reduced to the following eigen-equation.

$$\int G(s,t)\phi(t)dt = \lambda\phi(s) \tag{3}$$

where ϕ is an eigenfunction and λ is an eigenvalue. This continuous functional eigen-analysis problem can be solved by an approximately equivalent matrix eigen-analysis task. For more details, see (James O. Ramsay and B. W. Silverman 1997) chapter 8.4.

It can be also shown that each curve of $v_i(t), i = 1, \ldots, N$, is approximated by the expansion in terms of a small number of orthonormal basis functions ϕ's by the following form.

$$v_i(t) \cong \mu(t) + \sum_{k=1}^{K} \xi_{ik}\, \phi_k(t) \tag{4}$$

where $\mu(t) = \sum_{i=1}^{N} v_i(t)\,/N$ is the mean function, $\xi_{ik} = \int \phi_k(t) v_i(t) dt$ is the k-th functional principal component (FPC) score, and $\phi_k(t)$ is the k-th eigenfunction.

The same procedure can be conducted to the current and temperature curves. In our research, we extracted five FPC scores from each curve of voltage, current, and temperature. These 15 FPC scores for each cycle are used as the input features of the battery capacity prediction model.

3.4. Lasso Regression:

The lasso regression has become widely used to estimate parameters in regression problems with a large number of covariates. It is extensively employed due to shrinking the vector of regression coefficients toward zero, setting some coefficients equivalently equal to zero, which ensures simultaneous and variable estimation procedure (Tibshirani 1996) . The lasso is a form of penalized least squares that minimizes the residual sum of squares while controlling the $\mathrm{L}-1$-norm of the coefficient vector β. For a generalized linear regression model coefficient β_0 and $\beta = (\beta_1, \beta_{2,} \dots \dots, \beta_p)$ are measured by using following equation-

$$\left(\widehat{\beta_0}, \widehat{\beta^{lasso}}\right) = argmin\,\{\sum_{i=1}^{N} \left(Y_i - (\beta_0 + \beta X_i^T)\right)^2 + \lambda \sum_{j=1}^{p} |\beta_j|\} \tag{5}$$

Where Y and X are the capacity of lithium-ion battery and FPC scores respectively. λ is a non-negative parameter or penalty that regulate the amount of shrinkage that needed to incorporate in the model. The penalty is the form of regularization, a method to solve an ill-poised problem or prevent the overfitting data into the model.

Chapter 4

Results and Analysis

This section discusses the application of the mathematical models to a real-life problem regarding the capacity prediction of lithium-ion battery applied employing functional performance measures such as voltage, current and temperature. The model identifies the features of functional data that have been used to predict the capacity using lasso regression. The model also compares the predicted capacity with the actual capacity through lasso regression. The following sections comprises of detailed discussion regarding the model outcomes.

4.1. Data visualization

The data of lithium-ion battery was received from a public data repository of NASA Ames Prognostic Center of Excellence (PCoE). We were concerned with the actual data structure of performance measure that obtained while experimenting in the laboratory under specified conditions. Since the objective is to observe the capacity degradation, the discharge phase of the battery has the main responsibility. We wanted to realize and understand the data structure of performance measures such as voltage, current and temperature among other parameters. Figure 4.1 shows the first smoothen 100 cycles of voltage rating against time. The cycles are showing a notable pattern from the initial cycle to the end. The voltage is showing declining behavior from the beginning to end the cycle while each cycle started around 4.0 V and ends at 2.5V. Different colors are used in the plot to distinguish the cycles and its behavior.

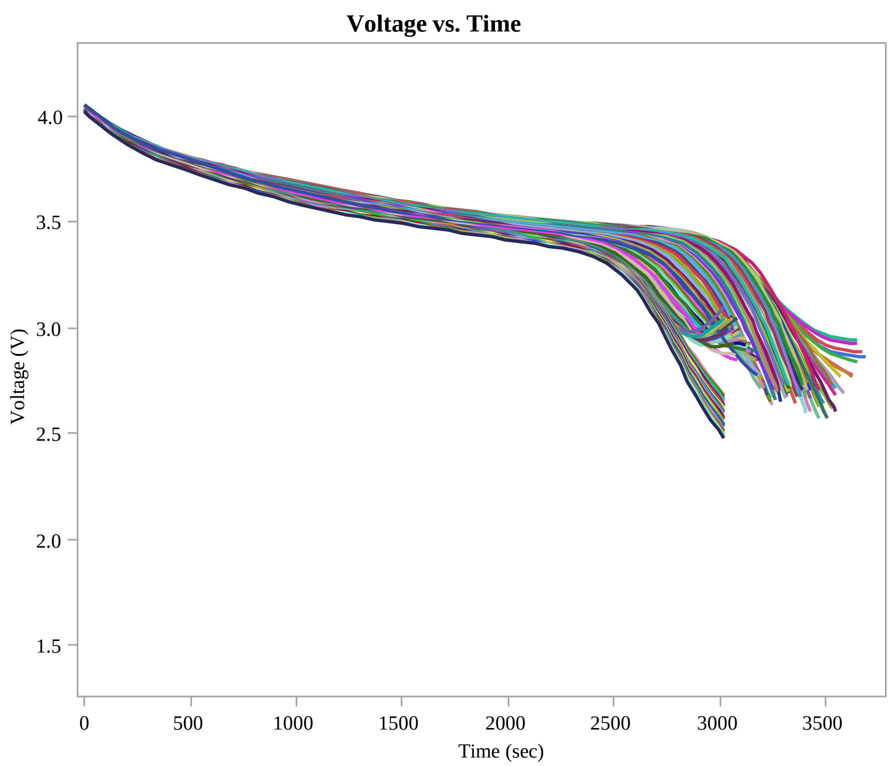


Figure 4. 1: The first 100 cycles of voltage rating against time steps.

There is other variable such as current, we wanted to observe the behavior while discharging the capacity. Figure 4.2 demonstrates the behavior of current rating as cycles vs. time points. If we see the plot, the voltage rating remains constant and kept negative rating for a while from the beginning of each cycle. However, it goes to zero voltage rating while the cycle came to an end of each cycle. The following plot is showing the smoothen curve of first 100 cycles of current rating against time.

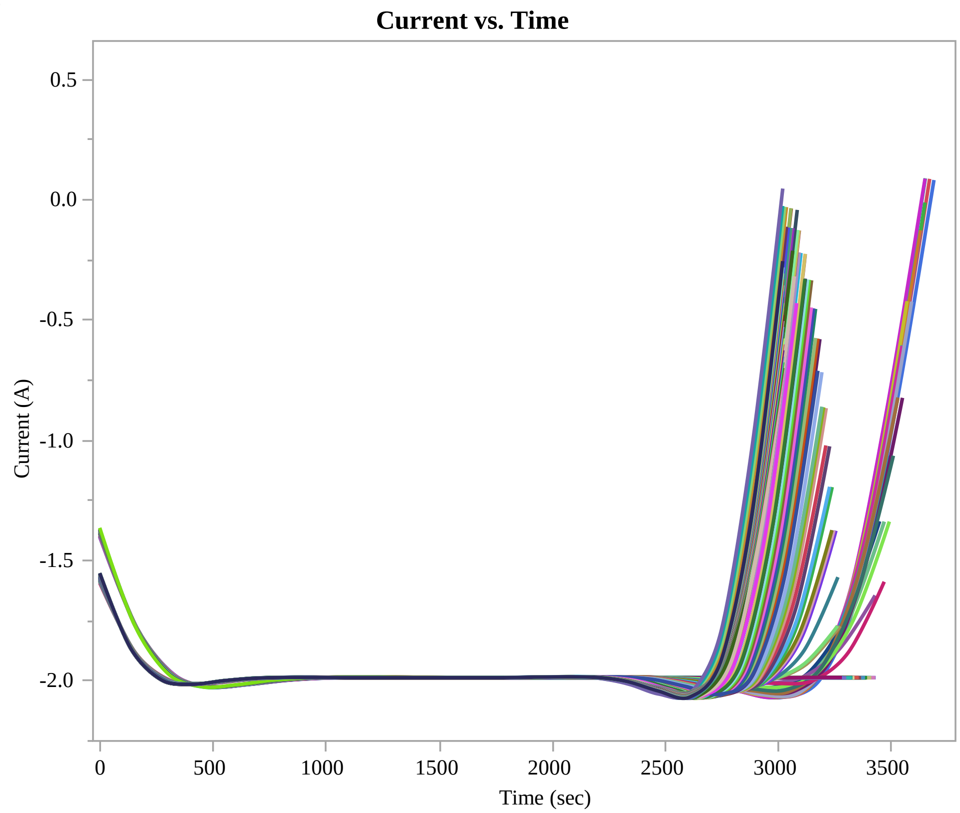


Figure 4. 2: The first 100 cycles of current rating against time steps.

The temperature is another performance measure that we wanted to investigate before the model development. The temperature is showing opposite trend to the voltage. While voltage rating was showing declining tendency, the temperature is demonstrating increasing inclination as capacity degrade over time. Figure 4.3 depicts the first 100 cycles of temperature against time period. The temperature of each cycle starts around 25˚C while ends at 45˚C. The following plot shows the smoothen curve of initial 100 cycles of temperature vs. time duration.

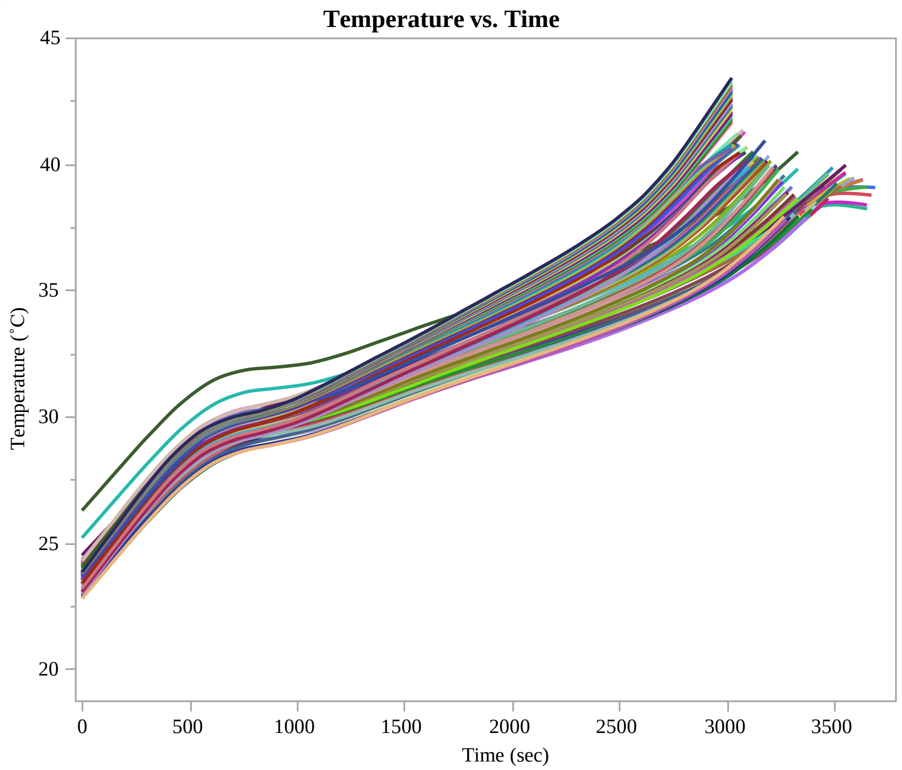


Figure 4. 3: The first 100 cycles of temperature against time steps.

The capacity degradation over time is the main focus of our research interest. The battery has specified capacity at the beginning of the experiment. However, after successive discharge cycles the capacity degrade significantly. The experiment ran 168 overall cycle before reached to its end of life criteria. The following figure 4.4 depicts the capacity degradation plot over number of cycles. Each discharge cycle reduces the capacity significantly.

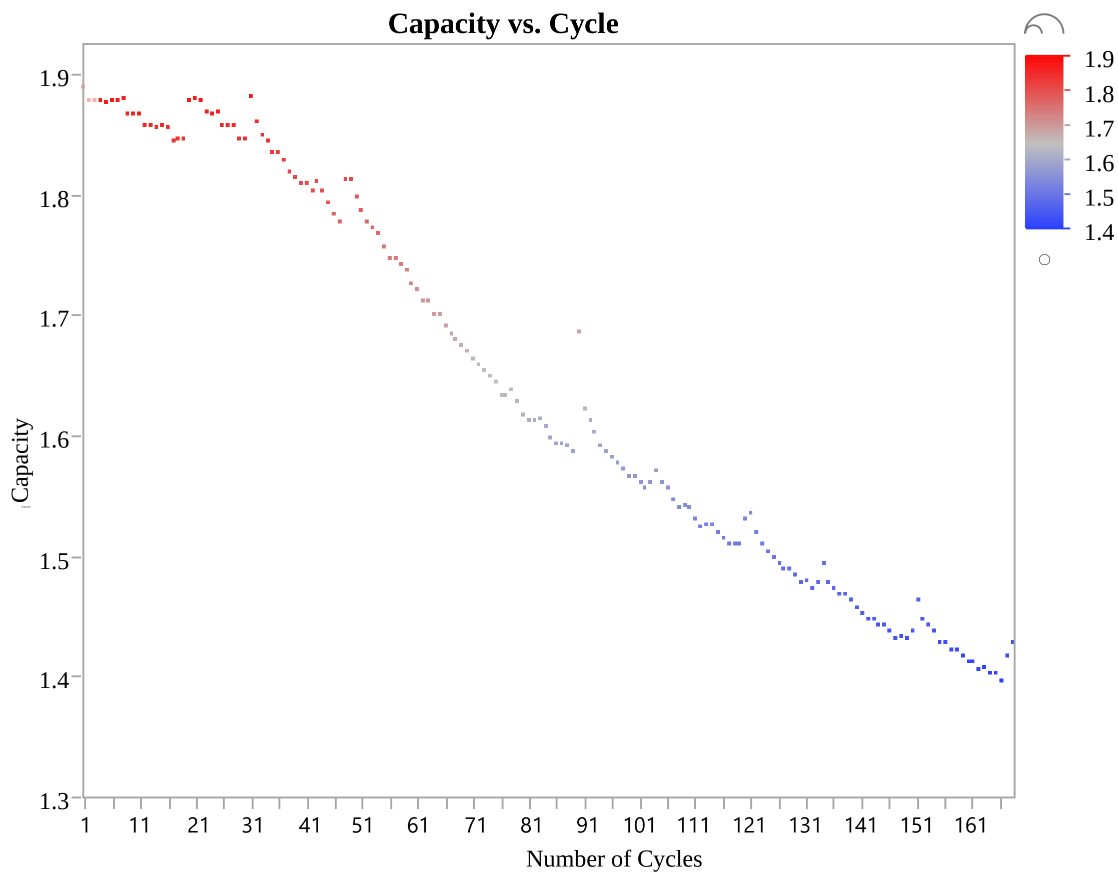


Figure 4. 4: The degradation of capacity over the number of discharge cycle of lithium-ion battery.

The capacity plot gives us an impression of degradation rate of each cycle. We want to investigate the influence of aforementioned performance measures such as voltage, current and temperature value to the capacity of lithium-ion battery. Following section will discuss the findings of the developed model of FPCA for capacity degradation.

4.2. Functional principal component analysis model

The FPCA model was employed to analyze the initial 100 cycles of data of voltage, current, and temperature respectively to provide FPC scores and eigenfunctions. Based upon the BIC value, the cubic spline was adopted to fit the model. We have set the number of knots

35 which has provided better results among other options. Figure 4.5 illustrates examples of the fitted B-spline basis expansion model for the first 50 cycles of voltage curves.

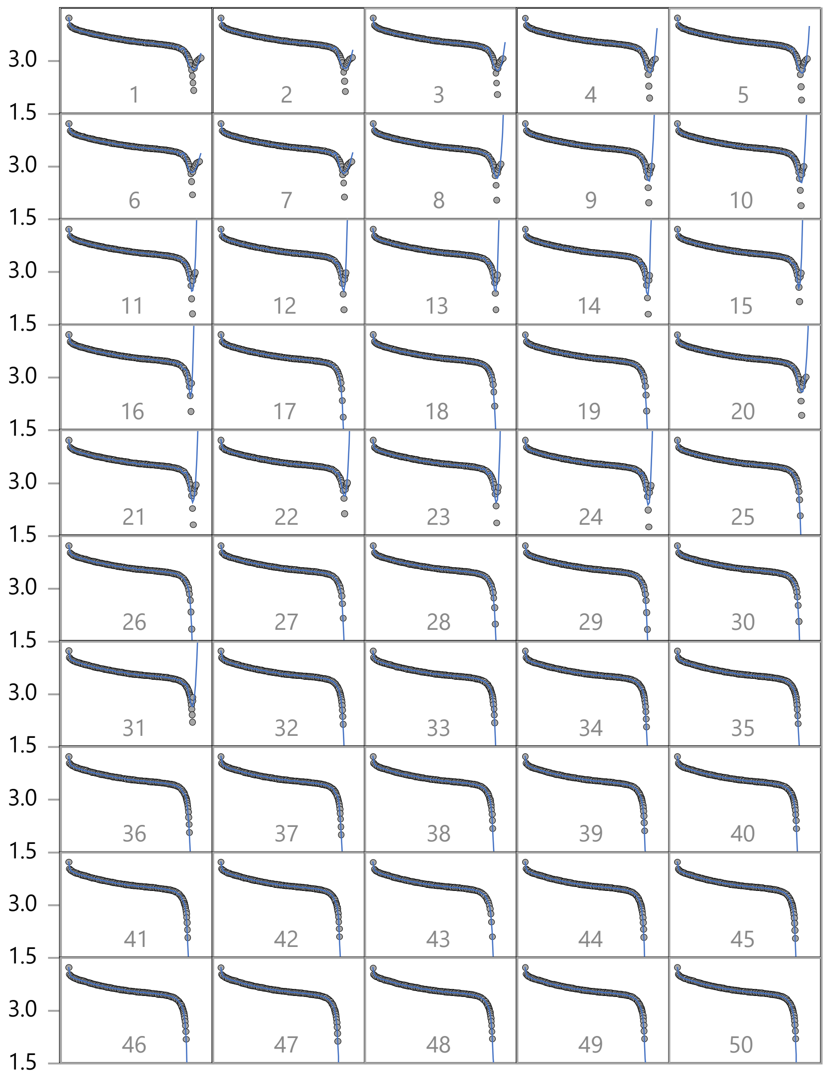


Figure 4. 5: B-spline fitted for the first 50 cycles of voltage curves.

The model diagnostic has provided the actual voltage vs. voltage predicted curve where we can validate the model fitting. Figure 4.6 demonstrates the voltage vs voltage predicted curve. Most of the data point lies around the tangent line which validate the model.

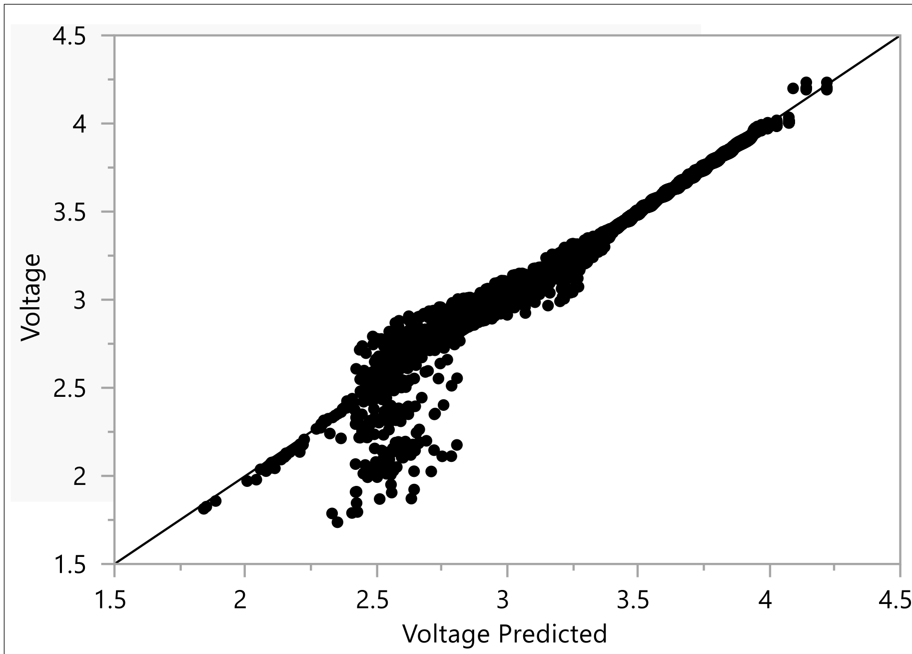


Figure 4. 6: Model fitting for actual voltage vs. voltage predicted.

The FPCA for voltage provided three separate FPC scores and eigenfunctions while the first FPC was responsible for 75.4% data variation and the second FPC was accountable for 21.1% data difference. According to the eigenvalues and their contributions towards data variation, the first FPC is the dominant function among other FPC's for the variations. Figure 4.7 illustrates the mean functions of initial 100 cycles including three eigenfunctions of voltage that we obtained using FPCA.

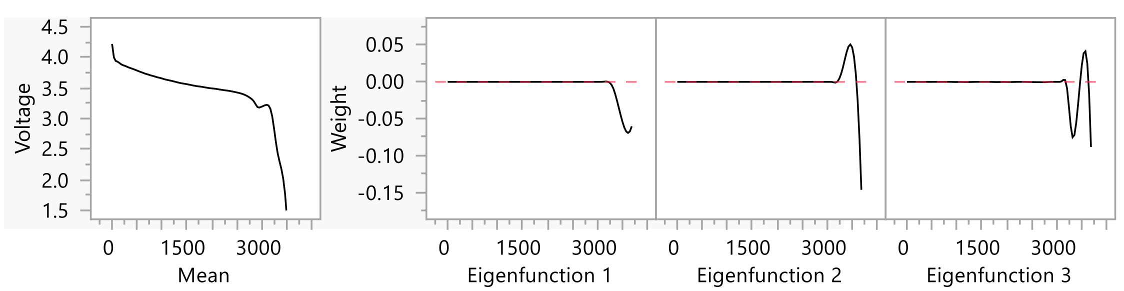


Figure 4. 7: The mean and eigenfunctions of voltage cycles.

As voltage, we have conducted the same analysis for other performance measures such as current and temperature. Figure 4.8 illustrates examples of the fitted B-spline basis expansion model for the first 50 cycles of current curves.

Figure 4. 8: B-spline fitted for the first 50 cycles of current curves.

The FPCA generates three distinct eigenfunctions and corresponding eigenvalues for the functional data of current. Regardless of the mean function, the first eigenfunction corresponding to 65% of data variation while the second eigenfunction is responsible for 30% variation. The rest of the 5% data variation is responsible for the third eigenfunction. Figure 4.9 illustrates the mean functions of initial 100 cycles including three eigenfunctions from current that we obtained using FPCA.

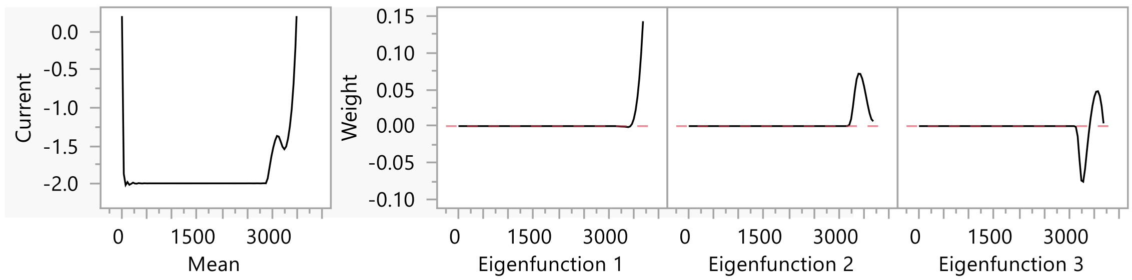


Figure 4. 9: The mean and eigenfunctions of current cycles.

Similarly, the temperature is also subjected to the B-spline fitted model which is illustrated in the following 4.10 figure. As aforementioned performance measures, we have shown 50 initial temperature cycles for comprehensible understanding. The model fitted so well that the based on the diagnostic plot, most of the points lies on the tangent line. The diagnostic plot of actual temperature against the predicted temperature plot is depicted in the subsequent plot shown in 4.11 figure.

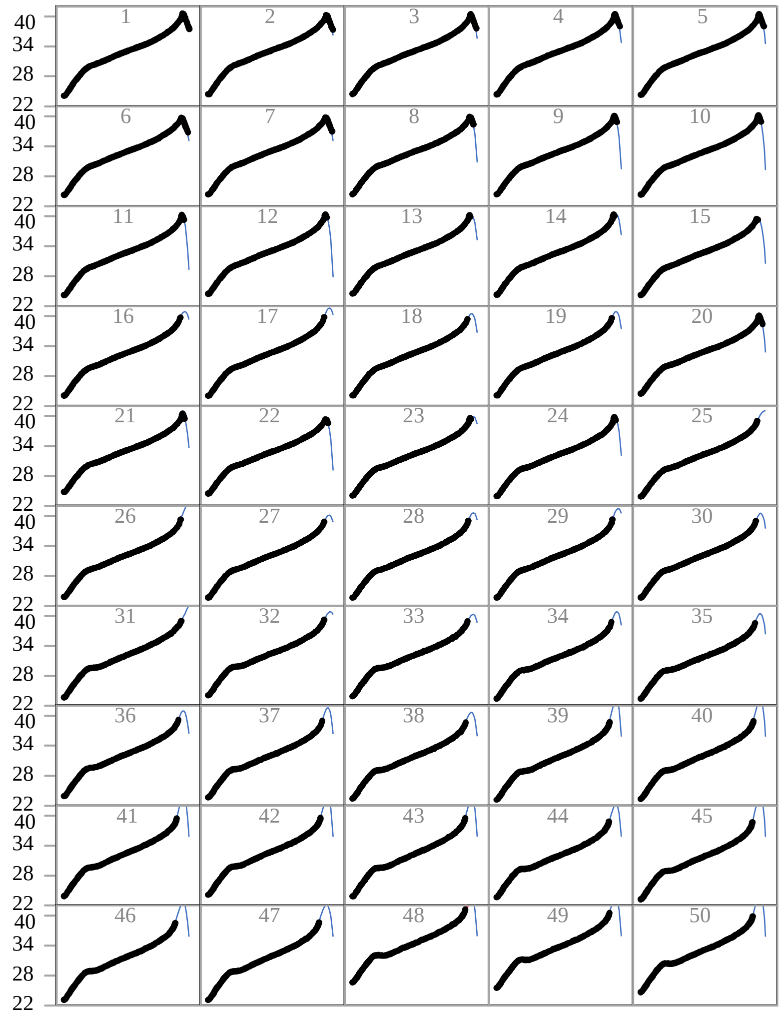


Figure 4. 10: B-spline fitted for the first 50 cycles of temperature curves.

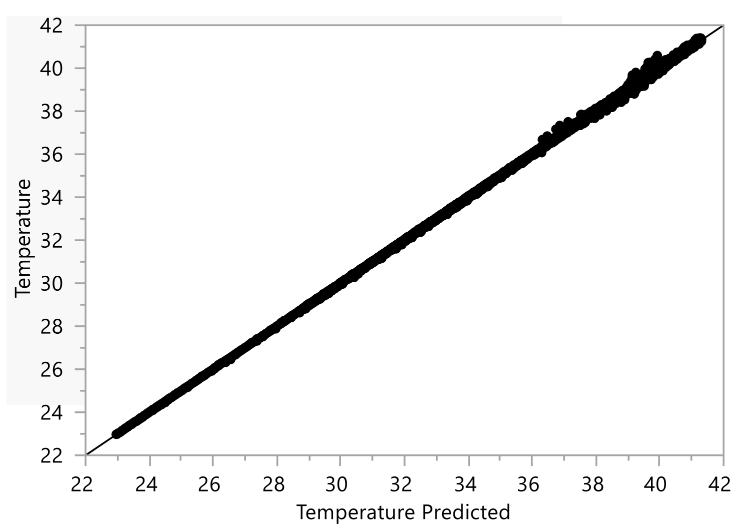


Figure 4. 11: Model fitting for actual temperature vs. temperature predicted

Unlike to the previous variables such as current and voltage, the temperature generates six separate eigenfunctions. Based upon its percentile, it is evident that no eigenfunction dominant for the existing variations which is different from other parameters. From these 6 functions, the first one contributes around 42% variation whereas the second one contributes 33% and rest of them are respectively 15.5%, 3.95%, 3.49% and 1.81%. Figure 4.12 shows the mean and eigenfunctions for the temperature.

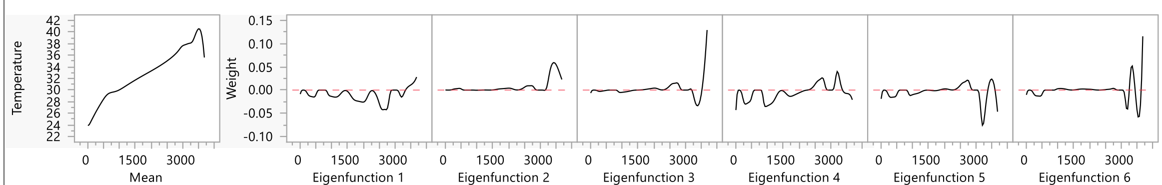


Figure 4. 12: The mean and eigenfunctions of current cycles.

Each eigenfunctions of current, voltage and temperature generate corresponding FPC scores for each cycle. We can simulate each function with the help of those corresponding FPC scores and mean function of each performance measures that are subjected to the investigation. We have extracted FPC scores for training dataset which is 100 cycles of current, voltage and temperature. Following table 4.1 demonstrates obtained FPC scores for initial 20 cycles of voltage, current and temperature with corresponding capacity of each cycle.

Table 4. 1: FPC scores for initial 20 cycles of voltage, current and temperature.

Cycle	**Capacity**	**Voltage**			**Current**			**Temperature**					
		FPC 1	FPC 2	FPC 3	FPC 1	FPC 2	FPC 3	FPC 1	FPC 2	FPC 3	FPC 4	FPC 5	FPC 6
1	1.89	-50.94	-26.7	-8.49	-55.8	-3.91	-3.88	5.49	-8.07	1.88	-10.16	-3.87	6.13
2	1.88	-51.89	-26.9	-8.42	-55.0	-5.54	-4.75	-2.35	-11.44	-4.29	-12.19	-5.00	7.31
3	1.88	-52.15	-28.0	-8.84	-54.9	-5.51	-4.73	-1.55	-12.84	-6.81	-11.81	-3.90	4.93
4	1.88	-52.89	-30.0	-9.43	-53.9	-5.07	-4.64	1.49	-15.20	-10.48	-9.85	-1.95	2.43
5	1.88	-53.08	-30.4	-9.38	-53.9	-5.31	-4.74	4.13	-16.84	-11.69	-8.30	-0.96	2.32
6	1.88	-53.39	-27.0	-8.07	-55.1	-8.00	-5.82	1.05	-22.50	-13.13	-7.19	-3.13	9.86
7	1.88	-53.23	-27.2	-7.97	-55.2	-8.19	-5.88	-0.85	-20.21	-12.45	-8.37	-3.61	9.16
8	1.88	-62.15	-38.6	-11.08	-43.4	-3.51	-5.01	-4.95	-26.60	-33.26	-5.61	1.68	0.09
9	1.87	-71.35	-48.6	-12.83	-27.1	1.75	-4.31	-5.03	-29.30	-40.69	-3.66	3.78	-2.00
10	1.87	-72.16	-50.2	-13.04	-27.3	1.79	-4.26	-2.27	-29.38	-40.79	-2.54	4.65	-3.01
11	1.87	-89.04	-67.0	-15.86	1.0	10.72	-3.12	1.13	-29.14	-40.22	-0.87	5.80	-3.46
12	1.86	-133.77	-102.6	-19.34	56.2	22.83	-3.16	-10.51	-27.81	-50.52	-2.34	4.32	-3.04
13	1.86	-164.09	-127.2	-22.58	91.8	31.94	-2.64	2.99	-9.76	-7.49	-9.92	-2.72	2.00
14	1.86	-181.38	-143.1	-24.48	92.0	31.91	-2.67	10.08	-6.04	-0.75	-8.61	-1.77	1.49
15	1.86	-207.44	-155.0	-23.77	97.0	27.09	-5.18	1.65	-31.12	-37.50	-0.16	2.86	2.75
16	1.86	-223.96	-169.5	-25.33	75.0	22.46	-5.08	21.28	-0.26	16.85	-8.95	-3.51	4.10
17	1.85	182.81	121.0	2.79	-17.5	9.19	-1.95	23.40	7.32	23.16	-9.87	-3.78	2.64
18	1.85	154.90	97.8	0.39	-12.0	9.54	-2.32	14.87	-6.08	0.20	-5.66	-1.80	3.36
19	1.85	180.47	111.4	1.31	-11.9	9.51	-2.33	16.65	0.44	5.28	-6.87	-1.80	1.89
20	1.88	-58.16	-36.7	-11.71	-44.4	-1.38	-3.96	3.58	-24.94	-20.25	-8.18	1.36	-0.65

4.3. Linear regression model

The battery capacity prediction begins with the forthcoming cycle prediction using simple linear regression. This model will predict the cycle as testing dataset for the above model where we have considered 100 cycles as training dataset. However, we have randomly chosen 20 cycles as testing dataset for the model. Linear regression model produces similar shapes of cycles as preceding 100 cycles that has been considered for the FPCA model. With the aid of these simulated cycles of current, voltage and temperature, we have extracted the FPC scores of similar to the training dataset. Following figures 4.13-4.15 demonstrates the predicted cycles of voltage, current and temperature respectively.

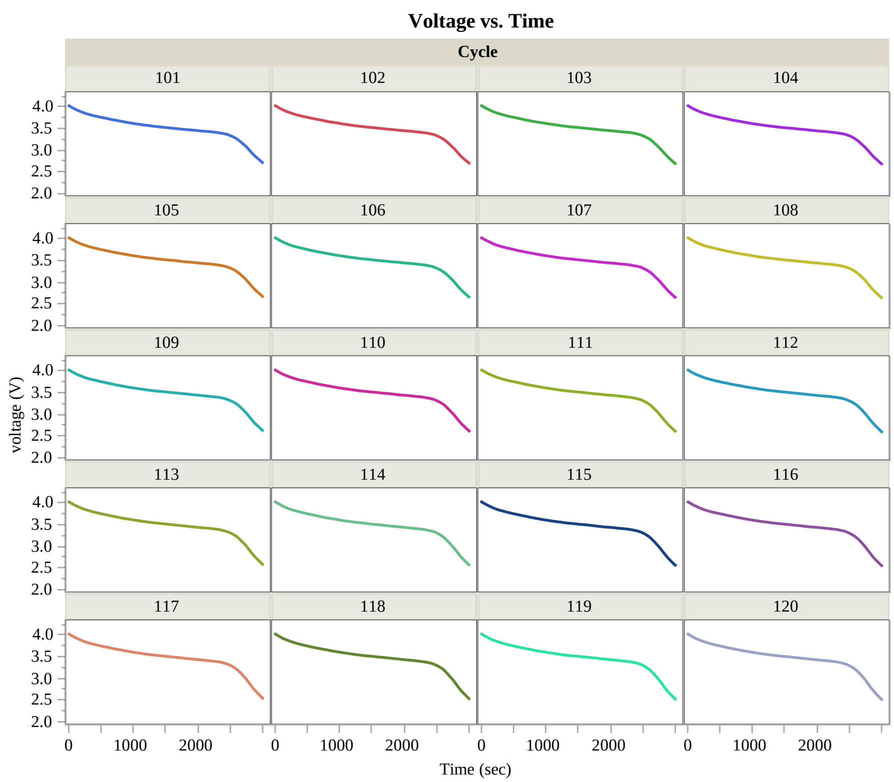


Figure 4. 13: The predicted 20 cycles of testing dataset for voltage.

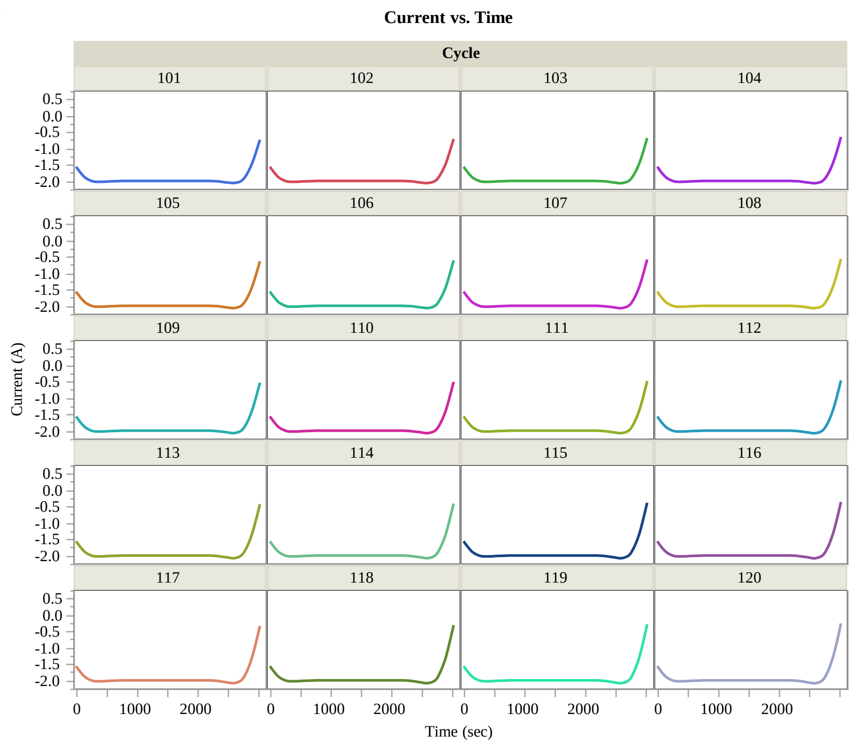


Figure 4. 14: The predicted 20 cycles of testing dataset for current.

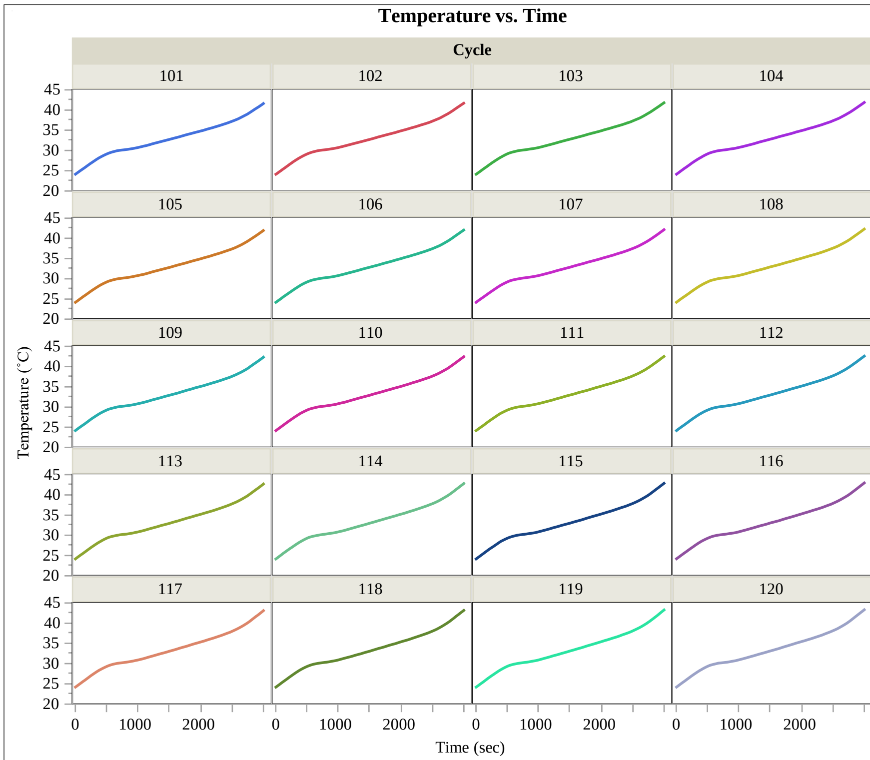
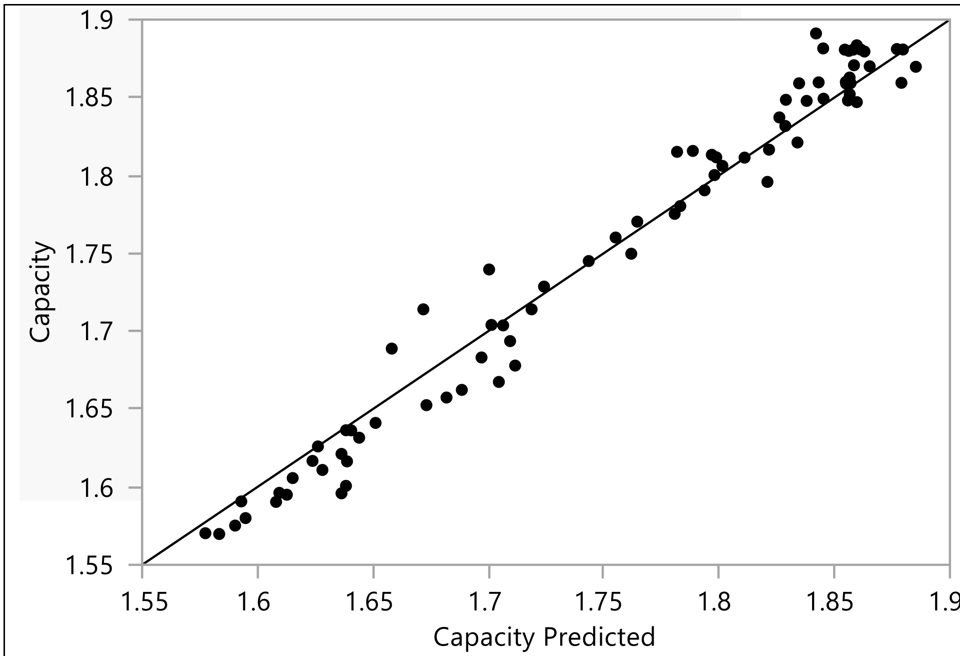


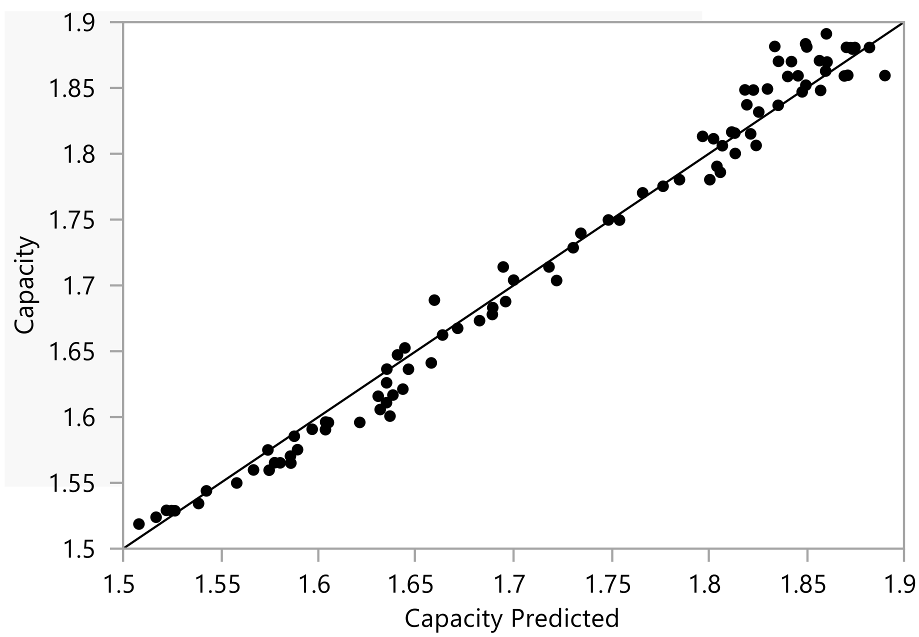
Figure 4. 15: The predicted 20 cycles of testing dataset for temperature.

4.4. Lasso regression model

The FPC scores that we obtained through the model are the input for the generalized regression model where we have used lasso regression with 5-folds cross-validation. The lasso was used to train the preceding FPC scores with their corresponding capacity and help to predict the capacity based upon the FPC scores of predicted cycles. The experiment was conducted under three different conditions. We have evaluated three cases for splitting the dataset. First case where we have examined 100 cycles for the training purpose and subsequent 20 cycles for testing the model while those cycles were predicted using linear regression. For second case, designed 120 cycles for training and next 20 as testing purpose. Finally, inspected the model with 140 cycles of performance measure as dataset and next 20 cycles for the testing the model. These number of cycles were considered randomly. As Initially, we have considered the initial 100 cycles as a training dataset while predicted 20 cycles for model testing. Figure 4.16 shows actual versus predicted capacity for training and testing dataset for two cases while (a) represents 100 cycles as training dataset and (b) demonstrates the case where 120 cycles are considered as training dataset.



(a)



(b)

Figure 4. 16: Actual versus predicted capacity (a) 100 cycles as training (b) 120 cycles as training for training and testing dataset

The predicted outcomes as capacity prediction are randomly distributed around the tangent line which corroborates the model validity and facilitates credible prediction outcomes. The plot provides evidence that the capacity prediction outcomes from the developed model appears to be closer to the actual value and the model would provide crucial result for our investigated problem. As mentioned above, the model has feds 100 cycles as an input at the beginning as training dataset and predicted the capacity for the subsequent 20 cycles which we obtained using linear regression. After a while getting the results as capacity prediction value, we have continued the process and kept doing the prediction process considering 120 cycles as training and predicted the following 20 cycles for testing purpose. Finally, considered 140 cycles of actual dataset of current, voltage and temperature as training data and predicted capacity for ending 20 cycles of lithium-ion battery. Using the process of replication with different sets of datasets, the model helped us to check the model and verified its accuracy of prediction. Figure 4.17- 4.18 demonstrates the result of the model as capacity for first two cases. The fitted capacity curve is color coated where fitted capacity and predicted capacity is apparent.

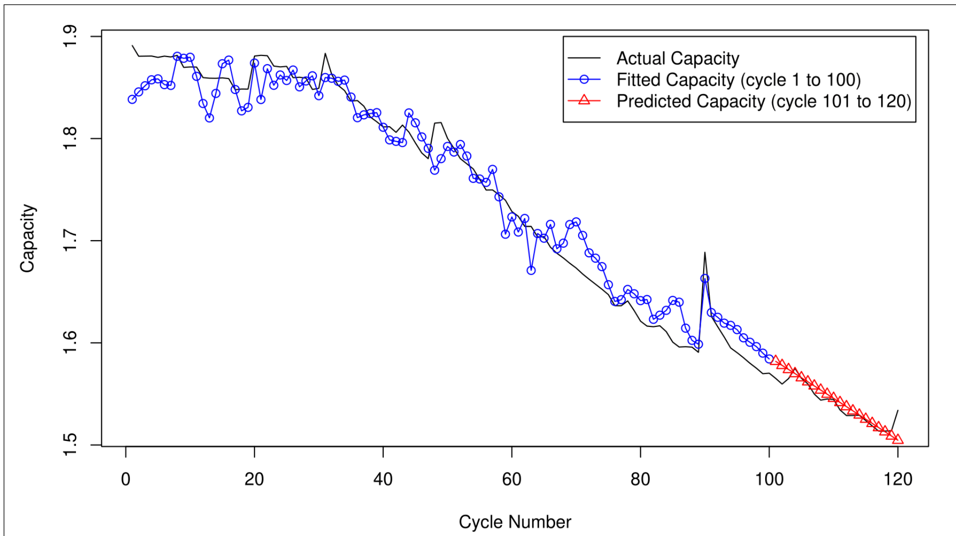


Figure 4. 17: Fitted curve for capacity prediction of initial 120 cycles

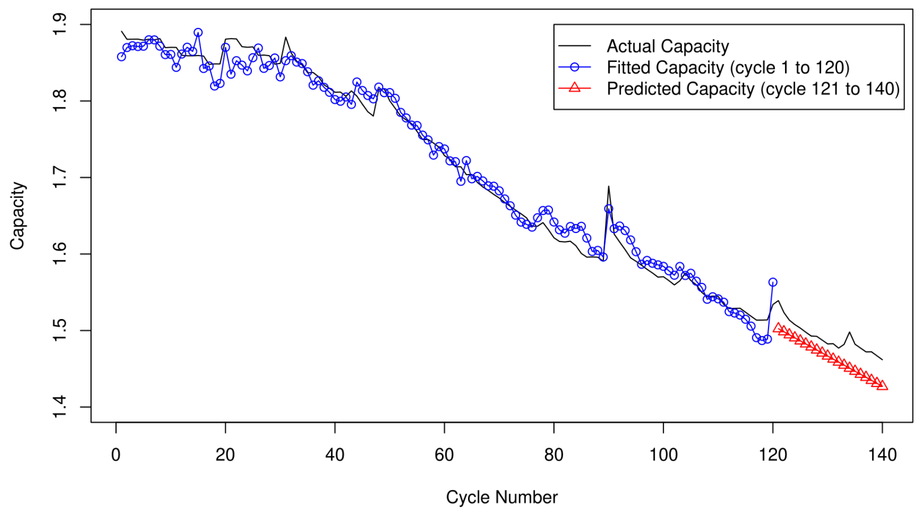


Figure 4. 18: Fitted curve for capacity prediction of initial 140 cycles.

The root-mean-square error (RMSE) and mean-absolute-percentage-error (MAPE) is used to evaluate the accuracy of the model. The RMSE and MAPE values of the capacity data and prediction error between the actual measurements and the estimated battery capacity are calculated to evaluate the accuracy of the proposed method. In order to check prediction accuracy, we have conducted both RMSE and MAPE. The errors such as root mean square error (RMSE) and mean absolute percentage error (MAPE) shows that the model works well at the beginning while both the RMSE and MAPE is minimal. However, both the error increases as we predict end cycles. Table 4.2 illustrates the prediction performances by using different training cycles for three separate cases. The RMSE and MAPE both values of the capacity data for three distinct cases exhibit a decreasing trend, whereas the predicted capacity becomes more and more approximate to the actual lifetime, which indicates that the model applied to FPCA has higher accuracy than other preceding proposed models (F. K. Wang and Mamo 2018).

Note that the small error values from MAPE and RMSE represent better accuracy as we are seeing from Table 4.2. The results indicate that the model using current, voltage, and temperature applied to the FPCA produces fairly accurate capacity compare to the actual capacity.

Table 4. 2: Prediction performance of battery No 7 for three separate cases.

Criteria	Case 1	Case 2	Case 3
	Training cycle 1-100 Testing cycle 101-120	Training cycle 1-120 Testing cycle 121-140	Training cycle 1-140 Testing cycle 141-160
RMSE	.009	.02	.04
MAPE (%)	.44	1.74	3.18

Finally, after rigorous analysis using proposed model, the tables and plots are providing a comprehensive and ratifying view regarding the result of the model. We can conclude with the result that we are getting quite accurate prediction of capacity of lithium-ion battery using our proposed FPCA model.

Chapter 5

Conclusions and Future Work

5.1. Conclusions

In this thesis, we have attempted to address and fill-up some research gap regarding capacity prediction using battery performance measures or monitoring data such as current, voltage and temperature. As mentioned in the literature review chapter, apparently most of the research based on capacity degradation data rather considering these performance measures. Moreover, the data we have engaged to investigate in a sequential form while past study has considered average value of each cycle. However, numerous studies have been conducted using the same dataset since the repository is publicly available for further research. Apparently, none of the preceding studies attempted to address these research gap. Thus, these factors persuaded us to confront and solve above mentioned problems. To solve these issues, the capacity prediction model is proposed using voltage, current and temperature of lithium-ion battery that applied to FPCA and lasso regression. In the methodology section, the detailed working procedure along with the basic concept of FPCA and lasso regression is described with visual representation. In the result section, the capacity prediction method for the lithium-ion battery based on FPCA using lasso regression is demonstrated. The results are represented with table and plots. The crucial facts are discussed to provide an inclusive knowledge regarding the model and its accuracy towards capacity prediction.   The prediction model is applied using FPCA which helped us to predict the capacity of the lithium-ion battery. The lasso regression is employed to explore the FPC scores that we obtained from the FPCA model and provide predicted cycle capacity accurately. From the experiment results, we observed that the proposed model based upon FPCA could effectively predict lithium-ion battery capacity. We have introduced RMSE and MAPE to assess and verify the results. Moreover, the result needs to be compared with preceding studies to check the robustness of the model. The statement verified by the error performance such as RMSE and MAPE, the minimal value of error performance proved that the model can deliver higher prediction accuracy with designed experiment.

Some limitations of the model are also observed during analysis since the model did not delivered consistent error performance for all cases. While RMSE provides decent error performance for all the scenarios, the MAPE seems showing some irregularities at the last scenario. The error of 3.18 as MAPE shows that there is significant prediction error between the actual measurements and the estimated battery capacity. This irregularity might be result of testing dataset while linear regression is considered. The ending actual capacity and cycles have irregular deterioration. However, the linear regression failed to grasp the trend which resultant predicted dissimilar cycles of current, voltage and temperature.

5.2. Future Work

The thesis was intended to develop a model that can address and explore a field of prediction maintenance of lithium-ion battery using its sequential monitoring data which is the function of its capacity. The success can bring an immense change in the capacity prediction research field and its imminent implementation process. However, the proposed model still needs to be refined before its execution to real life application. There are several factors that came to our notice while doing research with proposed model of FPCA.

The capacity degradation phenomenon of lithium-ion battery shows an irregular behavior while degrading the capacity. There are multiple occasions while capacity increased suddenly while discharging process was running and it reached to a pick from where the capacity degraded ahead time period. This path is known as degeneration trajectory from preceding study (L. Li et al. 2016). While we have extracted FPC scores from functional monitoring data from each variable, we saw the similar behavior from the dominant FPC score of each variable. These discoveries persuaded us to conduct further studies and investigate and explore the probability of accumulating the degeneration trajectory of capacity degradation process into the model so that the proposed model can capture more information from historical data of capacity and predict as per the trend of actual capacity value.

Lastly, we have detected the linear regression has limitation predicting the testing dataset based upon the training dataset as per the model. Future works are projected to be conducted considering more sophisticated curve predictions for voltage, current, and temperature rather than simple linear regression. The future goal is to use recurrent neural network (RNN) for analyzing the sequential data of training dataset and predict the testing dataset which would be more accurate than linear regression. In conclusion, we believe that the proposed FPCA model based on sequential data of battery monitoring data such as voltage, current, and temperature is a potentially useful tool that can deliver more accurate and reliable capacity prediction results.

BIBLIOGRAPHY

1. *(12) United States Patent45) Date of Patent: (54) METHOD AND APPARATUS FOR DETERMINING THE STATE OF CHARGE OF A LITHIUM-ON BATTERY*. 2004.
2. "• Projection of Lithium Demand Globally for Batteries by Type 2030 | Statista." https://www.statista.com/statistics/452010/projected-demand-for-lithium-in-batteries-by-type-globally/ (October 30, 2020).
3. An, D, JH Choi, NH Kim - Reliability Engineering & System Safety, and undefined 2013. "Prognostics 101: A Tutorial for Particle Filter-Based Prognostics Algorithm Using Matlab." *Elsevier*. https://www.sciencedirect.com/science/article/pii/S0951832013000549 (November 7, 2020).
4. Arorat, Pankaj, Ralph E White, and Marc Doyle. *Capacity Fade Mechanisms and Side Reactions in Lithium-Ion Batteries*.
5. Aurbach, Doron et al. 2007. "Review on Electrode-Electrolyte Solution Interactions, Related to Cathode Materials for Li-Ion Batteries." *Journal of Power Sources* 165(2): 491–99.
6. Aurbach, Doron, Ella Zinigrad, Yaron Cohen, and Hanan Teller. 2002. "A Short Review of Failure Mechanisms of Lithium Metal and Lithiated Graphite Anodes in Liquid Electrolyte Solutions." In *Solid State Ionics*, , 405–16.
7. B. Saha and K. Goebel. 2007. "Battery Data Set." *NASA Ames Prognostics Data Repository*. http://ti.arc.nasa.gov/project/prognostic-data-repository.
8. Barré, Anthony et al. 2013. "A Review on Lithium-Ion Battery Ageing Mechanisms and Estimations for Automotive Applications." *Journal of Power Sources* 241: 680–89.
9. "Battery Supplies Ran Dangerously Low in Iraq." https://www.nationaldefensemagazine.org/articles/2003/9/1/2003september-battery-supplies-ran-dangerously-low-in-iraq (January 29, 2020).
10. Besse, Philippe C., Herve Cardot, and David B. Stephenson. 2000. "Autoregressive

Forecasting of Some Functional Climatic Variations." *Scandinavian Journal of Statistics* 27(4): 673–87.

11. Birkl, Christoph R. et al. 2017. "Degradation Diagnostics for Lithium Ion Cells." *Journal of Power Sources* 341: 373–86.
12. Broussely, M. et al. 2005. "Main Aging Mechanisms in Li Ion Batteries." In *Journal of Power Sources*, Elsevier, 90–96.
13. Cardot, Hervé, and Pacal Sarda. 2005. "Estimation in Generalized Linear Models for Functional Data via Penalized Likelihood." *Journal of Multivariate Analysis* 92(1): 24–41.
14. Chang, Yang, Huajing Fang, and Yong Zhang. 2017. "A New Hybrid Method for the Prediction of the Remaining Useful Life of a Lithium-Ion Battery." *Applied Energy* 206: 1564–78.
15. Chen, Lin et al. 2016. "Prediction of Lithium-Ion Battery Capacity with Metabolic Grey Model." *Energy* 106: 662–72.
16. Cheng, Yujie, Chen Lu, Tieying Li, and Laifa Tao. 2015. "Residual Lifetime Prediction for Lithium-Ion Battery Based on Functional Principal Component Analysis and Bayesian Approach." *Energy* 90: 1983–93.
17. Chonde Carlos Parra Chia-Jung Chang Harold, Seifu, and Inge Marcus. *MINIMIZING FLOW-TIME AND TIME-TO-FIRST-TREATMENT IN AN EMERGENCY DEPARTMENT THROUGH SIMULATION*.
18. Coffey, N, AJ Harrison, OA Donoghue - Human movement science, and undefined 2011. "Common Functional Principal Components Analysis: A New Approach to Analyzing Human Movement Data." *Elsevier*.
19. Donà, Giulia et al. 2009. "Application of Functional Principal Component Analysis in Race Walking: An Emerging Methodology." *Sports Biomechanics* 8(4): 284–301.
20. Eddahech, Akram, Olivier Briat, and Jean Michel Vinassa. 2015. "Performance Comparison of Four Lithium-Ion Battery Technologies under Calendar Aging." *Energy* 84: 542–50.
21. Einhorn, M., F. V. Conte, C. Kral, and J. Fleig. 2010. "A Method for Online Capacity Estimation of Lithium Ion Battery Cells Using the State of Charge and

the Transferred Charge." In *2010 IEEE International Conference on Sustainable Energy Technologies, ICSET 2010,*.

22. Epifanio, Irene, Carolina Ávila, Álvaro Page, and Carlos Atienza. 2008. "Analysis of Multiple Waveforms by Means of Functional Principal Component Analysis: Normal versus Pathological Patterns in Sit-to-Stand Movement." *Medical and Biological Engineering and Computing* 46(6): 551–61.
23. Gao, Lijun, Shengyi Liu, and Roger A. Dougal. 2002a. "Dynamic Lithium-Ion Battery Model for System Simulation." In *IEEE Transactions on Components and Packaging Technologies,* , 495–505.
24. Gao, Lijun, Shengyi Liu, and Roger A Dougal. 2002b. "Dynamic Lithium-Ion Battery Model for System Simulation." *IEEE TRANSACTIONS ON COMPONENTS AND PACKAGING TECHNOLOGIES* 25(3): 495.
25. Goebel, Kai et al. 2008. "Prognostics in Battery Health Management." *IEEE Instrumentation and Measurement Magazine* 11(4): 33–40.
26. Guo, Jian, and Zhaojun Li. 2017. "Prognostics of Lithium Ion Battery Using Functional Principal Component Analysis." In *2017 IEEE International Conference on Prognostics and Health Management, ICPHM 2017*, Institute of Electrical and Electronics Engineers Inc., 14–17.
27. Han, Xuebing et al. 2019. "A Review on the Key Issues of the Lithium Ion Battery Degradation among the Whole Life Cycle." *eTransportation* 1: 100005.
28. Hu, X, SE Li, Y Yang - IEEE Transactions on Transportation, and undefined 2015. "Advanced Machine Learning Approach for Lithium-Ion Battery State Estimation in Electric Vehicles." *ieeexplore.ieee.org*.
29. Hu, Xiaosong, Jiuchun Jiang, Dongpu Cao, and Bo Egardt. 2016. "Battery Health Prognosis for Electric Vehicles Using Sample Entropy and Sparse Bayesian Predictive Modeling." *IEEE Transactions on Industrial Electronics* 63(4): 2645–56.
30. Hu, Yu, Xuming He, Jian Tao, and Ningzhong Shi. 2009. "Modeling and Prediction of Children's Growth Data via Functional Principal Component Analysis." *Science in China, Series A: Mathematics* 52(6): 1342–50.
31. Ingrassia, Salvatore, and G. Damiana Costanzo. 2005. "Functional Principal

Component Analysis of Financial Time Series." In *Studies in Classification, Data Analysis, and Knowledge Organization*, Springer Berlin Heidelberg, 351–58.

32. James, Gareth M., and Catherine A. Sugar. 2003. "Clustering for Sparsely Sampled Functional Data." *Journal of the American Statistical Association* 98(462): 397–408.
33. James O. Ramsay, Jim O. Ramsay, Ramsay, James O. Ramsay, B. W. Silverman. 1997. *Functional Data Analysis - Google Books*. https://www.google.com/books/edition/_/vYXCsgEACAAJ?hl=en&sa=X&ved=2ahUKEwi7v57Yp7HtAhVFaq0KHV8WDWoQre8FMBZ6BAgTEAY (December 3, 2020).
34. Jannesari, H., M. D. Emami, and C. Ziegler. 2011. "Effect of Electrolyte Transport Properties and Variations in the Morphological Parameters on the Variation of Side Reaction Rate across the Anode Electrode and the Aging of Lithium Ion Batteries." *Journal of Power Sources* 196(22): 9654–64.
35. Johnson, Bradley A., and Ralph E. White. 1998. "Characterization of Commercially Available Lithium-Ion Batteries." *Journal of Power Sources* 70(1): 48–54.
36. Jolliffe, I. T. 1986. "Principal Components in Regression Analysis." In Springer, New York, NY, 129–55.
37. Jolliffe, Ian T., and Jorge Cadima. 2016. "Principal Component Analysis: A Review and Recent Developments." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374(2065): 20150202.
38. Kanamura, Kiyoshi, Nao Akutagawa, and Kaoru Dokko. 2005. "Three Dimensionally Ordered Composite Solid Materials for All Solid-State Rechargeable Lithium Batteries." In *Journal of Power Sources*, , 86–89.
39. Kayano, Mitsunori, and Sadanori Konishi. 2009. "Functional Principal Component Analysis via Regularized Gaussian Basis Expansions and Its Application to Unbalanced Data." *Journal of Statistical Planning and Inference* 139(7): 2388–98.
40. Kipp, Kristof, Samuel T. Johnson, and Mark A. Hoffman. 2011. "Functional Principal Component Analysis of H-Reflex Recruitment Curves." *Journal of*

Neuroscience Methods 197(2): 270–73..

41. Lee, S. J., J. H. Kim, J. M. Lee, and Bo H. Cho. 2007. “The State and Parameter Estimation of an Li-Ion Battery Using a New OCV-SOC Concept.” In *PESC Record - IEEE Annual Power Electronics Specialists Conference*, , 2799–2803.
42. Lee, Seungchul, Mohammad Rezvanizaniani, Harry Cui, and Jun Ni. 2012. “Battery Prognostics: SOC and SOH Prediction.” In *ASME 2012 International Manufacturing Science and Engineering Conference Collocated with the 40th North American Manufacturing Research Conference and in Participation with the Int. Conf., MSEC 2012*, American Society of Mechanical Engineers Digital Collection, 689–95.
43. Li, Lingling et al. 2016. “Remaining Useful Life Prediction for Lithium-Ion Batteries Based on Gaussian Processes Mixture” ed. Yuanquan Wang. *PLOS ONE* 11(9): e0163004.
44. Li, Yi et al. 2019. “Data-Driven Health Estimation and Lifetime Prediction of Lithium-Ion Batteries: A Review.” *Renewable and Sustainable Energy Reviews* 113: 109254.
45. Liu, Binghe et al. 2020. “Safety Issues and Mechanisms of Lithium-Ion Battery Cell upon Mechanical Abusive Loading: A Review.” *Energy Storage Materials* 24: 85–112.
46. Liu, Jie et al. 2010. Conference of the Prognostics and Health Management Society *An Adaptive Recurrent Neural Network for Remaining Useful Life Prediction of Lithium-Ion Batteries*.
47. Liu, Ping et al. 2010. “Aging Mechanisms of LiFePO[Sub 4] Batteries Deduced by Electrochemical and Structural Analyses.” *Journal of The Electrochemical Society* 157(4): A499.
48. Long, B et al. “An Improved Autoregressive Model by Particle Swarm Optimization for Prognostics of Lithium-Ion Batteries.” *Elsevier*.
49. Lu, Chen, Laifa Tao, and Huanzhen Fan. 2014. “Li-Ion Battery Capacity Estimation: A Geometrical Approach.” *Journal of Power Sources* 261: 141–47.
50. Ma, S., and Y. Dai. 2011. “Principal Component Analysis Based Methods in Bioinformatics Studies.” *Briefings in Bioinformatics* 12(6): 714–22.

51. Miao, Yu, Patrick Hynan, Annette von Jouanne, and Alexandre Yokochi. 2019. “Current Li-Ion Battery Technologies in Electric Vehicles and Opportunities for Advancements.” *Energies* 12(6): 1074.
52. Musoro, Jammbe Z. et al. 2014. “Validation of Prediction Models Based on Lasso Regression with Multiply Imputed Data.” *BMC Medical Research Methodology* 14(1): 1–13.
53. Patillon, Jean-Noel, Paris ; Florence D’alche-Buc, and Antony ; Jean-Pierre. 1999. *54 SYSTEM MONITORING THE DISCHARGING PERIOD OF THE CHARGING/DISCHARGING CYCLES OFA RECHARGEABLE BATTERY, AND HOST DEVICE INCLUDING ASMART BATTERY 75 Inventors*.
54. Peled, E., D. Golodnitsky, and G. Ardel. 1997. “Advanced Model for Solid Electrolyte Interphase Electrodes in Liquid and Polymer Electrolytes.” *Journal of the Electrochemical Society* 144(8).
55. Proceedings, JD Kozlowski - 2003 IEEE Aerospace Conference, and undefined 2003. “Electrochemical Cell Prognostics Using Online Impedance Measurements and Model-Based Data Fusion Techniques.” *ieeexplore.ieee.org*..
56. Rao, C. Radhakrishna. 1958. “Some Statistical Methods for Comparison of Growth Curves.” *Biometrics* 14(1): 1. https://www.jstor.org/stable/2527726?origin=crossref (February 6, 2020).
57. Rice, John A., and B. W. Silverman. 1991. “Estimating the Mean and Covariance Structure Nonparametrically When the Data Are Curves.” *Journal of the Royal Statistical Society: Series B (Methodological)* 53(1): 233–43.
58. Saha, Bhaskar, Scott Poll, Kai Goebel, and Jon Christophersen. 2007. “An Integrated Approach to Battery Health Monitoring Using Bayesian Regression and State Estimation.” In *AUTOTESTCON (Proceedings)*, , 646–53.
59. Sarasketa-Zabala, E. et al. 2015. “Cycle Ageing Analysis of a LiFePO4/Graphite Cell with Dynamic Model Validations: Towards Realistic Lifetime Predictions.” *Journal of Power Sources* 275: 573–87.
60. ———. 2016. “Realistic Lifetime Prediction Approach for Li-Ion Batteries.” *Applied Energy* 162: 839–52.
61. Sarasketa-Zabala, E, E Martinez-Laserna, M Berecibar - Applied energy, and

undefined 2016. “Realistic Lifetime Prediction Approach for Li-Ion Batteries.” *Elsevier*..

62. Särkkä, S. 2013. *Bayesian Filtering and Smoothing*. www.cambridge.org/sarkka. (November 7, 2020).
63. Shamsudduha, Mohammad, Richard G. Taylor, and Richard E. Chandler. 2015. “A Generalized Regression Model of Arsenic Variations in the Shallow Groundwater of Bangladesh.” *Water Resources Research* 51(1): 685–703. (February 12, 2020).
64. Stroe, Daniel Ioan et al. 2014. “Accelerated Lifetime Testing Methodology for Lifetime Estimation of Lithium-Ion Batteries Used in Augmented Wind Power Plants.” In *IEEE Transactions on Industry Applications*, Institute of Electrical and Electronics Engineers Inc., 4006–17.
65. Tibshirani, Robert. 1996. “Regression Shrinkage and Selection Via the Lasso.” *Journal of the Royal Statistical Society: Series B (Methodological)* 58(1): 267–88.
66. “Types of Battery Cells; Cylindrical Cell, Button Cell, Pouch Cell.” https://batteryuniversity.com/learn/article/types_of_battery_cells (December 2, 2020).
67. Vetter, J. et al. 2005a. “Ageing Mechanisms in Lithium-Ion Batteries.” *Journal of Power Sources* 147(1–2): 269–81.
68. ———. 2005b. “Ageing Mechanisms in Lithium-Ion Batteries.” *Journal of Power Sources* 147(1–2): 269–81.
69. Wang, Fu Kwun, and Tadele Mamo. 2018. “A Hybrid Model Based on Support Vector Regression and Differential Evolution for Remaining Useful Lifetime Prediction of Lithium-Ion Batteries.” *Journal of Power Sources* 401: 49–54.
70. Wang, Jane-Ling, Jeng-Min Chiou, and Hans-Georg Müller. “Review of Functional Data Analysis.” www.annualreviews.org (November 8, 2020).
71. Xing, Y et al. “An Ensemble Model for Predicting the Remaining Useful Performance of Lithium-Ion Batteries.” *Elsevier*.
72. Yao, Fang et al. 2003. “Shrinkage Estimation for Functional Principal Component Scores with Application to the Population Kinetics of Plasma Folate.” *Biometrics* 59(3): 676–85.
73. Yao, Fang, and Thomas C. M. Lee. 2006. “Penalized Spline Models for Functional

Principal Component Analysis.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68(1): 3–25. (November 8, 2020).

74. Yao, Fang, Hans-Georg Müller, and Jane-Ling Wang. 2004. *Functional Data Analysis for Sparse Longitudinal Data Short Title: FDA for Sparse Longitudinal Data*.
75. Zhang, Jingliang, and Jay Lee. 2011. “A Review on Prognostics and Health Monitoring of Li-Ion Battery.” *Journal of Power Sources* 196(15): 6007–14.
76. Zhou, Y, M Huang - Microelectronics Reliability, and undefined 2016. “Lithium-Ion Batteries Remaining Useful Life Prediction Based on a Mixture of Empirical Mode Decomposition and ARIMA Model.” *Elsevier*.).