UNIVERSITY OF CALIFORNIA SAN DIEGO

Design optimization approach to estimate the second life lithium-ion batteries life cycle prediction

A thesis submitted in partial satisfaction of the
requirements for the Master of Science

in

Mechanical and Aerospace Engineering

by

Kwangwoo Yeum

Committee in charge:

Professor Jan Kleissl, Chair
Professor Michael Davidson
Professor Patricia Hidalgo-Gonzalez

2022

The thesis of Kwangwoo Yeum is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would love to express my gratitude to my advisors, DR. Jan Kleissl and Dr. Antoni Tong, who guided me throughout this thesis. I would also like to thank my friends and family who supported me and offered deep insight into the broad study.

ABSTRACT OF THE THESIS

Design optimization approach to estimate the second life lithium-ion batteries life cycle prediction

by

Kwangwoo Yeum

Master of Science in Mechanical and Aerospace Engineering

University of California San Diego, 2022

Professor Jan Kleissl, Chair

Lithium ion (Li-ion) batteries degrade with cyclic usage and storage duration. Batteries close to their end-of-life can no longer fulfill their performance requirements, and have an increased likelihood of catastrophic failures. Different usage conditions, complex manufacturing, and lack of essential data contribute to the complex degradation of second life batteries and hinder accurate analysis of battery capacity degradation. Therefore, a quick and precise diagnosis of used batteries has become an important research area for battery management, specifically in large-scale power storage systems. This thesis illustrates potential approaches for diagnosing battery degradation, considering both a physical-based and a data-driven model. The main objective is to

boost the degradation prediction with the data-driven model by leveraging artificially generated data from the physical model.

The approach is divided into two steps with two different battery cell models. A gradient-free optimization approach is introduced with the most widely used battery model (21700), which comes with published data of its battery cell structure, to optimize inside-of-the-cell structure. We estimate the physical battery cell structure to acquire artificial data and compare the performance of the estimated battery against the original battery. Then the degradation prediction is investigated with the A123 battery, which has extensive and high quality of battery degradation data but lacks exact cell physical structural information. We acquire artificial degradation data by estimating the physical battery cell structure with the optimization approach, and then utilize these data to boost the original battery life cycle prediction. Linear regression and backpropagation methods (resilient backpropagation, conjugate backpropagation, and bayesian regularization backpropagation) are applied along with various test matrices and cross-validation to compare the life cycle predictions. The life cycle prediction was first conducted with only the actual data. Then the artificial data was added to the training sets to conduct the life cycle prediction. The life cycle prediction error with actual data was improved with resilient backpropagation and a sigmoid activation function by 4 - 6% compared to linear regression. Backpropagation performs better on test sets containing a large amount of extreme values than the linear regression. Lastly, the prediction results with artificial data and actual data were compared by using the regression and resilient backpropagation. Adding artificial data reduced the life cycle prediction error by 0.3 - 1.3% in the regression method; there was a greater improvement when the life cycle prediction error with actual data was over 15%. However, with the backpropagation method, adding artificial data resulted in larger prediction errors.

Chapter 1

Introduction

1.1 Motivation

The nearly exhausted fossil energy sources, the overall population growth, and the significant number of electric devices in the world have necessitated the search for alternative sources of energy and research in recycle energy [12]. Every year new technology amazes human beings, and we enjoy our developed presents. However, these trends have various side effects that damage the environment. Due to large amounts of energy consumption, reducing energy loss has become one of the main issues regarding our planet that people are determined to resolve. Recently, Lithium-ion batteries (LIB) have drawn the most interest compared to other battery types due to their long life and high energy density following the growth of global battery demand [61]. Many researchers make an effort to reuse wasted battery cells in the energy grid, as well as improve the battery performance to withstand the future global battery demand, shown in figure 1.1. Figure 1.2 estimates the second life battery capacity to increase as the global electrical vehicle demand grows.

One major challenge for the second-life LIB research is predicting the battery life cycle accurately [45]. There are many on going efforts to understand the degradation process and physical LIB model. However, diverse aging mechanisms, significant manufacturing variability, varying operating conditions, and lack of actual battery lifetime data remain significant challenges to improve our understanding of the second life battery degradation [54][14]. Furthermore, aging

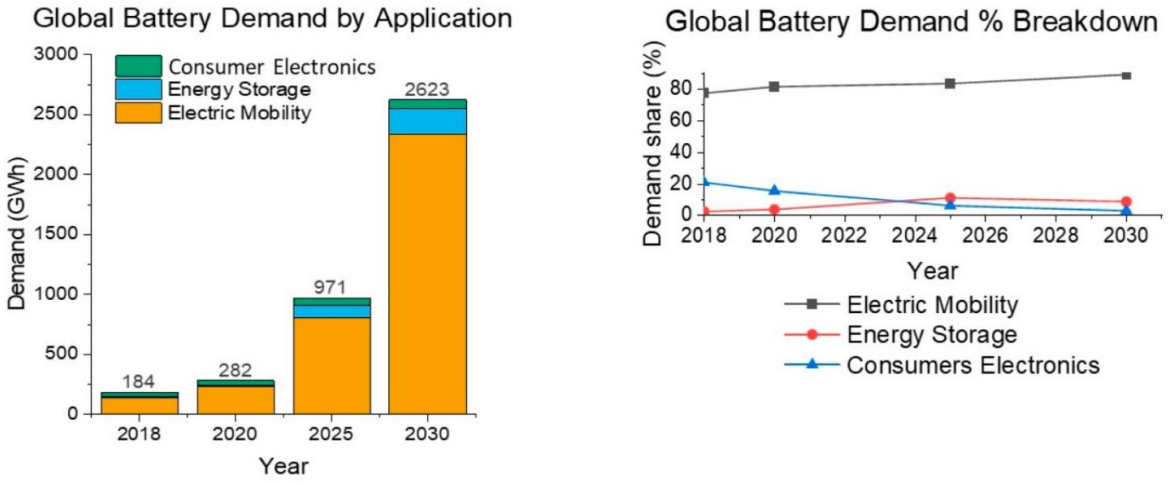


Figure 1.1. Global battery demand by application and percentage breakdown from 2018 to 2030. Figure from [2].

tests for the new battery can take up to six months to understand degradation performance [10], which also makes it difficult to forecast the battery degradation. For this reason, we studied a design optimization method to estimate the unknown physical structure, like a second life battery cell, and utilize the result for a battery life cycle prediction.

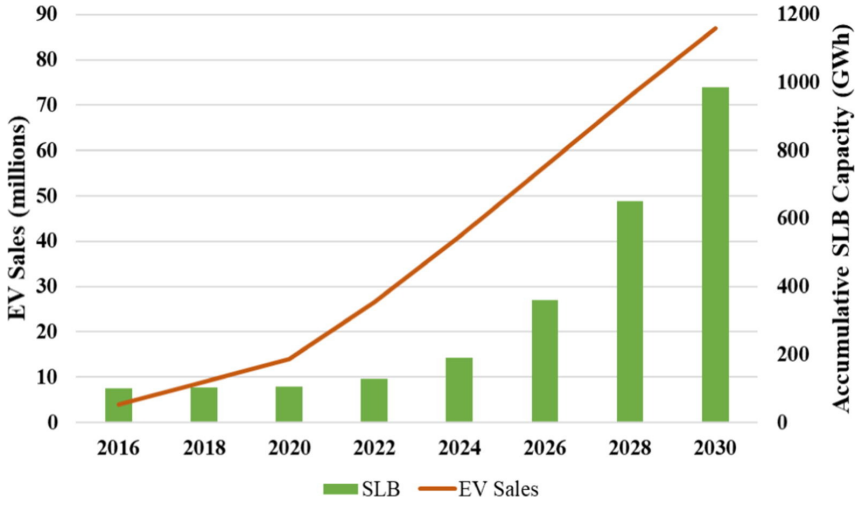


Figure 1.2. Estimation of the second life battery (SLB) capacity increases due to the global battery demand in electrical vehicle (EV) from 2018 to 2030. Figure from [19].

1.2 Thesis organization

This thesis studies both a physical-based and a data-driven model for diagnosing battery degradation. The main objective is to boost the life cycle prediction with the data-driven model by leveraging artificially generated data from the physical model. A design optimization approach to estimate inside of the cell structure is introduced with a 21700 (LGM50) battery model with published data of its battery cell structure. Then the degradation prediction is explored with the A123 battery, which has extensive and high quality of battery degradation data but lacks cell structural information. This paper is organized as follows: Chapter 2 covers the basic overview of the physical and data-driven models in battery degradation research and explains the battery degradation modes. Chapter 3 illustrates the battery life cycle prediction with various test matrices and methods with actual data and generated data by an electrochemical battery model. In Appendix C, we describe a design optimization approach as a tool to understand the unknown cell structure of the battery used in chapter 3.

Chapter 2

Overview of the battery life cycle prediction methodology and the Li-ion physical battery models

2.1 Battery life cycle prediction methodology

The LIB life cycle has been a crucial subject of battery performance since its introduction to the commercial vehicle [13]. Figure 2.1 illustrates the comparison of different lifetime modeling methodologies. First of all, the data-driven models are often built with fitting mathematical equations or trained with learning algorithms. These black-box modeling approaches can forecast the whole life in terms of battery capacity fade. For example, Severson et al. used data-driven black-box modeling to predict the battery life cycle with 10% error. They extracted the capacity feature in the early life cycle and performed prognostic using machine learning algorithm [43]. Fortunately, the data-driven models in figure 2.1 require little or no knowledge of the battery chemistry, but a significant amount of characterization data is necessary to structure the model response. The model performance is also heavily reliant on the quality and quantity of the investigated data set.

On the contrary, LIB electrochemical models can identify the loss of lithium inventory and active materials by analyzing the key degradation mechanisms such as solid-electrolyte interphase and lithium plating [32]. The Doyle-Fuller-Newman (DFN) model is a popular

electrochemistry-based lithium-ion battery model which represents solid-state and electrolyte diffusion dynamics and accurately predicts the current and voltage response [11]. Many model proposals and parameter studies were introduced following this model, such as the single-particle model (SPM) and single-particle with electrolyte model. However, we cannot consider all the side effects of battery usage using the LIB physics-based model and have a lack of applicability to practical purposes.

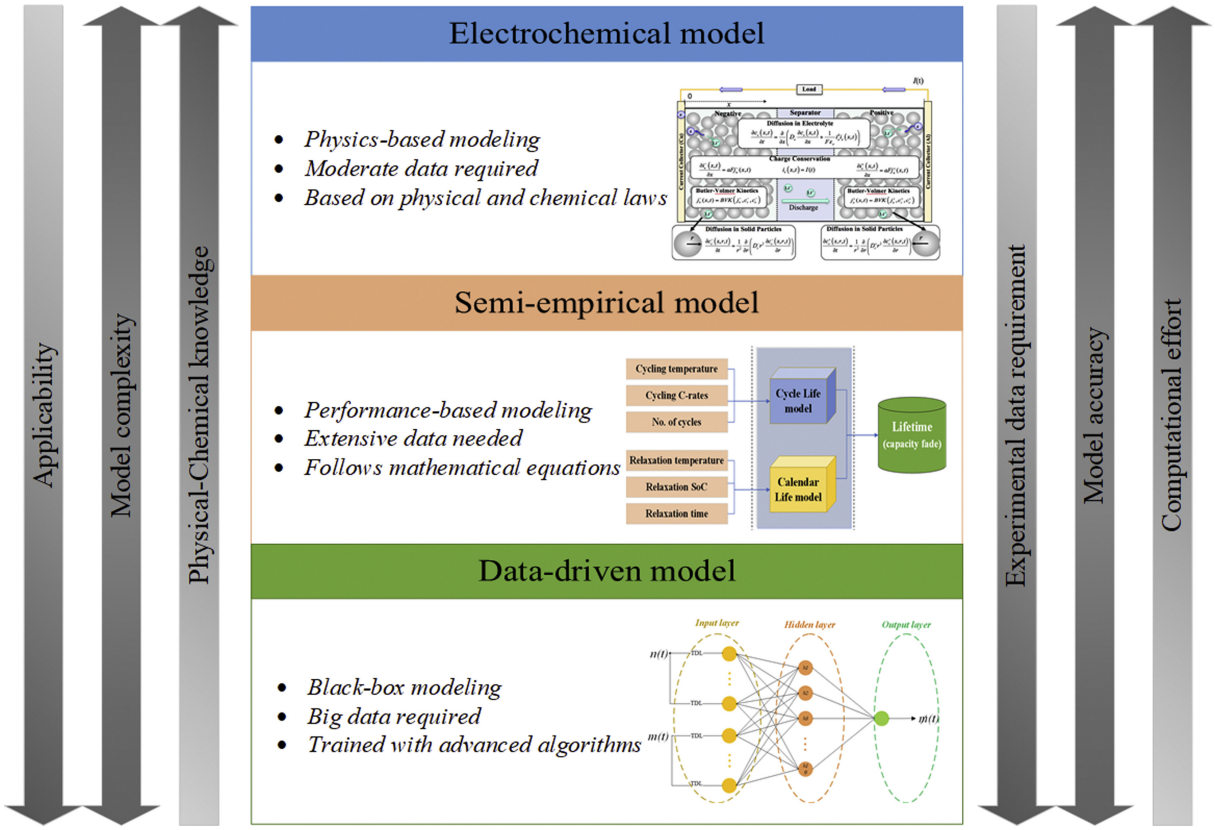


Figure 2.1. Methodologies of the battery life time prediction with electrochemical, semi-empirical, and data-driven models. Each strength and weakness is described in the figure. Figure from [3].

There has been remarkable progress in estimating the behavior of Li-ion batteries through data-driven and electrochemical models. The data-driven model offers a flexible statistical framework that yields applicable predictions but lacks a physical description of degradation modes which can restrict its reliability and have a disadvantage of a significant time needed for generating experiment data [17]. Physics-based models provide a mechanical framework

to capture the influence of degradation physics on electrochemical behavior, but challenges include difficult-to-identify degradation modes. It has become clear that the electrochemical and data-driven models provide distinct but complementary qualities to each other [4]. Thus, based on their properties, the following step would be investigating a proper blend for an enhanced battery life cycle prediction. It is facilitated as a proper approach for a better battery life cycle prediction to develop a combined model for LIB life cycle studies.

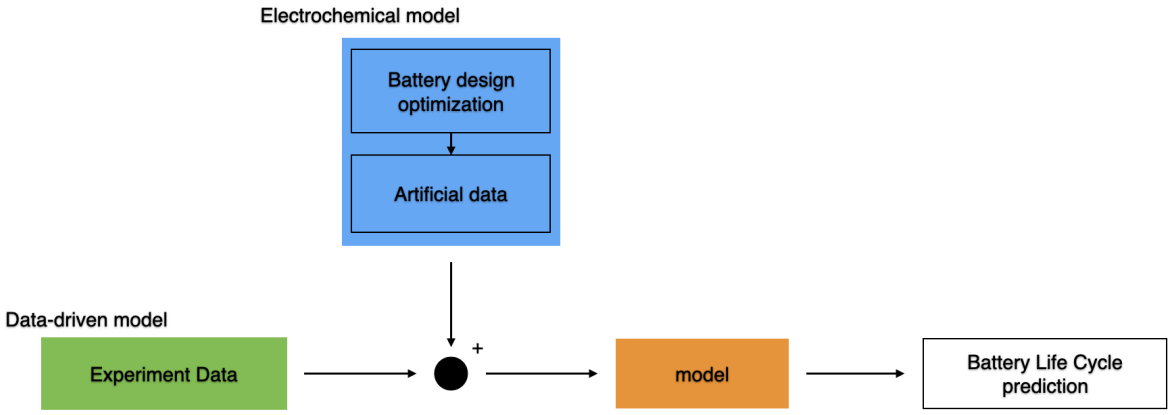


Figure 2.2. Flow chart of the thesis structure. The main objective is to implement artificial data from electrochemical models to boost the life cycle prediction.

We have learned various methodologies of predicting Li-ion battery life span, however, predicting the performance of a second life battery is not achieved by one of the previously introduced methods. Due to the limited information on a second-life battery cell structure using physical models and increased time to complete experiments with data-driven models, we approach the design optimization to estimate the battery cell structure with published material data and apply those data to predict battery life cycle. Figure 2.2 shows the flow chart in this thesis to help the battery life cycle prediction. This chapter discusses a standard electrochemical model of a single battery cell and the diversity of battery degradation modes. The detailed battery model and simulation solver information are addressed in Appendix A. The optimization methodology is described in Appendix B.

2.2 Physical electrochemical battery model

Battery cell degradation has been typically considered as the nonlinear behavior due to the nonlinear governing equations with variables of the cell voltage, energy density, and power output [60]. In addition, factors such as activation polarization, concentration polarization, and internal resistance contribute to the cell's potential capacity drop and impact the degradation [35]. Thus, it is not easy to simplify the degradation problem equations for the LIB. However, a cell model has to account for all the major internal processes occurring in design structures to understand the LIB behavior accurately.

2.2.1 Lithium ion battery cell models

A typical LIB is composed of 1) electrodes (anode and cathode), 2) a separator between the two electrodes, and 3) an electrolyte that fills the remaining space of the battery. The anode and cathode can store Li-ions. Energy is collected and delivered as Li-ions which travel between these electrodes through the electrolyte. Figure 2.3 shows the standard battery cell structure.

The detailed cell model in figure 2.4 contains a homogeneous electrode formulation with a concentrated solution theory [33]. The porous electrodes consist of solid active materials and liquid electrolytes. The state variables solved in the model are constructed from the ion concentration in the electrolyte and solid matrix, the current density in the electrolyte, the interfacial current density at the solid matrix surface, and the potentials in the electrolyte and solid phases [33]. These state variables, in turn, provide the cell properties that are used to evaluate the cell energy density and power output. The governing equations with the associated transport parameters are defined in all three regions of the electrochemical cell: the positive electrode (cathode), the negative electrode (anode), and the separator. The influence of porosity is accounted for using Bruggeman's relation for spherical particles [6].

Many electrochemical models, such as the Doyle-Fuller-Newman (DFN) and the Single Particle Model (SPM) models, have been developed throughout the years. The DFN model

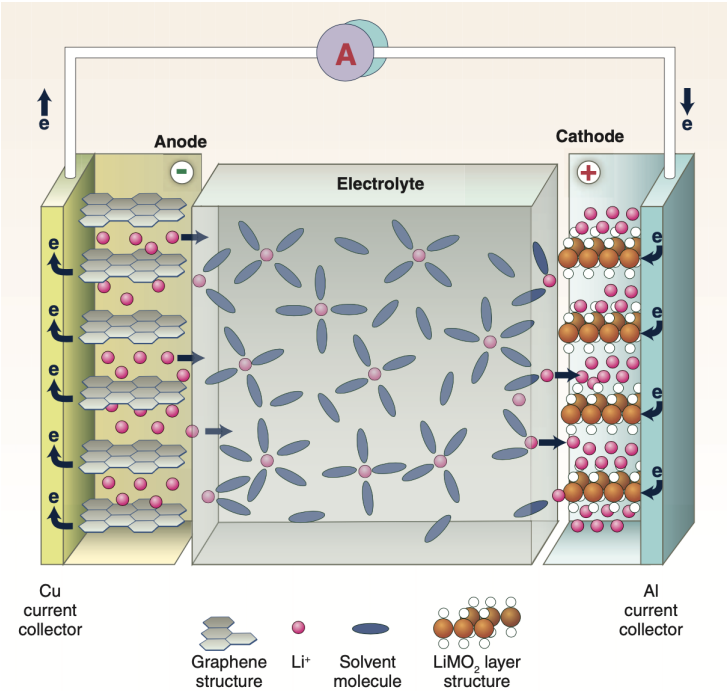


Figure 2.3. Standard Li-ion battery cell structure. Li-ion cell mainly consists of the anode, separator, cathode, and electrolyte. Figure from [55].

covers the most detailed electrochemical reactions, including the complex electrode diffusion from Newman [32]. However, many researchers have been using SPM models, which contain a relatively simplified governing equation, yet are still applicable with lower accuracy. Marquis et al. has improved the accuracy and range of applicability of the SPM model by including electrolyte effects [27]. This model, known as the Single Particle Model with electrolyte (SPMe), allows the user to quantify the error in the reduced model and lets users analyze essential electrolyte reactions [27][31]. The SPMe model simplifies the solid phase Li concentration in each electrode, and the electrode concentration is uniformly distributed in time. In our paper, the SPMe model with Pybamm open source is implemented for the simulation. A detailed description of this model has been presented in [23].

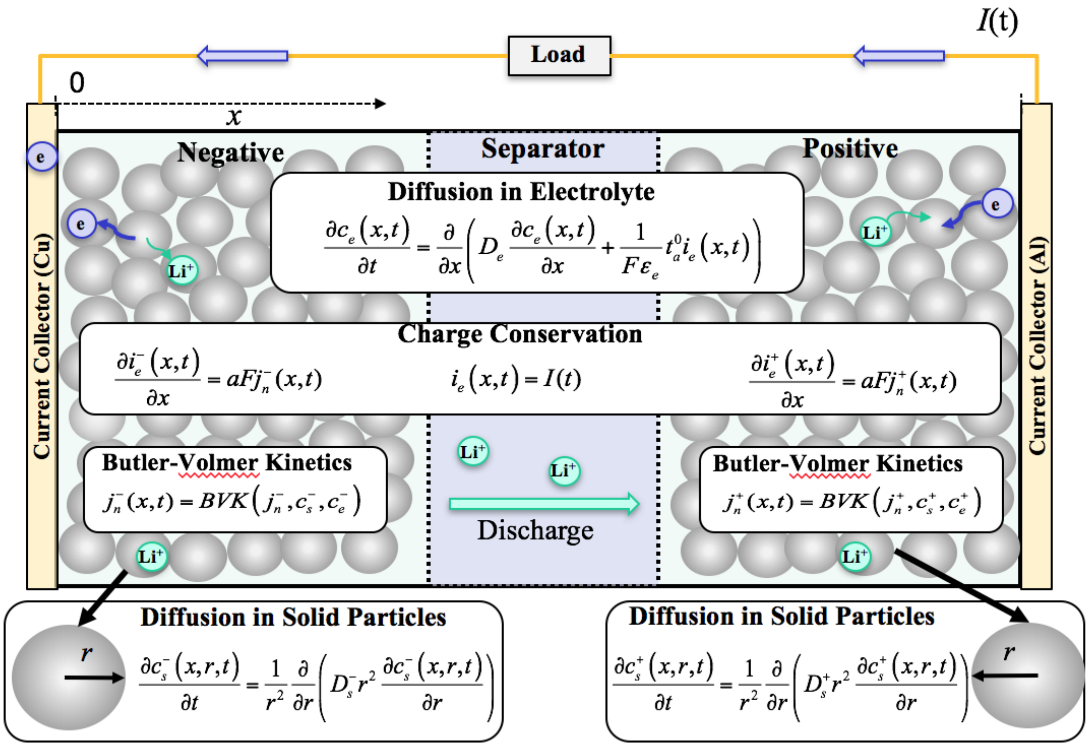


Figure 2.4. A structure of a physical battery model and governing equations. The governing equations consist of diffusion in electrolytes, charge conservation, butler-volmer kinetics, and diffusion in particles. Figure from [21].

2.2.2 Degradation modes of lithium ion battery

Various mechanisms are related to each other within the process of battery degradation. Figure 2.5 represents the complexity of the degradation process, which can be driven by the time, temperature, voltage, current load, mechanical stress, etc. Solid Electrolyte Interphase (SEI) degradation mechanisms are the most knowledge-demanding during the battery cycle, and are followed by many kinds of research. Figure 2.6 shows the schematic illustration of the 2D SEI degradation and the SEI thickness increases during the Li-ion travel [3]. This SEI formation is an irreversible process that consumes cyclable lithium and is one of the primary causes of lithium loss in the cell [49]. However, still, the formation and growth mechanism of the entire model of SEI are yet to be understood entirely due to their complex structure and lack of reliability in the experiment, which has required an empirical trial-and-error process [52][58].

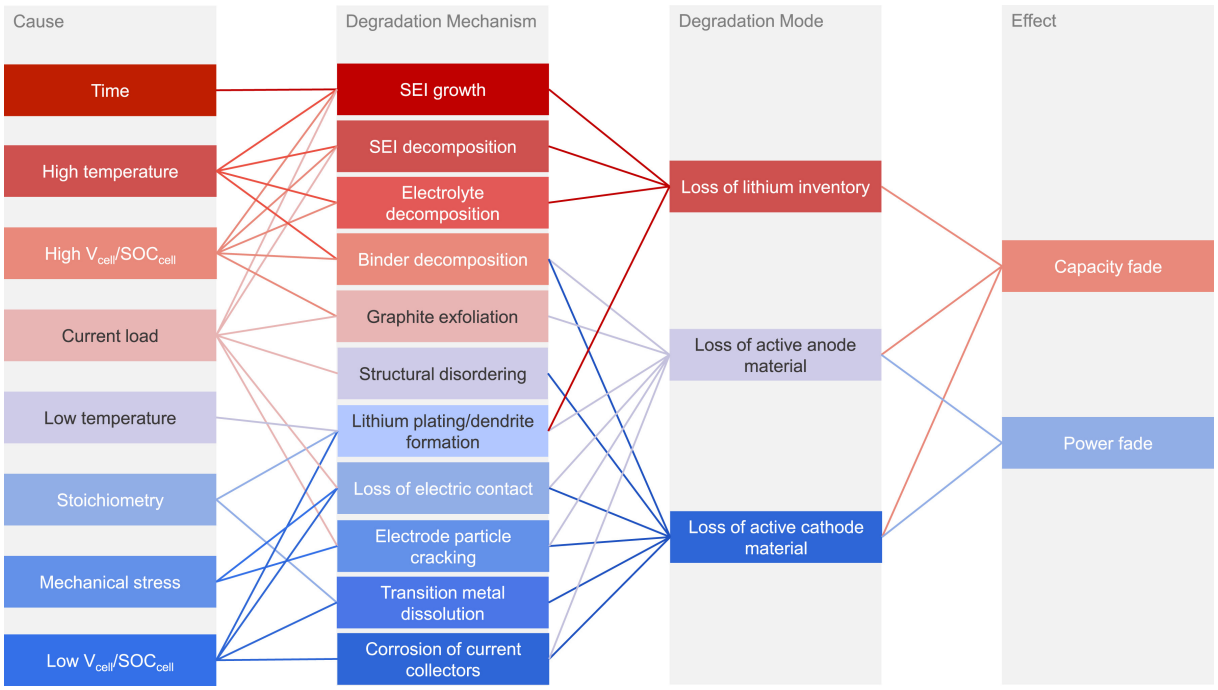


Figure 2.5. Degradation modes in a Li-ion battery. Various factors cause the degradation mechanism. Figure from [5].

To better understand LIB degradation, the battery physical cell model should include a physics-based interaction between electrode active material and electrolytes, which influences the SEI thickness and lithium loss. In the SPMe model we covered in 2.2.1, various assumptions exist to explain SEI layers, such as neglecting the solid deformation and constant reaction rate from the electrolyte, and other sources of the capacity fade can be neglected [37]. We use 'First-order solvent decomposition kinetics' in the SPMe model to generate degradation data, which assumes the rate of SEI formation is proportional to the concentration of the reacting electrolyte species in a battery cell [15]. Recently, Sulzer et al. proposed an adaptive inter-cycle extrapolation algorithm that can simulate the overall life cycle of Li-ion batteries, and shown the approach to work well in a Single Particle Model with SEI formation and loss of active material. In his approach, he used SEI kinetic rate constant as one of the key degradation parameters in his accelerated battery lifetime simulations [46][57]. Thus, we generated artificial data by assuming these information as key variables. In this thesis, the SEI growth mechanism is considered for

the loss of lithium inventory to generate the artificial degradation data [52][34].

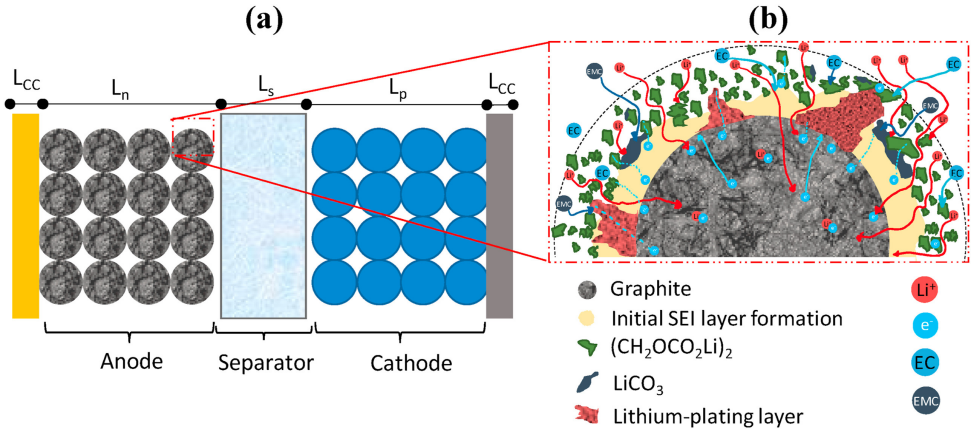


Figure 2.6. A schematic illustration of the SEI layer behavior with active materials in a battery cell. (a) A schematic illustration of the 2D representation in a battery (b) shows the illustration of SEI layer reaction with active materials on anode in the battery ageing mechanism. Figure from [3].

Chapter 3

Battery life cycle prediction and performance assessment with artificial data

3.1 Early prediction of a battery life cycle with artificial data

3.1.1 Background

The latest data-driven approaches have shown great possibilities in forecasting the battery life cycle using battery performance features. Severson et al. illustrated a fairly accurate life cycle prediction with the first 100 cycle data of the discharge voltage curve. In their study, the 124 A123 (APR18650M1A) cells were cycled between full charge and full discharge with identical nominal temperatures and discharge C-rates but varying charge C-rates. Severson et al. selected six extracted features from early life cycle capacities in A123 battery and used those features for linear regression to predict a battery life cycle. Additional researches have been performed to improve the prediction result of this A123 battery life cycle with a machine learning. Fei et al. showed six different machine learning models by generating a set of 42 features from Severson et al.'s A123 battery data [16]. However, in these studies, we noticed that data partitions are needed as an improvement due to an imbalanced sample distribution. To improve this imbalanced data distribution, we suggested generating the artificial data from a physical model with a design optimization approach and battery degradation assumptions. We added a cross-validation with linear regression and compared the proposed feed-forward neural network to understand actual

data performance. Then we used the generated artificial data to examine the boost effect of the prediction. This work aims to understand the proposed approach which can be beneficial for degradation prediction rather than improving the prognostics algorithm itself.

3.1.2 Data-driven prediction of the battery life cycle and feature extraction

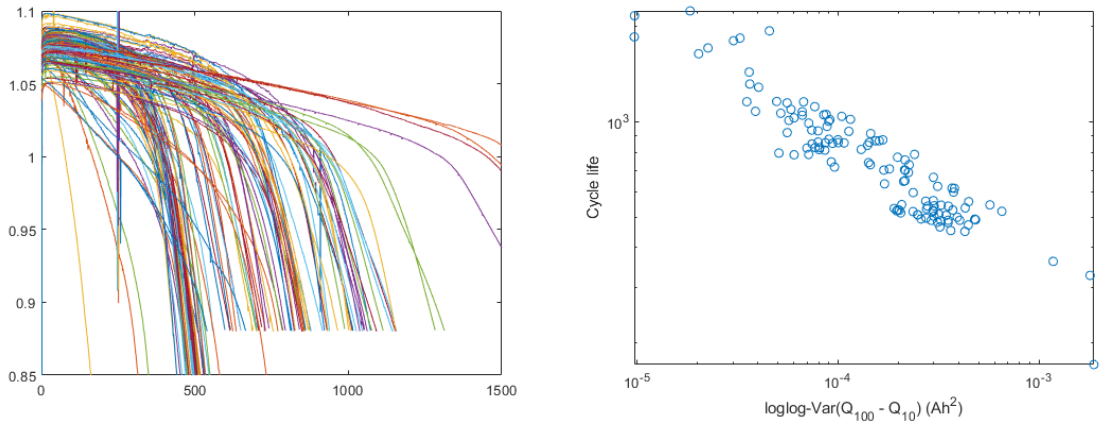
For this chapter, we used Phosphate/graphite A123 battery degradation data from [43] for the battery life cycle prediction and compared it with the artificial degradation data. Severson et al. generated a data set of 124 cells with life cycles ranging from 150 to 2,300 using 72 different fast-charging conditions where the end of battery life is defined as 80% of the nominal capacity. The commercial Phosphate/graphite cells were cycled under varied fast-charging conditions and identical discharging conditions. Figure 3.1(b) shows the strong correlation in Severson et al.'s proposed feature using the early cycled data such as the variance of the capacity ($Q_{100} - Q_{10}$). This correlation stands for the charge-discharge voltage curves containing more essential information for the degradation prediction than simply providing the capacity of the cell [43].

Table 3.1. Extracted features of the A123 Li-ion battery from Severson et al. which work as a guideline to generate artificial data with the electrochemical battery model.

Extracted feature	Contents
1	Capacity log variance of difference cycles 10 and 100
2	minimum of log variance of difference cycles 10 and 100
3	Slope of linear fit to capacity fade curve from cycles 2 to 100
4	Intercept of linear fit to capacity fade curve from cycles 2 to 100
5	Discharge capacity at cycle 2
6	Average charge time over first 5 cycles
7	Minimum internal resistance between cycles 2 and 100
8	Internal resistance difference between cycles 2 and 100

The cycle-to-cycle evolution of Q(V), the discharge voltage curve as a function of voltage for a given cycle, is a vital feature. In figure 3.1(a), the capacity fade accelerates near the end of

the life while the capacity fade is negligible in the first 100 cycles. The proposed eight essential features from Severson et al. are shown in table 3.1.



(a) **(b)**

Figure 3.1. A degradation curve and a graph representing linear relations between log-based life cycle and variance of capacity differences in the early cycles of the A123 battery. (a) shows the degradation curve of discharge capacity throughout the cycle. (b) represents the linear relationship between the log-based life cycle and the variance of capacity difference between 10th and 100th cycle. Data from [43].

3.1.3 Artificial data generation and feature extraction

In this chapter, we estimate the unknown physical battery cell structure of an A123 battery by adopting the design optimization method covered in Appendix C. Unlike the 21700 battery cell introduced in Appendix C, the published data of the cell structure for the A123 battery is not available. Likewise, other second life batteries are not equipped with these kind of published data in general because the battery parameterization is not a simple task as illustrated in Appendix A. Thus, we started from referencing the published data of the cathode and anode with same material parameters from Prada et al. and Chen et al [38][8]. Then we implemented the optimization method described in Appendix C as a tool to estimate the battery cell structure. Additional adjustment constraining the specific required power (over $1,850W/kg$) to keep the nominal capacity was made. As a result, we acquired the below result (Table 3.2) for the degradation simulation.

In chapter 2.2.2, we discussed the complex degradation modes of Li-ion batteries. There were three main degradation modes as shown in figure 2.5. Among those modes, we mainly consider the loss of lithium inventory for generating artificial data in this thesis. Table 3.3 shows the essential parameters of the SPMe model to understand the degrading performance referenced from [41]. We noticed that the degradation result largely depends on SEI kinetic rate constant in the simulation and used as a variable like we discussed in 2.2.2.

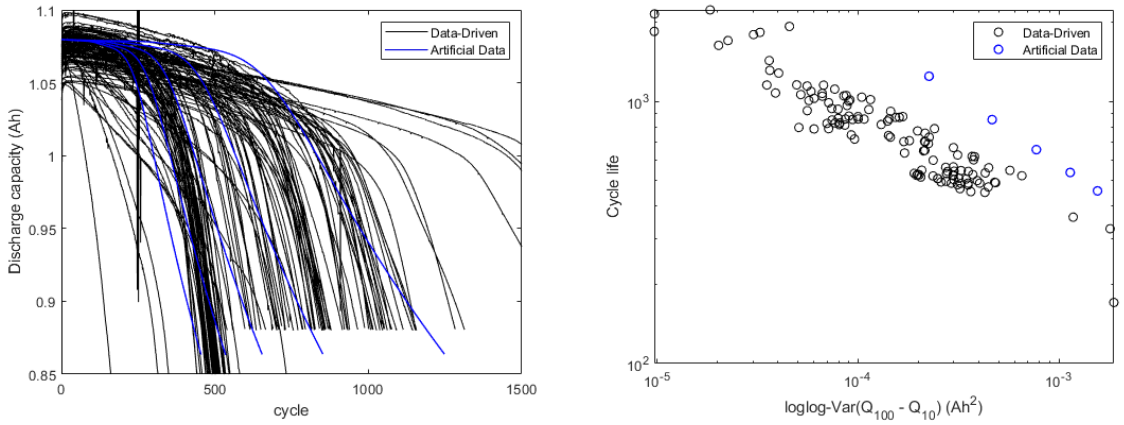
Table 3.2. The result of the estimated A123 battery cell structure by the optimization approach. The chemical parameters of battery materials came from Prada et al. and Chen et al. Detailed parameters can be found in [38][8].

Optimized Parameter	Unit	Anode	Cathode	Separator
Material		Graphite	LiFePo4	polyolefin
Electrode Density	kg/m^3	1657	2341	397
Electrode thickness	μ m	54	24	12
Electrode Porosity		0.25	0.3	0.4
Collector thickness	μm	10	19	18
Collector Density	kg/m^3	8960	2700	
Overall length	cm	178		
Overall width	cm	6.49		

Figure 3.2 (a), (b) shows both artificial and real data in one figure. The SEI kinetic rate constants are selected to $1.0 \cdot 10^{-16}$, $1.5 \cdot 10^{-16}$, $2 \cdot 10^{-16}$, $2.5 \cdot 10^{-16}$, and $3 \cdot 10^{-16}$. We observed an offset exists due to a not perfect battery structure estimation and degradation related assumptions, however, we can see the linear relationship remains. Based on this linearity, we generated the degradation data to improve the imbalanced data set and planned to see the boosting effect for the life cycle prediction in the following section. We adopted Severson et al.'s charge and discharge policy for the degradation cycle simulation. The cells were cycled with various fast-charging policies (Supplementary Table 9 from [43]) but discharged following one identical policy. The voltage cutoffs (C/50) used in this work were determined by the recommendation of the A123 manufacturer.

Table 3.3. SEI variables for the battery degradation simulation. Detailed parameters can be found in [25][46].

SEI default variables	Unit	Value
Initial SEI thickness	m	$2.5 \cdot 10^{-9}$
SEI kinetic rate constant	m·s^{-}1	$1 \cdot 10^{-16}$
EC initial concentration in electrolyte	mol·m^{-}3	4541.0
Outer SEI solvent diffusivity	m^2·s^{-}1	$2 \cdot 10^{-18}$
SEI open-circuit potential	V	0.5



(a) **(b)**

Figure 3.2. A degradation curve (figure 3.1) with artificial data added. The black line represents actual data of the A123 battery and blue line is artificial data generated by a physical battery model. The SEI parameters were selected from table 3.3. The SEI kinetic rate constant for each artificial data is $1.0 \cdot 10^{-16}$, $1.5 \cdot 10^{-16}$, $2 \cdot 10^{-16}$, $2.5 \cdot 10^{-16}$, and $3 \cdot 10^{-16}$, respectively.

3.2 Machine learning algorithm for the battery prediction

3.2.1 Linear regression and artificial neural network

Researchers often use linear regression to characterize the relationship between the inputs and outputs of both natural and social sciences. An artificial neural network (ANN) is also often used for prediction and classification. ANN is a computational model formed by hidden units interconnected by weights that can be updated according to the quality parameters, which evaluate the proximity between the critical response and the one obtained [59]. The networks are

arranged in layers; the first layer takes an input and the last layer produces outputs. The middle layers have no connection with the external world, and that is why they are called hidden layers [20]. ANN consists of a feed-forward neural network such as backpropagation. Based on the backpropagation, the recurrent neural networks and convolutional neural networks are widely used to understand huge black box models. The neuron and a basic model of ANN structure are shown in figure 3.3.

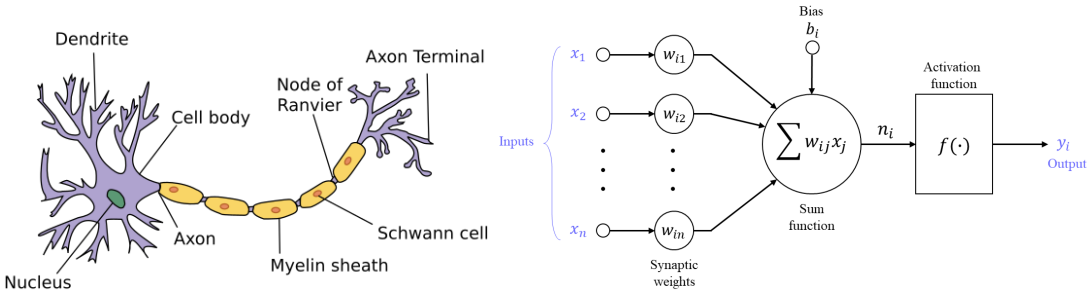


Figure 3.3. A neuron and a basic model of an artificial neural network structure. Figure from [18].

3.2.2 Test matrix and life cycle prediction methodology

Test matrix

The life cycle prediction is conducted in this chapter following table 3.4. There are four different test modes with different partitions of test data. These four test modes with different percentages of test sets are to see the impact of the training data size on the test result. Test mode 4 used 30% of data for the test set, which has a similar test partition to Severson et al. and Fei et al. [43][16]. The cross-validation is conducted in each test with five folds, and the results are averaged by the number of folds as shown in figure 3.4. This cross-validation exclusively partitions the test, validation, and train data set. The performance matrices are evaluated by two types of errors—the root mean square error (RMSE) and mean absolute percentage error (MAPE).

Table 3.4. A test matrix for the life cycle prediction. Different partitions of test, valid, train data are constructed. Cross-validation was applied per trial, and each test mode consists of a total of five trials.

Test Mode	Mode 1	Mode 2	Mode 3	Mode 4
Test set	10%	15%	20%	30%
Validation set	10%	15%	20%	20%
Training set	80%	70%	60%	50%
Cross validation fold	5	5	5	5
Trial	5	5	5	5

The equation below demonstrates the calculation of the RMSE:

$$RMSE = \sqrt{(\frac{1}{n}) \sum_{i=1}^{n} (y_i - x_i)^2}$$

The percentage average of all errors was calculated to quantify the results. The vector containing the absolute difference between the actual and estimated output values is calculated for the error percentage. The equation of the MAPE is :

$$MAPE = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{y_i - x_i}{y_i} \right|$$

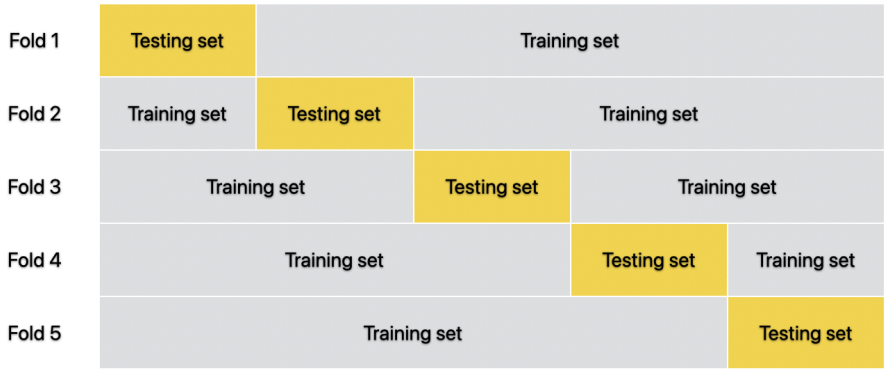


Figure 3.4. Structure of the cross-validation. The error is calculated by the average of all folds.

Methodology with a linear regression model

Severson et al. presented the regression methods to understand the LIB life cycle performance [43]. His method contains model fitting and selection and employs the regularization technique for a secure prediction. This linear regression method was brought to this thesis as a baseline of the battery life time prediction.

$$y_i = w^T x_i + \beta$$
$$\hat{w} = min_w(||y - Xw - \beta||_2)^2 + \lambda P_\alpha(w)$$

where,

$$\lambda P_\alpha(w) = \frac{1-\alpha}{2}||w||_2^2 + \alpha||w||_1$$

A linear model of the layout was proposed where y_i is the foreseen number of cycles for battery i, x_i is a set of feature vectors for battery i, and w^T is a set of model coefficient vectors, and β is a scalar intercept. The linear model is regularized using the elastic net (net58) by MATLAB to handle high correlations between the features. A penalty term is combined to the least-squares optimization equation to avoid overfitting. The training data is used to determine the hyperparameters α and λ, and choose values of the coefficients, w^T. The above regularization solves the problem for α between 0 and 1, and λ, and penalizes the weight to make it small in proportion to its size.

Methodology with the artificial neural network

We propose an additional approach using a backpropagation feed-forward neural network. ANN is an interconnected group of nodes, which is frequently used in engineering applications. This method is also a widely used and efficient data-driven modeling tool for dynamic nonlinear systems and identification for its universal estimation capability to capture complex behaviors [50]. The ANN consists of a number of neurons, training functions, and the activation (transfer) functions. In figure 3.5 (a), each circular node represents an artificial neuron, and an arrow

represents a connection from an output of one neuron to an input. The input layer is composed of electrochemical features extracted from the differential analysis by Severson et al.

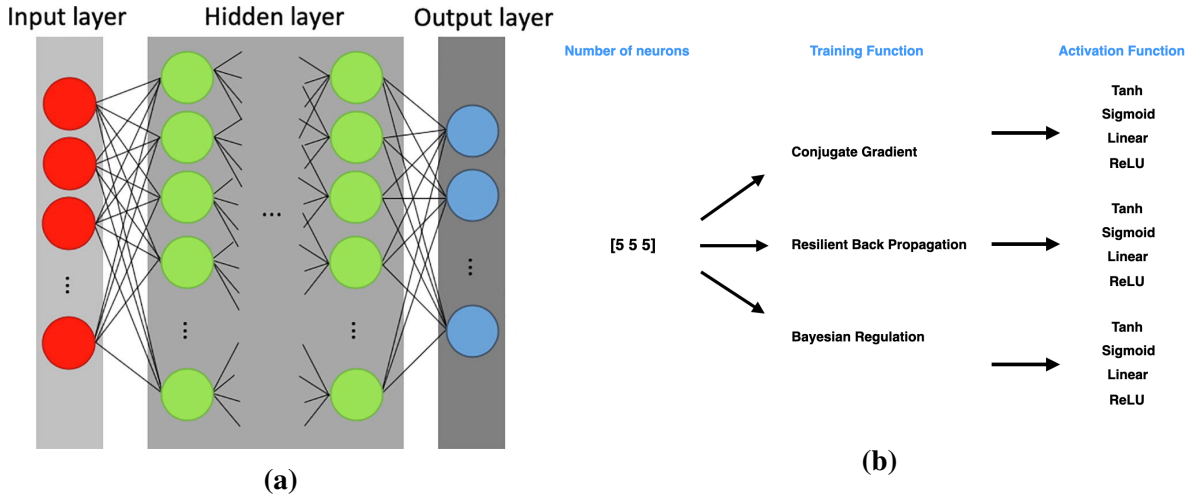


Figure 3.5. Schematics of the feed-forward neural net structure. (a) shows a general structure for a hidden layer and neural network. Figure from [36]. (b) represents the architecture of backpropagation.

This thesis implements three different optimizing methods and four different activation functions using a feed-forward net from MATLAB and the stochastic gradient descent as an optimizer. We used a three-layer fold structure with five hidden layers each in figure 3.5 (b) with the limit of 500 epoch. These parameters of the layer structure were empirically selected. Conjugate gradient backpropagation (CGP), Bayesian regulation backpropagation (BR), and resilient backpropagation (RPROP) were selected for training function. The transfer functions (activation fuction) we use are log-sigmoid (logsig), tan-sigmoid (tansig), linear transfer (purelin), and positive linear (poslin). Figure 3.6 shows the activation function for ANN training.

Bayesian regularization minimizes a linear combination of squared errors and weights. It modifies the linear combination so that the resulting network has good generalization qualities at the end of the training [26]. In MATLAB, BR updates the weight and bias values according to Levenberg-Marquardt optimization [30]. The Marquardt adjustment parameter has defaulted to 0.005. The decrease and increase factors are 0.1 and 10, respectively—the default value of the minimum performance gradient is 10^{-7}. The RPROP training algorithm eliminates the harmful

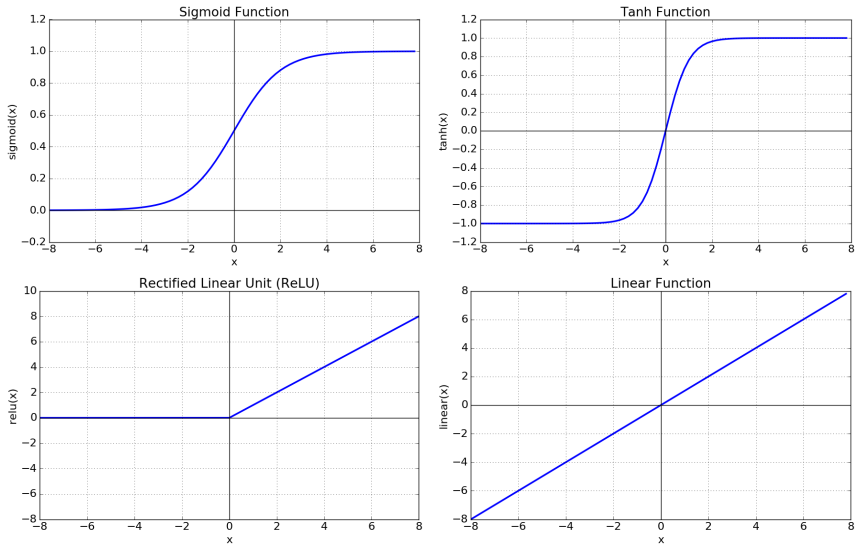


Figure 3.6. The four activation functions (transfer functions). Sigmoid, Tanh, ReLU, and linear activation functions are used for each backpropagation training function.

effects of the magnitudes of the partial derivatives. Only the sign of the derivative can determine the direction of the weight update since the magnitude of the derivative does not affect the weight update. The default value of the learning rate is 0.01. The increment in weight change and decrement in weight change are 1.2 and 0.5, respectively. The initial weight change is set to 0.07, and the maximum weight change is 50.0. In the conjugate gradient (CGP) algorithms, a search is performed along with conjugate directions with the second-order gradient. We use Polak-Ribiére variation of conjugate gradient. For the line search parameter of a conjugate method in MATLAB, we specified four parameters: the scale factor that determines a sufficient step reduction (α), the scale factor that determines a sufficiently large step size (β), a parameter to avoid small reductions in performance (γ), and an initial step size in the interval location step (δ). α, β, γ, and δ are set to 0.001, 0.1, 0.1, and 0.01, respectively. The rest of the undefined parameters in this section follows the default setting in the MATLAB net function. More detail information for training function is described in Appendix B.

3.3 Result and discussion

Battery life cycle prediction with actual data

This section will discuss the result of life cycle prediction with the methods described in section 4.2. First of all, we validated the result of Severson et al. by using the same test partitions provided from Severson et al. which were split into 40, 41, and 43 samples for test, validation, and training set, respectively. Additionally, due to a small size of training samples in Severson et al.'s partition, we randomly distributed the validation sets and training sets with a ratio of 20:80 from 50:50 to observe the performance improvement. Figure 3.7 shows an intuitive prediction result between the ANN (BR with Tanh) and the linear regression models. Being close to the centerline of scattering points corresponds to a better prediction result. Table 3.5 shows the averaged result of five runs for each linear regression and the backpropagation method.

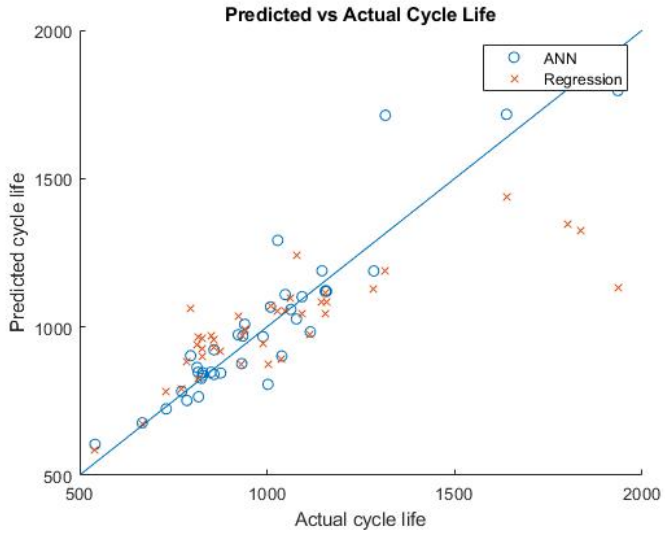


Figure 3.7. A direct comparison between linear regression and backpropagation. The blue circle shows the life cycle prediction by ANN, and the red cross represents the life cycle prediction by linear regression. The test set configuration follows Severson et al.'s specified data sets.

In table 3.5, we can see two test trials with different training and validation set ratios. Test 1 represents the result of the same partition from Severson et al. We can observe the best test performance, 9.92%, which is close to the value of the previous author provided (9.98%).

However, the validation MAPE (21.81%) is more than double of the train MAPE (9.84%). We can assume that the test sets were too easily compared to overall training as mentioned above, and the over-fitting issue might exist for the training sets. We run an additional trial (Test 2) with a different ratio by allocating more data for training. Test 2 result shows the over-fitting improved, however, the overall test error increased.

Table 3.5. Result of linear regression and ANN with Severson et al.'s specified partitions of train, test, and validation data. Test 1 has a ratio of 50:50 between valid and train set which equals to Severson et al.'s test set partitions. Test 2 has a different ratio of 20:80 between valid and train set. The unit of RMSE and MAPE are cycles and %, respectively.

Configuration	TEST		VALID		TRAIN	
ERROR	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
Test 1 ratio 50 : 50						
Regression	210	9.92	181	21.81	88	9.84
ANN (BR)	177	10.48	123	14.45	259	7.56
Test 2 ratio 20 : 80						
Regression	257	13.20	216	11.32	99	11.49
ANN (BR)	185	11.27	137	7.79	60	8.67

Furthermore, we performed the backpropagation with the specified data set. As we can see a small volume of training set in test 1 in table 3.5, the backpropagation with hidden layers does not perform well compared to linear regression. This result shows that using various hidden layers with the small training sets might not be appropriate. In test 1, we achieved an error of 10.48% and could see the prediction overfits to the training set. The validation MAPE (14.45%) is more than double of the train MAPE (7.56%). In test 1 and 2, the delta, the difference between the MAPE of training and validation set, was improved from 6.89 to 0.88. Thus, the overfitting was improved and the backpropagation exhibited an enhanced result in terms of the error compared to linear regression. In this study, we have three different train functions (RPROP, BR, and CGP) and four different activation functions (poslin, logsig, purelin, and tansig) introduced in 3.2.2. From test 1 and 2, the candidates of training functions are RPROP

and BR based on the stability and error result in figure 3.8. However, in this trial with Severson et al.'s test configuration in table 3.5, we observed that errors fluctuated only by changing the different ratio of data sets. For this reason, unlike Severson et al, we conducted further prognosis with cross-validation following table 3.4.

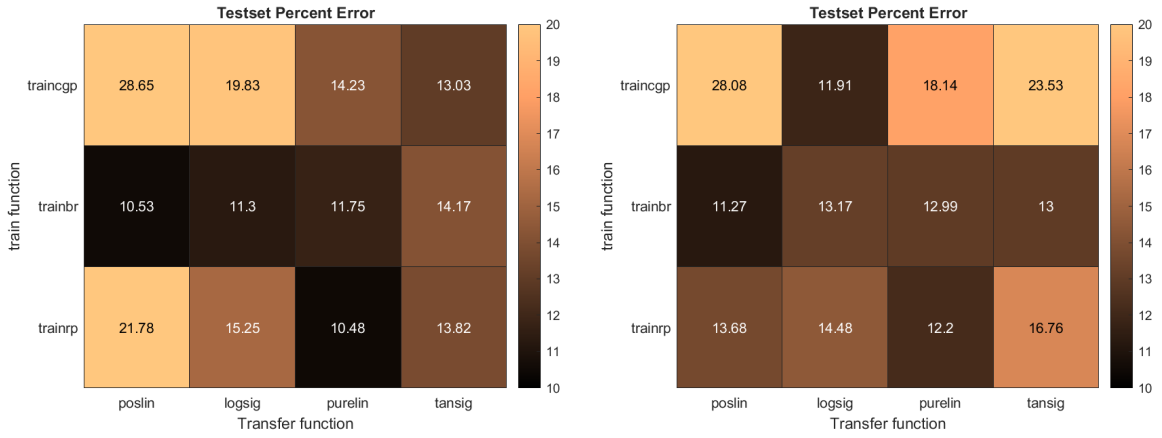


Figure 3.8. Heatmap results of MAPE using ANN networks with the specified test set configurations following table 3.5. (a) shows the MAPE for the test set 1, and (b) represents the MAPE for the test set 2 in table 3.5. The train function and transfer function are shown on each axis.

Table 3.6 shows the result of cross-validation. In comparison with table 3.5, we randomly distribute the data sets to understand the prediction performance. Two different methods (Linear regression and backpropagation) were directly compared by the same data partitions in each trial. In test mode 1, we obtained the best training RMSE, an average of 134 cycles (11.12% MAPE), and test RMSE, an average of 170 cycles (16.58% MAPE) with the linear regression. For the back propagation methods we mainly compared the best result RPORP and BR methods. The training RMSE and test RMSE for RPROP with logsig is an average of 81 cycles (7.44% MAPE) and 126 cycles (11.66% MAPE), respectively. The training RMSE and test RMSE for BR with tansig is an average of 60 cycles (5.43% MAPE) and 128 cycles (11.12% MAPE), respectively. The rest of each test mode's result is exhibited in table 3.6. Table 3.6 shows that the RPORP and BR methods result in lower errors than linear regression. Moreover, the lower MAPE standard deviation was achieved in backpropagation. The MAPE standard deviation for linear regression

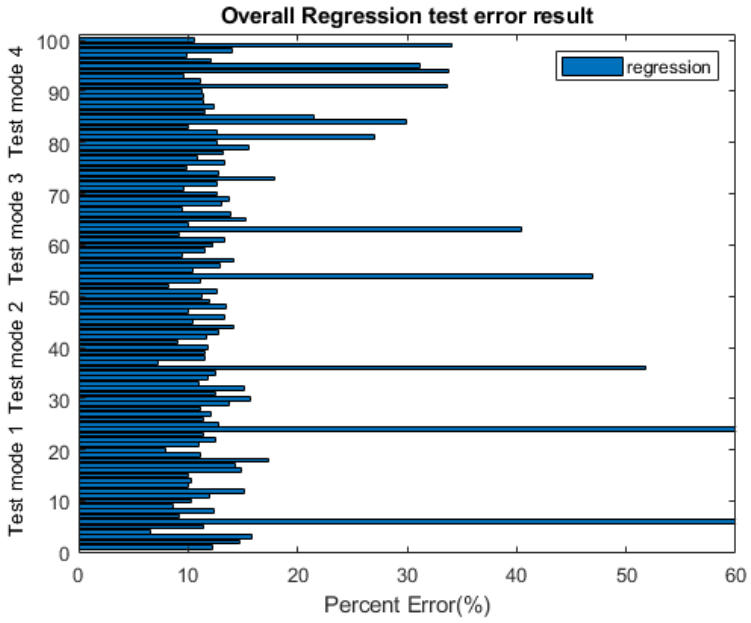
in each test mode is 14.4, 11.3, 10.1, and 9.5. In each test mode, the MAPE standard deviations for the RPORP and BR were 6.2, 5.2, 5.7, 5.1, and 4.3, 6.2, 5.7, 5.7, respectively.

Figure 3.9 (a) shows the regression test result of the overall trials. Figure 3.9 (b) shows the direct comparison by the overlaid graph between the regression and RPROP. We can observe that the linear regression has a weakness in predicting the complex test sets comparing to RPROP.

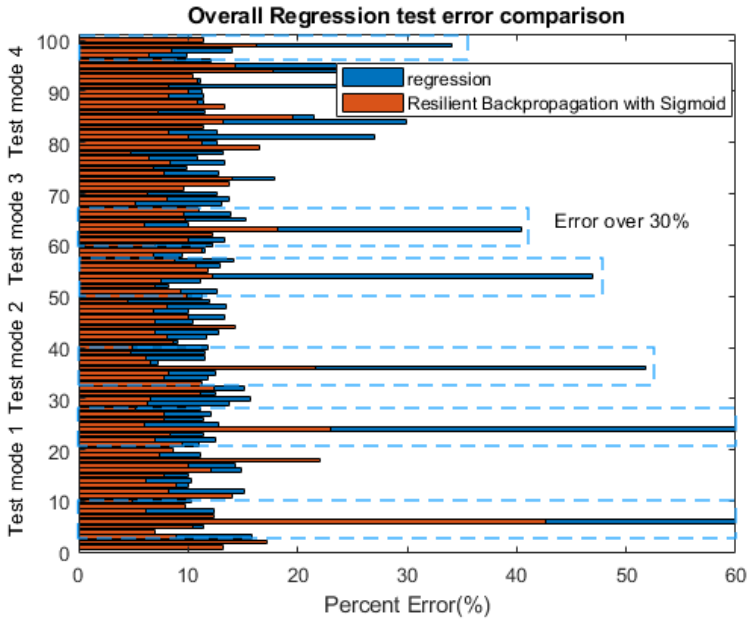
Table 3.6. Test result of the regression and backpropagation with cross-validation. The test follows test matrices in table 3.4 in 3.2.2. The provided results are the two best performance combinations of ANN from figure 3.8, the resilient backpropagation with sigmoid activation function and Bayesian regularization backpropagation with Tanh activation function. The units of RMSE and MAPE are cycles and %, respectively.

Configuration	TEST		VALID		TRAIN	
ERROR	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
Test Mode 1						
Regression	170	16.58	200	18.32	134	11.12
ANN(RPROP)	126	11.66	137	11.26	81	7.44
ANN(BR)	128	11.12	130	10.28	60	5.43
Test Mode 2						
Regression	149	13.55	196	16.44	133	11.58
ANN(RPROP)	107	8.52	129	9.72	81	7.57
ANN(BR)	110	8.78	120	8.42	65	6.51
Test Mode 3						
Regression	155	14.54	146	12.29	142	11.88
ANN(RPROP)	109	9.63	112	9.31	88	8.05
ANN(BR)	114	9.43	105	8.46	80	7.07
Test Mode 4						
Regression	171	17.00	166	18.06	143	11.54
ANN(RPROP)	123	10.89	105	9.81	97	8.36
ANN(BR)	155	11.77	100	9.02	80	6.70

In conclusion, with cross-validation, we find the regression might be weak to predict the unknown and difficult test sets while backpropagation shows a more stable prediction outcome. Still, linear regression could be powerful with a small volume of training data sets. Also, an additional regularization such as dropout can be implemented to improve the overfitting.



(a)



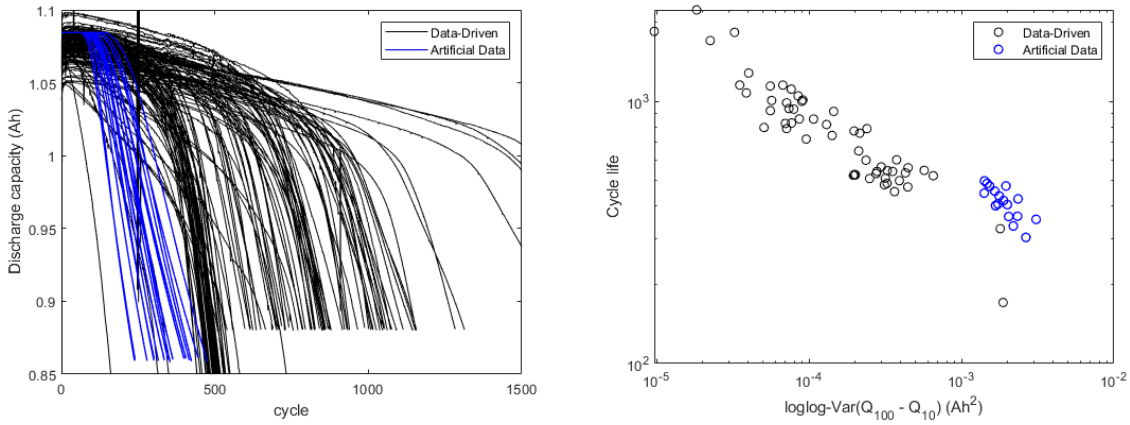
(b)

Figure 3.9. Overall test result from test modes 1 to 4. Results of 100 test sets are shown (five runs of the five-fold cross-validation for four different test modes). (a) represents the error of linear regression. (b) shows the overlaid graph comparing the linear regression and the RPORP with a sigmoid. The blue dashed box is the area where the error is over 30% which can be considered as a difficult test set. In this box, the linear regression shows a relatively high error than RPROP meaning that the linear regression has a weakness in difficult test sets than resilient backpropagation.

However, we can see that there is a possibility to use backpropagation in the battery life cycle prediction research with a sufficient amount of training data. In our experiment, the backpropagation achieved a lower prediction error when assigning more than 50 training sets out of 124 degradation data sets comparing to linear regression. Moreover, the result could be comparable with the feed-forward neural network result from Fei et al [16]. Fei et al. conducted the life cycle prediction using the same A123 data with similar test partitions as the test mode 4 in table 3.6. Fei et al. shared the optimized test result with Adam optimizer. Comparing to his result, a resilient backpropagation with a logsig activation function decreased the MAPE from 17.1% to 10.89%.

Battery life cycle prediction with artificial data

Until now, we performed the studies only using actual data. This section discusses the life cycle prediction with artificial data. In 3.1.3, we generated the discharge capacity graph by varying specific SEI kinetic constant and discussed the influence of these parameters. Figure 3.10 shows the low life cycle artificial data obtained by modifying the SEI kinetic constant. The SEI kinetic constants are randomly distributed between $4 \cdot 10^{-16}$ to $6 \cdot 10^{-16}$ for generating low life cycle degradation data. The charging policy varies following Severson et al.'s experiment policy. As we can observe a linear relation trend in the figure 3.10, we can expect an improved life cycle prediction result by adding this artificial data to training sets.

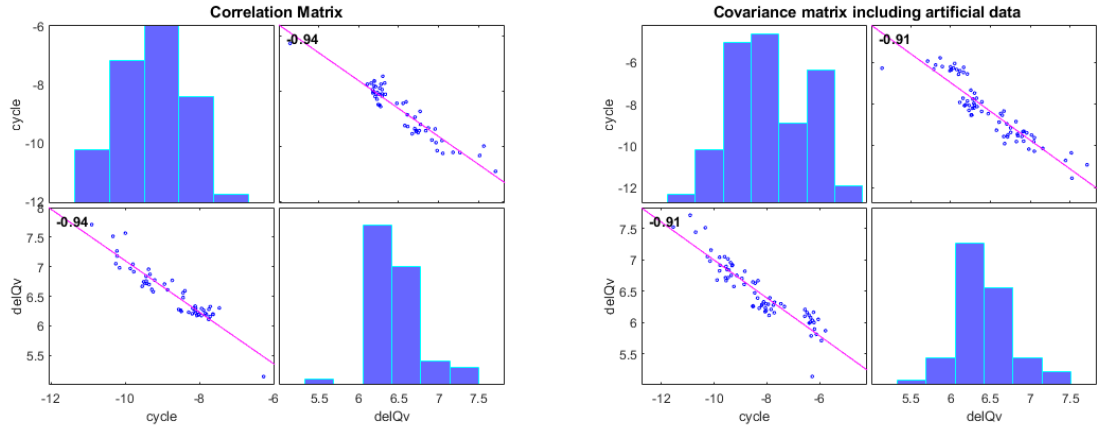


(a) **(b)**

Figure 3.10. A degradation curve with artificial data having low life cycles added. Artificial data with low life cycle was generated from varying SEI variables. The black line represents actual data, and blue line shows artificial data having low life cycles. We generated the data to improve the imbalanced data for the machine learning algorithm. (a) and (b) corresponds to the figure 3.1 (a) and (b), respectively.

Figure 3.11 shows the Pearson correlation graph between actual data and the combined data of actual and artificial data between log-based life cycle and log-based variance of the difference between the 10th cycle and 100th cycle. In figure 3.11, we can see the Pearson value, which defined as a reference point to discuss the result in this thesis, dropped from -0.94 to -0.91. Note that we generated artificial data using the single-cell model simulation, and we have

not acquired all the Severson et al.'s feature data in table 3.1. The internal resistance, internal resistance difference, and the average time in table 3.1 are not used for the training.



(a) **(b)**

Figure 3.11. Pearson correlation graphs for artificial data and actual data. Graphs show the correlation of the log-based life cycle with log variance of the difference between the 10th cycle and 100th cycle. (a) represents the result of A123's actual data. (b) represents the result of combining data (original and artificial data). The difference of the Pearson values between the two is 0.03.

We have two experiments with different artificial data generation sets. Each of experiment has 10% of 124 data sets and 20% of 124 data sets added to training sets. Table 3.7 shows the directly compared result of the battery life cycle prediction between actual data and artificial data using linear regression. We used the test matrices covered in table 3.4. Tests with four different partitions were conducted to check the influence of the training data size on the test result. By adding artificial data, we observed the averaged error reduction of 0.49 - 1.34% with boost (10%) and found the averaged error reduction of 0.37 - 1.32% with boost (20%) in table 3.7.

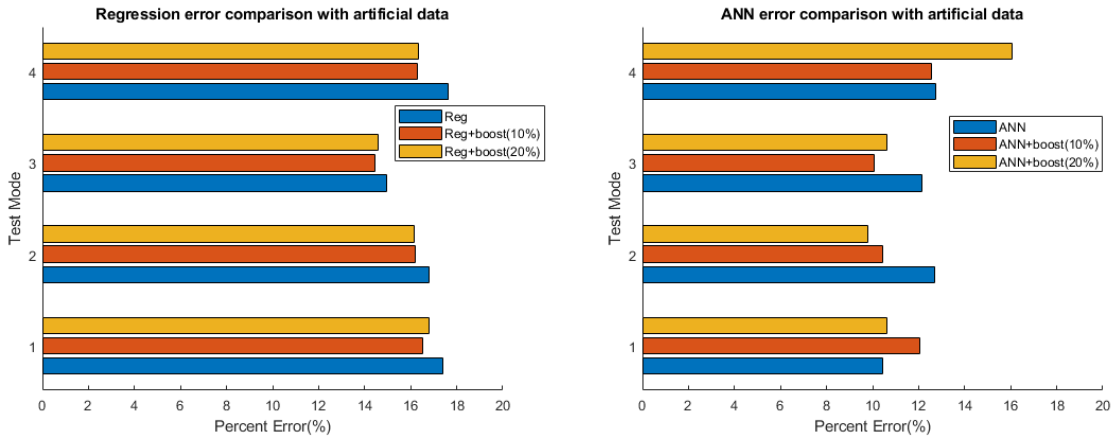
Figure 3.12 compares the averaged five trials of actual data and two generation sets of artificial data (10% and 20%). Figure 3.12 (a) and (b) shows the result with the linear regression and resilient backpropagation, respectively. All results of test modes were improved with linear regression while less precise results were obtained with resilient backpropagation. Figure 3.13

Table 3.7. Test results of regression with artificial data. Two experiments with two different ratios of data augmentation were performed. Each of 10% and 20% of the total datasets was added to the training set to form augmentation of artificial data. The units of RMSE and MAPE are cycles and %, respectively.

Configuration	TEST		VALID		TRAIN	
ERROR	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
Test Mode 1						
Regression	158	17.41	173	22.18	153	13.30
Boost (10%)	162	16.51	164	21.08	141	13.72
Boost (20%)	147	16.78	165	21.32	135	13.00
Test Mode 2						
Regression	164	16.81	149	16.22	157	13.49
Boost (10%)	163	16.22	138	15.24	141	13.88
Boost (20%)	152	16.14	140	15.40	135	13.19
Test Mode 3						
Regression	164	14.95	192	14.37	151	13.46
Boost (10%)	156	14.46	162	13.46	138	13.55
Boost (20%)	151	14.58	164	13.61	132	12.82
Test Mode 4						
Regression	187	17.64	190	18.29	148	13.06
Boost (10%)	173	16.27	173	16.97	135	13.16
Boost (20%)	167	16.32	175	17.21	128	12.36

represents all trial results from figure 3.12. The results were sorted into two classes. One is when the original error is above 15% , and the other is when the original error is below 15%. In figure 3.12 (a), we found the error was mainly improved when the error was over 15% in linear regression with actual data. However, no distinct characteristic was observed when using resilient backpropagation.

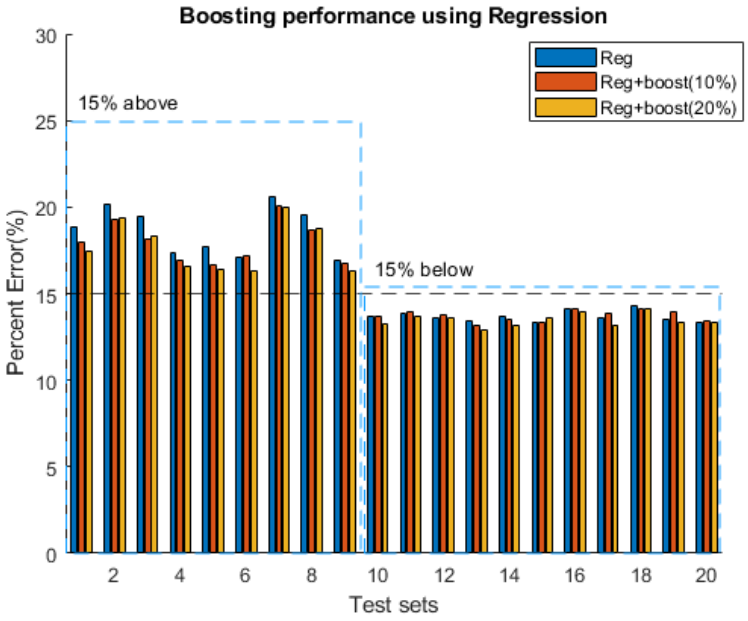
This section analyzed the effect of artificial data using the linear regression and resilient backpropagation. We observed the improvement in linear regression with artificial data; however, we found an unstable prediction in resilient backpropagation. We assume that this result comes from insufficient information from artificially generated data. As discussed in 3.1.3, the internal resistance, internal resistance difference, and the average time in feature extraction in table 3.1 are not used in degradation prediction in this section. This is because the current simulation



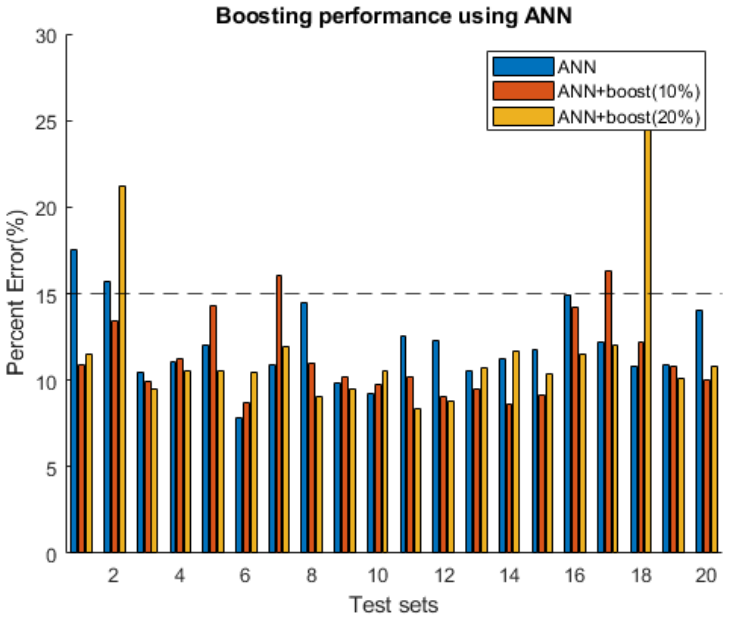
(a) **(b)**

Figure 3.12. Averaged error bar plots for each test mode between actual data and artificial data. See table 3.4 for detailed data partitions for each test mode. (a) represents the result of linear regression. (b) represents the result of resilient backpropagation. We observed the overall improvement in regression with artificial data and less precise results with backpropagation.

do not consider the parallel battery charging and discharging policy, which the actual battery followed. However, we used the same feed-forward network architecture to directly compare actual data test and artificial data test with reduced features. The averaged backpropagation prediction error increased by 1 - 2 %, comparing to the result of life cycle prediction with actual data from table 3.6, which can be interpreted as having a not efficient architecture with reduced feature. Furthermore, since we only used five out of eight features from table 3.1, we also noticed a lack of training data to use backpropagation. Linear regression, however, was mainly impacted by the linear appearance in discharge capacity in figure 3.10 corresponding to Severson et al.'s algorithm that selects the most-influencing six features out of eight. To sum up, the efficiency of backpropogation is proportional to the number of the features while the number of features for linear regression is limited. In the future, we can perform additional studies to overcome the current limit but first we need to optimize the feed-forward layer architecture for the backpropagation with artificial data. Second, we can generate artificial data with battery pack-level simulation (parallel charge and discharge) with high computing power to obtain internal resistance and average time data.



(a)



(b)

Figure 3.13. Sorted bar plots of the percent error for regression and backpropagation using actual and artificial data. (a) shows the bar plot sorted into two performance levels. One has over 15% original error. The other is below 15% original error. (a) represents the result of the regression method. When the original regression had a higher error (over 15%), the boost decreased the percent error of the regression model; however, the boost had no impact on the model performance in the lower percent error class. (b) shows the direct comparison of the result of the resilient backpropagation. Unlike figure 3.13 (a), no distinct feature was observed.

Chapter 4

Conclusion and future work

The growing demand for electrical devices and increasing need of power infrastructure are propelling the growth of the energy storage industry. Abundant battery models are available in the market, however, the unknown and non-linearly degrading behavior of a second life battery creates a big hurdle to reuse it for a new application. Therefore, this thesis focused on the overall procedure of a life cycle prediction using two different battery models considering both the electrochemical and data-driven models. We first selected a 21700 cylindrical battery cell (LGM50) with published data of its detailed battery cell structure. Then we introduced the design optimization approach to estimate inside of the battery cell and compared it with its actual data in Appendix C. By using optimization, we reduced the time of iterations for estimating unknown battery cell structures and performances. On the basis of the estimated structure of a battery, we discussed the methods of predicting battery life cycle with data-driven model and generated artificial data.

In the beginning, we only used actual data to validate the linear regression and three different backpropagation methods. We noticed the necessity of additional prognostic and conducted cross-validation with four different test matrices. Resilient backpropagation with a sigmoid activation function performs well compared to linear regression from the previous author. According to the experiment, backpropagation generally takes more computing time than linear regression; however, it could improve battery life cycle prediction. Lastly, we implemented

the life cycle prediction by adding two different ratios of artificial data. With 10% of artificial data, the linear regression was improved by 0.49% - 1.34%. With 20% of artificial data, the error was improved by 0.37% - 1.32%. We observed that the error was mainly enhanced when the regression with actual data was over 15%. However, we encountered a larger and unstable prediction error with the backpropagation by adding artificial data.

Currently, two distinct predicting battery degradation approaches are widely used for lifetime prediction: the data-driven and physical models. Data-driven methods find a correlation between input and output, not depending on the underlying physical understanding. On the other hand, physical models, which are constructed from physical principles, require relatively less experimental data than the data-driven methods. Unfortunately, neither of these strategies can accurately predict the life cycle of a complex battery alone. Therefore, a hybrid approach combining the data-driven and the physics-based approach is needed. For the second-life batteries, it is even harder to get complete information on the inside of the battery's physical structure, and it takes significant time to experiment with the data-driven model. Therefore, with more studies, it is desired to reduce the duration of data-driven experiments for life cycle simulation, and the advancements in the battery life cycle prediction can bring a more practical design, cost-effective maintenance, and additional opportunities for second-life battery applications.

4.1 Future work

This thesis addresses a potential approach for the better understanding of the second life battery degradation and future usage. However, still, there are many limits we can overcome, and more studies are required. First, the optimization constraints could be more realistically considered such as cost, power requirement, and manufacturability to better estimate the inside of the unknown battery cell. For example, the battery cell's power density should be appropriately selected because electronic devices have a specific power requirement that must be satisfied. Likewise, the material cost would also be a critical factor to design a battery model.

Therefore, optimizing a specific energy density while maintaining a specific power requirement and estimated cost requirement would improve the battery cell structure estimation. Moreover, some design variables, such as a particle size of cathode materials, are not used due to the computing efficiency according to the analysis of variance from Xue et al. [56]. However, those design variables have the possibility to contribute the battery degradation. Thus, a more accurate estimation of the cell can be achieved by adding aforementioned design variables with high computing power. Second, we simulated the artificial degradation data using a single cell-based model by assuming SEI kinetic constant in 3.1.3. It is currently not generalized and conducted how to measure the SEI parameters during the aging cycle even though we understand the SEI growth is the primary factor for battery degradation. Thus, the life cycle prediction could be improved by parameterizing the SEI variables and using actual SEI parameters. Lastly, we observed a limitation for backpropagation due to a lack of artificial data such as the internal resistance and average charge time due to the simulation limit. With high computing power, generating parallel charge-discharge simulation would enhance the life cycle prediction using artificial data.

Appendix A

Physical battery model parameterization

Parameterization for the battery architecture is not a simple task, and the tear-down process of a cell have a possibility to damage the cell component. Chen et al. shared the parameterization protocols for the LGM50 21700 commercial cell and published battery parameters[8]. He introduced analytical methods to determine information on the geometry, chemistry, and electrode microstructure. Equations of the SPMe model parameterization are shown in figure A.1. The Python Battery Mathematical Modeling (PyBaMM) package [44] was used to solve these SPMe model simulations, and Dualfoil.py [11] from John newman was modified to use in this thesis. This open source can solve continuum models for batteries using both numerical methods and asymptotic analysis. A more detailed way to parameterize the battery for an electrochemical model can be obtained from [8] [1][38].

Description	Equation	Boundary conditions
	Electrodes	
Mass conservation	$\frac{\partial c_{s,k}}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 D_{s,k}\frac{\partial c_{s,k}}{\partial r}\right)$	$\left.\frac{\partial c_{s,k}}{\partial r}\right\|_{r=0} = 0, -D_{s,k}\left.\frac{\partial c_{s,k}}{\partial r}\right\|_{r=R_k} = \frac{J_k}{a_k F}$
Charge conservation	$\frac{\partial}{\partial x}\left(\sigma_{s,k}\frac{\partial \phi_{s,k}}{\partial x}\right) = J_k$	$-\sigma_{s,n}\left.\frac{\partial \phi_{s,n}}{\partial x}\right\|_{x=0} = -\sigma_{s,p}\left.\frac{\partial \phi_{s,p}}{\partial x}\right\|_{x=L} = i_{\text{app}}$
		$-\sigma_{s,n}\left.\frac{\partial \phi_{s,n}}{\partial x}\right\|_{x=L_n} = -\sigma_{s,p}\left.\frac{\partial \phi_{s,p}}{\partial x}\right\|_{x=L-L_p} = 0$
	Electrolyte	
Mass conservation	$\varepsilon_k \frac{\partial c_{e,k}}{\partial t} = \frac{\partial}{\partial x}\left(\varepsilon_k^b D_e \frac{\partial c_{e,k}}{\partial x}\right) + (1 - t^+)\frac{J_k}{F}$	$\left.\frac{\partial c_{e,n}}{\partial x}\right\|_{x=0} = \left.\frac{\partial c_{e,p}}{\partial x}\right\|_{x=L} = 0$
Charge conservation	$\frac{\partial}{\partial x}\left(\varepsilon_k^b \sigma_{e,k}\left(\frac{\partial \phi_{e,k}}{\partial x} - \frac{2(1 - t^+)R\,T}{F}\frac{\partial \log c_{e,k}}{\partial x}\right)\right) = -J_k$	$\left.\frac{\partial \phi_{e,n}}{\partial x}\right\|_{x=0} = \left.\frac{\partial \phi_{e,p}}{\partial x}\right\|_{x=L} = 0$
	Reaction kinetics	
Butler–Volmer	$J_k = \begin{cases} a_k j_{0,k} \sinh\left(\frac{1}{2}\frac{R\,T}{F}\eta_k\right), & k \in \{n, p\}, \\ 0, & k = \text{s}. \end{cases}$	
Exchange current	$j_{0,k} = k_k \sqrt{c_{e,k} c_{s,k} (c_{s,k}^{\max} - c_{s,k})}\|_{r=R_k}$	
Overpotential	$\eta_k = \phi_{s,k} - \phi_{e,k} - U_k(c_{s,k}\|_{r=R_k}), \quad k \in \{n, p\}$	
	Initial conditions	
Initial conditions	$c_{s,k} = c_{k0}$, $c_{e,k} = c_{e0}$	
	Terminal voltage	
Terminal voltage	$V = \phi_{e,p}\|_{x=L} - \phi_{e,n}\|_{x=0}$	

Figure A.1. Description of the SPMe model. Figure from [8]

Appendix B

Optimization methodology

Optimization methods are used in diverse areas of study to find the best result such as maximizing the volume or minimizing the loss. The advantage of the optimization process lies in the ability to quickly approach to the solution for the multiple alternative designs. The optimization has three major fundamentals. The first factor is an objective function, which we would like to maximize or minimize. The second element is a collection of variables, which are quantities whose values can be manipulated in order to optimize the objective. The third element is a set of constraints, which are restrictions on values that variables can take. Before we solve the optimization problem, we need to define the problem first for efficiency. Generally, the optimization methods are classified by objective functions, design variables, and constraints, like in Table B.1 [28].

The design optimization algorithm can be broadly divided into two categories, gradient-based and gradient-free optimization according to the encountered problem that needs to be optimized. Gradient-based optimizers are advantageous to finding local minima for high-dimensional, nonlinear, and convex problems. However, most gradient-based optimizers have difficulties in solving noisy and discontinuous functions and discrete design variables. On the other hand, gradient-free methods can resolve the previously mentioned problems despite the lower efficiency and accuracy than gradient-based methods. In addition to that, these methods can find multiple local optima while searching for the global optimum.

Table B.1. Classification of the optimization problems

Classification	Objective function	Design variables	Equality constraints	Inequality constraints
LP	linear	continuous	linear	linear
QP	quadratic	continuous	linear	linear
NLP	nonlinear	continuous	nonlinear	nonlinear
ILP	linear	integer	linear	linear
MILP	linear	mixed	linear	linear
MINLP	nonlinear	mixed	nonlinear	nonlinear

Many of the existing problems are considered as nonlinear problems. In a battery pack design, the problem presented in this study is classified as MINLP in table B.1 since the number of the cell in the pack is an integer, which makes us consider the design variable as mixed. A general nonlinear optimization problem can be written as below where $f(x)$ is the objective function to be optimized with respect to the bounded variables x, subject to equality constraints $c(x)$ and inequality constraints $\hat{c}(x)$. Both gradient-based and gradient-free optimizers can solve the battery pack optimization problem. Appendix C uses gradient-free optimization to solve the MINLP problem by searching the global minimum with mixed design variables.

$$Minimize f(x),\ f : \mathbf{R^n} \rightarrow \mathbf{R}$$

$$subject\ to$$

$$x_{min} \leq x \leq x_{max}$$

$$c_j(x) = 0,\ where\ j = 1, \cdots, m$$

$$c_k(x) \leq 0,\ where\ k = 1, \cdots, \hat{m}$$

Machine learning algorithms can also be considered as optimization algorithm in terms of finding the minimum loss. We use the backpropagation (BP) algorithm for the battery life cycle prediction in chapter 3. BP is a widely used algorithm for training feedforward neural networks. In accommodating a neural network, it calculates the gradient of the loss function concerning the network's weights for a single input-output example efficiently. This efficiency makes it feasible to use gradient methods for training multilayer networks updating weights. The

backpropagation algorithm works by computing the gradient of the loss function for each weight by the chain rule, computing the gradient one layer at a time, and iterating backward from the last layer to avoid redundant calculations of intermediate terms in the chain rule [22].

Backpropagation is the algorithm to train a neural network to produce outputs close to those given by the training set. It consists of the following contents. Many supporting algorithms exist, such as resilient backpropagation and bayesian regulation, since it is easy to fall into the trap of abstracting away the standard learning process.

Back propagation process:

1. Calculating the outputs based on inputs and a set of weights (forward pass)
2. Comparing these outputs to the target values via a loss function
3. Calculating gradients to reduce the loss (backward pass)
4. Changing the weights according to the gradients

The below equation is the matrix form of the standard backpropagation algorithm from Rumelhart, Hinton, and Williams. A more detailed explanation can be attained from [42]. k is the number of output units in the last layer. E(w) is the loss function.

$$\nabla W_k E(W) = \delta_k h_{k-1}$$

where $h_0 = x$

$$\delta_k = [\frac{\partial g(a_k)}{\partial a_k}]M$$

if k is an output layer : $M = y - d$

otherwise : $M = W_{k+1}\delta_{k+1}$

x is a set of training input vectors. d is a set of desired target vectors. y is a set of all predictions. $a_k = W_k h_k$ is a set of all weighted sum. $h_k = g(a_k)$ is a set of all hidden outputs. W_k is a set of weights at layer k.

In chapter 3, we introduce three different backpropagation methods for the neural network: Resilient backpropagation (RPROP), Conjugate gradient backpropagation with Polak-Ribiere (CGP) updates, and Bayesian Regulation backpropagation (BR). The RPROP training algorithm eliminates these harmful effects of the magnitudes of the partial derivatives. Only the sign of the derivative is used to determine the direction of the weight update; thus, the magnitude of the derivative does not affect the weight update [39]. CGP is one of the conjugate gradient methods, which is a mathematical way to optimize linear and nonlinear systems [39]. It is an efficient way to deal with the symmetric and sparse linear system. This neural network is a mathematical process that converts a nonlinear regression into a statistical problem in the manner of a regression [7]. BR minimizes a linear combination of weights and squared errors. It modifies the linear combination so that the resulting network has good generalization qualities at the end of the training process.

Appendix C

Battery cell design optimization approach to generate artificial data in the second life battery

C.1 Single cell design optimization

In a real market, diverse batteries exist for their purposes. One company makes a battery considering the energy capacity most while another focuses on a lower cost. It might be possible to know which kind of chemical materials are used for a battery cell. However, it is hard to acquire all the cell's physical structural information, such as the thickness of anode, cathode, and porosity. Likewise, diverse aging mechanisms, manufacturing variability, and lack of actual battery performance information remain significant challenges to improve the battery degradation prediction in the second life battery. This appendix shows the estimating process for second-life batteries using design optimization as a tool in the electrochemical physical model. We used the most widely used and well-known 21700 battery cell to estimate the inside structure and compare with the original structure to examine the method. We will maximize the specific energy density as our objective and constrain the battery design to examine the optimization approach. To estimate the battery physical structure, this objective can be adjusted and the constraints can be added.

C.1.1 Governing equation and objective

The main objective of Appendix C is to maximize the specific energy density while constraining the battery volume to optimize the battery cell structure. We will examine the optimization techniques in Appendix C and will reduce iteration process to estimate the battery cell structure. As we discussed in chapter 2, there are many design variables in one single cell physical model. Kim et al. and Xue et al. conducted the analysis of variance (ANOVA) to select a high influencing design variable for a computing efficiency [56][24]. We adopted important design variables for optimizing a specific energy density, which is covered in table C.2.

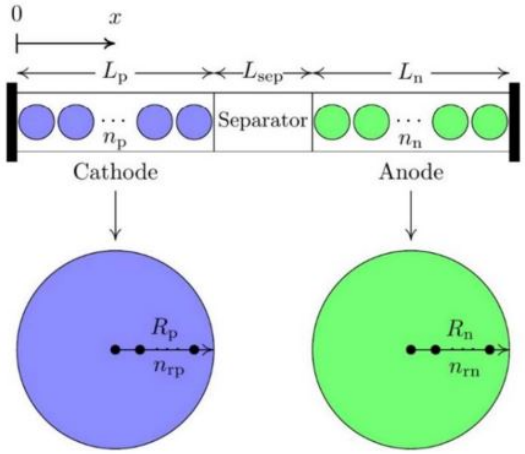


Figure C.1. An illustration of a single battery cell structure for the SPMe model. Figure from [1].

$$Find : x1, x2, x3, x4, x5$$

$$E_{specific} = \frac{1}{M_{cell}} \int_0^{t_{end}} V_{cell} \cdot i_{cell} dt \approx I \sum_{k=0}^{n} \frac{V_{k+1} + V_k}{2} (t_{k+1} - t_k)$$

$$P_{specific} = \frac{1}{M_{cell}} \int_0^{t_{end}} V_{cell} \cdot i_{cell} dt \cdot \frac{1}{t_{end}}$$

Table C.1. Essential parameters for LGM50 battery cell structure. LGM50 uses graphene for the anode material, ceramic polyolefin for the separator, and LiNi for the cathode. Detailed and overall parameters are in [8]

Battery cell parameter	Unit	Anode	Cathode	Separator
Material		Graphite	LiNix-y	polyolefin
Electrode density	kg/m^3	1657	2341	397
Electrode thickness	μm	85.2	75.6	12
Electrode porosity		0.25	0.33	0.47
Collector thickness	μm	12	16	0.4
Collector density	kg/m^3	8960	2760	
Particle size (radius)	μm	5.86	5.22	
Overall length	cm	158		
Overall width	cm	6.5		

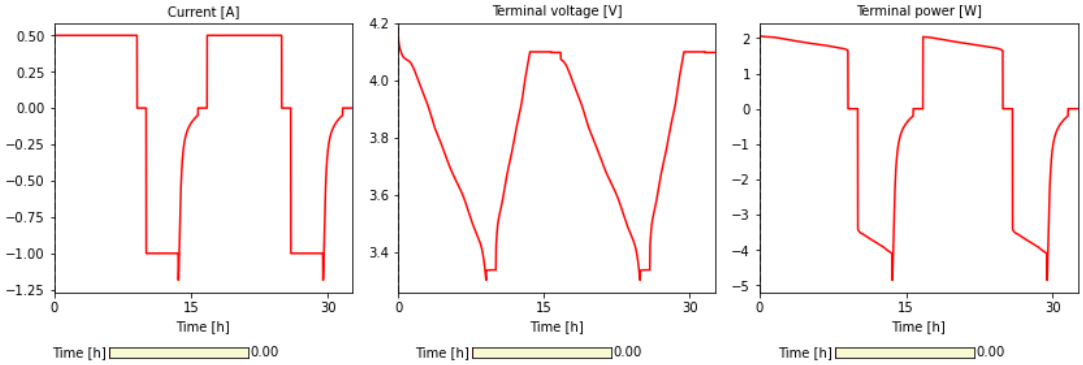
$$M_{cell} = M_{cathode} + M_{separator} + M_{anode}$$

Figure C.1 shows a simple schematic of the SPMe model for the optimization governing equation example. Design variables included an anode thickness (x1), cathode thickness (x2), anode porosity (x3), cathode porosity (x4), and separator thickness (x5). These design variables were chosen as the most significant factors affecting the specific energy capacity. Kim et al. came to this conclusion using the ANOVA [24]. The cell's energy density is calculated with the help of trapezoidal approximation to the integral over-discharge time. Table C.2 shows the design and constrained variables, and table C.1 represents LGM50 battery cell parameters.

Table C.2. Selected design variables for simplifying the battery optimization [56]. Cell thickness is the constrained design variable from the cylindrical 21700 battery cell design.

Design variables	Unit	Lower limit	Upper limit
Anode thickness	μm	60	140
Cathode thickness	μm	60	140
Anode porosity	-	0.2	0.6
Cathode porosity	-	0.2	0.6
Separator thickness	μm	10	30
Cell thickness(d_{ascs})	μm	205	

Using the process denoted above, we calculated the specific energy density using input current and the terminal voltage obtained from the SPMe model. The trapezoidal approximation used for calculation. An example of simulation data is plotted in figure C.2. For the cycling experiment in this section, the following steps are implemented: 1) Discharge a battery at C/10 for about 10 hours or until reaching 3.3V, 2) Rest for 1 hour, 3) Charge a battery at C/10 until 4.1V, and rest for 1 hour. The cell performance is obtained through galvanostatic discharge until reaching a cutoff voltage of 3.0V. After generating the data, we used these data to calculate the governing equation covered in C.1.1 and iterated the following particle swarm optimization method covered in C.1.3.



(a) Input current **(b)** Open Circuit Voltage **(c)** Terminal Power

Figure C.2. The result of the battery cell cycle simulation following the discharge-charge cycle step. Two cycles of simulation are presented. The terminal voltage and terminal power calculated from SPMe battery model along with the input current cycles.

C.1.2 Design variables and constraints limits

In this section, the realistic design limits are considered. Currently, there are four typical types of lithium-ion manufacturing processes, including single-sheet stacking, Z-stacking, cylindrical winding, and prismatic winding process as shown in figure C.3.

For single-sheet stacked cells (Fig C.3a), the stacks of sheet separators and sheet electrodes are alternately stacked one on top of the other. The four edges of the stacked cell without confinement increase the chances of cell shorting caused by Li deposition on the sides. The Z-

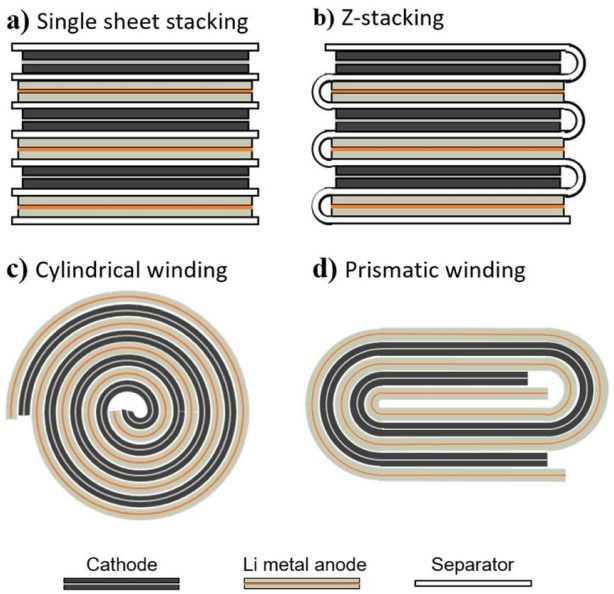


Figure C.3. Schematics of four typical types of Li metal battery manufacturing processes. (a) Single sheet stacking, (b) Z-stacking, (c) Cylindrical winding, and (d) Prismatic winding. Currently, the cylindrical winding is widely used in the EV industry. Figure from [53].

stacking (Fig C.3b) generates less stress and enhances a uniform stress distribution in the stacked cell. As the Z-stacking process continuously feeds the separator in a z-style folding pattern while adding sheet electrodes in a discrete location, only the top and bottom of stacked cells are open, and separators wrap the other two sides. The jelly roll from the cylindrical winding (Fig C.3c) and prismatic winding (Fig C.3d) usually have internal stress resulting from winding tension, which may induce cell deformation during the repeated cycling [53]. However, the cylindrical cell is one of the most widely used packaging styles for primary and secondary batteries. The advantages are ease of manufacture and good mechanical stability. Also, the jellyroll gives heat energy advantages such as a thermally balanced battery performance. Therefore, we chose the cylindrical design for the optimization.

A 21700 battery has a cylindrical cell-based structure, and it refers to a cell casing with a diameter of 21mm and a width of 70mm (Fig C.4b). In our battery optimization problem, we must ensure that the layers of electrodes and separators fit within the casing. The configuration of layers is shown in figure C.4c. Once these layers are stacked, they are wound around a cell

core with a diameter (D_{core}) of 4mm, which is depicted in figure C.4a. When the electrodes are wrapped around the cell core, they have the shape of an Archimedean spiral with a length that equals to the cathode length. We chose to have at least 0.5mm of clearance from the inner edge of the casing to ensure that the electrodes and separators would fit within the casing. This means the maximum diameter (D_{max}) is 20mm (Fig C.4a). With this dimensional requirement, we can calculate the maximum usable volume [53]. Once the maximum usable volume (V_{usable}) is calculated, the overall thickness d_{ascs} (anode, cathode, separator one, separator two) can be calculated by dividing V_{usable} by the effective area of the electrodes, $Area_{electrode}$ (Fig C.4c).

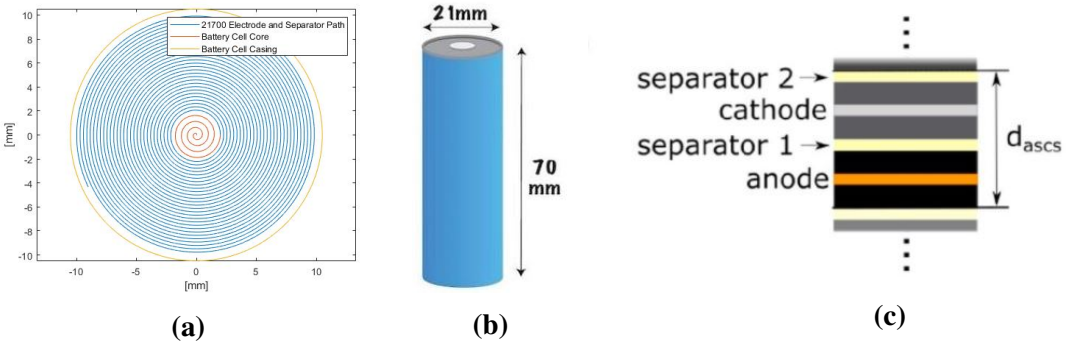


Figure C.4. The 21700 (LGM50) cylindrical cell design. a) shows a cross-sectional view of battery contents. b) represents the width (h) and diameter (D) of the battery casing. c) corresponds to the overall thickness d_{acsc} which is the sum of the thicknesses of the electrodes and separators [51]. The two separators are included in one thickness due to cylindrical winding.

To avoid high manufacturing costs, we set the lower limit of the electrodes to $60\mu m$. The upper limit of the electrodes, porosity of the electrodes, and the thickness of the separator were chosen based on previously reported data of actual battery dimensions.

$$Volume_{usable} = \pi w \frac{(D_{max}^2 - D_{core}^2)}{4}$$

$$d_{ascs} \leq \frac{Volume_{usable}}{Area_{electrode}}$$

C.1.3 Particle swarm optimization

In this paper, the battery pack design problem is a nonlinear mixed-integer problem with nonlinear constraints. The optimizer must handle a discontinuous constrained design space and efficiently manage nonlinear problems. Particle Swarm Optimization (PSO) algorithms in figure C.5 are simple yet robust; they are easy to implement for these kinds of problems due to their computing efficiency. PSO algorithms can also perform well without a large population, and each particle retains knowledge of reasonable solutions. For these reasons, we decided to use the PSO algorithm for this optimization problem [40]. PSO is a gradient-free and population-

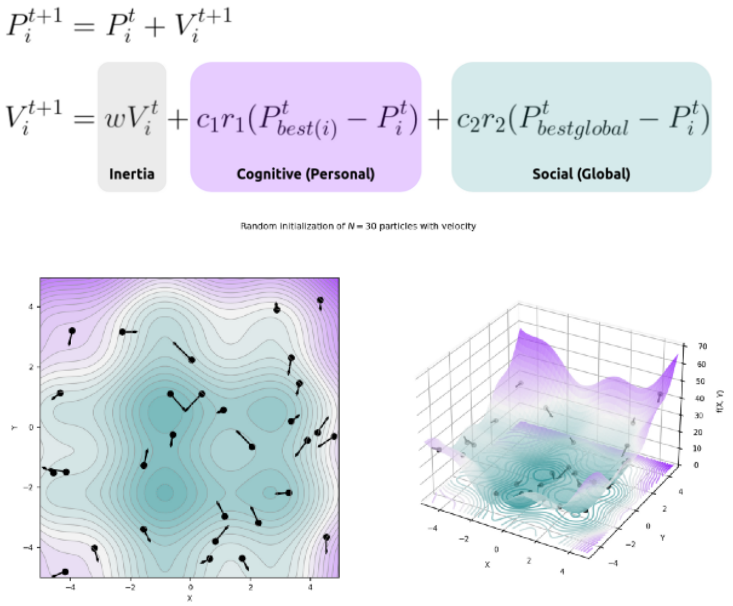


Figure C.5. Particle swarm optimization (PSO) equations and visual representation for the particle swarm. PSO consists of inertia, cognitive, and social section to search optimized variables. Figure from [40].

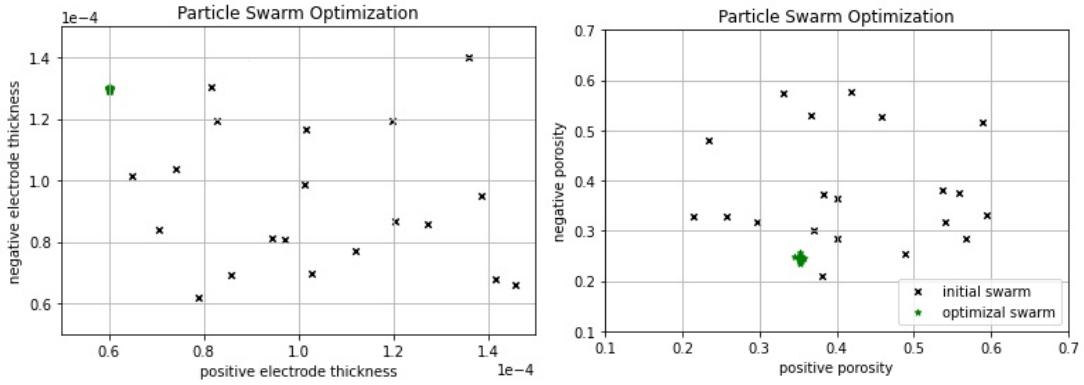
based algorithm that is based on the concept of applying swarm intelligence. Design points are interpreted as particles with cognitive and social characteristics that search for the best solution. The cognitive (c_1) and social (c_2) parameters in the PSO algorithm also provide the opportunity to tune global and local search behavior. Unlike genetic algorithms, each particle is directly updated and influenced by its local best-known position. Each particle is also affected to move

towards the best-known position in the search space. The initial swarm is randomly selected, and the swarm particle position (P) is updated by the velocity (V). In a constrained optimization problem, the PSO algorithm will penalize particles for avoiding the infeasible space [48].

C.2 Result and discussion

C.2.1 Optimization result

Figure C.2 shows the input current, open-circuit voltage, and terminal power. We can calculate the power and specific energy density from this data. Figure C.6 shows the particle convergence, which used the specific energy density as the objective. By trial and error, we selected values of 0.7 for inertia (ω), 0.5 for c_1, and 0.5 for c_2 for the PSO algorithm. This ω , c_1, and c_2 correspond to formulas shown in C.1.3 optimization methods.

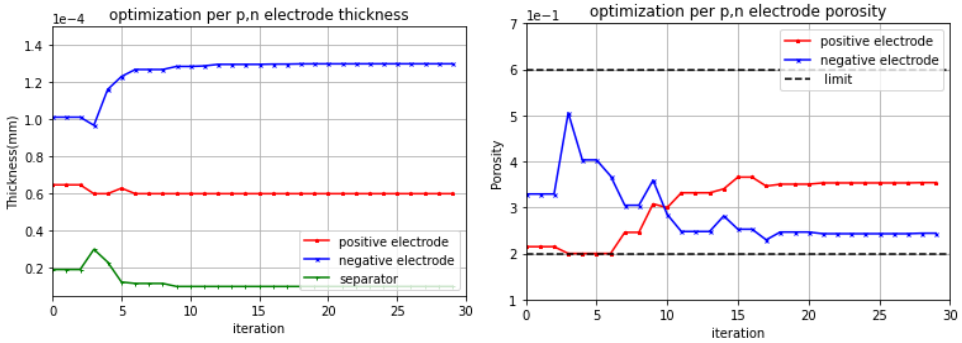


(a) electrode thickness **(b)** porosity

Figure C.6. The result of the particle swarm optimization. (a) describes negative and positive electrode thickness, and (b) shows negative and positive porosity. Black mark shows the initial swarm particle position. Green mark represents points where the swarm converges into.

The commercially manufactured LGM50 battery has a thickness of 85.2μm for an anode and 75.6μm for a cathode. Our results show that the optimal thickness of the anode and cathode are 130μm and 60μm, respectively; the sum of the electrode thicknesses has increased. This aligns with the findings of Mei et al., which indicate that thicker electrodes provide higher energy density [29]. The results in this paper indicate that the optimal ratio between the thickness of the

anode and cathode is 2:1.



(a) **(b)**

Figure C.7. The result of the optimized electrode thickness and porosity. (a) shows the thickness of the positive (p) and negative (n) electrode changes during optimization. (b) indicates the electrode positive and negative porosity changes. Separator thickness moves to the lower limit while n electrode moves to the upper limit.

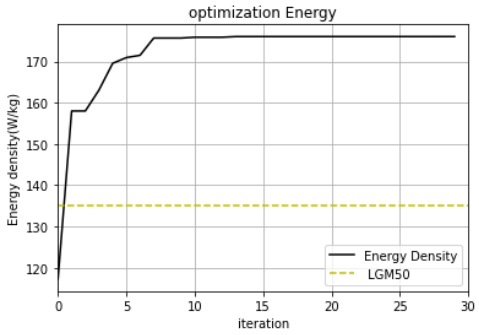


Figure C.8. The result of the maximized specific energy density compared with the original parameters.

In figure C.7, the separator thickness decreased to $10\mu m$. As the thickness of the separator increases, both of the distances (the cathode to end of the separator, the anode to end of the separator) increase which negatively affects the specific energy density [24]. We also found that the optimized porosity of the cathode and anode are 0.25 and 0.35, respectively, and these values are close to the lower boundary. Theoretically, a decrease in the porosity increases the amount of active material within an electrode, which results in energy increase [24].

We used 30 swarm particles per iteration and chose the best swarm particle for the plot in figure C.7. Figure C.8 shows that the energy density was increased by 28% (from 137 to 176)

while satisfying the volume constraints. The optimization converged after 25 iterations with a tolerance of 10^{-14}. The optimal solution satisfies the overall cell thickness constraint. This result only considers the specific capacity, not considering the power requirement for the battery cell and the estimated cost.

C.2.2 Artificial data generation with the optimized simulation result

In this section, with the optimized battery structure in C.1.4, we generated the artificial data with five experimental cycles (Fig. C.9) and the Urban Dynamometer Driving Schedule (UDDS) federal driving cycle (Fig. C.10). Figure C.9 shows the changes in the open-circuit voltage (OCV) with respect to current cycles. Figure C.10 represents results from the UDDS drive cycle input. Battery experiments and simulation protocols are essential to understand battery cell behavior in large-scale systems such as energy grids and general electrical battery research. This is because the different discharge cycles might impact the battery performance and be highly contributable to the battery research community.

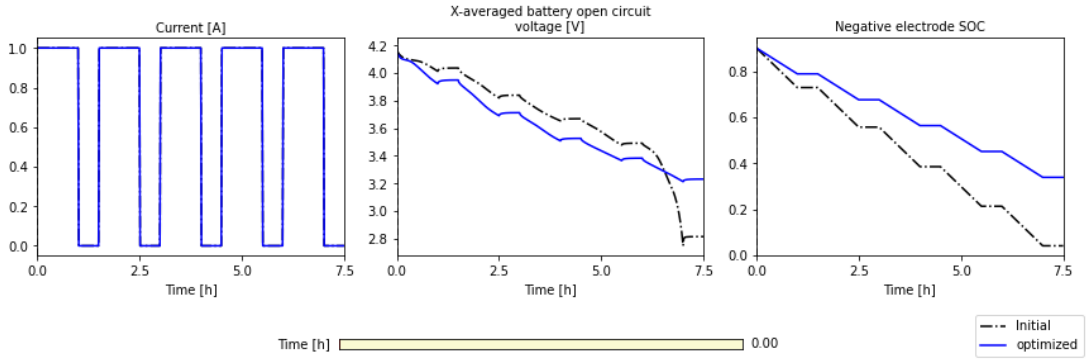


Figure C.9. Results of the five cycles of simulation with optimized battery cell parameters. The optimized data show an overall increase in capacity while exhibiting a more extensive reduction in high voltage levels. The black dash line represents the original LGM50 battery performance, and the blue line shows the optimized battery performance.

During the five experimental cycles, there is a significant voltage drop for the LGM50 battery after six hours. The optimized battery can maintain a relatively high voltage longer than the LGM50, which indicates that the optimized battery has a higher capacity. This result aligns

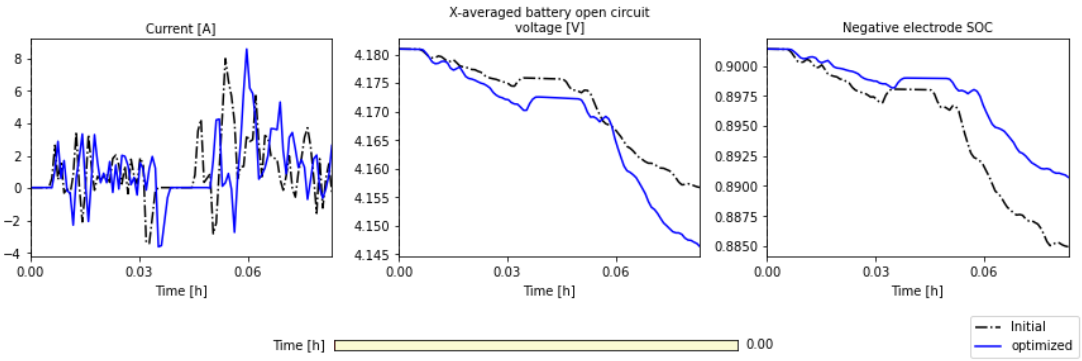


Figure C.10. Result of the federal drive simulation with optimized battery cell parameters. The black dash line represents the original LGM50 battery performance, and the blue line shows the optimized battery performance. The optimized result shows a delayed response compared to the initial since the optimized cell reduces the voltage, taking more time to hit the target power requirement.

with our previous optimization results. However, the OCV for the optimized battery is lower than that of the LGM50. The OCV is directly related to a terminal power, inferring that the optimized battery has a lower terminal power. This analysis is supported by the Ragone plot, which shows the trade-off relation between power and capacity [9]. A ragone plot is being used to compare the performance of various devices for energy storage [47]. In Appendix C, we used the most widely used cylindrical battery cell (LGM50) to optimize the specific energy density from published data and compare it with original result. As we discussed, the second life battery is not easy to gather its structural cell information. By constraining realistic conditions such as the power requirement, cost, and manufacturability of second life batteries, we can expect to estimate the battery cell structure close to the original battery.

Bibliography

[1] Weilong Ai, Ludwig Kraft, Johannes Sturm, Andreas Jossen, and Billy Wu. Electrochemical thermal-mechanical modelling of stress inhomogeneity in lithium-ion pouch cells. *Journal of The Electrochemical Society*, 167(1):013512, oct 2019.

[2] Global Battery Alliance. A vision for a sustainable battery value chain in 2030: Unlocking the full potential to power sustainable development and climate change mitigation. In *Geneva, Switzerland: World Economic Forum*, 2019.

[3] Selcuk Atalay, Muhammad Sheikh, Alessandro Mariani, Yu Merla, Ed Bower, and W. Dhammika Widanage. Theory of battery ageing in a lithium-ion battery: Capacity fade, nonlinear ageing and lifetime prediction. *Journal of Power Sources*, 478:229026, 2020.

[4] Muratahan Aykol, Chirranjeevi Balaji Gopal, Abraham Anapolsky, Patrick K. Herring, Bruis van Vlijmen, Marc D. Berliner, Martin Z. Bazant, Richard D. Braatz, William C. Chueh, and Brian D. Storey. Perspective—combining physics and machine learning to predict battery lifetime. *Journal of The Electrochemical Society*, 168(3):030525, mar 2021.

[5] Christoph R Birkl, Matthew R Roberts, Euan McTurk, Peter G Bruce, and David A Howey. Degradation diagnostics for lithium ion cells. *Journal of Power Sources*, 341:373–386, 2017.

[6] Von DAG Bruggeman. Berechnung verschiedener physikalischer konstanten von heterogenen substanzen. i. dielektrizitätskonstanten und leitfähigkeiten der mischkörper aus isotropen substanzen. *Annalen der physik*, 416(7):636–664, 1935.

[7] Frank Burden and Dave Winkler. *Bayesian Regularization of Neural Networks*, pages 23–42. Humana Press, Totowa, NJ, 2009.

[8] Chang-Hui Chen, Ferran Brosa Planella, Kieran O'Regan, Dominika Gastol, W. Dhammika Widanage, and Emma Kendrick. Development of experimental techniques for parameterization of multi-scale lithium-ion battery models. *Journal of The Electrochemical Society*, 167(8):080534, may 2020.

[9] Thomas Christen and Martin W. Carlen. Theory of ragone plots. *Journal of Power Sources*, 91(2):210–216, 2000.

[10] Lysander De Sutter, Yousef Firouz, Joris De Hoog, Noshin Omar, and Joeri Van Mierlo. Battery aging assessment and parametric study of lithium-ion batteries by means of a fractional differential model. *Electrochimica Acta*, 305:24–36, 2019.

[11] Marc Doyle, Thomas F. Fuller, and John Newman. Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell. *Journal of The Electrochemical Society*, 140(6):1526–1533, jun 1993.

[12] Peter Duffy, Colin Fitzpatrick, Thomas Conway, and Robert P. Lynch. Energy sources and supply grids – the growing need for storage. In *Energy Storage Options and Their Environmental Impact*, pages 1–41. The Royal Society of Chemistry, 2019.

[13] Madeleine Ecker, Nerea Nieto, Stefan Käbitz, Johannes Schmalstieg, Holger Blanke, Alexander Warnecke, and Dirk Uwe Sauer. Calendar and cycle life study of li(nimnco)o2-based 18650 lithium-ion batteries. *Journal of Power Sources*, 248:839–851, 2014.

[14] Jacqueline S. Edge, Simon O'Kane, Ryan Prosser, Niall D. Kirkaldy, Anisha N. Patel, Alastair Hales, Abir Ghosh, Weilong Ai, Jingyi Chen, Jiang Yang, Shen Li, Mei-Chin Pang, Laura Bravo Diaz, Anna Tomaszewska, M. Waseem Marzook, Karthik N. Radhakrishnan, Huizhi Wang, Yatish Patel, Billy Wu, and Gregory J. Offer. Lithium ion battery degradation: what you need to know. *Phys. Chem. Chem. Phys.*, 23:8200–8221, 2021.

[15] Henrik Ekström and Göran Lindbergh. A model for predicting capacity fade due to SEI formation in a commercial graphite/LiFePO4cell. *Journal of The Electrochemical Society*, 162(6):A1003–A1007, 2015.

[16] Zicheng Fei, Fangfang Yang, Kwok-Leung Tsui, Lishuai Li, and Zijun Zhang. Early prediction of battery lifetime via a machine learning based framework. *Energy*, 225:120205, 2021.

[17] Donal P. Finegan, Juner Zhu, Xuning Feng, Matt Keyser, Marcus Ulmefors, Wei Li, Martin Z. Bazant, and Samuel J. Cooper. The application of data-driven methods and physics-based learning for improving battery safety. *Joule*, 5(2):316–329, 2021.

[18] In-Ji Han, Tian-Feng Yuan, Jin-Young Lee, Young-Soo Yoon, and Joong-Hoon Kim. Learned prediction of compressive strength of ggbfs concrete using hybrid artificial neural network models. *Materials*, 12(22), 2019.

[19] Mohammed Hussein Saleh Mohammed Haram, Jia Woon Lee, Gobbi Ramasamy, Eng Eng Ngu, Siva Priya Thiagarajah, and Yuen How Lee. Feasibility of utilising second life ev batteries: Applications, lifespan, economics, environmental impact, assessment, and challenges. *Alexandria Engineering Journal*, 60(5):4517–4536, 2021.

[20] Simon S. Haykin. *Neural networks and learning machines*. Pearson Education, Upper Saddle River, NJ, third edition, 2009.

[21] Wei He, Michael Pecht, David Flynn, and Fateme Dinmohammadi. A physics-based electrochemical model for lithium-ion battery state-of-charge estimation solved by an optimised projection-based method and moving-window filtering. *Energies*, 11(8):2120, 2018.

[22] ROBERT HECHT-NIELSEN. Iii.3 - theory of the backpropagation neural network**based on "nonindent" by robert hecht-nielsen, which appeared in proceedings of the international joint conference on neural networks 1, 593–611, june 1989. © 1989 ieee. In Harry Wechsler, editor, *Neural Networks for Perception*, pages 65–93. Academic Press, 1992.

[23] Ali Jokar, Barzin Rajabloo, Martin Désilets, and Marcel Lacroix. Review of simplified pseudo-two-dimensional models of lithium-ion batteries. *Journal of Power Sources*, 327:44–55, 09 2016.

[24] Ji-San Kim, Dong-Chan Lee, Jeong-Joo Lee, and Chang-Wan Kim. Optimization for maximum specific energy density of a lithium-ion battery using progressive quadratic response surface method and design of experiments. *Scientific Reports*, 10(1):15586, Sep 2020.

[25] Ming Li, Wenqiang Du, and Fuzhong Nian. An adaptive particle swarm optimization algorithm based on directed weighted complex network. *Mathematical Problems in Engineering*, 2014:434972, Apr 2014.

[26] David JC MacKay. Bayesian interpolation. *Neural computation*, 4(3):415–447, 1992.

[27] Scott G. Marquis, Valentin Sulzer, Robert Timms, Colin P. Please, and S. Jon Chapman. An asymptotic derivation of a single particle model with electrolyte, 2019.

[28] Joaquim R. R. A. Martins and Andrew Ning. *Engineering Design Optimization*. Cambridge University Press, 2021.

[29] Wenxin Mei, Haodong Chen, Jinhua Sun, and Qingsong Wang. The effect of electrode design parameters on battery performance and optimization of electrode thickness based on the electrochemical–thermal coupling model. *Sustainable Energy Fuels*, 3:148–165, 2019.

[30] Jorge J. Moré. The levenberg-marquardt algorithm: Implementation and theory. In G. A. Watson, editor, *Numerical Analysis*, pages 105–116, Berlin, Heidelberg, 1978. Springer Berlin Heidelberg.

[31] Scott J Moura, Federico Bribiesca Argomedo, Reinhardt Klein, Anahita Mirtabatabaei, and Miroslav Krstic. Battery state estimation for a single particle model with electrolyte dynamics. *IEEE Transactions on Control Systems Technology*, 25(2):453–468, 2016.

[32] John Newman. Optimization of porosity and thickness of a battery electrode by means of a reaction-zone model. *Journal of The Electrochemical Society*, 142(1):97–101, jan 1995.

[33] John Newman and Karen E Thomas-Alyea. *Electrochemical systems*. John Wiley & Sons, 2012.

[34] Simon EJ O'Kane, Weilong Ai, Ganesh Madabattula, Diego Alonso Alvarez, Jacqueline Sophie Edge, Billy Wu, Gregory J Offer, and Monica Marinescu. Lithium-ion battery degradation: how to model it. *arXiv preprint arXiv:2112.02037*, 2021.

[35] V.J. Ovejas and A. Cuadras. State of charge dependency of the overvoltage generated in commercial li-ion cells. *Journal of Power Sources*, 418:176–185, 2019.

[36] Gustavo Paneiro and Manuel Rafael. Artificial neural network with a cross-validation approach to blast-induced ground vibration propagation modeling. *Underground Space*, 6(3):281–289, 2021.

[37] Matthew B. Pinson and Martin Z. Bazant. Theory of SEI formation in rechargeable batteries: Capacity fade, accelerated aging and lifetime prediction. *Journal of The Electrochemical Society*, 160(2):A243–A250, dec 2012.

[38] E. Prada, D. Di Domenico, Y. Creff, J. Bernard, V. Sauvant-Moynot, and F. Huet. A simplified electrochemical and thermal aging model of LiFePO4-graphite li-ion batteries: Power and capacity fade simulations. *Journal of The Electrochemical Society*, 160(4):A616–A628, 2013.

[39] Martin Riedmiller and Heinrich Braun. A direct adaptive method for faster backpropagation learning: The rprop algorithm. In *IEEE international conference on neural networks*, pages 586–591. IEEE, 1993.

[40] Nie Ru and Yue Jianhua. A ga and particle swarm optimization based hybrid algorithm. In *2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*, pages 1047–1050, 2008.

[41] Felipe Salinas and Julia Kowal. Discharge rate capability in aged li-ion batteries. *Journal of The Electrochemical Society*, 167(14):140519, oct 2020.

[42] L. E. Scales. Introduction to non-linear optimization. 1985.

[43] Kristen A. Severson, Peter M. Attia, Norman Jin, Nicholas Perkins, Benben Jiang, Zi Yang, Michael H. Chen, Muratahan Aykol, Patrick K. Herring, Dimitrios Fraggedakis, Martin Z. Bazant, Stephen J. Harris, William C. Chueh, and Richard D. Braatz. Data-driven prediction of battery cycle life before capacity degradation. *Nature Energy*, 4(5):383–391, May 2019.

[44] Marquis S.G. Timms R. Robinson M. Sulzer, V. and S.J. Chapman. Python battery mathematical modelling (pybamm). *Journal of The Electrochemical Society*, page 14, April 2021.

[45] Valentin Sulzer, Peyman Mohtat, Antti Aitio, Suhak Lee, Yen T. Yeh, Frank Steinbacher, Muhammad Umer Khan, Jang Woo Lee, Jason B. Siegel, Anna G. Stefanopoulou, and David A. Howey. The challenge and opportunity of battery lifetime prediction from field data. *Joule*, 5(8):1934–1955, 2021.

[46] Valentin Sulzer, Peyman Mohtat, Sravan Pannala, Jason B Siegel, and Anna G Stefanopoulou. Accelerated battery lifetime simulations using adaptive inter-cycle extrapolation algorithm. *Journal of The Electrochemical Society*, 168(12):120531, 2021.

[47] Bengt Sundén. Chapter 4 - battery technologies. In Bengt Sundén, editor, *Hydrogen, Batteries and Fuel Cells*, pages 57–79. Academic Press, 2019.

[48] Gerhard Venter. *Review of Optimization Techniques*. 12 2010.

[49] J. Vetter, P. Novák, M.R. Wagner, C. Veit, K.-C. Möller, J.O. Besenhard, M. Winter, M. Wohlfahrt-Mehrens, C. Vogler, and A. Hammouche. Ageing mechanisms in lithium-ion batteries. *Journal of Power Sources*, 147(1):269–281, 2005.

[50] Steven Walczak and Narciso Cerpa. Artificial neural networks. In Robert A. Meyers, editor, *Encyclopedia of Physical Science and Technology (Third Edition)*, pages 631–645. Academic Press, New York, third edition edition, 2003.

[51] Thomas Waldmann, Rares-George Scurtu, Karsten Richter, and Margret Wohlfahrt-Mehrens. 18650 vs. 21700 li-ion cells – a direct comparison of electrochemical, thermal, and geometrical properties. *Journal of Power Sources*, 472:228614, 2020.

[52] Aiping Wang, Sanket Kadam, Hong Li, Siqi Shi, and Yue Qi. Review on modeling of the anode solid electrolyte interphase (sei) for lithium-ion batteries. *npj Computational Materials*, 4(1):15, Mar 2018.

[53] Bingbin Wu, Yang Yang, Dianying Liu, Chaojiang Niu, Mark Gross, Lorraine Seymour, Hongkyung Lee, Phung M. L. Le, Thanh D. Vo, Zhiqun Daniel Deng, Eric J. Dufek, M. Stanley Whittingham, Jun Liu, and Jie Xiao. Good practices for rechargeable lithium metal batteries. *Journal of The Electrochemical Society*, 166(16):A4141–A4149, jan 2019.

[54] Bolun Xu, Alexandre Oudalov, Andreas Ulbig, Göran Andersson, and Daniel S. Kirschen. Modeling of lithium-ion battery degradation for cell life assessment. *IEEE Transactions on Smart Grid*, 9(2):1131–1140, 2018.

[55] Kang Xu. Nonaqueous liquid electrolytes for lithium-based rechargeable batteries. *Chemical Reviews*, 104(10):4303–4418, 2004. PMID: 15669157.

[56] Nansi Xue, Wenbo Du, Amit Gupta, Wei Shyy, Ann Marie Sastry, and Joaquim R. R. A. Martins. Optimization of a single lithium-ion battery cell with a gradient-based algorithm. *Journal of The Electrochemical Society*, 160(8):A1071–A1078, 2013.

[57] Xiao-Guang Yang, Yongjun Leng, Guangsheng Zhang, Shanhai Ge, and Chao-Yang Wang. Modeling of lithium plating induced aging of lithium-ion batteries: Transition from linear to nonlinear aging. *Journal of Power Sources*, 360:28–40, 2017.

[58] Toshihiro Yoshida, Michio Takahashi, Satoshi Morikawa, Chikashi Ihara, Hiroyuki Katsukawa, Tomoyuki Shiratsuchi, and Jun ichi Yamaki. Degradation mechanism and life prediction of lithium-ion batteries. *Journal of The Electrochemical Society*, 153(3):A576, 2006.

[59] Aston Zhang, Zachary C. Lipton, Mu Li, and Alexander J. Smola. Dive into deep learning. *CoRR*, abs/2106.11342, 2021.

[60] Ruifeng Zhang, Bizhong Xia, Baohua Li, Libo Cao, Yongzhi Lai, Weiwei Zheng, Huawen Wang, Wei Wang, and Mingwang Wang. A study on the open circuit voltage and state of charge characterization of high capacity lithium-ion battery under different temperature. *Energies*, 11(9), 2018.

[61] Yanyan Zhao, Oliver Pohl, Anand I. Bhatt, Gavin E. Collis, Peter J. Mahon, Thomas Rüther, and Anthony F. Hollenkamp. A review on battery market trends, second-life reuse, and recycling. *Sustainable Chemistry*, 2(1):167–205, 2021.