

**STATE ESTIMATION MODEL DEVELOPMENT
FOR LITHIUM-ION BATTERY PACKS**

TEOW YEE TING

UNIVERSITI TUNKU ABDUL RAHMAN

**STATE ESTIMATION MODEL DEVELOPMENT
FOR LITHIUM-ION BATTERY PACKS**

TEOW YEE TING

**A project report submitted in partial fulfilment of the
requirements for the award of Bachelor of Mechatronics
Engineering with Honours**

**Lee Kong Chian Faculty of Engineering and Science
Universiti Tunku Abdul Rahman**

October 2024

DECLARATION

I hereby declare that this project report is based on my original work except for citations and quotations which have been duly acknowledged. I also declare that it has not been previously and concurrently submitted for any other degree or award at UTAR or other institutions.

Signature :



Name :

Teow Yee Ting

ID No. :

1903674

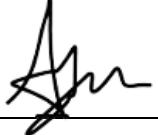
Date :

1 OCTOBER 2024

APPROVAL FOR SUBMISSION

I certify that this project report entitled **STATE ESTIMATION MODEL DEVELOPMENT FOR LITHIUM-ION BATTERY PACKS** was prepared by **TEOW YEE TING** has met the required standard for submission in partial fulfilment of the requirements for the award of Bachelor of Mechatronics Engineering with Honours at Universiti Tunku Abdul Rahman.

Approved by,

Signature : 

Supervisor : Dr. Hau Lee Cheun

Date : 7 OCTOBER 2024

The copyright of this report belongs to the author under the terms of the copyright Act 1987 as qualified by Intellectual Property Policy of Universiti Tunku Abdul Rahman. Due acknowledgement shall always be made of the use of any material contained in, or derived from, this report.

© 2024, TEOW YEE TING. All right reserved.

ACKNOWLEDGEMENTS

I would like to thank everyone who had contributed to the successful completion of this project. I would like to express my gratitude to my research supervisor, Dr. Hau Lee Cheun for his invaluable advice, guidance and his enormous patience throughout the development of the research.

In addition, I would also like to express my gratitude to my loving parents and friends who have helped and given me encouragement through the two stressful semesters to complete this final year project.

ABSTRACT

Accurate state of charge (SoC) estimation is essential for ensuring reliable battery management systems (BMS) in autonomous mobile robots (AMRs) across various industries. Precise SoC estimation enables AMRs to optimize performance by efficiently managing charging and discharging cycles, ensuring the effective use of lithium-ion batteries (LIBs), which are the primary power source. LIBs typically use lithium compounds, such as lithium cobalt oxide or lithium iron phosphate, as the cathode and graphite or carbon-based materials as the anode. Lithium-ion battery packs (LiBats) consist of multiple batteries assembled into a single unit. However, SoC estimation for LIB and LiBat systems is challenging due to the complexity of factors like electrochemical processes, thermal and kinetic effects, and battery aging. These factors limit the effectiveness of conventional SoC estimation methods—such as model-based, lookup table-based, and Coulomb counting approaches—which often rely on linear models and static assumptions, resulting in significant estimation errors in real-world conditions. As battery systems scale from individual cells to full packs, the complexity of SoC estimation increases, further exacerbated by variability in manufacturing and operating conditions, making accurate SoC predictions even more difficult. Some research suggests that data-driven SoC estimation techniques, particularly artificial neural network (ANN) models, can greatly enhance the performance of LIB and LiBat systems. By leveraging real-time data from local sensors and online databases, and utilizing artificial intelligence (AI) to adapt and refine models continuously, these techniques provide a more accurate understanding of the internal dynamics of LIBs and LiBats. This study emphasizes the importance of data-driven SoC estimation methods, focusing on ANN models such as backpropagation neural networks (BPNN), convolutional neural networks (CNN), long short-term memory (LSTM), and gated recurrent units (GRU). These models are compared with conventional SoC estimation techniques, specifically model-based approaches like the extended Kalman filter (EKF) and unscented Kalman filter (UKF). Estimation performance is evaluated using mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2), by comparing the experimental SoC, determined through the Coulomb counting method, with

the estimated SoC. Results show that the BPNN model outperforms all other methods, with an MAE of 6.1325, an RMSE of 10.1701, and an R^2 of 0.8806. The EKF method ranks second with an MAE of 6.3182, an RMSE of 8.9174, and an R^2 of 0.8608, followed by LSTM, GRU, and CNN. The UKF method performs the worst, with an MAE of 7.4654, an RMSE of 10.0908, and an R^2 of 0.8279. These findings underscore the pivotal role of data-driven SoC estimation techniques in improving the accuracy and reliability of LIB and LiBat systems in AMR applications.

TABLE OF CONTENTS

DECLARATION	i
APPROVAL FOR SUBMISSION	ii
ACKNOWLEDGEMENTS	iv
ABSTRACT	v
TABLE OF CONTENTS	vii
LIST OF TABLES	x
LIST OF FIGURES	xii
LIST OF SYMBOLS / ABBREVIATIONS	xv
LIST OF APPENDICES	xx

CHAPTER

1	INTRODUCTION	1
1.1	General Introduction	1
1.2	Importance of the Study	3
1.3	Problem Statement	4
1.4	Aim and Objectives	7
1.5	Scope and Limitation of the Study	7
1.6	Contribution of the Study	8
1.7	Outline of the Report	9
2	LITERATURE REVIEW	11
2.1	Introduction	11
2.2	Overview of LIB and LiBat Systems Technologies	11
	2.2.1 Internal Structure and Mechanism of LIB and LiBat Systems	11
	2.2.2 Industrial AMR Applications of LIB and LiBat Systems	13
2.3	Overview of Conventional SoC Estimation Methods	16

2.3.1	Look-Up Table Method	16
2.3.2	Coulomb Counting Method	17
2.3.3	Model Based Method	18
2.4	Overview of Data-Driven SoC Estimation Methods	32
2.4.1	Machine Learning Method	32
2.4.2	Fuzzy Logic Method	38
2.5	Summary	44
3	METHODOLOGY AND WORK PLAN	46
3.1	Introduction	46
3.2	Collecting LIB and LiBat Systems Internal Parameter Datasets	46
3.3	Developing LIB and LiBat Systems SoC Estimation Models	55
3.4	Fine-tuning LIB and LiBat Systems SoC Estimation Models	56
3.5	Evaluating LIB and LiBat Systems SoC Estimation Models	57
3.6	Final Year Project Gantt Chart	58
3.7	Summary	60
4	RESULTS AND DISCUSSION	62
4.1	Introduction	62
4.2	Model Optimization	62
4.2.1	Conventional Optimization	62
4.2.2	Data-Driven Optimization	71
4.3	Model SoC Estimation Results	74
4.3.1	Conventional SoC Estimation Results	74
4.3.2	Data-Driven SoC Estimation Results	84
4.3.3	Overall SoC Estimation Results	90
4.4	Challenges and Limitation of Study	94
4.5	Summary	96
5	CONCLUSIONS AND RECOMMENDATIONS	98
5.1	Conclusions	98
5.2	Recommendations for future work	98

REFERENCES	100
APPENDICES	104

LIST OF TABLES

Table 2.1: Characteristics of Various LIB and LiBat Systems (Liu, Placke, and Chau, 2022; McNulty et al., 2022). Reprinted with permission from Copyright 2022 Elsevier.	15
Table 4.1: EKF Parameter Settings for Li-NMC Graphite Battery System.	62
Table 4.2: EKF Parameter Settings for Li-Ion Graphite Battery System.	63
Table 4.3: EKF Parameter Settings for Li-NCA Graphite Battery System.	64
Table 4.4: EKF Parameter Settings for Li-PO Graphite Battery System.	64
Table 4.5: EKF Parameter Settings for Li-FP Graphite Battery System.	65
Table 4.6: UKF Parameter Settings for Li-NMC Graphite Battery System.	65
Table 4.7: UKF Parameter Settings for Li-Ion Graphite Battery System.	66
Table 4.8: UKF Parameter Settings for Li-NCA Graphite Battery System.	67
Table 4.9: UKF Parameter Settings for Li-PO Graphite Battery System.	67
Table 4.10: UKF Parameter Settings for Li-FP Graphite Battery System.	68
Table 4.11: BPNN Hyperparameter Settings for LIB and LiBat Systems.	71
Table 4.12: LSTM Hyperparameter Settings for LIB and LiBat Systems.	71
Table 4.13: GRU Hyperparameter Settings for LIB and LiBat Systems.	72
Table 4.14: CNN Hyperparameter Settings for LIB and LiBat Systems.	72

Table 4.15:	EKF Average SoC Estimation Results for the LG 18650HG2 3000mAh 3.6V Li-NMC Graphite Battery.	74
Table 4.16:	EKF Average SoC Estimation Results for the Samsung IN21700-30T 3000mAh 3.6V Li-Ion Graphite Battery.	74
Table 4.17:	EKF Average SoC Estimation Results for the SB LiMotive 5200mAh 3.7V Li-NCA Graphite Battery.	74
Table 4.18:	EKF Average SoC Estimation Results for the Turnigy Graphene 5000mAh 3.7V Li-PO Graphite Battery.	75
Table 4.19:	EKF Average SoC Estimation Results for the A123 1100, 2000, and 4400mAh 3.2V Li-FP Graphite Batteries.	75
Table 4.20:	Comprehensive Summary of EKF Overall SoC Estimation Results for LIB and LiBat Systems.	75
Table 4.21:	Average SoC Estimation Results Using the UKF Method for the LG 18650HG2 3000mAh 3.6V Li-NMC Graphite Battery.	75
Table 4.22:	Average SoC Estimation Results Using the UKF Method for the Samsung IN21700-30T 3000mAh 3.6V Li-Ion Graphite Battery.	76
Table 4.23:	Average SoC Estimation Results Using the UKF Method for the SB LiMotive 5200mAh 3.7V Li-NCA Graphite Battery.	76
Table 4.24:	Average SoC Estimation Results Using the UKF Method for the Turnigy Graphene 5000mAh 3.7V Li-PO Graphite Battery.	76
Table 4.25:	Average SoC Estimation Results Using the UKF Method for the A123 1100, 2000, and 4400mAh 3.2V Li-FP Graphite Batteries.	76
Table 4.26:	Comprehensive Summary of Average and Overall SoC Estimation Results Using the UKF Method for LIB and LiBat Systems.	76
Table 4.27:	SoC Estimation Results Using ANN Models for LIB and LiBat Systems Testing Datasets.	84
Table 4.28:	Overall Summary of SoC Estimation Results for LIB and LiBat Systems: EKF, UKF, BPNN, LSTM, GRU, and CNN.	90

LIST OF FIGURES

<p>Figure 1.1: General Function of a BMS (Xiong et al., 2018). Reprinted with permission from Copyright 2018 IEEE.</p> <p>Figure 2.1: General Structure of LIBs and LiBats (McNulty et al., 2022). Reprinted with permission from Copyright 2022 Elsevier.</p> <p>Figure 2.2: Radar Chart of Characteristics for Various LIB and LiBat Systems (Liu, Placke, and Chau, 2022; McNulty et al., 2022).</p> <p>Figure 2.3: OCV vs. SoC Curve of Li-FP Battery (Mendoza et al., 2017). Reprinted with permission from Copyright 2017 Elsevier.</p> <p>Figure 2.4: Schematic of the Model-Based Method (Ali et al., 2019). Reprinted with permission from Copyright 2019 MDPI.</p> <p>Figure 2.5: Schematic of the P2D Model (Petit, Calas, and Bernard, 2020). Reprinted with permission from Copyright 2020 Elsevier.</p> <p>Figure 2.6: Schematic of the ECM (Alter, Gu, and Cong, 2022). Reprinted with permission from Copyright 2022 Github.</p> <p>Figure 2.7: Schematic of the Impedance Circuit (Xiong et al., 2018). Reprinted with permission from Copyright 2018 IEEE.</p> <p>Figure 2.8: Terminal Voltage vs. SoC Charging Curves for the (a) LAB and (b) LIB systems (Man et al., 2012). Reprinted with permission from Copyright 2012 IEEE.</p> <p>Figure 2.9: General Structure of Black-Box Model (Natakarnkitkul, 2021b). Reprinted with permission from Copyright 2019 Medium.</p> <p>Figure 2.10: Structure of ANN for SoC Estimation (Chang, 2013). Reprinted with permission from Copyright 2013 Wiley.</p> <p>Figure 2.11: General Portrait of the SVM (Patle, and Chouhan, 2013). Reprinted with permission from Copyright 2013 IEEE.</p> <p>Figure 2.12: General Portrait of the Decision Boundaries (Soft Margins) and Hyperplane (Pannakkong et al., 2022). Reprinted with permission from Copyright 2022 Wiley.</p>	<p style="margin-top: 100px;">2</p> <p>12</p> <p>15</p> <p>16</p> <p>19</p> <p>20</p> <p>22</p> <p>23</p> <p>26</p> <p>33</p> <p>34</p> <p>36</p> <p>37</p>
---	---

Figure 2.13:	General Structure of FLC (Nur, Omaç, and Öksüztepe, 2017b). Reprinted with permission from Copyright 2017 Sigma.	40
Figure 2.14:	2D MF of FLC for SoC Estimation (Du et al., 2014). Reprinted with permission from Copyright 2014 IEEE.	40
Figure 2.15:	General Structure of ANFIS (Dai et al., 2015). Reprinted with permission from Copyright 2015 Elsevier.	42
Figure 3.1:	Discharge and Charging Curves for the LG 18650HG2 (Li-NMC) Battery: Terminal Voltage, Load Current, Ambient Temperature, and SoC at (a) -20°C, (b) -10°C, (c) 0°C, (d) 10°C, (e) 25°C, and (f) 40°C (Kollmeyer et al., 2020).	52
Figure 3.2:	OCV vs. SoC Curves: (a) LG 18650HG2 (Li-NMC), (b) Samsung IN21700-30T (Li-Ion), (c) SB LiMotive (Li-NCA), (d) Turnigy Graphene (Li-PO); and Terminal Voltage vs. SoC Curves: (e) A123 (Li-FP) Battery (Kollmeyer et al., 2020; Kollmeyer, and Skells, 2020, 2023; Xing et al., 2014).	55
Figure 3.3:	Gantt Chart of FYP Part 1.	58
Figure 3.4:	Gantt Chart of FYP Part 2.	59
Figure 4.1:	LG 18650HG2 (Li-NMC) Battery: Comparison of SoC Curves (Actual, EKF, and UKF) and Error Curves (EKF and UKF) at (a, b) 40°C, (c, d) 25°C, (e, f) 10°C, (g, h) 0°C, (i, j) -10°C, and (k, l) -20°C for the First 35,000 Samples.	81
Figure 4.2:	Comparison of SoC Curves between Actual, BPNN, LSTM, GRU, and CNN models, along with Error Curves for BPNN, LSTM, GRU, and CNN models at (a, b) the First 10 Samples, (c, d) 50 Samples, and (e, f) 100 Samples.	86
Figure 4.3:	Comparison of Loss Curves During Training and Validation for (a) BPNN, (b) LSTM, (c) GRU, and (d) CNN Models Over 30 Epochs.	88
Figure 4.4:	Comparison of SoC Curves between Actual, BPNN, LSTM, GRU, CNN, EKF, and UKF models, along with Error Curves for BPNN, LSTM, GRU, CNN EKF, and UKF models at (a, b) the First 10 Samples, (c, d) 50 Samples, and (e, f) 100 Samples.	92

Graph A-1:	Discharge and Charging Curves for the Samsung IN21700-30T (Li-Ion) Battery: Terminal Voltage, Load Current, Ambient Temperature, and SoC at (a) -20°C, (b) -10°C, (c) 0°C, (d) 10°C, (e) 25°C, and (f) 40°C (Kollmeyer, and Skells, 2023).	106
Graph A-2:	Discharge and Charging Curves for the SB LiMotive (Li-NCA) Battery: Terminal Voltage, Load Current, Ambient Temperature, and SoC at (a) -20°C, (b) -10°C, (c) 0°C, (d) 10°C, (e) 25°C, and (f) 40°C (Kollmeyer, and Skells, 2023).	109
Graph A-3:	Discharge and Charging Curves for the Turnigy Graphene (Li-PO) Battery: Terminal Voltage, Load Current, Ambient Temperature, and SoC at (a) -20°C, (b) -10°C, (c) 0°C, (d) 10°C, (e) 25°C, and (f) 40°C (Kollmeyer, and Skells, 2020).	112
Graph A-4:	Discharge and Charging Curves for the A123 (Li-FP) Batteries: Terminal Voltage, Load Current, Ambient Temperature, and SoC at (a) 0°C~5°C, (b) 10°C, (c) 20°C~25°C, (d) 30°C~35°C, and (e) 40°C (Xing et al., 2014; Catenaro, and Onori, 2021b; and Redondo-Iglesias, and Tian, 2024).	115
Graph A-5:	Samsung IN21700-30T (Li-Ion) Battery: Comparison of SoC Curves (Actual, EKF, and UKF) and Error Curves (EKF and UKF) at (a, b) 40°C, (c, d) 25°C, (e, f) 10°C, (g, h) 0°C, (i, j) -10°C, and (k, l) -20°C for the First 35,000 Samples.	119
Graph A-6:	SB LiMotive (Li-NCA) Battery: Comparison of SoC Curves (Actual, EKF, and UKF) and Error Curves (EKF and UKF) at (a, b) 40°C, (c, d) 25°C, (e, f) 10°C, (g, h) 0°C, (i, j) -10°C, and (k, l) -20°C for the First 35,000 Samples.	123
Graph A-7:	Turnigy Graphene (Li-PO) Battery: Comparison of SoC Curves (Actual, EKF, and UKF) and Error Curves (EKF and UKF) at (a, b) 40°C, (c, d) 25°C, (e, f) 10°C, (g, h) 0°C, (i, j) -10°C, and (k, l) -20°C for the First 35,000 Samples.	127
Graph A-8:	A123 (Li-FP) Battery: Comparison of SoC Curves (Actual, EKF, and UKF) and Error Curves (EKF and UKF) at (a, b) 40°C, (c, d) 35–30°C, (e, f) 25–20°C, (g, h) 10°C, and (i, j) 5–0°C for the First 35,000 Samples.	131

LIST OF SYMBOLS / ABBREVIATIONS

A_k	input system matrix of the system at sampling time k
B_k	input system matrix of the system at sampling time k
C_p	transient response capacitance at sampling time t , F
CPE_2	ZARC element impedance, Ω
C_k	constant matrix of the linearized system at sampling time k / input system matrix of the system at sampling time k
C_6	graphite anode
D_k	input system matrix of the system at sampling time k
e^-	negatively charged electron
$f(\text{net}_{ij})$	activation function of input layer i and hidden layer j
f	true format output estimated value
h_j	output value from hidden layer j
H_k	gain matrix of the system at sampling time k / output system matrix of system at sampling time k
$I_C(t)$	corrected load current at sampling time t , A
$I_L(t)$	load current at sampling time t , A
$I(t)$	load current at sampling time t , A
I_{PI_k}	continuous feedback function at sampling time k
$\dot{I_{PI_k}}$	rate of change of continuous feedback function at sampling time k
k_n	descrete sampling time
K_k	Kalman gain of the system at sampling time k
K_{p_k}	propotional gain matrix of the system at sampling time k
K_{i1_k}	integral gain matrix 1 at sampling time k
K_{i2_k}	integral gain matrix 2 at sampling time k
$K(x_i, x_j)$	kernel function
k_2	corrected load current constant
k_1	corrected load current constant
k_0	corrected load current constant
$LiCoO_2$	lithium cobalt oxide cathode

Li^+	positively charged lithium ion
LiC_6	lithium carbide compound
L_k	Luenberger gain matrix of the system at sampling time k
η	coulomb efficiency
n	degree of polynomial of the kernel function
net_{ij}	sum of the multiplication between the w_{ij} and x_i that are connected to the hidden layer j
p_i	coefficients of linear equation from fourth layer at node i
p_k	switching gain matrix at sampling time k
\hat{P}_k	updated error covariance of the system at sampling time k
P_{k+1}	error covariance of the system at sampling time $k + 1$
$Q(t)$	current available capacity at sampling time t , mAH
$Q_{max}(t)$	maximum available capacity at sampling time t , mAH
Q_{max}	maximum available capacity, As
q_i	coefficients of the linear equation from fourth layer at node i
Q_k	system's predicted covariance matrix at sampling time k
r_i	coefficients of linear equation from fourth layer at node i
R_k	system's corrected covariance matrix at sampling time k
R_1	internal resistance, Ω
R_0	internal resistance, Ω
R_p	transient response resistance at sampling time t , Ω
$SoC(t)$	state-of-charge at sampling time t
$SoC(t_k)$	state-of-charge at sampling time t_k
$SoC(t_0)$	state-of-charge at initial time t_0
$sgn(y_k - \widetilde{y}_k)$	discontinuous feedback at sampling time k
Δt	sampling time interval, s
$u_{A_i}(x)$	fuzzy linguistic input value of x from first layer at node A_i
$u_{B_i}(y)$	fuzzy linguistic input value of y from first layer at node B_i
u_k	input of the system at sampling time k
\hat{u}_k	updated input of the system at sampling time k
u_k	input of the system at sampling time k
$U_t(t)$	terminal voltage at sampling time t , V
U_{ocv}	open-circuit voltage, V

$U_p(t)$	transient response voltage at sampling time t , V
v_k	output noise of the system at sampling time k
v_k	output noise of the system at sampling time k
$V_0(t)$	terminal voltage at sampling time t , V
V_{ocv}	open-circuit voltage, V
$\bar{w}_l f_i$	output value of the linear equation at node i
\bar{w}_l	weighting factor from the third layer at the node i
w_i	weight of the second layer at the node i
W	Warburg element impedance of the battery, Ω
w_k	input noise of the system at the sampling time k
w_k	input noise of the system at the sampling time k
w_{ij}	weight of the neuron that is connected between the input layer i and the hidden layer j
x_k	state input of the system at sampling time k
\hat{x}_k	updated state input of the system at sampling time k
\hat{X}_k	updated state input of the system at sampling time k
$\dot{\tilde{x}}_k$	rate of change of filtered state input of the system at sampling time k
\widetilde{x}_k	filtered state input of the system at sampling time k
x_i	input value from the input layer i
x_i	vector of the input data i
x_j	vector of the input data j
x	true format input value x
∂x_k	derivative of the state input of the system at the sampling time k
y_{k+1}	output of the system at sampling time $k + 1$
y_k	output of the system at sampling time k
\widetilde{y}_k	filtered output of the system at sampling time k
y	true format input value y
∂y_k	derivative of the output of the system at the sampling time k
γ	kernel parameter of the kernel function
AMR	autonomous mobile robot

AI	artificial intelligence
AUC	area under the curve
ANFIS	adaptive neuro-fuzzy inference system
ANN	artificial neural network
BMS	battery management system
BPNN	backpropagation neural network
CNN	convolutional neural network
EMF	electromotive force
ECM	equivalent circuit model
EIM	electrochemical impedance model
EM	electrochemical model
EKF	extended Kalman filter
EV	electric vehicle
FYP	final year project
FLC	fuzzy logic controller
GRU	gated recurrent unit
HPPC	hybrid pulse power characterization
IS	impedance spectroscopy
IR	internal resistance
IoT	Internet of Thing
KF	Kalman filter
LSTM	long short-term memory
LAB	lead-acid battery
LiBat	lithium-ion battery pack
LIB	lithium-ion battery
Li-CO	lithium cobalt oxide
Li-FP	lithium iron phosphate
Li-Ion	lithium-ion
Li-NMC	lithium nickel manganese cobalt oxide
LMB	lithium-metal battery
Li-MO	lithium manganese oxide
Li-NCA	lithium nickel cobalt aluminium oxide
Li-NO	lithium nickel dioxide
Li-PO	lithium polymer

LO	Luenberger observer
MAE	mean absolute error
MF	membership function
NBB	nickel-based battery
OCV	open circuit voltage
PIO	proportional integral observer
P2D	pseudo-two-dimensional
Q3D	quasi-three-dimensional
R ²	coefficient of determination
RSME	root mean square error
ROC	receiver operating characteristic
RC	resistor-capacitor
SIB	sodium-ion battery
SoC	state of charge
SoE	state of energy
SoH	state of health
SoL	state of life
SoP	state of power
SoT	state of temperature
SVM	support vector machine
SMO	sliding mode observer
UTAR	Universiti Tunku Abdul Rahman
UKF	unscented Kalman filter
UT	unscented transform
1D	one-dimensional
3IR	third industrial revolution
4IR	fourth industrial revolution

LIST OF APPENDICES

Appendix A: Graphs	104
Appendix B: Tables	132
Appendix C: Open Access to Image Rights	133

CHAPTER 1

INTRODUCTION

1.1 General Introduction

In the early 2010s, a significant acceleration marked the transition from the third industrial revolution (3IR) to the fourth industrial revolution (4IR), primarily catalyzed by Germany's influence. This shift propelled the widespread adoption of autonomous mobile robots (AMRs) across diverse global industries, deployed in industrial, commercial, social, public, and personal sectors. AMRs address pressing challenges, notably the economic downturn due to escalating labor costs and workforce reductions attributed to aging demographics and declining birth rates. Their unmatched proficiency in repetitive tasks, spanning productivity, accuracy, and computation costs, distinguishes them from human counterparts (McNulty et al., 2022).

Furthermore, within the context of the 4IR, there has been a significant acceleration in demand for flexible work environments, leading to the rapid integration of AMRs across diverse industries. These AMRs leverage a wide array of Internet of Things (IoT) technologies, including sensors, artificial intelligence (AI), automation, cloud computing, and data science. This comprehensive integration empowers AMRs to excel in tasks involving locomotion, perception, and navigation, enabling seamless adaptation to the dynamic work scopes characteristic of the 4IR era. Unlike the static setups prevalent in conventional manufacturing during the 3IR, which relied on fixed tasks performed by pre-trained workers and pre-designed machines, the contemporary work landscape demands adaptability.

To optimize the performance of AMRs in various industries, a portable power source is essential. Unlike stationary machines powered by electrical cables, AMRs rely on rechargeable battery packs for mobility. Among these options, lithium-ion batteries (LIBs) are favored over lead-acid batteries (LABs), nickel-based batteries (NBBs), and sodium-ion batteries (SIBs) due to their superior key performance indicator (KPI) in specific energy, energy density, fast charging capabilities, safety, and cost-effectiveness (Liu, Placke, and Chau, 2022). Then, to fulfill the power needs of AMRs, several LIBs are linked in

parallel and series connections to create a lithium-ion battery pack (LiBat), guaranteeing the voltage and capacity for functioning.

Additionally, since AMRs rely on multiple LIBs operating together as a LiBat, integrating a battery management system (BMS) becomes essential to optimize LIB performance. Any failure of a LIB within the LiBat could have severe consequences for the AMRs. The BMS employs a variety of technologies, including sensors, actuators, processors, and controllers. Initially, the sample circuit collects internal battery data, such as load current, terminal voltage, and ambient temperature. This data is then processed by the control circuit as input. Using state estimation models, the control circuit assesses the performance of each LIB. Subsequently, this performance data is transmitted to the vehicle controller, enabling informed energy management decisions to ensure safe operation, extend range, and reduce costs, thereby enhancing LIB functionality in AMRs, as shown in Figure 1.1 (Xiong et al., 2018).

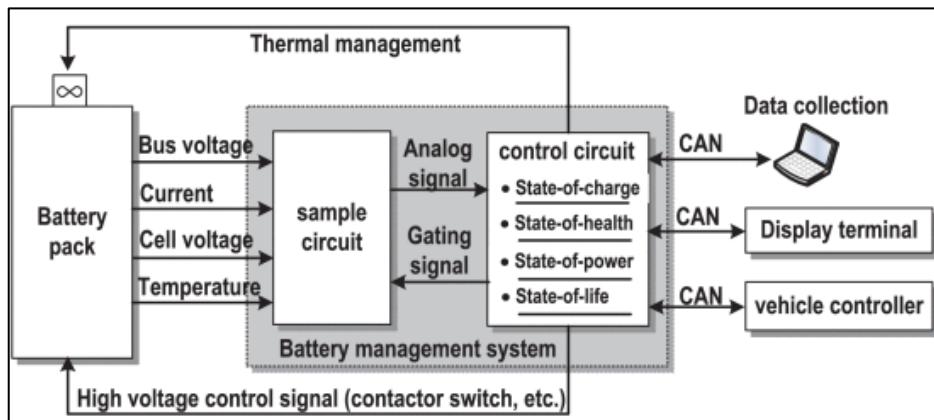


Figure 1.1: General Function of a BMS (Xiong et al., 2018). Reprinted with permission from Copyright 2018 IEEE.

Moreover, the state performance of the LIB and LiBat systems encompasses six key categories: state of charge (SoC), state of health (SoH), state of power (SoP), state of life (SoL), state of temperature (SoT), and state of energy (SoE). Among these, SoC stands out as the most vital and functional parameter. This is because the SoC is essential for accurately monitoring AMR operations and managing efficient charging and discharging schedules. The SoC represents the percentage of remaining capacity relative to the maximum available capacity, which can be expressed as below:

$$SoC(t) = \frac{Q(t)}{Q_{max}(t)} \quad (1.1)$$

where

$SoC(t)$ = state-of-charge at the sampling time t

$Q(t)$ = current available capacity at the sampling time t , mAh

$Q_{max}(t)$ = maximum available capacity at the sampling time t , mAh

Therefore, the primary objective of this research is to develop a state estimation model using data-driven methods for the LIB and LiBat systems utilized in AMRs, specifically focusing on the SoC aspect. To execute this data-driven state estimation model, practical work will be conducted on LIB and LiBat systems available at Universiti Tunku Abdul Rahman (UTAR), employing various sensors to measure internal parameters such as ambient temperature, terminal voltage, and load current during battery charge and discharge phases. The collected data will undergo a data sampling process using an Arduino Microcontroller Nano to convert continuous analog input data into discrete digital input data. Subsequently, the discrete digital input data will be transmitted to a computer or laptop embedded with pre-built data-driven models developed using Python for SoC estimation. Finally, the estimated SoC output values will be transmitted to the vehicle controller of the AMR for operational adjustments and displayed on monitor screens.

Importance of the Study

The research outcomes of this project hold immense potential to significantly impact battery state performance management by providing precise and effective estimates of the SoC value of LIB and LiBat systems. As previously mentioned, the integration of AMRs has become inevitable in companies across various industry sectors because of the significant acceleration of the global transition from the 3IR to the 4IR era. However, this rapid transition to the 4IR era has outpaced many companies capabilities to effectively incorporate AMR technologies into their industries. Therefore, these companies often confine themselves to strictly adhering to the guidelines and data provided in AMR data sheets and utilizing pre-installed scale indicators on AMRs for battery state

performance management rather than fully leveraging IoT technologies to improve AMR operations. This is particularly evident in the scheduling of AMR charging and discharging. However, these methods are often unreliable due to their lack of comprehensiveness in multiple key aspects.

Primarily, these scale indicators display SoC values only as whole numbers, lacking detailed decimal values, thereby hindering users from accurately assessing the rate of LIB and LiBat systems discharging and charging due to the absence of constant changes in SoC values as a reference facilitated by decimal precision. Furthermore, these data sheets serve only as general guidelines and data for monitoring the AMR's operational duration of the LIBs and LiBats. This hinders the user's ability to effectively and efficiently schedule the daily operation of the AMR and periodically schedule AMR charging and discharging for the LIB and LiBat systems, as manufacturing errors can cause individual performance to vary. Subsequently, since the data sheets are generally generated based on both theoretical and experimental scenarios, they are expected to provide limited guidance for the actual operation of the AMR.

Moreover, as time passes, the reliability of the data sheets and scale indicators will also decrease dramatically due to battery degradation and aging, which cause a change in the internal parameters such as ambient temperature, terminal voltage, and load current of the LIB and LiBat systems. Therefore, the AMR scale indicator's state estimation model primarily relies on conventional methods, potentially resulting in inaccurate SoC estimation due to the constantly changing internal parameters of the LIB and LiBat systems. The AMR's data sheets typically rely on the LIB and LiBat systems operating under fresh and optimized conditions, making them unsuitable for guiding the operation duration of the AMR with degraded and aged LiBats. In conclusion, the results of this study have the potential to help AMR implementers in many different areas get a better idea of the SoC values of the LIB and LiBat systems. This will help them plan the charging and discharging times of the AMR in a way that makes it work better and more efficiently.

1.2 Problem Statement

Estimating the SoC of the LIB and LiBat systems is a tedious and challenging task. According to Xiong et al. (2018), this task presents significant challenges,

primarily because conventional estimation approaches, including model-based, lookup table-based, and amp-hour integral methods, often yield results that differ significantly from real-world scenarios due to their limited adaptability to dynamic conditions that are full of uncertainties.

This problem stems from the inherent complexity of LIBs, which are intricate electrochemical devices with dynamic physical and chemical structures. The conversion mechanism of chemical energy into electrical energy is far from static; it's a deeply dynamic process influenced by numerous internal and external factors affecting electrochemical reactions, such as kinetic and thermodynamic processes. As a result, conventional estimation methods rely heavily on linear mathematical models to represent battery mechanisms and structures, rendering them unreliable, particularly in dynamic conditions marked by measurement errors from random variations.

Moreover, as batteries degrade over time, their dynamic properties experience further fluctuations in internal parameters. This complicates the reliability of SoC estimation using conventional methods, which assume that parameters are constant over time. For example, lookup table-based methods derive SoC estimates directly from the relationship tables of internal parameters such as open circuit voltage (OCV), electromotive force (EMF), impedance spectroscopy (IS), internal resistance (IR), ambient temperature, terminal voltage, and load current. Similarly, the amp-hour integral method employs an integral equation combining load current, maximum available capacity, and initial SoC to calculate the current SoC value of the battery.

Meanwhile, model-based methods integrate aspects of both amp-hour integral and lookup table-based methods. They achieved this by utilizing battery mathematical models such as the equivalent circuit model (ECM), electrochemical impedance model (EIM), and electrochemical model (EM), as well as measurement models, to obtain the terminal voltage parameter. Adaptive filter-based or observer-based models will then process the error values between the measured and calculated terminal voltage to estimate the SoC. However, these methods often overlook the impact of battery aging and random variations on these relationships, leading to inaccuracies in SoC estimation. In reality, the variability introduced by degradation and internal and external factors is inevitable, undermining the accuracy of conventional SoC estimation models.

Furthermore, as the estimation level scales from a single battery to the entire battery pack, it becomes even more difficult to accurately estimate the SoC using conventional estimation methods. This is because the battery manufacturing process is inconsistent, resulting in batteries with different operating conditions, degradation rates, and initial parameters. Since the battery packs usually contain more than one battery connected in parallel and series, this makes the overall mechanism and structure of the battery packs even more complex and unpredictable. In addition, given that most batteries' SoC estimation devices on the market rely on conventional estimation methods, the adoption of data-driven estimation methods is even more crucial to their practical usage for both industries and research purposes.

Thus, this project implements data-driven estimation methods, which are a much more accurate approach to solving the problem. Conventional estimation methods, unlike data-driven estimation methods, rely heavily on ideal mathematical models and assumptions, making them unsuitable for handling systems with dynamic structures and numerous uncertainties. While the data-driven estimation methods stand out as they only require retrieving the system's input data from real-time datasets from local sensors and online databases, with the help of AI, it processes the ability to fine-tune itself over time, offering greater potential to learn and uncover the internal dynamics of batteries using data sets compared to their counterparts.

However, the majority of the datasets for LIB and LiBat systems found in open-source repositories do not include the necessary internal parameters of SoC datasets. These parameters are crucial for constructing input-output databases used to train data-driven SoC estimation models. This is due to the inability to directly measure the SoC of the LIB and LiBat systems. As a result, within this project, there is an extra task of manually preprocessing the collected LIB and LiBat datasets. This is done by following the instructions provided by the experiment conductors, which are documented in the laboratory manual and published journal reports. The amp-hour integral method and the look-up table method are used to accurately create SoC datasets by eliminating internal and external noise.

1.3 Aim and Objectives

This project aims to develop a state estimation model for LIB and LiBat systems using both conventional methods, such as the extended Kalman filter (EKF) and unscented Kalman filter (UKF), and data-driven approaches, including backpropagation neural network (BPNN), convolutional neural network (CNN), long short-term memory (LSTM), and gated recurrent unit (GRU). The research will provide a detailed comparison of these SoC estimation methods based on coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE), determining the necessity for implementing data-driven models and the best-performing data-driven model for LIB and LiBat systems in the AMR sector. To achieve this, the following objectives must be completed:

- 1) Examine existing conventional and data-driven estimation models for SoC estimation in LIB and LiBat systems.
- 2) Identify and collect relevant datasets for developing the SoC estimation models for LIB and LiBat systems.
- 3) Develop and optimize SoC estimation models to achieve an R^2 greater than 0.8, a MAE less than 10, and a RMSE below 15.
- 4) Evaluate the accuracy of the optimised SoC estimation models using key metrics such as R^2 , MAE, and RMSE.
- 5) Identify the best-performing data-driven model, evaluating its necessity for implementation for LIB and LiBat systems in AMR sector based on R^2 , MAE, and RMSE.

1.4 Scope and Limitation of the Study

The primary goal of this final year project (FYP) is to develop both conventional and data-driven state estimation models for LIB and LiBat systems, with a specific focus on SoC estimation. The project aims to provide a detailed comparison of the performance of each SoC estimation method that undergoes moderate optimization to achieve a R^2 greater than 0.8, a MAE below 10, and a RMSE below 15, highlighting the necessity of implementing data-driven models and identifying the best-performing ones. This will be achieved by leveraging existing codes from open-source libraries and platforms, along with internal parameter datasets of the LIB and LiBat systems. Since the state estimation model is treated as a regression model, the performance will be

benchmarked without incorporating classification metrics like confusion matrices, area under the curve (AUC), and receiver operating characteristic (ROC).

Notably, this FYP will not incorporate advanced algorithms such as reinforcement learning, transformers, deep learning architectures, Bayesian methods, graph neural networks, or generative adversarial networks into the data-driven estimation models. Additionally, SoC estimation values will be transmitted solely from a computer or laptop equipped with pre-built data-driven models to the vehicle controller of the AMR for operational adjustments via a physical connection, with no involvement of IoT. However, the internal parameter datasets of the LIB and LiBat systems, such as internal resistance, dynamic resistance, and capacitance, are excluded from this study because they are specifically designed for advanced conventional and data-driven estimation models. Furthermore, this study does not include the assessment of the state performance of LIB and LiBat systems in AMR sectors, including SoT, SoE, SoH, SoP, and SoL.

SoH is defined as the percentage of the remaining maximum available capacity relative to the original maximum capacity of fresh LIB and LiBat systems. Similarly, SoL represents the percentage of the remaining discharge and charge cycles compared to the original maximum cycles of a fresh system. SoE and SoP are defined as the percentages of the remaining available energy and power relative to their original maximum values in fresh LIB and LiBat systems. Each of these metrics is essential for monitoring the operating performance of LIB and LiBat systems as they undergo degradation and aging, ensuring effective management before reaching the end of their usable service life. Additionally, SoT is defined as the percentage of the operating temperature relative to the maximum safe operating temperature of LIB and LiBat systems. It is used to monitor temperature conditions, ensuring they remain below the thermal runaway threshold to mitigate impacts on performance, safety, and lifespan.

1.5 Contribution of the Study

This FYP secured second place in the 2024 Tech 4 Good Challenge Competition, earning the 1st runner-up position in the Agriculture category with the title

Revolutionizing Agriculture Through Autonomous Technologies and Efficient Battery Management. The project successfully developed reliable and accurate SoC estimation models using the EKF method, delivering a detailed and precise battery management process for LIB and LiBat systems in industrial agricultural AMRs. For more details on the award, visit: [2024 Tech 4 Good Challenge - Agriculture Track, 2nd place](#). The main contributions of this study are summarized as follows:

- i. This study presents a series of reliable and accurately pre-processed internal parameter datasets for LIB and LiBat systems. The datasets were generated using coulomb counting and look-up methods, supported by guidance from experiment conductors, to accurately produce the experimental SoC values.
- ii. This study highlights the importance and necessity of data-driven SoC estimation methods in LIB and LiBat systems for industrial AMR applications, effectively addressing the complexities of battery management. Data-driven models provide superior performance and reliability compared to conventional models, which can be ranked as follows: BPNN, EKF, LSTM, GRU, CNN, and UKF.
- iii. This study found that the reliability of the UKF method is lower than that of the EKF method, contradicting the theoretical expectation that UKF should outperform EKF in the field of SoC estimation for LIB and LiBat systems. Additionally, as the operational temperature decreases, the reliability of the conventional models, including both EKF and UKF, declines, highlighting their sensitivity to lower temperature conditions.

1.6 Outline of the Report

The FYP is structured into five chapters. Chapter 2 begins with a literature review divided into two subsections: the first examines advancements in LIB and LiBat systems, focusing on their internal structures, mechanisms, comparisons with conventional batteries, and their applications in industrial AMR systems. The second subsection provides an overview of SoC estimation approaches, highlighting both conventional and data-driven techniques, with an emphasis on internal parameter dataset regression for LIB and LiBat systems. Chapter 3 details the methodology and work plan across three subsections: the

collection and pre-processing of internal parameter datasets, the development and optimization of SoC estimation models, and the benchmarking of these models using performance metrics such as R^2 , MAE, and RMSE. Chapter 4 presents the results and discussion, organized into three parts: an evaluation of conventional SoC estimation models, an evaluation of data-driven models, a comprehensive comparison of both, and the challenge and limitations faced in this research study. Finally, Chapter 5 concludes the study by summarizing the key findings and offering recommendations for future improvement and research.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This literature review examines the development of state estimation models for LIB and LiBat systems. It begins with an evaluation of the current advancements in the LIB and LiBat systems, focusing on their internal composition, mechanisms, and applications in the current industrial application of AMR. The goal of the first part is to assess current LIB and LiBat system technology in the context of industrial AMR applications, aiming to enhance our understanding of the current trend of the advancement of LIB and LiBat systems through the research process.

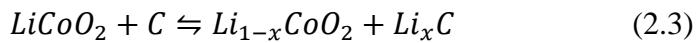
On the other hand, the second part of the review explores different methods of SoC estimation for LIB and LiBat systems, including conventional and data-driven approaches. It assesses the strengths and weaknesses of each SoC estimation model by comparing their respective advantages and disadvantages. Additionally, the review analyses various conventional and data-driven algorithms for SoC estimation models, evaluating their working principles, benefits, and limitations as described in the existing literature. The goal of the second part is to provide a detailed comparison between conventional and data-driven SoC estimation models and to determine the necessity of data-driven SoC estimation approaches.

2.2 Overview of LIB and LiBat Systems Technologies

2.2.1 Internal Structure and Mechanism of LIB and LiBat Systems

According to Wagner et al. (2013), the LIBs and LiBats are an advanced type of battery system that includes a positive electrode (anode), a negative electrode (cathode), and a separator. These components are immersed in an electrolyte. Each component and part of the LIBs and LiBats collaborate to execute the complex internal electrochemical thermal process of converting chemical energy to electrical energy during the discharging phase and converting electrical energy to chemical energy during the charging phase. Based on McNulty et al. (2022), take the lithium cobalt oxide (Li-CO) graphite battery as

an example. The complete redox electrochemical reaction equations of the cathode, anode, and overall are shown, respectively:



where

$LiCoO_2$ = lithium cobalt oxide cathode

C_6 = graphite anode

Li^+ = positively charged lithium ion

e^- = negatively charged electron

LiC_6 = lithium carbide compound

Anode, cathode, and electrolyte are the main components of a LIBs and LiBats. They work together to facilitate the electrochemical reaction that occurs by moving electrons and negative and positive ionic charges between the cathode and anode during the charging and discharging phases of the battery. This process will produce a current flow in both directions. Then, the electrolyte present within the cells will enhance the movement of electrons and negative and positive ions, ensuring a productive and efficient electrochemical reaction. Furthermore, a separator made up of electrically insulating material between the anode and cathode will be installed to mitigate the risk of short-circuit incidents within the battery, as shown in Figure 2.1.

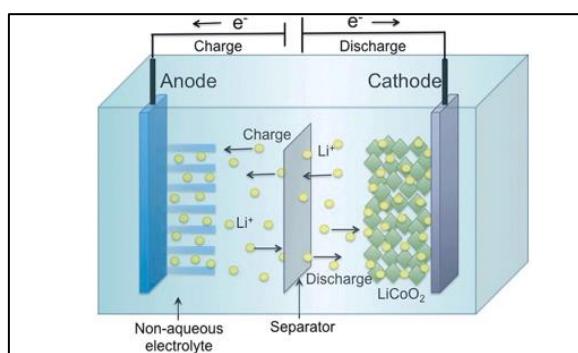


Figure 2.1: General Structure of LIBs and LiBats (McNulty et al., 2022).

Reprinted with permission from Copyright 2022 Elsevier.

2.2.2 Industrial AMR Applications of LIB and LiBat Systems

According to Mikołajczyk et al. (2023) and McNulty et al. (2022), the manufacturers primarily use LIBs and LiBats battery systems to incorporate industrial AMR applications in various industry sectors worldwide, with the exception of exceptional manufacturers like Fetch Robotics and Panasonic, which continue to use sealed lead-acid battery systems. This is because, according to Liu, Placke, and Chau (2022), the LIBs and LiBats battery systems are capable of providing a substantial number of KPIs in specific energy, energy density, fast charging capabilities, safety, and cost-effectiveness that are required by industrial AMR applications for long-term, continuous, and mobile operations (Mikołajczyk et al., 2023; McNulty et al., 2022). In contrast, conventional battery types such as acidic LABs and alkaline NBBs are unable to provide the sufficient amount of KPIs as mentioned previously to sustain industrial AMR applications. As a result, these devices have restricted use and are only applicable in certain situations involving small electrical and electronic systems and machinery (Křepelková, 2017).

Apart from that, other than the LIB and LiBat systems, the lithium-based battery systems can also be categorized into lithium-metal battery (LMB) systems, with the LMB system using metallic lithium as the anode material. Various cathode materials can be used, with manganese dioxide, iodine, and pyrite being common examples. However, the LIB and LiBat systems do not incorporate metallic lithium in their design. Instead, the cathode is made of lithium compounds such as lithium cobalt oxide, lithium iron phosphate, or other lithium-metal oxides. The anode is commonly composed of graphite or other carbon compounds. Despite the fact that LMB systems, being the counterparts of LIB and LiBat systems, are capable of delivering extraordinarily high theoretical KPI values in terms of theoretical capacity, the application of LMB is impeded by challenges such as unmanageable dendritic formation, inadequate cycling efficiency, and significant safety risks including short circuits, thermal runaway, electrolyte combustion, and the possibility of cell explosion. The obstacles have hindered its extensive acceptance, particularly as a rechargeable battery that may be frequently and continuously utilized for industrial AMR applications. At present, the LMB system is commercially used in single-disposable applications (McNulty et al., 2022).

Based on observations and research from Mikołajczyk et al. (2023) and McNulty et al. (2022), the most frequently discussed LIB and LiBat battery systems for implementation in industrial AMR applications by manufacturers are lithium iron phosphate (Li-FP), lithium-ion (Li-Ion), and lithium nickel manganese cobalt oxide (Li-NMC) graphite battery systems. Although there is a limited range of LIB and LiBat systems available for industrial AMR applications, many other applications, such as electric vehicles (EVs), and drones, now utilise various types of LIB and LiBat systems. These systems have the potential to be implemented in industrial AMR applications. Potential candidates include lithium nickel cobalt aluminium oxide (Li-NCA), lithium polymer (Li-PO), lithium nickel dioxide (Li-NO), and lithium manganese oxide (Li-MO) graphite battery systems (Liu, Placke, and Chau, 2022; Dileepan, Madhavan, and Jayakumar, 2017).

Each of these battery systems has its own unique advantages, making them relevant to automotive applications like industrial AMRs. First, the S-tier LIB and LiBat battery systems are Li-FP graphite battery systems due to their thermal stability, safety, high cycle life, and low cost (cobalt-poor/free) despite their low energy densities. They are popular among Magazino, Aethon, and Matthews Automation. Then, the A-tier LIB and LiBat battery systems are Li-NCA and Li-NMC graphite battery systems due to their lower thermal stability, safety, and higher cost (cobalt-medium) compared to Li-FP graphite battery systems, despite their higher energy densities, which are used by Mobile Industrial Robots and extensively in Tesla EVs. Furthermore, the B-tier LIB and LiBat battery systems, which are used in Hyundai Motor Company, Bolloré, and Kia Motors EVs, are Li-PO graphite battery systems due to their high energy densities, similar to those of Li-NCA and Li-NMC graphite battery systems, but with a lower weight and cost (cobalt-free), despite their lower thermal stability and safety. Finally, the C-tier LIB and LiBat battery systems are made of Li-NO and Li-MO graphite battery systems. They have lower energy densities than the Li-NCA and Li-NMC graphite battery systems, despite having similar thermal stability, safety, and lower cost (cobalt-free). Therefore, they are often outpaced by the Li-NCA and Li-NMC graphite battery systems in terms of popularity (McNulty et al., 2022; Liu, Placke, and Chau, 2022; Dileepan, Madhavan, and Jayakumar, 2017).

Moreover, Table 2.1 displays the characteristics of the KPI values for various types of LIB and LiBat systems, including Li-FP, Li-NMC, Li-NCA, Li-PO, Li-NO, and Li-MO graphite battery systems. The KPI values are characterized by the nominal voltage, discharge rate, charge rate, battery life cycle, specific energy, and thermal runway. Additionally, Figure 2.2 will display an illustration of various aspects of the KPI values for different types of LIB and LiBat systems in the form of a radar chart, as stated in Table 2.1.

Table 2.1: Characteristics of Various LIB and LiBat Systems (Liu, Placke, and Chau, 2022; McNulty et al., 2022). Reprinted with permission from Copyright 2022 Elsevier.

Cell	Norminal Voltage (V)	Discharge Rate (C)	Charge Rate (C)	Life Cycle (k)	Specific Energy (Wh/kg)	Thermal Runway (°C)
FP	3.20~3.30	1.0	1.0	1~2	90~130	270
NMC	3.65~4.00	1.0	0.7~1.0	1~2	130~241	210
NCA	3.60~3.65	1.0	0.7	0.5	200~310	150
PO	3.70~3.90	1.0	0.7~1.0	0.5~1	150~200	150
NO	3.65~3.70	1.0	0.7~1.0	0.3	150~200	150
MO	3.70~4.00	1.0	0.7~1.0	0.3~0.7	100~150	250

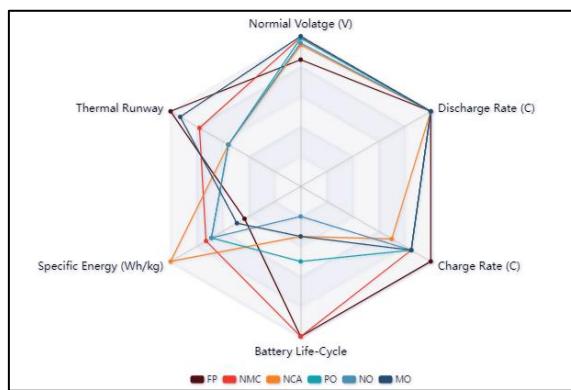


Figure 2.2: Radar Chart of Characteristics for Various LIB and LiBat Systems (Liu, Placke, and Chau, 2022; McNulty et al., 2022).

2.3 Overview of Conventional SoC Estimation Methods

2.3.1 Look-Up Table Method

The look-up table method is one of the most widely used methods for determining the SoC of LIB and LiBats systems. It relies on the relationship between the measurement of the internal parameters of the LIB and LiBat systems and the SoC.

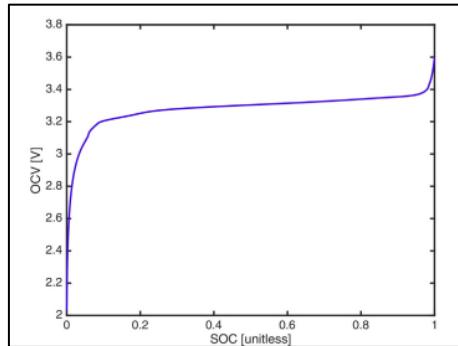


Figure 2.3: OCV vs. SoC Curve of Li-FP Battery (Mendoza et al., 2017).

Reprinted with permission from Copyright 2017 Elsevier.

Generally, there are seven different techniques commonly employed for obtaining the internal parameters, including EMF, IS, IR, ambient temperature, terminal voltage, load current, and OCV approaches. By employing lookup tables derived from these techniques, the correlation between internal parameters and the SoC of the LIB and LiBat systems, as seen in Figure 2.3, can be utilized for immediate estimation using measured internal parameter values (Ali et al., 2019).

Although the look-up table method is widely used in different kinds of industrial applications because of its simplicity and low computing cost, it has numerous limitations that make it impractical for real-world use. A significant constraint is the inability to obtain internal parameters such as the EMF and OCV in dynamic battery states, leaving it appropriate solely for laboratory settings. The need to repeatedly switch off power and allow the battery to rest for extended periods in order to achieve equilibrium further hampers its applicability (Xiong et al., 2018). Then, the internal parameters of IS and IR are dependent on the specific types of measurement instruments and techniques;

thus, they are difficult to implement for running industrial AMR applications (Ali et al., 2019).

Furthermore, based on Huang et al. (2017), as the cycle of the batteries increases, the correlation between their SoC and internal parameters alters due to the degradation and aging of batteries, leading to a decrease in the precision of predicting SoC using this approach. While the relationship between SoC and internal parameters also varies among batteries due to manufacturing variances, Thus, in order to improve the dependability of the look-up table approach, it is imperative to incorporate other types of conventional state estimation models, such as coulomb counting and model-based methods, by comparing their state estimation results representatively.

2.3.2 Coulomb Counting Method

The coulomb counting method is generally recognized as one of the most often used approaches for measuring the SoC of LIB and LiBat systems. It relies on a mathematical formula expressed as:

$$SoC(t_k) = SoC(t_0) - \frac{\eta}{Q_{max}} \int_{t_0}^{t_k} I_L(t) dt \quad (2.4)$$

where

$SoC(t_k)$ = state-of-charge at the sampling time t_k

$SoC(t_0)$ = state-of-charge at the initial time t_0

Q_{max} = maximum available capacity of the battery, As

η = coulomb efficiency

$I_L(t)$ = load current of the battery at the sampling time t , A

By utilizing Equation 2.4, the coulomb counting method accurately and reliably estimates the SoC by integrating the battery's charging or discharging load current over time. According to Xiong et al. (2018), this estimation is effective when the initial values, such as maximum available capacity, coulomb efficiency, and initial SoC of a battery, are known and the load current of charging and discharging can be measured precisely. As a result, the coulomb counting method offers the advantages of simple implementation and low

computational costs for estimating SoC. Moreover, its efficiency in estimating short-term values has made it a popular choice among industrial AMR applications.

Despite its advantages, the accuracy of the coulomb counting method is affected by various factors. The accuracy of the charging or discharging load current sensors is critical, as disturbances, noise, and drift caused by internal and external battery factors will impact precision. Apart from that, uncertainties surrounding the values of the initial SoC, coulomb efficiency, and the fluctuating maximum available capacity due to battery aging considerations further contribute to potential inaccuracies.

Besides, to enhance the performance of the coulomb counting method, which is similar to the look-up table method, it is often necessary to collaborate with other types of conventional state estimation models, such as look-up tables and model-based methods, and compare their state estimation results representatively. In addition, Chang (2013) noted that the modified coulomb counting method is acknowledged as a viable approach to enhance the precision of the coulomb counting method. The modified coulomb counting method's mathematical expression is as follows:

$$SoC(t_k) = SoC(t_0) - \frac{\eta}{Q_{max}} \int_{t_0}^{t_k} I_C(t) dt \quad (2.5)$$

$$I_C(t) = k_2 I_L(t)^2 + k_1 I_L(t) + k_0 \quad (2.6)$$

where

$I_C(t)$ = corrected load current of the battery at the sampling time t , A

k_2 , k_1 and k_0 = corrected load current constant from practical experiment data

2.3.3 Model Based Method

The model-based method is a complex approach used to estimate the SoC of LIB and LiBat systems. It employs mathematical models, such as the EIM, ECM, and EM methods, to represent the internal chemical, electrical, and thermal structure of a battery. The model will be translated into state equations based on their respective theories to calculate the battery's terminal voltage output at different sampling times. Next, the terminal voltage error between

measured and calculated values is processed by the adaptive SoC estimation methods at various sampling times using nonlinear state estimation algorithms to derive the estimated SoC, as depicted in Figure 2.4.

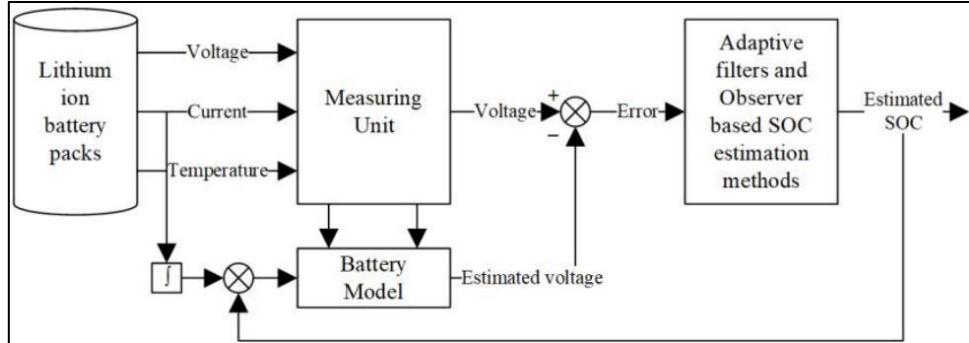


Figure 2.4: Schematic of the Model-Based Method (Ali et al., 2019). Reprinted with permission from Copyright 2019 MDPI.

The commonly used adaptive SoC estimation methods for the LIB and LiBat systems in model-based approaches include filter-based methods such as Kalman filter (KF), EKF, UKF, as well as observer-based methods such as Luenberger observer (LO), sliding mode observer (SMO), and proportional integral observer (PIO). Each of these methods can improve the robustness and accuracy of the estimated SoC values of the battery by utilizing their inherent capability to self-correct in response to varying input signals caused by the battery's dynamic behavior, battery degradation, external environmental disturbances, and sensor precision.

On the other hand, the model-based method can also be used in conjunction with other conventional methods to further improve the accuracy of these methods, as mentioned by Xiong et al. (2018) and Liu, Placke, and Chau (2022). Chang (2013) noted that merging the dual EKF with the OCV-based look-up table methods will significantly improve the accuracy of the estimated SoC by combining the measured and calculated values to get the weighted values. Chang (2013) also mentioned that by combining the coulomb counting with the KF and EKF methods, the initial values used in the coulomb counting method will be corrected by combining measured and calculated initial values to converge them as close as possible to the real values.

2.3.3.1 Electrochemical Model (EM) Method

The EM method precisely replicates the complex internal electrothermal dynamics of the battery, which are determined by the electrochemical and thermal reactions that govern the movement of positively charged ions, electrons, and other materials within the electrolytes, anode, and cathode during the conversion of energy between electrical and chemical forms. By governing the transit of mass, energy, and momentum across each phase within the battery, the terminal voltage can be calculated, which may be described using partial differential equations as displayed in Figure 2.5 (Petit, Calas, and Bernard, 2020).

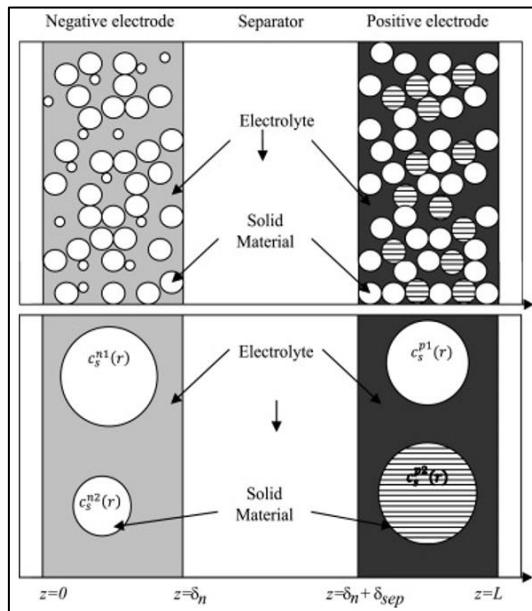


Figure 2.5: Schematic of the P2D Model (Petit, Calas, and Bernard, 2020).

Reprinted with permission from Copyright 2020 Elsevier.

Despite the fact that the EM method offers moderate accuracy and reliability, its implementation can be costly because of the challenging interpretation of the electrothermal processes occurring within the battery. In order to handle complexity, it is necessary to describe the internal behavior of batteries using abundance and intricate partial differential equations, along with the pseudo-two-dimensional (P2D), one-dimensional (1D), quasi-three-dimensional (Q3D) full-order physical, or first principle models (Xiong et al., 2018). Furthermore, the implemented model is subjected to periodic calibration

and validation using experiment data. Therefore, the EM method is only preferred in battery design instead of industrial applications of AMR.

2.3.3.2 Equivalent Circuit Model (ECM) Method

The ECM method provides an accurate description of the complex internal structure and behavior of batteries by utilizing electrical circuit components such as resistors, capacitors, and voltage sources. These components represent the internal resistance, OCV, and dynamic transient characteristics of the battery. As shown in Figure 2.6, the ECM method constructs circuit networks using resistor-capacitor (RC) and Thevenin circuits. Subsequently, according to Xiong et al. (2018) and Alter, Gu, and Cong (2022), the circuit equations can be described as follows:

$$U_t(t) = U_{ocv} - U_p(t) - R_0 I(t) \quad (2.7)$$

$$\dot{U_p}(t) = -\frac{1}{R_p C_p} U_p(t) + \frac{1}{C_p} I(t) \quad (2.8)$$

where

$U_t(t)$ = terminal voltage of the battery at the sampling time t , V

U_{ocv} = open-circuit voltage of the battery, V

$U_p(t)$ = transient response voltage of the battery at the sampling time t , V

R_0 = internal resistance of the battery, Ω

$I(t)$ = load current of the battery at the sampling time t , A

R_p = transient response resistance of the battery at the sampling time t , Ω

C_p = transient response capacitance of the battery at the sampling time t , F

The solution of the differential equation (2.8) is as follows :

$$U_p(t) = \exp\left(\frac{-t}{R_p C_p}\right) C + R_p I(t) \quad (2.9)$$

The arbitrary constant, C can be determined by employing the Hybrid Pulse Power Characterization (HPPC) test. The discrete representations of various states can be combined in the following manner:

$$U_p(k_{n+1}) = \exp\left(\frac{-\Delta t}{R_p C_p}\right) * (U_p(k_n) - \dots \\ \dots I(k_n) R_p) + R_p I(k_n) \quad (2.10)$$

$$U_t(k_{n+1}) = U_{ocv} - U_p(k_{n+1}) - R_0 I(k_{n+1}) \quad (2.11)$$

where

k_n = discrete sampling time of the battery

Δt = sampling time interval of the battery, s

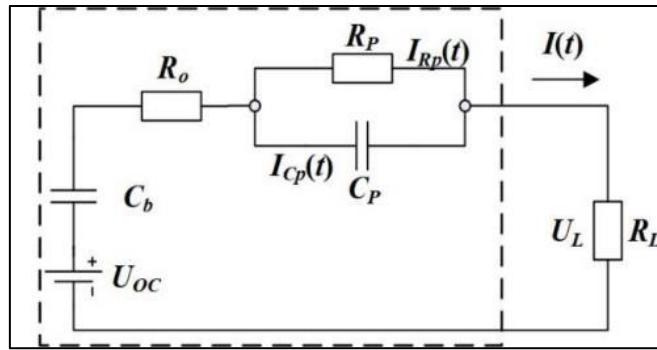


Figure 2.6: Schematic of the ECM (Alter, Gu, and Cong, 2022). Reprinted with permission from Copyright 2022 Github.

Based on Liu, Placke, and Chau (2022), despite the fact that the ECM method has the same level of accuracy and reliability as the EM method, it is preferred in industrial AMR applications that utilize model-based approaches for the SoC estimation of batteries because of its low implementation cost and straightforward and simple interpretation nature.

2.3.3.3 Electrochemical Impedance Model (EIM) Method

The EIM method represents the complex internal structure and behavior of batteries by leveraging porous electrode theory. It employs electrical components, including resistors, capacitors, and voltage sources, represented in impedance form. The components include the Warburg and Zarc elements, internal resistance, OCV, and the battery's dynamic transient characteristics. As shown in Figure 2.7, the EIM uses RC and Thevenin circuits to represent the impedance circuit networks. Consequently, the circuit mathematical equations can be used to describe the EIM of the battery as follows:

$$V_0(t) = V_{ocv} - R_1 I(t) - CPE_2 I(t) - WI(t) \quad (2.12)$$

$$V_0(t) = V_{ocv} - V_1(t) - V_2(t) - V_3(t) \quad (2.13)$$

where

$V_0(t)$ = terminal voltage of the battery at the sampling time t , V

V_{ocv} = open-circuit voltage of the battery, V

R_1 = internal resistance of the battery, Ω

$I(t)$ = load current of the battery at the sampling time t , A

CPE_2 = ZARC element impedance of the battery, Ω

W = Warburg element impedance of the battery, Ω

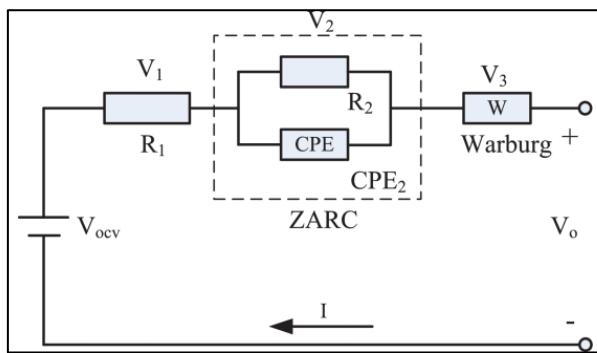


Figure 2.7: Schematic of the Impedance Circuit (Xiong et al., 2018). Reprinted with permission from Copyright 2018 IEEE.

Due to the fact that the EIM method is capable of identifying parameters through harmonic small-signal excitation at different frequencies of current across the battery, it enables direct measurement of various SoC points (Xiong et al., 2018). However, despite having a similar methodology, Liu, Placke, and Chau (2022) claim that the accuracy of the ECM suffers because it requires a large number of RC branches for the circuit networks. The accuracy of the ECM method increases with the number of RC circuit parts, albeit at the cost of increased model complexity. Thus, under the similar structural level of the circuit networks, the EIM is far more accurate than the ECM, making it widely used in diverse industrial AMR applications. However, due to its perceived complexity compared to the ECM method, the EIM method may be less preferred in certain contexts.

2.3.3.4 Kalman Filter (KF) Method

The KF method is employed in circumstances where both the measurement and calculation models of the system are linear. Its purpose is to eliminate the inaccuracies within the input data of the system. Based on Adil Khan (2020), the KF method utilizes a series of discrete recursive equations to analyze the error values between measurements and calculations in the input data of the system, as shown below:

$$x_{k+1} = A_k x_k + B_k u_k + w_k \quad (2.14)$$

$$y_{k+1} = H_k x_k + v_k \quad (2.15)$$

where

x_k = state input of the system at the sampling time k

u_k = input of the system at the sampling time k

y_{k+1} = output of the system at the sampling time $k + 1$

A_k = input system matrix ($n \times n$) of the system at the sampling time k

B_k = input system matrix ($n \times I$) of the system at the sampling time k

H_k = output system matrix ($n \times m$) of the system at the sampling time k

w_k = input noise of the system at the sampling time k

v_k = output noise of the system at the sampling time k

Equation (2.14) is processed through the predicted process to remove the system noise factors as follows:

$$\hat{x}_{k+1} = A_k \hat{x}_k + B_k \hat{u}_k \quad (2.16)$$

where

\hat{x}_k = updated state input of the system at the sampling time k

\hat{u}_k = updated input of the system at the sampling time k

Equation (2.16) is processed through the predicted process to produce the error covariance equation as follows:

$$P_{k+1} = A_k P_k A^T k + Q_k \quad (2.17)$$

where

P_{k+1} = error covariance of the system at the sampling time $k + 1$

Q_k = predicted covariance matrix ($n \times n$) of the system at the sampling time k

Equation (2.17) is processed through the corrected process to produce the Kalman gain equation as follows:

$$K_k = P_k H^T k (H_k P_k H^T k + R_k)^{-1} \quad (2.18)$$

where

K_k = Kalman gain of the system at the sampling time k

R_k = corrected covariance matrix ($n \times n$) of the system at the sampling time

Equation (2.18) is being applied to the equations (2.17 and 2.16) through the corrected process to produce the updated state input and error covariance equation as follows:

$$\hat{X}_k = \hat{x}_k + K_k (y_k - H_k \hat{x}_k) \quad (2.19)$$

$$\hat{P}_k = (I - K_k H_k) P_k \quad (2.20)$$

where

\hat{X}_k = updated state input of the system at the sampling time k

\hat{P}_k = updated error covariance of the system at the sampling time k

Based on the series of discrete recursive equations mentioned above, the KF method will then be able to obtain the weighted SoC estimation values between the input data of measurements and calculation values. However, according to Moussalli, Brahim Sedra, and Ait Laachir (2018), the KF technique is exclusively appropriate for linear battery systems, such as the LAB, where the terminal voltage and the SoC have a proportional relationship that can be easily expressed using a mathematical function. An increase in the terminal

voltage leads to an increase in the SoC, and vice versa. In contrast, it is not suitable for non-linear LIB and LiBat systems, where the relationship between the terminal voltage and the SoC is constantly changing. These systems exhibit a proportional relationship between 20% to 40% and 80% to 100% of the SoC, while the rest of the SoC displays horizontal line relationships, leading to high levels of uncertainty as shown in Figure 2.8, which makes it less popular in the field of industrial AMR applications.

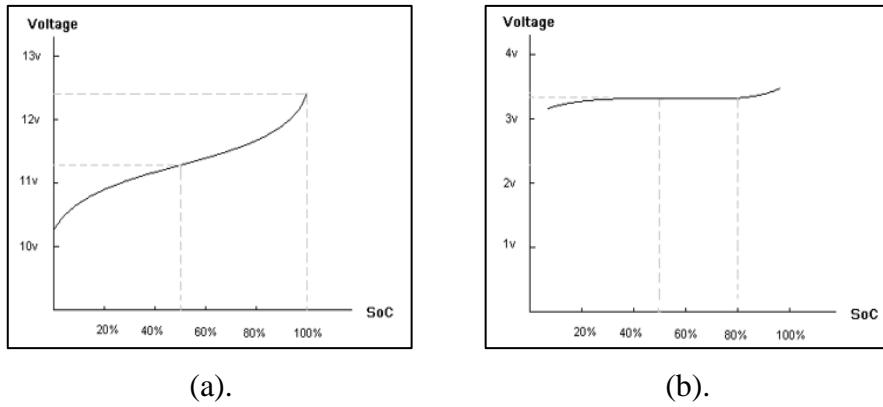


Figure 2.8: Terminal Voltage vs. SoC Charging Curves for the (a) LAB and (b) LIB systems (Man et al., 2012). Reprinted with permission from Copyright 2012 IEEE.

2.3.3.5 Extended Kalman Filter (EKF) Method

The EKF method is used when both the measurement and calculation models of the system are non-linear. The purpose of this system is to address the issue of the KF method by utilizing linearization mathematical models to linearize the non-linear measurement and calculation models and eliminate inaccuracies in the input data. Based on Alter, Gu, and Cong (2022), the EKF method included the first-order Taylor series expansion to provide the approximate estimation values of the non-linear system models, as shown below:

$$C_k = \frac{\partial y_k}{\partial x_k} \quad (2.21)$$

where

C_k = constant matrix of the linearized system at the sampling time k

∂y_k = derivative of the output of the system at the sampling time k

∂x_k = derivative of the state input of the system at the sampling time k

Equation (2.21) is being applied to equations (2.18 and 2.19) through the linearization process to produce the linearized Kalman gain and updated state input equation as follows:

$$K_k = P_k C^T k (C_k P_k C^T k + R_k)^{-1} \quad (2.22)$$

$$\hat{X}_k = \hat{x}_k + K_{k+1}(y_k - C_k \hat{x}_k) \quad (2.23)$$

Equations (2.22 and 2.23) are being utilized to produce the linearized error covariance equation as follows:

$$\hat{P}_k = P_k - K_k C_k P_k \quad (2.24)$$

Henceforth, the EKF method will obtain more accurate weighted estimation values of SoC between the input data of measurements and calculation values. The ECM will introduce an additional RC branch within the model-based method to implement the EKF method (Ali et al., 2019). This makes the EKF method an excellent choice for industrial AMR applications that utilize LIB and LiBat systems in addition to the UKF method.

This is because, under the real-world scenario, the reliability of the EKF method may be compromised as the working principle of the EKF method will result in a significant amount of linearization error. As the non-linearity of the system model increases, the linearization error will increase (Moussalli, Brahim Sedra, and Ait Laachir, 2018). This becomes the EKF method's biggest weakness in dealing with LIB and LiBat systems that operate in real-world scenarios. Under the real-world scenario, the LIB and LiBat systems will show an even higher amount of non-linearity than in an experimental environment. This is because during the practical operation, numerous external factors will be encountered, including sensor noise, temperature drifting, battery degradation, manufacturing variation, dynamic working environment, and the types of LIB and LiBat systems used.

2.3.3.6 Unscented Kalman-Filter (UKF) Method

The UKF method is employed in situations where both the measurement and calculation models of the system exhibit significant non-linearity. Such significant non-linearity can be constituted by the LIB and LiBat systems that operate in real-world scenarios with numerous external and internal factors, as mentioned in Section 2.3.3.5. Therefore, the objective of the UKF method is to resolve the problem of high linearization error in the EKF method when working with the LIB and LiBat systems, which are much more non-linear.

The UKF method will utilize the unscented transform (UT), also known as the "sigma point filter" approach. It employs the Gaussian distribution to propagate through the nonlinear system models, with the aim of capturing a collection of sample points known as sigma points. Next, the acquired sigma points will be propagated together with their propagated covariance and mean. Furthermore, the measurement function will be applied to the propagated sigma points in order to produce the updated propagated sigma points. Based on the updated propagated sigma points, the updated propagated covariance and mean will also be produced (Moussalli, Brahim Sedra, and Ait Laachir, 2018).

In addition, the estimated covariances will then be produced by combining the updated and non-updated propagated sigma points. The estimated covariances will be combined to calculate the Kalman gain, which is crucial for generating the covariance and mean of the linearized measurement and calculation models of the LIB and LiBat systems. Based on the covariance and mean of the linearized models of LIB and LiBat systems mentioned above, the UKF method will then be able to obtain the weighted SoC estimation values between the input data of measurements and calculation values.

Therefore, the UKF method can achieve more robust and accurate weighted SoC estimations by significantly reducing the impact of linearization error. As the UKF method accuracy surpasses the EKF, it can predict high-order nonlinear system states accurately. This makes it popular among the industrial AMR applications that apply LIBs and LiBats. It can effectively counter the issue of LIB and LiBat systems experiencing higher non-linearity when operating in real-world scenarios. However, as the UKF method is more complex in terms of modeling, the computation and implementation cost become its biggest weaknesses (Ali et al., 2019).

2.3.3.7 Luenberger Observer (LO) Method

Similar to the KF method, the LO method is a technique that is used in situations where both the measurement and calculation models of the system are linear (Zou et al., 2014). Thus, dealing with the LIB and LiBat systems with the LO method will be difficult due to its non-linear and dynamic nature, as mentioned in Section 2.3.3.4. Based on Hu, Sun, and Zou (2010), the LO method will accurately utilize a series of state equations to examine the error values between measurements and calculations in the input data of the battery system. Throughout the estimation process of the SoC of the battery system, with the help of the integral of the filtered state input equation, the Luenberger gain in the state equations will be continuously updated in order to reduce the inaccuracies in the input data of the battery system, as shown below:

$$x_{k+1} = A_k x_k + B_k u_k + w_k \quad (2.25)$$

$$y_k = C_k x_k + D_k u_k + v_k \quad (2.26)$$

where

x_k = state input of the system at the sampling time k

u_k = input of the system at the sampling time k

y_k = output of the system at the sampling time k

A_k, B_k, C_k and D_k = input system matrixs of the system at the sampling time k

w_k = input noise of the system at the sampling time k

v_k = output noise of the system at the sampling time k

Equations (2.25 and 2.26) are processed through the Luenberger filtering process to remove the system noise factors as follows:

$$\dot{\widetilde{x}_k} = A_{k-1} \widetilde{x}_{k-1} + B_{k-1} u_{k-1} + L_k (y_k - \widetilde{y}_k) \quad (2.27)$$

$$\widetilde{y}_k = C_k \widetilde{x}_{k-1} + D_k u_k \quad (2.28)$$

where

$\dot{\widetilde{x}_k}$ = rate of change of filtered state input of the system at the sampling time k

\widetilde{x}_k = filtered state input of the system at the sampling time k

\widetilde{y}_k = filtered output of the system at the sampling time k

L_k = Luenberger gain matrix of the system at the sampling time k

2.3.3.8 Sliding Mode Observer (SMO) Method

The SMO method is the technique that is utilized based on the framework of the LO method, as shown in Section 2.3.3.7. Similar to the EKF method, the purpose of this method is to address the issue of the LO method being unable to be used to eliminate the inaccuracies in the input data of the LIB and LiBat systems that have non-linear and dynamic measurement and calculation models. Based on Hu, Sun, and Zou (2010), the SMO method included the discontinuous feedback function at the back of the input state equation of the LO method, as shown below:

$$\begin{aligned}\dot{\widetilde{x}}_k &= A_{k-1} \widetilde{x}_{k-1} + B_{k-1} u_{k-1} + \dots \\ &\dots H_k (y_k - \widetilde{y}_k) + p_k \cdot sgn(y_k - \widetilde{y}_k)\end{aligned}\quad (2.29)$$

$$sgn(y_k - \widetilde{y}_k) = \begin{cases} -1, & (y_k - \widetilde{y}_k) < 0 \\ 1, & (y_k - \widetilde{y}_k) \geq 0 \end{cases} \quad (2.30)$$

where

H_k = gain matrix of the system at the sampling time k

p_k = switching gain matrix at the sampling time k

$sgn(y_k - \widetilde{y}_k)$ = discontinuous feedback at the sampling time k

By implementing the discontinuous feedback function into the input state equation, the SMO method can provide a sliding regime on the error values between the measured and calculated data of the battery system. The user can further adjust the switching gain value of the discontinuous feedback function to manipulate the effect of the sliding regime. This will enhance the robustness and minimize the inaccuracies of the battery system. Therefore, the SoC of the battery system can be accurately estimated based on the input data from the calculation and measurement models, as the discontinuous feedback function effectively compensates for modeling inaccuracies in the non-linear battery system (Ali et al., 2019).

2.3.3.9 Proportional Integral Observer (PIO) Method

Alternatively, the PIO method is the second technique that is being utilized to tackle the issue of the LO method being unable to be used in the non-linear measurement and calculation models of the LIB and LiBat systems. Therefore, similar to the SMO method, the PIO method will also work within the framework of the LO method. According to Hu, Sun, and Zou (2010), the PIO method included the integral continuous feedback function at the back of the input state equation of the LO method, as shown below:

$$\begin{aligned}\dot{\widetilde{x}_k} = & A_{k-1}\widetilde{x}_{k-1} + B_{k-1}u_{k-1} + \dots \\ & \dots K_{p_k}(y_k - \widetilde{y}_k) + K_{i1_k}I_{PI_k}\end{aligned}\quad (2.31)$$

$$\dot{I_{PI_k}} = K_{i2_k}(y_k - \widetilde{y}_k) \quad (2.32)$$

where

K_{p_k} = proportional gain matrix of the system at the sampling time k

K_{i1_k} = integral gain matrix 1 at the sampling time k

K_{i2_k} = integral gain matrix 2 at the sampling time k

I_{PI_k} = continuous feedback function at the sampling time k

$\dot{I_{PI_k}}$ = rate of change of continuous feedback function at the sampling time k

Similar to the SMO method, the implementation of the integral continuous feedback function into the input state equation by the PIO method can provide a potential integral tuning on the error values between the measured and calculated data of the battery system. The user is allowed to further adjust the two different integral gain values on the filtered state input equation and the integral continuous feedback function repeatedly to manipulate the effect of the tuning to enhance the robustness and minimize inaccuracies of the battery system. There are two different parameters to be tuned for the PIO method, which also gives it a higher degree of freedom for the user while utilizing it to accurately estimate the SoC value in the non-linear battery system compared to the SMO method (Ali et al., 2019).

2.4 Overview of Data-Driven SoC Estimation Methods

The data-driven estimation method is an advanced and intelligent approach capable of providing precise SoC estimation for LIB and LiBat systems. According to Xiong et al. (2018), the complex internal factors such as the thermodynamics and electrochemical behaviors and unpredictable external conditions of LIB and LiBat systems pose challenges for accurately estimating the SoC values by simply using the conventional SoC estimation methods that utilize the linear mathematical models mentioned in Section 2.3. Henceforth, data-driven estimation methods will apply non-linear estimation models to effectively resolve these challenges.

As the name of the data-driven estimation method suggests, the non-linear estimation models operate by analyzing the interpretation of the data sets that demonstrate complex connections between various input and output parameters, such as ambient temperature, load current, terminal voltage, and SoC values of the LIB and LiBat systems, without relying on any specific mathematical models or assumptions. Therefore, the data-driven estimation method has gained popularity in recent years in industrial AMR applications. The most commonly seen data-driven estimation methods in the current day include machine learning and fuzzy logic techniques.

2.4.1 Machine Learning Method

The machine learning methods are data-driven estimation approaches that utilize the black-box model with the capabilities of parallel distributed processing and AI learning to provide precise SoC estimation for the LIB and LiBat systems. Based on Xiong et al. (2018), the black-box model employs nonlinear estimation models that consist of the ability to self-learn, correct, and tune to accurately represent the complex connections between various input and output parameters of the LIB and LiBat systems.

Among all the data-driven estimation approaches, machine learning methods have gained the most popularity in recent years in industrial AMR applications. This is due to the black-box model's parallel distributed processing and AI learning ability, which allow for high computation rates. The machine learning methods can quickly provide the estimated value of the SoC by simply providing the battery status data into the pre-trained black box, as shown in

Figure 2.9. The popular machine learning methods among industrial AMR applications include artificial neural networks (ANN) and support vector machines (SVM) methods.

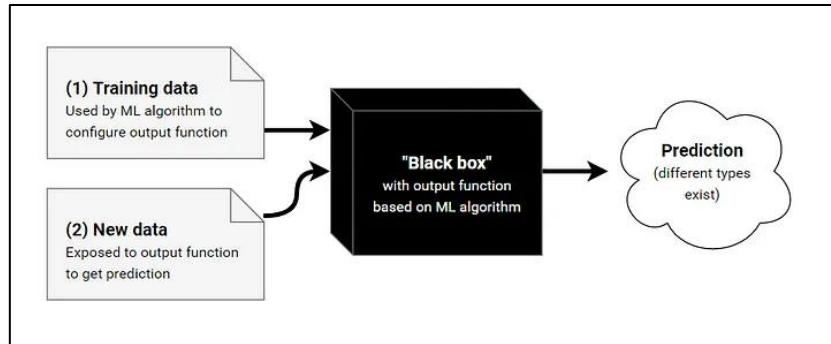


Figure 2.9: General Structure of Black-Box Model (Natakarnkitkul, 2021b).

Reprinted with permission from Copyright 2019 Medium.

However, the machine learning method is highly sensitive to the quality of the training samples and the hyperparameters of the black box models. This could result in poor convergent performance, leading to either overfitting or underfitting of the models. Apart from that, the machine learning methods are purely based on the system data base; relatively larger data storage will be needed as well. Therefore, this also highlights the primary limitation of the machine learning methods, as they necessitate a comparatively higher level of computational effort and expense.

2.4.1.1 Artificial Neural Network (ANN) Method

ANNs have quickly become one of the most widely used machine learning methods, forming the backbone of modern algorithms. Inspired by the human brain's intricate structure, ANNs aim to replicate its information processing capabilities. As shown in Figure 2.10, an ANN consists of three layers: input, hidden, and output. The input layer, similar to sensory cells, is responsible for receiving input signals. These signals are then passed through the hidden layers, which consist of interconnected neurones that perform a series of complex computations, mimicking the neural networks. Finally, the output layer, comparable to end-effector cells, produces the processed information,

converting it into output signals. This layered structure enables ANNs to efficiently learn and adapt, making them indispensable tools in various fields.

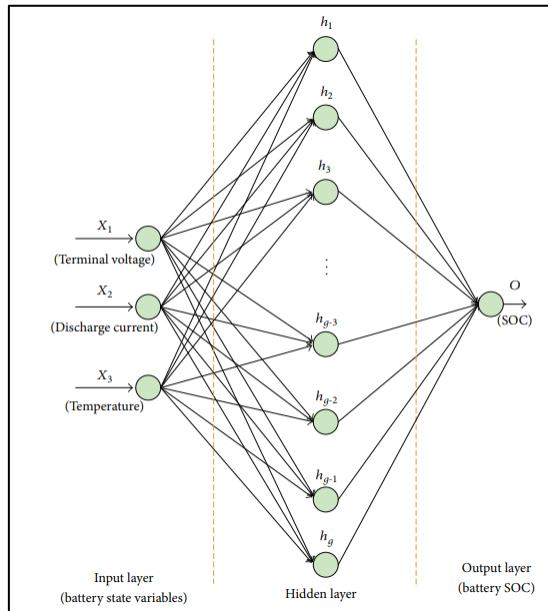


Figure 2.10: Structure of ANN for SoC Estimation (Chang, 2013).
Reprinted with permission from Copyright 2013 Wiley.

Subsequently, according to Chang (2013), the structure of the ANN model in terms of SoC estimation can be described as follows: An input layer, an output layer, and multiple hidden layers form the structure of the ANN. For the input layer, it will be representing the battery internal parameters, such as the terminal voltage, load current, and ambient temperature, with multiple neurons scattered and connected to the hidden layers. While the hidden layers will shrink down to the output layer with only one neuron for the value of the SoC of the battery, as shown in Figure 2.10. In between the layers, the neuron will multiply the incoming value from the previous layer with its own weight and combine through the summarization equation shown below:

$$net_{ij} = \sum_1^n x_i w_{ij} \quad (2.33)$$

where

x_i = input value from the input layer of i

w_{ij} = weight of the neuron that is connected between the input layer of i and the hidden layer of j

net_{ij} = sum of the multiplication between the w_{ij} and x_i that are connected to the hidden layer of j

Then, the activation equation will process the summation of the multiplication between each input value from the input layer and the weight of the neurons connected to the hidden layers. The activation equation can be represented as the hyperbolic tangent and sigmoid functions based on different types of ANN methods, including BPNN and feedforward neural network, as shown below, respectively:

$$h_j = f(net_{ij}) = \frac{1-e^{-2net_{ij}}}{1+e^{-2net_{ij}}} \quad (2.34)$$

$$h_j = f(net_{ij}) = \frac{1}{1+e^{-net_{ij}}} \quad (2.35)$$

where

h_j = output value from the hidden layer of j

$f(net_{ij})$ = activation function of the input layer of i and hidden layer of j

In addition, the similar process outlined previously will also go through the hidden layers, output layers, and neurons connected between them using the summarization and activation equations to get the estimated SoC value of the battery. Other than the general ANN methods, the more advanced and modern ANN methods, such as the radial basis neural network, wavelet neural network, deep neural network, LSTM, GRU and CNN that utilize softmax, radial basis, wavelet, and relu transform functions, will be applied for higher accuracy and speed of estimation performance in exchange for the complexity of the models. The ANN method is highly applicable for highly dynamic and non-static applications, like industrial AMR applications.

Apart from that, based on Pannakkong et al. (2022), the performance of the ANN method can be ensured with properly tuned hyperparameters such as the learning rate, number of bias inputs, regularization value, number of

hidden layers, epoch times, and type of sample training (online, batch, and mini-batch trainings). Although there is no absolute formula or equation for defining the correct hyperparameters for the ANN method, with a simple and straightforward process for tuning the hyperparameters, the effect of this disadvantage can be omitted.

2.4.1.2 Support Vector Machine (SVM) Method

The SVM method has gained significant attention in recent years as an effective machine learning approach, particularly for tasks involving classification and regression. Like the ANN method, SVM relies on advanced pattern recognition techniques to analyze the input-output relationships within a battery system's database. What sets SVM apart is its use of a kernel function, which enables it to handle complex, non-linear data more effectively.

Initially, the SVM method applies the kernel function to transform the mixed input-output data—such as terminal voltage, load current, ambient temperature, and SoC—from a seemingly inseparable lower-dimensional, two-dimensional (2D) space into a higher-dimensional, three-dimensional (3D) space. This transformation allows the SVM to separate data points that appear inseparable in the original 2D space, as illustrated in Figure 2.11 (Patle, and Chouhan, 2013). By mapping the data into this higher dimension, the SVM method creates a clear separation between different classes, making it an effective tool for evaluating and predicting battery system behaviour.

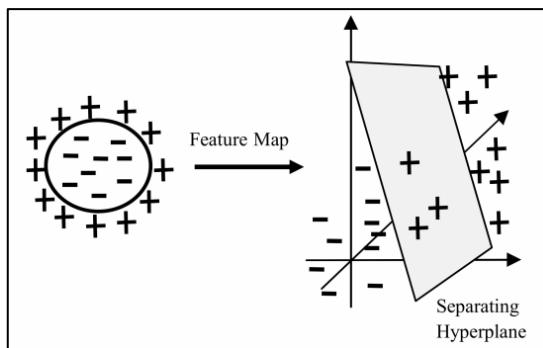


Figure 2.11: General Portrait of the SVM (Patle, and Chouhan, 2013).
Reprinted with permission from Copyright 2013 IEEE.

In addition, Figure 2.12 illustrates that the classification algorithms will identify the hyperplane and decision boundaries (soft margin) for the LIB and LiBat systems. The hyperplane implies that the nearest points in the input-output database for each feature and the separation plane between them are maximized in terms of distance. While the decision boundaries (soft margin) will filter out the data points that are offset from the input-output database by noise, disturbance, and drift within the LIB and LiBat systems, this enables effective segregation and categorization of the input-output database of the LIB and LiBat systems without bias.

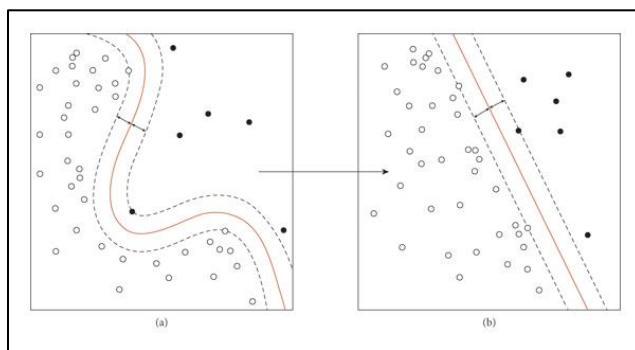


Figure 2.12: General Portrait of the Decision Boundaries (Soft Margins) and Hyperplane (Pannakkong et al., 2022). Reprinted with permission from Copyright 2022 Wiley.

Henceforth, based on the kernel function, hyperplane, and decision boundaries (soft margin) that are being applied to the non-linear and dynamic input-output databases of the LIB and LiBat systems, the linearized LIB and LiBat systems will be able to present the regression relationship between the input-output databases and produce the estimated SoC value of the LIB and LiBat systems, shown in Figure 2.12.

Nevertheless, unlike the ANN method, the SVM method presents challenges in tuning the hyperparameters. This is because the SVM method required a selection of the optimal kernel function for SoC estimation in LIB and LiBat systems. However, based on Patle and Chouhan (2013), the kernel function can take various forms, such as linear, polynomial, radial basis, and sigmoid, as shown below, respectively:

$$K(x_i, x_j) = (x_i^T \cdot x_j) \quad (2.36)$$

$$K(x_i, x_j) = (1 + x_i^T \cdot x_j)^n \quad (2.37)$$

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2} \quad (2.38)$$

$$K(x_i, x_j) = \tanh(j x_i^T \cdot x_j + r) \quad (2.39)$$

where

x_i = vector of the input data of i

x_j = vector of the input data of j

n = degree of polynomial of the kernel function

γ = kernel parameter of the kernel function

$K(x_i, x_j)$ = kernel function

In addition, based on Pannakkong et al. (2022), after the optimal kernel function is selected, it is also essential for the SVM method to tune around with the kernel parameter for the kernel function and the size parameter of the decision boundaries (soft margin) to employ the most optimum parameter values. This will enable the creation of an accurate linearized battery system with minimal linearized error.

Therefore, it is difficult to apply a precise pattern recognition process to the input-output database of the LIB and LiBat systems. The SVM method must employ not only the appropriate kernel function but also the optimal kernel parameter and decision boundaries (soft margin). In addition, the SVM approach has a higher complexity compared to other methods due to the requirement of sophisticated matrix operations and quadratic programming. Thus, the SVM method is less likely to be applicable for industrial AMR applications compared to the ANN method due to its difficulty in tuning the hyperparameters and programming.

2.4.2 Fuzzy Logic Method

The fuzzy logic method has been widely recognized as one of the most popular data-driven estimating methods in recent years. The popular fuzzy logic methods among industrial AMR applications include fuzzy logic controller (FLC) and adaptive neuro-fuzzy inference system (ANFIS) methods. In general,

the overall structure and working principles of the fuzzy logic method are directly influenced by the framework of human cognition and reasoning in decision-making.

The fuzzy logic methods adapt the concept of "degrees of truth" logic instead of the traditional computational logic of "true or false" (1 or 0) Boolean logic to provide precise SoC estimation for the LIB and LiBat systems. The fuzzy logic methods employ various input and output parameters of the LIB and LiBat systems and categorize them into different classes from 1 to 0, such as "very high" (0.9), "high" (0.75), "medium" (0.5), "low" (0.25), and "very low" (0.1), according to the concept of "degrees of truth" logic instead of the Boolean logic that categorized the input and output parameters of the LIB and LiBat systems only into the two classes.

Among all the data-driven estimation methods, fuzzy logic methods have gained quite a lot of popularity in industrial AMR applications, second only to machine learning methods. This is because the fuzzy logic method is less sensitive to the quality of the training samples compared to machine learning. This allows it to provide an estimated value of the SoC by simply providing the battery status data to the fuzzy system due to its ability to model human cognition and reasoning in making decisions and produce a better reflection of the non-linear, complex, and dynamic nature of real-world working situations than the classical computation method.

However, according to Ali et al. (2019), similar to machine learning methods, fuzzy logic methods also require a significant amount of data storage, computational effort, and expense. Even though the fuzzy logic method can work on indeterministic datasets, the output results are often a compromise estimate rather than an actual estimated value. Then, unlike machine learning methods that can adjust the hyperparameters of the black box models to guarantee their fit, fuzzy logic methods rely solely on the types of selected membership functions (MF) and fuzzy rule bases to ensure their performance, making them less appealing than machine learning methods.

2.4.2.1 Fuzzy Logic Controller (FLC) Method

The FLC method is one of the most popular types of fuzzy logic in recent years, and it is also one of the fundamental frameworks for many of the modern fuzzy

logic algorithms. According to Umair Ali et al. (2018), the FLC method operates in three distinct processes, which are fuzzification, fuzzy computation (fuzzy rule base and fuzzy inference), and defuzzification processes, as shown in Figure 2.13.

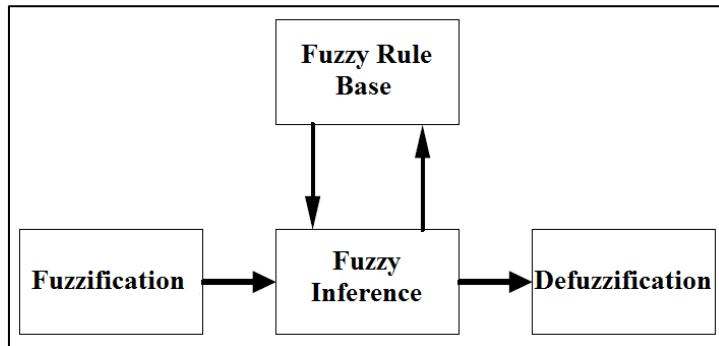


Figure 2.13: General Structure of FLC (Nur, Omaç, and Öksüztepe, 2017b).
Reprinted with permission from Copyright 2017 Sigma.

Firstly, Figure 2.14 demonstrates the fuzzification process, which involves transforming the input-output database of the LIB and LiBat systems, including terminal voltage, load current, ambient temperature, and SoC, from a true value format to a fuzzy linguistic set format (1 to 0). This is achieved by employing a MF and fuzzy controller, both of which are based on the "degrees of truth" logic previously mentioned in Section 2.4.2. Furthermore, the fuzzy rule base and inference engine cooperate with each other to perform the fuzzy computation process on the input-output fuzzy linguistic data sets.

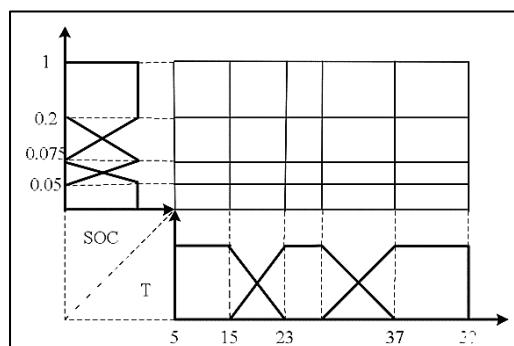


Figure 2.14: 2D MF of FLC for SoC Estimation (Du et al., 2014).
Reprinted with permission from Copyright 2014 IEEE.

The fuzzy rule base serves as the primary control logic and parameters of the fuzzy logic system, which are made out of a series of IF-THEN statements. The fuzzy inference engine then functions as the fuzzy operator within the fuzzy system. The engine will process the input-output fuzzy linguistic data sets using control logic and parameters from the fuzzy rule base, resulting in the production of the output fuzzy linguistic data sets. Finally, the defuzzification process, which is the opposite of the fuzzification process, will occur, producing the output data by converting the output fuzzy linguistic data sets back into their original format values.

Nevertheless, as previously mentioned in Section 2.4.2, compared to the simple and straightforward tuning of the hyperparameters of the machine learning methods, the FLC method presents challenges in undergoing simple tuning over their performance. This is because the FLC method required a selection of the optimal fuzzy controller and MF for SoC estimation in LIB and LiBat systems. However, the MF can take various forms, such as singleton, triangular, trapezoidal, sinusoidal, and Gaussian, and is further classified into types 1 and 2, each offering different degrees of freedom to the fuzzy system (Ren, Balazinski, and Baron, 2011).

Other than that, it is also difficult to select the most optimal fuzzy rule base for the FLC method, as there is no absolute approach to defining the correct series of IF-THEN statements to be coded inside the fuzzy system, except based on the opinion and experience of the past researchers and numerous attempts at the try-and-error approach. Therefore, in order to achieve high accuracy and optimal speed in SoC estimation, spending effort in coding the optimal fuzzy rule base and choosing a fuzzy controller and MF with the correct degree of freedom is crucial and critical.

2.4.2.2 Adaptive Neuro-Fuzzy Inference System (ANFIS) Method

According to Ali et al. (2019), the ANFIS method, also known as fuzzy neural networks, is a type of fuzzy logic method that developed from the idea of the Takagi-Sugeno fuzzy inference system, which combines both the FLC and ANN method frameworks. Additionally, the ANFIS method shares the same structure as the FLC method; however, it further divides each of these parts into multiple layers and nodes, akin to the ANN method as discussed in Section

2.4.1.1. There are five distinct layers, which are the fuzzification (IF part), fuzzy computation (rules, norms, and THEN parts), and defuzzification layers (output part), as shown in Figure 2.15 (Yeom, and Kwak, 2018).

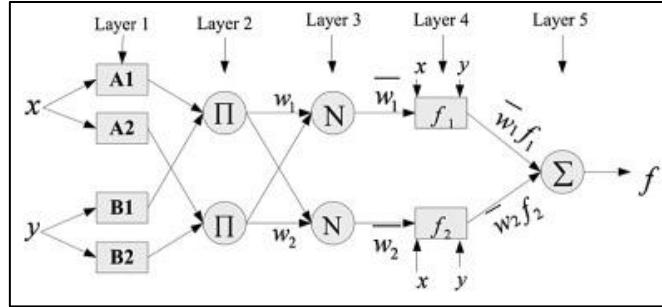


Figure 2.15: General Structure of ANFIS (Dai et al., 2015). Reprinted with permission from Copyright 2015 Elsevier.

Firstly, the first layer of the ANFIS method, similar to the FLC method mentioned in Section 2.4.2.1, will transform the input-output database of the LIB and LiBat systems, including terminal voltage, load current, ambient temperature, and SoC, from a true value format to a fuzzy linguistic set format (1 to 0) by employing an MF and fuzzy controller, as shown below:

$$\begin{pmatrix} x \\ y \\ \vdots \end{pmatrix} \rightarrow \begin{pmatrix} u_{A_i}(x), i = 1, 2, \dots \\ u_{B_i}(y), i = 1, 2, \dots \\ \vdots \end{pmatrix} \quad (2.40)$$

where

x = true format input value of x

y = true format input value of y

$u_{A_i}(x)$ = fuzzy linguistic input value of x from the first layer at the node A_i

$u_{B_i}(y)$ = fuzzy linguistic input value of y from the first layer at the node B_i

Moreover, at the second layer of the ANFIS method, the input-output fuzzy linguistic data sets will be connected to respective nodes in the rules part. Each fuzzy linguistic input-output data set with the same "degrees of truth" will multiply with each other, producing a weight for its respective nodes based on the multiplication equation, as shown below:

$$w_i = u_{A_i}(x) \times u_{B_i}(y) \times \dots, \quad i = 1, 2, \dots \quad (2.41)$$

where

w_i = weight of the second layer at the node i

Then, at the third layer of the ANFIS method, the weight at the rules part will be connected to the respective nodes at the norm part. Each of the weights will produce the ratio between the weight of the node from the rules part and the total weight and represent the normalized value for its respective nodes based on the weighting factor equation, as shown below:

$$\bar{w}_i = \frac{w_i}{\sum_1^n w_i} \quad (2.42)$$

where

\bar{w}_i = weighting factor from the third layer at the node i

In addition, at the fourth layer of the ANFIS method, the weighting factor at the norm part will be connected to the respective nodes at the THEN part. Each of these weighting factors will then go through the first-order linear equation for its respective nodes, as shown below:

$$\bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (2.43)$$

where

p_i, q_i and r_i = coefficients of linear equation from fourth layer at node i

$\bar{w}_i f_i$ = output value of linear equation at node i

Lastly, at the fifth layer of the ANFIS method, the output value of the linear equation at the THEN part will be connected to the single node at the output part. Each of these output values of the linear equation will then be summed up based on the summarization equation, as shown below:

$$f = \sum_1^n \bar{w}_i f_i \quad (2.44)$$

where

f = true format output estimated value

Henceforth, due to the unique structure of the ANFIS method as mentioned previously, it is able to obtain the advantages of both fuzzy logic and machine learning methods to provide a much more accurate value of the estimated SoC for the LIB and LiBat systems compared to the FLC method as mentioned in Section 2.4.2.1. This is because, by combining these two methods, the issue of the FLC method being unable to select the most optimal fuzzy rule base can be solved. By assigning simple and straightforward first-order linear equations and utilizing the power of the ANN method in self-learning, correcting, and tuning, the ANFIS method can effectively and efficiently set up the rule base (Chang, 2013). Hence, the ANFIS method has also become one of the hot picks for fuzzy logic systems in recent years. The only drawback is that the ANFIS method requires a higher amount of computation power than the FLC method due to its more complex structure.

2.5 Summary

To summarize, the literature review introduces that, among the battery systems, the LIB and LiBat systems are more preferable for industrial AMR applications in terms of safety concerns. Currently, the most commonly used LIB and LiBat systems are the lithium-ion, Li-FP, and Li-NMC graphite battery systems. However, a variety of LIB and LiBat systems, including Li-NCA, Li-NCA, and Li-PO graphite battery systems, hold significant potential for industrial AMR applications in the future. Li-NO and Li-MO graphite battery systems are not as advantageous as the others.

In addition, it also introduced the topic of the commonly used and accessible forms of conventional SoC estimate models by categorizing them into three main groups: look-up tables, Coulomb counting, and model-based techniques. Within the category of look-up table methods, there are seven different approaches to choose from: EMF, IS, IR, ambient temperature, terminal voltage, load current, and OCV. Furthermore, the model-based methods comprise two separate components: the mathematical models for the battery systems and the adaptive system. The mathematical models of battery

systems can be classified into three categories: EM, ECM, and EIM models. Among these, the ECM model is the most widely used due to its simplicity in producing the mathematical models. There exist two distinct categories of models for adaptive systems: observer-based models and filter-based models. The filter-based model is preferred due to its greater flexibility to handle both deterministic and uncertain input datasets for LIB and LiBat systems.

Apart from that, the popular types of data-driven SoC estimation models for the LIB and LiBat systems are machine learning and fuzzy logic methods. The research looks at two types of SoC estimation models: conventional and data-driven. The data-driven SoC estimation models can match themselves to the battery system's input-output data sets by analyzing their interpretation. Conversely, the battery system's various internal and external factors have a significant impact on conventional SoC estimation models, preventing them from accurately estimating SoC values from uncertain input datasets. In exchange for the more complicated modeling process, relatively larger data storage, and a comparatively higher level of computational effort and expense, data-driven SoC estimation models allow for more accurate estimation results than conventional models.

Then, the popular machine learning methods consist of ANN and SVM methods, while the fuzzy logic methods include FLC and ANFIS methods. Out of these two methods, the machine learning method is the most preferred approach, as opposed to the fuzzy logic method. This is due to the fact that the machine learning methods consist of the AI's ability to self-tune and fix their internal parameters to match the input-output data sets of the LIB and LiBat systems. Then, its parallel distributed processing allows for higher computation rates, enabling it to quickly provide the estimated value of the SoC using the provided battery status data. Last but not least, one of the most important reasons for its applicability in industrial AMR applications is its simple and straightforward tuning of hyperparameters in exchange for performance, despite its oversensitivity to the quality of the input-output data sets of the LIB and LiBat systems compared to fuzzy logic methods.

CHAPTER 3

METHODOLOGY AND WORK PLAN

3.1 Introduction

A systematic approach with several project steps is essential for creating accurate SoC estimation models for LIB and LiBat systems. This chapter outlines the process as follows: Firstly, it involves identifying and collecting relevant internal parameter datasets during charging and discharging to form input-output and conventional model parameter databases sourced from various sources. Next, the focus shifts to leveraging existing codes from the open-sourced platform to develop optimal conventional and data-driven SoC estimation models. Following the fine-tuning of the SoC estimation models, the models will undergo training, testing, and modeling using the collected internal parameter datasets. Then, the SoC estimation models will be validated and evaluated using R^2 , MAE, and RMSE performance metrics, comparing the values of the estimated SoC to the experimental SoC. As a result, the optimization process will involve scaling the input-output databases and fine-tuning hyperparameters and parameters of the SoC estimation models. The goal is to achieve optimal estimation performance with an R^2 greater than 0.8, MAE below 10, and RMSE under 15, ensuring reliable benchmarking between the SoC estimation models. Finally, the chapter will conclude by benchmarking the performance of each SoC estimation model to identify the best-performing data-driven model and determine its necessity.

3.2 Collecting LIB and LiBat Systems Internal Parameter Datasets

The primary task of this project is to identify input-output and conventional model parameter databases that accurately represent the internal parameters of LIB and LiBat systems during charging and discharging cycles. These databases include Li-FP, Li-NMC, Li-NCA, Li-PO, and Li-Ion graphite battery systems, as detailed in Section 2.2.2. These battery systems are commonly used in industrial AMR and other applications. The comprehensive input-output databases will be utilized to train, validate, and test data-driven SoC estimation models. These models aim to provide accurate SoC estimates for the different

battery systems under various conditions, such as working rates, states, and environmental temperatures. The study will use these databases to model and evaluate conventional SoC estimation methods.

The data collection process for LIB and LiBat systems involved gathering internal parameter datasets from various repositories, including Mendeley Data, Borealis, the Centre for Advanced Life-Cycle Engineering (CALCE), and Recherche Data Gouv. These datasets were provided by Kollmeyer et al. (2020), Kollmeyer and Skells (2020), Kollmeyer and Skells (2023), Xing et al. (2014), Catenaro and Onori (2021b), and Redondo-Iglesias and Tian (2024). The input datasets encompass measurements of load current, terminal voltage, and ambient temperature, while the conventional model parameter datasets include OCV capacity and terminal voltage-load current measurements. Both datasets were obtained during the charging and discharging cycles of several LIB and LiBat systems. The output experimental SoC across all datasets was calculated using coulomb counting and lookup tables, following the methodologies outlined by the respective authors.

The first dataset features the LG 18650HG2 3000mAh 3.6V Li-NMC graphite battery, sourced from Kollmeyer et al. (2020) via Mendeley Data. Tests were conducted at six different temperatures (40°C, 25°C, 10°C, 0°C, -10°C, and -20°C), with a steady 1C charge applied after each test. The battery underwent a dynamic four-pulse discharge HPPC test (1C, 2C, 4C, and 6C discharge; 0.5C, 1C, 1.5C, and 2C charge), a steady C/20 discharge and charge test, and additional steady discharge tests at 0.5C, 2C, and two 1C rates. Dynamic drive cycles (UDDS, HWFET, LA92, and US06), along with eight dynamic mixed drive cycles, were also conducted to generate the input datasets. OCV-capacity datasets were derived from the steady C/20 charging and discharging tests under the same temperature conditions.

The second dataset concerns the Samsung IN21700-30T 3000mAh 3.6V Li-Ion graphite battery, provided by Kollmeyer and Skells (2023) via Mendeley Data. Tests were conducted under the same temperature conditions and steady charging protocols as the first dataset. The battery was subjected to a dynamic four-pulse discharge HPPC test (1C, 2C, 6C, and 12C discharge; 0.5C, 1C, 2C, and 4C charge), steady discharge tests at 0.5C, 2C, two 1C rates, and a C/20 rate, and a steady C/20 charge test. Then, dynamic drive cycles

(UDDS, HWFET, LA92, and US06) and eight dynamic mixed drive cycles were conducted to generate the input datasets. OCV-capacity datasets were similarly derived from identical tests under the same temperature settings.

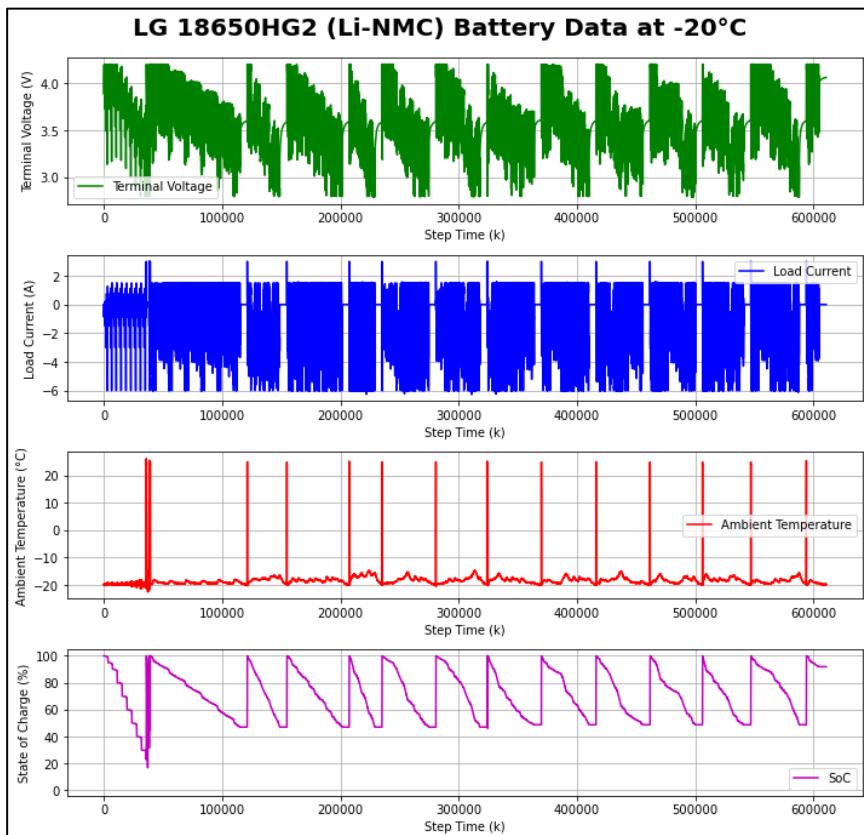
The third dataset, sourced from Borealis and provided by Kollmeyer and Skells (2023), features the SB LiMotive 5200mAh 3.7V Li-NCA graphite battery. Testing was conducted under consistent temperature conditions, with a steady 1C charge rate following each test. The battery underwent a four-pulse dynamic discharge HPPC test (1C, 2C, 6C, and 12C discharge; 0.5C, 1C, 2C, and 4C charge), a steady C/20 discharge and charge test, and steady discharge tests at 0.5C, 2C, and two 1C rates—one before the dynamic UDDS cycle and another before the dynamic HWFET, LA92, and US06 cycles. Dynamic drive cycles (UDDS, HWFET, LA92, and US06) and eight dynamic mixed drive cycles were used to generate the input datasets. OCV-capacity datasets were derived from similar measurements as previously described.

The fourth dataset includes the Turnigy Graphene 5000mAh 3.7V Li-PO graphite battery, sourced from Kollmeyer and Skells (2020) via Mendeley Data. Testing followed the same temperature and steady charging protocols as the previous datasets, with a four-pulse dynamic discharge HPPC test (1C, 2C, 5C, and 10C), a steady C/20 discharge and charge test, steady discharge tests at 0.5C, 2C, and two 1C rates, dynamic drive cycles (UDDS, HWFET, LA92, and US06), and eight dynamic mixed drive cycles. OCV-capacity datasets were obtained through similar measurements as described above.

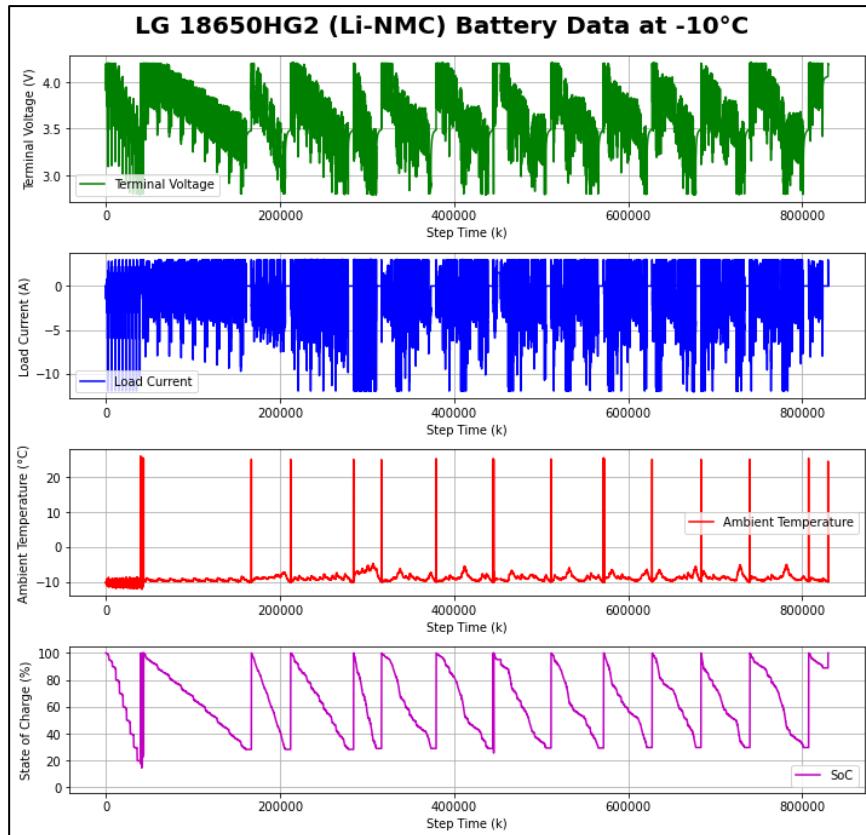
The fifth dataset includes A123 1100mAh, 2000mAh, and 4400mAh 3.2V Li-FP graphite batteries, sourced from Xing et al. (2014), Catenaro and Onori (2021b), and Redondo-Iglesias and Tian (2024) via the Centre for Advanced Life-Cycle Engineering (CALCE), Mendeley Data, and Recherche Data Gouv repositories, respectively. The A123 1100mAh battery was tested at six temperatures (40°C, 30°C, 25°C, 20°C, 10°C, and 0°C) with a steady 1C charge following each test. The tests included dynamic drive cycles (DST, FUDS, and US06) to generate input datasets. Terminal voltage-load current datasets were obtained through steady low-rate charging and discharging tests at five temperatures (40°C, 30°C, 25°C, 10°C, and 0°C), with OCV-capacity datasets derived from average terminal voltage during these cycles.

Apart from that, for the A123 2000mAh battery, six units were tested at three temperatures (35°C , 25°C , and 5°C) with steady 1C charging, followed by steady discharge tests at various rates (0.05C, 1C, 3C, 5C, 10C, 15C, and 20C). Lastly, the A123 4400mAh battery underwent testing at three temperatures (40°C , 25°C , and 0°C), including steady discharge and charge at a 1C rate, a dynamic seven-pulse HPPC test (0.5C to 8C discharge and charge), and dynamic drive cycles (UDDS and LA92) to complete the dataset.

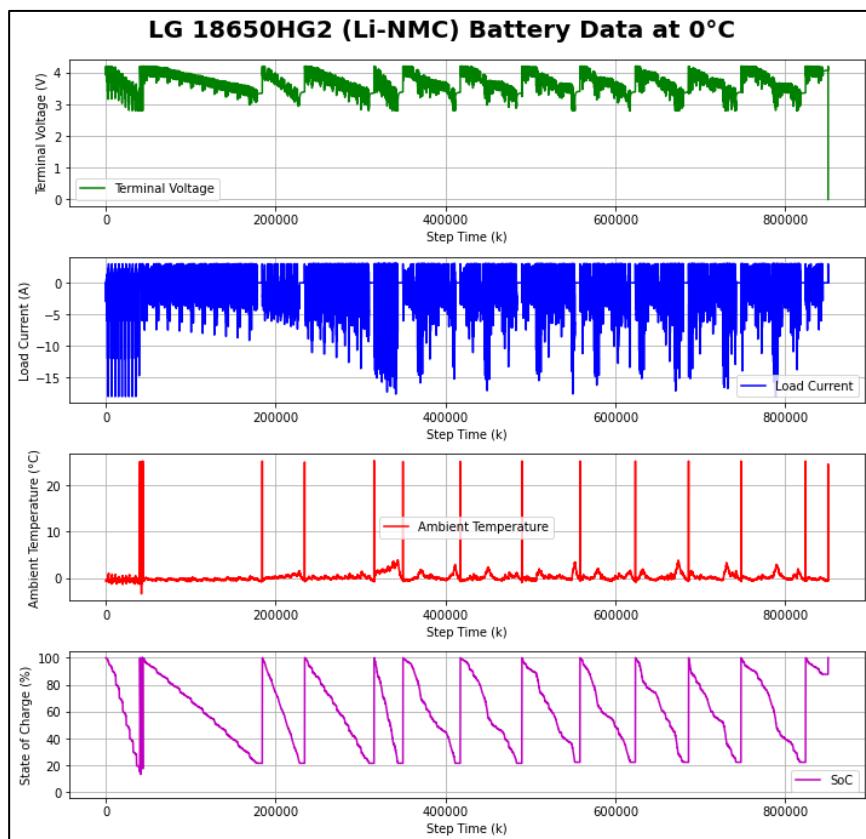
Figures 3.2 illustrates the relationship between OCV and terminal voltage as a function of SoC during both charging and discharging at various temperatures. This analysis includes five graphite battery systems: LG 18650HG2 (3000mAh 3.6V Li-NMC), Samsung IN21700-30T (3000mAh 3.6V Li-Ion), SB LiMotive (5200mAh 3.7V Li-NCA), Turnigy Graphene (5000mAh 3.7V Li-PO), and A123 (1100mAh 3.2V Li-FP). For additional insights into terminal voltage, load current, ambient temperature, and SoC, please refer to Figures 3.1 and Graphs A-1 to A-4, which also include data for A123's 2000mAh and 4400mAh Li-FP batteries.



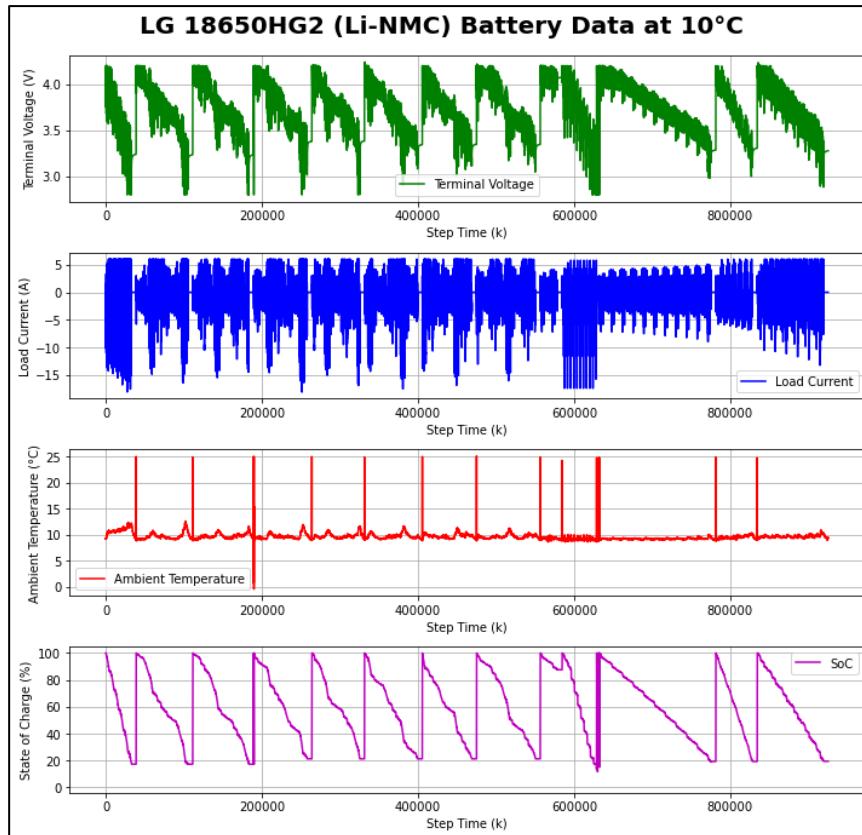
(a).



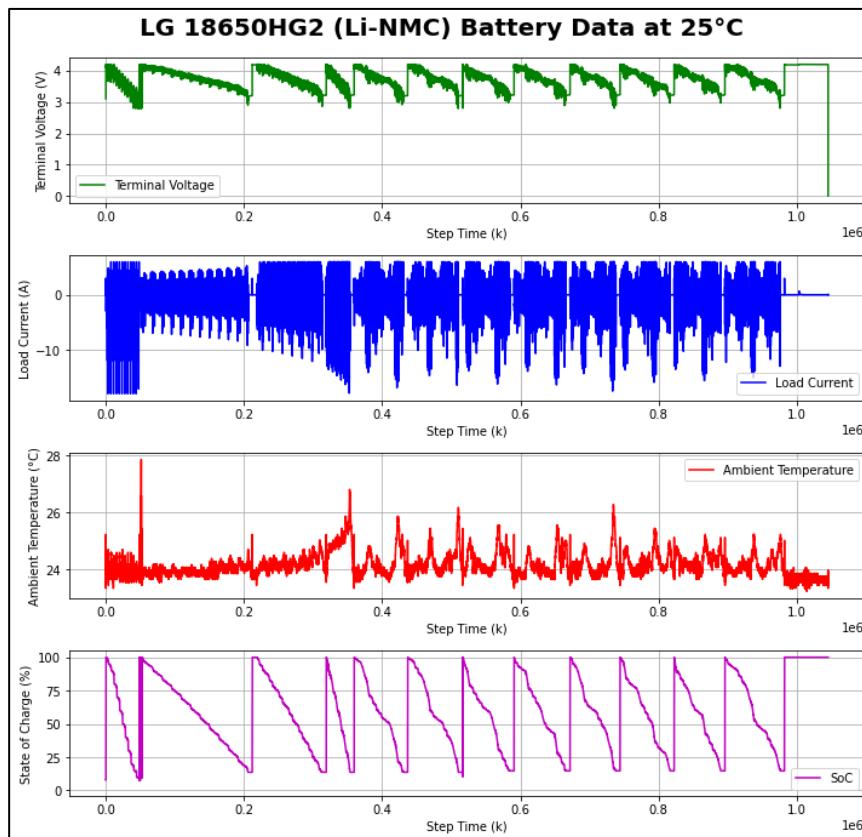
(b).



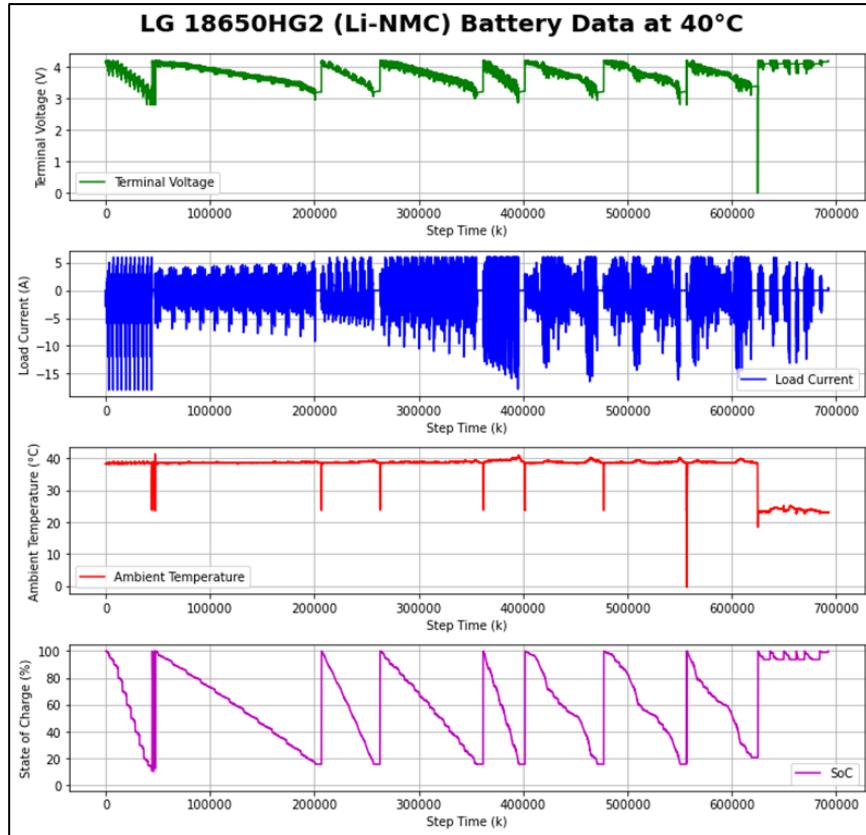
(c).



(d).

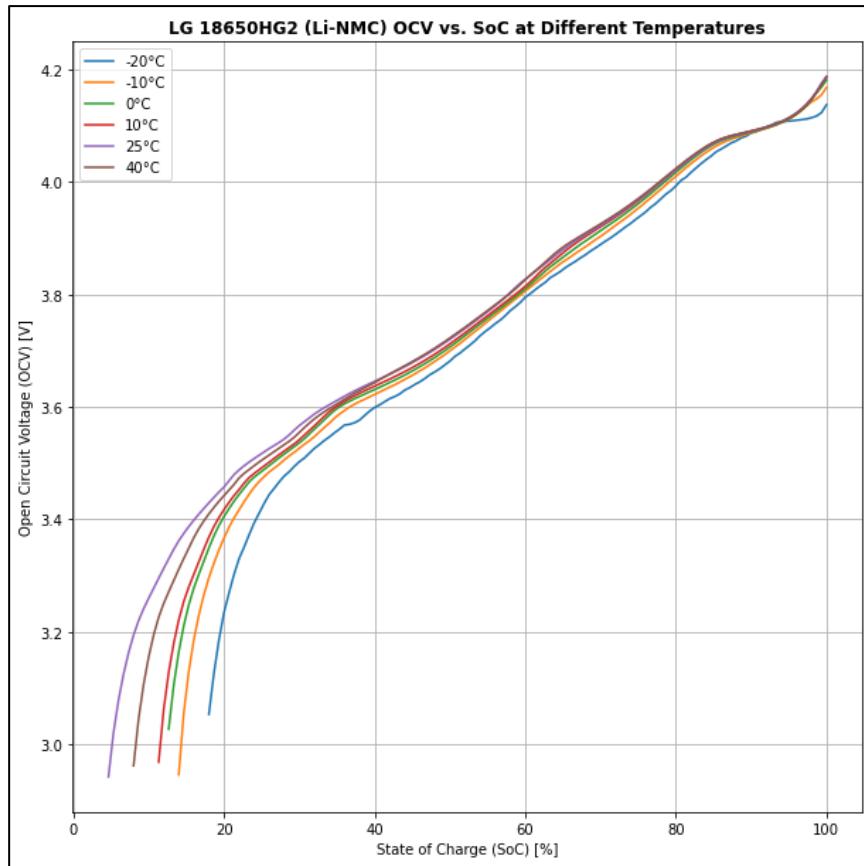


(e).

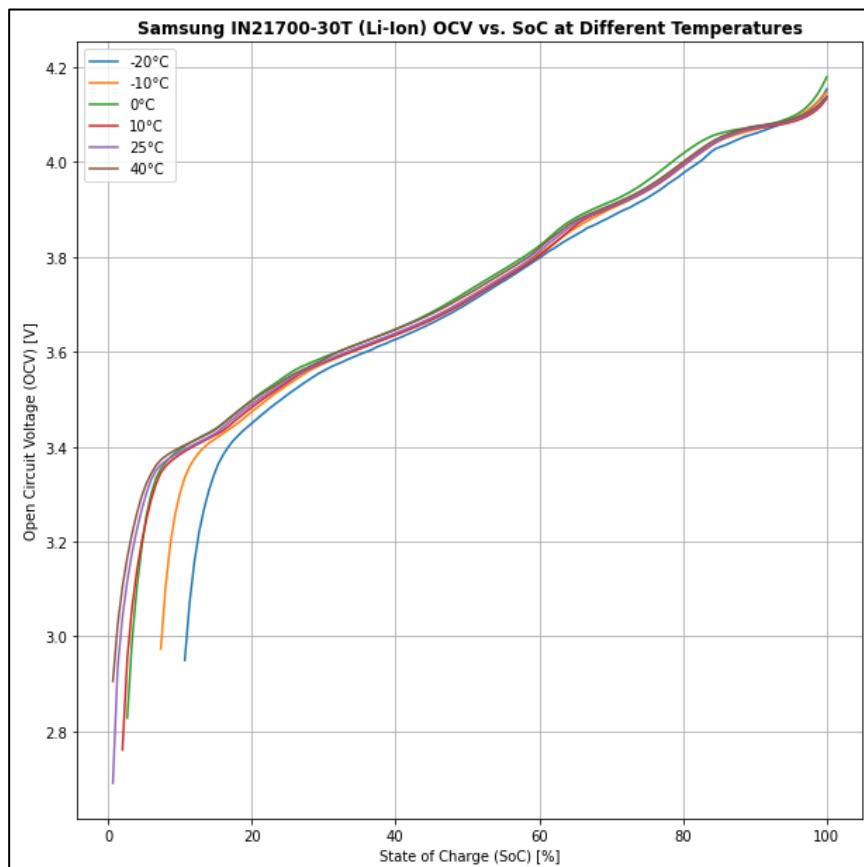


(f).

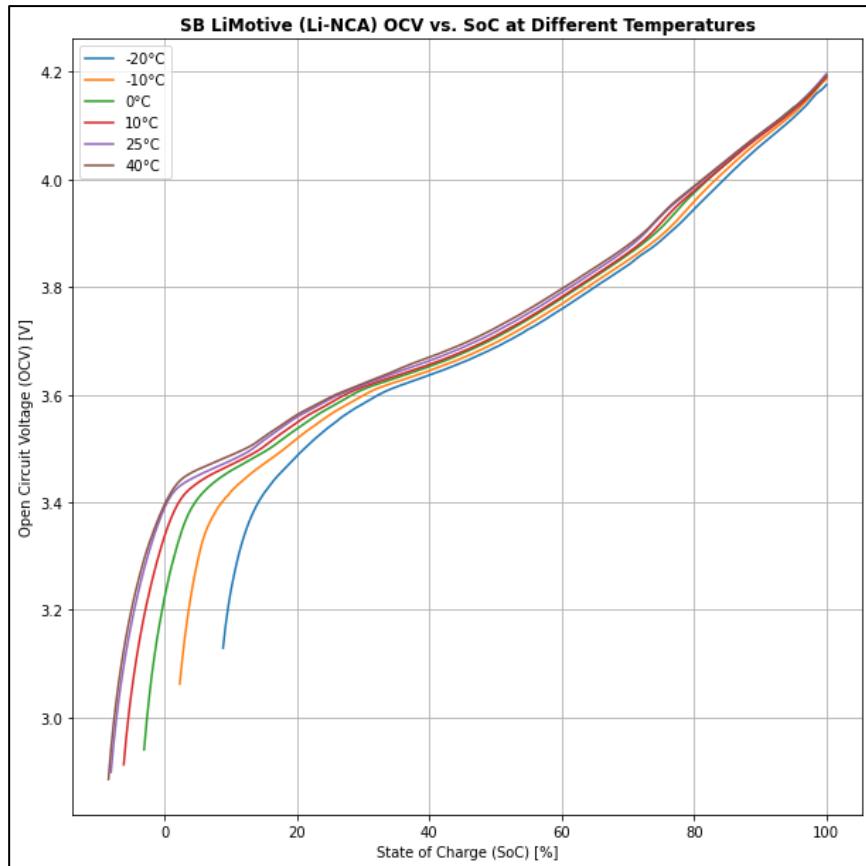
Figure 3.1: Discharge and Charging Curves for the LG 18650HG2 (Li-NMC) Battery: Terminal Voltage, Load Current, Ambient Temperature, and SoC at (a) -20°C, (b) -10°C, (c) 0°C, (d) 10°C, (e) 25°C, and (f) 40°C (Kollmeyer et al., 2020).



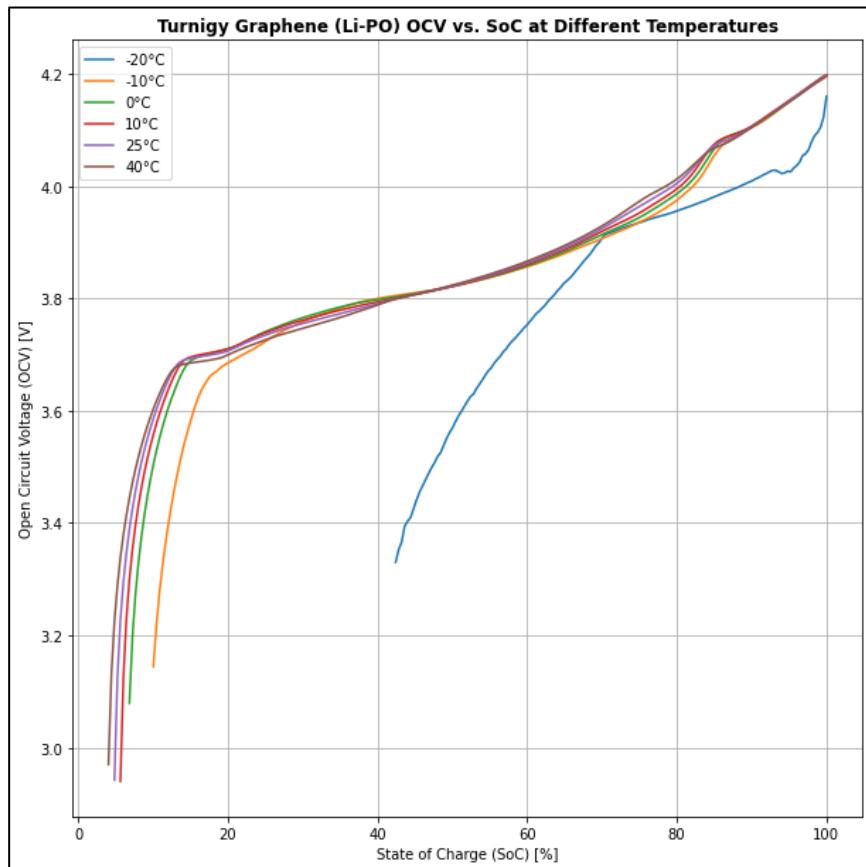
(a).



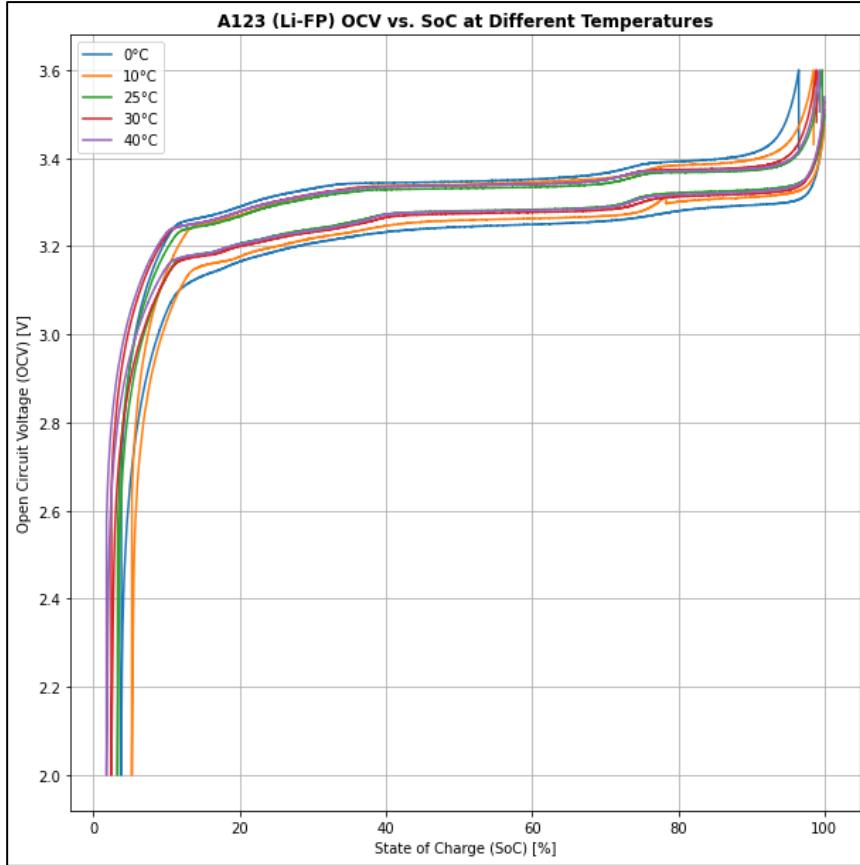
(b),



(c).



(d).



(e).

Figure 3.2: OCV vs. SoC Curves: (a) LG 18650HG2 (Li-NMC), (b) Samsung IN21700-30T (Li-Ion), (c) SB LiMotive (Li-NCA), (d) Turnigy Graphene (Li-PO); and Terminal Voltage vs. SoC Curves: (e) A123 (Li-FP) Battery (Kollmeyer et al., 2020; Kollmeyer, and Skells, 2020, 2023; Xing et al., 2014).

3.3 Developing LIB and LiBat Systems SoC Estimation Models

Additionally, the next focus of this project involves leveraging existing codes from the reliable open-source PyTorch library and GitHub repositories to develop the optimal conventional and data-driven SoC estimation models for producing estimated SoC values for the LIB and LiBat systems. The reliable and high-performing existing codes for the data-driven and conventional SoC estimation models will be provided by professional and experienced developers in the open-source library and repositories.

Literature research suggests that two distinct conventional SoC estimation models should be developed: model-based EKF and UKF methods,

both for LIB and LiBat systems. These models will be implemented on the Python platform, following the approaches outlined by Larchuto (2021) and Alter, Gu, and Cong (2022) in the Github repositories. Both the EKF and UKF methods will integrate with the ECM for accurate and straightforward SoC estimation. Meanwhile, it also indicates that among data-driven SoC estimation models, machine learning methods have proven to be more effective than fuzzy logic models. The most suitable and widely used machine learning models are the BPNN, CNN, LSTM, and GRU methods, which are extensions of the ANN method; all of these models will be developed using the Python platform. The development of the models will be based on the machine learning library PyTorch for performing tensor computation and mathematical operations on neural networks.

In contrast, the look-up table and coulomb counting methods, used as fundamental techniques for obtaining experimental SoC, will not be tied to any specific platform. Instead, these methods will provide the baseline SoC datasets within the input-output and conventional model parameter databases, which will be used to model the EKF and UKF models and evaluate their reliability and accuracy in producing the estimated SoC values. On the other hand, the input-output databases will be split into 80% as training data to train the BPNN, CNN, LSTM, and GRU models while being validated by 10% of the validation data and tested by 10% of the testing data.

3.4 Fine-tuning LIB and LiBat Systems SoC Estimation Models

In addition, the third focus of this project involves optimising the performance of the implemented SoC estimation models that were trained and modelled with the collected internal parameter datasets mentioned previously in Sections 3.2 for the LIB and LiBat systems. The performance of the trained and modelled SoC estimation models will be validated, tested, and evaluated using the metric values of R^2 , MAE, and RMSE, where the estimated SoC will be compared to the experimental SoC. Furthermore, to ensure reliable benchmarking between the trained data-driven and modelled conventional SoC estimation models, the performance metric values of these models will be optimised to achieve minimum thresholds of R^2 greater than 0.8, MAE below 10, and an RMSE under 15.

First and foremost, optimising conventional SoC estimation models involves calibrating internal parameters and refining the mathematical assumptions of linear models. According to Larchuto (2021) and Alter, Gu, and Cong (2022), ten key parameters are adjusted for the model-based EKF method: coulomb efficiency, maximum available capacity, initial SoC, internal resistance, dynamic resistance, capacitance of the ECM, measurement noise covariance, state error covariance, process noise covariance matrices, and the polynomial degree for simulating the OCV-SoC relationship. For the model-based UKF method, in addition to these ten parameters, four additional parameters—dimension, alpha, beta, and lambda values—are tuned.

The optimization of data-driven SoC estimation models involves scaling input-output datasets and calibrating hyperparameters for machine learning models, including BPNN, CNN, LSTM, and GRU. For all four models, common scaling methods like MinMaxScaler and StandardScaler are applied. Key hyperparameters such as learning rate, number of epochs, batch size, optimiser, and loss function are adjusted across all models. Then, the BPNN method requires tuning three exclusive hyperparameters: the number of hidden layers, activation functions, and initial weights. For the CNN method, four specific hyperparameters are optimized: input, output channel filter layers, filter kernel, and max-pooling window sizes. The LSTM and GRU methods involve five exclusive hyperparameters: input, hidden, output feature sizes, recurrent layers, and input sequence length.

3.5 Evaluating LIB and LiBat Systems SoC Estimation Models

The fourth focus of this project is to compare and evaluate the performance of the fine-tuned SoC estimation models, optimizing factors like input-output database scaling, hyperparameters, internal parameters, and mathematical assumptions for optimal estimation performance. This evaluation will be based on average metrics such as R^2 , MAE, and RMSE.

In order to assess and compare the estimation performance of conventional SoC estimation models, namely the EKF and UKF methods, benchmarking will be carried out. In a similar vein, the benchmarking procedure will encompass data-driven SoC estimation models, meticulously evaluating each ANN approach (BPNN, CNN, LSTM, and GRU) to ascertain the most

accurate and reliable performance. The objective of this comparison is to determine the most effective model within its respective category.

After evaluating each category of SoC estimation models, a comprehensive benchmarking process will rank them based on performance for the LIB and LiBat systems. Additionally, this benchmarking will determine the necessity of implementing and identify the best-performing data-driven models, particularly for LIB and LiBat systems in industrial AMR applications.

3.6 Final Year Project Gantt Chart

This section delineates the research planning and management for constructing a data-driven SoC estimation model tailored for LIB and LiBat systems in industrial AMR applications. The project plan encompasses major milestones to be achieved over two consecutive semesters of the FYP. Figures 3.3 and 3.4 will illustrate the five major milestones that will be undertaken during FYP Part 1 and the six major milestones that will be undertaken during Part 2.

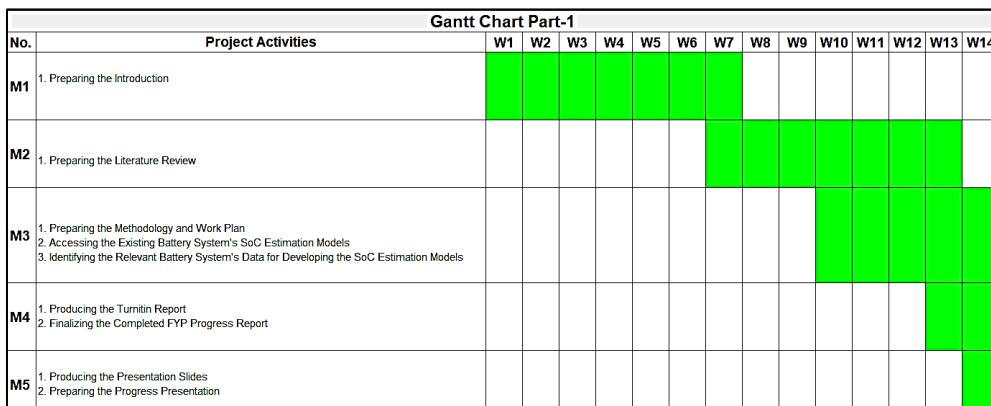


Figure 3.3: Gantt Chart of FYP Part 1.

First and foremost, FYP Part 1 consists of five major milestones: project initialisation, project research and study, project flow development, project report writing, and project presentation. Firstly, during the project initialisation phase, the introduction will be prepared by defining the problem statement, significance of the study, aims and objectives, and scope and limitations. This involves discussions with supervisor to understand the challenges and issues faced by the global industry in SoC estimation for LIB and LiBat systems used in industrial AMR applications. Then, in the project

research and study phase, a comprehensive literature review will be conducted, focussing on advancements in the technology of LIB and LiBat systems and the available conventional and data-driven SoC estimation models for LIB and LiBat systems in the context of industrial AMR applications.

Thirdly, during the project flow development phase, the input and output and conventional model parameter databases of the internal parameter datasets during charging and discharging for the LiB and LiBat systems will be collected from Mendeley Data, Borealis, the CALCE, and Recherche Data Gouv open-source repositories. Existing codes for the SoC estimation models for LIB and LiBat systems will be leveraged from the GitHub open-source repositories. Next, the methodology and work plan for the project, which will be used in the preparation of FYP Part 2, will be developed. Finally, in the project report writing and presentation phase, the project will be finalised by standardising the format and completing the necessary details of the progress report for submission. Simultaneously, presentation slides that provide a detailed overview of the progress report will be prepared for the forthcoming presentation session of FYP Part 1.

Gantt Chart Part-2															
No.	Project Activities	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14
M1	1. Preparing the Relevant Battery System's Data for SoC Estimation Model Development 2. Preparing the Existing Conventional Battery System's SoC Estimation Models 3. Fine-tuning the Existing Conventional Battery System's SoC Estimation Models 4. Testing and Validating the Performance of the Existing Conventional Battery System's SoC Estimation Models with the Collected Battery System's Data														
M2	1. Preparing the Existing Data-Driven Battery System's SoC Estimation Models 2. Improving the Representation of the Collected Battery System's Data 3. Fine-tuning the Existing Data-Driven Battery System's SoC Estimation Models 4. Testing and Validating the Performance of the Existing Data-Driven Battery System's SoC Estimation Models with the Collected Battery System's Data														
M3	1. Comparing and Evaluating the Performance of the Existing Battery System's SoC Estimation Models														
M4	1. Finalizing the Completed FYP Final Report 2. Preparing the FYP Documents 3. Producing the FYP Poster Design 4. Preparing the Submission of the FYP Poster Design														
M5	1. Producing the FYP Presentation Slides 2. Preparing the FYP Final Presentation														
M6	1. Preparing the Submission of the FYP Final Report and Documents														

Figure 3.4: Gantt Chart of FYP Part 2.

Afterward, FYP Part 2 consists of six major milestones. First, during the project preparation phase, the internal parameter datasets for LIB and LiBat systems, including Li-FP, Li-NMC, Li-NCA, Li-PO, and Li-Ion graphite systems, will be processed using coulomb counting and lookup table methods to produce the experimental SoC, following the methodologies outlined by the respective authors. Next, the existing codes for conventional SoC estimation models, gathered during FYP Part 1 from the GitHub open-source repositories,

will be organised and fine-tuned to optimise their performance, which will be evaluated using the processed input-output and conventional model parameter databases.

Second, the existing codes for data-driven SoC estimation models, available in the PyTorch library, will be prepared and organized. The input-output databases will be scaled using the MinMaxScaler and StandardScaler methods. The data-driven models will be fine-tuned for hyperparameters to enhance performance, followed by validation and testing with the scaled input-output databases. Third, a comparison of the estimation performance of each SoC estimation model will be conducted to identify the best-performing data-driven models and assess the need for their implementation. Fourth, the project will proceed with preparing the final reports, posters, and key FYP documents, followed by the submission of the posters. Fifth, the presentation slides will be prepared for the FYP presentation. Finally, the project will conclude with the submission of the final reports and all essential FYP documents, marking the completion of the project's milestones.

3.7 Summary

To summarize, the project methodology and work plan outline a multi-phase approach to SoC estimation for various LIB and LiBat systems, including Li-FP, Li-NMC, Li-NCA, Li-PO, and Li-Ion graphite batteries. The initial phase focuses on identifying and utilizing input-output and conventional model parameter databases that capture internal parameters during charging and discharging cycles. These databases, sourced from repositories like Mendeley Data, Borealis, CALCE, and Recherche Data Gouv, include datasets for batteries such as the LG 18650HG2, Samsung IN21700-30T, SB LiMotive, Turnigy Graphene, and A123. The datasets will be processed to establish baseline SoC values using coulomb counting and lookup tables, following the methodologies outlined by their respective authors.

The second phase focuses on developing optimal conventional SoC estimation models using existing GitHub open-source codes, including model-based EKF and UKF methods integrated with the ECM based on approaches from Larchuto (2021) and Alter, Gu, and Cong (2022). Concurrently, the optimal data-driven SoC estimation models such as BPNN, CNN, LSTM, and

GRU methods will be developed using the PyTorch library. The third phase aims to optimize these SoC estimation models by adjusting key parameters, mathematical assumptions, dataset scaling, and hyperparameters. Performance will be assessed using metrics like R^2 , MAE, and RMSE, with targets of R^2 above 0.8, MAE below 10, and RMSE below 15 to ensure reliable benchmarking between the SoC estimation models.

The final phase focuses on a comparative evaluation of the optimized SoC estimation models. Performance benchmarking will identify the most effective models within each category, covering both conventional and data-driven approaches. This comprehensive assessment will determine the top-performing models for the LIB and LiBat systems and evaluate the potential necessity of data-driven methods in industrial AMR applications.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

This section presents a comprehensive internal and overall evaluation of the SoC estimation performance, using metrics such as MAE, RMSE, and R², for two pre-modeled conventional methods (EKF and UKF) and four pre-trained data-driven models (BPNN, LSTM, GRU, and CNN). The evaluations are based on five datasets, each representing different batteries from the LIB and LiBat systems, tested under various operating temperatures, states, and rates. A discussion is provided on the parameters and hyperparameters used to optimize both the conventional and data-driven models. Visual comparisons of the estimated SoC and SoC error values for each model throughout the cycle of the LIB and LiBat systems are illustrated to support further analysis. Additionally, the challenges and limitations encountered are outlined to highlight the scope and difficulties faced during the FYP study.

4.2 Model Optimization

4.2.1 Conventional Optimization

Tables 4.1 to 4.5 outline the ten key parameter settings employed in the EKF SoC estimation method, which yield optimized performance metrics for the Li-NMC, Li-Ion, Li-NCA, Li-PO, and Li-FP graphite battery systems, respectively. Correspondingly, Tables 4.6 to 4.10 provide a detailed overview of the fourteen parameter settings utilized in the UKF SoC estimation method across comparable LIB and LiBat systems.

Table 4.1: EKF Parameter Settings for Li-NMC Graphite Battery System.

	40°C	25°C	10°C	0°C	-10°C	-20°C
Dynamic Resistance, Ω				1×10^{-4}		
Capacitance, F				1×10^5		
State Error				2.25×10^{-8}		

Measurement Noise	2.25×10^{-8}					
Process Noise	2.25×10^{-13}					
Maximum Capacity, Ah	3.0					
Initial SoC	0.5					
Polynomial Degree	20					
Coulomb Efficiency	1.3	1.7	2.3	3.2	5.0	8.5
Internal Resistance, Ω	0.001	0.002	0.004	0.008	0.016	0.032

Table 4.2: EKF Parameter Settings for Li-Ion Graphite Battery System.

	40°C	25°C	10°C	0°C	-10°C	-20°C
Dynamic Resistance, Ω	1×10^{-3}					
Capacitance, F	1×10^5					
State Error	2.25×10^{-4}					
Measurement Noise	2.25×10^{-4}					
Process Noise	2.25×10^{-7}					
Maximum Capacity, Ah	3.0					
Initial SoC	0.5					
Polynomial Degree	20					
Coulomb Efficiency	1.5	3.0	4.0	7.0	10.0	14.0
Internal Resistance, Ω	0.001	0.002	0.004	0.008	0.016	0.032

Table 4.3: EKF Parameter Settings for Li-NCA Graphite Battery System.

	40°C	25°C	10°C	0°C	-10°C	-20°C
Dynamic Resistance, Ω				1×10^{-3}		
Capacitance, F				1×10^5		
State Error				2.25×10^{-4}		
Measurement Noise				2.25×10^{-4}		
Process Noise				2.25×10^{-7}		
Maximum Capacity, Ah				5.2		
Initial SoC				0.5		
Polynomial Degree				20		
Coulomb Efficiency	1.5	2.0	3.0	5.0	9.0	15.0
Internal Resistance, Ω	0.001	0.002	0.004	0.008	0.016	0.032

Table 4.4: EKF Parameter Settings for Li-PO Graphite Battery System.

	40°C	25°C	10°C	0°C	-10°C	-20°C
Dynamic Resistance, Ω				1×10^{-3}		
Capacitance, F				1×10^5		
State Error				2.25×10^{-4}		
Measurement Noise				2.25×10^{-4}		
Process Noise				2.25×10^{-7}		
Maximum Capacity, Ah				5.0		
Initial SoC				0.5		
Polynomial Degree				20		

Coulomb Efficiency	2.0	2.5	5.0	8.0	19.0	9.0
Internal Resistance, Ω	0.001	0.002	0.004	0.008	0.016	0.032

Table 4.5: EKF Parameter Settings for Li-FP Graphite Battery System.

	40°C	35~30°C	25~20°C	10°C	5~0°C
Dynamic Resistance, Ω			1×10^{-4}		
Capacitance, F			1×10^5		
State Error			2.25×10^{-10}		
Measurement Noise			2.25×10^{-10}		
Process Noise			2.25×10^{-15}		
Maximum Capacity, Ah			$1.1 / 2.0 / 4.4$		
Initial SoC			0.5		
Polynomial Degree			1		
Coulomb Efficiency			0		
Internal Resistance, Ω	0.001	0.0015	0.002	0.004	0.004

Table 4.6: UKF Parameter Settings for Li-NMC Graphite Battery System.

	40°C	25°C	10°C	0°C	-10°C	-20°C
Dynamic Resistance, Ω			1×10^{-4}			
Capacitance, F			1×10^5			
State Error			1×10^{-7}			
Measurement Noise			1×10^{-6}			
Process Noise			4×10^{-9}			

Dimension	1					
Alpha	0.04					
Beta	2					
Lambda	1.5					
Maximum Capacity, Ah	3.0					
Initial SoC	0.5					
Polynomial Degree	3					
Coulomb Efficiency	0					
Internal Resistance, Ω	-0.03	-0.03	-0.05	-0.07	-0.1	-0.2

Table 4.7: UKF Parameter Settings for Li-Ion Graphite Battery System.

	40°C	25°C	10°C	0°C	-10°C	-20°C
Dynamic Resistance, Ω				1×10^{-3}		
Capacitance, F				1×10^5		
State Error				1×10^{-7}		
Measurement Noise				1×10^{-6}		
Process Noise				4×10^{-9}		
Dimension	1					
Alpha	0.04					
Beta	2					
Lambda	1.5					
Maximum Capacity, Ah	3.0					
Initial SoC	0.5					
Polynomial Degree	3					

Coulomb Efficiency	0					
Internal Resistance, Ω	-0.02	-0.02	-0.02	-0.03	-0.04	-0.06

Table 4.8: UKF Parameter Settings for Li-NCA Graphite Battery System.

	40°C	25°C	10°C	0°C	-10°C	-20°C
Dynamic Resistance, Ω	1×10^{-3}					
Capacitance, F	1×10^5					
State Error	1×10^{-7}					
Measurement Noise	1×10^{-6}					
Process Noise	4×10^{-9}					
Dimension	1					
Alpha	0.04					
Beta	2					
Lambda	1.5					
Maximum Capacity, Ah	5.2					
Initial SoC	0.5					
Polynomial Degree	3					
Coulomb Efficiency	0					
Internal Resistance, Ω	-0.03	-0.04	-0.06	-0.08	-0.1	-0.2

Table 4.9: UKF Parameter Settings for Li-PO Graphite Battery System.

	40°C	25°C	10°C	0°C	-10°C	-20°C
Dynamic Resistance, Ω	1×10^{-3}					
Capacitance, F	1×10^5					

State Error	1×10^{-7}					
Measurement Noise	1×10^{-6}					
Process Noise	4×10^{-9}					
Dimension	1					
Alpha	0.04					
Beta	2					
Lambda	1.5					
Maximum Capacity, Ah	5.0					
Initial SoC	0.5					
Polynomial Degree	3					
Coulomb Efficiency	0					
Internal Resistance, Ω	-0.01	-0.01	-0.02	-0.02	-0.05	-0.03

Table 4.10: UKF Parameter Settings for Li-FP Graphite Battery System.

	40°C	35~30°C	25~20°C	10°C	5~0°C
Dynamic Resistance, Ω	1×10^{-4}				
Capacitance, F	1×10^5				
State Error	1×10^{-7}				
Measurement Noise	1×10^{-6}				
Process Noise	4×10^{-9}				
Dimension	1				
Alpha	0.04				
Beta	2				
Lambda	1.5				
Maximum Capacity, Ah	1.1 / 2.0 / 4.4				

Initial SoC	0.5				
Polynomial Degree	1				
Coulomb Efficiency	9	11	11	14	25
Internal Resistance, Ω	0.01	0.015	0.02	0.03	0.08

The optimized EKF and UKF SoC estimation models are developed through analytical and empirical optimization of ten and fourteen key parameters, respectively. However, because of the variations in battery specifications, as well as differences in the internal and external physical and chemical structures within LIB and LiBat systems, the models are divided into five distinct parts, each tailored to represent the unique characteristics of these systems. Additionally, the EKF and UKF SoC estimation models rely on different mathematical approaches, leading to distinct sets of optimized parameter values for each model.

The first part of the EKF SoC estimation model is designed for the LG 18650HG2 3000mAh 3.6V Li-NMC graphite battery system. Table 4.1 shows the seven key parameters are set as follows: dynamic resistance at $1 \times 10^{-4} \Omega$, capacitance at $1 \times 10^5 \text{ F}$, state error and measurement noise covariance at 2.25×10^{-8} , process noise covariance at 2.25×10^{-13} , initial SoC at 0.5, and the OCV-SoC polynomial relationship calibrated at 20. The coulomb efficiency and internal resistance are adjusted based on temperature: $1 \times 10^{-3} \Omega$ and 1.3 at 40°C , $2 \times 10^{-3} \Omega$ and 1.7 at 25°C , $4 \times 10^{-3} \Omega$ and 2.3 at 10°C , $8 \times 10^{-3} \Omega$ and 3.2 at 0°C , $16 \times 10^{-3} \Omega$ and 5 at -10°C , and $32 \times 10^{-3} \Omega$ and 8.5 at -20°C .

The second part of the EKF SoC estimation model is designed for the Samsung IN21700-30T 3000mAh 3.6V Li-Ion graphite battery system. Table 4.2 shows seven key parameters are set as follows: dynamic resistance at $1 \times 10^{-3} \Omega$, capacitance at $1 \times 10^5 \text{ F}$, state error and measurement noise covariance at 2.25×10^{-4} , process noise covariance at 2.25×10^{-7} , initial SoC at 0.5, and the OCV-SoC polynomial relationship calibrated at 20 degrees. The coulomb efficiency and internal resistance vary with temperature: $1 \times 10^{-3} \Omega$ and 1.5 at

40°C , $2 \times 10^{-3} \Omega$ and 3 at 25°C , $4 \times 10^{-3} \Omega$ and 4 at 10°C , $8 \times 10^{-3} \Omega$ and 7 at 0°C , $16 \times 10^{-3} \Omega$ and 10 at -10°C , and $32 \times 10^{-3} \Omega$ and 14 at -20°C .

The third part of the EKF SoC estimation model is designed for the SB LiMotive 5200mAh 3.7V Li-NCA graphite battery system, maintaining the same key parameters and assumptions as the second part. However, coulomb efficiency varies with temperature: 1.5 at 40°C , 2 at 25°C , 3 at 10°C , 5 at 0°C , 9 at -10°C , and 15 at -20°C , as shown in Table 4.3. Similarly, the fourth part, focused on the Turnigy Graphene 5000mAh 3.7V Li-PO graphite battery systems, follows the same parameters but with adjusted coulomb efficiency: 2 at 40°C , 2.5 at 25°C , 5 at 10°C , 8 at 0°C , 19 at -10°C , and 9 at -20°C , as shown in Table 4.4. Finally, the fifth part addresses the A123 1100, 2000, and 4400mAh 3.2V Li-FP graphite battery systems, maintaining consistency in key parameters and assumptions as in the first part, with the OCV-SoC polynomial degree fixed at 1 and a constant coulomb efficiency of 0 across all conditions as shown in Table 4.5.

The first part of the UKF SoC estimation model is designed for the LG 18650HG2 3000mAh 3.6V Li-NMC graphite battery system. As outlined in Table 4.6, twelve key parameters are set as follows: dynamic resistance at $1 \times 10^{-4} \Omega$, capacitance at $1 \times 10^5 \text{ F}$, state error at 1×10^{-7} , measurement noise at 1×10^{-6} , and process noise covariances at 4×10^{-9} , initial SoC at 0.5, dimension at 1, alpha at 0.04, beta at 2, lambda at 1.5, OCV-SoC polynomial relationship calibrated to degree 5, and a coulomb efficiency of 0. The internal resistance is adjusted based on temperature: -0.03Ω at 40°C , -0.03Ω at 25°C , -0.05Ω at 10°C , -0.07Ω at 0°C , -0.1Ω at -10°C , and -0.2Ω at -20°C .

Similarly, the second, third, and fourth parts of the UKF SoC estimation model, designed for the Samsung IN21700-30T 3000mAh 3.6V Li-Ion, SB LiMotive 5200mAh 3.7V Li-NCA, and Turnigy Graphene 5000mAh 3.7V Li-PO graphite battery systems, maintain the same key parameters and assumptions as the first part. Nevertheless, the internal resistance varies with temperature for each system: for the Li-Ion, it is -0.02Ω at 40°C , -0.02Ω at 25°C , -0.02Ω at 10°C , -0.03Ω at 0°C , -0.04Ω at -10°C , and -0.06Ω at -20°C ; for the Li-NCA, it is -0.03Ω at 40°C , -0.04Ω at 25°C , -0.06Ω at 10°C , -0.08Ω at 0°C , -0.10Ω at -10°C , and -0.20Ω at -20°C ; and for the Li-PO, it is -0.01

Ω at 40°C, -0.01 Ω at 25°C, -0.02 Ω at 10°C, -0.02 Ω at 0°C, -0.05 Ω at -10°C, and -0.03 Ω at -20°C as outlined in Tables 4.7 to 4.9.

Lastly, the fifth part of the UKF SoC estimation model, designed for the A123 1100, 2000, and 4400mAh 3.2V Li-FP graphite battery systems, retains the same key parameters and assumptions as the previous parts, with the exception of the OCV-SoC polynomial relationship, which is calibrated to degree 1. As detailed in Table 4.10, the coulomb efficiency and internal resistance vary with temperature as follows: 9 and 0.01 Ω at 40°C, 11 and 0.015 Ω at 35~30°C, 11 and 0.02 Ω at 25~20°C, 14 and 0.03 Ω at 10°C, and 25 and 0.08 Ω at 5~0°C.

4.2.2 Data-Driven Optimization

Tables 4.11 to 4.14 provide a comprehensive overview of the hyperparameter settings used in the data-driven SoC estimation methods, including BPNN, LSTM, GRU, and CNN models. These configurations have been fine-tuned to achieve optimized performance metrics for the LIB and LiBat systems.

Table 4.11: BPNN Hyperparameter Settings for LIB and LiBat Systems.

Data Scaling Method	MinMaxScaler
Batch Size	96
Activation Function	ReLU
Hidden Layer	7
Number of Neuron	256 (4) \rightarrow 512 (10)
Input Layer	3
Output Layer	1
Epoch	30
Loss Function	Mean Square Error (MSE)
Optimizer	Adam
Learning Rate	0.001

Table 4.12: LSTM Hyperparameter Settings for LIB and LiBat Systems.

Data Scaling Method	MinMaxScaler
Batch Size	96

Activation Function	ReLU
Hidden Layer	7
Number of Neuron	32 (2) → 64 (2) → 128 (2) → 256 (4) → 512 (4)
Input Feature Size	3
Hidden Feature Size	10
Output Feature Size	1
Recurrent Layer	1
Epoch	30
Loss Function	Mean Square Error (MSE)
Optimizer	Adam
Learning Rate	0.001

Table 4.13: GRU Hyperparameter Settings for LIB and LiBat Systems.

Data Scaling Method	MinMaxScaler
Batch Size	96
Activation Function	ReLU
Hidden Layer	7
Number of Neuron	32 (2) → 64 (2) → 128 (2) → 256 (4) → 512 (4)
Input Feature Size	3
Hidden Feature Size	10
Output Feature Size	1
Recurrent Layer	1
Epoch	30
Loss Function	Mean Square Error (MSE)
Optimizer	Adam
Learning Rate	0.001

Table 4.14: CNN Hyperparameter Settings for LIB and LiBat Systems.

Data Scaling Method	MinMaxScaler
Batch Size	96
Activation Function	ReLU

Hidden Layer	8
Number of Neuron	$10 \rightarrow 128(4) \rightarrow 256(4) \rightarrow 512(8)$
Input Channel Filter Layer	3
Filter Kernel Size	1
Output Channel Filter Layer	10
Max-Pooling Window Size	1
Epoch	30
Loss Function	Mean Square Error (MSE)
Optimizer	Adam
Learning Rate	0.001

Similar to the EKF and UKF SoC estimation models discussed in Section 4.2.1, the data-driven SoC estimation models undergo a rigorous analytical and empirical hyperparameter optimization process. Key hyperparameters such as data scaling method, learning rate, number of epochs, batch size, optimizer, loss function, and activation function are fine-tuned using well-established choices. Specifically, MaxMinScaler is used for scaling, the learning rate is set at 0.001, with 30 epochs, a batch size of 96, the Adam optimizer, MSE as the loss function, and ReLU as the activation function.

Moreover, for the BPNN method, four exclusive hyperparameters are optimized: 7 hidden layers, 3 input layers, 1 output layer, and a neuron configuration comprising two layers with 256 neurons and five layers with 512 neurons. The LSTM and GRU methods share six exclusive hyperparameters, including 7 hidden layers, 3 input features, 1 output feature, 10 hidden features, 1 recurrent layer, and neuron configurations with one layer of 32 neurons, one layer of 64, one layer of 128, two layers of 256, and two layers of 512 neurons. Lastly, for the CNN method, four exclusive hyperparameters are optimized as follows: 3 input channel filters, 10 output channel filters, a kernel size of 1, and a max-pooling window size of 1.

4.3 Model SoC Estimation Results

4.3.1 Conventional SoC Estimation Results

Tables 4.15 to 4.19 present the SoC estimation results using the EKF method, detailing performance across various datasets for different LIB and LiBat systems at varying operating temperatures. These datasets include Li-FP, Li-NMC, Li-NCA, Li-PO, and Li-Ion graphite battery systems, with performance evaluated using metrics such as MAE, RMSE, and R² values. Table 4.20 provides a comprehensive summary of the EKF method's effectiveness across all datasets, offering an overall assessment of its SoC estimation capabilities for these LIB and LiBat systems.

Table 4.15: EKF Average SoC Estimation Results for the LG 18650HG2 3000mAh 3.6V Li-NMC Graphite Battery.

	40°C	25°C	10°C	0°C	-10°C	-20°C	Average
MAE	3.0476	3.5105	4.0850	5.2216	7.0799	9.6169	<u>5.4269</u>
RMSE	5.9107	6.4150	6.4441	7.5349	9.1164	11.8678	<u>7.8815</u>
R²	0.9519	0.9464	0.9346	0.9068	0.8404	0.5544	<u>0.8558</u>

Table 4.16: EKF Average SoC Estimation Results for the Samsung IN21700-30T 3000mAh 3.6V Li-Ion Graphite Battery.

	40°C	25°C	10°C	0°C	-10°C	-20°C	Average
MAE	2.3704	2.0783	3.1938	5.2651	6.7616	8.5677	<u>4.7062</u>
RMSE	4.0027	3.6954	4.5535	6.7344	9.4746	11.8968	<u>6.7262</u>
R²	0.9850	0.9862	0.9778	0.9482	0.8900	0.8163	<u>0.9339</u>

Table 4.17: EKF Average SoC Estimation Results for the SB LiMotive 5200mAh 3.7V Li-NCA Graphite Battery.

	40°C	25°C	10°C	0°C	-10°C	-20°C	Average
MAE	2.4948	2.7147	4.4267	5.1977	6.6735	8.6113	<u>5.0198</u>
RMSE	4.1248	4.3096	5.8234	7.0946	9.1984	12.0920	<u>7.1071</u>
R²	0.9850	0.9831	0.9652	0.9477	0.8872	0.7687	<u>0.9228</u>

Table 4.18: EKF Average SoC Estimation Results for the Turnigy Graphene 5000mAh 3.7V Li-PO Graphite Battery.

	40°C	25°C	10°C	0°C	-10°C	-20°C	Average
MAE	1.8558	1.9378	2.5258	4.4794	7.2477	14.6749	<u>5.4536</u>
RMSE	3.0296	3.2343	3.9683	6.2631	9.8570	18.7000	<u>7.5087</u>
R²	0.9877	0.9855	0.9778	0.9394	0.8192	0.2086	<u>0.8197</u>

Table 4.19: EKF Average SoC Estimation Results for the A123 1100, 2000, and 4400mAh 3.2V Li-FP Graphite Batteries.

	40°C	35~30°C	25~20°C	10°C	5~0°C	Average
MAE	11.6892	7.0826	9.5513	16.0106	10.5888	<u>10.9845</u>
RMSE	14.2137	11.4158	13.6192	20.1130	17.4568	<u>15.3637</u>
R²	0.7655	0.9120	0.8459	0.5417	0.7927	<u>0.7716</u>

Table 4.20: Comprehensive Summary of EKF Overall SoC Estimation Results for LIB and LiBat Systems.

	Li-NMC	Li-Ion	Li-NCA	Li-PO	LI-FP	Overall
MAE	5.4269	4.7062	5.0198	5.4536	10.9845	<u>6.3182</u>
RMSE	7.8815	6.7262	7.1071	7.5087	15.3637	<u>8.9174</u>
R²	0.8558	0.9339	0.9228	0.8197	0.7716	<u>0.8608</u>

Similarly, Tables 4.21 to 4.25 present the SoC estimation results using the UKF method across various LIB and LiBat systems at different temperatures. The datasets include five different graphite battery systems, with performance assessed using metrics like MAE, RMSE, and R². Table 4.26 provides a comprehensive summary of the overall effectiveness of the UKF method in estimating SoC across the LIB and LiBat systems.

Table 4.21: Average SoC Estimation Results Using the UKF Method for the LG 18650HG2 3000mAh 3.6V Li-NMC Graphite Battery.

	40°C	25°C	10°C	0°C	-10°C	-20°C	Average
MAE	3.5453	4.4391	5.2803	6.9032	10.0065	13.5846	<u>7.2932</u>
RMSE	4.3684	5.2893	6.5541	8.5008	11.8928	16.1747	<u>8.7967</u>

R²	0.9738	0.9636	0.9324	0.8814	0.7284	0.1722	<u>0.7753</u>
----------------------	--------	--------	--------	--------	--------	--------	---------------

Table 4.22: Average SoC Estimation Results Using the UKF Method for the Samsung IN21700-30T 3000mAh 3.6V Li-Ion Graphite Battery.

	40°C	25°C	10°C	0°C	-10°C	-20°C	Average
MAE	4.8279	4.5031	4.1286	6.2897	8.0175	10.1622	<u>6.3215</u>
RMSE	6.9308	6.0822	5.2921	8.3210	10.6265	13.9323	<u>8.5308</u>
R²	0.9551	0.9627	0.9700	0.9209	0.8616	0.7481	<u>0.9031</u>

Table 4.23: Average SoC Estimation Results Using the UKF Method for the SB LiMotive 5200mAh 3.7V Li-NCA Graphite Battery.

	40°C	25°C	10°C	0°C	-10°C	-20°C	Average
MAE	3.3998	3.6295	4.6995	5.1831	7.2819	7.6709	<u>5.3108</u>
RMSE	4.9048	5.1443	6.4269	6.9868	9.1330	9.8399	<u>7.0726</u>
R²	0.9787	0.9759	0.9576	0.9493	0.8888	0.8469	<u>0.9329</u>

Table 4.24: Average SoC Estimation Results Using the UKF Method for the Turnigy Graphene 5000mAh 3.7V Li-PO Graphite Battery.

	40°C	25°C	10°C	0°C	-10°C	-20°C	Average
MAE	4.2128	4.1504	4.7127	6.1457	9.2405	13.9479	<u>7.0678</u>
RMSE	5.5377	5.5654	6.6359	8.4837	12.3211	18.5400	<u>9.5140</u>
R²	0.9586	0.9570	0.9379	0.8889	0.7176	0.2221	<u>0.7804</u>

Table 4.25: Average SoC Estimation Results Using the UKF Method for the A123 1100, 2000, and 4400mAh 3.2V Li-FP Graphite Batteries.

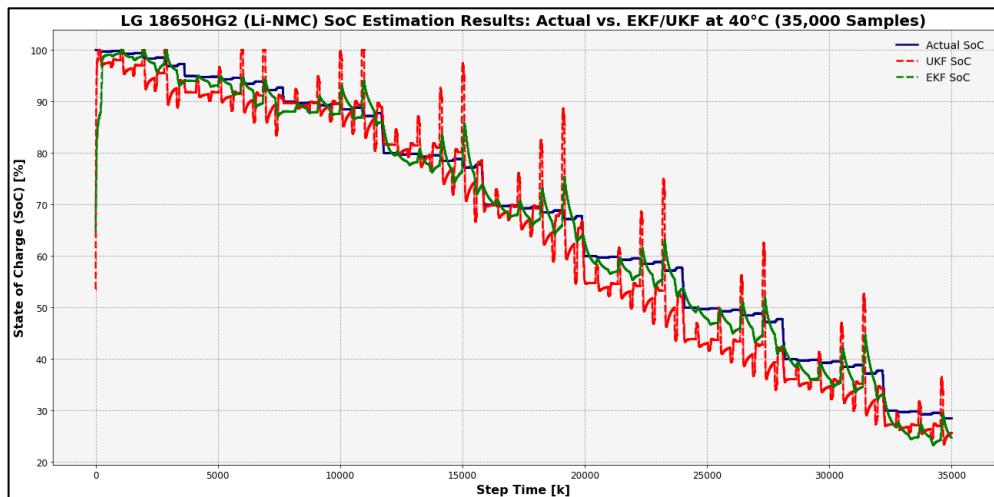
	40°C	35~30°C	25~20°C	10°C	5~0°C	Average
MAE	11.6114	8.1954	9.9940	15.4603	11.4082	<u>11.3339</u>
RMSE	14.2880	14.3978	15.1348	19.9140	18.9638	<u>16.5397</u>
R²	0.7631	0.8601	0.8096	0.5507	0.7554	<u>0.7478</u>

Table 4.26: Comprehensive Summary of Average and Overall SoC Estimation Results Using the UKF Method for LIB and LiBat Systems.

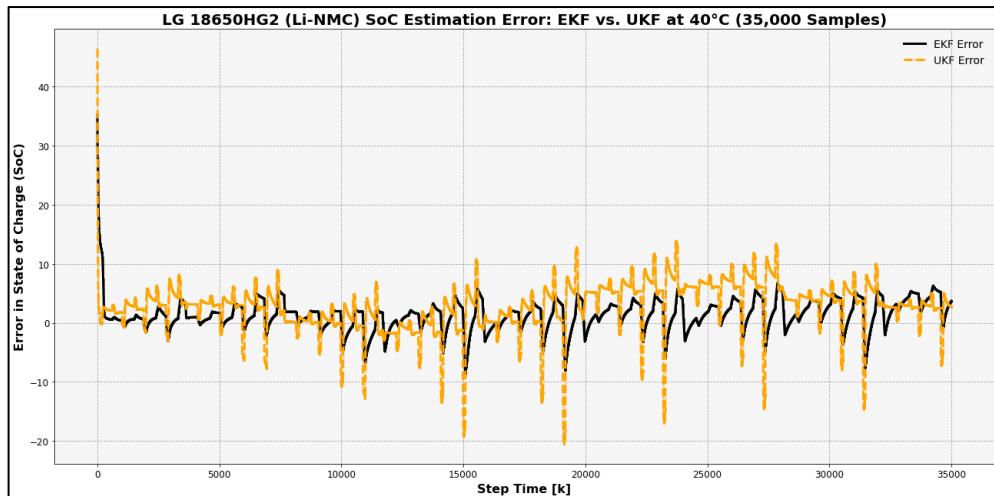
	Li-NMC	Li-Ion	Li-NCA	Li-PO	Li-FP	Overall
--	---------------	---------------	---------------	--------------	--------------	----------------

MAE	7.2932	6.3215	5.3108	7.0678	11.3339	<u>7.4654</u>
RMSE	8.7967	8.5308	7.0726	9.5140	16.5397	<u>10.0908</u>
R²	0.7753	0.9031	0.9329	0.7804	0.7478	<u>0.8279</u>

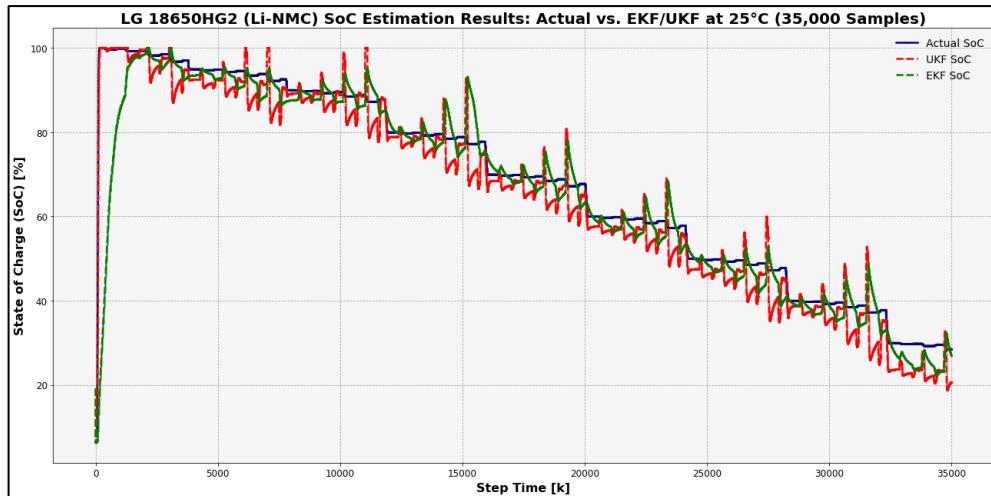
Figure 4.1 and Graphs A-5 to A-8 show the estimated SoC values from EKF and UKF methods alongside experimental SoC, comparing estimation errors across 35,000 samples for various LIB and LiBat systems (Li-FP, Li-NMC, Li-NCA, Li-PO, and Li-Ion) under different temperatures.



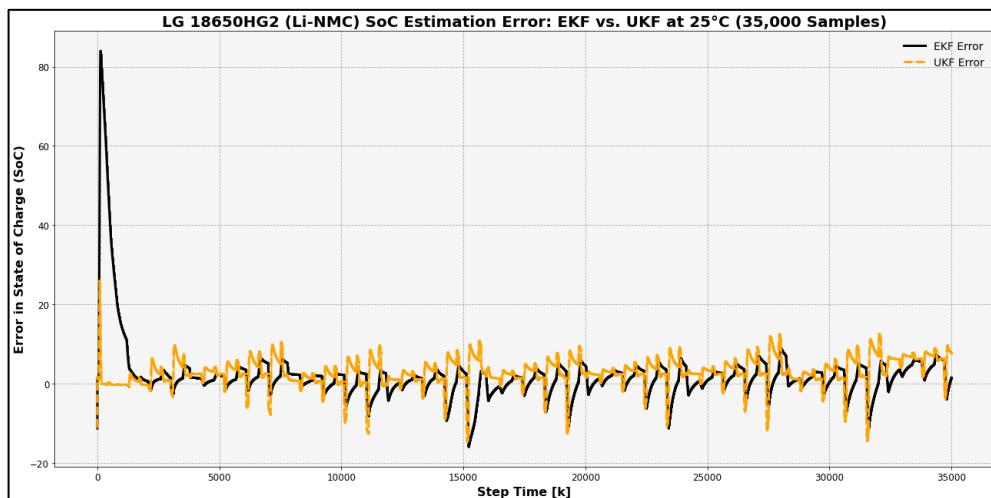
(a).



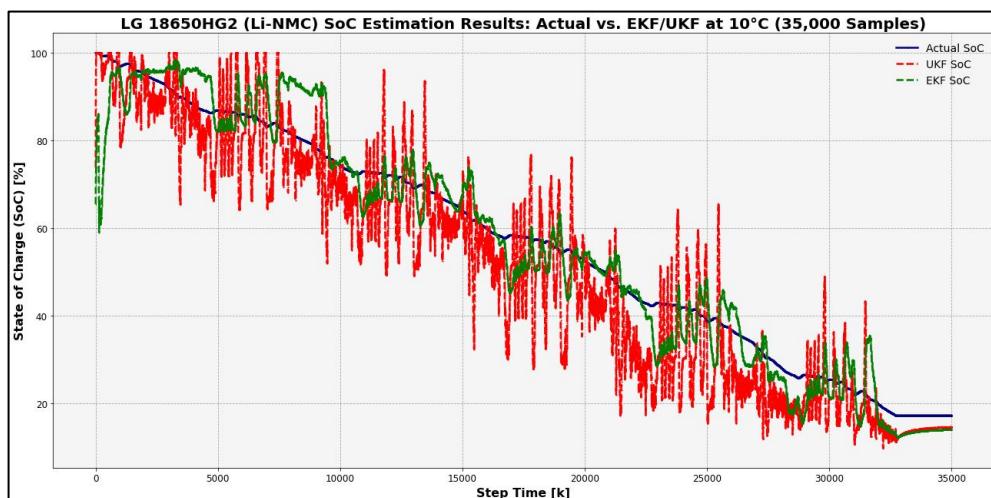
(b).



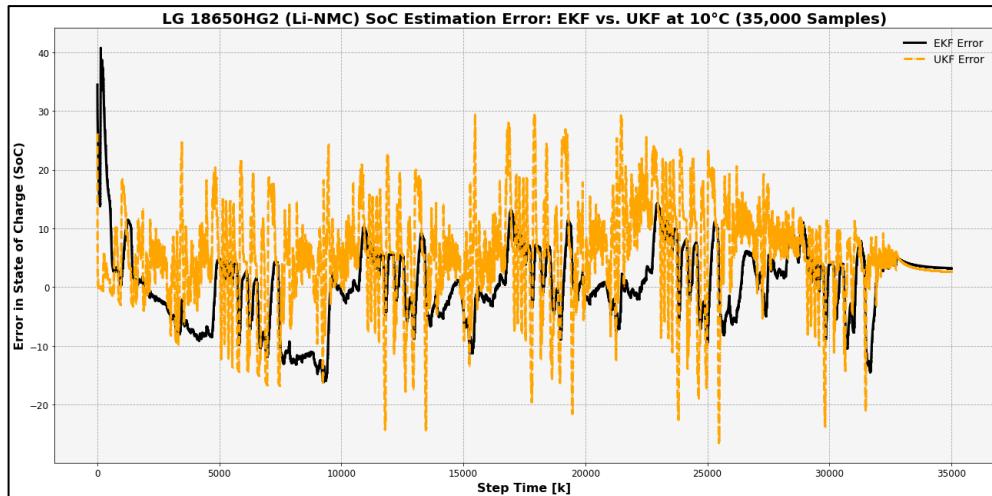
(c).



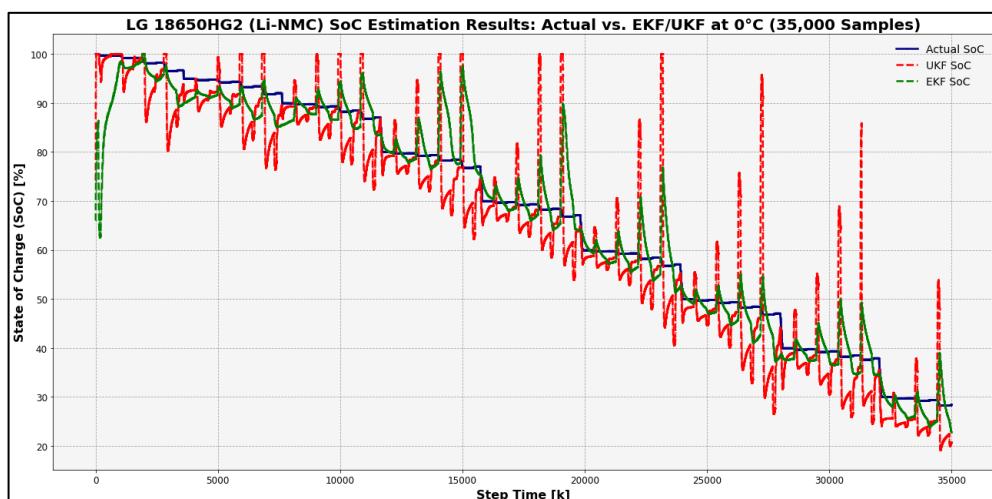
(d).



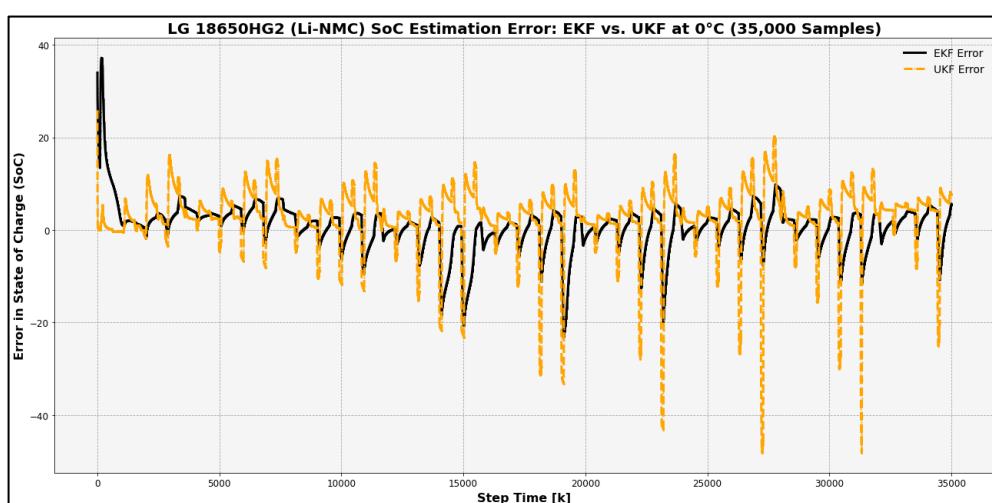
(e).



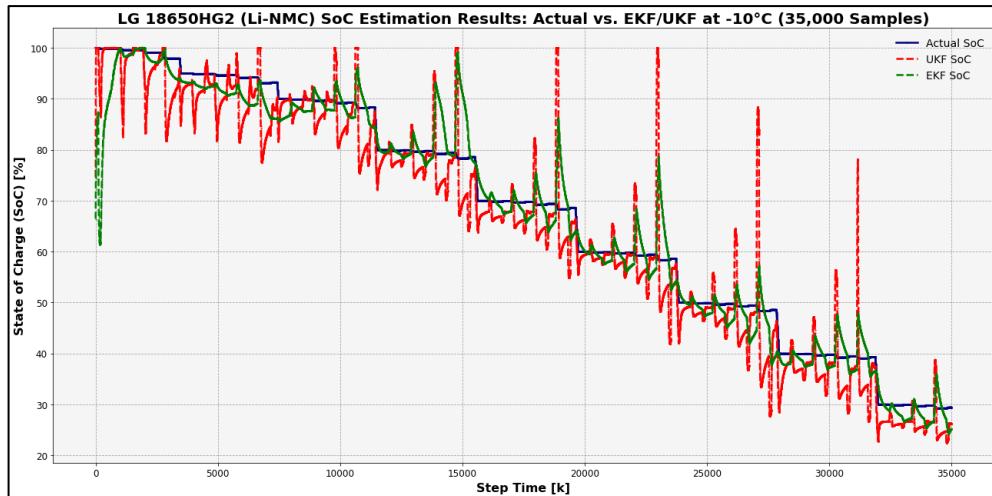
(f).



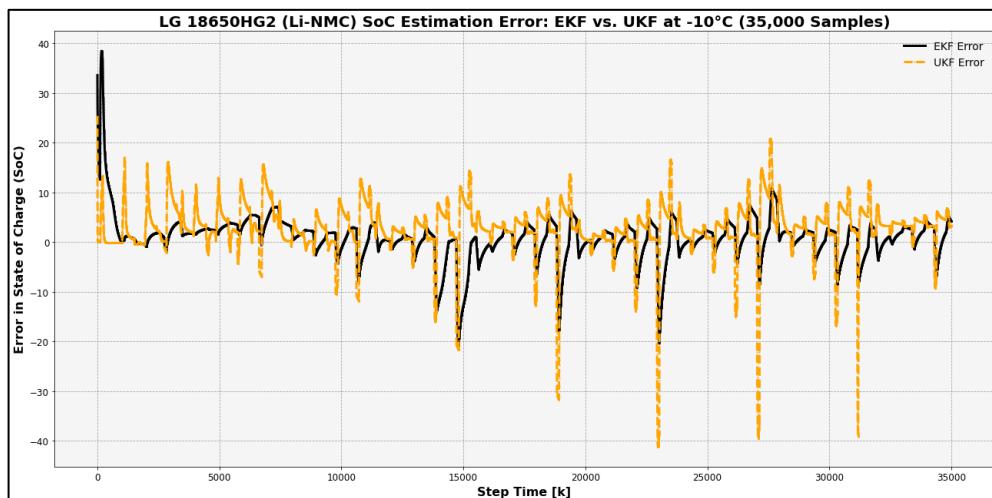
(g).



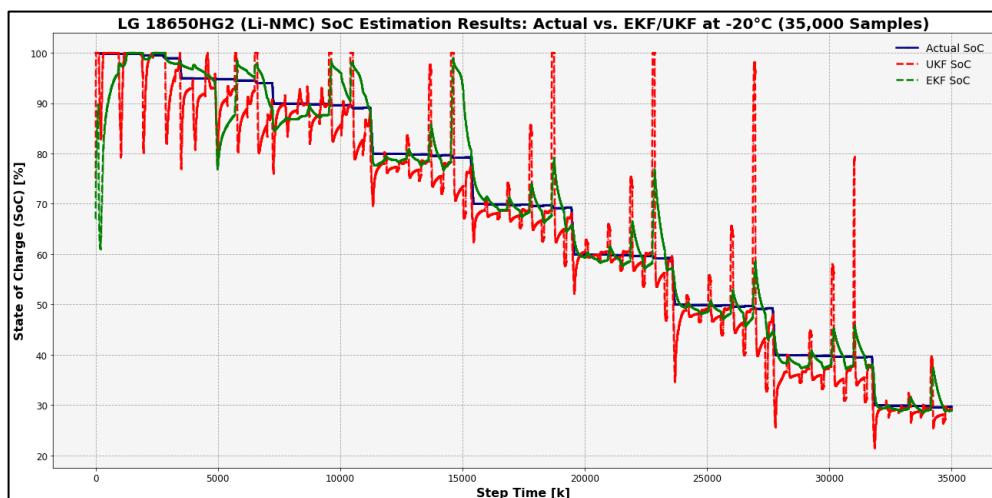
(h).



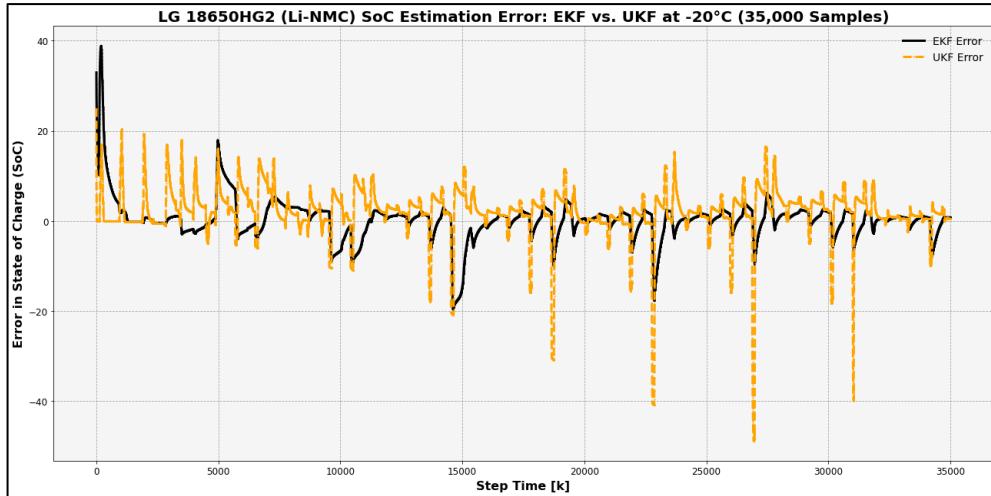
(i).



(j).



(k).



(l).

Figure 4.1: LG 18650HG2 (Li-NMC) Battery: Comparison of SoC Curves (Actual, EKF, and UKF) and Error Curves (EKF and UKF) at (a, b) 40°C, (c, d) 25°C, (e, f) 10°C, (g, h) 0°C, (i, j) -10°C, and (k, l) -20°C for the First 35,000 Samples.

Based on the results presented in Tables 4.15 to 4.19 and Tables 4.21 to 4.25, show that the optimized EKF and UKF SoC estimation models maintain stable performance across Li-NMC, Li-NCA, Li-PO, and Li-Ion graphite battery systems within the operational temperature range of 0°C to 40°C. Within this range, both methods display relatively consistent metrics for MAE, RMSE, and R². Specifically, the EKF method achieves MAE values between 1.8558 and 5.2651, RMSE values between 3.0296 and 7.5349, and R² values between 0.9068 and 0.9977. In comparison, the UKF method yields MAE values ranging from 3.3998 to 6.9032, RMSE from 4.3684 to 8.5008, and R² between 0.8814 and 0.9787.

However, as the temperature drops from 0°C to -20°C, a noticeable deterioration in performance of both methods emerges. For the EKF method, the error increases significantly, with MAE values rising from 4.4794 to 14.6749, RMSE from 6.2631 to 18.7, and R² sharply declining from 0.9482 to 0.2086. The UKF method follows a similar trend, with MAE values ranging from 5.1831 to 13.9479, RMSE from 6.9868 to 18.54, and R² decreasing from 0.9493 to 0.1722. These results highlight a clear relationship between operating temperatures and the accuracy of both EKF and UKF SoC estimation models.

across the evaluated battery systems. As temperatures decrease, the models struggle to capture the internal parameter dynamics during discharge and charge cycles, leading to an exponential decline in accuracy below 0°C. This trend demonstrates the significant challenges low-temperature conditions pose for reliable SoC estimation.

In contrast, the Li-FP graphite battery system, shown in Tables 4.19 and 4.25, presents more pronounced fluctuations even within the 0°C to 40°C range. For the EKF method, R^2 drops from 0.9120 to 0.5417, RMSE increases from 11.4158 to 20.1130, and MAE rises from 7.0826 to 16.0106. Similarly, the UKF method exhibits R^2 decreasing from 0.8601 to 0.5507, RMSE rising from 14.288 to 19.914, and MAE increasing from 8.1954 to 15.4603. Unlike other systems, the Li-FP system does not follow a linear decline in accuracy. Instead, it shows unpredictable behavior with marked fluctuations across temperatures, particularly in the sequence of 35–30°C, 25–20°C, 5–0°C, 40°C, and 10°C. This indicates considerable instability in capturing the internal dynamics of this system under varying temperature conditions, posing additional challenges for SoC estimation.

Additionally, Tables 4.20 and 4.26 outline the average optimized performance of the EKF and UKF models across LIB and LiBat systems. The EKF method shows R^2 values ranging from 0.7716 to 0.9339, RMSE from 6.7262 to 15.3637, and MAE from 4.7062 to 10.9845. Similarly, the UKF method presents R^2 values between 0.7478 and 0.9329, MAE from 5.3108 to 11.3339, and RMSE from 7.0726 to 16.5397. A clear performance hierarchy emerges across battery systems, with Li-Ion at the top, followed by Li-NCA, Li-NMC, Li-PO, and finally Li-FP graphite systems. This linear decline in performance underscores the model's superior performance with Li-Ion graphite systems and its least effective results with Li-FP systems.

In conclusion, while the EKF method demonstrates a generally robust performance, the UKF method shows greater performance deficiencies, particularly at lower temperatures. The EKF method achieves an MAE of 6.3182, an RMSE of 8.9174, and an R^2 of 0.8608, surpassing the baseline by 36.82% for MAE, 40.55% for RMSE, and 7.6% for R^2 , with an average surpassion of 28.32%. In contrast, the UKF method only outperforms the baseline by 25.346% for MAE at 7.4654, 32.728% for RMSE at 10.0908, and

3.4875% for R^2 at 0.8279, with an average surpassion of 20.5205%. These findings demonstrate that the EKF method outperforms the UKF method, contrasting with theoretical predictions in Sections 2.3.3.5 and 2.3.3.6.

Figure 4.1 and Graphs A-5 to A-8, illustrate that at the commencement of the discharge cycle, both EKF and UKF methods initialize the SoC at 50%. As the cycle progresses, the estimated SoC values exhibit gradual deviations from this initial value but generally converge towards the actual SoC values over time. This observation underscores the effectiveness of both EKF and UKF methods in integrating prediction steps, similar to the look-up method, and updating steps akin to the coulomb counting method by leveraging measured parameters such as OCV, terminal voltage, load current, and ambient temperature, as described in Sections 2.3.3.5 and 2.3.3.6.

The initial SoC for both the EKF and UKF methods is deliberately set at 50%, as this represents the midpoint of the SoC range from 0% to 100%. This choice ensures an unbiased estimation process, allowing both methods an equal starting point. Setting the initial SoC at 50% enables more efficient convergence toward the actual SoC values, achieving faster convergence compared to other starting points. The figures indicate that the UKF method demonstrates a superior ability to achieve faster convergence toward actual SoC values, following a linear trend compared to the EKF method, which converges along a more curved trajectory when the initial estimates are misaligned with the actual SoC. This offers a significant advantage in the rapid self-recalibration of SoC estimates for LIB and LiBat systems in the AMR sector, particularly when the initial state of the battery is unknown.

Additionally, the fluctuations and spiking in the estimated SoC values produced by the EKF method are significantly smaller than those from the UKF method. The EKF consistently maintains deviations within a narrower range of SoC error values, closely tracking the actual SoC trend over time. In contrast, the UKF method exhibits more unpredictable behavior, with greater variation in offset drift, rates, and magnitude throughout the estimation process, resulting in a wider range of SoC error values. This demonstrates the superior robustness of the EKF method in accurately reflecting the SoC dynamics during the cycle of LIB and LiBat systems compared to the UKF method. In addition, the figures reveal that fluctuations in the estimated SoC values from both EKF and UKF

methods increase as the operating temperature of LIB and LiBat systems decreases from 40°C to -20°C. These fluctuations and spikes become more pronounced in terms of randomness, rate, offset drift, and magnitude, resulting in a wider range of SoC error values.

According to Graphs A-5 (a,b), Graphs A-6 (a,b), Graphs A-7 (k,l), and Graphs A-8, in contrast to the dynamic state cycle that characterizes the majority of LIB and LiBat systems, the SoC error values during the steady state cycle present a distinct behavior. Initially, the error values gradually increase, reach a peak, and subsequently decline. This trend suggests that both estimation methods demonstrate reduced accuracy as SoC values increase from 0% to 50% and decrease from 100% to 50%, while accuracy improves as SoC values rise from 50% to 100% and decrease from 50% to 0%. In comparison, the SoC error values in the dynamic state cycle do not exhibit a curve-like pattern; instead, they display irregular fluctuations and spikes throughout. Overall, the fluctuations and spikes vary by battery type, with Li-Ion showing the most variability, followed by Li-NCA, Li-NMC, Li-PO, and Li-FP graphite, which exhibits the least fluctuation.

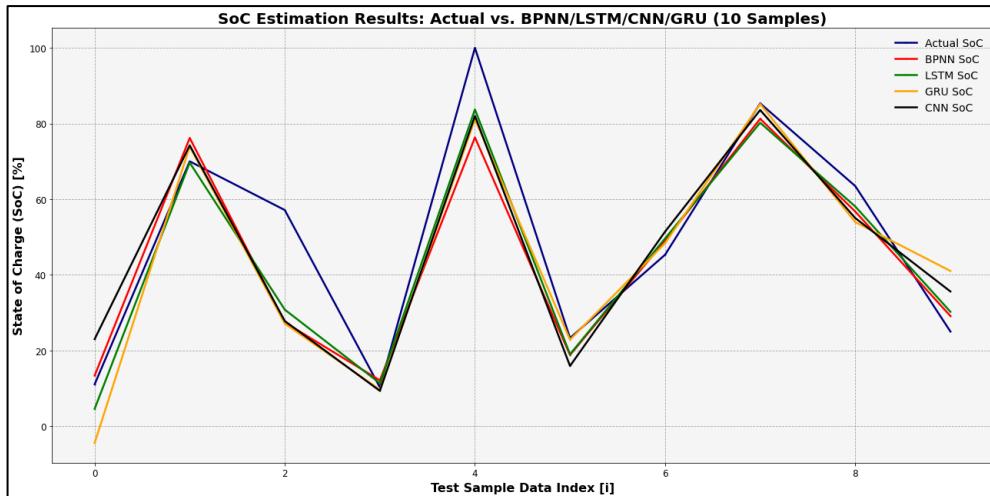
4.3.2 Data-Driven SoC Estimation Results

Table 4.27 presents the testing results for SoC estimation values from various ANN methods using BPNN, CNN, GRU, and LSTM models, providing performance metrics (MAE, RMSE, R²) for the testing datasets.

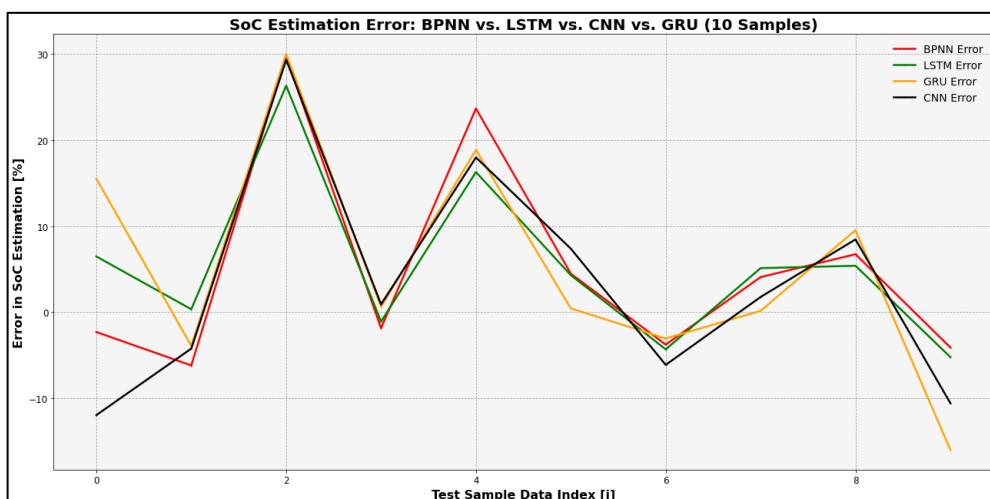
Table 4.27: SoC Estimation Results Using ANN Models for LIB and LiBat Systems Testing Datasets.

	BPNN	CNN	LSTM	GRU
MAE	6.1325	6.8554	6.4907	6.5692
RMSE	10.1701	11.2293	10.4336	10.8746
R²	0.8806	0.8544	0.8743	0.8634

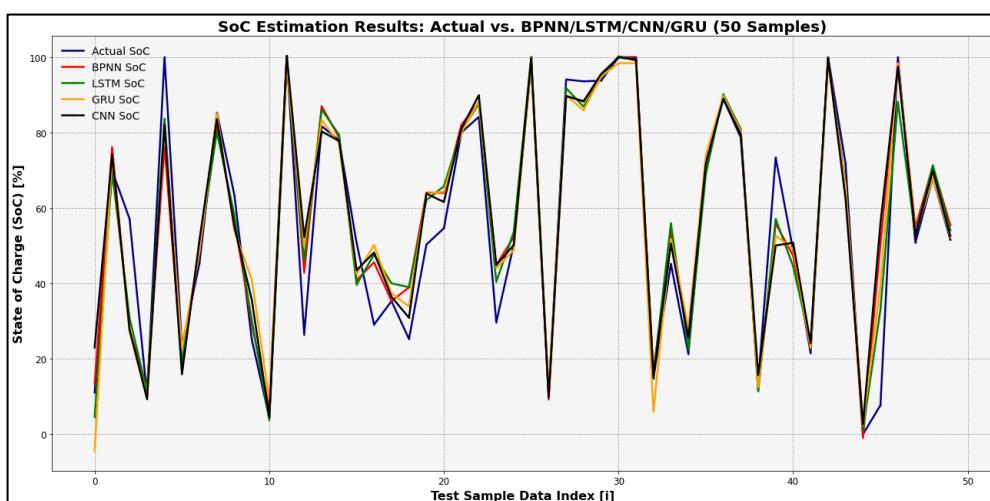
Figure 4.2 presents the estimated SoC values from various ANN methods (BPNN, CNN, GRU, LSTM) alongside with experimental SoC values, highlighting estimation errors values for the first 10, 50, and 100 samples, representing the testing datasets.



(a).



(b).



(c).

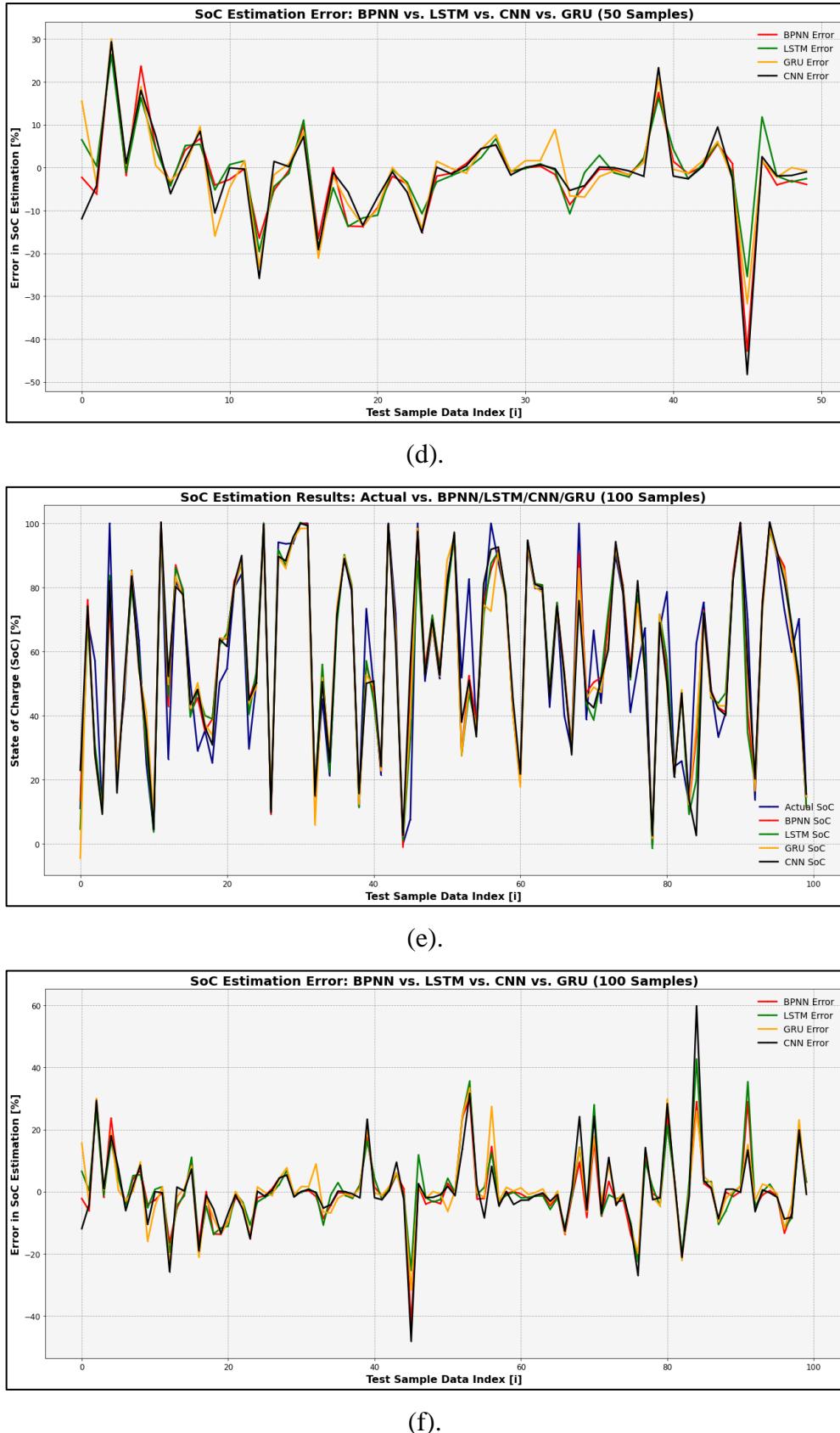
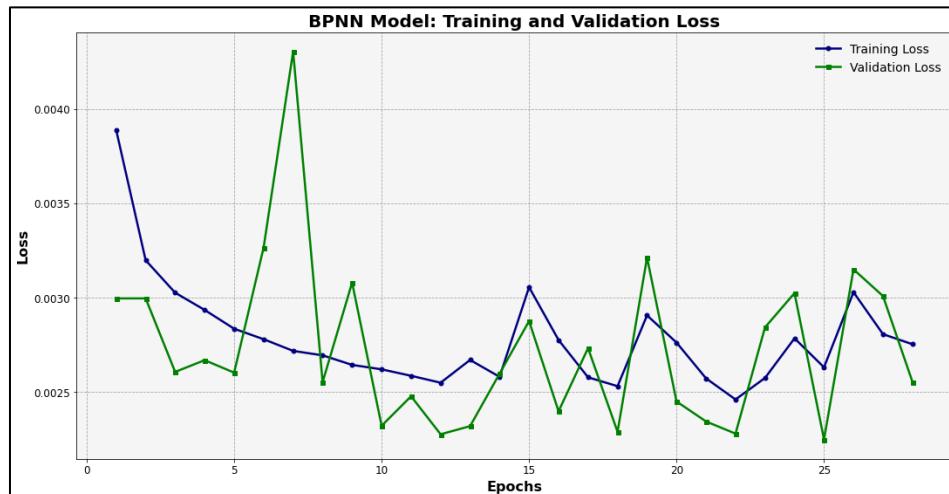


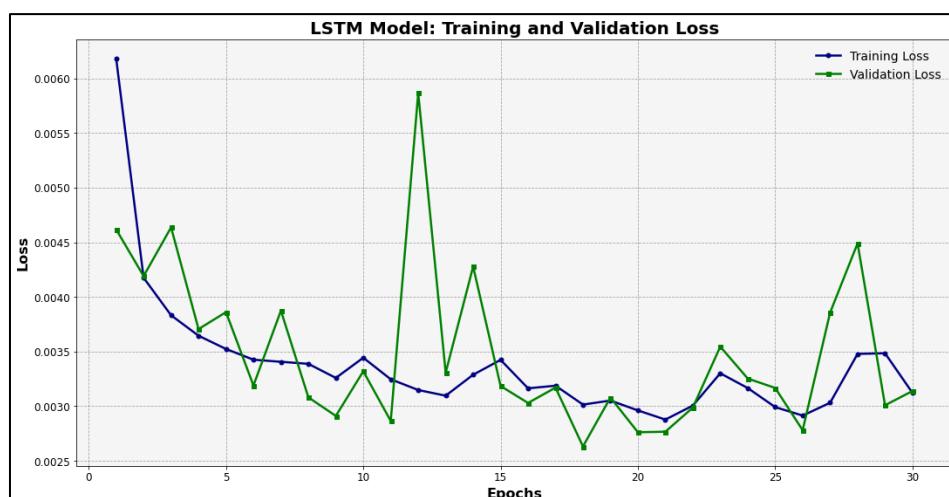
Figure 4.2: Comparison of SoC Curves between Actual, BPNN, LSTM, GRU, and CNN models, along with Error Curves for BPNN, LSTM, GRU, and CNN.

and CNN models at (a, b) the First 10 Samples, (c, d) 50 Samples, and (e, f) 100 Samples.

Figure 4.3 illustrates the training and validation loss values obtained from various ANN methods, including BPNN, CNN, GRU, and LSTM models throughout the training and validation process over 30 epochs. The figures display the loss values for training and validation datasets, which represent 80 and 10% of the input-output databases, respectively. These visualizations provide clear insights into the training and validation processes for each ANN SoC estimation model.



(a).



(b).

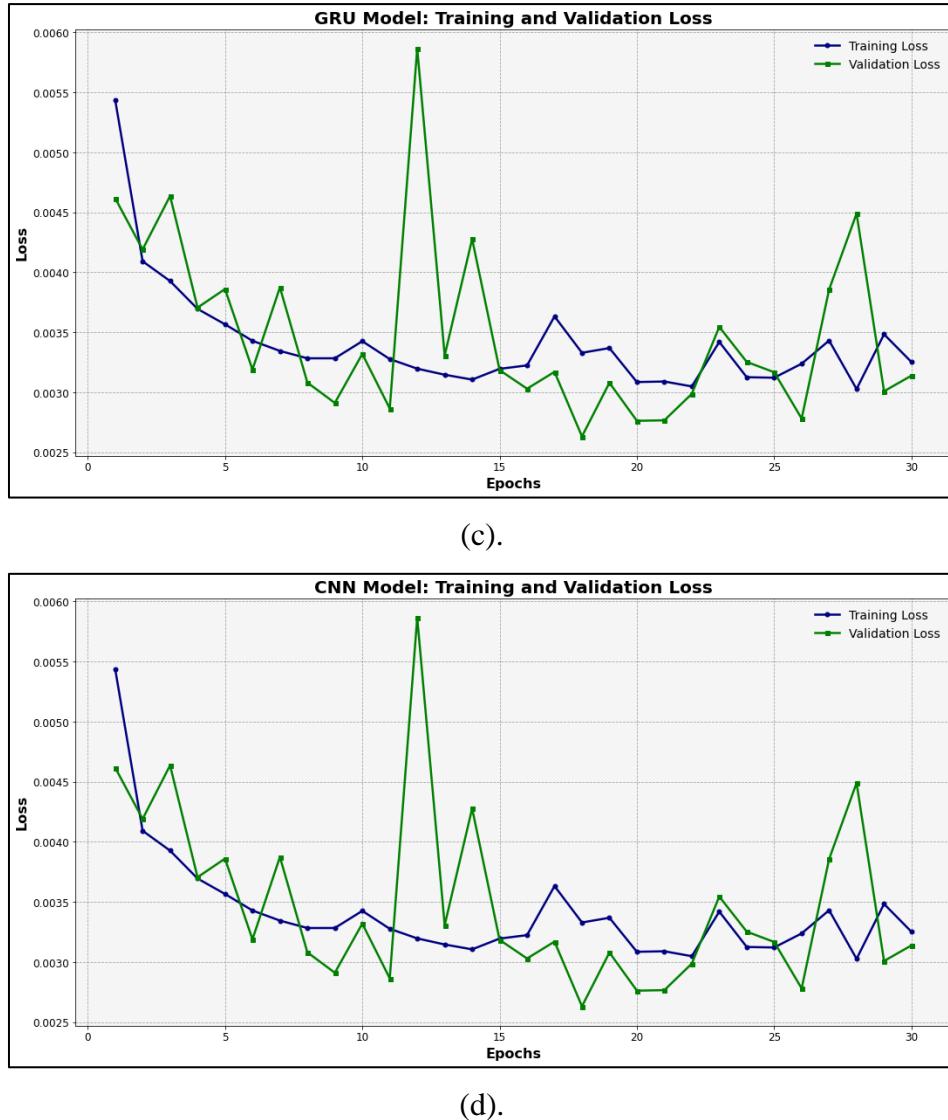


Figure 4.3: Comparison of Loss Curves During Training and Validation for (a) BPNN, (b) LSTM, (c) GRU, and (d) CNN Models Over 30 Epochs.

The results presented in Table 4.27 show that the optimized ANN SoC estimation models, including BPNN, LSTM, GRU, and CNN, perform effectively across the testing datasets, which account for 10% of the input-output databases. Each model exceeds the baseline performance metrics, achieving an MAE of less than 10, an RMSE below 15, and an R^2 greater than 0.8. These findings highlight the robustness of ANN models for SoC estimation across various LIB and LiBat systems, such as Li-NCA, Li-NMC, Li-PO, Li-FP, and Li-Ion graphite batteries, operating under different temperature conditions, states, and rates. However, the performance of each ANN model

varies, with BPNN demonstrating the highest accuracy, followed by LSTM, GRU, and CNN in descending order.

The BPNN method achieves an MAE of 6.1325, an RMSE of 10.1701, and an R^2 of 0.8806, exceeding the baseline by 38.68% for MAE, 32.20% for RMSE, and 10.08% for R^2 , with an average surpassion of 26.98%. The LSTM method achieves an MAE of 6.4907, an RMSE of 10.4336, and an R^2 of 0.8743, exceeding the baseline by 35.09% for MAE, 30.44% for RMSE, and 9.29% for R^2 , with an average surpassion of 24.94%. Similarly, the GRU method achieves an MAE of 6.5692, an RMSE of 10.8746, and an R^2 of 0.8634, exceeding the baseline by 34.31% for MAE, 27.50% for RMSE, and 7.93% for R^2 , with an average surpassion of 23.25%. Finally, the CNN method achieves an MAE of 6.8554, an RMSE of 11.2293, and an R^2 of 0.8544, exceeding the baseline by 31.45% for MAE, 25.14% for RMSE, and 6.80% for R^2 , with an average surpassion of 21.13%.

Furthermore, Figure 4.2 illustrates the relationship between the estimation performances of the various ANN models. All the ANN models produce estimated SoC values that follow similar trends, deviating slightly from the actual SoC values—either marginally higher or lower. This consistent pattern highlights the comparable behavior of each ANN model in relation to the actual SoC data. Hence, this results in ANN models producing SoC error values that fluctuate and spike along the axis, within a similar range. Nevertheless, it is observed that among all the ANN models, the CNN method produces SoC error values range with the largest magnitude of fluctuations and spikes. This is followed by the LSTM and GRU methods, which exhibit a higher fluctuation rate than the CNN method but with lower overall magnitudes. Lastly, the BPNN method demonstrates a moderate level of both fluctuation rate and magnitude compared to the other methods.

Figure 4.3 illustrates the relationship between the training and validation loss values for the various ANN models throughout the entire training and validation process for SoC estimation in LIB and LiBat systems. At the beginning of the training process, the training loss values start relatively high and gradually decrease, following a curved trend. Notably, the BPNN method starts with a training loss value that is relatively lower than the other models, followed by the GRU, CNN, and LSTM methods. Subsequently, after the 9th

epoch for the LSTM, GRU, and CNN methods and the 11th epoch for the BPNN method, the training loss enters a fluctuating and spiking phase. Throughout the training process, the training loss values remain relatively close to the validation loss values, suggesting that none of the models are overfitting to the training datasets.

In contrast, the validation process exhibits a different pattern. The validation loss values start relatively high and gradually decrease, characterized by a series of fluctuations and spikes. After reaching a peak, the amplitude of these fluctuations diminishes, entering a stable region and eventually converging towards a minimum value. Notably, the BPNN method begins with a validation loss that is lower than that of the other models, followed by LSTM, GRU, and CNN. Subsequently, after the 12th epoch for the LSTM, GRU, and CNN methods, and the 5th epoch for the BPNN method, the validation loss transitions into a fluctuating and spiking phase. This enables the BPNN method to reach a local minimum validation loss earlier than the other methods within the 30 epochs.

4.3.3 Overall SoC Estimation Results

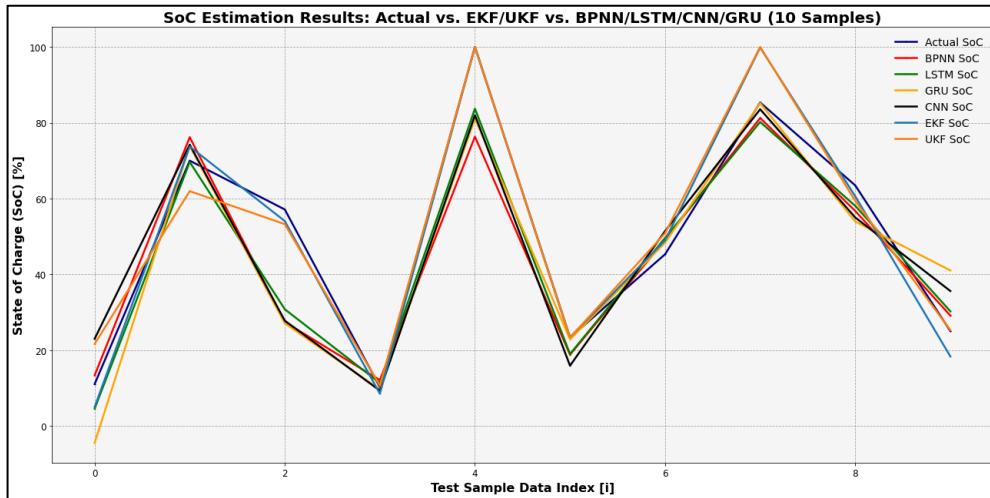
Table 4.28 summarizes SoC estimation results from both conventional (EKF, UKF) and data-driven (BPNN, CNN, GRU, LSTM) methods, providing testing performance metrics for data-driven models and overall metrics (MAE, RMSE, R²) for conventional methods.

Table 4.28: Overall Summary of SoC Estimation Results for LIB and LiBat Systems: EKF, UKF, BPNN, LSTM, GRU, and CNN.

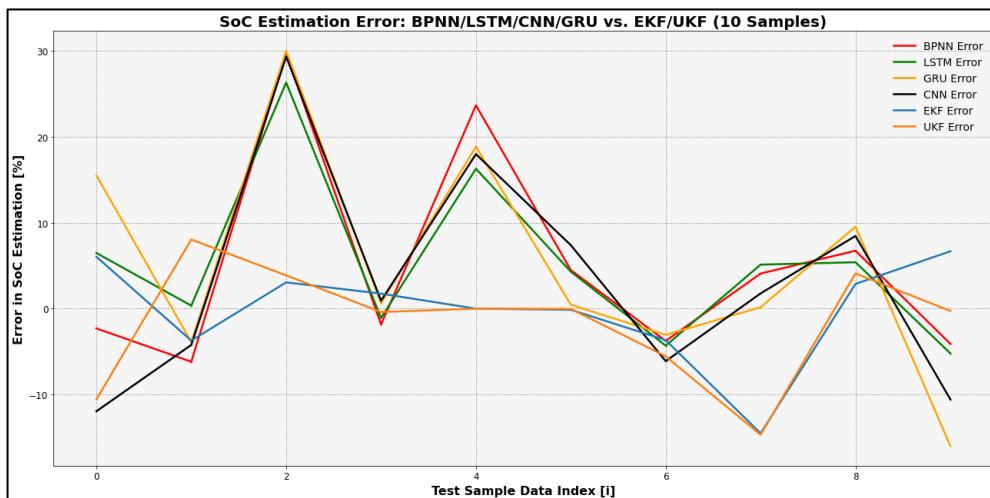
	EKF	UKF	BPNN	GRU	LSTM	CNN
MAE	6.3182	7.4654	6.1325	6.5692	6.4907	6.8554
RMSE	8.9174	10.0908	10.1701	10.8746	10.4336	11.2293
R²	0.8608	0.8279	0.8806	0.8634	0.8743	0.8544

Figure 4.4 shows the estimated SoC values from conventional (EKF, UKF) and data-driven (BPNN, CNN, GRU, LSTM) methods alongside experimental SoC for the first 10, 50, and 100 samples of the testing datasets,

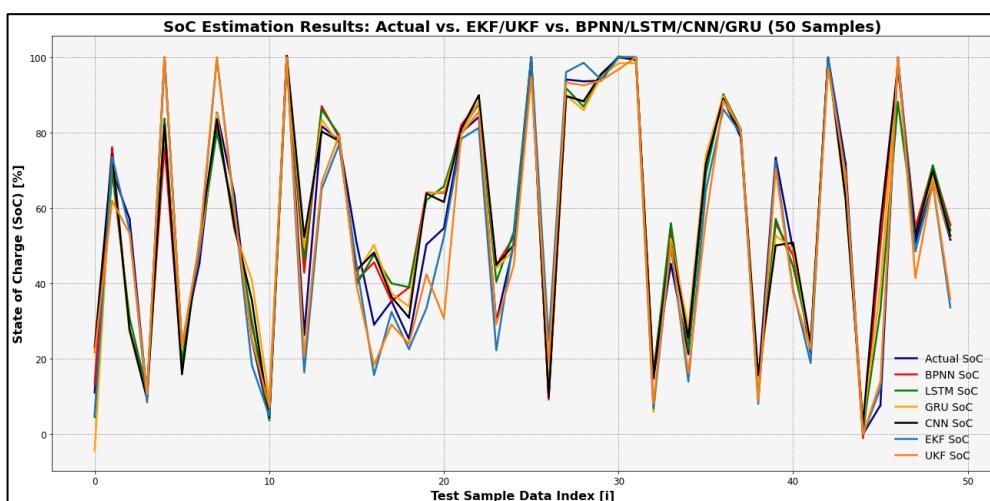
highlighting the estimation error values and offering insights into the models' performance across the LIB and LiBat systems.



(a).



(b).



(c).

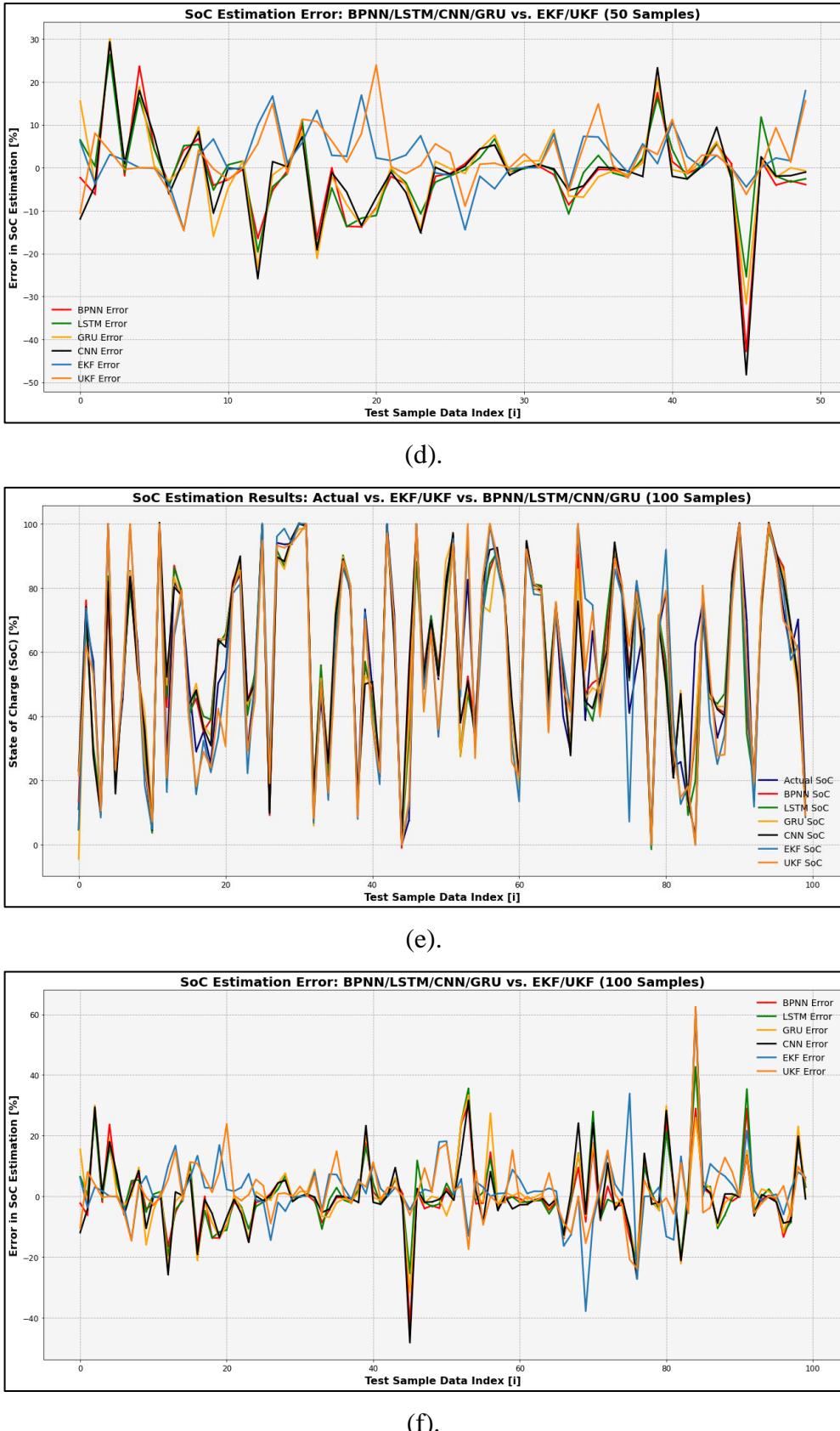


Figure 4.4: Comparison of SoC Curves between Actual, BPNN, LSTM, GRU, CNN, EKF, and UKF models, along with Error Curves for BPNN, LSTM, GRU, CNN, EKF, and UKF.

LSTM, GRU, CNN EKF, and UKF models at (a, b) the First 10 Samples, (c, d) 50 Samples, and (e, f) 100 Samples.

Based on Table 4.28, the EKF method achieves an MAE of 6.3182, an RMSE of 8.9174, and an R^2 of 0.8608, yielding the highest average surpassion of 28.32%. The BPNN method follows with an MAE of 6.1325, an RMSE of 10.1701, and an R^2 of 0.8806, recording the second-highest average surpassion of 26.98%. The LSTM method, with an MAE of 6.4907, an RMSE of 10.4336, and an R^2 of 0.8743, ranks third, with an average surpassion of 24.94%. The GRU method achieves an MAE of 6.5692, an RMSE of 10.8746, and an R^2 of 0.8634, rank fourth, with an average surpassion of 23.25%. The CNN method shows an MAE of 6.8554, an RMSE of 11.2293, and an R^2 of 0.8544, rank fifth, with an average surpassion of 21.13%. Finally, the UKF method, outperforming only in MAE at 7.4654, RMSE at 10.0908, and R^2 at 0.8279, achieves the lowest average surpassion of 20.52%.

These findings demonstrate that conventional SoC estimation models, such as EKF and UKF, exhibit relatively higher MAE and lower R^2 values but perform well in terms of RMSE compared to data-driven models like BPNN, LSTM, GRU, and CNN. This suggests that conventional models can maintain low SoC error values close to zero throughout the LIB and LiBat system cycles. However, despite their ability to minimize error, conventional models struggle to produce SoC estimates that closely align with the actual SoC values of these systems, limiting their overall accuracy. Even though the EKF method outperforms all other methods in terms of overall SoC estimation performance, with the highest average surpassion across metrics—indicating its suitability for SoC estimation in LIB and LiBat systems—this advantage is primarily due to its relatively lower RMSE. However, the BPNN method surpasses EKF in both MAE and R^2 , highlighting that BPNN may be the superior choice for tasks where accuracy and consistency in error reduction are critical.

Apart from that, it is also important to note that, unlike the BPNN method, which only requires tuning around ten hyperparameters to acquire an optimized SoC estimation model, the EKF method necessitates fine-tuning ten different parameters for five distinct battery types across six operational temperatures—resulting in a total of three hundred key parameters. Moreover,

many of these key parameters, such as dynamic resistance, internal resistance, capacitance, and OCV, must be obtained through a series of detailed experiments on the LIB and LiBat systems. Therefore, optimizing an EKF SoC estimation model involves a tedious process of tuning key parameters, collecting accurate data for the battery system, and simulating its internal structure and operation through Thevenin circuit modeling.

Figure 4.4 further elaborates on the SoC estimation performance of the EKF and UKF methods, demonstrating that while the SoC error values tend to fluctuate around the actual SoC values, there are specific points during the estimation where the EKF and UKF methods experience significant spikes, reaching both high maximum and minimum SoC error values. Notably, the EKF and UKF methods tend to estimate SoC values that are relatively higher or lower than the actual SoC values at these points. In contrast, the ANN models exhibit more stable fluctuations and spiking throughout the entire cycle of the LIB and LiBat systems, maintaining a more consistent estimation performance. This highlights the distinct estimation patterns between conventional and data-driven methods.

4.4 Challenges and Limitation of Study

The first challenge in this study was the lack of available datasets in open-source repositories detailing the relationship between key parameters and the SoC of Li-FP, Li-NCA, Li-NMC, Li-PO, and Li-Ion graphite batteries. Critical parameters such as internal resistance, dynamic resistance, and capacitance, which are vital for conventional SoC estimation models like EKF and UKF, were missing. Additionally, the Li-FP graphite battery lacked key datasets related to OCV, making it difficult to accurately simulate the OCV-SoC relationship using terminal voltage-SoC datasets.

Henceforth, in this research study, the EKF and UKF methods could only be moderately optimized by tuning key parameters to an assumed constant value throughout the entire cycle of the LIB and LiBat systems. This approach was necessary due to the lack of complete conventional model parameter datasets, which prevented the simulation of actual dynamic values based on the SoC relationship. As a result, this limitation hampers the study's effectiveness, as the overall estimation performance of the conventional SoC estimation

models, when compared to data-driven models, does not achieve the optimal potential of these methods.

Another key challenge encountered in this study arose from the data-driven SoC estimation models. The total size of the input-output databases for the LIB and LiBat systems consisted of 25.7 million data points (25,750,534), creating a significantly large dataset. Due to the sheer volume of this data, the training time required for the BPNN, LSTM, GRU, and CNN models was relatively long, with an average training time of 10 hours per attempt. Consequently, during this research, each model was only optimized to a moderate level of estimation performance, which did not fully demonstrate their optimal potential when compared to conventional models.

The third major challenge encountered in this study was the limited availability of datasets for LIB and LiBat systems. The input-output databases consisted only of internal parameter datasets derived from a single battery for each type, which restricted the comprehensiveness of the research. This limitation impacted the evaluation of SoC estimation performance for both conventional and data-driven models. Additionally, the internal parameter datasets for the Li-FP graphite battery were particularly constrained, covering an only operational temperature range of 0°C to 40°C, which further limited the evaluation of the models, especially for the Li-FP graphite battery.

The fourth challenge arose from the lack of well-preprocessed input-output databases for SoC estimation models of LIB and LiBat systems. To obtain reliable input-output databases, all internal parameter datasets had to be manually preprocessed using coulomb counting, look-up methods, and guidance from the dataset providers, adding complexity to the study. Lastly, the fifth major challenge was the limited research time. Due to this constraint, only four data-driven models and two conventional models were developed and tested, leading to a limited comprehensive evaluation of their estimation performance. This time limitation hindered the exploration of additional models, restricting the depth of analysis and comparison between the data-driven and conventional SoC estimation methods.

4.5 Summary

The study examines SoC estimation models through extensive optimization of both conventional and data-driven approaches. Conventional models, specifically EKF and UKF, undergo meticulous tuning of key parameters to accommodate various battery systems, including LG 18650HG2, Samsung IN21700-30T, SB LiMotive, Turnigy Graphene, and A123. These parameters include dynamic resistance, capacitance, state error, and temperature-related variations. Data-driven models—BPNN, LSTM, GRU, and CNN—are optimized by adjusting hyperparameters such as data scaling, learning rates, and activation functions, with specific configurations to enhance performance.

The analysis reveals the EKF method stands out as the most effective conventional model. Then, EKF and UKF models perform consistently across battery systems at temperatures from 0°C to 40°C, with EKF showing MAE values between 1.8558 and 5.2651, RMSE from 3.0296 to 7.5349, and R² between 0.9068 and 0.9977. UKF demonstrates MAE from 3.3998 to 6.9032, RMSE from 4.3684 to 8.5008, and R² between 0.8814 and 0.9787. However, performance deteriorates from 0°C to -20°C, with significant increases in MAE and RMSE and decreases in R² for both methods, highlighting the challenge of low operational temperatures. The Li-FP battery system shows notable instability compared to other battery types.

Among the data-driven models, BPNN excels with the highest accuracy, achieving an MAE of 6.1325, RMSE of 10.1701, and R² of 0.8806, surpassing the baseline by significant margins. It is followed by LSTM, GRU, and CNN in terms of performance. The results indicate that while all ANN models follow similar SoC estimation trends, CNN exhibits the largest magnitude of fluctuations and spikes in SoC error, whereas BPNN shows moderate fluctuation levels and overall superior performance, while LSTM and GRU show the highest rate of fluctuations and spikes in SoC error.

Lastly, the EKF method achieves the highest average performance surpassion of 28.32%, surpassing BPNN, LSTM, GRU, and CNN. Despite its effectiveness in maintaining low RMSE, EKF has higher MAE, lower R² values, and requires complex parameter tuning, limiting its overall rating to second-best among SoC estimation models. Additionally, both EKF and UKF methods experience significant SoC error spikes, while data-driven models exhibit more

stable performance throughout battery cycles. This stability underscores the advantages of data-driven approaches, particularly BPNN, which consistently provides more accurate and reliable SoC estimations compared to traditional methods.

The limitations of the study include several critical challenges impacting the optimization and evaluation of SoC estimation models. The primary issue was the lack of comprehensive open-source datasets detailing key parameters such as internal resistance, dynamic resistance, and capacitance for various battery types, which hampered the accurate calibration of conventional models like EKF and UKF. This data deficiency led to the moderate optimization of these models, with key parameters set to constant values throughout battery cycles. Additionally, the extensive dataset of 25.7 million data points for data-driven models resulted in long training times, limiting the extent of model optimization and performance. The availability of limited datasets from only a single battery type for each category further restricted the evaluation, particularly for the Li-FP battery with a narrow temperature range. Compounding these issues was the need for manual preprocessing of datasets, adding complexity to the study, and a constrained research timeline that allowed for the development and testing of only a few models. These factors collectively constrained the depth of analysis and the ability to fully explore the potential of both SoC estimation methods.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

In conclusion, the FYP study demonstrates that despite the EKF method showing the highest average deviation of 28.32%, its performance metrics, including MAE and R^2 , are relatively higher and lower, respectively, when compared to the BPNN model. This indicates that EKF has limited capability in providing a highly accurate overall SoC estimation. Moreover, the process of optimizing the EKF method is tedious, time-consuming, and costly, with over 300 parameters to be optimized. On the other hand, the BPNN model emerges as the best-performing SoC estimation model for LIB and LiBat systems, despite its second-highest average deviation of 26.98%. With MAE and R^2 showing lower and higher values, respectively, the BPNN model proves to be more reliable in overall SoC estimation, even with a moderate level of fine-tuning of ten hyperparameters. This study underscores the necessity of implementing data-driven SoC estimation models for LIB and LiBat systems in the AMR sector.

This FYP study reveals that the overall estimation performance of the SoC estimation models can be ranked as follows: BPNN is the best-performing model, followed by EKF, LSTM, GRU, CNN, with UKF ranking lowest in performance. Among the conventional models, EKF demonstrates significantly more robust performance compared to UKF, where UKF shows greater deficiencies, contrary to theoretical expectations. Additionally, conventional models tend to perform better in higher operational temperature environments. For the data-driven SoC estimation models, despite similarities in the trends and patterns of estimated SoC and SoC error values, the BPNN model is identified as the best-performing model among them.

5.2 Recommendations for future work

Future development of this study could focus on several key areas to improve the accuracy and comprehensiveness of both conventional and data-driven SoC

estimation models. Future research should prioritize collecting data from multiple batteries of the same type, rather than relying on a single battery per system, to improve model generalization. Additionally, more detailed datasets of LIB and LiBat systems, covering a broader range of operational temperatures, states, and discharge/charge rates, should be gathered. This will enhance the robustness and reliability of SoC estimation models across different operating conditions.

Second, it is essential to address the lack of available datasets for conventional models in open-source repositories. Future research should focus on developing and collecting comprehensive datasets through detailed experiments and sourcing from open repositories. These datasets should include critical parameters such as internal resistance, dynamic resistance, and capacitance for various battery chemistries, including Li-FP, Li-NCA, Li-NMC, Li-PO, and Li-Ion graphite batteries. Expanding these datasets to include OCV data, particularly for Li-FP graphite batteries. Furthermore, the development dynamic parameter of the Thevenin circuit model captures real-time variations based on SoC, allowing conventional methods like EKF and UKF to reach their full potential.

Third, a deeper and more detailed optimization of the training process through fine-tuning data-driven models like BPNN, LSTM, GRU, and CNN is a critical area for improvement. This would enable more extensive model optimization, unlocking the full potential of these models and providing a more comprehensive evaluation of SoC estimation performance. Additionally, expanding research to explore a wider range of both data-driven and conventional SoC estimation methods, such as the LO method, SMO method, PIO method, radial basis neural networks, wavelet neural networks, and deep neural networks, would allow for more thorough investigation. This approach would lead to deeper insights and a more robust comparison of performance across different models. By addressing these areas, future work can significantly enhance the accuracy and reliability of the research study for the SoC estimation models for LIB and LiBat systems.

REFERENCES

- Adil Khan, M., 2020. SOC estimation of Li-ion battery using Kalman filter. *GitHub*, [online] 23 September. Available at: <https://github.com/adilkhan095/SOC-Estimation-of-Li-ion-battery-using-Kalman-filter/blob/master/Adil%20Conference%20paper.pdf> [Accessed 12 April 2024].
- Ali, M. U., Zafar, A., Nengroo, S. H., Hussain, S., Alvi, M. J., and Kim, H., 2019. Towards a smarter battery management system for electric vehicle applications: A critical review of lithium-ion battery state of charge estimation. *Energies*, 12(3), p. 446. <https://doi.org/10.3390/en12030446>.
- Alter, W.L., Gu, P.C., and Cong, J., 2022. Battery state of charge estimation using Kalman filter. *GitHub*, [online] 9 December. Available at: https://github.com/AlterWL/Battery_SOC_Estimation [Accessed 11 April 2024].
- Catenaro, E., and Onori, S., 2021b. Experimental data of lithium-ion batteries under galvanostatic discharge tests at different rates and temperatures of operation. *Data in Brief*, 35, p. 106894. <https://doi.org/10.1016/j.dib.2021.106894>.
- Chang, W., 2013. The state of charge estimating methods for battery: A review. *ISRN Applied Mathematics*, 2013, pp. 1–7. <https://doi.org/10.1155/2013/953792>.
- Dai, H., Guo, P., Wei, X., Sun, Z., and Wang, J., 2014. ANFIS (adaptive neuro-fuzzy inference system) based online SOC (state of charge) correction considering cell divergence for the EV (electric vehicle) traction batteries. *Energy*, 80, pp. 350–360. <https://doi.org/10.1016/j.energy.2014.11.077>.
- Dileepan, V. M., Madhavan, K. T., and Jayakumar, J., 2017. Performance analysis of lithium polymer battery and super capacitor. In *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, Chennai, India, 2017, pp. 2029–2033. <https://doi.org/10.1109/ICECDS.2017.8389805>.
- Du, J., Liu, Z., Wang, Y., and Wen, C., 2014. A fuzzy logic-based model for Li-ion battery with SOC and temperature effect. In *11th IEEE International Conference on Control & Automation (ICCA)*, Taichung, Taiwan, 2014, pp. 1333–1338. <https://doi.org/10.1109/ICCA.2014.6871117>.
- Hu, X., Sun, F., and Zou, Y., 2010. Estimation of state of charge of a lithium-ion battery pack for electric vehicles using an adaptive Luenberger observer. *Energies*, 3(9), pp. 1586–1603. <https://doi.org/10.3390/en3091586>.

Huang, S., Tseng, K., Liang, J., Chang, C., and Pecht, M., 2017b. An online SOC and SOH estimation model for lithium-ion batteries. *Energies*, 10(4), p. 512. <https://doi.org/10.3390/en10040512>.

Kollmeyer, P. and Skells, M., 2023. Samsung INR21700 30T 3Ah Li-ion battery data. *Mendeley Data*, [online] 4 December, V2. Available at: <https://doi.org/10.17632/9xyvy2njj3.1> [Accessed 8 August 2024].

Kollmeyer, P. and Skells, M., 2020. Turnigy Graphene 5000mAh 65C Li-ion battery data. *Mendeley Data*, [online] 14 August, V1. Available at: <https://doi.org/10.17632/4fx8cjprxm.1> [Accessed 8 August 2024].

Kollmeyer, P. and Skells, M., 2023. Multi temperature Li-ion battery data - SB LiMotive 5Ah. *Borealis*, [online] 23 August, V1. Available at: <https://doi.org/10.5683/SP3/LFPKAS> [Accessed 8 August 2024].

Kollmeyer, P., Vidal, C., Naguib, M. and Skells, M., 2020. LG 18650HG2 Li-ion battery data and example deep neural network xEV SOC estimator script. *Mendeley Data*, [online] 6 March, V3. Available at: <https://doi.org/10.17632/cp3473x7xv.3> [Accessed 8 August 2024].

Křepelková, M., 2017. Evolution of batteries: From experiments to everyday usage. *Poster 2017*, [online] 23 May. Available at: https://poster.fel.cvut.cz/poster2017/proceedings/Poster_2017/Section_HS/HS_034_Krepelkova.pdf [Accessed 27 April 2024].

Larchuto, 2021. Python implementation. *GitHub*, [online] 4 November. Available at: <https://github.com/larchuto/Battery-Kalman> [Accessed 12 April 2024].

Liu, W., Placke, T., and Chau, K., 2022. Overview of batteries and battery management for electric vehicles. *Energy Reports*, 8, pp. 4058–4084. <https://doi.org/10.1016/j.egyr.2022.03.016>.

Man, K. L., Ting, T. O., Krilavičius, T., Wan, K., Chen, C., Chang, J., and Poon, S. H., 2012. Towards a hybrid approach to SoC estimation for a smart battery management system (BMS) and battery-supported cyber-physical systems (CPS). In *2012 2nd Baltic Congress on Future Internet Communications*, Vilnius, Lithuania, 2012, pp. 113–116. <https://doi.org/10.1109/BCFIC.2012.6217989>.

McNulty, D., Hennessy, A., Li, M., Armstrong, E., and Ryan, K. M., 2022. A review of Li-ion batteries for autonomous mobile robots: Perspectives and outlook for the future. *Journal of Power Sources*, 545, p. 231943. <https://doi.org/10.1016/j.jpowsour.2022.231943>.

Mendoza, S., Liu, J., Mishra, P., and Fathy, H., 2017. On the relative contributions of bias and noise to lithium-ion battery state of charge estimation errors. *Journal of Energy Storage*, 11, pp. 86–92. <https://doi.org/10.1016/j.est.2017.01.006>.

Mikołajczyk, T., Mikołajewski, D., Kłodowski, A., Łukaszewicz, A., Mikołajewska, E., Paczkowski, T., Macko, M., and Skornia, M., 2023. Energy sources of mobile robot power systems: A systematic review and comparison of efficiency. *Applied Sciences*, 13(13), p. 7547. <https://doi.org/10.3390/app13137547>.

Moussalli, Z., Brahim Sedra, M., and Ait Laachir, A., 2018. State of charge estimation algorithms in lithium-ion battery-powered electric vehicles. In *2018 International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*, Kenitra, Morocco, 2018, pp. 1–6. <https://doi.org/10.1109/ICECOCS.2018.8610646>.

Natakarnkitkul, S., 2021b. Show me the black box - towards AI. *Medium*, [online] 13 December. Available at: <https://pub.towardsai.net/show-me-the-black-box-3495dd6ff52c> [Accessed 20 April 2024].

Nur, A., Omaç, Z., and Öksüztepe, E., 2017b. Fuzzy logic based indirect vector control of squirrel cage induction motor. *Sigma Journal of Engineering and Natural Sciences – Sigma Mühendislik Ve Fen Bilimleri Dergisi*, 8, p. 65. <https://sigma.yildiz.edu.tr/article/499>.

Pannakkong, W., Thiwa-Anont, K., Singthong, K., Parthanadee, P., and Buddhakulsomsiri, J., 2022. Hyperparameter tuning of machine learning algorithms using response surface methodology: A case study of ANN, SVM, and DBN. *Mathematical Problems in Engineering*, 2022, pp. 1–17. <https://doi.org/10.1155/2022/8513719>.

Patle, A. and Chouhan, D. S., 2013. SVM kernel functions for classification. In *2013 International Conference on Advances in Technology and Engineering (ICATE)*, Mumbai, India, 2013, pp. 1–9. <https://doi.org/10.1109/ICAdTE.2013.6524743>.

Petit, M., Calas, E., and Bernard, J., 2020. A simplified electrochemical model for modelling Li-ion batteries comprising blend and bidispersed electrodes for high power applications. *Journal of Power Sources*, 479, p. 228766. <https://doi.org/10.1016/j.jpowsour.2020.228766>.

Ren, Q., Balazinski, M., and Baron, L., 2011. Type2 TSK fuzzy logic system and its Type1 counterpart. *International Journal of Computer Applications*, 20(6), pp. 8–13. <https://doi.org/10.5120/2440-3292>.

Redondo-Iglesias, E. and Tian, B., 2024. Characterisation tests of A123 4.4Ah cell. *Recherche Data Gouv*, [online] 20 August, V1. Available at: <https://doi.org/10.57745/XIDTZH> [Accessed 8 August 2024].

Umair Ali, M., Hussain Nengroo, S., Adil Khan, M., Zeb, K., Ahmad Kamran, M., and Kim, H.-J., 2018. A real-time Simulink interfaced fast-charging methodology of lithium-ion batteries under temperature feedback with fuzzy logic control. *Energies*, 11(5), p. 1122. <https://doi.org/10.3390/en11051122>.

Wagner, R., Preschitschek, N., Passerini, S., Leker, J., and Winter, M., 2013. Current research trends and prospects among the various materials and designs used in lithium-based batteries. *Journal of Applied Electrochemistry*, 43(5), pp. 481–496. <https://doi.org/10.1007/s10800-013-0533-6>.

Xiong, R., Cao, J., Yu, Q., He, H., and Sun, F., 2018. Critical review on the battery state of charge estimation methods for electric vehicles. In *IEEE Access*, vol. 6, pp. 1832–1843. <https://doi.org/10.1109/ACCESS.2017.2780258>.

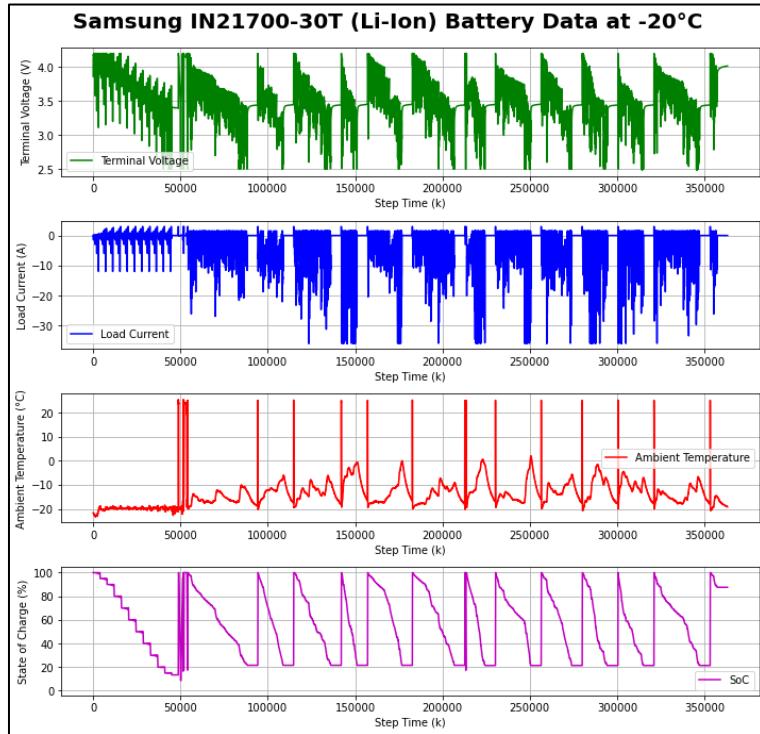
Xing, Y., He, W., Pecht, M., and Tsui, K. L., 2014. State of charge estimation of lithium-ion batteries using the open-circuit voltage at various ambient temperatures. *Applied Energy*, 113, pp. 106–115. <https://doi.org/10.1016/j.apenergy.2013.07.008>.

Yeom, C.-U., and Kwak, K.-C., 2018. A performance comparison of ANFIS models by scattering partitioning methods. In *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Vancouver, BC, Canada, 2018, pp. 814–818. <https://doi.org/10.1109/IEMCON.2018.8614898>.

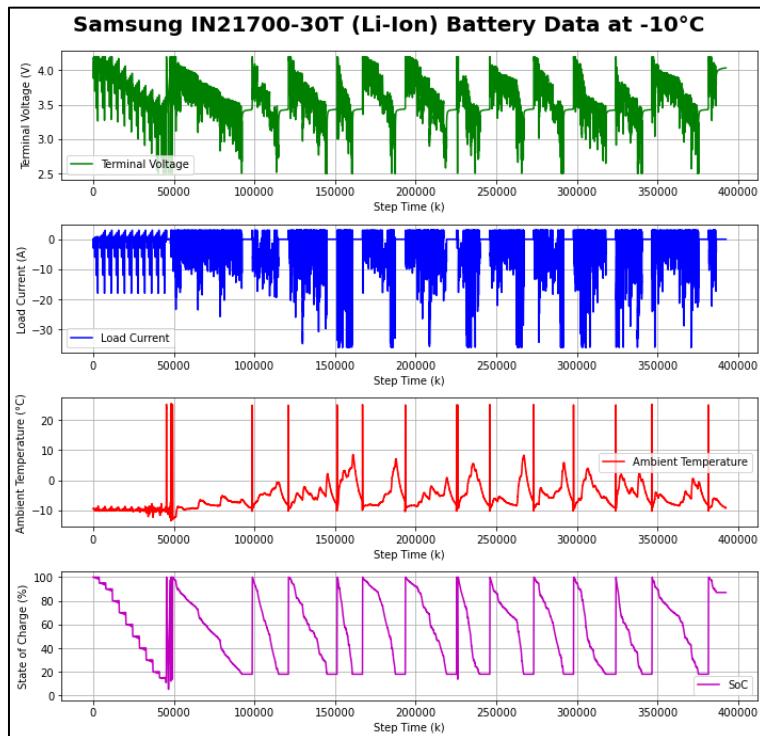
Zou, Z., Xu, J., Mi, C., Cao, B., and Chen, Z., 2014. Evaluation of model-based state of charge estimation methods for lithium-ion batteries. *Energies*, 7(8), pp. 5065–5082. <https://doi.org/10.3390/en7085065>.

APPENDICES

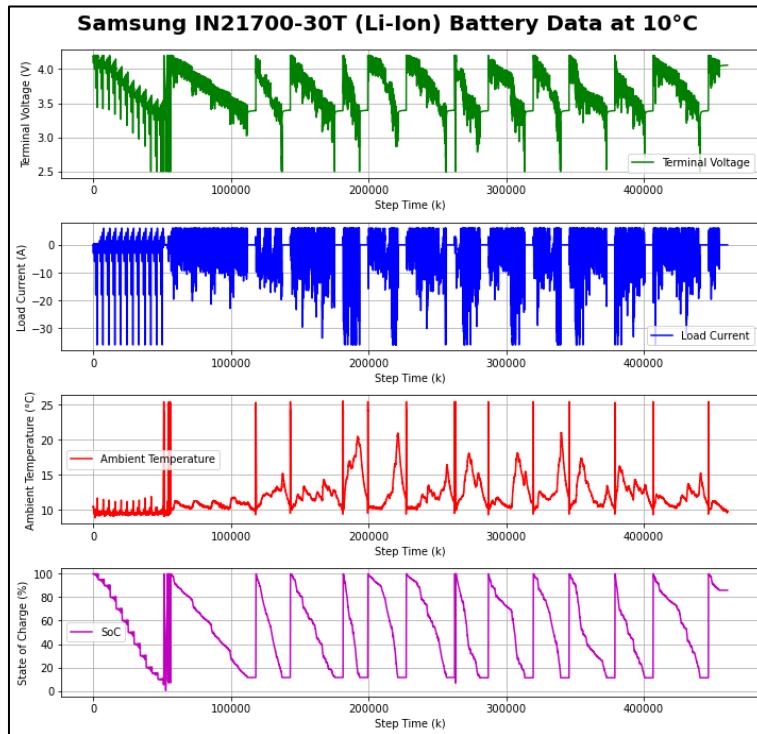
Appendix A: Graphs



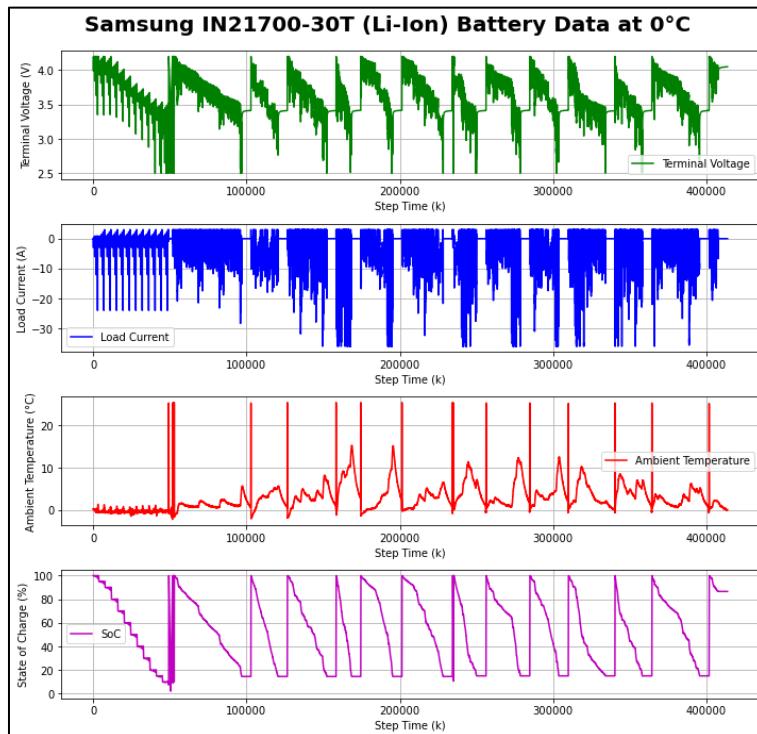
(a).



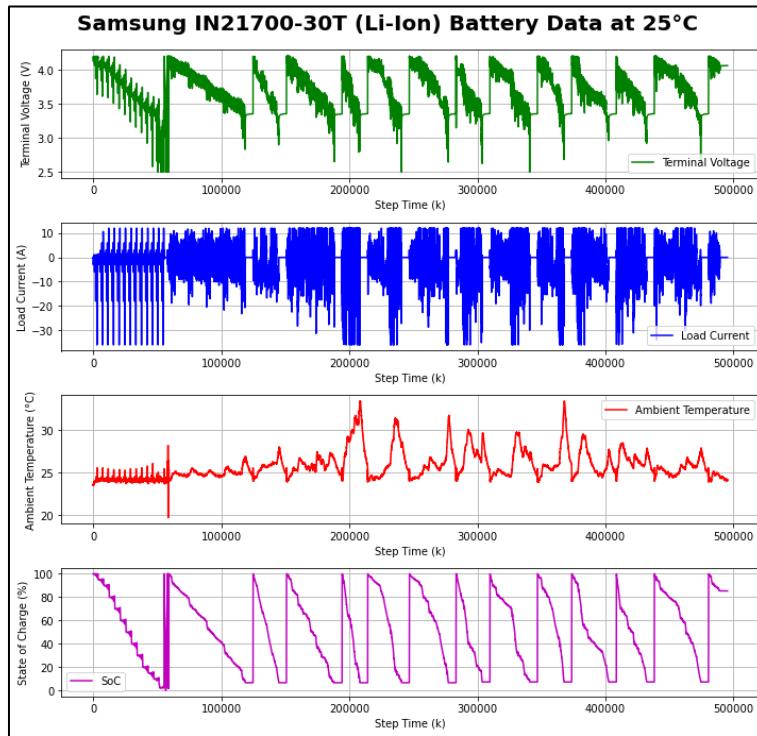
(b).



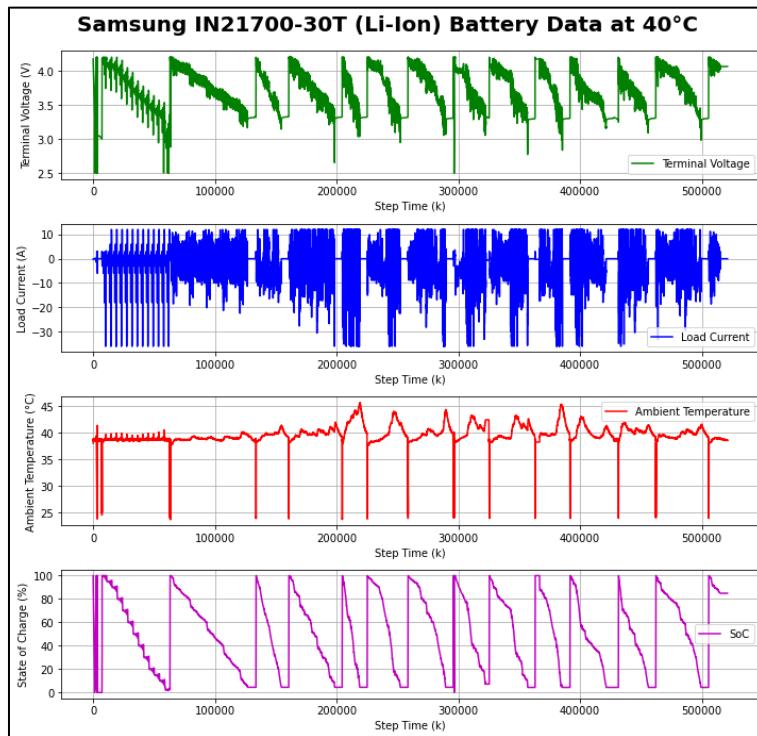
(c).



(d).



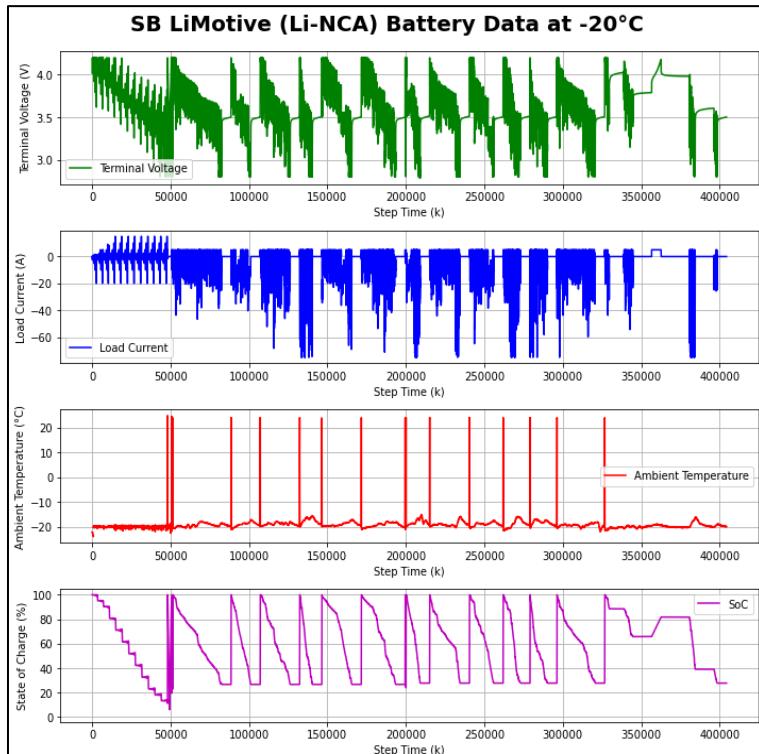
(e).



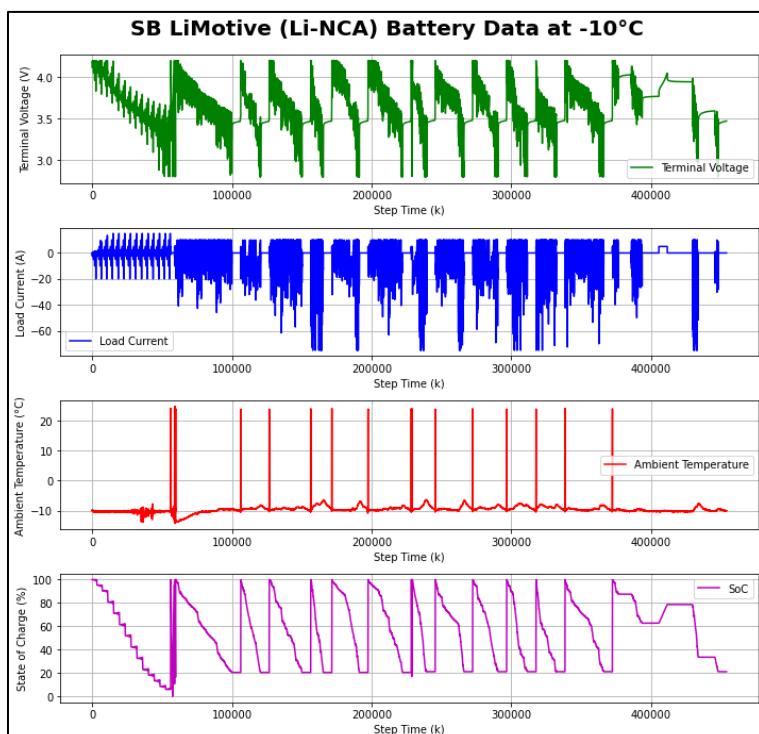
(f).

Graph A-1: Discharge and Charging Curves for the Samsung IN21700-30T (Li-Ion) Battery: Terminal Voltage, Load Current, Ambient

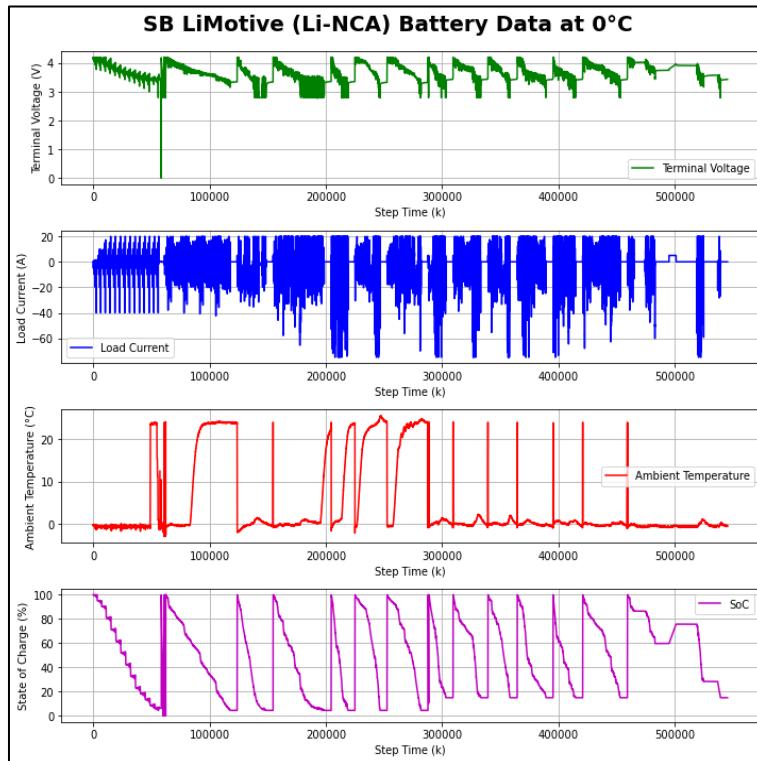
Temperature, and SoC at (a) -20°C, (b) -10°C, (c) 0°C, (d) 10°C, (e) 25°C, and (f) 40°C (Kollmeyer, and Skells, 2023).



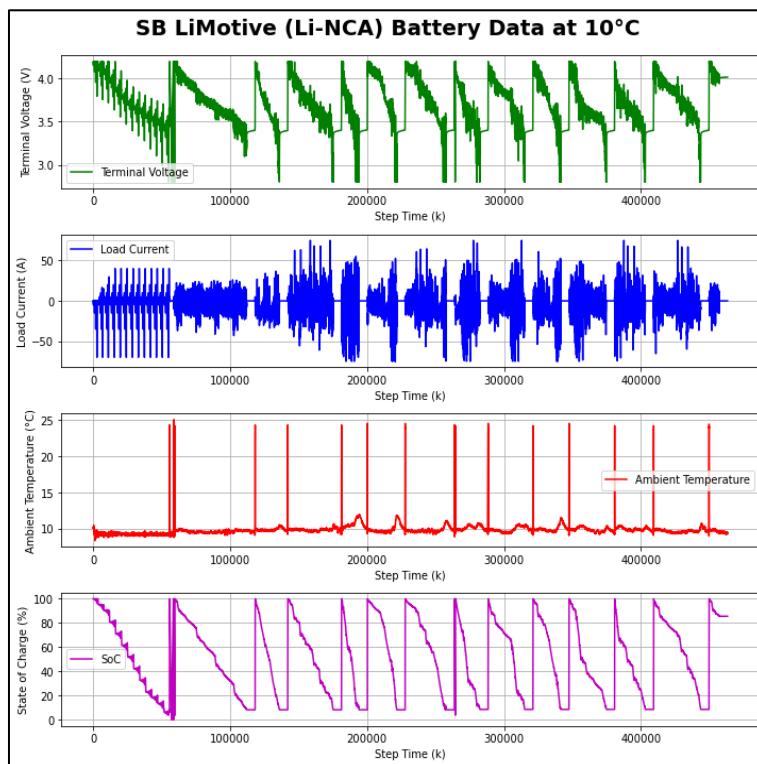
(a).



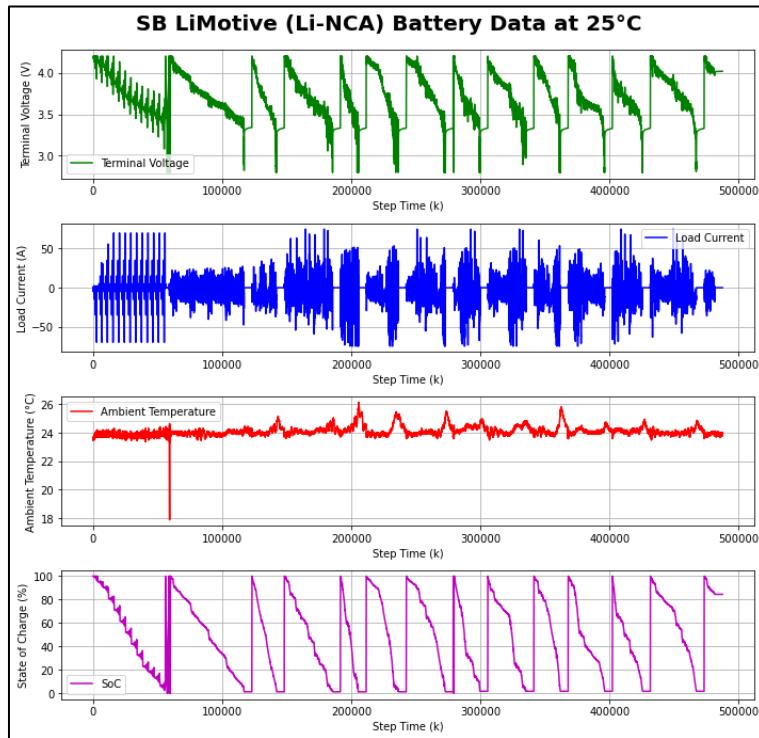
(b).



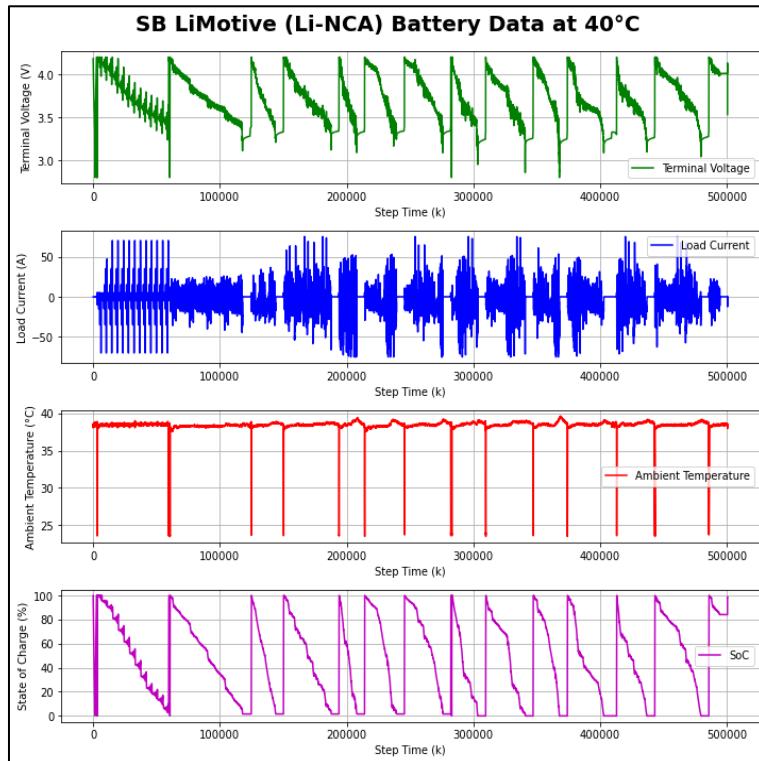
(c).



(d).



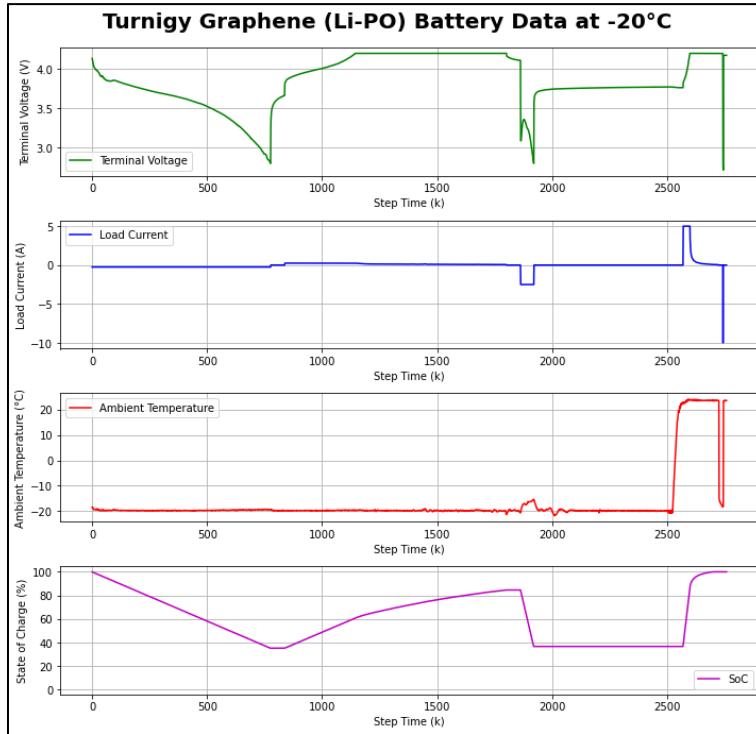
(e).



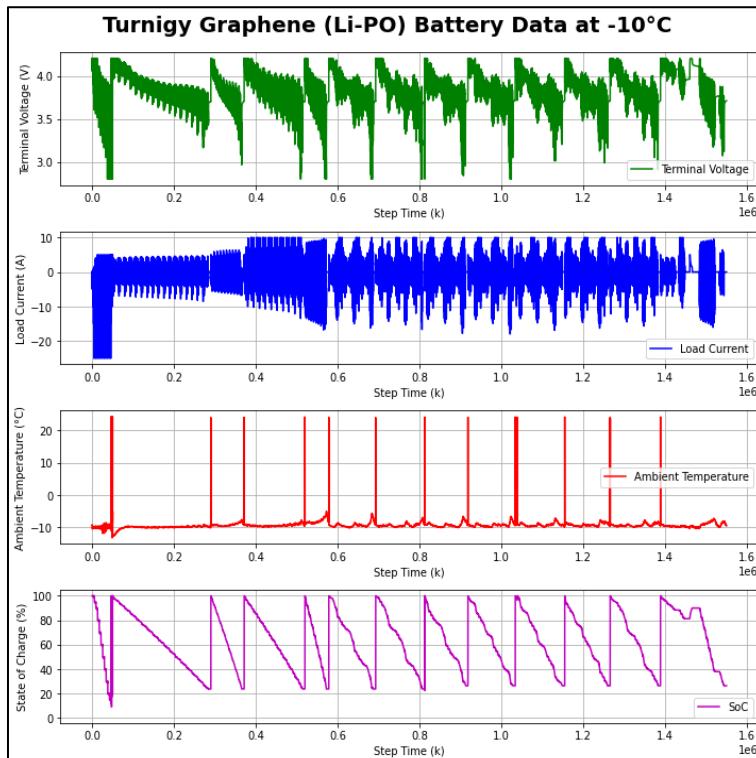
(f).

Graph A-2: Discharge and Charging Curves for the SB LiMotive (Li-NCA)
 Battery: Terminal Voltage, Load Current, Ambient Temperature,
 State of Charge

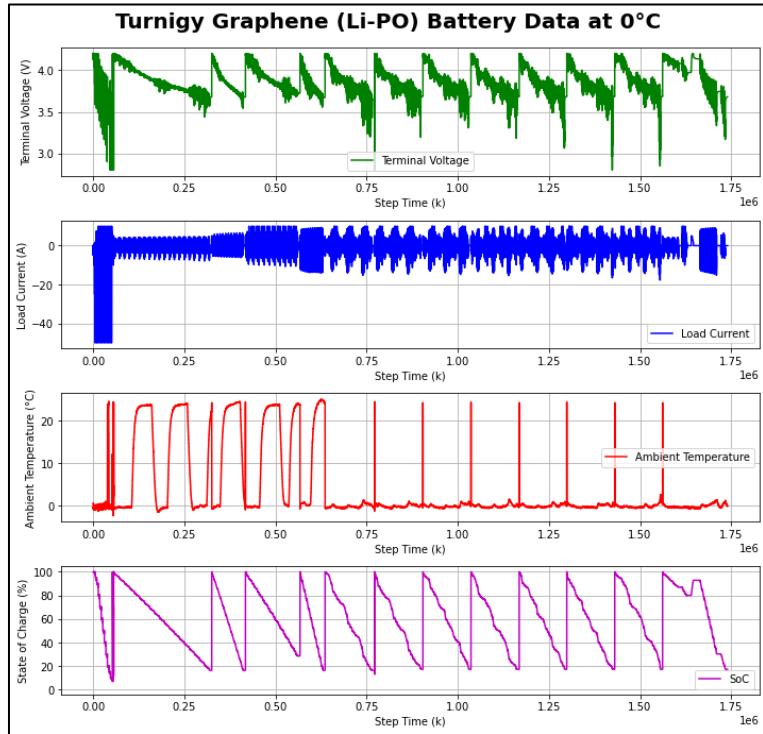
and SoC at (a) -20°C, (b) -10°C, (c) 0°C, (d) 10°C, (e) 25°C, and (f) 40°C (Kollmeyer, and Skells, 2023).



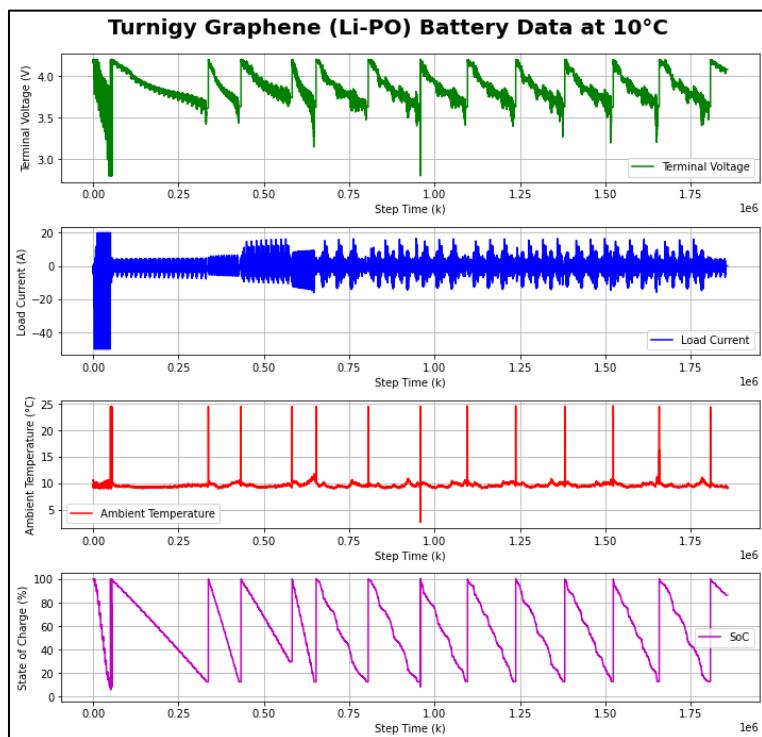
(a).



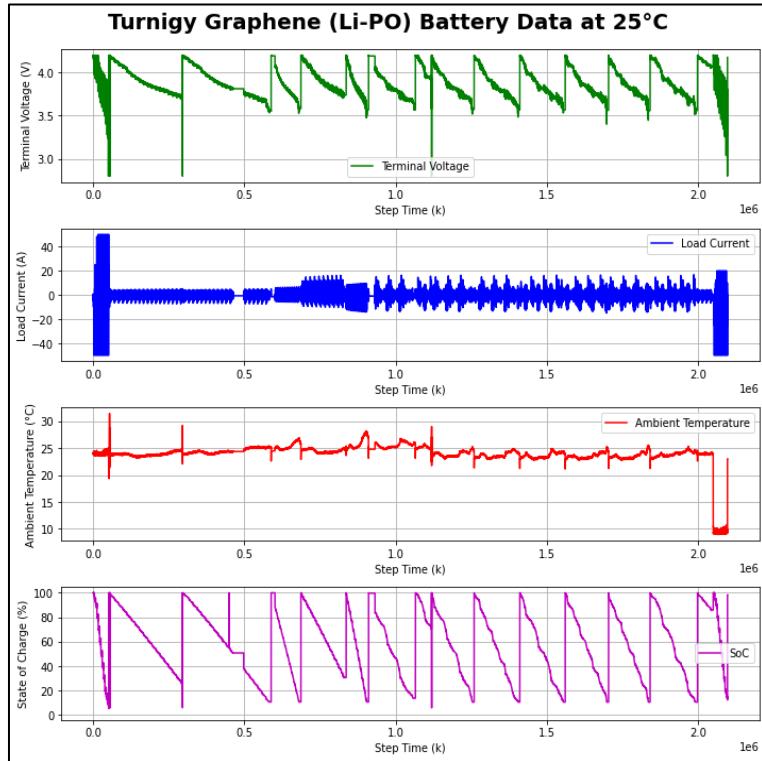
(b).



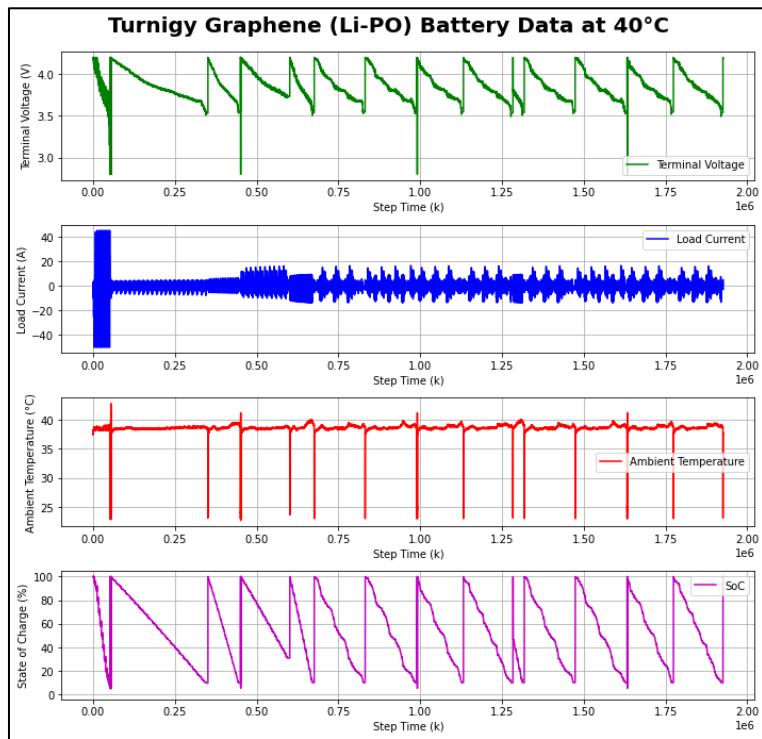
(c).



(d).



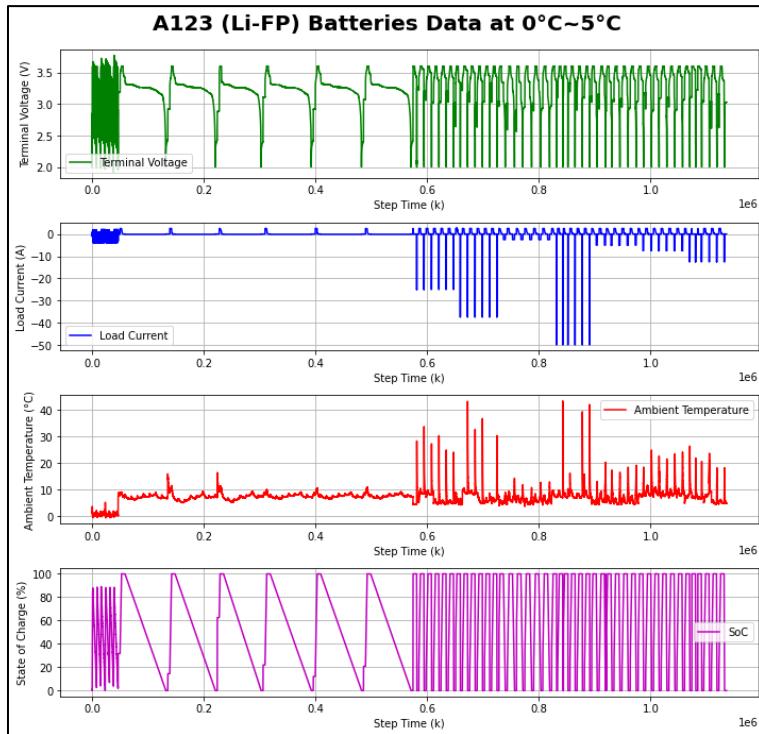
(e).



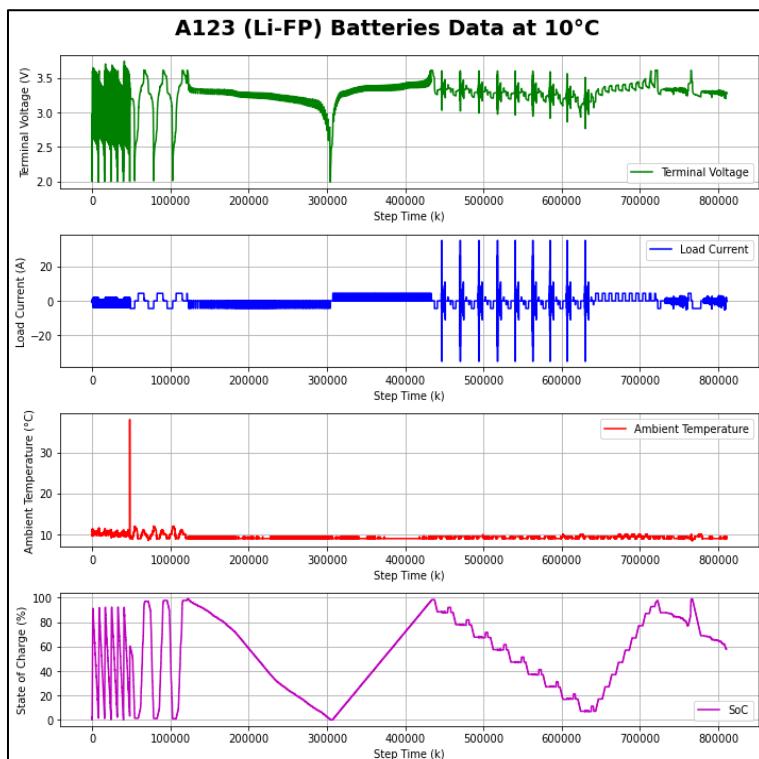
(f).

Graph A-3: Discharge and Charging Curves for the Turnigy Graphene (Li-PO)
 Battery: Terminal Voltage, Load Current, Ambient Temperature,

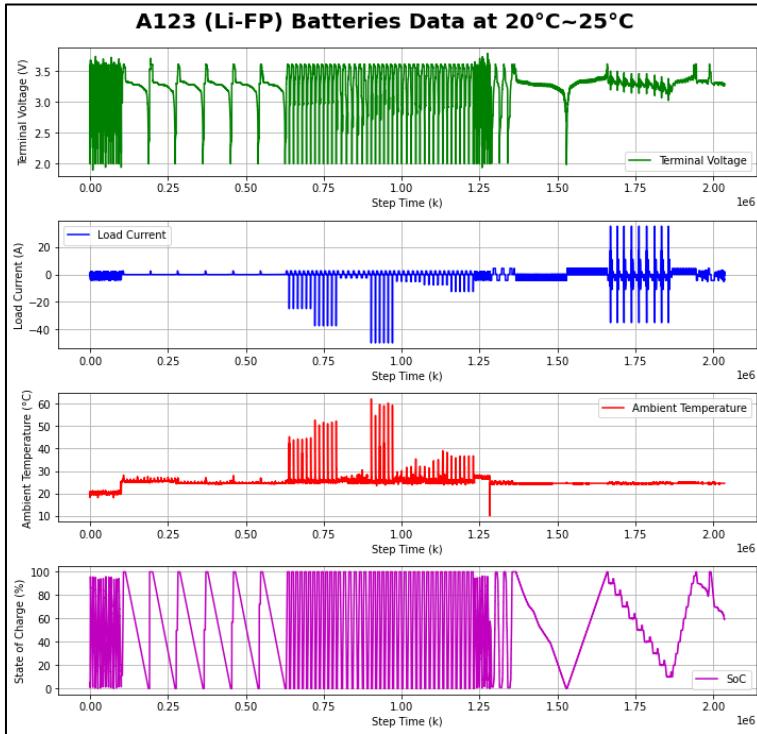
and SoC at (a) -20°C, (b) -10°C, (c) 0°C, (d) 10°C, (e) 25°C, and (f) 40°C (Kollmeyer, and Skells, 2020).



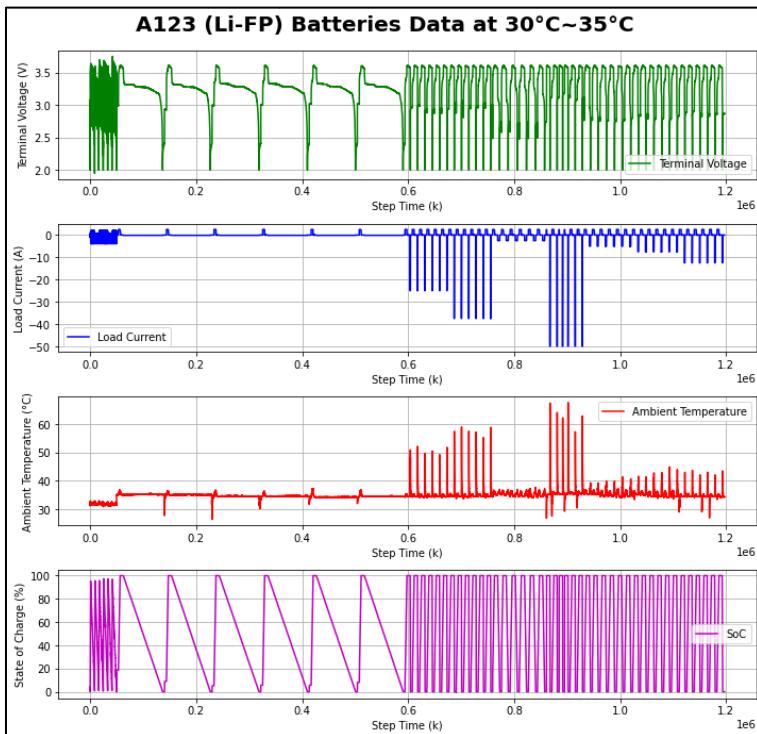
(a).



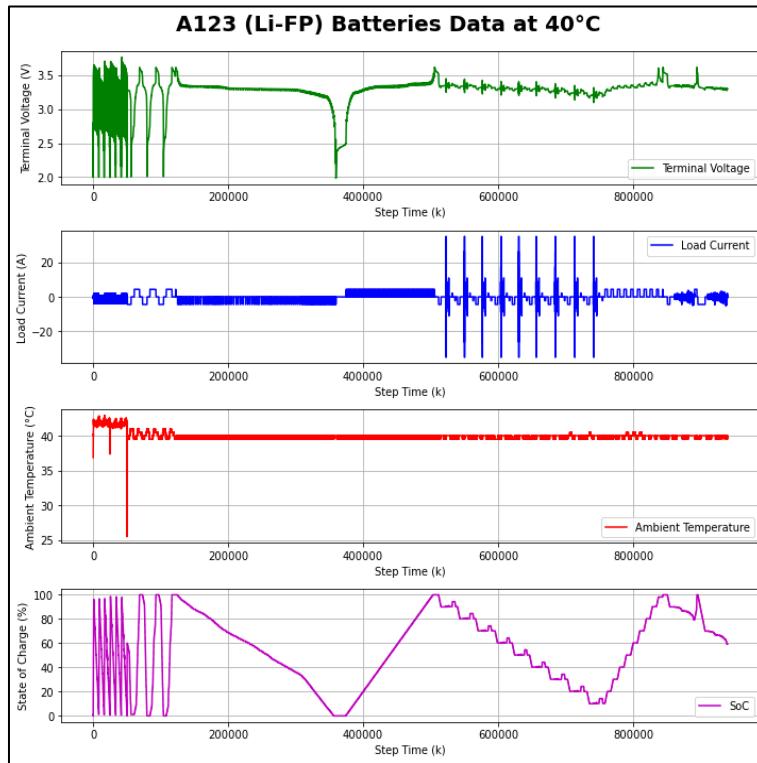
(b).



(c).

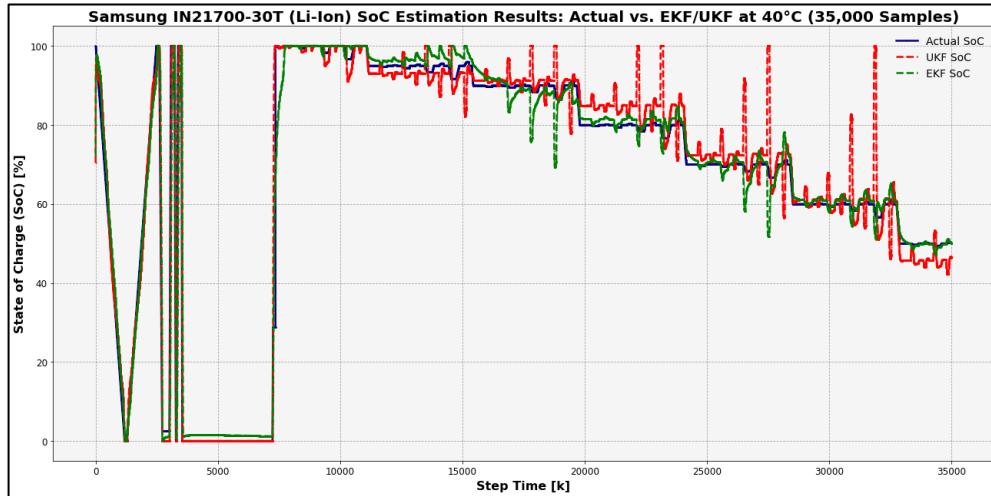


(d).

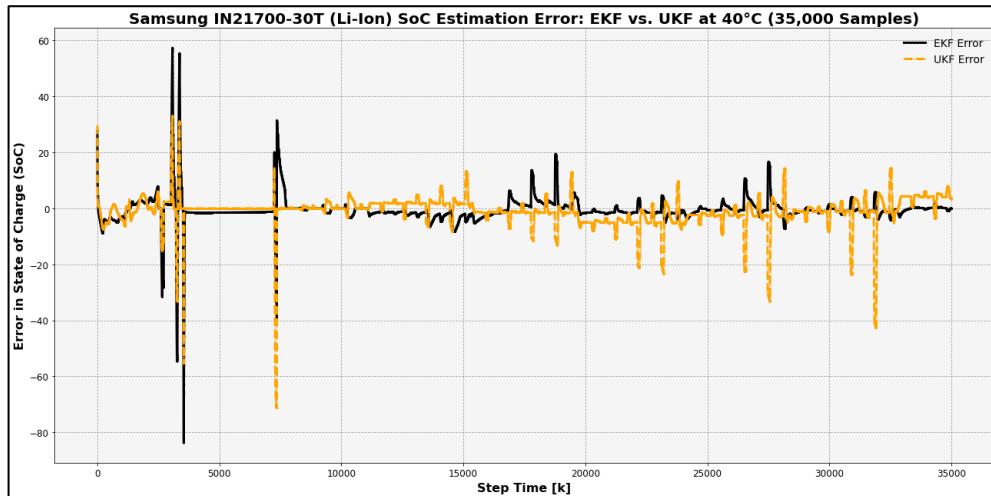


(e).

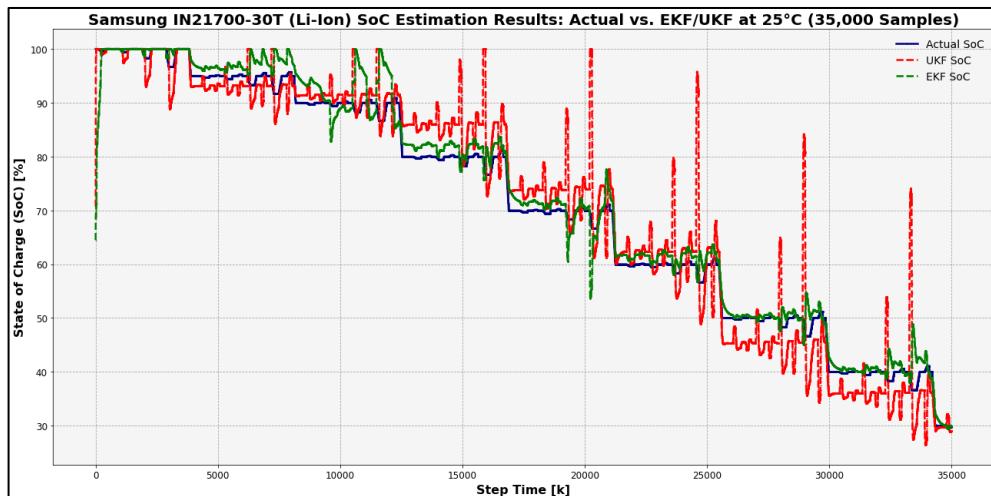
Graph A-4: Discharge and Charging Curves for the A123 (Li-FP) Batteries:
 Terminal Voltage, Load Current, Ambient Temperature, and SoC at
 (a) 0°C~5°C, (b) 10°C, (c) 20°C~25°C, (d) 30°C~35°C, and (e)
 40°C (Xing et al., 2014; Catenaro, and Onori, 2021b; and Redondo-
 Iglesias, and Tian, 2024).



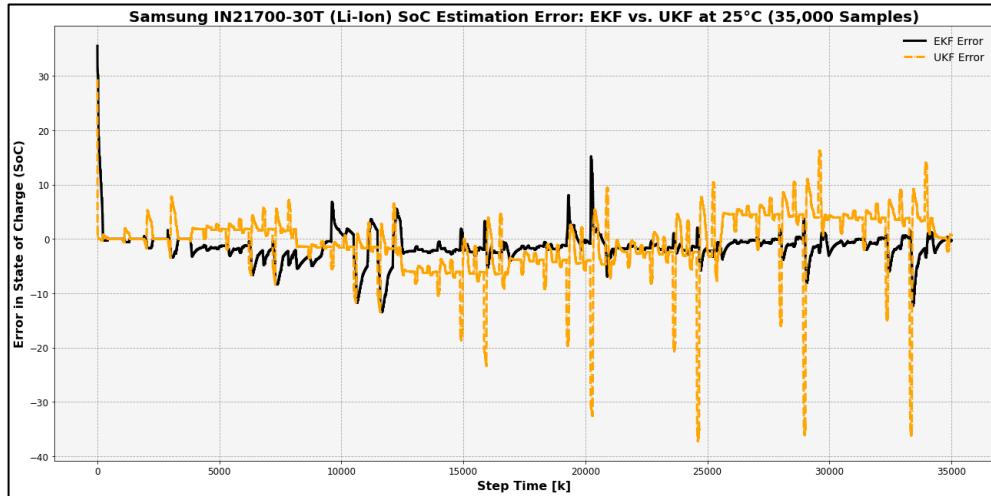
(a).



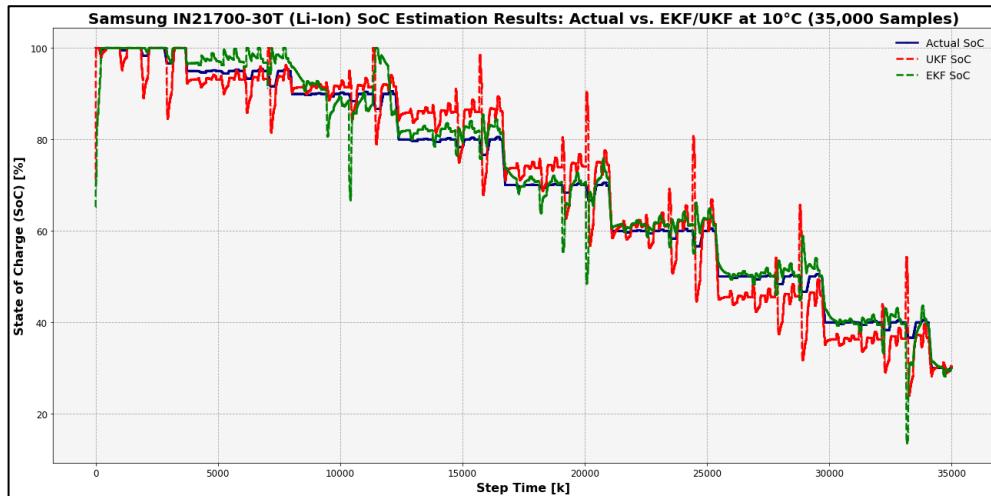
(b).



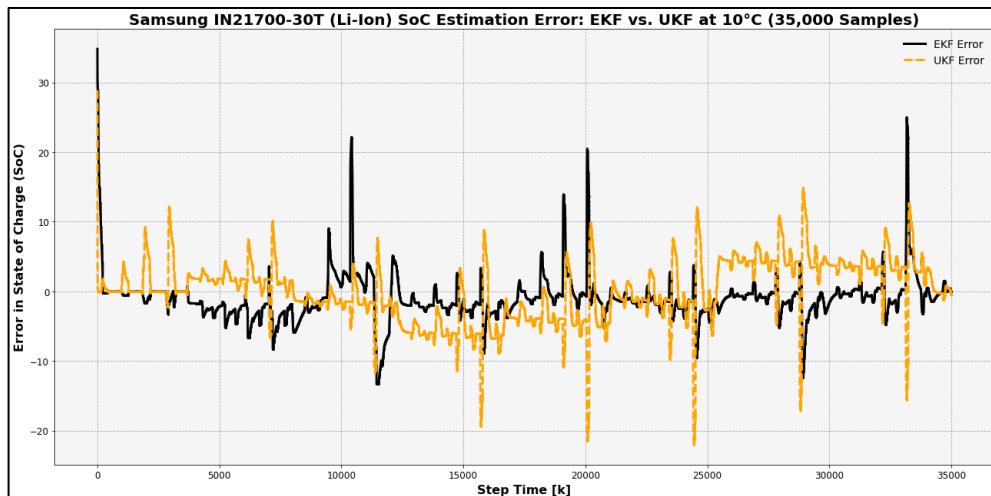
(c).



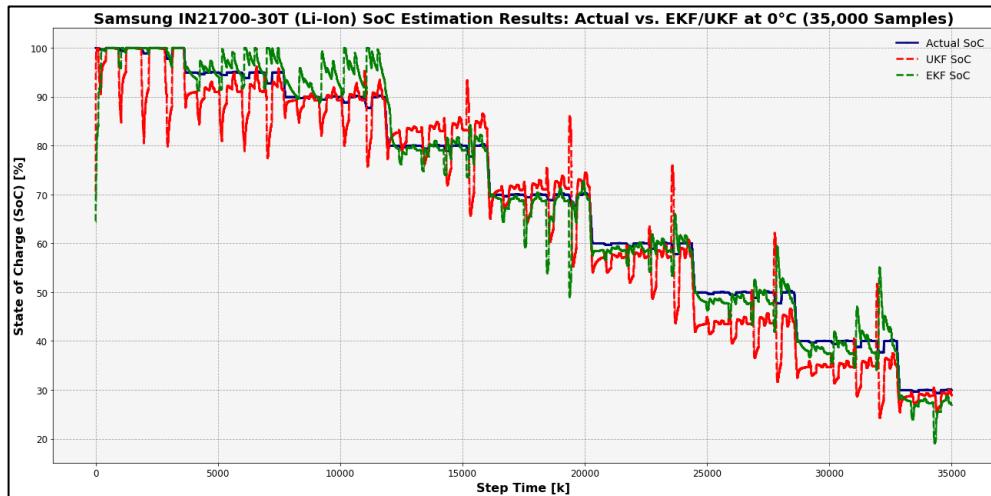
(d).



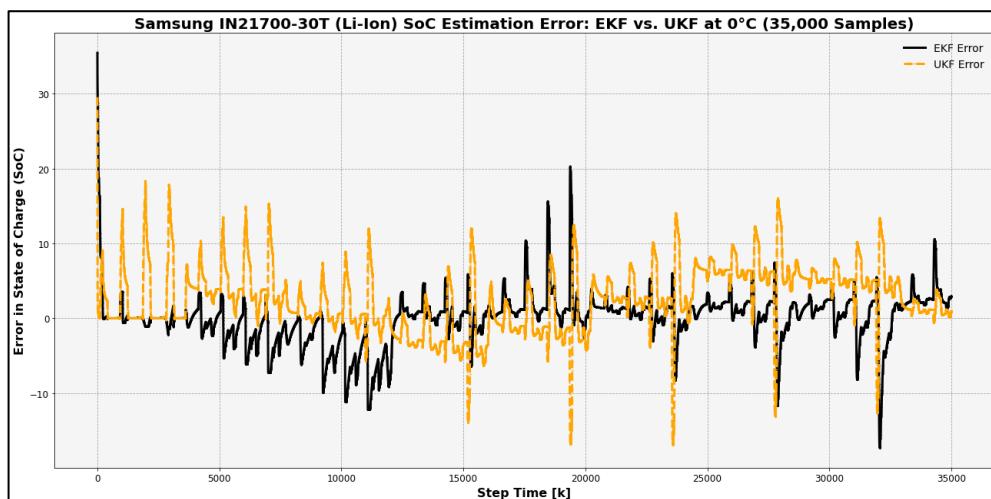
(e).



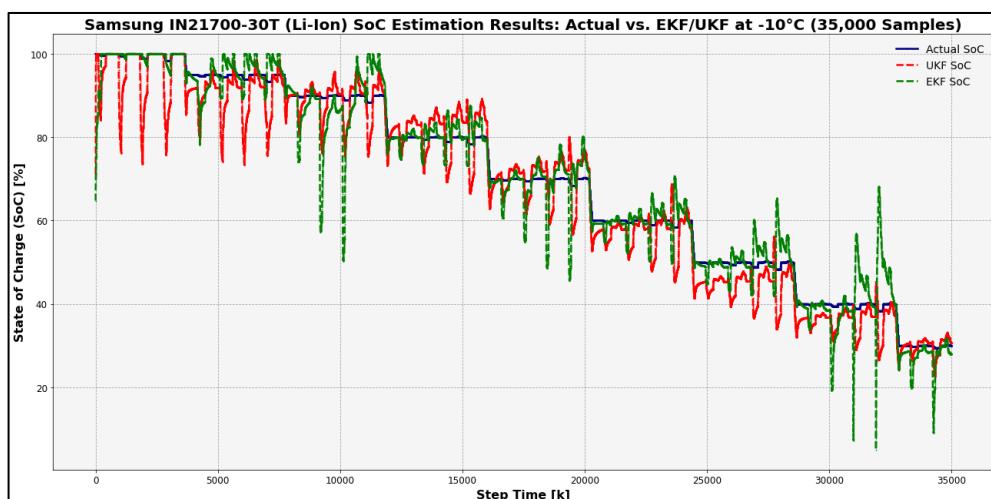
(f).



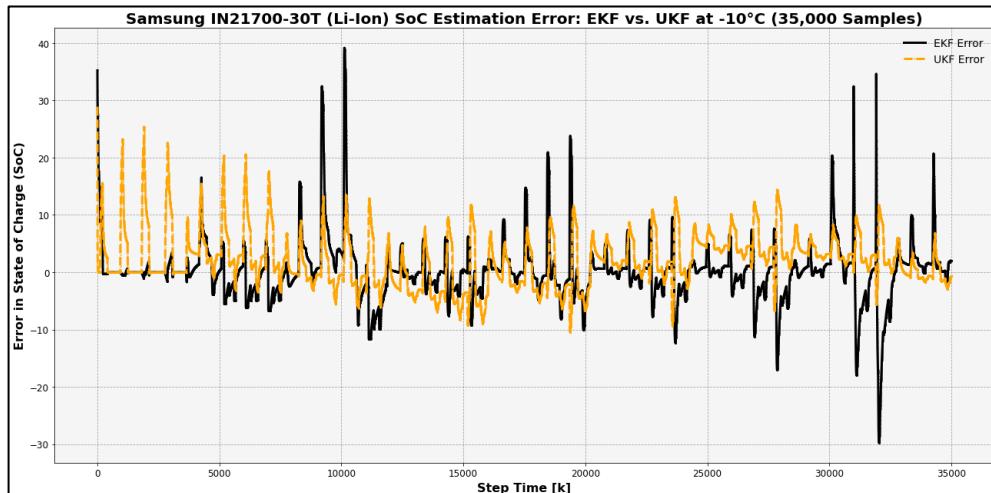
(g).



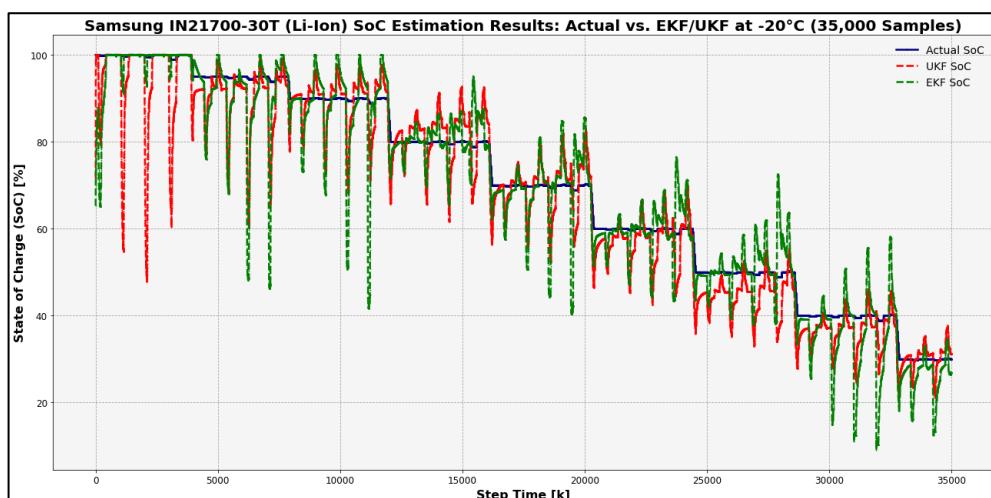
(h).



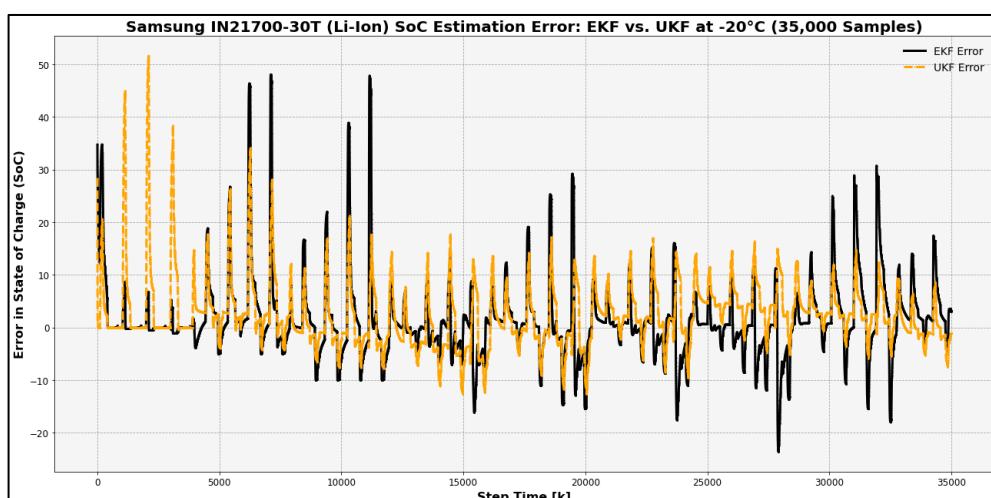
(i).



(j).



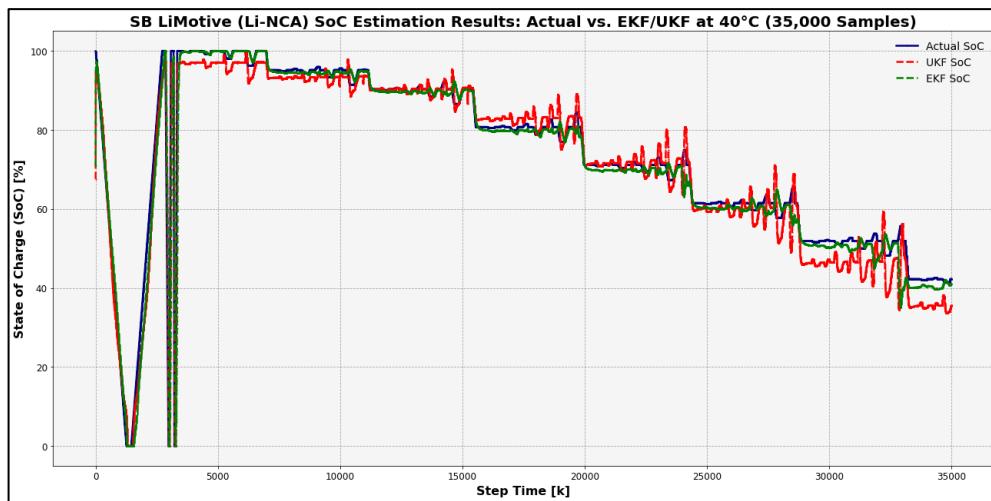
(k).



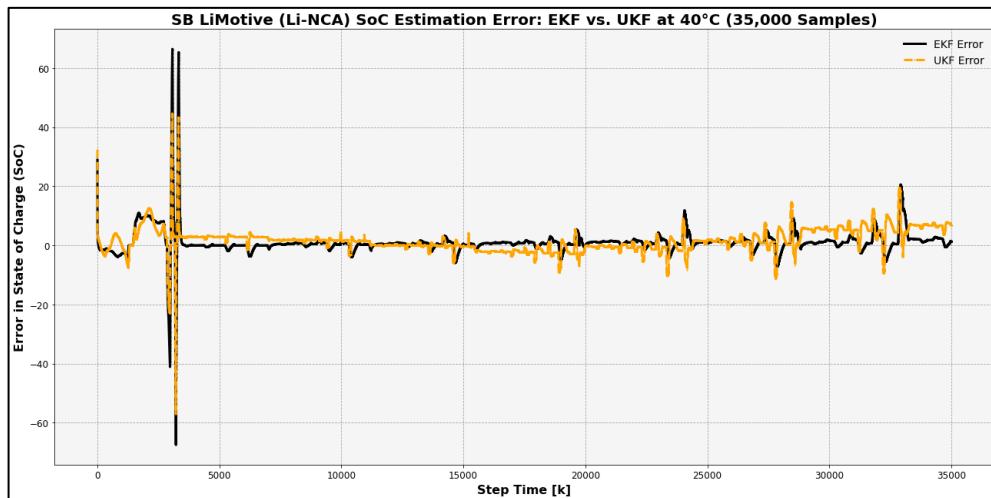
(l).

Graph A-5: Samsung IN21700-30T (Li-Ion) Battery: Comparison of SoC Curves (Actual, EKF, and UKF) and Error Curves (EKF and UKF)

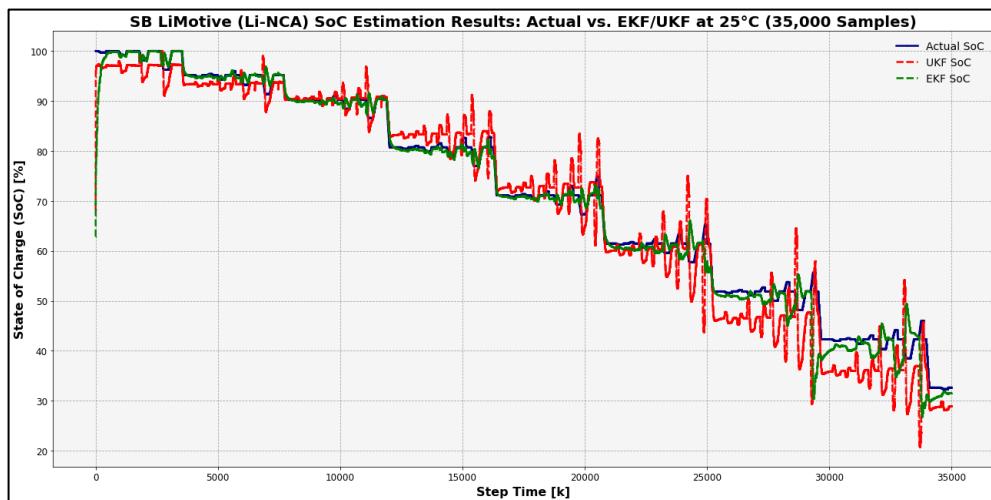
at (a, b) 40°C, (c, d) 25°C, (e, f) 10°C, (g, h) 0°C, (i, j) -10°C, and (k, l) -20°C for the First 35,000 Samples.



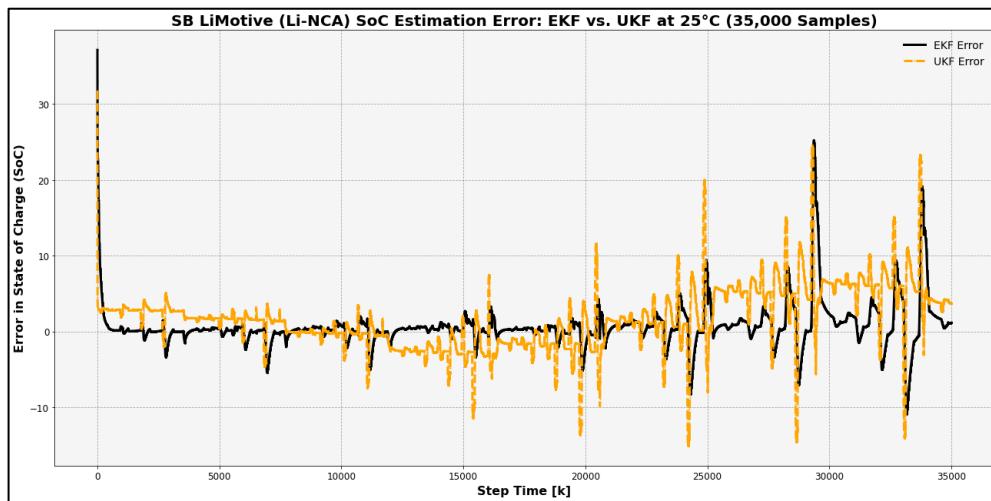
(a).



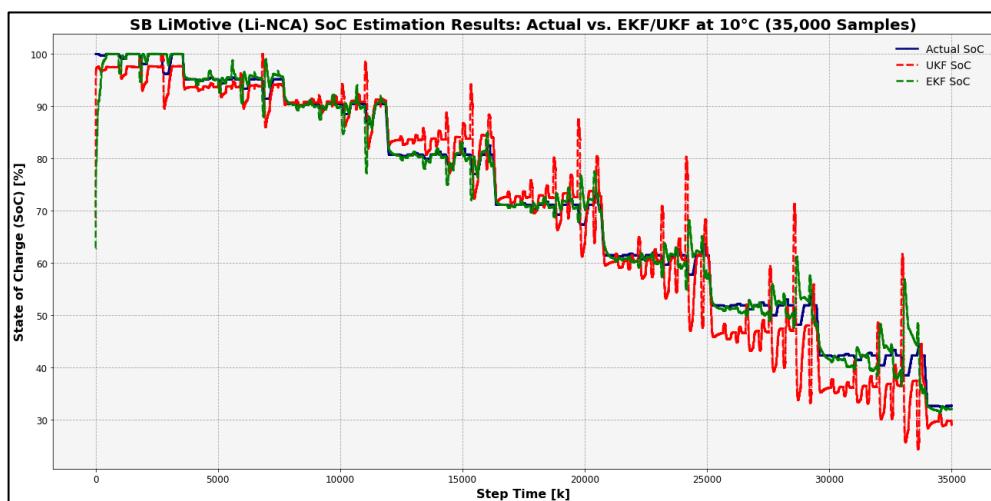
(b).



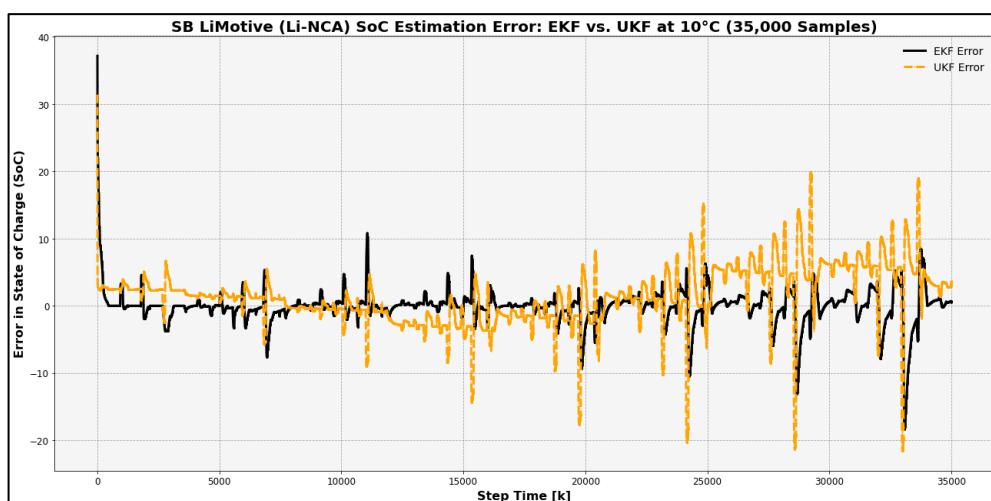
(c).



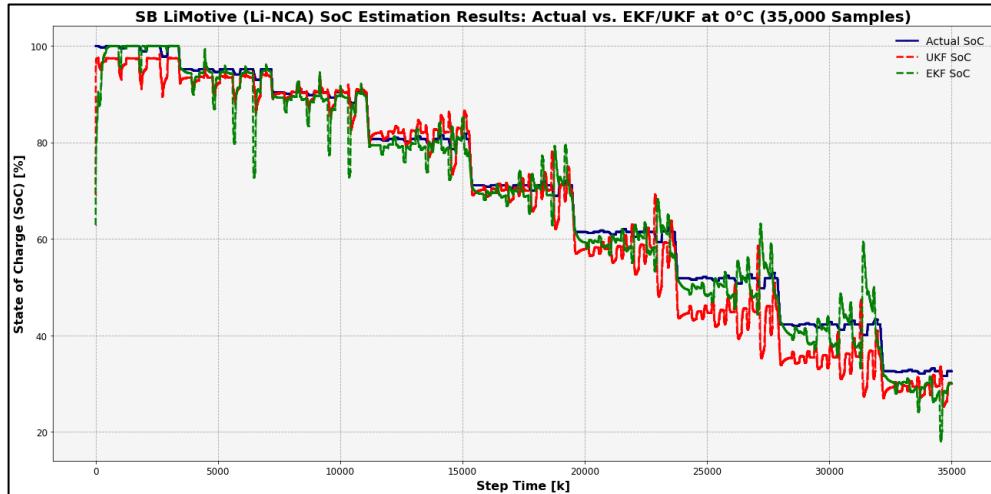
(d).



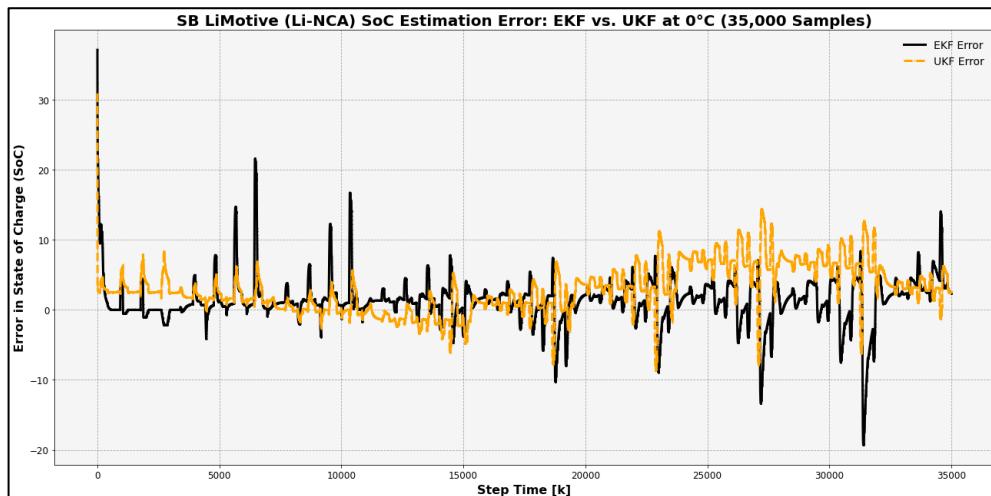
(e).



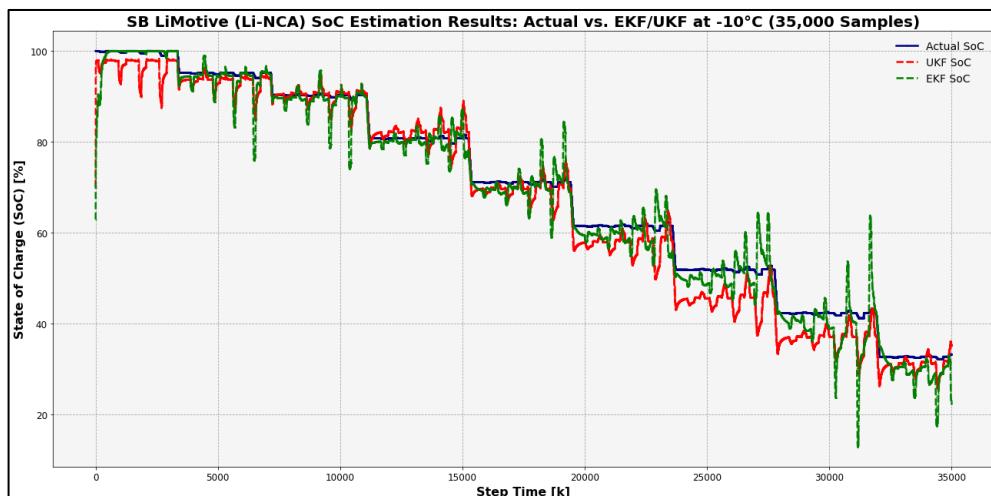
(f).



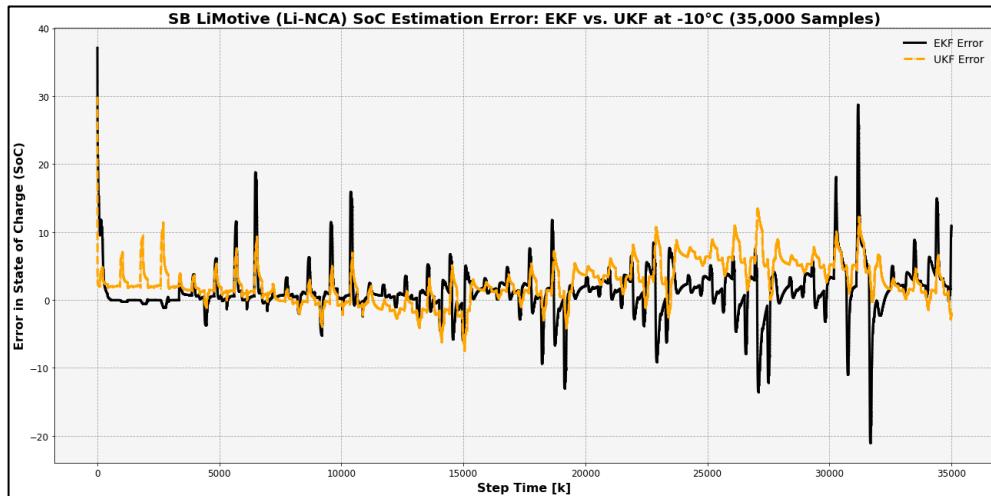
(g).



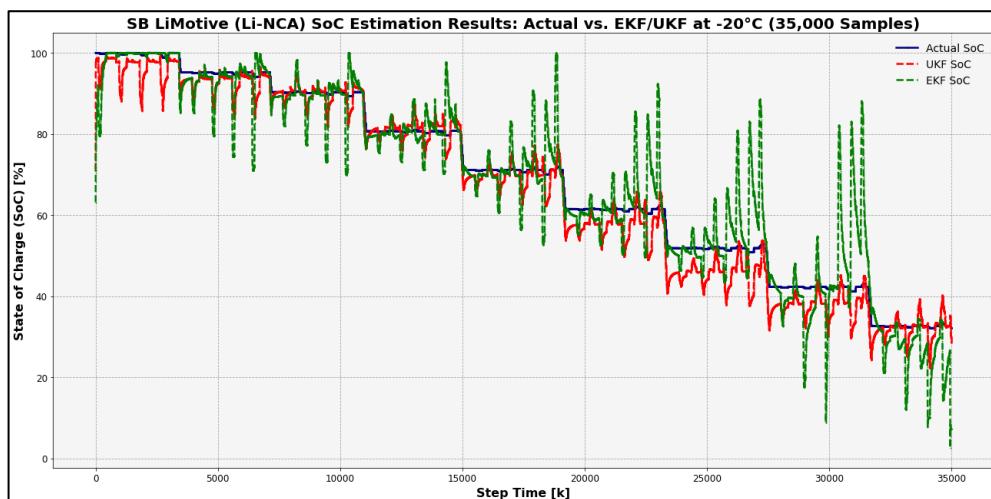
(h).



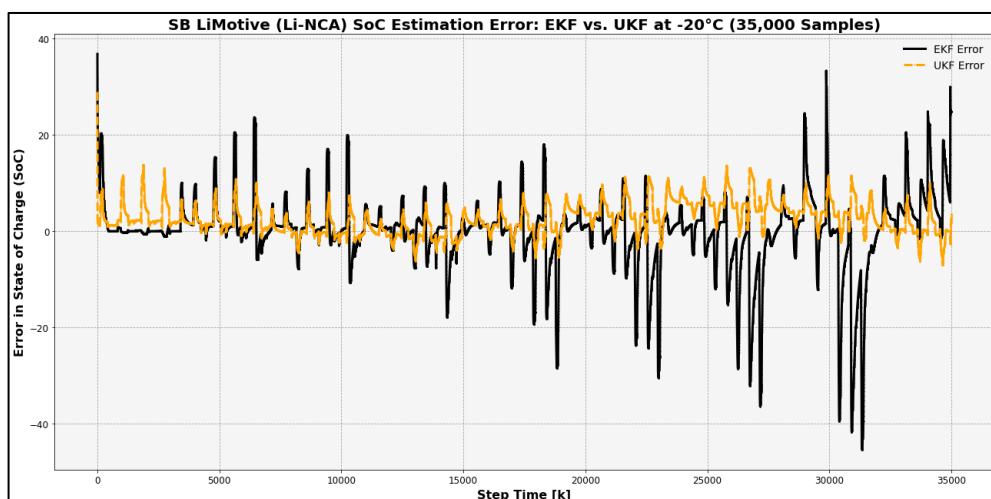
(i).



(j).



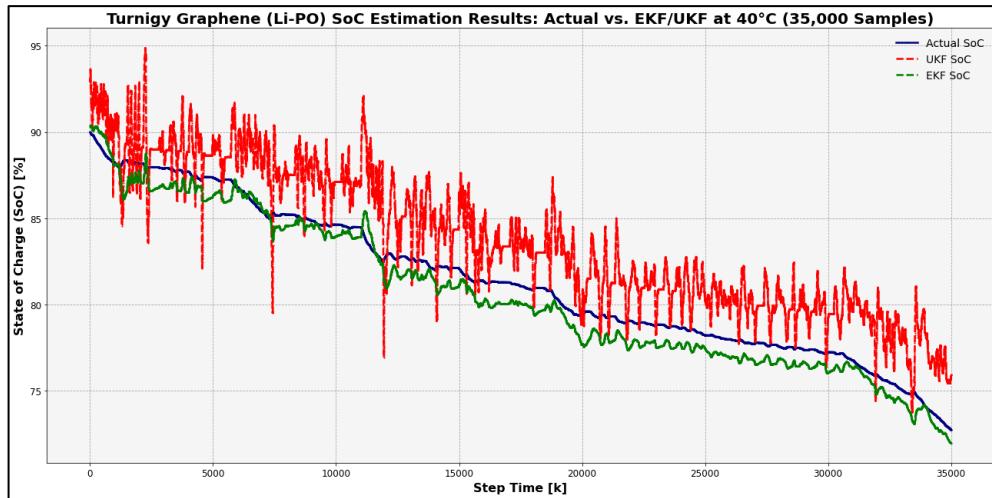
(k).



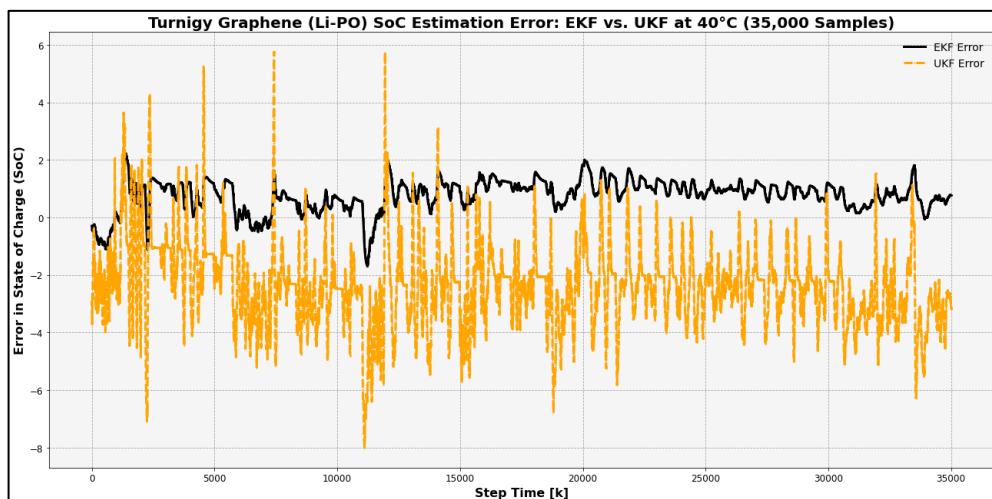
(l).

Graph A-6: SB LiMotive (Li-NCA) Battery: Comparison of SoC Curves (Actual, EKF, and UKF) and Error Curves (EKF and UKF) at (a, b)

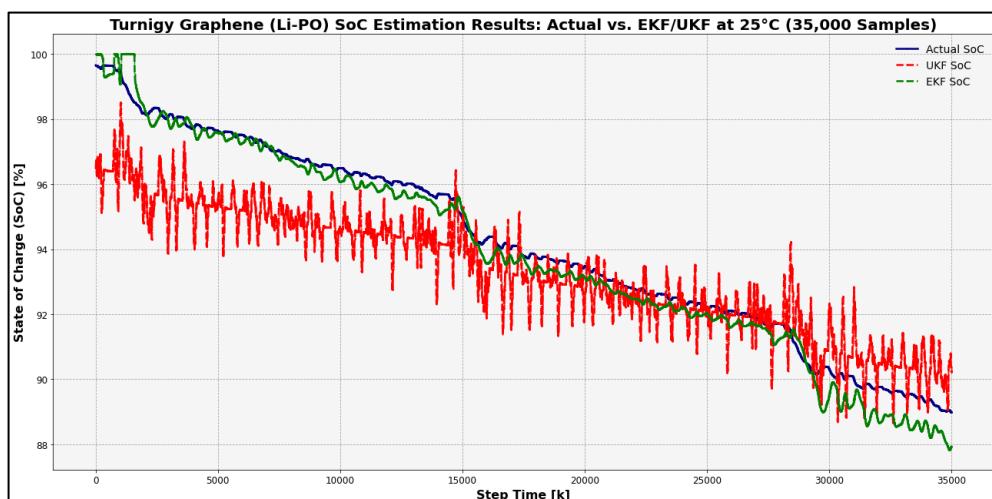
40°C, (c, d) 25°C, (e, f) 10°C, (g, h) 0°C, (i, j) -10°C, and (k, l) -20°C for the First 35,000 Samples.



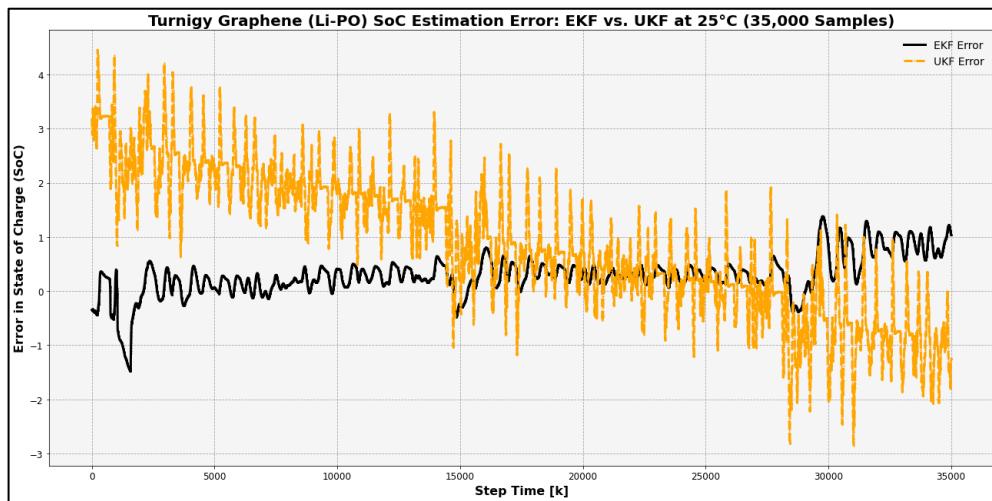
(a).



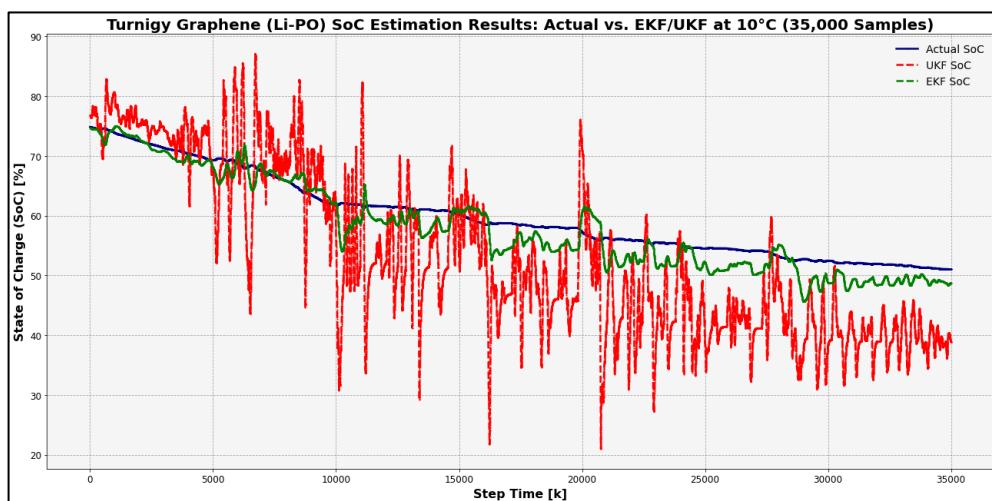
(b).



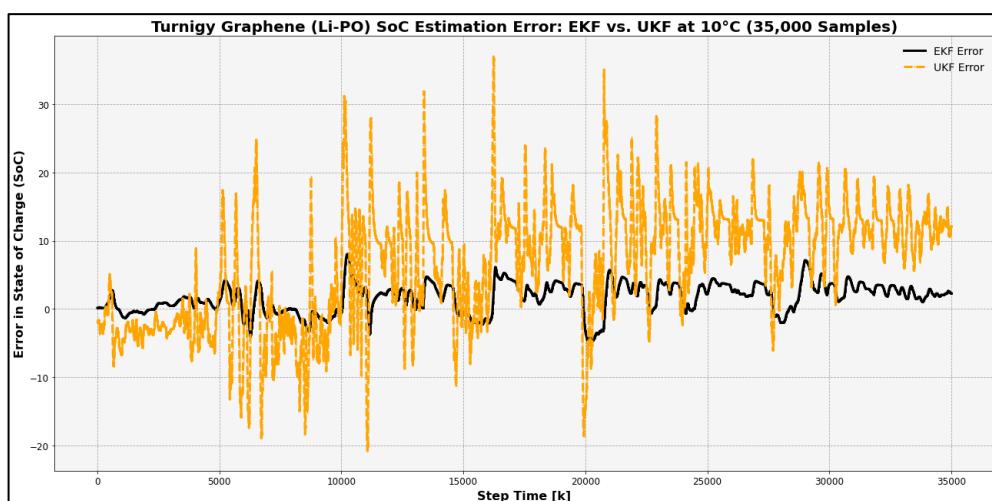
(c).



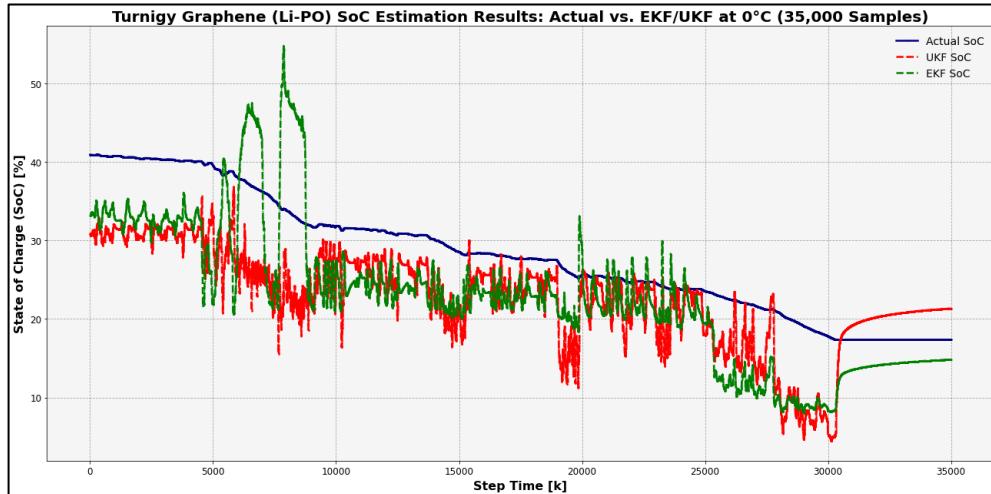
(d).



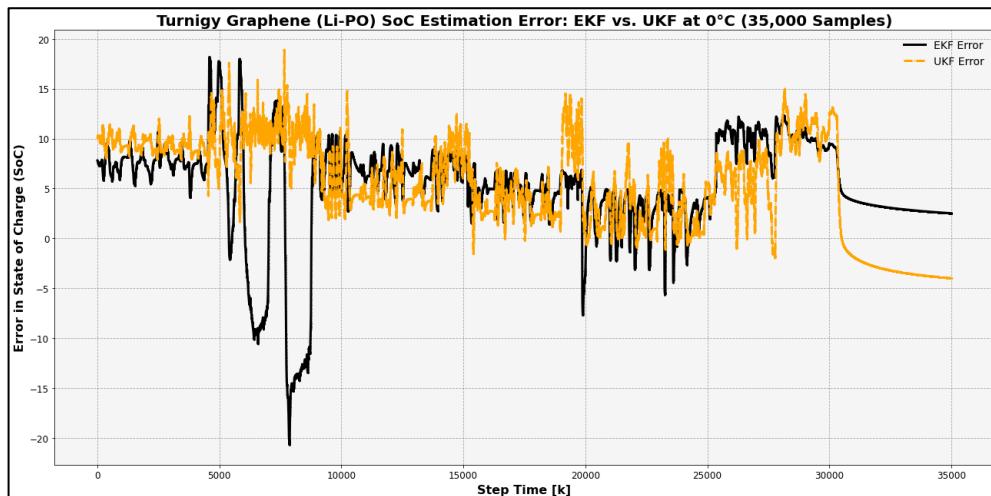
(e).



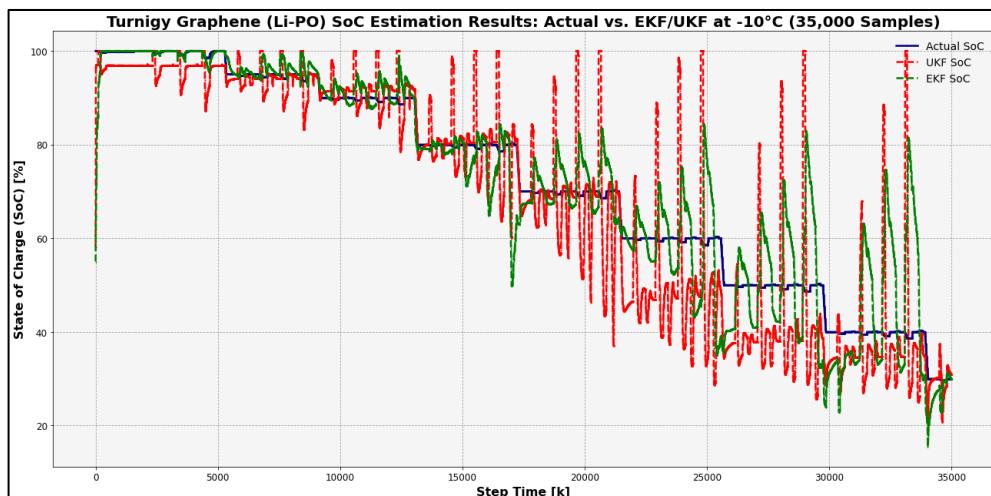
(f).



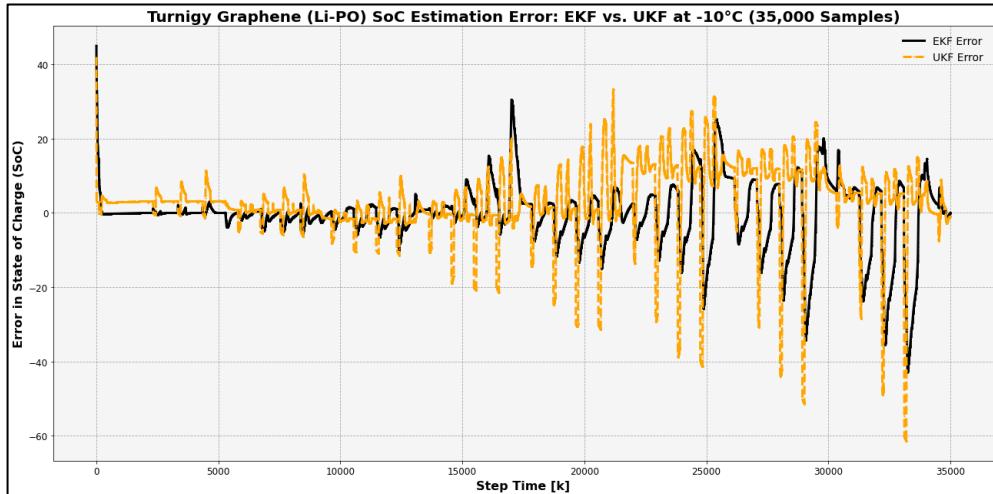
(g).



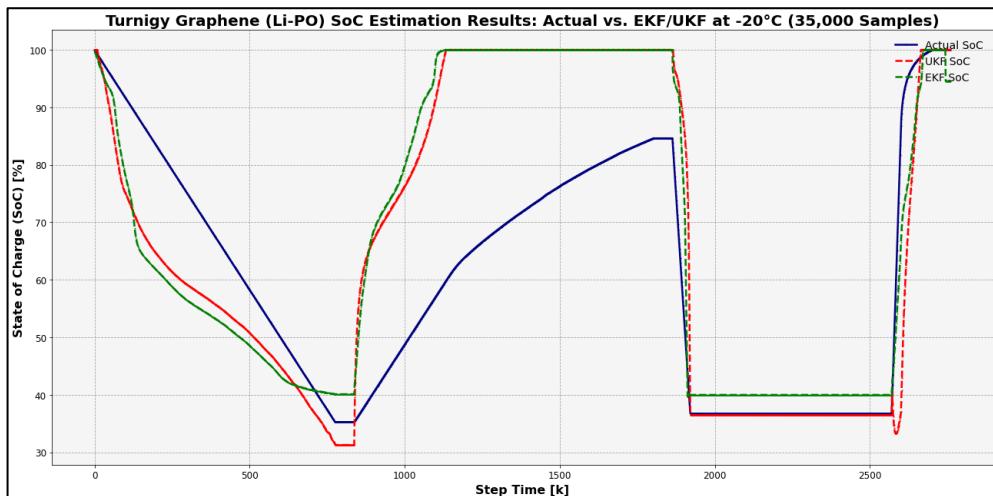
(h).



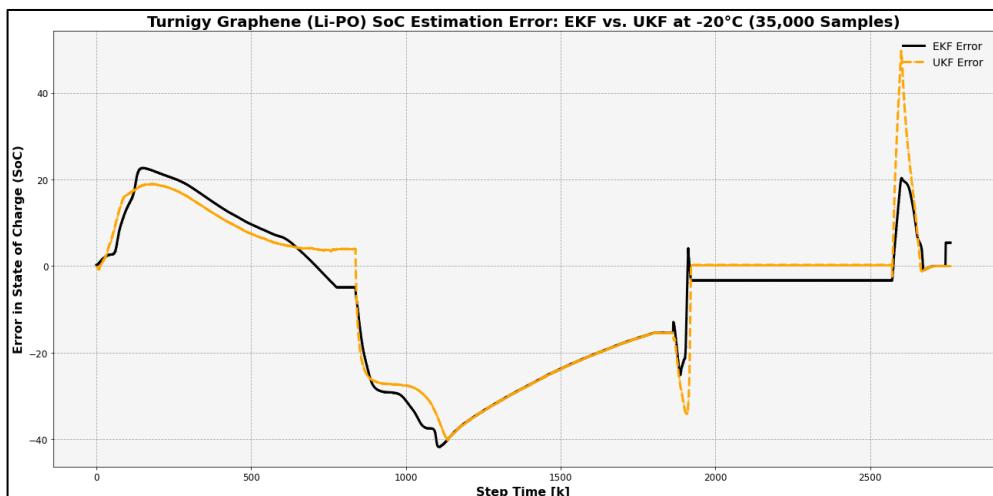
(i).



(j).



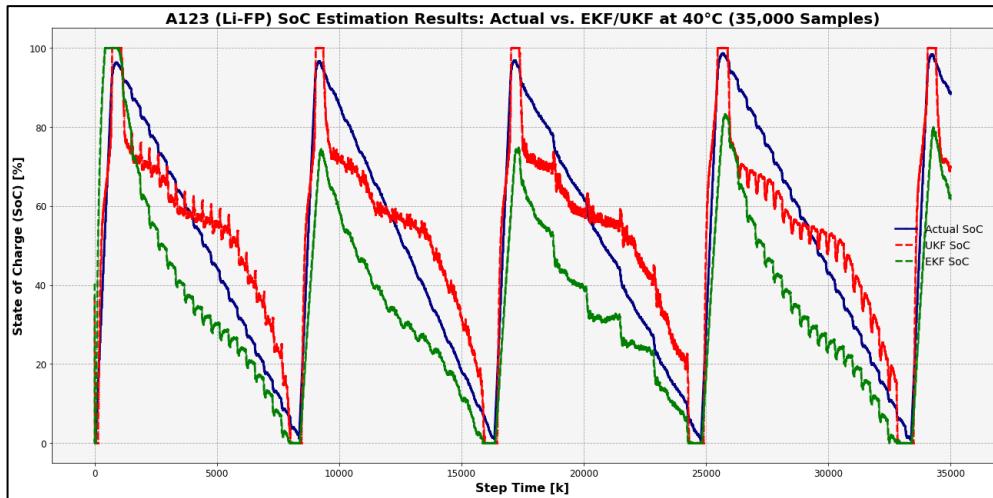
(k).



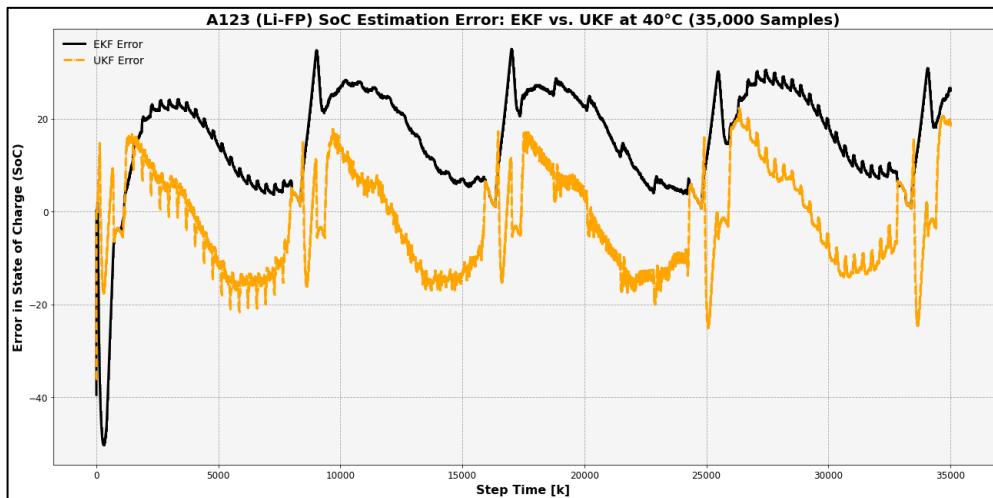
(l).

Graph A-7: Turnigy Graphene (Li-PO) Battery: Comparison of SoC Curves (Actual, EKF, and UKF) and Error Curves (EKF and UKF) at (a, b)

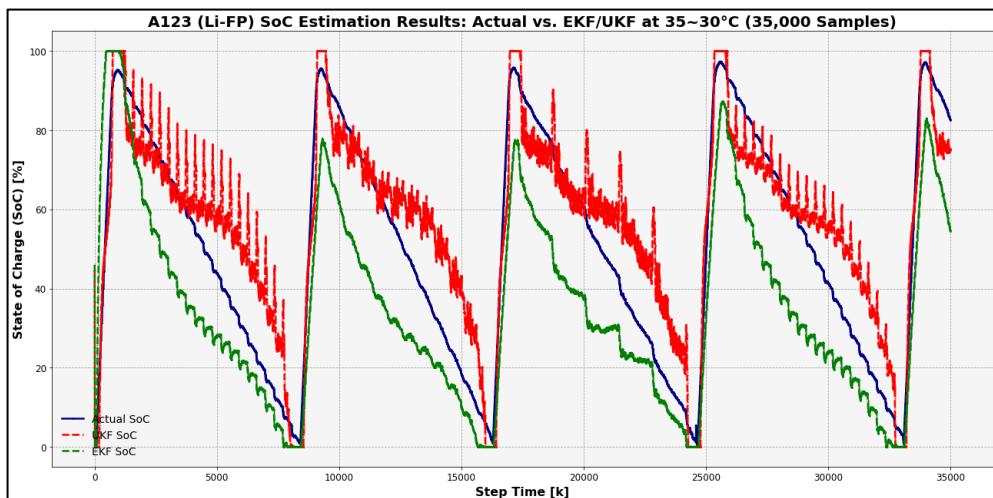
40°C, (c, d) 25°C, (e, f) 10°C, (g, h) 0°C, (i, j) -10°C, and (k, l) -20°C for the First 35,000 Samples.



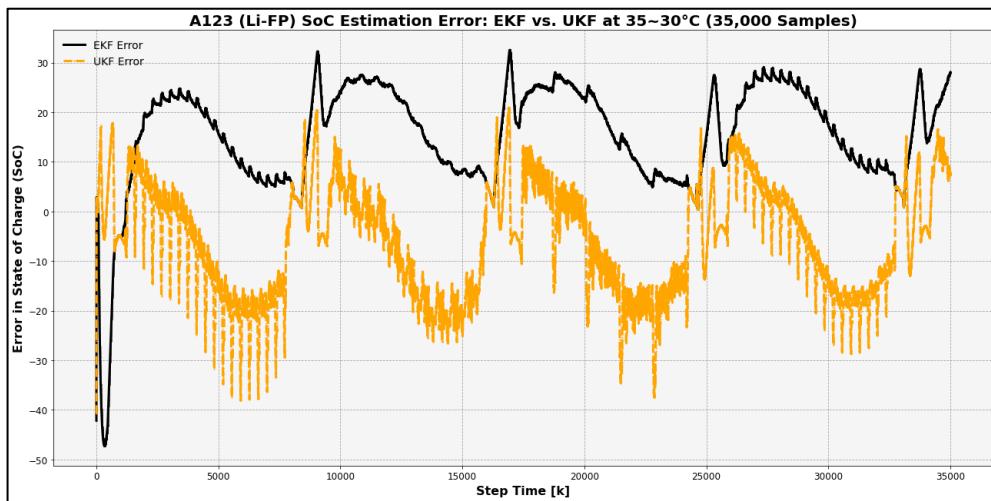
(a).



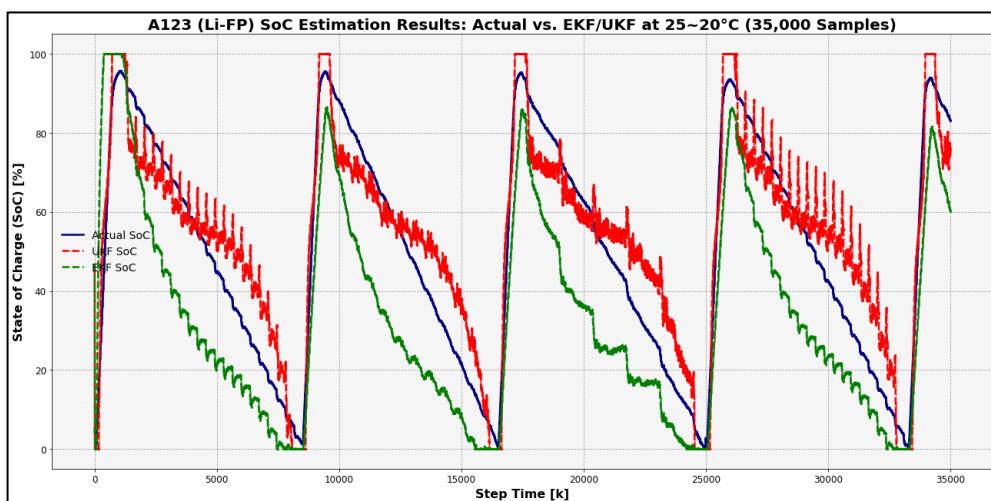
(b).



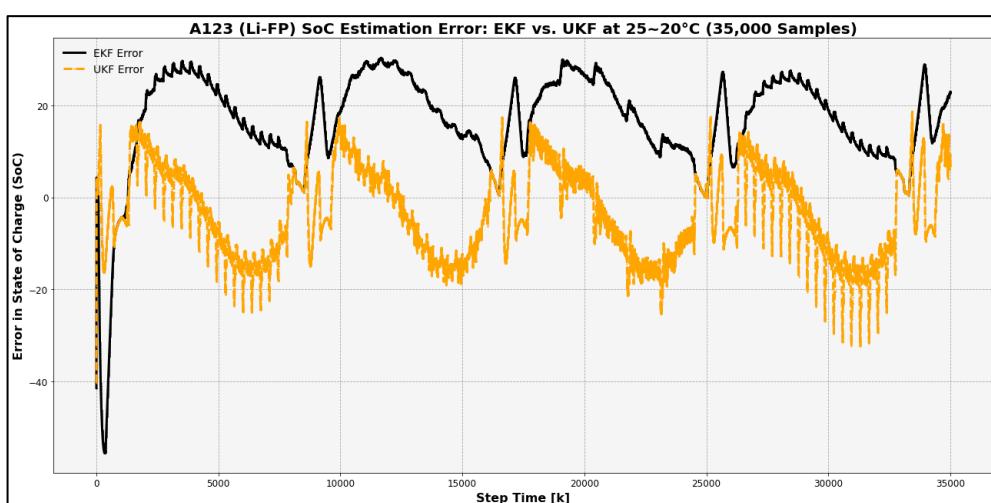
(c).



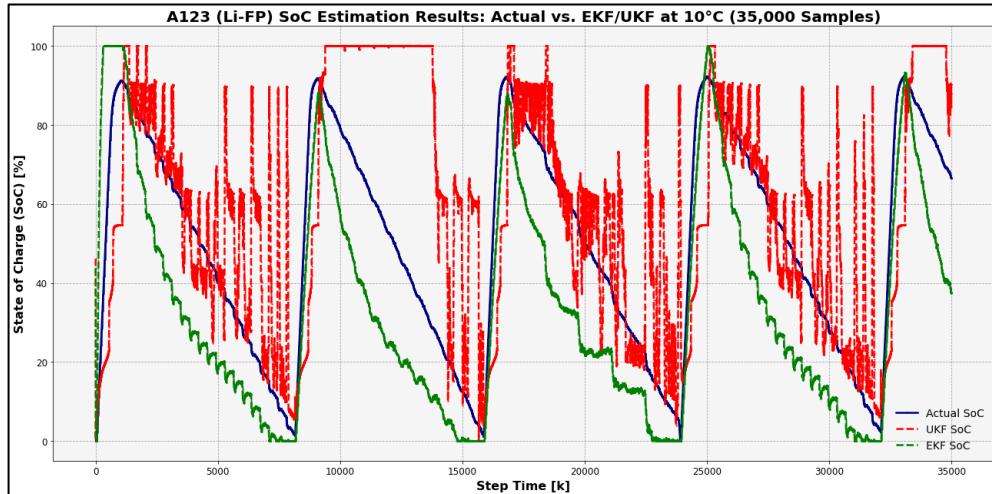
(d).



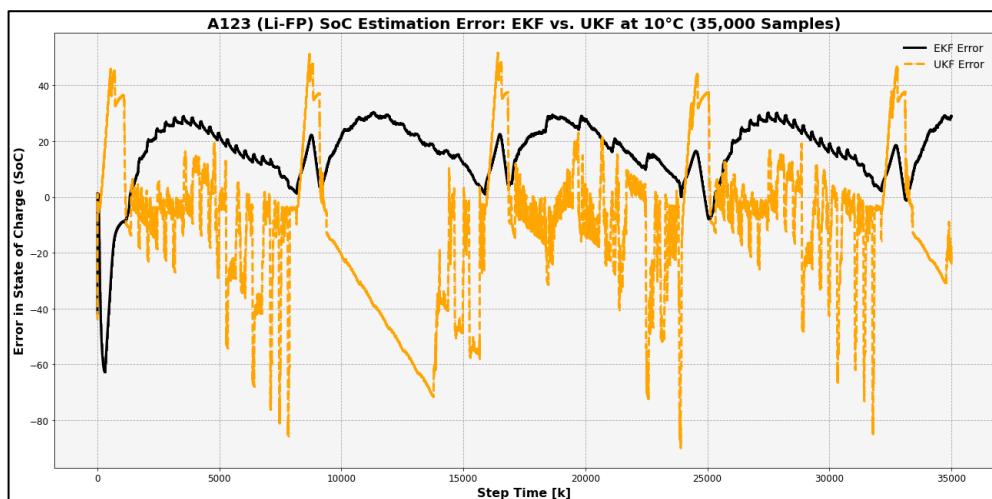
(e).



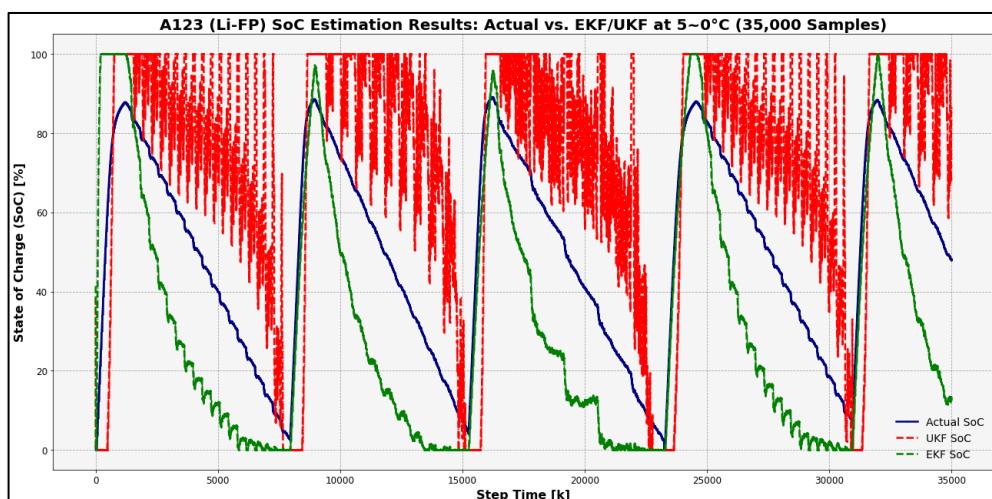
(f).



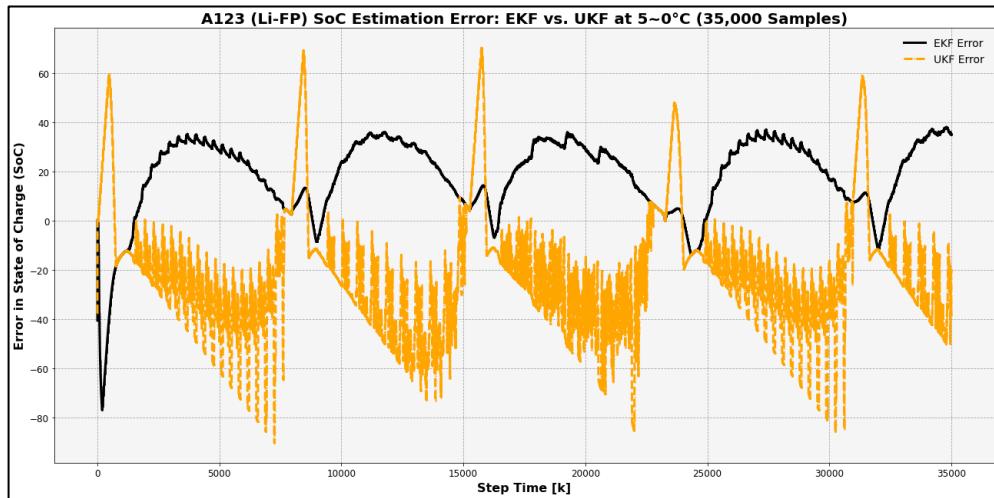
(g).



(h).



(i).



(j).

Graph A-8: A123 (Li-FP) Battery: Comparison of SoC Curves (Actual, EKF, and UKF) and Error Curves (EKF and UKF) at (a, b) 40°C, (c, d) 35–30°C, (e, f) 25–20°C, (g, h) 10°C, and (i, j) 5–0°C for the First 35,000 Samples.

Appendix B: Tables

Table B-1: Website Links for Conventional SoC Estimation Models, Data-Driven SoC Estimation Models Source Code, LIB and LiBat System Datasets.

Items	Website Link
Source Codes	EKF Method: https://github.com/larchuto/Battery-Kalman
	UKF Method: https://github.com/AlterWL/Battery_SOC_Estimation
	ANN Models: https://colab.research.google.com/drive/1KkYprjWm3QDoHpZ2NAOcj5Cyxoc53OcF
Battery Datasets	LG 18650HG2 3000mAh 3.6V Li-NMC: https://doi.org/10.17632/cp3473x7xv.3
	Samsung IN21700-30T 3000mAh 3.6V Li-Ion: https://doi.org/10.17632/9xyvy2njj3.1
	SB LiMotive 5200mAh 3.7V Li-NCA: https://doi.org/10.5683/SP3/LFPKAS
	Turnigy Graphene 5000mAh 3.7V Li-PO: https://doi.org/10.17632/4fx8cjprxm.1
	A123 1100mAh 3.2V Li-FP: https://data.mendeley.com/datasets/kxsbr4x3j2/2
	A123 2000mAh 3.2V Li-FP: https://doi.org/10.1016/j.apenergy.2013.07.008
	A123 4400mAh 3.2V Li-FP: https://doi.org/10.57745/XIDTZH

Appendix C: Open Access to Image Rights

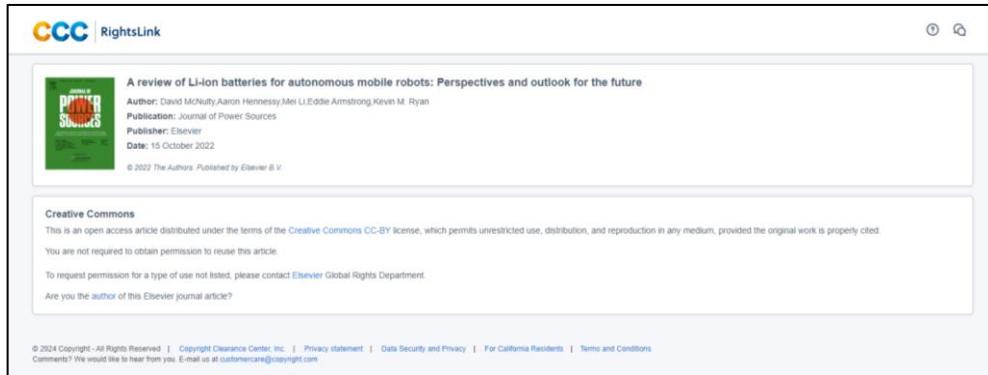


Figure C-1: General Structure of LIBs and LiBats (McNulty et al., 2022) and Characteristics of Various LIB and LiBat Systems (Liu, Placke, and Chau, 2022; McNulty et al., 2022). Reprinted with permission from Copyright 2022 Elsevier.



Figure C-2: Characteristics of Various LIB and LiBat Systems (Liu, Placke, and Chau, 2022; McNulty et al., 2022). Reprinted with permission from Copyright 2022 Elsevier.

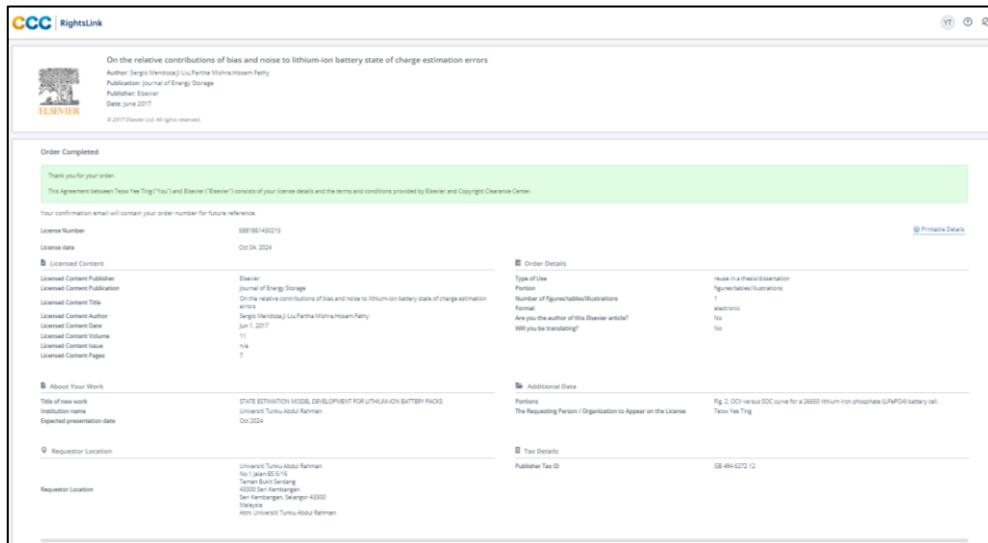


Figure C-3: OCV vs. SoC Curve of Li-FP Battery (Mendoza et al., 2017).

Reprinted with permission from Copyright 2017 Elsevier.

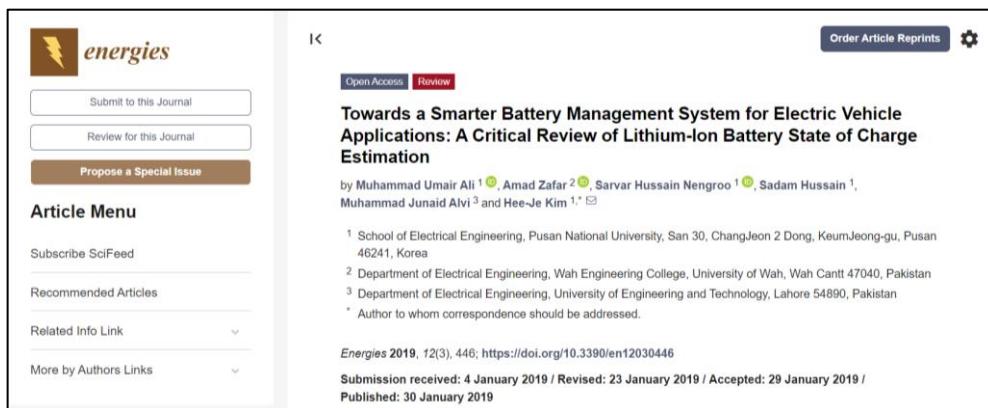


Figure C-4: Schematic of the Model-Based Method (Ali et al., 2019). Reprinted

with permission from Copyright 2019 MDPI.

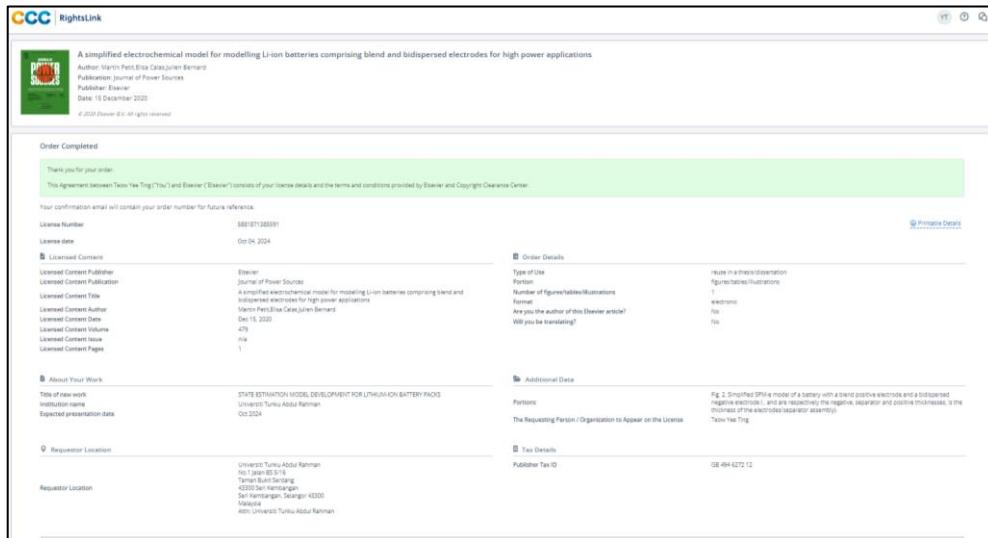


Figure C-5:Schematic of the P2D Model (Petit, Calas, and Bernard, 2020).

Reprinted with permission from Copyright 2020 Elsevier.



Figure C-6:General Function of a BMS and Schematic of the Impedance Circuit (Xiong et al., 2018). Reprinted with permission from Copyright 2018 IEEE.

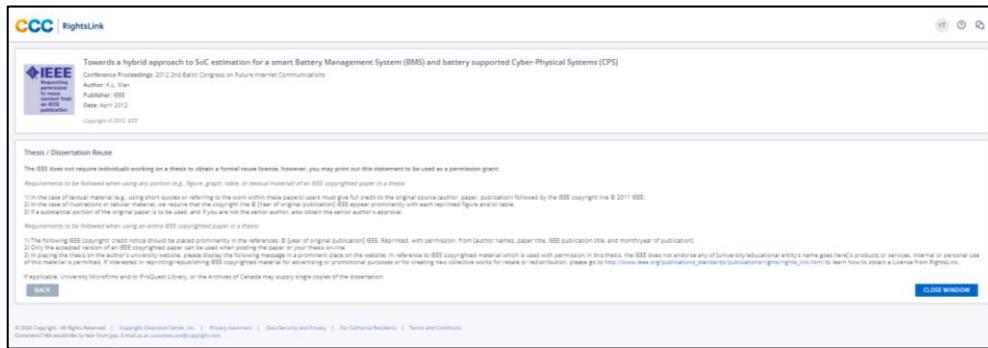


Figure C-7: Terminal Voltage vs. SoC Charging Curves for the (a) LAB and (b) LIB systems (Man et al., 2012). Reprinted with permission from Copyright 2012 IEEE.

WILEY Online Library

Search [\[Search icon\]](#) Login / Register

International Scholarly Research Notices

All articles

Review Article | [Open Access](#) | [\[Creative Commons icon\]](#)

The State of Charge Estimating Methods for Battery: A Review

Wen-Yea Chang

First published: 23 July 2013 | <https://doi.org/10.1155/2013/953792> | Citations: 354

Academic Editor: M. Brüning
Academic Editor: E. Di Nardo

SECTIONS [PDF](#) [TOOLS](#) [SHARE](#)

Abstract

An overview of new and current developments in state of charge (SOC) estimating methods for battery is given where the focus lies upon mathematical principles and practical implementations. As the battery SOC is an important parameter, which reflects the battery performance, so accurate estimation of SOC cannot only protect battery, prevent overcharge or discharge, and improve the battery life, but also let the application make rationally control strategies to achieve the purpose of saving energy. This paper gives a literature survey on the categories and mathematical methods of SOC estimation. Based on the assessment of SOC estimation methods, the future development direction of SOC estimation is proposed.

[Figures](#) [References](#) [Related](#) [Information](#)

Recommended

[A Stable Quasi-Solid-State Sodium-Sulfur Battery](#)
Dong Zhou, Yi Chen, Baohua Li, Hongbo Fan, Faliang Cheng, Devaraj Shanmukaraj, Teofilo Rojo, Michel Armand, Guoxiu Wang

[Angewandte Chemie International Edition](#)

[Comparison of Nonlinear Filtering Methods for Estimating the State of Charge of Li_xTi₅O₁₂ Lithium-Ion Battery](#)
Jianping Gao, Hongwen He

[Mathematical Problems in Engineering](#)

Figure C-8: Structure of ANN for SoC Estimation (Chang, 2013). Reprinted with permission from Copyright 2013 Wiley.

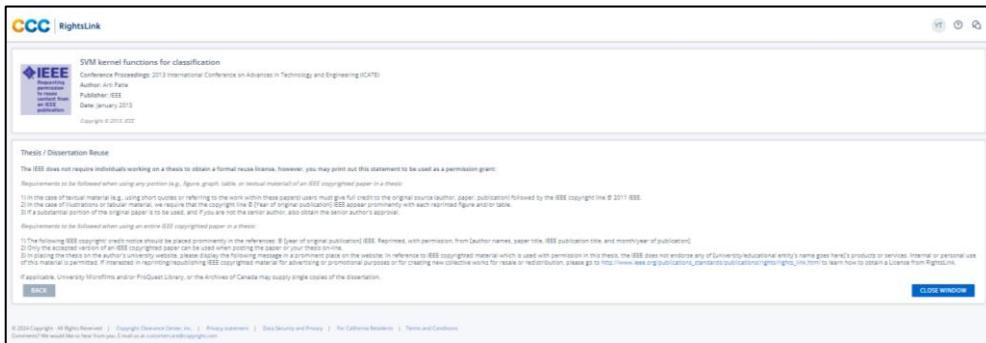


Figure C-9: General Portrait of the SVM (Patle, and Chouhan, 2013). Reprinted with permission from Copyright 2013 IEEE.



Figure C-10: General Portrait of the Decision Boundaries (Soft Margins) and Hyperplane (Pannakkong et al., 2022). Reprinted with permission from Copyright 2022 Wiley.

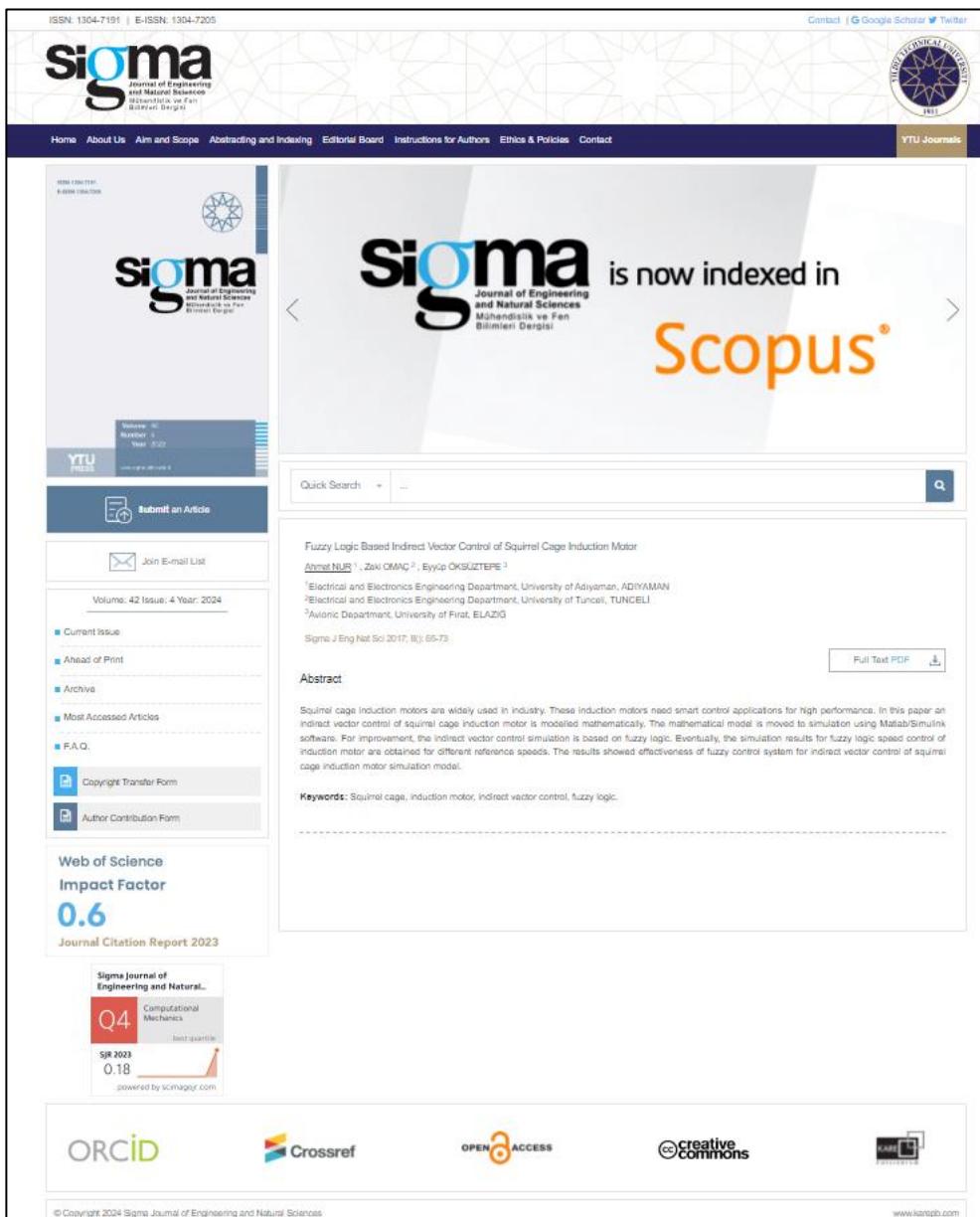


Figure C-11: General Structure of FLC (Nur, Omaç, and Öksüztepe, 2017b).
Reprinted with permission from Copyright 2017 Sigma.

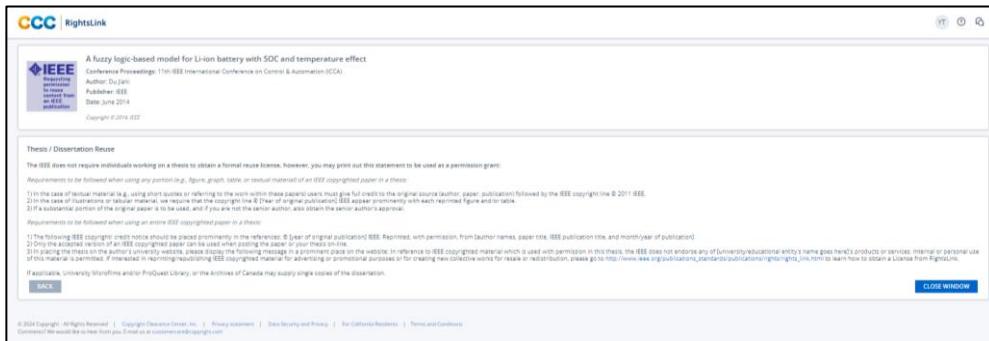


Figure C-12: 2D MF of FLC for SoC Estimation (Du et al., 2014). Reprinted with permission from Copyright 2014 IEEE.

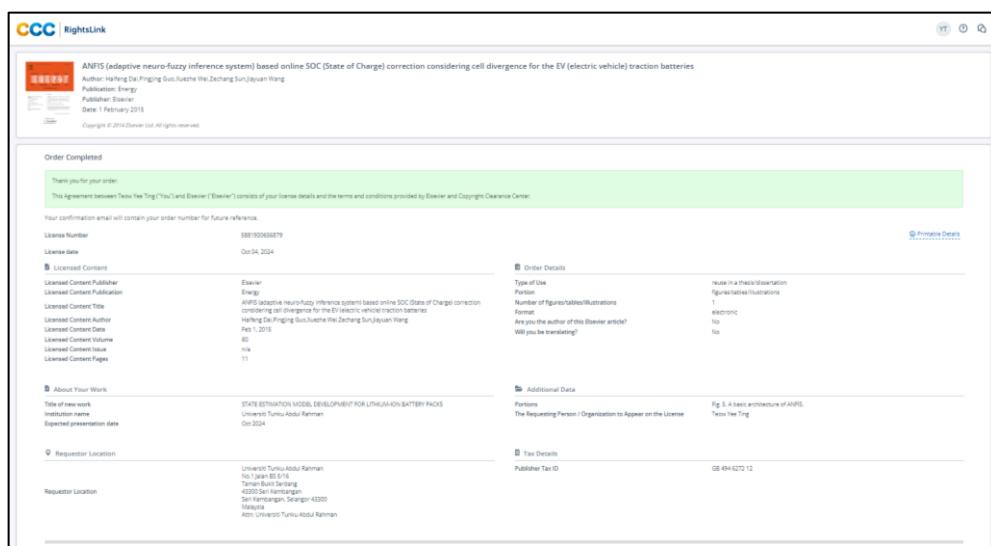


Figure C-13: General Structure of ANFIS (Dai et al., 2015). Reprinted with permission from Copyright 2015 Elsevier.