Robust State of Health Estimation for Lithium-Ion Batteries Using Machines Learning

Sui, Xin

DOI (link to publication from Publisher):
10.54337/aau466409590

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

ROBUST STATE OF HEALTH ESTIMATION FOR LITHIUM-ION BATTERIES USING MACHINE LEARNING

BY
XIN SUI

DISSERTATION SUBMITTED 2021

AALBORG UNIVERSITY
DENMARK

ROBUST STATE OF HEALTH ESTIMATION FOR LITHIUM-ION BATTERIES USING MACHINE LEARNING

by

Xin Sui

Department of Energy Technology

Aalborg University, Denmark

AALBORG UNIVERSITY
DENMARK

Dissertation submitted December, 2021

Dissertation submitted: December, 2021

PhD supervisor: Prof. Remus Teodorescu,
Aalborg University

Assistant PhD supervisor: Assoc. Prof. Daniel-Ioan Stroe,
Aalborg University

PhD committee: Professor Huai Wang (chair)
Aalborg University, Denmark

Professor Xiansong Hu
Chongqing University, China

Professor Ali Emadi
McMaster University, Canada

PhD Series: Faculty of Engineering and Science, Aalborg University

Department: Department of Energy Technology

ISSN (online): 2446-1636
ISBN (online): 978-87-7573-972-1

Published by:
Aalborg University Press
Kroghstræde 3
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2022

CURRICULUM VITAE

Xin Sui received the B.Eng. degree in Electrical Engineering from Northeast Electric Power University, Jilin, China, in 2015, and the M.Sc. degree in Electrical Engineering from Institute of Electrical Engineering, Chinese Academy of Sciences, Beijing, China, in 2018. She is currently pursuing the Ph.D. degree in Electrical Engineering at Aalborg University, Aalborg, Denmark. Her research interests include battery state of health estimation, lifetime extension, and machine learning algorithms.

ENGLISH SUMMARY

Machine learning (ML) technologies have gained considerable attention for state of health (SOH) estimation of Lithium-ion (Li-ion) batteries due to their advantages in learning the behavior of non-linear systems. ML methods do not require the battery physical modeling processes but rather map external characteristics of the battery to the loss in capacity. However, from the application perspective, there are still some challenges that need to be addressed, including the impact of data noise and data size on the estimation performance, the failure of features under variable operation conditions, the dependency on big data, and the difficulty of implementing complex algorithms in low-cost microprocessors. To cope with these issues, a systematic ML-based Li-ion battery SOH estimation framework is developed in this Ph.D. project, which has strong robustness to data size, data noise, and degradation conditions.

Because batteries are complex electrochemical systems, their aging process is closely related to the operating conditions, and the SOH feature will be invalid under different conditions. Fuzzy entropy (FE) of voltage, from a short-term pulse test, is proposed as a novel SOH feature. FE-based method is flexible in parameter selection, robust to noise and test conditions, and only requires less training data. Furthermore, the interaction between the test conditions, entropy features, and estimation performance is studied. The results can be used to select the appropriate voltage datasets and enhance the performance of entropy-based SOH estimation. For further improvement of entropy-based SOH estimation, various noise suppression methods are used before and after feature extraction. The experiments provide evidence that the smoothing step is effective in improving the estimation accuracy and simulation speed of the ML model.

Another solution to avoid the failure of manual-extracted features is to use neural networks especially deep learning methods. This method allows high-dimensional input and automatically-extracted features through hidden layers, but at the same time, this method also relies on relatively large data. Therefore, a bagging-based ensemble method is developed, enabling the model trained on limited data to achieve higher accuracy and better generalization performance. Finally, the effectiveness of the proposed methods in this project is verified by experiments including the cyclic aging test and calendar aging test under different temperature conditions.

DANSK RESUME

Machine learning (ML) teknologier har i stigende grad vundet opmærksomhed når state of health (SOH) af litium-ion (Li-ion) batterier skal estimeres, på grund af deres evne til at lære adfærden af ikke-lineære systemer. ML modellerer batteriets nedbrydning ved at kortlægge eksterne features mod tabet i kapacitet, og dermed undgå hele batterimodelleringsprocessen. Set fra et anvendelsesperspektiv er der dog stadig visse udfordringer, der skal løses, herunder betydning af støj og datastørrelse på estimationsevnen, fejl i features under variable driftsbetingelser, afhængigheden af store mængder af data og vanskeligheder i implementering af komplekse algoritmer i billige mikroprocessorer. For at håndtere disse spørgsmål udvikles, i dette Ph.D. projekt, et systematisk ML baseret framework til estimering af SOH i Li-ion-batterier, som har en stærk robusthed over for datastørrelse, støj og driftsbetingelser.

Da batterier er komplekse elektrokemiske systemer, er deres aldringsproces tæt forbundet med driftsbetingelserne, og features udviklet under et sæt af betingelser vil ikke nødvendigvis være gyldige under andre forhold. Fuzzy entropy (FE) af spændingen for kortvarige pulstest, foreslåes som en ny feature til prædiktion af SOH. FE metoden er fleksibel i valget af parametre, robust over for støj og testbetingelser og kræver kun små data til træning af modellen. Desuden analyseres interaktionen mellem testbetingelserne, entropy og præcisionen af estimationen. Resultaterne bruges til at vælge det korrekte spændingsinterval og forbedre nøjagtigheden af den entropy baserede SOH estimation. For at yderligere forbedre præcisionen af de entropy baserede metoder, bruges forskellige støjundertrykkelsesmetoder både før og efter ekstraktion af de entropy baserede features. De eksperimentelle resultater viser effektiviteten af udjævningstrinnet da det leder til højere præcisionen af estimation og hurtigere simuleringshastighed af ML modellen.

En anden løsning til at undgå potentielle fejl som kan opstå i mere manuelt ekstraherede features er ved at bruge neurale netværk, mere specifikt deep learning. Disse metoder tillader højdimensionelle input og kan automatisk ekstrahere features gennem såkaldte skjulte lag, men er på samme tid ekstremt afhængig af størrelsen af data. Derfor foreslås en bagging baseret ensemble metode som gør det muligt træne modeller på begrænsede mændger af data, så de opnår høj præcision og stadig har en gode generaliseringsevne. Endelig verificeres effektiviteten af den foreslåede metode ved forskellige eksperimenter, herunder den cykliske og kalender baserede aldringstest under forskellige temperaturforhold.

ACKNOWLEDGEMENTS

The work presented in this dissertation is a summary of the outcome from the Ph.D. project "Robust State of Health Estimation for Lithium-ion Batteries Using Machine Learning", which was carried out at AAU Energy, Aalborg University. This Ph.D. project is supported by Aalborg University, EUDP Denmark (under the CloudBMS project, grant number: 64017-05167), and Otto Mønsteds Fund. I would like to express my gratitude to the above-mentioned institutions.

My greatest respect and sincere thanks go to my main supervisor, Professor Remus Teodorescu, for his endless patience, guidance, and encouragement. During this three-year Ph.D. study journey, he has provided me with research opportunities, shared with me tremendous research experience, and guided me to move closer to my final goal step by step. I would like to thank him for his suggestions and support for my future career development. His pragmatic and serious attitude toward scientific research and life experiences inspired me, and this will be valuable during my entire life.

Commensurately, I would like to thank my co-supervisor, Associate Professor Daniel-Ioan Stroe, for all of his time and energy with my research. It must be said that he sacrificed countless precious weekends for me to revise my papers and provide valuable feedback. I am very grateful to him for tolerating my limitations and encouraging me to keep moving forward at all time.

I would also like to thank Professor Carlo Cecati from University of L'Aquila, Italy and Associate Professor Jinhao Meng from Sichuan University, China for providing me their external cooperation and valuable comments on my research work. I want to take this opportunity to thank my colleague Dr. Søren Byg Vilsen for using his expertise and thorough knowledge to provide suggestions and feedback, thank Dr. Anirudh Budnar Acharya for helping me in the laboratory, and also thank Dr. Maciej Swierczynski for help during data collection.

Special thanks to my colleagues and friends Gustavo Figueiredo Gontijo, Jinkui He, Dr. Hong Gong, Mengfan Zhang, Monika Sandelic, Roberta Di Fonso, and Xuewei Wu for bringing me joyful and relaxing times and bringing me all kinds of help. In particular, I would like to thank all my colleagues and secretaries at AAU Energy of Aalborg University for their kind help and detailed answers to any questions that I encountered during my Ph.D. studies. I would also like to express my gratitude to the assessment committee for the Ph.D. defense.

Last but not least, my sincerest gratitude goes to my entire family in China for their endless love, continuous support, and full understanding. Thanks to my husband, Shan He, for his patience, companionship, love in daily life, and for the valuable comments on my research work.

Xin Sui

Aalborg University, November, 2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

SOH	State of heath
SOC	State of charge
EOL	End-of-life
SEI	Solid electrolyte interface
EIS	Electrochemical impedance spectroscopy
LII	Loss of lithium-ion inventory
LAM	Loss of anode/cathode active materials
ICA	Incremental capacity analysis
DVA	Differential voltage analysis
APE	Absolute percentage error
MAPE	Mean absolute percentage error
RMSE	Root mean squared error
AE	Approximate entropy
SE	Sample entropy
FE	Fuzzy entropy
MSE	Multiscale sample entropy
CP	Conditional probability
ML	Machine learning
DL	Deep learning
LR	Linear regression
SVM	Support vector machine
GPR	Gaussian process regression
ANN	Artificial neural network
ELM	Extreme learning machine
EL	Ensemble learning
EMD	Empirical mode decomposition
IMF	Intrinsic mode function
TCs	Test conditions

HPPC	Hybrid pulse power characterization
CC	Constant current mode
CV	Constant voltage mode

LIST OF NOMENCLATURE

E_a	The activation energy of the chemical reaction
R_g	The gas constant
N	The number of data
r	The tolerance for accepting matrices
m	The dimension of vectors
D_N	The dataset containing N samples
D_N^b	The generated datasets by randomly sampling
$\mathbf{X}$	Input feature matrix
$\mathbf{Y}$	Output SOH values
$\mathbf{x}_i$	The ith sample point
x_{ij}	The jth feature value of ith sample point
y_i	Measured SOH value
$\hat{y}_i$	Predicted SOH value
$g(\cdot)$	Activation function
$\psi(\cdot)$	Mapping from input space to feature space
$K(\cdot)$	Kernel function
$f(\cdot)$	Established model
E_w	The sum of squared errors between the model and the output
e_i	The error variable
$\mathbf{W}$	Weight matrix
$\mathbf{W_0}$	The bias of hidden nodes
$\boldsymbol{\beta}$	Output weight matrix
$\mathbf{H}$	The hidden layer output matrix
O	The computational complexity
B	The number of bootstrap samples
M	The number of neurons of an ELM structure
$length_V$	The length of a voltage sequence
V_{start}	The starting point of a voltage sequence

k	Time step
V	Voltage
I	Current
T	Temperature
t	Time

PART I. REPORT

CHAPTER 1. INTRODUCTION

1.1. BACKGROUND

Global energy systems are altering in an effort to reduce carbon emissions and combat climate change [1]. Fossil fuels have been gradually replaced by the rapid development of renewable energy such as wind energy, solar power, and hydropower, etc. As renewable energy continues to integrate into the grid, battery energy storage systems play an important role [2]. A similar change can be observed in the transportation sector, where electric vehicles, mostly based on batteries, are gaining an important market share [3, 4]. Since the 1990s, Li-ion batteries have become an integral part of daily routines, and currently have the highest growth rate and the share of the market investments [J2, 5]. Due to the high power (up to 1500 W/kg) and energy density (up to 250 Wh/kg), the high energy efficiency (more than 95%), and also the relatively long cycle life (more than 3000 cycles), Lithium-ion (Li-ion) battery is a significant way to enhance the flexibility of clean energy and power EVs [J1, 6].

Li-ion batteries, similar to other energy storage systems, degrade with usage, resulting in a limited lifetime [7]. There are physical and chemical side reactions inside the battery that contribute to the degradation of the battery, i.e., graphite exfoliation, loss of electrolyte, solid electrolyte interface (SEI) layer formation, and continuous thickening, lithium plating, etc. [J2, 8]. Observed on a macroscopic level, battery aging manifests itself as the loss in capacity and power, which will seriously shorten the service life and decrease the efficiency of the system [9, 10]. It is therefore essential to determine the battery's state of health (SOH) at any moment during its operation. According to the estimated SOH, the battery management system can control the battery to run in the best state, so as to maximize its lifetime [J1, 11].

State-of-the-art of SOH estimation

SOH is a figure of merit that indicates the condition of the battery throughout its service life. SOH is usually expressed by two parameters, i.e., the battery capacity, and/or resistance (power) [J4]. Generally, battery aging data (i.e., voltage (V), current (I), temperature (T), and time (t),) can be measured and stored through various laboratory tests, as shown in Fig. 1.1 [J1]. For example, the constant current (CC) charging is for calculating the current capacity, the pulse current test is used to extract DC resistance, and the electrochemical impedance spectroscopy (EIS) test is used to measure the AC impedance. By utilizing the mentioned parameters, four types of SOH estimation methods are proposed [12-14]. Their performance comparison is summarized in TABLE 1.1.

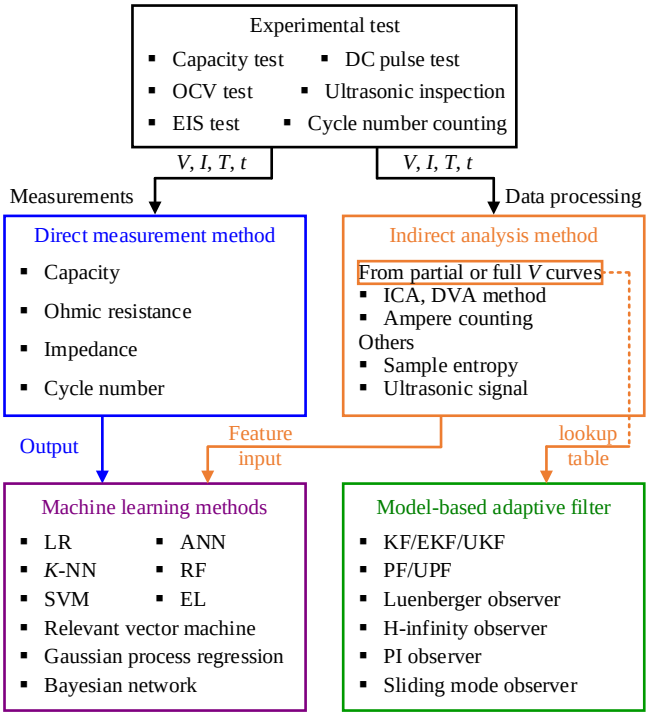


Fig. 1.1. Summary of existing methods for Li-ion battery SOH estimation. Source: [J1].

TABLE 1.1. Comparison of SOH estimation of lithium-ion batteries using different methods.

Method	Advantages	Disadvantages
Experimental methods	▪ Be suitable for investigating the aging mechanisms ▪ Provide a theoretical foundation for model-based approaches	▪ Be less feasible in real applications
Model-based filters	▪ Can be applied to any form of state-space model ▪ Good for Gaussian and non-Gaussian systems ▪ Provide on-line estimation	▪ Rely on an accurate battery model ▪ High calculation complexity
ML-based methods	▪ Does not need a battery model ▪ Be simple and feasible ▪ Good for all linear or non-linear systems	▪ Lack of interpretability ▪ Sensitive to features ▪ Complex training process

To explore the aging mechanisms, experimental approaches, i.e., measuring the charge throughput [15-17] or analyzing the health indicators [18-20] can be applied. However, the viability of these approaches in real-world applications is restricted. On the one hand, this method relies on high-precision current sensors. On the other hand,

the battery must be stopped from the regular operation so that the capacity test or EIS test can be conducted. Furthermore, due to the limitation in real-life systems, certain measurements are unable to be performed. For example, DC pulses with high currents are not permitted by the BMS of the EV since they are considered abnormal operating conditions [J1]. Various model-based filters have been designed, and the SOH, as an internal state variable, can be estimated by establishing the state-space model. For example, particle filter [21], Kalman filter [22], and nonlinear predictive filter [23] are established based on semi-empirical models, electrochemical models, or equivalent-circuit models for battery SOH estimation [J1]. The model-based filters can capture the dynamic character of the system, but they require an accurate battery model. That means a cumbersome task of battery parameters identification will be introduced.

Because Machine learning (ML) technologies do not rely on specific battery models, they possess immense potential in inferring battery SOH. According to laboratory data and learning and optimization theory, the mapping from features (i.e., health indicators containing the age-related information, such as *V*, *I*, *T*, and their statistics) to SOH (i.e., capacity, internal resistance) can be offline established. The general process of training an ML model for battery SOH estimation is illustrated in Fig. 1.2.

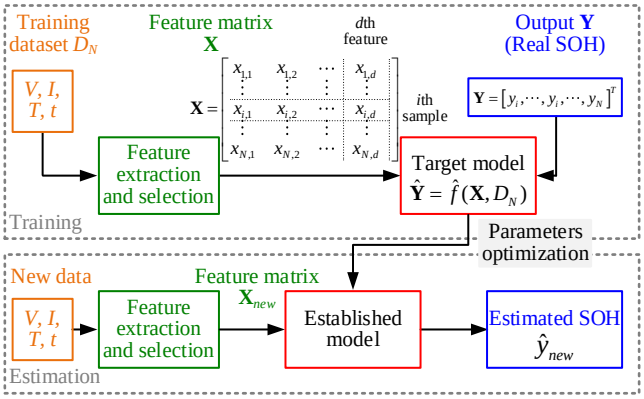


Fig. 1.2. Schematic diagram of ML-based SOH estimation for Li-ion batteries. Source: [J1].

Currently, extensive research has been carried out to apply ML technologies in battery SOH estimation, such as some classical ML methods including linear regression (LR) [24], support vector machine (SVM) [25], relevance vector machine [26-30], artificial neural network (ANN) [31-33], ensemble learning (EL) [34, 35], Gaussian process regression (GPR) [36], sparse Bayesian learning [37], etc. Furthermore, cloud computing provides benefits including global access, scalability, and increased computation power. This reduces the demands on the microprocessor and improves the SOH estimation accuracy as well, making the modern ML, i.e. deep learning (DL) with big data, a great solution for battery SOH estimation [J1, 34].

Comparison of various ML-based SOH estimation methods

The ML algorithms can be divided into the probabilistic-based methods and the non-probabilistic methods. The first type of ML methods, e.g., GPR and Bayesian networks, is useful for predicting the remaining useful lifetime of the battery over a long period of time, while the second type of ML method, e.g., LR, SVM, ANN, DL, and EL are mainly tasked for battery SOH estimation [J1]. Consequently, from the aspects of estimation performance (see Fig. 1.3) and publication trend (see Fig. 1.4), five non-probabilistic ML methods are comprehensively compared. As discussed in [J1], SVM has good approximation capability which is suitable for a small dataset, and is efficient in computing high-dimensional features. The EL algorithm represented by random forest has better accuracy and generalization performance, and can achieve more stable estimation results. The SVM and EL algorithms are identified to be suitable for SOH estimation and are the current research hotspots. Therefore they are selected as target models in this Ph.D. project.

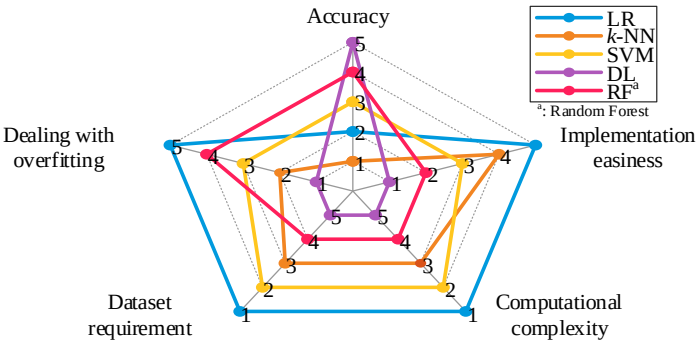


Fig. 1.3. Comparison of the five non-probabilistic ML algorithms across five metrics (i.e., accuracy, implementation easiness, computation complexity, dataset requirement, and dealing with overfitting. Source: [J1].

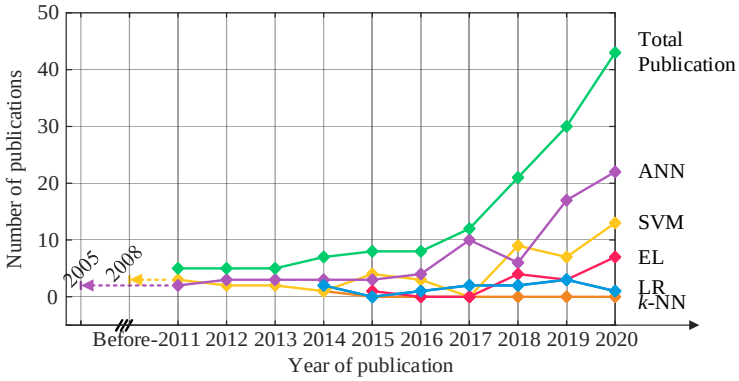


Fig. 1.4. The trend of publications of battery SOH estimation using non-probabilistic ML in over the last ten years according to the Elsevier Scopus international databases. Source: [J1].

1.2. PROJECT MOTIVATION

As aforementioned, ML represents an appealing solution for SOH estimation because it can extract aging information (i.e., features) from measurements and relate them to battery performance parameters, avoiding the complex battery modeling process. From the application perspective, there are nevertheless still some challenges that need to be addressed in ML-based battery SOH estimation, which motivated this Ph.D. project. The challenge mainly comes from three aspects: data collection and noise suppression, robust feature extraction, and algorithm optimization [J1].

- Many times, features are extracted from the electrochemical impedance spectroscopy measurements or full charging/discharging voltage curves, which are difficult to obtain in real-life applications because of specific requirements and restriction of the measurement devices [J4]. Utilizing the short-term pulse test, some differential features and geometric features can be extracted. However, these types of features face the problem of invalidation when the battery operating conditions change. As a consequence, robust and accurate assessment of the battery's SOH remains a challenge [12, 13]. Therefore, it is essential to explore effective SOH features that can be conveniently obtained and are robust to aging conditions.
- The measured battery data is often disturbances by a certain degree of noise from i.e., the inner electrochemical reactions and external environmental conditions [C3]. In this case, data smoothing techniques are needed to reduce the noise in the original data. At the same time, it is necessary to ensure that valid aging information contained in the data is not filtered out.
- DL with big data provides a possible choice for accurate SOH estimation under different aging conditions. SOH features can be automatically extracted from high-dimensional data. However, they depend on the availability of big data sets, which increases the laboratory effort. Therefore, expanding the existing data sets by increasing data diversity would be a convenient and quick solution. Besides, as an emerging substitute of DL, the EL of simple base learners make a trade-off between data size and accuracy [J1].

1.3. PROJECT OBJECTIVES AND LIMITATIONS

1.3.1. PROJECT OBJECTIVES

Based on the above consideration, the main objective of the Ph.D. project is defined:

- **Identify and optimize robust ML-based SOH estimation algorithms for Li-ion batteries.**

In order to achieve the main objective, the following technical objectives are defined from three aspects, which are robust features selection, data noise suppression, and improved model establishing, respectively.

- O1: Identify the suitable ML methods for battery SOH estimation by analyzing the capacity fade behavior and deriving the principles of different ML algorithms.
- O2: Propose effective SOH features which can be extracted easily and are particularly robust to aging conditions, data noise, and data size.
- O3: Propose a data noise suppression method to effectively reduce the influence of data noise on the estimation performance while not filtering out the useful aging information.
- O4: Propose an improved ML-based SOH estimation model which has competitive performance on high-dimensional input while it relies on reduced training data.

1.3.2. PROJECT LIMITATIONS

This Ph.D. project proposes robust SOH estimation algorithms for Li-ion batteries based on ML methods. However, this project is also subject to several limitations:

- Because aging data of battery packs is not available, the research on the SOH estimation is still at the battery cell level. The effectiveness of the proposed methods for the SOH estimation of battery pack level needs further demonstration.
- Due to the different design of the aging experiment and the tested battery chemistry, other data sets maybe not be suitable for the proposed methods. Besides the data obtained in our laboratory, no other data sets are used for verification.
- The probabilistic ML methods will be very useful for long-term battery remaining useful life prediction as they can provide the estimate of the uncertainty in the prediction as well as a clear probabilistic interpretation. In this project, non-probabilistic algorithms are basically adopted because they are capable to solve the problem of SOH estimation.

1.4. MAIN CONTRIBUTIONS

- **MC1: An in-depth review of ML-based SOH estimation methods is provided, as well as a discussion of future prospects.**

The basic principles of existing non-probabilistic ML-based SOH estimation algorithms are rigorously derived. Then, the challenges that remain unresolved and the corresponding improvement methods of these ML-based methods are introduced. Additionally, the training modes of these methods are summarized and their estimation performance are fully compared. Finally, to inspire the researchers, the outlook of the future research focus on SOH estimation methods is given. The relevant contents have been published in [J1].

- **MC2: A nonlinear regression method with two-step fitting is proposed to establish a long-term aging model of Li-ion battery under multiple impact factors.**

By conducting calendar aging tests on Li-ion batteries over a period of up to 43 months, the long-term aging behavior, which is beyond the typical range of 20% capacity fade, is analyzed. A lifetime model on the basis of the nonlinear regression approach is established. The model can simultaneously study the effect of storage temperature and state of charge (SOC) level on the rate of calendar degradation. In addition, the proposed model reflects the piecewise variation of capacity fade, and the lifetime prediction results are more reliable than those from the traditional semi-empirical model. The relevant contents have been published in [J2].

- **MC3: A novel SOH feature based on fuzzy entropy (FE) is proposed for Li-ion battery SOH estimation, which is not only convenient to obtain but also very effective.**

From the short-pulse test, the FE feature can be easily obtained which is time-saving and convenient for both laboratory measurements and real application. The experimental results proved that compared to traditional SE-based estimator, the FE-based approach is freer in parameter selection, more robust against noise and aging test conditions (TCs), and requires less training data. Furthermore, three SOH estimation models considering the aging temperature are proposed, which verify the robustness of the FE feature to the aging conditions. The relevant contents have been published in [J4].

- **MC4: A data sets selection method is proposed to improve ML-based battery SOH estimation.**

To begin with, by changing the SOC level and the current direction, six types of voltage data sets are collected. Then the effect of TCs on the performance of entropy-based SOH estimation is examined. Finally, the SOH estimation results under cyclic aging conditions prove that extracting entropy features from the polarization zone

helps increase the SOH estimation accuracy. The relevant contents have been published in [J3].

- **MC5: A data noise suppression method is proposed and evaluated for further enhancing the robustness of ML-based SOH estimation.**

Two types of data noise suppression methods, i.e., the adaptive iterative algorithms and the regression-based linear smoothers are evaluated and combined with the ML-based SOH estimation model. The estimation results show that a simple data noise suppression step can improve not only the SOH estimation accuracy but also the simulation speed. The relevant contents have been published in [C3] and [C4].

- **MC6: An EL method is proposed to reduce the dependence of ML-based SOH estimation method on the data size.**

A bagging-based extreme learning machine (ELM) is proposed. By resampling the original data to generate more subsets, it is possible to train multiple base-level ELMs on a limited data set. Additionally, the hyperparameters of the proposed algorithm are optimally selected. The experimental results demonstrate that the presented SOH estimation method maintains the advantages of ELM less computational complexity, and fast operation as well as achieve higher accuracy and generalization performance. The relevant contents have been published in [C5].

1.5. THESIS OUTLINE

The summary of the outcomes of this Ph.D. project are documented in the form of article-based Ph.D. thesis, including a ***Report*** and a collection of the ***Selected Publications*** through the entire study. The thesis structure is shown in Fig. 1.5 to illustrate how to connect the content in the ***Report*** to the ***Selected Publications***.

The ***Report*** is organized into six chapters. ***Chapter 1*** introduces the background and motivation of the Ph.D. project, after which the research objectives and limitations of the Ph.D. study are discussed. Then the following four chapters deal with the robust machine learning-based SOH estimation for Li-ion batteries from three levels, namely, data analysis and preprocessing, robust feature extraction, and algorithm optimization. ***Chapter 2*** aims at data analysis and presents the calendar aging and cyclic aging tests that batteries are subjected to. Then a lifetime model is proposed to investigate the effect of test conditions on the long-term calendar degradation behavior. The main focus of ***Chapter 3*** and ***Chapter 4*** is the robust feature extraction for ML-based SOH estimation. In ***Chapter 3***, entropy is selected as SOH feature for the following research. The effects of datasets selection on the performance of various entropy features are analyzed. FE is proposed as a novel SOH feature and its estimation performance is validated with experiments. ***Chapter 4*** uses the proposed FE feature for SOH estimation under various temperature conditions. The performance of the feature is further improved by data noise suppression methods. ***Chapter 5*** focuses on improving the SOH estimation performance of ML for batteries from the level of algorithm

optimization. The SOH estimation results of using a bagging-based ELM method is presented. Finally, ***Chapter 6*** concludes the findings of this Ph.D. project and proposes topics for future work.

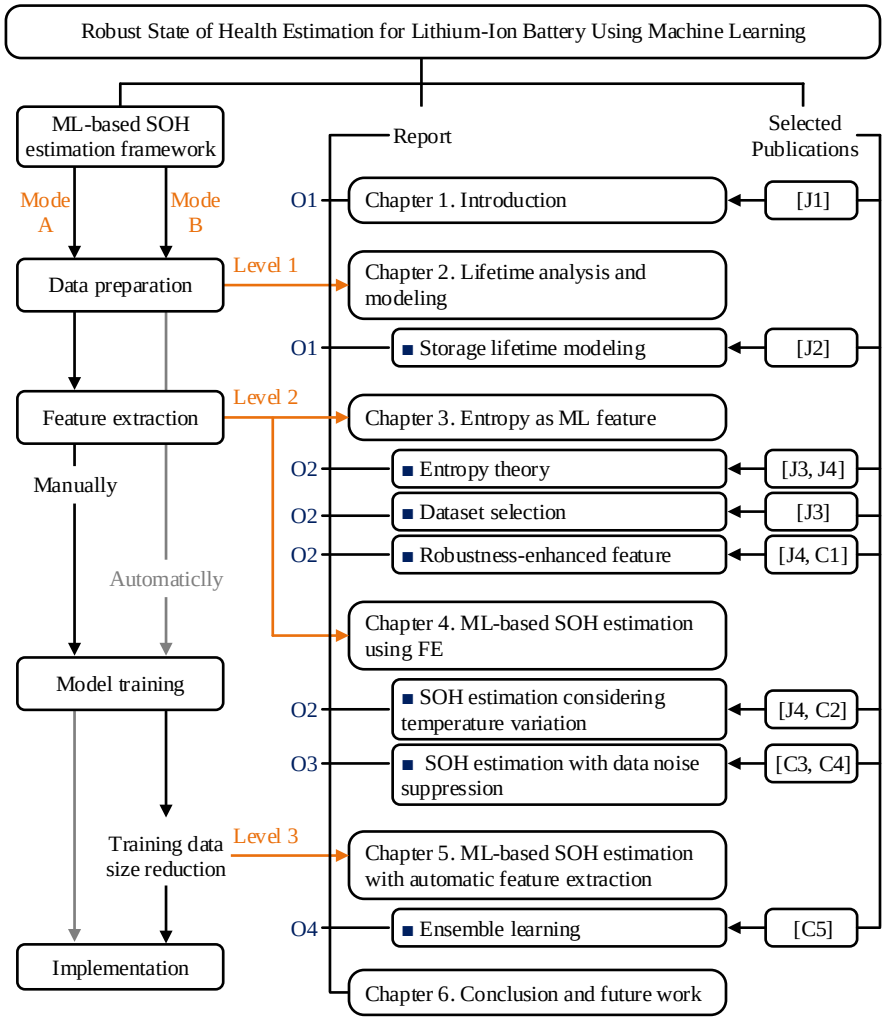


Fig. 1.5. Thesis structure with related topics and research outcomes for each chapter; where "O" denotes objective, while "J" and "C" refer to a journal or conference publication, respectively. Level 1, level 2, and level 3 represent the data level, the feature level, and the algorithm level in the proposed ML framework. Correspondingly, the overall performance of ML-based SOH estimation is improved from three aspects: data analysis and preprocessing, robust feature extraction, and algorithm optimization.

1.6. LIST OF PUBLICATIONS

The research outcome during the Ph.D. study has been disseminated in the form of journal papers and conference publications, as listed below. Parts of these publications are summarized and used in the Ph.D. dissertation (***Report***) as previously listed in Fig. 1.5.

Journal Papers:

J1. **X. Sui**, S. He, S.-B. Vilsen, J. Meng, R. Teodorescu, and D.-I. Stroe, "A review of non-probabilistic machine learning-based state of health estimation techniques for Li-ion battery," *Appl. Energy*, vol, 300, pp. 117346, Oct. 2021.

J2. **X. Sui**, M. Świerczyński, R. Teodorescu, and D.-I. Stroe, "The degradation behavior of $LiFePO_4$/C batteries during long-term calendar aging," *Energies*, vol. 14, no. 6, pp. 1732, Mar. 2021.

J3. **X. Sui**, D.-I. Stroe, S. He, X. Huang, J. Meng, and R. Teodorescu, "The effect of voltage dataset selection on the accuracy of entropy-based capacity estimation methods for lithium-ion batteries," *Appl. Sci.*, vol. 9, no. 19, pp. 4170, Oct. 2019.

J4. **X. Sui**, S. He, J. Meng, R. Teodorescu, and D.-I. Stroe, "Fuzzy entropy-based state of health estimation for Li-ion batteries," *IEEE Trans. Emerg. Sel. Topics Power Electron.*, vol. 9, no. 4, pp. 5125-5137, Aug. 2021.

Conference Papers:

C1. **X. Sui**, S. He, D.-I. Stroe, and R. Teodorescu, "State of health estimation for lithium-ion battery using fuzzy entropy and support vector machine," in *Proc. IEEE ECCE Asia*, pp. 1417-1422, 2020.

C2. **X. Sui**, S. He, J. Meng, R. Teodorescu, and D.-I. Stroe, "Fuzzy entropy-based state of health estimation of $LiFePO_4$ batteries considering temperature variation," in *Proc. IEEE ECCE*, pp. 4401-4406, 2020.

C3. **X. Sui**, S. He, D.-I. Stroe, and R. Teodorescu, "Lithium-ion battery state of health estimation using empirical mode decomposition sample entropy and support vector machine," in *Proc. IEEE APEC*, pp. 3424-3429, 2020.

C4. **X. Sui**, S. He, X. Huang, R. Teodorescu, and D.-I. Stroe, "Data smoothing in fuzzy entropy-based battery state of health estimation," in *Proc. IEEE IECON*, pp. 1779-1784, 2020.

C5. **X. Sui**, S. He, S.-B Vilsen, R. Teodorescu, and D.-I. Stroe, "Fast and robust estimation of Lithium-ion batteries state of health using ensemble learning," in *Proc. IEEE ECCE*, pp. 1393-1399, 2021.

Other Publications not included in the Ph.D. thesis:

OP1. **X. Sui**, S. He, J. Meng, D.-I. Stroe, X. Huang, and R. Teodorescu, "Optimization of the discharge cut-off voltage in $LiFePO_4$ battery packs," in *Proc. IEEE ECCE Europe*, pp. P.1-P.8, 2019.

OP2. **X. Sui**, S. He, D.-I. Stroe, X. Huang, J. Meng, and R. Teodorescu, "A review of sliding mode observers based on equivalent circuit model for battery SoC estimation," in *Proc. IEEE ISIE*, pp. 1965-1970, 2019.

OP3. J. Meng, L. Cai, D.-I. Stroe, G. Luo, **X. Sui**, and R. Teodorescu, "Lithium-ion battery state-of-health estimation in electrical vehicles using optimized partial charging voltage profiles," *Energy*, vol. 185, pp. 1054-1062, 2019.

OP4. Y. Li, D.-I. Stroe, Y. Cheng, H. Sheng, **X. Sui**, and R. Teodorescu, "On the feature selection for battery state of health estimation based on charging–discharging profiles," *J. Energy Storage*, vol. 33, pp.102122, 2021.

OP5. X. Huang, Y. Li, A.B. Acharya, **X. Sui**, J. Meng, R. Teodorescu, and D.-I. Stroe, "A review of pulsed current technique for lithium-ion batteries," *Energies*, vol. 13, pp. 2458, 2020.

OP6. R. Teodorescu, A.B. Acharya, **X. Sui**, X. Huang, and D.-I. Stroe, "Smart battery technology for battery lifetime control," *Energies*, 2021. (***Major Revision***).

OP7. X. Huang, **X. Sui**, D.-I. Stroe, and R. Teodorescu, "A review of management architectures and balancing strategies in smart batteries," in *Proc. IEEE IECON*, pp. 5909-5914, 2019.

OP8. S.-B. Vilsen, **X. Sui**, and D-I. Stroe, "A time-varying Log-linear model for predicting the resistance of lithium-ion batteries", in *Proc. IEEE ECCE Asia*, 2020.

OP9. X. Huang, A. B. Acharya, J. Meng, **X. Sui**, D.-I. Stroe, and R. Teodorescu, "Wireless smart battery management system for electric vehicles," in *Proc. IEEE ECCE*, pp. 5620-5625, 2020.

OP10. X. Huang, Y. Li, J. Meng, **X. Sui**, R. Teodorescu, and D.-I. Stroe, "The effect of pulsed current on the performance of lithium-ion batteries," in *Proc. IEEE ECCE*, pp. 5633-5640, 2020.

OP11. S. Jin, X. Huang, **X. Sui**, S. Wang, R. Teodorescu, and D.-I. Stroe, "Overview of methods for battery lifetime extension," in *Proc. IEEE ECCE Europe*, pp. 1-8, 2021.

OP12. R. Teodorescu, **X. Sui**, A. B. Acharya, D.-I. Stroe, and X. Huang, "Smart battery concept, a battery that can breathe," in *5th E-Mobility Integration Symposium*, 2021.

OP13. R.-D. Fonso, **X. Sui**, A. B. Acharya, R. Teodorescu, and C. Cecati, "Multidimensional machine learning balancing in smart battery packs," in *Proc. IEEE IECON*, pp. 1-6, 2021.

CHAPTER 2. LIFETIME ANALYSIS AND MODELING

The content of this chapter is based on the results presented in [J2].

This chapter introduces the methods used to reach objective 1 and mainly answer the following questions:

- What types of aging batteries usually go through?
- What aging mechanism correspond to these types of aging?
- What data sets are used in this project for battery SOH estimation?
- How to model the capacity fade behavior of batteries?
- What are the effects of different conditions on the lifetime of batteries?

2.1. BACKGROUND

The Li-ion battery, taking lithium iron phosphate battery ($LiFePO_4$) as an example, is a type of electrochemical energy storage device with graphite as anode and $LiFePO_4$ as cathode material. It works based on the intercalation/deintercalation process of the lithium ions between the layered structure of the anode and the lattice of the cathode. Studies show that the loss of lithium-ion inventory (LII) and loss of anode/cathode active materials (LAM) constitute the main reasons for battery degradation [J2, 38, 39]. Commonly, there are two types of battery degradation under real-life operations: cyclic and calendar aging. The degradation of $LiFePO_4$ during cycling is mainly due to kinetic-inducted effects such as LAM, lithium plating at the anode, and reversibility at the cathode side [9, 11]. Conversely, there are independent degradation behaviors involved in the calendar aging process. Numerous parasitic reactions occur at the electrode-electrolyte interfaces that contribute to the degradation [J2, 11]. To separately study the two types of degradation, cyclic aging and the calendar aging tests are performed on commercial high-power, cylindrical Li-ion batteries in this project [40]. TABLE A.1 (see Appendix A) lists the main electrical parameters of the cells. The experimental setups used for the accelerated aging tests are given in Appendix B and the detailed testing procedure in Appendix C. The tests are conducted in climatic chambers under temperature control. The temperatures shown in this report are measured on the battery's surface.

2.2. ACCELERATED AGING DATA

Typically, a battery's SOH will be 1 at the time of manufacturing and will decrease with battery degradation. In this project, a battery's SOH is calculated by dividing its current available capacity by its initial available capacity [J4]. Because 20% capacity

fade is adopted as the end-of-life (EOL) criteria of the battery, the battery SOH at the end of aging reaches 0.8. The results of cyclic aging are shown in Fig. 2.1. The whole cyclic aging test lasts for 38 weeks and the pulse tests are conducted after each week of aging, therefore 38 voltage series will be collected in each TC. However, six voltage series are missing during the experiment, and the week numbers corresponding to these missing data are 3, 10, 18, 32, 35, and 36.

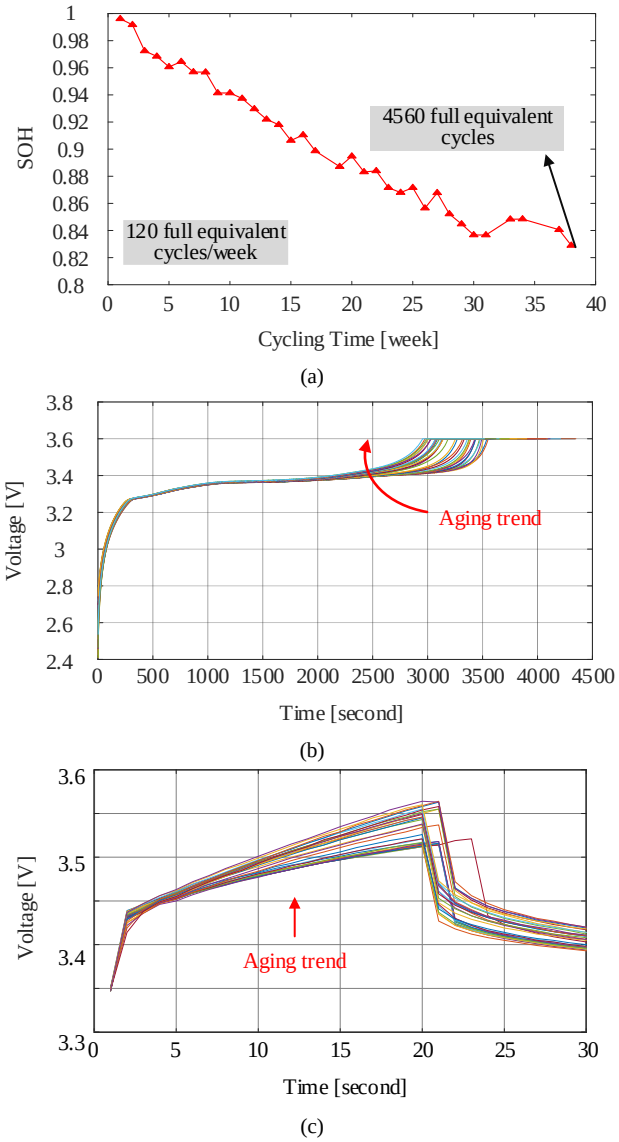


Fig. 2.1. Cyclic aging results: (a) SOH curve during the cyclic aging, (b) voltage responses during the CC discharging, and (c) voltage responses during the HPPC test (illustrate in the case of pulse charging at 80% SOC).

Fig. 2.2 displays the results of calendar aging test, including the measured capacity fade under different TCs and the voltage responses during the modified hybrid pulse power characterization (HPPC) tests.

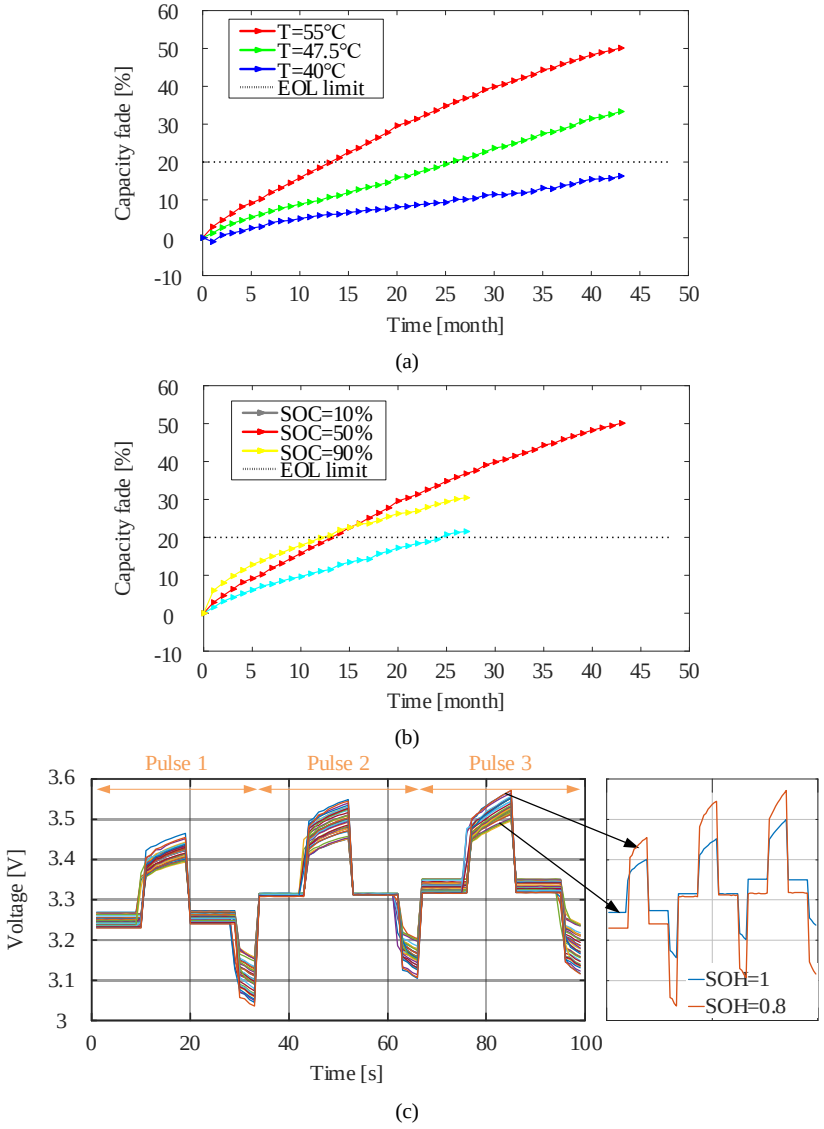


Fig. 2.2. Calendar aging results: (a) Capacity fade curves under different temperatures (at 50% SOC level), (b) Capacity fade curves under different SOC levels (at 55°C), and voltage datasets obtained during the HPPC tests.

2.3. CALENDAR LIFETIME MODELING

Calendar aging might be the primary part of battery degradation, especially for the applications where the batteries have been in an idle state or the depth of discharge and working current rates are relatively low. Three types of aging approaches including the pure-lifetime model, the physics-based model, and the mathematical models are proposed, as summarized in TABLE 2.1. The principles of each method are introduced, and the performance of these aging models are compared according to the complexity, accuracy, and interpretability.

TABLE 2.1. The summary of various lifetime models. Source: [J2].

Method	Complexity	Accuracy	Interpretability[a]	Exemplary applications
Pure-lifetime model ▪ Calculate the amount of throughput in charge or energy.	Low	Low	No	[41]
Physics-based electrochemical model ▪ Describe the thickness and conductivity of the SEI mathematically.	High	High	High	[42]
Physics-based empirical model ▪ Create a time-dependent power law to represent the aging of the calendar.	Medium	Medium	No	[43]-[44]
Physics-based semi-empirical model ▪ Consider the Arrhenius law, the Eyring law, or the underlying SEI growth.	Medium	Medium	Medium	[45]-[47]
Physics-based equivalent circuit model ▪ Add the "aging-dimension" to equivalent circuit models.	High	Medium	Medium	[10]
ML-based mathematical model ▪ Establish mapping between battery state and measurements based on evolution rules.	Low	High	No	[48]

[a] Interpretability refers to the representativeness of the model to the physicochemical properties.

The aging data collected during the above experiments are used in the following chapters for training and verification of the ML-based SOH estimation. In this chapter, to understand the long-term calendar aging behavior of the batteries, a nonlinear regression method with two-step fitting is proposed. The aging temperature and storage SOC level are considered as stress factors in the model created, therefore to study their influence on the rate of calendar degradation. Moreover, for model comparison and validation, a traditional semi-empirical model is also established.

2.3.1. TRADITIONAL SEMI-EMPIRICAL MODEL

In the traditional semi-empirical modeling, the Arrhenius equation, as given in (2.1) can be used [49, 50]. Because the calendar aging is primarily associated with the growth of SEI layer, the capacity fade of the batteries over time can be described as (2.2)

$$\ln k = \ln k_{ref} + \frac{-E_a}{R_g}\left(\frac{1}{T} - \frac{1}{T_{ref}}\right) \tag{2.1}$$

$$C_{fade\_semi}(t) = k \times t^{A} \tag{2.2}$$

where T and T_{ref} denote the absolute temperature and the reference temperature in Kelvin, respectively. k denotes the influence of temperature on calendar aging, and k_{ref} represents the coefficient corresponding to T_{ref}. The activation energy of the chemical reaction occurring at temperature T is denoted by E_a, and R_g is the gas constant. A is the time index, and it is usually set to 0.5~1 according to existing models. For better fitting the measurements, A is selected as 0.75. Fig. 2.3 shows the relationship between the Arrhenius coefficient in natural logarithm form and temperature. Then the comparison of experimental and semi-empirical modeling results is shown in Fig. 2.4. However, this model fails to study the dependency of the battery degradation on SOC level. Hence, a nonlinear regression model with two-step fitting is proposed, and the main content are summarized as follows.

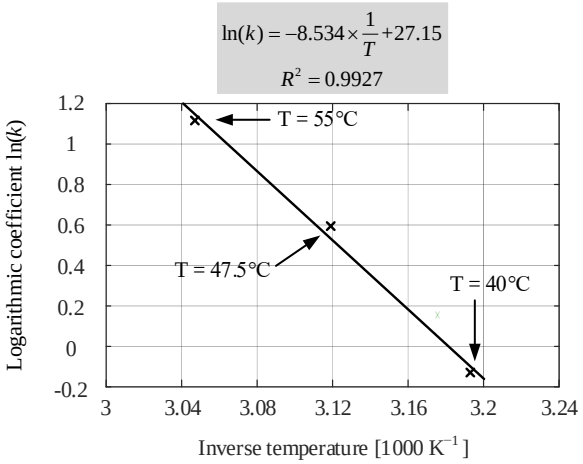


Fig. 2.3. Arrhenius equation for describing the effect of temperature on the capacity fade (SOC level is set at 50%). Source: [J2].

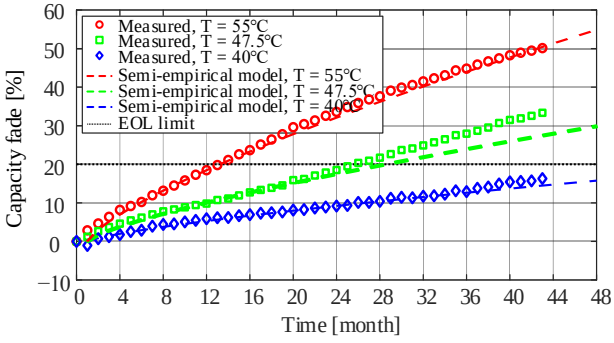


Fig. 2.4. Semi-empirical model of capacity fade with respect to temperature and storage time (SOC level is fixed to 50%). Source: [J2].

2.3.2. NONLINEAR REGRESSION MODEL WITH TWO-STEP FITTING

The proposed regression method separates the dependency of capacity fade on the aging time and the stress factors for analysis. The modeling process consists of two steps:

- Identifying the relationship $C_{fade}(t)$ between the capacity fade and the calendar aging time.
- Modeling the coefficients of $C_{fade}(t)$ by jointly considering the aging temperature and storage SOC.

To determine $C_{fade}(t)$ in the first step of battery lifetime modeling, four types of functions are used, as shown in TABLE 2.2 and Fig. 2.5. The power function provides the best fitting results, which is consistent with the method suggested by the semi-empirical model. For the power equation with a constant term 0.7, the small initial error of 0.7% can be ignored, and the function converges quickly. In addition, choosing a variable constant term in the power function will not bring obvious improvement of the accuracy, but increase the computational complexity. Therefore, for the above reasons, (2.3) is selected.

$$C_{fade}(t) = a \times t^{b} + 0.7 \tag{2.3}$$

TABLE 2.2. Comparison of different modeling results (Take the calendar aging test of Case 1 @55°C and 50% SOC as an example). Source: [J2].

Fitting Function Type		Number of Parameters	R^2
Logarithmic function	$C_{fade}(t) = a \times \ln(b \times t)$	2	0.9748
Polynomial function	$C_{fade}(t) = a \times t + b$	2	0.9876
Power function with a variable constant term	$C_{fade}(t) = a \times t^{b} + c$	3	0.9980
Power function with a fixed constant term	$C_{fade}(t) = a \times t^{b} + 0.7$	2	0.9974

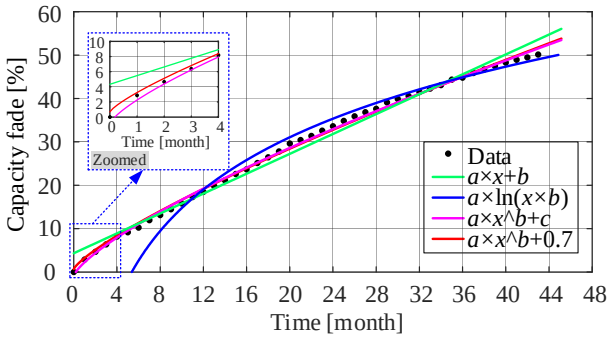


Fig. 2.5. Different modeling results showing the relationship between the capacity fade and storage time. (Take the calendar aging test of Case 1 @55°C and 50% SOC as an example). Source: [J2].

To study the effect of the stress factors (i.e., the storage temperature and SOC level) on the capacity fade in the second step, coefficients in the target model (2.3) will be modeled. The nonlinear regression method can be used to optimize the coefficients, and its structure of the regression method is shown in Fig. D.2. The detailed flows of nonlinear regression are shown in Fig. D.1 (see Appendix D.1). After obtaining the capacity fade function, the coefficient of determination R^2 is utilized to assess the quality of the results. R^2 can be expressed in (2.4), and the closer R^2 is to 1, the better the fit.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{2.4}$$

where SS_{res} and SS_{tot} are the sum of squares of residuals and the total sum of squares, respectively. SS_{res} represents the deviation between the measurements and their estimates, SS_{tot} represents the deviation between the measurements and their average values.

In the second step of battery lifetime modeling, the dependency of the capacity fade on the considered stress factors is modeled in turn. By varying the storage temperature from 55°C to 40°C while fixing the SOC level to 50%, the capacity fade can be determined as a function of the storage temperature. The general model (2.3) is detailed to (2.5).

$$C_{fade}(t) = a_T \times t^{b_T} + 0.7 \tag{2.5}$$

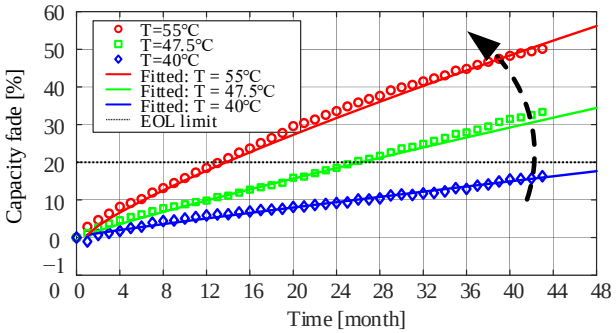


Fig. 2.6. Nonlinear regression model of capacity fade with respect to temperature and storage time (SOC level is fixed at 50%). Source: [J2].

TABLE 2.3. The coefficients of the nonlinear regression model that takes the storage temperature variation. (SOC level is fixed to 50%) into account. Source: [J2].

Test condition	Temperature (°C)	a_T	b_T	R^2
Case 1	55 °C	2.428	0.812	0.9974
Case 2	47.5 °C	1.08	0.897	0.9943
Case 3	40 °C	0.452	0.932	0.9862

Fig. 2.6 shows the fitted results for the $LiFePO_4$ battery cells from the aging tests Case 1, 2, and 3. For each case, the obtained coefficients are listed in TABLE 2.3. Consequently, the coefficient a_T increases as the temperature rises indicating that higher storage temperature accelerates the capacity fade. With b_T value of less than 1, the capacity fade decreases with time, which is in line with the aging characteristics of the battery. The aging trend described above is mainly related to the increase in SEI under the effect of the temperature. As discussed in [J2], LII and LAM the main cause for the calendar aging of the battery [39, 50]. The increasing temperature can cause the continuous thickening of the SEI film, thus lead to the LII. Nevertheless, the thickening of the SEI film will, in turn, limits side reactions in the battery, thus reducing the battery aging [J2].

By plotting parameters a_T and b_T on Fig. 2.7, the relationships between the aforementioned coefficients and temperature are determined, as given in (2.6) and (2.7), respectively.

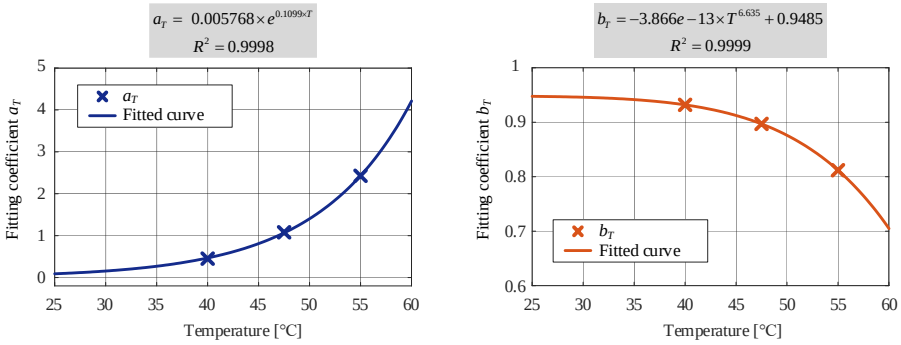
$$a_T = 0.005768 \times e^{0.1099 \times T} \tag{2.6}$$

$$b_T = -3.866 \times 10^{-13} \times T^{6.635} + 0.9485 \tag{2.7}$$

Next, (2.5), (2.6), and (2.7) are combined together, and a general model is obtained, as given in (2.8).

$$C_{fade}(t,T) = 0.005768 \times e^{0.1099 \times T} \times t^{-3.866 \times 10^{-13} \times T^{6.635} + 0.9485} + 0.7 \tag{2.8}$$

where t denotes the storage time, expressed in months, and T denotes the storage temperature, expressed in degrees Celsius.



(a) (b)

Fig. 2.7. The relationship between the coefficients of the nonlinear regression model and the relevant storage temperature: (a) Exponential function of a_T with respect to T, and (b) Power function of b_T with respect to T. Source: [J2].

A similar procedure is performed to model the capacity fade in terms of SOC level, as shown in (2.9). In this step, the aging data from Case 1, 4, and 5, where the temperature is all at 55 °C, are considered. The measured and modelled capacity fade

curve under different SOC levels are shown in Fig. 2.8, and the coefficients in (2.9) are summarized in TABLE 2.4.

$$C_{fade}(t, SOC) = a_{SOC} \times t^{b_{SOC}} + 0.7 \tag{2.9}$$

As observed from Fig. 2.8, the SOC level shows piecewise impacts on the capacity. Before the battery cells reach 20% capacity fade, a higher SOC will cause a faster capacity fade. As with the influence of temperature, there is a similar explanation. The high SOC leads to a deepening of the lithiation of the graphite anode, as well as an acceleration of side reactions. In contrast, SEI film becomes stable when the capacity fades higher than 20%. As a result, the speed of the battery degradation under this condition decreases [39, 50]. Then the parameters a_{SOC} and b_{SOC} are plotted on Fig. 2.9. According to and the relationships between the aforementioned coefficients and SOC level are as given in (2.10) and (2.11).

$$a_{SOC} = 1.087 \times e^{0.0169 \times SOC} \tag{2.10}$$

$$b_{SOC} = -4.853 \times 10^{-12} \times SOC^{5.508} + 0.823 \tag{2.11}$$

By combining (2.9), (2.10) and (2.11), a general model is established, as given in (2.12). This model describes the dependency of capacity fade in conjunction with storage time and SOC level at the same time.

$$C_{fade}(t, SOC) = 1.087 \cdot e^{0.0169 \cdot SOC} \cdot t^{-4.853 \cdot 10^{-12} \cdot SOC^{5.508} + 0.823} + 0.7 \tag{2.12}$$

where t represents the storage time given in months, and SOC is the storage SOC level, given in percentage.

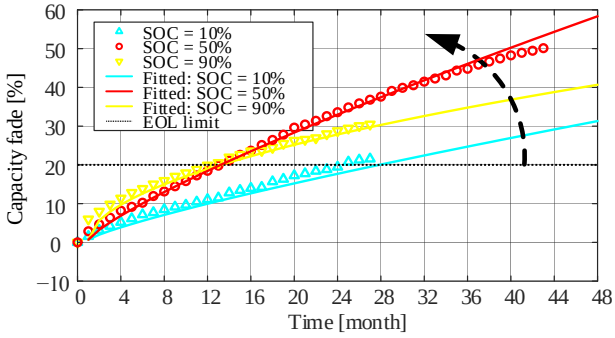
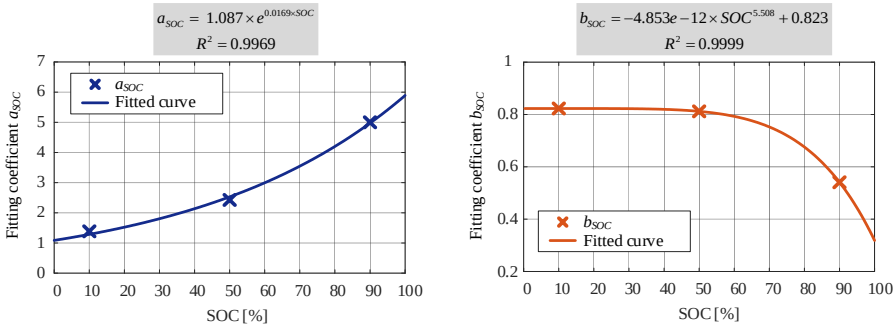


Fig. 2.8. Nonlinear regression model of capacity fade with respect to SOC level and storage time (at 55 °C). Source: [J2].



(a) (b)

Fig. 2.9. The relationship between the coefficients of the nonlinear regression model and the relevant storage SOC level: (a) Exponential function of a_{SOC} with respect to *SOC*, and (b) Power function of b_{SOC} with respect to *SOC*. Source: [J2].

TABLE 2.4. The coefficients of the nonlinear regression model that takes the storage SOC level variation (at 55 °C) into account. Source: [J2].

Test condition	SOC (%)	a_{SOC}	b_{SOC}	R^2
Case 4	10%	1.387	0.823	0.9973
Case 1	50%	2.428	0.812	0.9974
Case 5	90%	4.999	0.541	0.9990

As stated up to this point, the impact of the storage temperature and SOC level on the capacity fade are modelled separately. In order to fuse both contributions into one function, a scaling of the two obtained functions is used. It should be noted that the scaling process is considered to be accurate enough, because the possible interactions between two stress factors are ignored in this work. As a result, a general battery lifetime model is obtained, as presented in (2.13).

$$C_{fade}(t,T,SOC) = 0.0025 \times e^{0.1099 \times T} \times e^{0.0169 \times SOC} \times t^{(-3.866 \times 10^{-13} \times T^{6.635} - 4.853 \times 10^{-12} \times SOC^{5.508} + 0.9595)} + 0.7 \quad (2.13)$$

Based on the model obtained above, the calendar lifetime of the battery under different storage conditions can be inferred. As shown in Fig. 2.10, when not being used, making the batteries' SOC level lower than 20% or higher than 80% as well as cooler temperatures prolong the batteries' lifetime [J2]. Batteries stored at 10% SOC and 25 °C will be expected to withstand approximately 45.1 years until they reached the EOL threshold. However, when the storage conditions are changed to 50% SOC and 25 °C, the lifetime drops dramatically to 23.8 years, and to 8.7 years at 40 °C [J2].

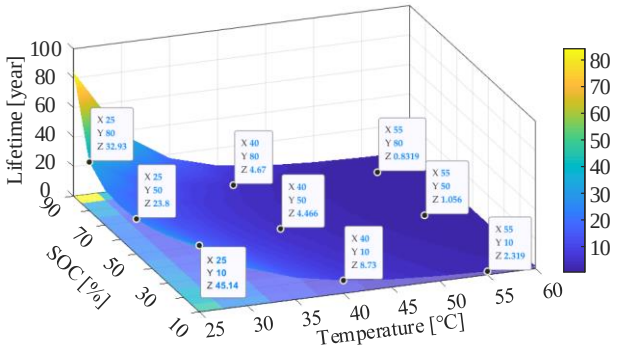


Fig. 2.10. The prediction results of the battery lifetime under different storage conditions using the nonlinear regression model (when EOL criterion of 20% capacity fade is reached). Source: [J2].

2.3.3. COMPARISON

Based on the above analysis, it can be found that compared to the traditional semi-empirical model, the proposed nonlinear regression model is more able to capture the effects of storage time, temperature, and SOC level simultaneously. In addition, the proposed method, as shown in (2.13), can well model the piecewise degradation behavior under the influence of SOC level by introducing the variable exponent. However, the Arrhenius equation can only be used to investigate the influence of the storage temperature on the capacity fade. In other words, the traditional semi-empirical model is unable to account for the piecewise effect of SOC on battery degradation.

A quantitative analysis is also conducted to compare the prediction accuracy of the two lifetime models when only time and temperature dependences are considered. Fig. 2.11 compares two kinds of capacity fade modeling method, and in TABLE 2.5 shows the lifetime prediction results. It is evident that the proposed model performs better in terms of prediction accuracy. Moreover, compared to the semi-empirical model, the proposed model gives underpredicted lifetime values which is good for predictive maintenance. Since the battery stored at 40 °C has not reached the EOL before stopping the entire aging test, the real lifetime values at this temperatures cannot be provided.

It should be noted that the battery under the condition of 40 degrees does not finally reach the EOL of 20% capacity attenuation, so it cannot provide the measurement life at lower temperature [J2].

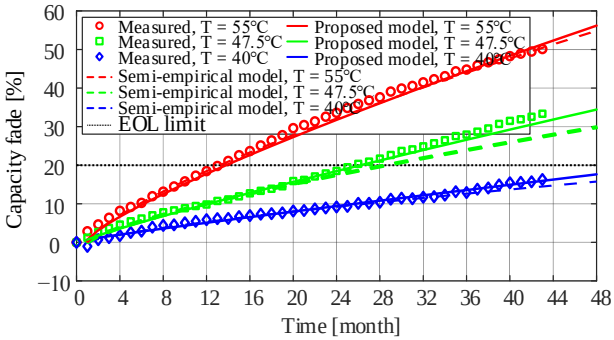


Fig. 2.11. Comparison of modeling results between semi-empirical model and the proposed nonlinear regression model (SOC level is fixed at 50% and both models only consider the temperature variation and time dependency). Source: [J2].

TABLE 2.5. Comparison of lifetime prediction results between semi-empirical model and the proposed nonlinear regression model (SOC level is fixed at 50% and both models only consider the temperature variation and time dependency; when EOL criterion of 20% capacity fade is reached). Source: [J2].

Temperature (°C)	Predicted Lifetime Using the Proposed Model (month)	Predicted Lifetime Using Semi-Empirical Model (month)	Measured Lifetime (month)
55 °C	12.5	12	13
47.5 °C	25	27.5	26
40 °C	53.5	64.5	/ [a]
25 °C	285.5	400.5	/

[a] The value is not available because the test did not reach the EOL.

2.4. SUMMARY

This chapter analyzes the long-term degradation behavior (i.e., beyond the typical range of 20% capacity fade) of $LiFePO_4$ batteries and studies the effect of aging conditions on the degradation behavior. It is found that the capacity fade shows a piecewise change with SOC variation. The traditional semi-empirical model based on the Arrhenius equations can only reflect the monotonic trend of capacity fade with temperature. A nonlinear regression lifetime model with two-step fitting is proposed that can simultaneously study the effect of storage temperature and SOC level on the rate of calendar degradation. Using the established battery degradation model, the calendar lifetime can be interpolated for different storage conditions. All the aging data introduced in this chapter will be used for ML model training and verification in the following chapters.

CHAPTER 3. ENTROPY AS FEATURE FOR MACHINE LEARNING SOH ESTIMATION

The content of this chapter is based on the results presented in [J3] and [J4].

This chapter introduces the methods to reach objective 2 and mainly answer the following questions:

- Why features extraction is needed in ML-based SOH estimation?
- What features are used in ML-based SOH estimation methods?
- What are the disadvantages of these existing features?
- Why entropy is selected as a feature?
- What's the performance of FE as a feature for SOH estimation?

3.1. BACKGROUND

The training of an ML model typically involves the following steps: data preparation (e.g., collecting data and suppressing the noise), feature dimensionality reduction (i.e., manually or automatically extracting and selecting features), and parameter optimization of the ML model [J1]. According to the algorithms used, two training modes are summarized, as presented in Fig. 3.1.

In order to simplify the computing process, the dimension of the measurements should be reduced first, and the features are extracted as input of the ML model. In order to increase the SOH estimation accuracy, the feature should include enough aging information, and the ease of feature extraction should be addressed in real-life applications [J4]. Various SOH features such as differential features, geometrical features, and statistical features are proposed in the existing publications. For example, through the differential calculation of voltage (i.e., dV/dt), current (i.e., IC curve) or temperature (i.e., dT/dt), some obvious SOH features such as the peak/valley value can be obtained [20, 51]. Geometrical features like knee points, extremum and curvature of voltage curves, as well as time associated with a fixed voltage interval are also powerful indicators of battery aging [J1, 25, 26]. However, extracting the aforementioned features manually brings some problems. For instance, extensive laboratory testing is required to obtain certain SOH features on the one hand. Some features, on the other hand, are reliant upon the test conditions, and they may become invalid when the temperature variation is considered. For these reasons, effective and robust SOH features should be proposed.

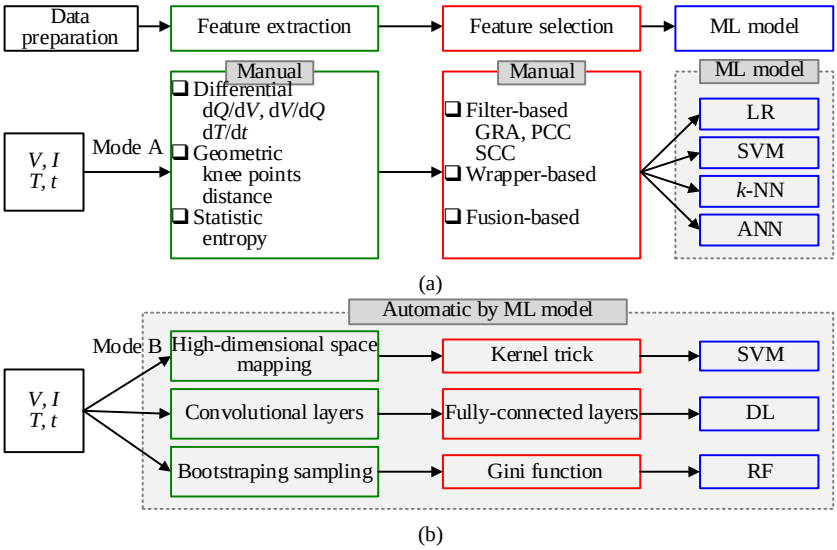


Fig. 3.1. Summary of ML model training modes: (a) Manual extraction and selection of SOH features (b) automatic extraction of SOH feature through the use of ML algorithms.

As an important category of SOH features, statistical features, namely the sample entropy and its enhanced forms, are used in this project. With a short pulse test, the difference in battery SOH can be determined by calculating entropy, a statistic indicator of signal complexity. The entropy has different variants, such as approximate entropy (AE), sample entropy (SE), fuzzy entropy (FE), and their multiscale form, i.e., multiscale entropy. In order to better apply entropy as a SOH feature, the theory behind these features is introduced and compared. Afterward, the effectiveness of the entropy feature is evaluated with regard to data set selection. Finally, FE is proposed as a robustness-enhanced SOH feature, and its effectiveness is validated by comparing with SE from the perspective of theory and performance.

3.2. ENTROPY THEORY

To determine the probability of producing a new pattern in a signal, AE was proposed [52, J3]. In AE(m, r, N), m (which is a positive integer) is the dimension of matrices, r (which is a positive real number) indicates the tolerance for accepting these matrices, and N specifies the total number of data. Moreover, SE is an effective way of quantifying the predictability degree of the time series, as well as their complexity [J3, 53]. It begins with the identical first two steps as the AE algorithm [J3] but does not compare the same vector to itself. SE is expressed as the negative natural logarithm of the conditional probability (CP). FE is the negative natural logarithm of the CP that a dataset of length N will repeat itself for m+1 points within a boundary after having repeated for m points [C3, 54]. Compared to SE, the similarity degree in FE calculation is computed by the exponential function instead of step function. As a result, FE can determine the data regularity more accurately. Furthermore, it is also

possible to compute the multiscale entropy using the coarse-grained procedure, which can indicate the degree of irregularity in a complicated time series [J3, 55, 56]. Fig. 3.2 shows the comparison of the AE, SE, FE, and multiscale entropy algorithms. The detailed steps of these algorithms are shown in Fig. E.1, Fig. E.2, Fig. E.3, and Fig. E.4, respectively (see Appendix E). The parameters m and r can be selected by minimizing the maximum sample entropy relative error, as presented in Fig. E.6 (see Appendix E.4) [J3, 57, 58].

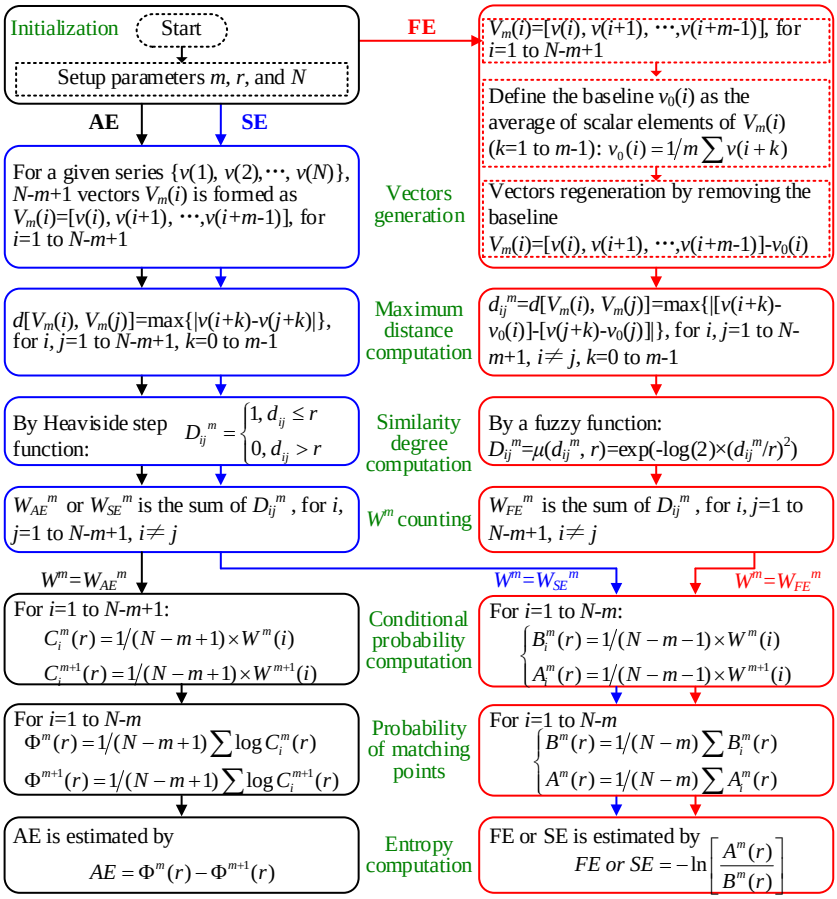


Fig. 3.2. The flowchart of the approximate entropy algorithm (black), the sample entropy algorithm(blue), and the fuzzy entropy algorithm (red). Source: [J3, J4].

3.3. DATASET SELECTION FOR ENTROPY CALCULATION

As voltage response differs under different TCs, it is imperative to evaluate the influence of voltage datasets [J3]. As illustrated in Appendix C.1, the cyclic aging test data are used in this chapter to validate the effectiveness of the proposed method. In order to generate different voltage profiles (see Fig. 3.3) and analyze their influence on the performance of entropy-based SOH estimation, the HPPC test, as illustrated Appendix C.3, can be divided into six unidirectional pulses at three SOC levels.

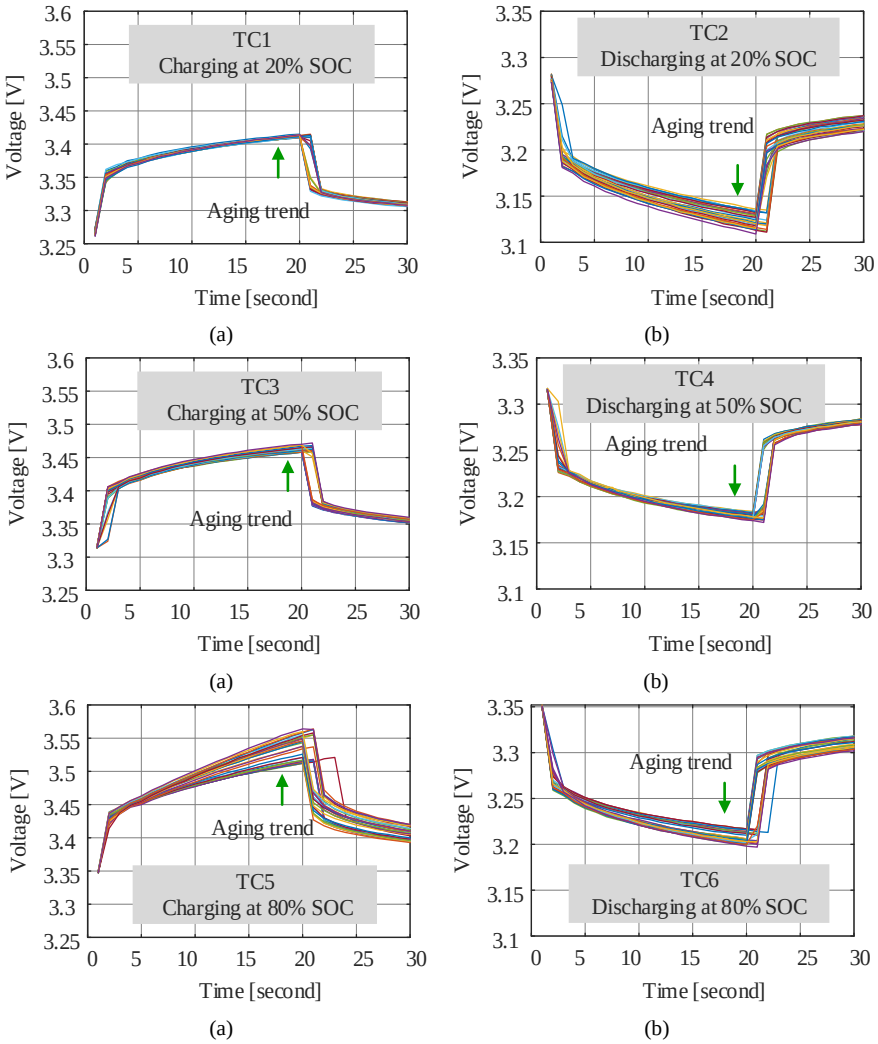


Fig. 3.3. Six voltage datasets acquired through HPPC tests under (a) TC1, (b) TC2, (c) TC3, (d) TC4, (e) TC5, and (f) TC6. Source: [J3].

The proposed dataset selection method is illustrated in Fig. 3.4. To begin, pulse voltage series under six TCs considering two impact factors (namely the current direction and SOC) are collected. Then AE and SE are chosen as SOH features, therefore the first-order polynomial model [53] can be established for mapping the features to SOH. By comparing the accuracy of the established estimators using different voltage datasets and different scales, the best suitable TC and scale can be identified, respectively.

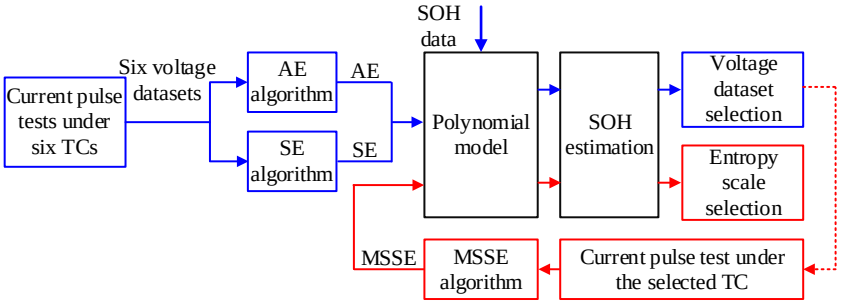


Fig. 3.4. An illustration of the proposed dataset selection strategy for entropy-based SOH estimation. Source: [J3].

3.3.1. DATASETS ANALYSIS

The classic terminal voltage characteristic curve of the $LiFePO_4$ battery is shown in Fig. 3.5. The slope of voltage under the 1C-rate constant current (CC) full charging test is also provided, and it is expressed as (3.1).

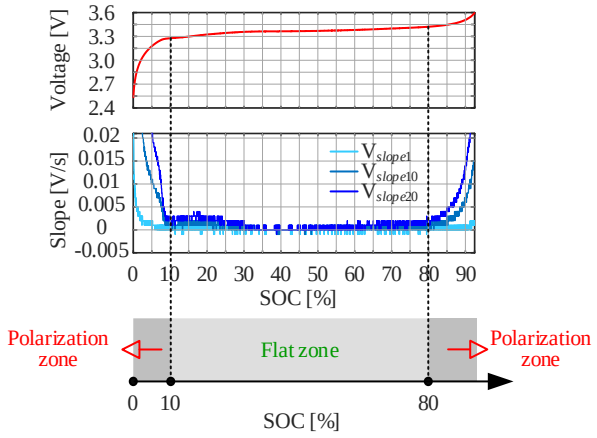


Fig. 3.5. Terminal voltage responses and the corresponding slope curve during 1C-rate CC charging. Source: [J3].

$$V_{slopei} = \frac{V(j+i)-V(j)}{t(j+i)-t(j)},\ i = 1, 10, 20 \tag{3.1}$$

where V_{slope1}, $V_{slope10}$, and $V_{slope20}$ are the slope corresponding to intervals of 1 s, 10 s, and 20 s, respectively. It can be seen from the slope curve that the polarization of the battery cell is severe at the polarization zone for which the SOC is either over 80% or less than 10%, while the variations in battery voltage at mid-SOC are flat (also known as the flat zone) [J3, 59]. Fig. 3.3 explains the test results that the voltage responses under TC2 and TC5 differ more obviously than those under other TCs.

3.3.2. ACCURACY COMPARISON

In the project, the root-mean-squared error (RMSE), the absolute percentage error (APE), and the mean absolute percentage error (MAPE) are the metrics used to evaluate the effectiveness of the proposed method [J4], defined as follows:

$$RMSE = \sqrt{\frac{1}{N_T}\sum_{i=1}^{N_T}\left(\hat{SOH}_i - SOH_i\right)^2} \tag{3.2}$$

$$APE_i = \frac{\left|\hat{SOH}_i - SOH_i\right|}{SOH_i} \times 100\% \tag{3.3}$$

$$MAPE = \frac{1}{N_T}\sum_{i=1}^{N_T}\left(\frac{\left|\hat{SOH}_i - SOH_i\right|}{SOH_i} \times 100\%\right) \tag{3.4}$$

where N_T denotes the total number of validation data, $\hat{SOH}_i$ and SOH_i are the estimated and real SOH of point *i*, respectively.

The results of the feature extraction and the SOH estimation are shown in Fig. 3.6, Fig. 3.7, and Fig. 3.8. According to these results, AE-based and SE-based estimators produce the best estimation under TC5 due to polarization. Likewise, from the perspective of the polarization, a low SOC value (e.g., smaller than 10%) can also yield improved performance. However, this case, where the battery is discharged to such a low SOC, is not common in the real operation of the batteries, therefore it is excluded from the experiment. Based on the findings of the experiments, it is concluded that the combined impacts of SOC level and current direction are responsible for the estimation accuracy. Feature extraction from the polarization zone helps improve the estimation performance [J3].

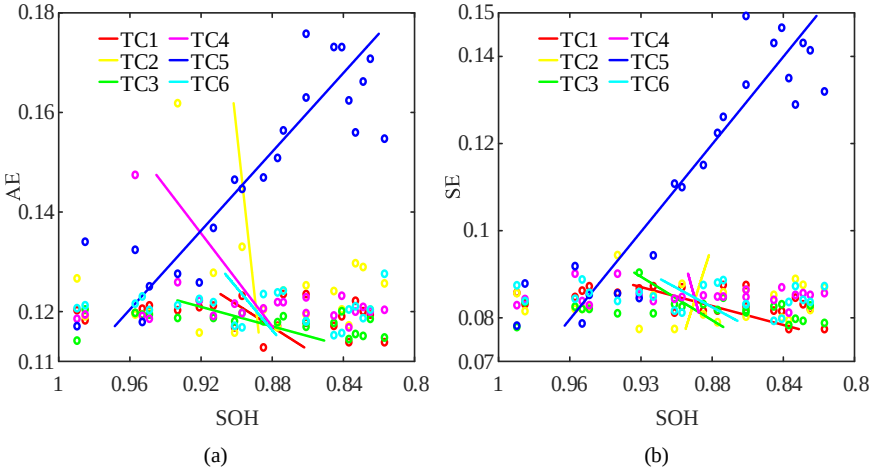
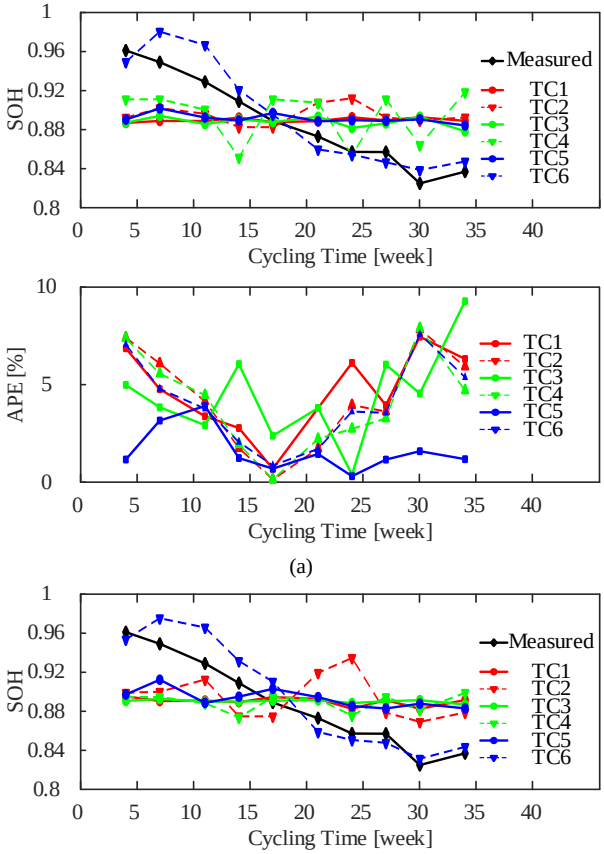
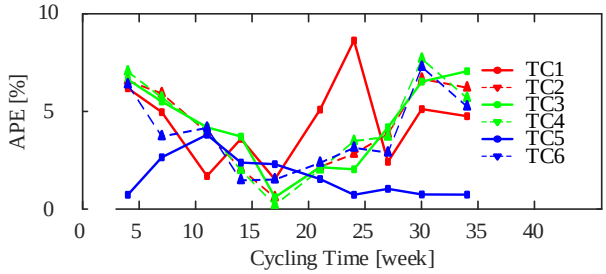


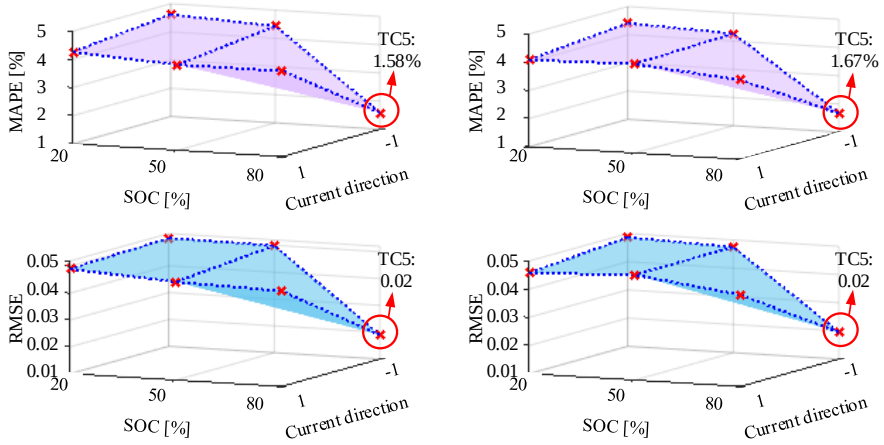
Fig. 3.6. The first-order fitting results of the relationship between SOH and (a) AE and (b) SE. Source: [J3].





(b)

Fig. 3.7. SOH estimation results under six TCs using (a) AE-based algorithm and (b) SE-based algorithm. Source: [J3].



(a) (b)

Fig. 3.8. The SOH estimation error using the datasets under different TCs (minus one and plus one denote the direction of charging and discharging, respectively): (a) AE-based method and (b) SE-based method. Source: [J3].

Next, how the scale affects the estimation performance of multiscale entropy-based (MSE-based) method was analyzed. A voltage dataset under TC5 is selected as a case study, and Fig. 3.9 shows the results of MSE (the scale varies from 1 to 3.) varies with SOH. As seen in Fig. 3.10 and TABLE 3.1, the estimation errors at scales 2 and 3 are large. The results show that it is no longer suitable to use MSE feature while using small samples for SOH estimation. Because in such situation, enlarging the scale factor will further filter the useful information, and eventually lead to a decrease in estimation accuracy.

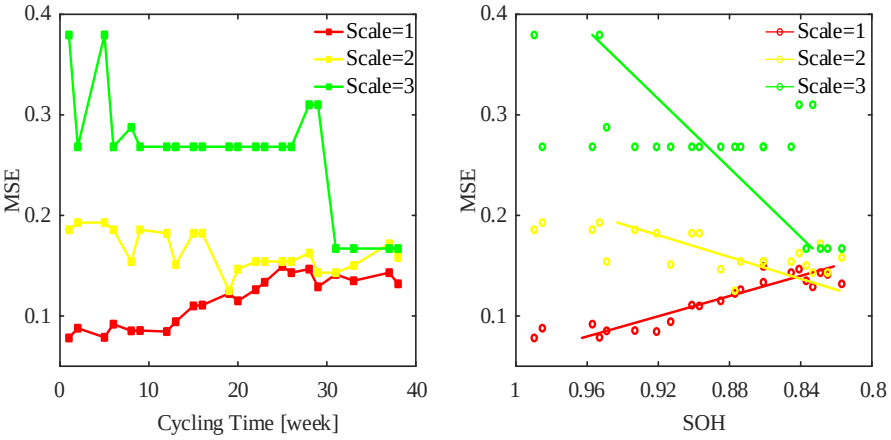


Fig. 3.9. The obtained MSE values with various scales: (a) MSE corresponding to different cycling time, and (b) the first-order fitting results of the relationship between SOH and MSE. Source: [J3].

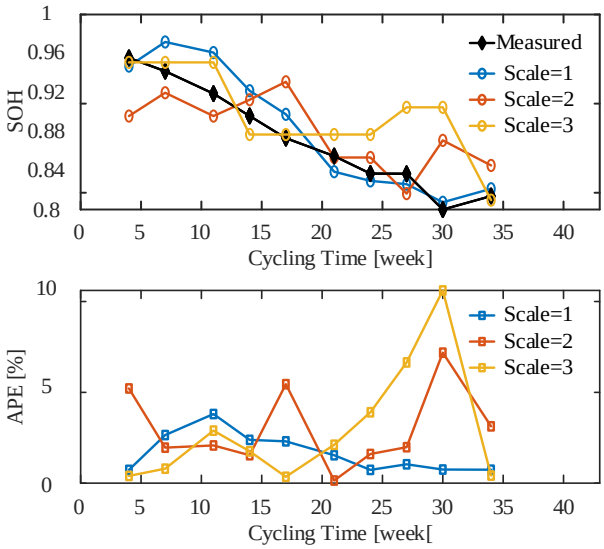


Fig. 3.10. SOH estimation results using the MSE-based algorithm. Source: [J3].

TABLE 3.1. The SOH estimation error using the MSE-based algorithm with various scales. Source: [J3].

Method	MAPE (%)	RMSE
Scale = 1	1.67	0.02
Scale = 2	3.03	0.04
Scale = 3	2.99	0.04

3.4. FE: ROBUSTNESS-ENHANCED SOH FEATURE

The previous subchapter selects the voltage datasets for entropy feature extraction, and the purpose is to improve the effectiveness of entropy features by defining the voltage interval. However, SE is highly sensitive to the parameter selection and noise because it uses the Heaviside step function in its computation of similarity degree [J4, 53]. To enhance the robustness of SOH estimation, FE feature is proposed and its performance is validated by conducting a comprehensive comparison from theoretical and experimental perspectives with the traditional method. The main content of [J4] is summarized as follows.

3.4.1. THEORETICAL COMPARISON

As shown in Fig. 3.2, FE has improved the SE algorithm in the two processes of vector generation and similarity degree calculation. On the one hand, by subtracting the mean value of the match templates, the shape of the vectors determines their similarity degree in FE. On the contrary, in SE, the similarity is calculated in absolute coordinates, which gives unreliable results. On the other hand, the exponential function is used in FE calculation as an alternative to Heaviside function used in SE. As shown in Fig. 3.11, FE is remarkable robust against data noise and TCs due to the use of fuzzy function, which greatly alleviate the sensitivity of SE to parameter variations. For more explanation, please refer to [J4]. The following work also highlights the effectiveness of FE.

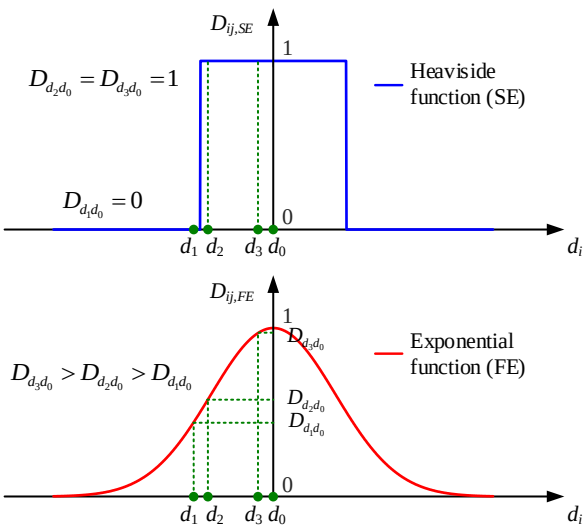


Fig. 3.11. The comparison of similarity calculation in SE and FE algorithms. Source: [J4].

3.4.2. PERFORMANCE COMPARISON

Using 38 weeks of cyclic aging data (as illustrated in Appendix C.1), this subchapter compares the performance of SE- and FE-based SOH estimation. Fig. 2.1(a) and Fig. 2.1(c) (see Chapter 2.2) illustrate the SOH curves and the evolution of the voltage responses under pulse test at 80% SOC. The self-validation approach is used, where the aging datasets of the tested battery are divided into a training group (28 of 38 feature-SOH pairs) and a validation group (the other 10 of 38 data pairs). The SVM model is selected as the target ML model, and four factors including the parameter *r* selection, noise, training data size, and TCs are examined for their effect on the estimation accuracy [J4].

A. Effect of parameter selection on estimation accuracy

To investigate the dependence of the estimation accuracy on the parameter *r*, *m* is set to 2 and *N* is 30. From the results (see Fig. 3.12 and Fig. 3.13) obtained by selecting different *r*, it can be seen that the variation of *r* has a more obvious influence on SE feature.

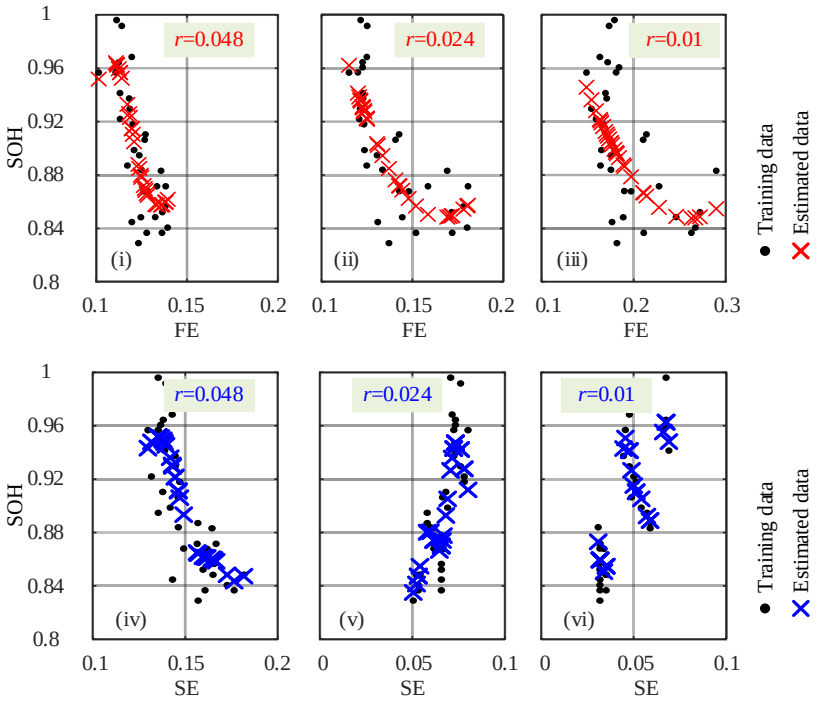


Fig. 3.12. SVM training results with different parameters *r* used in FE and SE algorithms. Source: [J4].

When *r* is set to 0.048, both methods provide good SOH estimation results with errors below 6%. With the decrease of *r*, the error of the SE method gradually increases and there are obvious fluctuations, until it fails when *r* is 0.01. On the contrary, FE is robust to parameter variation and can maintain a high estimation accuracy (i.e., lower than 4%).

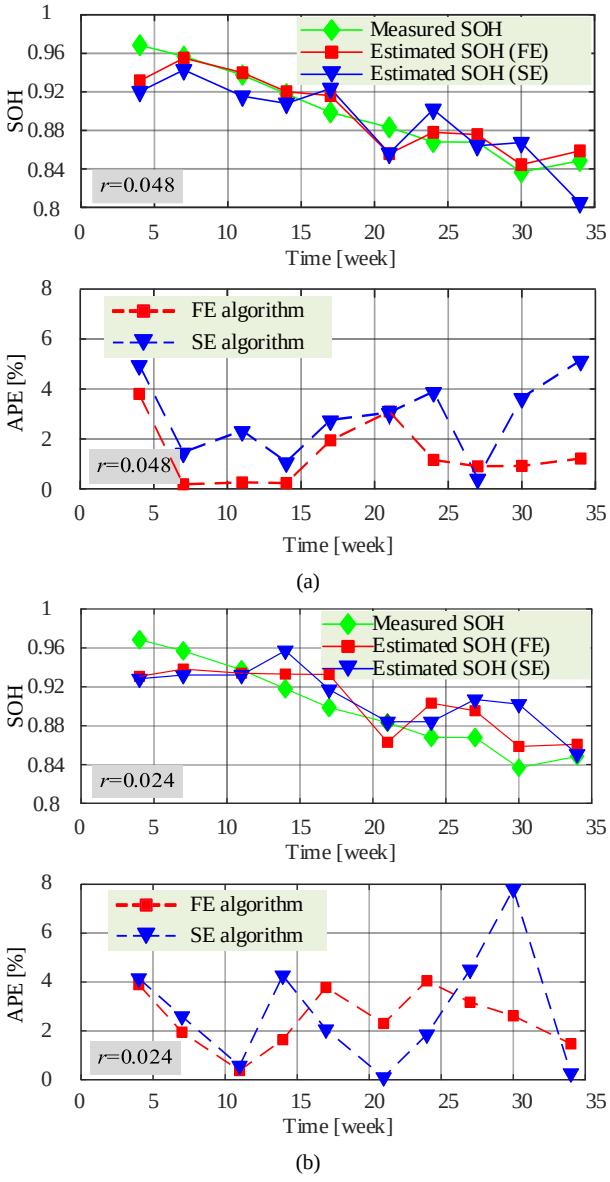


Fig. 3.13. SOH estimation results with different parameters *r* used in FE and SE algorithms: (a) *r*=0.048 and (b) *r*=0.024. Source: [J4].

B. Effect of noise on estimation accuracy

As a means of simulating voltage signals collected under real working conditions without a high-precision sensor, Gaussian noise with a signal-to-noise ratio of 50 dB is injected into original data. Here, *m* and *r* are set to 2 and 0.048 for both methods. The obtained noisy voltage, the corresponding entropy feature, and the comparison of the SOH estimation are shown in Fig. 3.14, Fig. 3.15, and Fig. 3.16, respectively. It can be observed that when the measured voltage is affected by noise pollution, FE is still an effective SOH feature. Conversely, the SE of the noisy data shows saltation and invalid for SOH estimation.

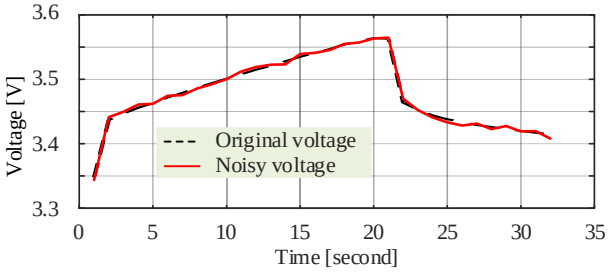


Fig. 3.14. Original voltage and noisy voltage used for entropy feature extraction. Source: [J4].

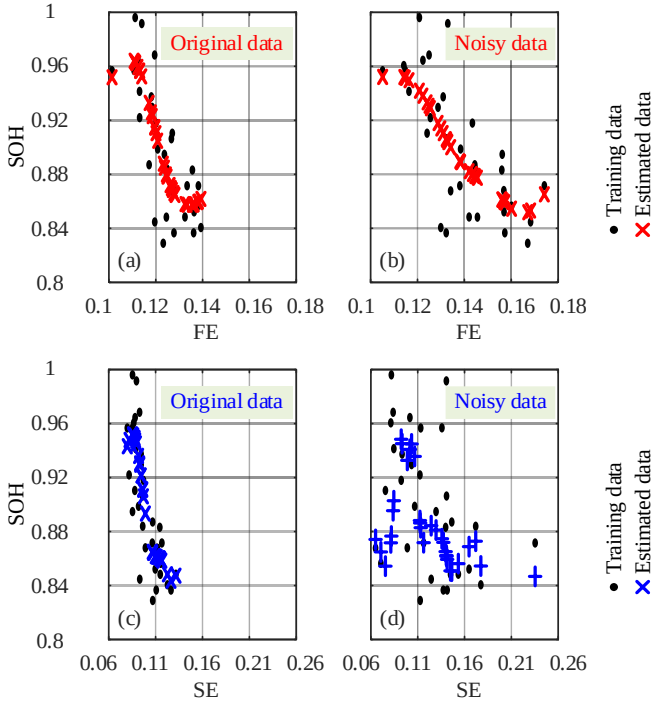


Fig. 3.15. The correspondence between SOH and different entropy features: (a) FE of original data, (b) FE of noisy data, (c) SE of original data, and SE of noisy data. Source: [J4].

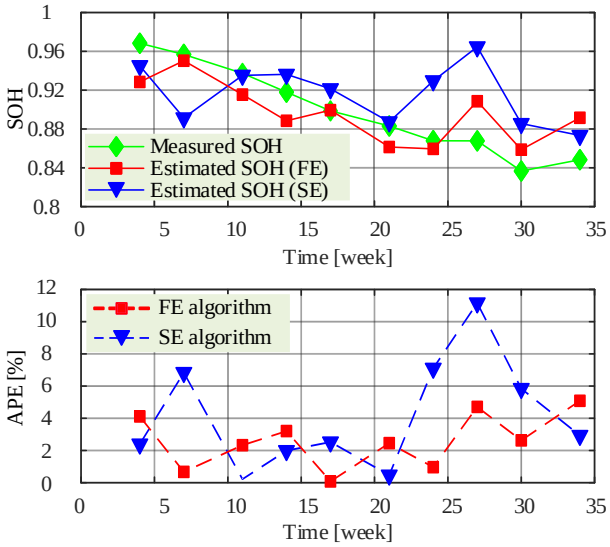
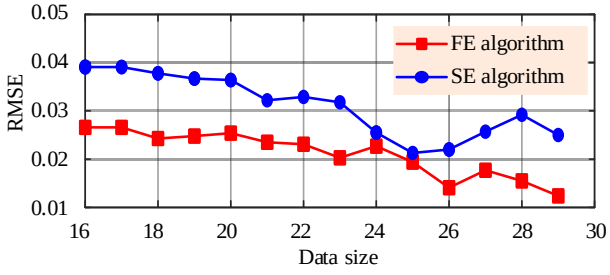


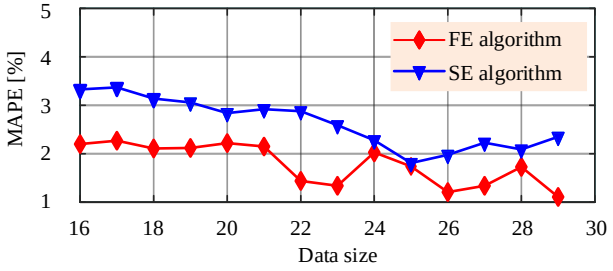
Fig. 3.16. SOH estimation results of FE-based and SE-based algorithms with Gaussian noise. Source: [J4].

C. Effect of data size on estimation accuracy

As suggested in part A in Chapter 3.4.2, *m* is still set to 2 while *r* is set to 0.048 for both methods in this part. By charging the training data size from 40% (i.e., 16 of 38 FE-SOH data pairs) to 75% (i.e., 29 of 38 FE-SOH data pairs) of the entire aging data, the dependency of the entropy-based SOH estimation on the training data size is studied. According to Fig. 3.17, FE-based method has higher estimation accuracy when the same amount of data is used for model training. In addition, for achieving the same estimation error, SE-based method requires 50% more training data than FE-based method. For example, 24 SE-SOH training data pairs are needed for SE-based method to make its RMSE less than 0.03. While in the case of using FE-based method, 16 FE-SOH training data is enough. It can be concluded that the FE-based method reduces the required data size effectively as compared with the traditional SE method, resulting in less effort for laboratory test [J4].



(a)

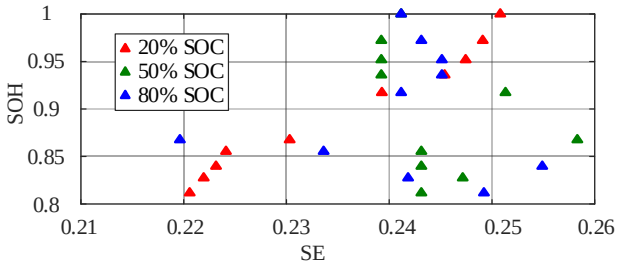


(b)

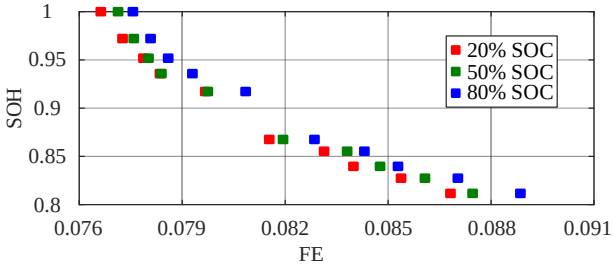
Fig. 3.17. SOH estimation error of FE-based and SE-based algorithms with different data size: (a) RMSE and (b) MAPE. Source: [J4].

D. Effect of SOC on estimation accuracy

To analyze the dependency of entropy-based SOH estimation method on the test condition, the voltage data which is used for SE and FE feature extraction is measured under three different SOC levels. The extracted entropy features under different test conditions are illustrated in Fig. 3.18. As shown in Fig. 3.18(a), the SE values obtained from 50% and 80% SOC levels are invalid in estimating the battery SOH. Moreover, the mapping of entropy to SOH obtained under 20% SOC is not consistent with the phenomenon revealed by entropy theory (i.e., the more disordered the data is, the more entropic it is considered). Contrarily, FE feature shows strong robustness against the test condition (i.e., the SOC level). As shown in Fig. 3.18(b), the FE value corresponding to a certain SOH point keeps almost constant with the varies of SOC. This shows that when using the FE feature to estimate the SOH, there is no need to accurately estimate the SOC in advance, which is very suitable for practical applications. This indicates that battery SOH estimation based on FE feature does not require the accurate estimation of SOC in advance. As a consequence, the suggested FE-based approach is applicable in real-world scenarios because the battery system does not need to be forced to reach a specified SOC before the application of a 30-second current pulse [J4].



(a)



(b)

Fig. 3.18. The correspondence between SOH and entropy features: (a) SE and (b) FE. (The features are extracted from the voltage under single pulse test at 47.5°C and different SOC levels). Source: [J4].

3.5. SUMMARY

This chapter studies the effectiveness of entropy as a feature for ML-based SOH estimation. To optimize the input voltage sequence and to guide the experimental test, the interaction between the TCs, entropy features (i.e., AE, SE, and MSE), and estimation accuracy are analyzed. It can be concluded that extracting the entropy feature from the polarization zone yields high SOH estimation. The proposed data sets selection method helps improve the ML-based battery SOH estimation in real applications. Additionally, FE is proposed as an effective feature and combined with SVM for SOH estimation of batteries. It is shown theoretically and experimentally that the FE-based method is superior in four aspects: parameter selection freedom, noise robustness, data size independence, and TC independence.

CHAPTER 4. ML-BASED SOH ESTIMATION USING FE

The content of this chapter is based on the results presented in [J3], [C3], and [C4].

This chapter introduces the methods to achieve objective 2 and 3, and mainly answer the following questions:

- How robust is the FE feature to temperature variation?
- What is the effect of data noise on SOH estimation using ML?
- How to improve SOH estimation performance through noise suppression or feature smoothing?

4.1. SOH ESTIMATION CONSIDERING TEMPERATURE VARIATION

As concluded in Chapter 3, SE is unable to represent the aging information accurately when the different TCs (i.e., the SOC level) is considered. In real applications, the temperature is regarded as an important disturbance variable and its effect on the entropy-based SOH estimation is evaluated. As shown in Fig. 4.1, the temperature is added to the SVM model as an independent variable. The calendar aging data under various temperature settings are used for training and validation of the FE-based method. More details can be referred to Appendix C.2. During the calendar aging test, Pulse1, Pulse 2, and Pulse 3 are performed at 20% SOC, 50% SOC, and 80% SOC, separately. As mentioned in Chapter 3, a short-term single pulse test is not enough for extracting effective SE features. Therefore, before performing the SOH estimation that considers the temperature variation, the minimum pulse time for FE/SE calculation is analyzed.

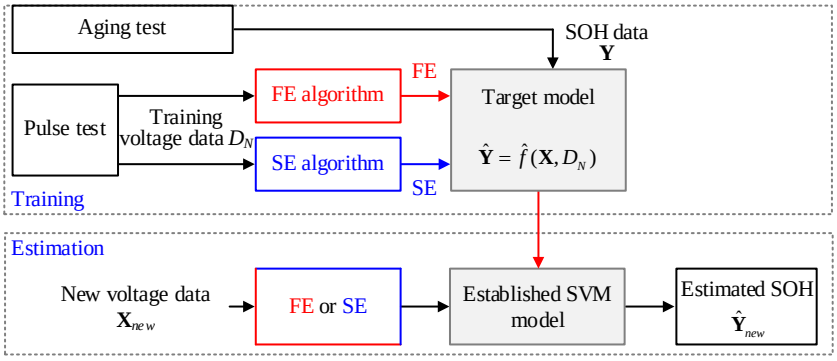


Fig. 4.1. Schematic diagram of the proposed FE-based SOH estimation method.

In Fig. 4.2, the SE and FE features of a single pulse test 80% SOC are displayed. These SE features do not accurately reflect the battery SOH. Meanwhile, FE is still an effective SOH feature, which shows a monotonous relationship with the SOH even in different temperatures. Certainly, it is obvious that the same result can be observed if a single pulse is applied at both 20% and 50% SOC. The SOH estimation results by using different single pulse test data and combined pulse test data are summarized in TABLE 4.1. It demonstrates that for SE-based method, a single pulse with 13-second of pulse current and 20-second rest time is too short to extract the aging information. However, a 33-second pulse test is sufficient for the FE-based method to provide a good SOH estimation (RMSE remains below 0.015).

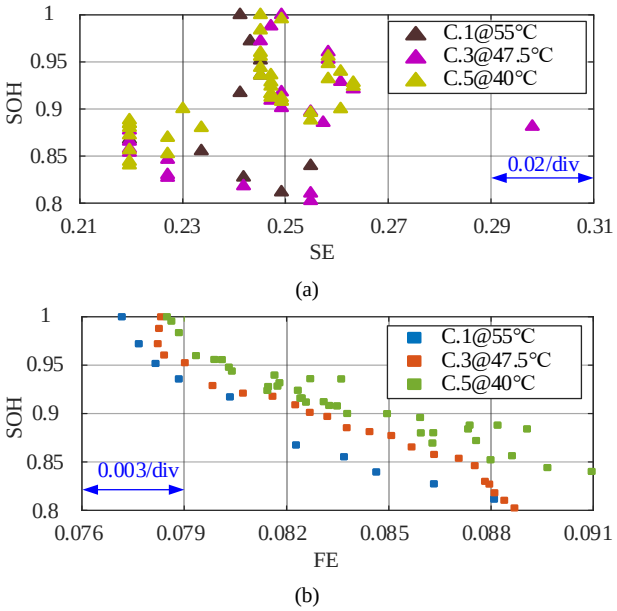


Fig. 4.2. The correspondence between SOH and entropy features: (a) SE and (b) FE. (Feature extraction is based on pulse voltage under 80% SOC and different temperatures). Source: [J4].

TABLE 4.1. Comparison of the training RMSE when extracting FE/SE features from various pulse voltages.

Feature	Battery cell	Pulse 1	Pulse 2	Pulse 3	Pulse 1+Pulse 2	Pulse 1+Pulse 2+Pulse 3
FE	C. 1@55℃	0.011	0.015	0.011	0.013	0.011
	C. 3@47.5℃	0.015	0.009	0.009	0.012	0.014
	C. 5@40℃	0.013	0.012	0.012	0.012	0.012
SE	C. 1@55℃	/ [a]	/	/	0.018	0.014
	C. 3@47.5℃	/	/	/	0.017	0.014
	C. 5@40℃	/	/	/	0.014	0.016

[a.] The SE extracted by the data under the specific condition is invalid for SOH estimation

Since the objective of this project is to demonstrate the robustness of entropy features against temperature variation, in the subsequent comparison, SE and FE features are extracted from the combined pulses voltage under three SOCs. Based on the parameter selection method (as illustrated in Appendix E.5), *m*, *r*, *N* were set to 2, 0.04, 99 for the FE algorithm and 2, 0.08, 99 for the SE algorithm [J4]. As it can be seen from Fig. 4.3, when the aging temperature changes from 40℃ to 55℃, the degrees of dispersion of the SE curve along the horizontal axis is ten times that of the FE curve. FE is, therefore, more robust to temperature variations than SE [J4].

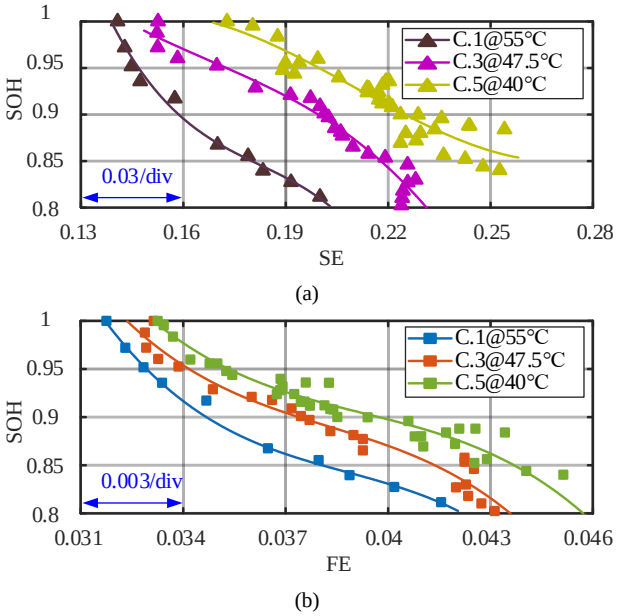


Fig. 4.3. The correspondence between SOH and entropy features: (a) SE and (b) FE. (Feature extraction is based on combined pulse voltages under three SOCs). Source: [J4].

4.1.1. SINGLE-TEMPERATURE MODEL

Fig. 4.4 and Fig. 4.5 illustrates the results of mutual validation for three single-temperature models. Overall when the validation data and training data come from the same temperature condition, both SE- and FE-based methods can achieve accurate estimation results. The main difference lies in the estimation towards low temperature (i.e., 40℃). By using FE-based method, the stability of the estimation result is enhanced. In another word, the FE-based method has stronger robustness against the temperature variation as compared to FE-based method [J4].

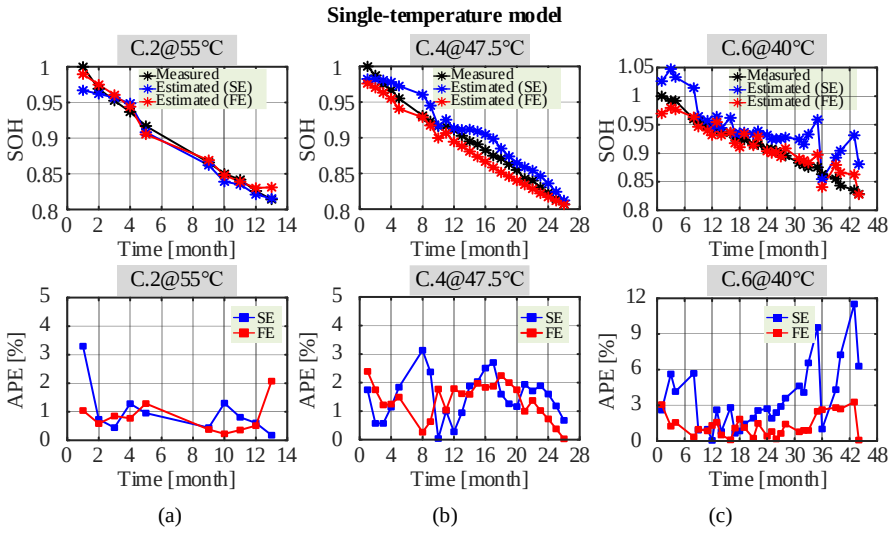


Fig. 4.4. SOH estimation results of single-temperature models: (a) C.2 at 55°C, (b) C.4 at 47.5°C, and (c) C.6 at 40°C. (Each model is trained using aging data at only one temperature, and verified using the aging data of another battery at the same temperature). Source: [J4].

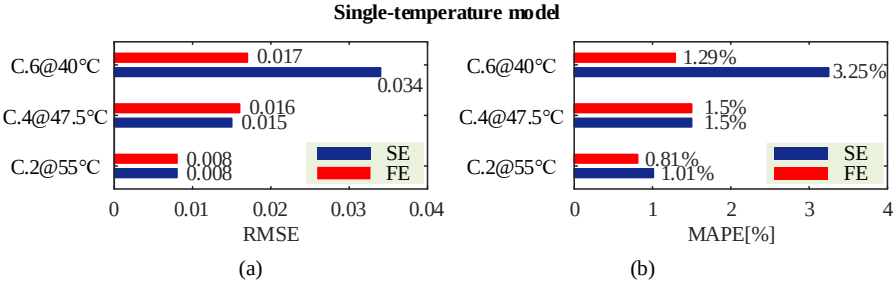


Fig. 4.5. SOH estimation errors of single-temperature models: (a) RMSE and (b) MAPE. Source: [J4].

4.1.2. FULL-TEMPERATURE MODEL

By using the aging data at all three temperatures, a full-temperature model is trained, and the estimation results for each model are illustrated in Fig. 4.6 and Fig. 4.7. It shows that the established full-temperature model based on the SE feature is more accurate in estimating SOH under the intermediate temperature condition. On the contrary, the estimation errors for the other two temperature conditions are relatively large, meanwhile, the estimation fluctuates significantly. However, such a phenomenon is alleviated by the FE-based method. On the one hand, the APE of FE-based SOH estimation keeps less than 3% (only one estimate of C.6 at 40°C is slightly larger than 3%, reaching 4%). On the other hand, there is no obvious deviation in RMSE and MAPE at the three temperatures. Based on the comparison results with

SE-based method, FE-based method shows stronger temperature and will result in relatively smaller estimation error regardless of the aging temperature. [J4].

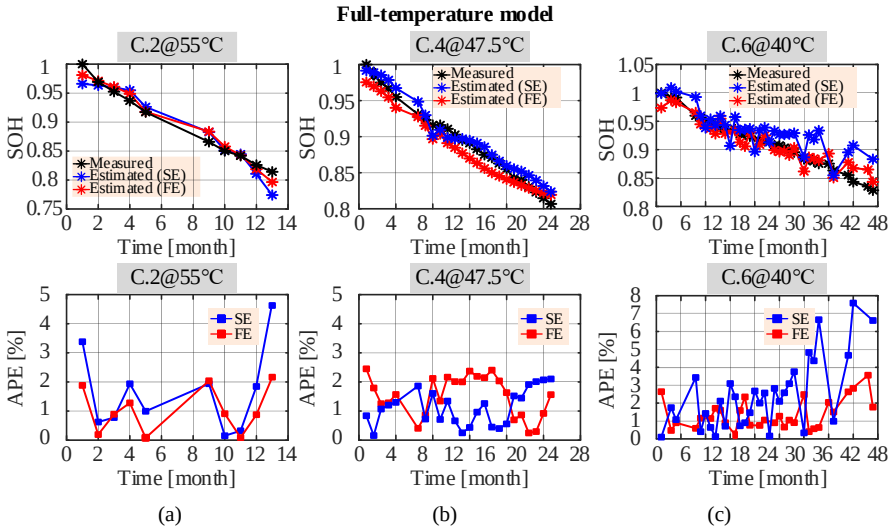


Fig. 4.6. SOH estimation results of full-temperature model: (a) C.2 at 55°C, (b) C.4 at 47.5°C, and (c) C.6 at 40°C. (The model is trained using the aging data of C.1@55°C, C.3@47.5°C, and C.5@40°C, and validated separately using the aging data of C.2@55°C, C.4@47.5°C, and C.6@40°C). Source: [J4].

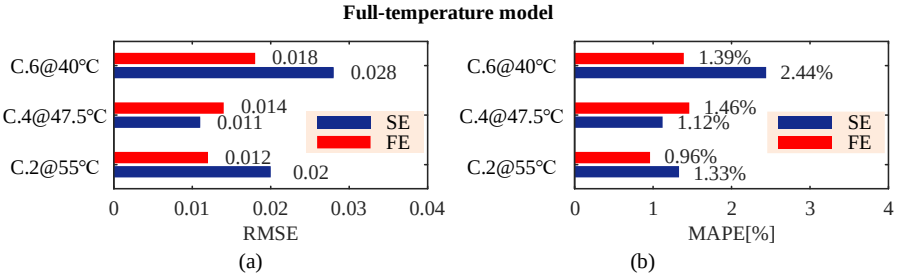


Fig. 4.7. SOH estimation errors of full-temperature model: (a) RMSE and (b) MAPE. Source: [J4].

4.1.3. PARTIAL-TEMPERATURE MODEL

The partial-temperature model is established by excluding the aging data at the intermediate temperature (using it as the validation data). As shown in Fig. 4.8 and Fig. 4.9, FE-based method provides a higher accuracy for most points, and MAPE and RMSE are reduced when compared to the SE-based method. This means the FE-based method involves a smaller amount of training data than the SE-based method [J4].

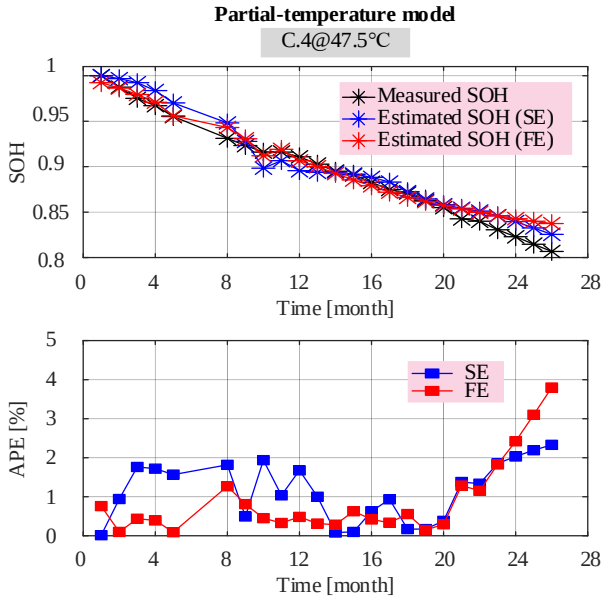


Fig. 4.8. SOH estimation results of partial-temperature model. (The model is trained using the aging data of C.1@55°C and C.5@40°C, and validated using the aging data of C.4@47.5°C). Source: [J4].

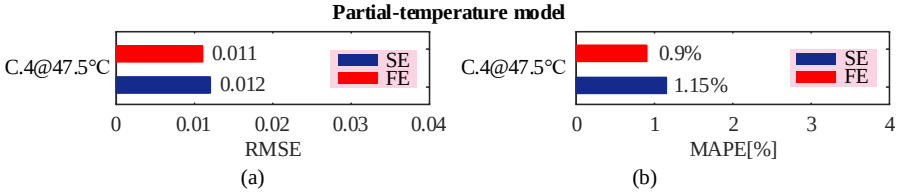


Fig. 4.9. SOH estimation errors of partial-temperature model for C.4@47.5°C: (a) RMSE and (b) MAPE. Source: [J4].

4.2. SOH ESTIMATION USING FURTHER IMPROVED FE

The entropy-based SOH estimation methods can be further improved by removing the noise from the original data and/or from the feature [60]. This Ph.D. project evaluated and compared two types of widely used data noise suppression methods, i.e., the adaptive iterative algorithm represented by empirical mode decomposition (EMD) and six regression-based smoothers. They are used before and after entropy feature extraction, respectively.

4.2.1. NOISE SUPPRESSION BEFORE FEATURE EXTRACTION

EMD was proposed by N.E. Huang et al. to decompose a time series into the sum of a finite number of intrinsic mode functions (IMF) [C3, 61]. The advantage of EMD technology is that it decomposes the signal from the data itself, so it can adaptively

separate the signal from the noise. The detailed steps of the EMD algorithm are illustrated in Fig. F.1 (see Appendix F.1). A battery SOH estimation method using EMD entropy is proposed, as illustrated in Fig. 4.10. The EMD algorithm is utilized to remove the noises in the original voltage data. Based on the decomposed voltage data, the SE feature is calculated for SOH estimation model training. Finally, SVM is used to obtain the optimal model by taking advantage of its ability in solving high-dimensional nonlinear problems.

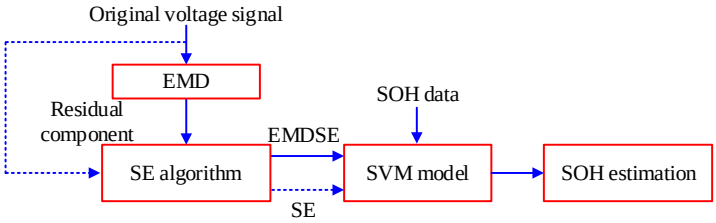


Fig. 4.10. Schematic diagram showing the noise suppression method based on EMD. Source: [C4].

The cyclic aging data under dynamic operation conditions, as illustrated in Appendix C.1, is used for features extraction and SVM model establishing. The obtained SOH curves and the CC discharging voltage of the two tested batteries are plotted in Fig. 4.11 and Fig. 4.12, respectively. The last 360 points in the CC discharge curve are selected to form an original voltage signal [C3]. The residual component (as shown in Fig. 5.) is obtained by processing the original signal through EMD. Correspondingly, based on these two sets of voltage signals, the SE/EMDSE features are obtained. In this case study, the parameters *m*, *r*, and *N* of the SE algorithm are selected as 2, 0.048, and 360 based [C3].

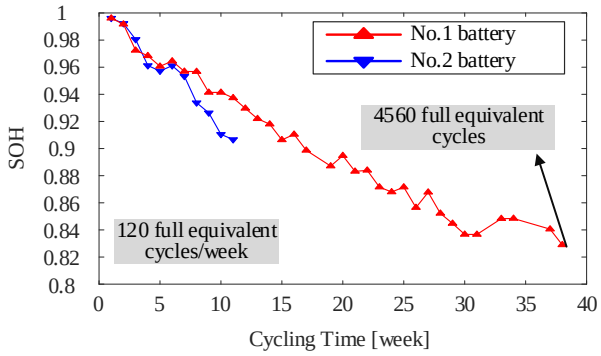


Fig. 4.11. SOH curve of two batteries under cyclic aging. Source: [C4].

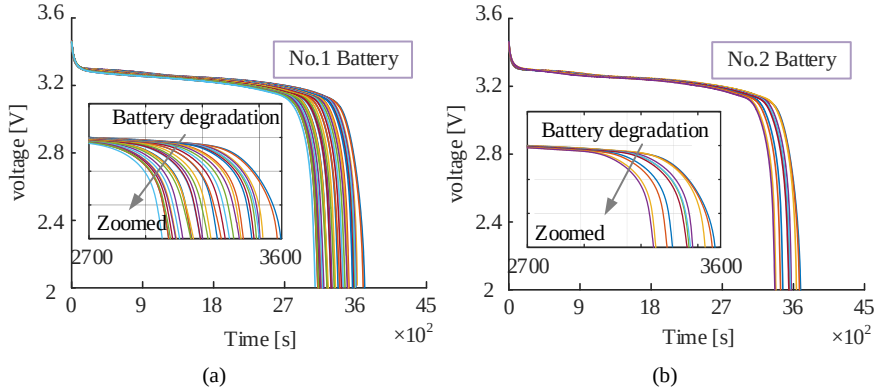


Fig. 4.12. Voltage responses during the CC discharging: (a) No.1 Battery and (b) No.2 Battery. Source: [C4].

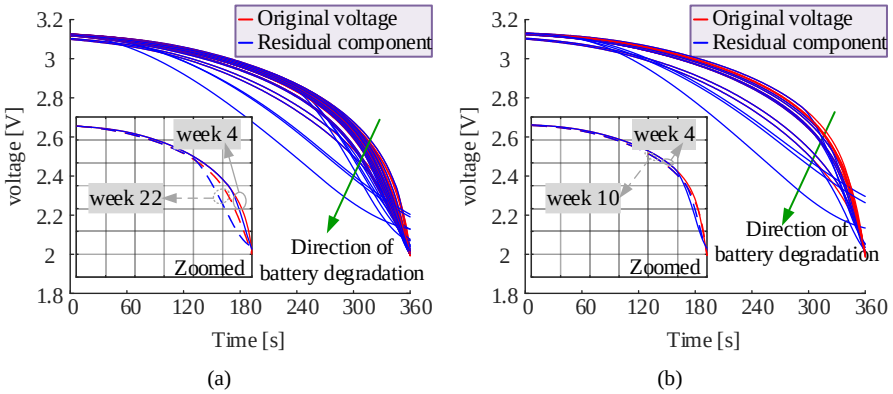
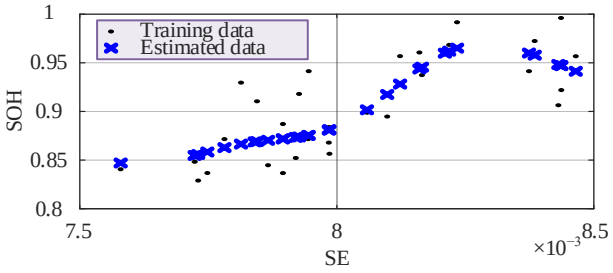
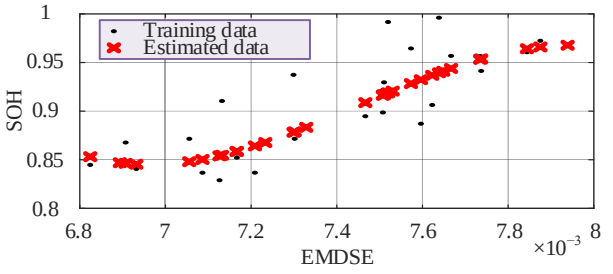


Fig. 4.13. Original voltage curve and the residual component obtained by EMD: (a) No.1 Battery and (b) No.2 Battery. Source: [C4].

Fig. 4.14 shows the SVM training results of SE and EMDSE methods. The more linear mapping between EMDSE and SOH indicates that EMDSE-based model is more accurate. The results of self-validation and mutual validation are shown in Fig. 4.15. The errors of these two methods are compared in Fig. 4.16. It can be seen that the EMDSE method improves the SOH estimation to a different extent for both validation methods. In summary, the EMD method can be applied to effectively reduce noise in the original voltage signal, thus improving the estimation accuracy [C3].

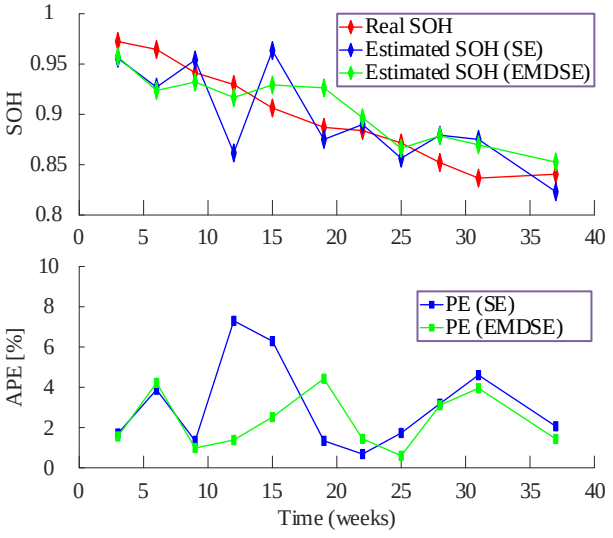


(a)

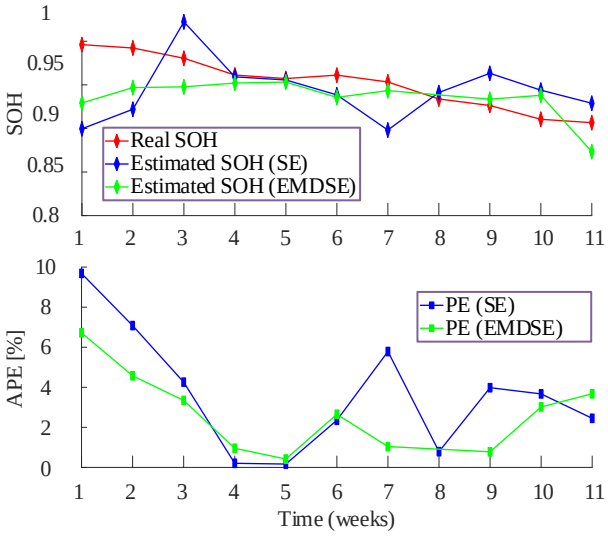


(b)

Fig. 4.14. SVM model training results using: (a) SE and (b) EMDSE feature. Source: [C3].

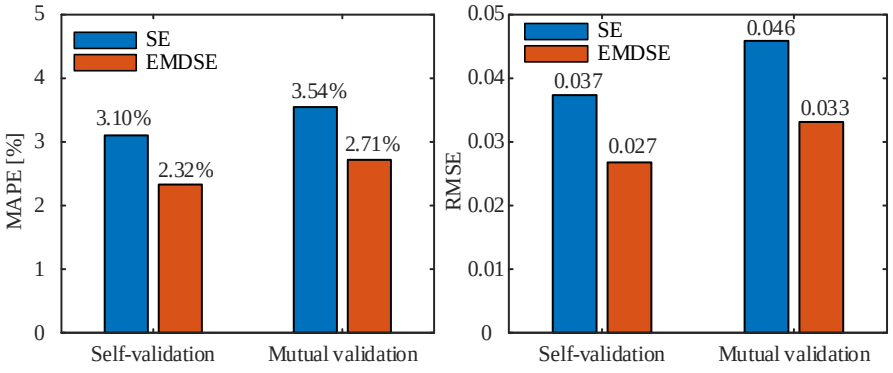


(a)



(b)

Fig. 4.15. SOH estimation results using (a) self-validation and (b) mutual validation. Source: [C3].



(a) (b)

Fig. 4.16. SOH estimation error using different SOH features: (a) MAPE and (b) RMSE. Source: [C3].

4.2.2. NOISE SUPPRESSION AFTER FEATURE EXTRACTION

In addition to the adaptive iterative algorithm (such as EMD), the regression-based smoothers are also very effective for noise suppression [62-64]. The previous subchapter studies the effect of noise suppression before SOH feature extraction, while in this part, the extracted features are smoothed and its effectiveness on the improvement of the SOH estimation is evaluated. In Fig. 4.17, an overview of the proposed analyzing method is given.

Two batteries' calendar aging data under 40°C, as illustrated in Appendix C.2, are used for features calculation. After the smoothing step, the SVM and GPR models are trained to map the smoothed feature to SOH. Finally, the effectiveness of noise suppression on the performance of SOH estimation is verified by experimental results. The SOH curve, the collected voltage data, and the extracted FE features for the tested two batteries are shown in Fig. 4.18. The mutual validation method is used to assess the model's performance. Training and validation data are taken from the No. 1 and No. 2 battery separately.

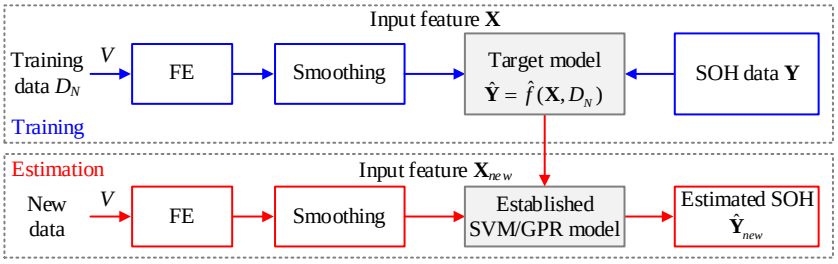
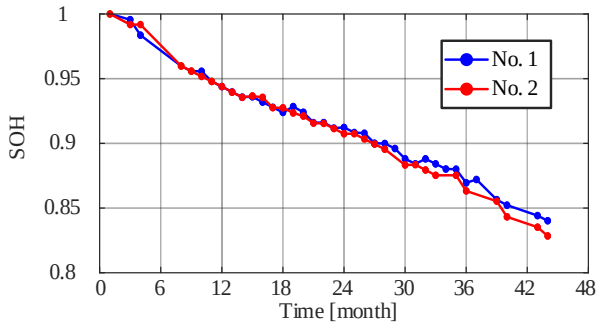
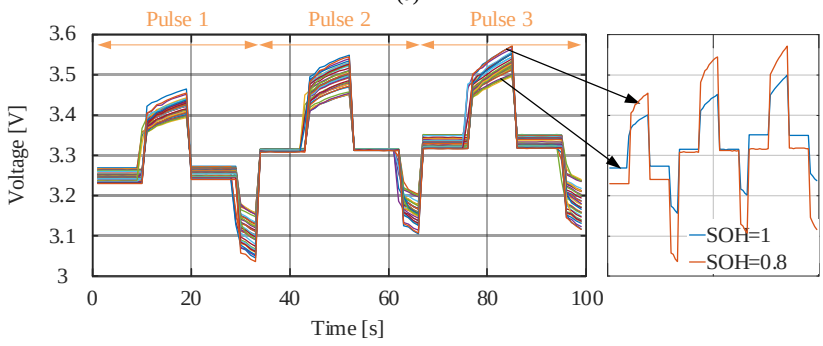


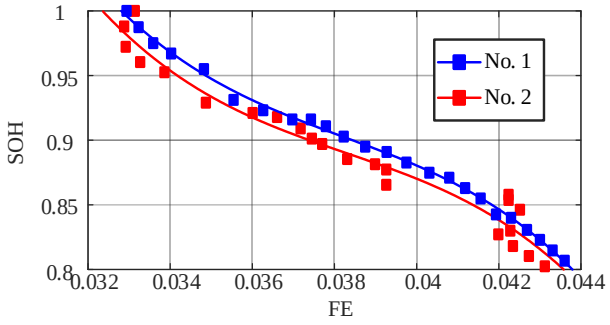
Fig. 4.17. Schematic diagram of the proposed feature smoothing algorithm. Source: [C4].



(a)



(b)



(c)

Fig. 4.18. Calendar aging results: (a) SOH curves at 40°C, (b) voltage responses during the HPPC test, and (c) the correspondence between SOH and FE. Source: [C4].

The SOH estimation results of the SVM model and the comparison of the performance in terms of MAPE, RMSE, R-squared, and simulation time are shown in Fig. 4.19 and Fig. 4.20. Likewise, the results of the GPR model are shown in Fig. 4.21 and Fig. 4.22. It can be seen from the results that for no matter which ML model without data smoothing, the estimation results show a relatively large fluctuation. Especially at the fifteenth month, the outlier shows up. Besides, the estimation error gradually increases at the end period of degradation. However, the smoothing methods have different degrees of improvement for both SVM and GPR models. Among the six smoothing methods considered, the lowess method is the most effective for performance improvement. Specifically, the MAPE of SVM model is reduced by the smoothing method from 0.9% to around 0.6%, while the RMSE is reduced from 0.016 to around 0.006. At the same time, the t_{sim} for SVM model is shortened from 9.56s to less than 3s. It is also worth mentioning that smoothing methods improve the performance of GPR more significantly.

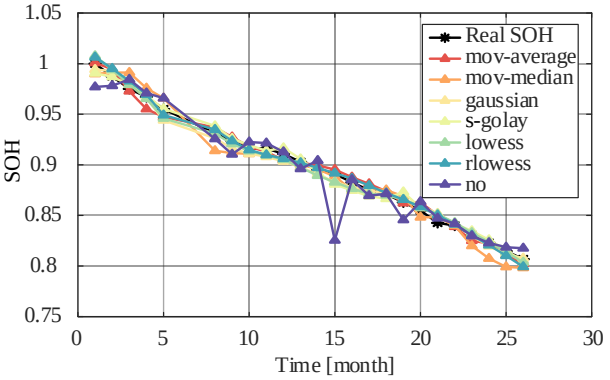


Fig. 4.19. SOH estimation results using different smoothing methods (when using SVM model). Source: [C4].

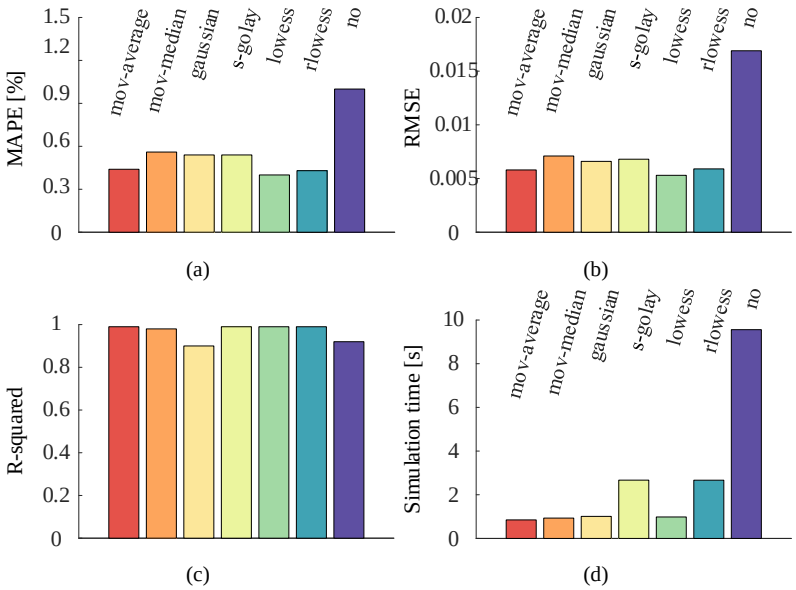


Fig. 4.20. Comparison of SOH estimation results using different smoothing methods (when using SVM model): (a) MAPE, (b) RMSE, (c) R-squared, and (d) Simulation time t_{sim}. Source: [C4].

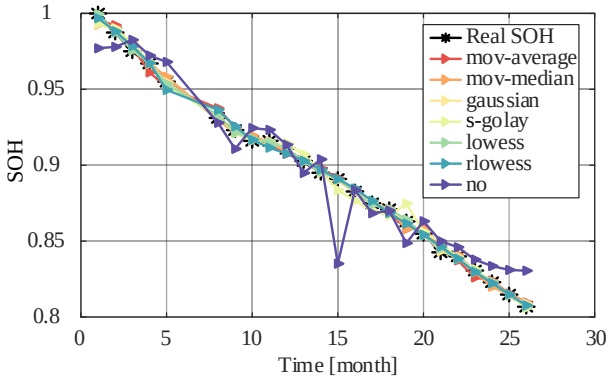
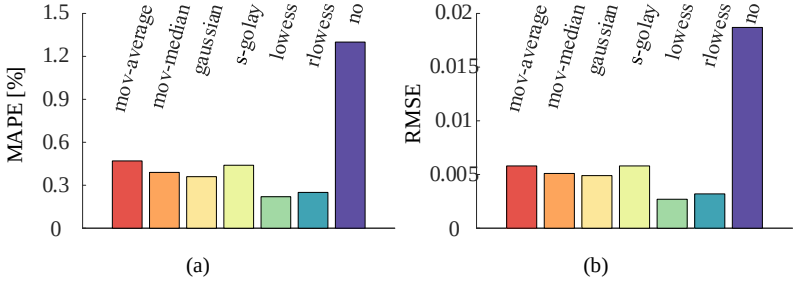


Fig. 4.21. SOH estimation results using different smoothing methods (when using GPR model). Source: [C4].



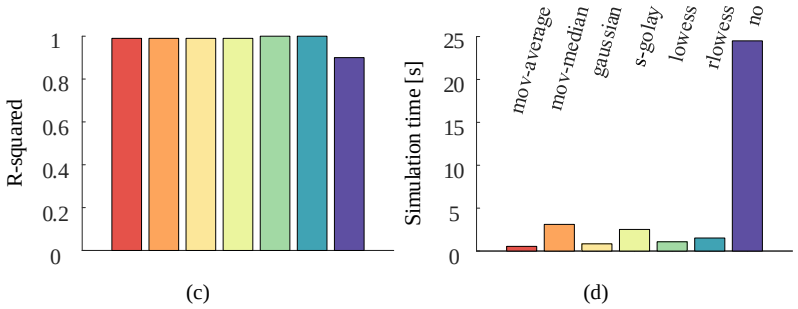


Fig. 4.22. Comparison of SOH estimation results using different smoothing methods (when using GPR model): (a) MAPE, (b) RMSE, (c) R-squared, and (d) Simulation time t_{sim}. Source: [C4].

4.3. SUMMARY

In this chapter, SOH estimation using FE as a feature is presented. Three SOH estimation models considering the aging temperature are established, and the results demonstrate the robustness of the FE feature against the aging conditions. For further improving the estimation performance, the EMD and six regression-based noise suppression methods are applied to proceed with the raw data and the extracted feature, separately. The experimental results prove that using noise suppression in ML improves accuracy as well as simulation speed.

CHAPTER 5. ML-BASED SOH ESTIMATION WITH AUTOMATIC FEATURE EXTRACTION

The content of this chapter is based on the results presented in [C5].

This chapter introduces the methods, which were used, to achieve objective 4, and mainly answer the following questions:

- How to reduce or even avoid the manual feature extraction process?
- What are the challenges of DL in battery SOH estimation?
- What is the EL method?
- Why use EL as an alternative to DL?

5.1. BACKGROUND

As introduced in Chapter 3, there are two training modes in ML. The previous chapters are dedicated to enhancing the robustness of features for ML methods that need manual feature extraction. Developing ML algorithms that can automatically derive SOH features is another way to avoid the failure of the feature. In other words, the measurement data can be used directly without dimensionality reduction. According to [J1], DL is a good candidate for SOH estimation since it has obvious advantages over other ML algorithms given a large amount of data. Using deep neural networks, for instance, global features can be derived from input with the help of multiple hidden layers. Convolutional neural networks are good at handling multi-dimensional data due to the use of convolutional techniques; thus the automatic feature extraction is realized. However, DL has the drawbacks of heavy computational burden and dependence on large datasets. When big data is not available, the overfitting becomes an issue. Amongst all ML methods, EL is emerging and it provides an alternative to DL, offering a trade-off between data size and accuracy.

EL refers to combing the results of multiple base learners, which is generally produce more accurate and robust results than a single base learner [J1]. Ensemble can take two forms: model- and data-level ensemble, as illustrated in Fig. 5.1. In the model-level ensemble, multiple different ML models are trained using the same data set. Based on a weighted average of base learners or through voting, the final output is given [66-69]. Obviously, training multiple ML models will increase the computational complexity. Except for model-level ensemble, a single base learning algorithm can also be used to set up the data-level EL method. Bagging and boosting can be used to resample the original data, allowing multiple homogeneous models to be built simultaneously [71, 72]. It should be noted that generating multiple subsets

only increases the diversity of the data, but does not enlarge the amount of data. Therefore, if complex algorithm, such as random forest, is selected as the base learner, overfitting will still be a problem. According to the above analysis, the data-level ensemble of simple base learners is a good solution. The bagging, also called bootstrap aggregating is selected because it can 1) avoid overfitting by reducing variance in the predicted outcome; 2) increase the computation slightly.

In this chapter, the methodology of the proposed BaggELM is presented first, followed by the hyper-parameters optimization, performance comparison with the traditional method, and experimental results. The summary of this chapter is given at last.

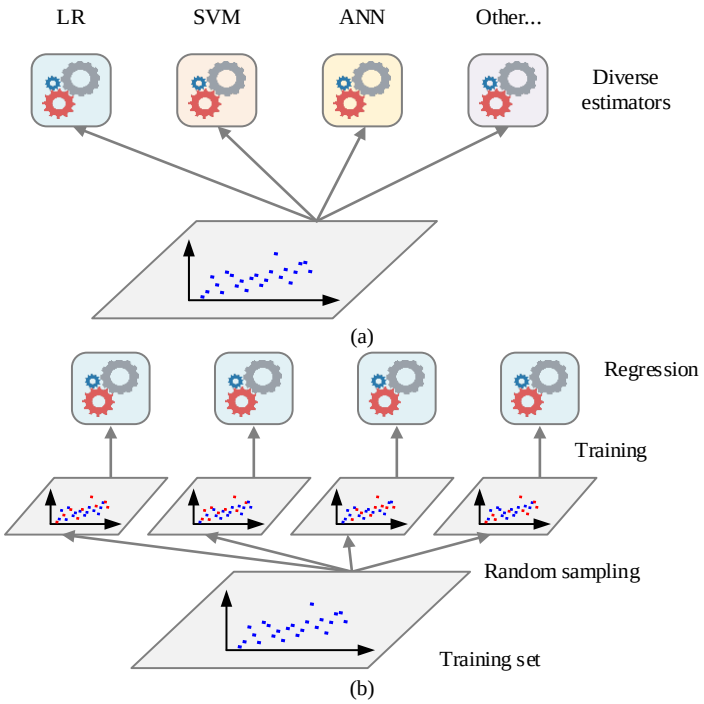


Fig. 5.1. Schematic diagram of two ways of ensemble in ML algorithm: (a) model-level ensemble and data-level ensemble (b). Source: [J1, 65].

5.2. SOH ESTIMATION USING ENSEMBLE LEARNING

Amongst commonly used ML algorithms, due to its computationally efficient, ELM is regarded as a promising method [73]. It requires only the predefinition of the network architecture. Compared to the algorithms whose weights are trained in an iterative way, the ELM method contains fewer trainable weights, resulting in low stability and accuracy of the ELM method. However, ELM can be improved by increasing the diversity of the datasets. Bagging allows the reconstruction of the original dataset on the basis of reducing subsets correlation [74]. Therefore, this

project applied bagging-based EL method to strengthen the generalization performance of ELM, following the structure presented in Fig. 5.2.

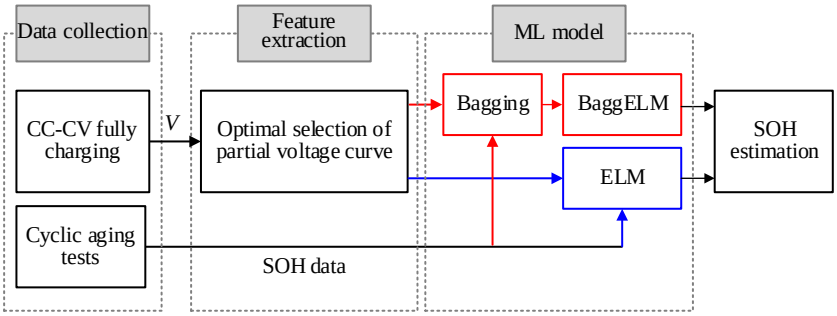


Fig. 5.2. The training process of the proposed EL algorithm. Source: [C5].

5.2.1. METHODOLOGY

In a typical feed-forward neural network, the parameters of the network will be optimized iteratively using for example gradient descent-based methods. The optimization is however a slow process, and parameters may gravitate toward the local optimum easily. Huang et al. [75] proposed ELM as a possible solution to this challenge. The detailed methodology of ELM are shown in Fig. D.6 and Fig. D.7 (see Appendix D.4). The output weights of ELM can be estimated through a pseudoinverse operation based on the randomly selected input weights and hidden biases. In contrast to iterative training, ELM requires much less computation and trains faster [J1].

As illustrated in Fig. 5.3, a bagging-based ELM ensemble is developed, which incorporates data resampling and model ensemble. Bagging is a process of generating B bootstrap samples, D_N^b, $b = 1, 2, \ldots, B$ based on random sampling with replacement over original training data, D_N [J1]. As a result, B ELMs can be trained and the estimated value can be obtained as

$$\hat{Y}_{new} = \frac{1}{B}\sum_{b=1}^{B} \hat{f}_b(\mathbf{x}_{new}) \tag{5.1}$$

where $\mathbf{x}_{\text{new}}$ denotes a new observation, $f_b(\cdot)$ is the bth trained ELM model, and B is the number of bagging.

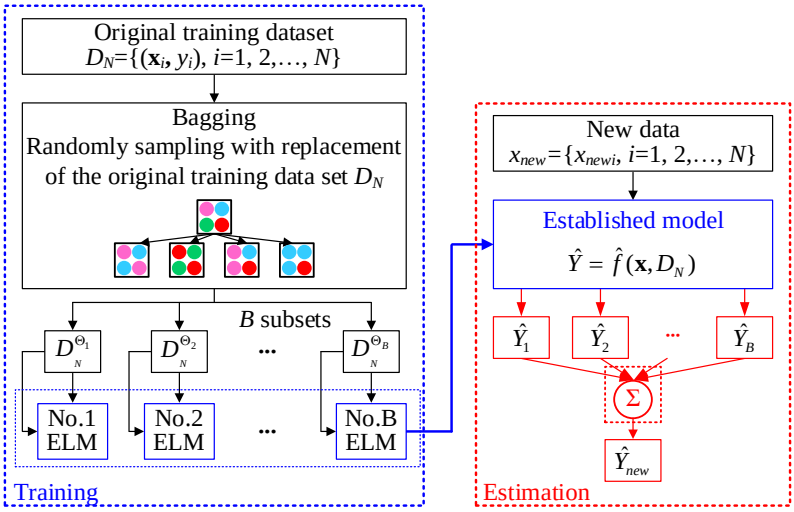


Fig. 5.3. Flowchart of the proposed BaggELM algorithm. Source: [C5].

5.2.2. EXPERIMENTAL TESTS

As presented in Appendix C.1, cyclic aging data of two $LiFePO_4$ batteries are used to evaluate the performance of the BaggELM method. Fig. 5.4 and Fig. 5.5 show the test results. It is not common for the battery to be fully charged or discharged in actual use, so the training data in this study are taken from the partial voltage curve. The details of input voltage selection are described in the following sub-chapter. Both self-validation and mutual validation are applied to verify the proposed method.

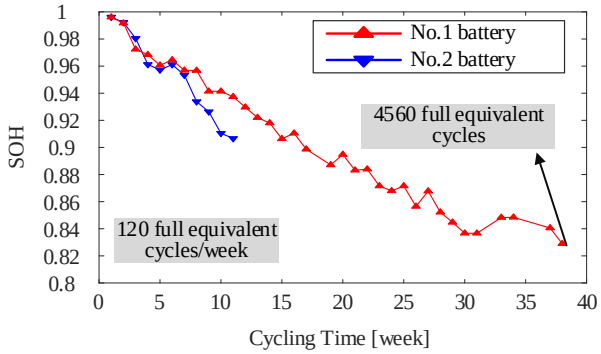


Fig. 5.4. SOH curves obtained from cyclic aging. Source: [C5].

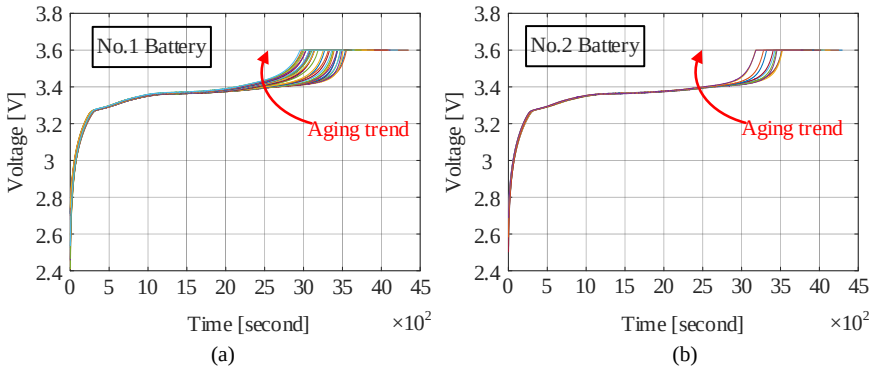


Fig. 5.5. Voltage responses under CC discharging: (a) No.1 and (b) No.2 Battery. Source: [C5].

5.2.3. HYPERPARAMETER OPTIMIZATION

According to the developed BaggELM method, several hyperparameters need to be optimized, including the length of voltage sequence ($length_V$), the starting point (V_{start}) of voltage sequence, the hidden neurons of the neural network (M), and the numbers of the bootstrap samples (B) [C5]. As depicted in Fig. 5.6, hyperparameter optimization is carried out. The input voltage was chosen from partial voltage curve for the following two reasons. Firstly, batteries will typically operate in a partial (between 10% and 90% SOC or even a narrower SOC range) rather than full SOC range [C5]. As a result, it's possible that the entire voltage curve won't be visible in real-world applications. Secondly, using a partial curve as input can shorten measurement time as well as simplify the computation.

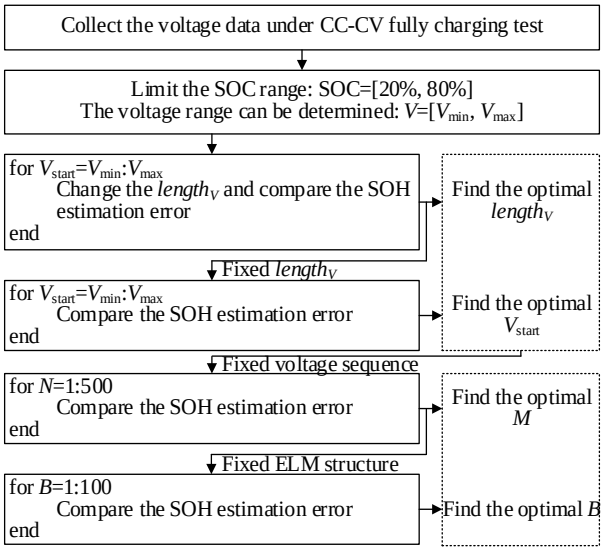


Fig. 5.6. Flowchart of parameters selection. Source: [C5].

Fig. 5.7 shows (a) current, (b) voltage, as well as (c) SOC curves of a fresh battery when it is being charged with CC-CV condition. When SOC varies from 20% to 80%, the corresponding voltage changes between 3.32V and 3.42V. Voltage sample points were taken every 2s and a set of voltage sequences was obtained to train BaggELM models. Fig. 5.8 shows the self-validation results observed for the No. 1 battery and shows how $length_V$ influences the accuracy of the SOH estimation. When $length_V$ is above 600, the proposed method is able to maintain a relative low error (i.e., RMSE and MAPE are around 0.01 and 1%, respectively). Fig. 5.9 presents the estimation results of $length_V$ from 600 to 850 for comparison to obtain the optimal value of $length_V$ [C5]. According to the results, the optimal $length_V$ was determined to be 750.

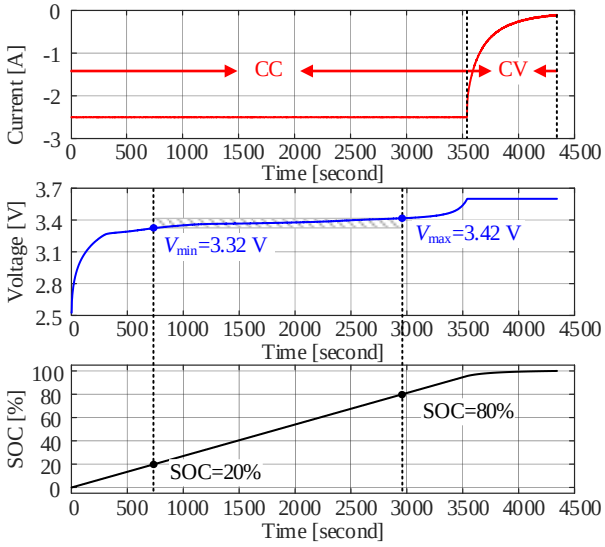
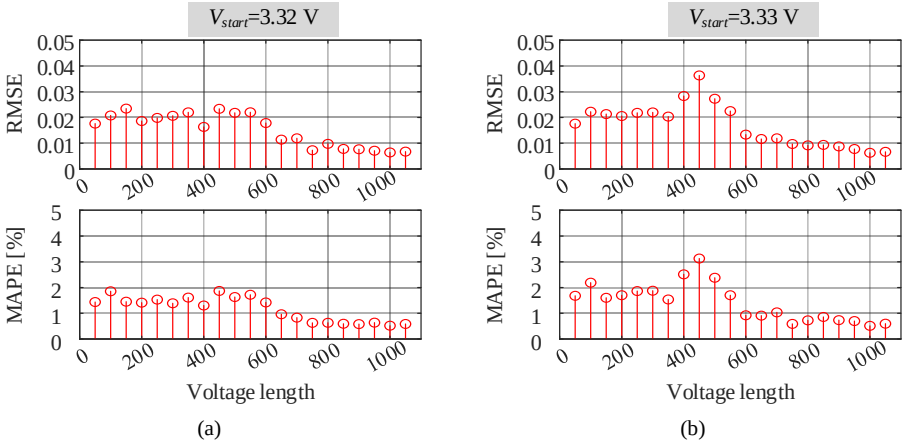


Fig. 5.7. The CC-CV charging curves for a fresh battery. Source: [C5].



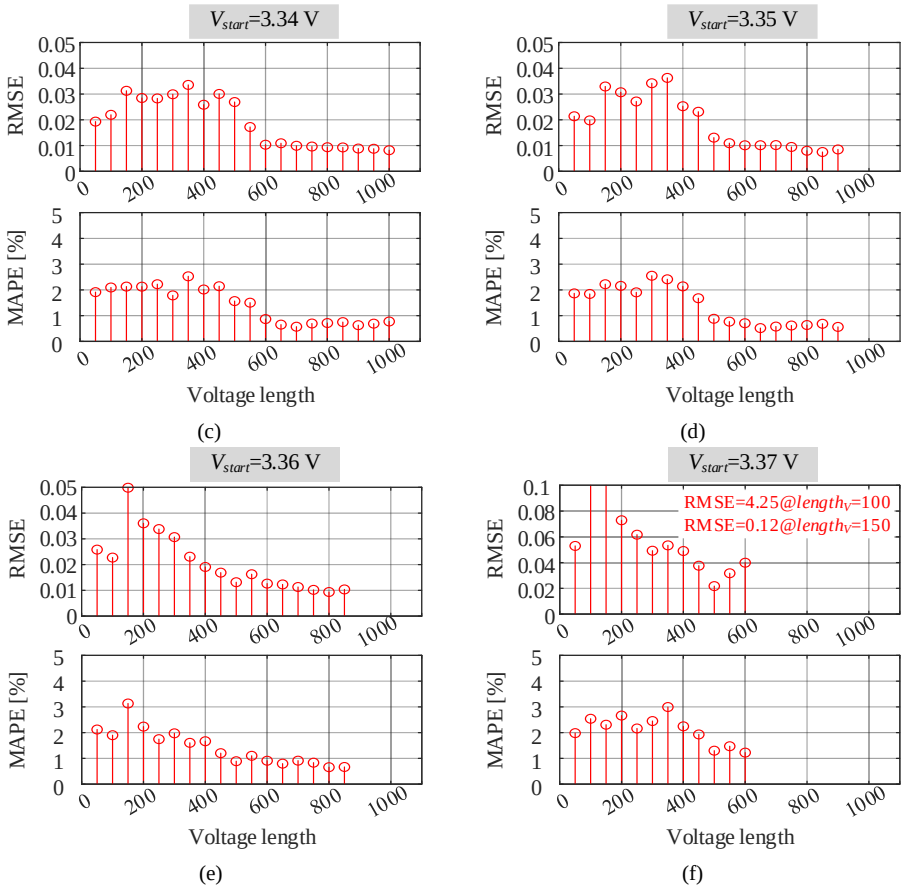


Fig. 5.8. Effect of the hyperparameter $length_V$ on the SOH estimation when V_{start} is set to: (a) 3.32 V, (b) 3.33 V, (c) 3.34 V, (d) 3.35 V, (e) 3.36 V, and (f) 3.37 V. Source: [C5].

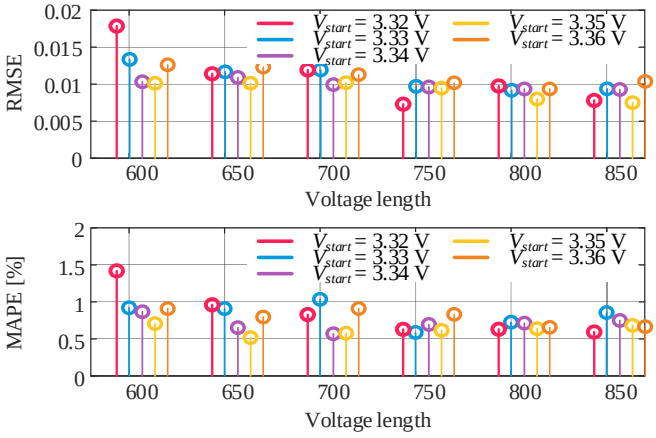


Fig. 5.9. The combination effect of $length_V$ and V_{start} on the SOH estimation. Source: [C5].

Given a fixed $length_V$ of 750, Fig. 5.10 shows the SOH estimation errors when V_{start} ranges from 3.32 to 3.36. V_{start} shows an unobvious impact on the estimation performance as is evident from Fig. 5.10. This is because of the unique and inherent flat voltage characteristic of $LiFePO_4$ batteries, in which a minor variation of V_{start} will not cause too much change in the shape of the voltage curve [C5]. V_{start} was therefore, set to 3.336 V at random. According to the optimization results of hyperparameters $length_V$ and V_{start}, the input voltage is eventually chosen, as depicted in Fig. 5.11.

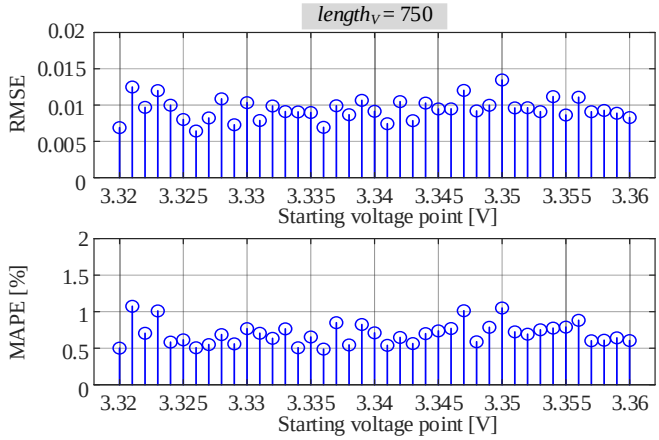


Fig. 5.10. The effect of the hyperparameter V_{start} ($length_V$ is fixed to 750) on the SOH estimation. Source: [C5].

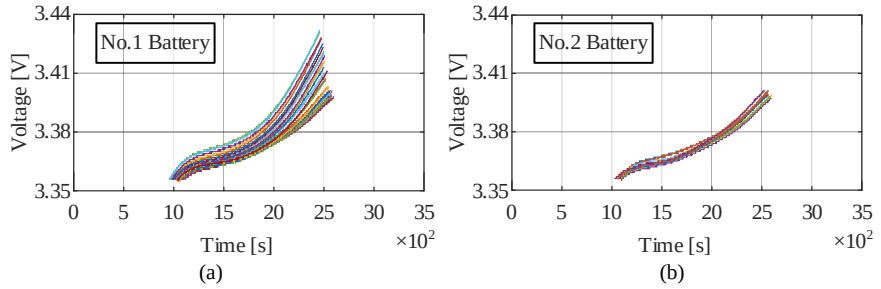


Fig. 5.11. The optimized voltage sequences (the sampled time is 2s): (a) No.1 Battery and (b) No.2 Battery. Source: [C5].

The other two hyperparameters M and B can be chosen by the cut-and-try method. For guaranteeing the global optimization of the hyperparameters, 5-fold cross-validation is applied in ML training. With the variation of M from 1 to 500, the validation results can be seen in Fig. 5.12. It shows that an excessively large M will cause an overfitting problem and reduce the estimation accuracy. That also means setting M in the range of 20 to 50 is reasonable. Additionally, the results obtained from the empirical formula can also verify this conclusion. The empirical formula indicates that M can be determined by the network structure, which is expressed as

$$M = \sqrt{a+b} + c \tag{5.2}$$

where a is the number of input nodes (750 in this study) and b is the number of output nodes (1 in this study), and c can be chosen at random between 1 ~ 13. M is in the range of 28 ~ 40, which is consistent with Fig. 5.12. M is set at 34 in this study.

B was chosen in a similar manner and the validation results are shown in Fig. 5.13. With the increase of B, the estimation error decreases first and stabilizes after B is about 40. Thus, B is adjusted to 40 [C5]. So far, the hyperparameter optimization has been completed, and the results are summarized in TABLE 5.1.

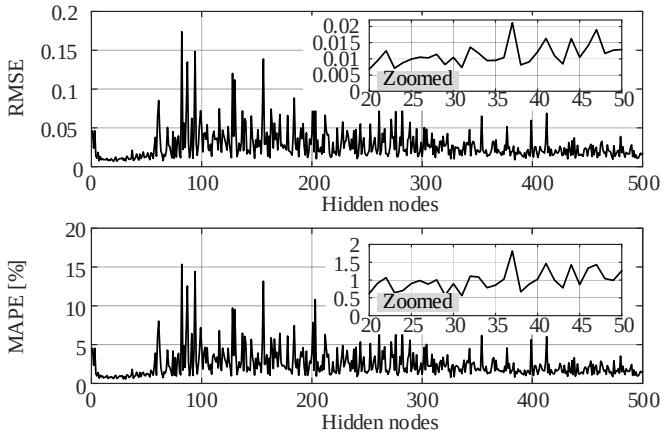


Fig. 5.12. SOH estimation error when using different hyperparameter M. Source: [C5].

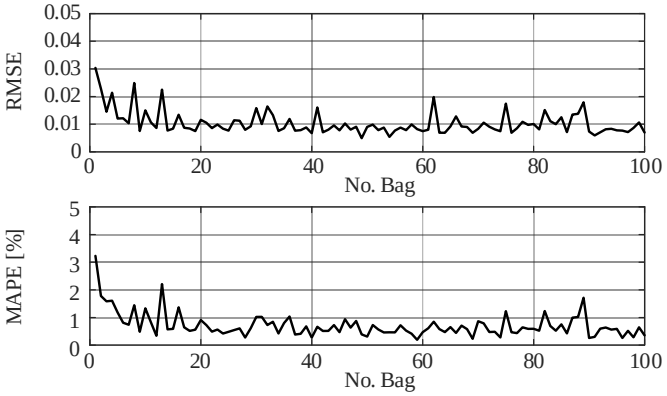


Fig. 5.13. SOH estimation error when using different hyperparameter B. Source: [C5].

TABLE 5.1. Summary of the optimized hyperparameters.

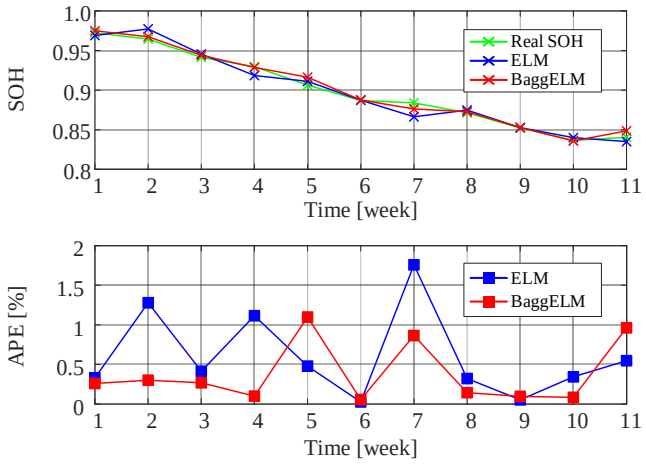
$length_V$	V_{start}	M	B
750	3.336	34	40

5.2.4. VALIDATION RESULTS

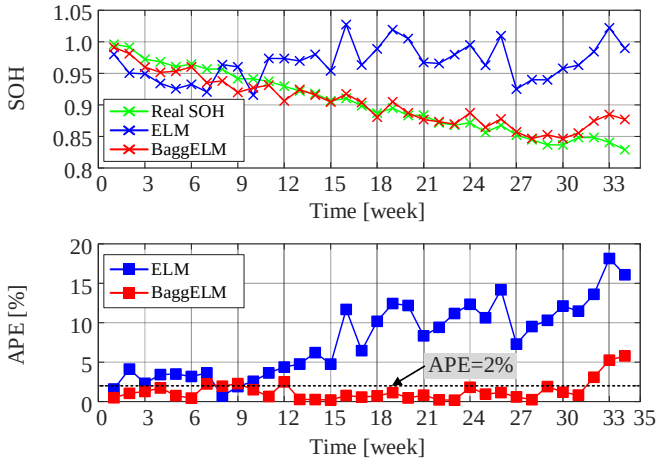
TABLE 5.2 and Fig. 5.14 show the validation and comparison results of both the traditional ELM method and the developed BaggELM method. In the case of self-validation, as observed in Fig. 5.14(a), both methods provide similar good results. BaggELM has a relatively stable estimate, and the maximum APE for ELM and BaggELM are 2% and 1%, respectively. While in the case of mutual validation, as observed in Fig. 5.14(b), BaggELM offers a significant performance advantage. Since the No.1 and No.2 batteries exhibit inconsistency in the degradation characteristics, the ELM method trained on No.2 battery is unable to reveal the SOH of No.1 battery accurately. By contrast, BaggELM provides a good generalization performance and its APE remains under 2% throughout battery lifetime (estimation error only increases towards the final aging stage). A good feature of BaggEML is that, in addition to having an accurate estimation, it is also robust against the inconsistencies between batteries [C5].

TABLE 5.2. Comparison of estimation errors. Source: [C5].

Method	Self-validation		Mutual validation	
	RMSE	**MAPE**	**RMSE**	**MAPE**
ELM	0.0080	0.44%	0.0913	9.00%
BaggELM	0.0048	0.26%	0.0161	0.89%



(a)



(b)

Fig. 5.14. SOH estimation results: (a) self-validation method and (b) mutual validation method. Source: [C5].

5.3. SUMMARY

This chapter studies the artificial neural networks that are capable of recognizing global features from the raw data and are able to cope with multi-dimensional data. To improve the generalization performance as well as accuracy, BaggELM method is proposed on limited data. Moreover, hyperparameters of the proposed method are optimized to guarantee the high accuracy of the estimation. Experimental results on two $LiFePO_4$ batteries with different modes degradation demonstrate the generalization of the proposed method. Furthermore, the BaggELM method has great application potential. Because compared to the estimation results obtained in Chapter 4, the proposed EL method can not only achieve the same accuracy (i.e., APE is lower than 2%) but also does not require manual feature definition and extraction.

CHAPTER 6. CONCLUSIONS AND FUTURE WORK

6.1. CONCLUSIONS

This Ph.D. project aims to identify and optimize robust ML-based SOH estimation algorithms for Li-ion batteries. The methods to improve the performance (i.e., the robustness, accuracy, and data size dependence) of ML-based SOH estimation have been systematically studied. The main findings of the Ph.D. thesis are summarized as follows:

- **Selection of ML-based SOH estimation methods.**

A comprehensive review of the applications of ML algorithms in battery SOH estimation is conducted. From the comparative analysis of estimation performance and training modes, it is founded that SVM and EL are suitable for battery SOH estimation. These methods are effective representative algorithms for the two training modes, of manual feature extraction and automatic feature extraction respectively, especially when the amount of data is limited.

- **A new lifetime model for calendar aging of Li-ion batteries.**

The calendar aging behavior of Li-ion batteries when their capacity fade exceeds 20% is analyzed. Traditionally, the semi-empirical model based on the Arrhenius equation can only study the effect of aging temperature on the capacity fade. Besides, the applicable range of the study is limited to within 20% capacity fade. To mitigate this problem, a nonlinear regression method with two-step fitting is developed. The model incorporates both the influence of storage conditions as well as the piecewise variation in capacity fade (exceeding 20%) [J2].

- **FE represents a robust feature for SOH estimation.**

FE-based SOH feature extracted from a short-term pulse test is proposed. It is proved that the FE-based method shows good performance in parameter selection, dependency on data size, robustness to TCs and noise. In addition, the strategy of dataset selection for entropy-based methods is studied and extracting entropy features from the polarization zone is recommended. Furthermore, multiscale entropy is not suitable when the voltage is short, because the information available in the voltage data will be filtered out in this case.

- **Data noise suppression improves the SOH estimation accuracy.**

An adaptive iterative algorithm and six regression-based smoothers are used before and after entropy-feature extraction, respectively. The proposed method can effectively reduce the noise while maintaining the useful aging information. Results prove adding a simple data noise suppression step improve the SOH estimation accuracy as well as simulation speed.

- **ELM with bagging technology is suitable for SOH estimation when limited training data is available.**

An EL method is proposed, which combines ELM with bagging technology, with the purpose of achieving good generalization performance. The developed BaggELM method is demonstrated to be able to reduce the dependence of ML-based SOH estimation on data size. Moreover, the feature extraction process is automatic, convenient, and efficient for real-world scenarios.

In light of the aforementioned findings, the following conclusions are finally drawn:

- **Battery SOH estimation can become more robust using ML.**

- **Efficient SOH estimation can be obtained even with limited and noisy data by using either fuzzy entropy or ensemble learning.**

6.2. FUTURE WORK

The entire work conducted in this Ph.D. project demonstrates the big potential of ML technologies in battery SOH estimation. As engineering applications develop, more advantages of using ML for SOH estimation can be explored in the future from the following five aspects:

- In view of the unique and complicated aging behavior of Li-ion batteries, ML estimates can be enhanced by incorporating the aging mechanism and modifying the algorithm hyperparameters [J1]. However, the convergence of the modified ML model still needs to be proved theoretically.
- The stability of the offline trained ML model for SOH estimation is still needed to be verified, because the training data, in the absence of big data technologies, are unable to include all the possible working conditions. Developing methods with self-learning and online updating is therefore necessary. In addition, the model-based adaptive filters can update the parameters, thus they can be combined with ML and to realize the real-time SOH estimation.
- Considering the easiness of raw data collection and the increase in computational capabilities of the hardware in practical applications, we can focus on developing

the DL algorithm, which is a very promising method in SOH estimation. On the one hand, it allows high-dimensional input, that is, without any manual intervention. On the other hand, it provides high accuracy and good generalization performance especially when big data is accessible.

- Prior to big data and cloud computing becoming mature, the EL algorithm can be used as an alternative to DL. It is flexible to choose some simple ML methods as base learners; therefore the computation complexity can be reduced. Meanwhile, the algorithm ensemble technology ensures the accuracy and stability of the SOH estimator.
- As a result of high computing resources required, hardware implementations of ML in battery SOH estimation are still in their infancy. From the literature review, most methods estimate SOH at macro-time, rather than in real-time. The advent and development of GPUs have made the inference of ML algorithms approachable and locally available. This provides a great opportunity for the hardware application of ML in battery SOH estimation [J1].

BIBLOGRAPHY

[1] D. Gielen, F. Boshell, D. Saygin, M.D. Bazilian, N. Wagner, and R. Gorini, "The role of renewable energy in the global energy transformation," *Energy Strategy Rev.*, vol. 24, pp.38-50, Apr. 2019.

[2] A. Lüth, J. M. Zepter, P. Crespo del Granado, and R. Egging, "Local electricity market designs for peer-to-peer trading: The role of battery flexibility," *Appl. Energy*, vol. 229, pp. 1233–1243, Nov. 2018.

[3] R. Schmuch, R. Wagner, G. Hörpel, T. Placke, and M. Winter, "Performance and cost of materials for lithium-based rechargeable automotive batteries," *Nat. Energy*, vol. 3, no. 4, pp. 267-268, Apr. 2018.

[4] X. Zeng, M. Li, D. Abd El-Hady, W. Alshitari, A. S. Al-Bogami, J. Lu, and K. Amine, "Commercialization of lithium battery technologies for electric vehicles," *Adv. Energy Mater.*, vol. 9, no. 27, pp. 1900161, June 2019.

[5] G. Zubi, R. Dufo-López, M. Carvalho, and G. Pasaoglu, "The lithium-ion battery: State of the art and future perspectives," *Sust. Energy Rev.*, vol. 89, pp. 292–308, June 2018.

[6] T. Chen, Y. Jin, H. Lv, A. Yang, M. Liu, B. Chen, Y. Xie, and Q. Chen, "Applications of lithium-ion batteries in grid-scale energy storage systems," *Trans. Tianjin Univ.*, vol. 26, no. 3, pp. 208-217, Feb. 2020.

[7] D. I Stroe, M. Swierczynski, A. Stroe, R. Laerke, P.C. Kjaer and R. Teodorescu, "Degradation behavior of lithium-ion batteries based on lifetime models and field measured frequency regulation mission profile," *IEEE Trans. Ind. Appl.*, vol. 52, no. 6, pp. 5009-5018, Nov. 2016.

[8] M. Dubarry, C. Truchot, and B. Y. Liaw, "Cell degradation in commercial $LiFePO_4$ cells with high-power and high-energy designs," *J. Power Sources*, vol. 258, pp. 408-419, July 2014.

[9] M. Broussely, P. Biensan, F. Bonhomme, P. Blanchard, S. Herreyre, K. Nechev, and R. J. Staniewicz, "Main aging mechanisms in Li ion batteries," *J. power sources*, vol. 146, no. 1-2, pp. 90-96, Aug. 2005.

[10] D. Stroe, M. Świerczyński, A. Stan, R. Teodorescu and S. J. Andreasen, "Accelerated lifetime testing methodology for lifetime estimation of lithium-ion batteries used in augmented wind power plants," *IEEE Trans. Ind. Appl.*, vol. 50, no. 6, pp. 4006-4017, Nov.-Dec. 2014.

[11] M. Berecibar, I. Gandiaga, I. Villarreal, N. Omar, J. Van Mierlo, and P. Van den Bossche, "Critical review of state of health estimation methods of Li-ion batteries for real applications," *Renew. Sust. Energ. Rev.*, vol. 56, pp. 572-587, Apr., 2016.

[12] M. H. Lipu, M. A. Hannan, A. Hussain, M. M. Hoque, P. J. Ker, M. H. M. Saad, and, A. Ayob, "A review of state of health and remaining useful life estimation methods for lithium-ion battery in electric vehicles: Challenges and recommendations," *J. Clean. Prod.*, vol. 205, pp.115-133, Dec. 2018.

[13] Y. Li, K. Liu, A. M. Foley, A. Zülke, M. Berecibar, E. Nanini-Maury, J. Van Mierlo, and H. E. Hoster, "Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review,". *Renew. Sust. Energ. Rev.*, vol. 113, pp.109254, Oct. 2019.

[14] H. Tian, P. Qin, K. Li, and, Z. Zhao, "A review of the state of health for lithium-ion batteries: Research status and suggestions," *J. Clean. Prod.*, vol. 261, pp. 120813, July 2020.

[15] W. Waag, S. Käbitz, D. U. Sauer, "Experimental investigation of the lithium-ion battery impedance characteristic at various conditions and aging states and its influence on the application," *Appl. Energy*, vol. 1, no. 102, pp. 885-97, Feb. 2013.

[16] K. S. Ng, C. S. Moo, Y. P. Chen, and Y. C. Hsieh, "Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries," *Appl. Energy*, vol. 86, pp. 1506-1511, Sep. 2009.

[17] D. I. Stroe, M. Swierczynski, A. I. Stroe, S. K. Kaer, and R. Teodorescu, "Lithium-ion battery power degradation modelling by electrochemical impedance spectroscopy," *IET Renew. Power Gener.*, vol. 11, no. 9, pp. 1136–1141, 2017.

[18] M. Dubarry and B. Y. Liaw, "Identify capacity fading mechanism in a commercial $LiFePO_4$ cell," *J. Power Sources*, vol. 194, no. 1, pp. 541– 549, 2009.

[19] K. Honkura, K. Takahashi, and T. Horiba, "Capacity-fading prediction of lithium-ion batteries based on discharge curves analysis," *J. Power Sources*, vol. 196, no. 23, pp. 10141–10147, 2011.

[20] D. I. Stroe and E. Schaltz, "Lithium-ion battery state-of-health estimation using the incremental capacity analysis technique," *IEEE Trans. Ind. Appl.*, vol. 56, no. 1, pp. 678-685, Jan.-Feb. 2020.

[21] J. Wei, G. Dong, and Z. Chen, "Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5634-5643, Dec. 2017.

[22] C. Zou, C. Manzie, D. Neši´c, and A. G Kallapur, "Multi-time-scale observer design for state-of-charge and state-of-health of a lithium-ion battery," *J. Power Sources*, vol. 335, pp. 121–130, Dec. 2016.

[23] Y. Hua, A. Cordoba-Arenas, N. Warner, and G. Rizzoni, "A multi time-scale state-of-charge and state-of-health estimation framework using nonlinear predictive filter for lithium-ion battery pack with passive balance control," *J. Power Sources*, vol. 280, pp. 293–312, Apr. 2015.

[24] Y. Li, M. Abdel-Monem, R. Gopalakrishnan, M. Berecibar, E. Nanini-Maury, N. Omar, P. van den Bossche, and, J. Van Mierlo, "A quick on-line state of health estimation method for Li-ion battery with incremental capacity curves processed by Gaussian filter," *J. Power Sources*, vol. 373, pp. 40–53, Jan. 2018.

[25] J. Meng, L. Cai, G. Luo, D. I. Stroe, and R. Teodorescu, "Lithium-ion battery state of health estimation with short-term current pulse test and support vector machine," *Microelectron. Reliab.*, vol. 88, pp. 1216-1220, Sep. 2018.

[26] J. Wu, Y. Wang, X. Zhang, and Z. Chen, "A novel state of health estimation method of Li-ion battery using group method of data handling," *J. Power Sources*, vol. 327, pp. 457–464, 2016.

[27] D. Yang, Y. Wang, R. Pan, R. Chen, and Z. Chen, "A neural network based state-of-health estimation of lithium-ion battery in electric vehicles," *Energy Procedia*, vol. 105, pp. 2059–2064, May, 2017.

[28] J. Wu, C. Zhang, and Z. Chen, "An online method for lithium-ion battery remaining useful life estimation using importance sampling and neural networks," *Appl. Energy*, vol. 173, pp. 134–140, July 2016.

[29] R. Lajara, J. J. Perez Solano, and J. Pelegri Sebastia, "Predicting the batteries state of health in wireless sensor networks applications," *IEEE Trans. Ind. Electron.*, vol. 65, no. 11, pp. 8936–8945, Nov. 2018.

[30] H. Li, D. Pan, and C.L.P. Chen, "Intelligent prognostics for battery health monitoring using the mean entropy and relevance vector machine," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 44, no. 7, pp. 851-862, Feb. 2014.

[31] H. Chaoui and C. C. Ibe-Ekeocha, "State of charge and state of health estimation for lithium batteries using recurrent neural networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 8773–8783, June 2017.

[32] G. W. You, S. Park, and D. Oh, "Diagnosis of electric vehicle batteries using recurrent neural networks," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 4885–4893, Feb. 2017.

[33] H. Lin, T. Liang, and S. Chen, "Estimation of battery state of health using probabilistic neural network," *IEEE Trans. Ind. Inform.*, Vol. 9, no. 2, pp. 679-685, May 2013.

[34] S. Shen, M. Sadoughi, M. Li, Z. Wang, and C. Hu, "Deep convolutional neural networks with ensemble learning and transfer learning for capacity estimation of lithium-ion batteries," *Appl. Energy*, vol. 260, pp. 114296, Feb. 2020.

[35] Y. Li, C. Zou, M. Berecibar, E. Nanini-Maury, J.C.W. Chan, P. van den Bossche, J. Van Mierlo, and N. Omar, "Random forest regression for online capacity estimation of lithium-ion batteries," *Appl. Energy*, vol. 232, pp. 197–210, Dec. 2018.

[36] D. Yang, X. Zhang, R. Pan, Y. Wang, and Z. Chen, "A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve," *J. Power Sources*, vol. 384, pp. 387-395, Apr. 2018.

[37] X. Hu, J. Jiang, D. Cao, and B. Egardt, "Battery health prognosis for electric vehicles using sample entropy and sparse Bayesian predictive modeling," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2645–2656, Apr. 2016.

[38] E. Sarasketa-Zabala, F. Aguesse, I. Villarreal, L. M. Rodriguez-Martinez, C. M. López, and P. Kubiak, "Understanding lithium inventory loss and sudden performance fade in cylindrical cells during cycling with deep-discharge steps," *J. Phys. Chem. C.* vol. 119, no. 2, pp. 896–906, Jan. 2015.

[39] X. Han,; L. Lu,; Y. Zheng,; X. Feng,; Z.; Li, J. Li, and M. Ouyang, "A review on the key issues of the lithium ion battery degradation among the whole life cycle," *ETransportation*, vol. 1, pp. 100005, Aug. 2019.

[40] D. I. Stroe, "Lifetime models for Lithium-ion batteries used in virtual power plant applications," Ph.D. dissertation, Dept. Energy Technol., Aalborg University, Aalborg, Denmark, 2014.

[41] H. Bindner, T. Cronin, P. Lundsager, J. F. Manwell, U. Abdulwahid, and I. Baring-Gould, "Lifetime modelling of lead acid batteries," *Risoe National Lab., Wind Energy Dept., Roskilde, Denmark*, 2005.

[42] T. R. Ashwin, A. Barai, K. Uddin, L. Somerville, A. McGordon, and J. Marco, "Prediction of battery storage ageing and solid electrolyte interphase property estimation using an electrochemical model," *J. Power Sources*, vol. 385, pp. 141–147, May 2018.

[43] M. Dubarry, N. Qin, and P. Brooker, "Calendar aging of commercial Li-ion cells of different chemistries–A review," *Curr. Opin. Electrochem.*, vol. 9, pp. 106–113, June 2018.

[44] E. Redondo-Iglesias, P. Venet, and S. Pelissier, "Modelling lithium-ion battery ageing in electric vehicle applications—calendar and cycling ageing combination effects," *Batteries*, vol. 6, no. 1, pp. 14, Mar. 2020.

[45] J. Schmalstieg, S. Käbitz, M. Ecker, and D. U. Sauer, "A holistic aging model for Li(NiMnCo)O2 based 18650 lithium-ion batteries," *J. Power Sources*, vol. 257, pp. 325–334, July 2014.

[46] E. Redondo-Iglesias, P. Venet, and S. Pelissier, "Eyring acceleration model for predicting calendar ageing of lithium-ion batteries," *J. Energy Storage*, vol. 13, pp. 176–183, Oct. 2017.

[47] S. L. Hahn, M. Storch, R. Swaminathan, B. Obry, J. Bandlow, and K. P Birke, "Quantitative validation of calendar aging models for lithium-ion batteries," *J. Power Sources*, vol. 400, pp. 402–414, Oct. 2018.

[48] K. Liu, Y. Li, X. Hu, M Lucu, and W. D. Widanage, "Gaussian process regression with automatic relevance determination kernel for calendar aging prediction of lithium-ion batteries," *IEEE Trans. Ind. Inform.*, vol. 16, no. 6, pp. 3767–3777, June 2020.

[49] S. Käbitz, J. B. Gerschler, M. Ecker, Y. Yurdagel, B. Emmermacher, D. André, T. Mitsch, and D. U. Sauer, "Cycle and calendar life study of a graphite| LiNi1/3Mn1/3Co1/3O2 Li-ion high energy system. Part A: Full cell characterization," *J. Power Sources*, vol. 239, pp. 572–583, Oct. 2013.

[50] M. Schimpe, M.E. von Kuepach, M. Naumann, H.C. Hesse, K. Smith, A. Jossen, "Comprehensive modeling of temperature-dependent degradation mechanisms in Lithium iron phosphate batteries," *J. Electrochem. Soc.*, vol. 165, no. 2, pp. A181, Jan. 2018.

[51] X. Li, Z. Wang, L. Zhang, C. Zou, and D. D. Dorrell, "State-of-health estimation for Li-ion batteries by combing the incremental capacity analysis method with grey relational analysis," *J. Power Sources*, vol. 410-411, pp.106-114, Jan. 2019.

[52] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proc. Natl. Acad. Sci. USA*, vol. 88, no. 6, pp. 2297–2301, Mar. 1991.

[53] J. S. Richman and J. R. Moorman, "Physiological time series analysis using approximate entropy and sample entropy," Amer. *J. Physiol.*, vol. 278, no. 6, pp. H2039–H2049, June 2000.

[54] W. Chen, Z. Wang, H. Xie and W. Yu, "Characterization of surface EMG signal based on fuzzy entropy," *IEEE Trans. Neural Syst. Rehabil. Eng*. vol. 15, no. 2, pp. 266-272, June 2007.

[55] M. Costa, A.L. Goldberger, and C.K. Peng, "Multiscale entropy analysis of complex physiologic time series," *Phys. Rev. Lett.*, vol. 89, no. 6, pp. 068102, July 2002.

[56] M. Costa, A.L. Goldberger, and C.K. Peng, "Multiscale entropy analysis of biological signals," *Phys. Rev. E*., vol. 71, no. 2, pp. 021906, Feb. 2005.

[57] J. M. Yentes, N. Hunt, K. K. Schmid, J. P. Kaipust, D. McGrath, and N. Stergiou, "The appropriate use of approximate entropy and sample entropy with short data sets," *Ann. Biomed. Eng.*, vol. 41, no. 2, pp. 349–365, Feb. 2013.

[58] D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman, "Sample entropy analysis of neonatal heart rate variability," *Am. J. Physiol. Regul. Integr. Comp. Physiol.* Vol. 283, no. 3, pp. R789–R797, Sep. 2002.

[59] J. Jiang, Q. Liu. C. Zhang, and W. Zhang, "Evaluation of acceptable charging current of power Li-ion batteries based on polarization characteristics," *IEEE Trans. Ind. Electron.*, vol. 61, no. 12, pp. 6844–6851, Apr. 2014.

[60] C. Yu, Y. Li, and M. Zhang, "Comparative study on three new hybrid models using Elman Neural Network and Empirical Mode Decomposition based technologies improved by Singular Spectrum Analysis for hour-ahead wind speed forecasting," *Energy Convers. Manage.* vol. 147, pp. 75-85, Sep. 2017.

[61] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C Yen, C. C Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *P. Roy. Soc. A-Math. Phy.*, vol. 454, no. 1971, pp. 903-995, Mar. 1998.

[62] R. R. Richardson, M. A. Osborne, and D. A. Howey, "Gaussian process regression for forecasting battery state of health," *J. Power Sources*, vol. 357, pp. 209-219, July 2017.

[63] Y. Li, C. Zou, M. Berecibar, E. Nanini-Maury, J. C. W. Chan, P. van den Bossche, J. Van Mierlo, and N. Omar, "Random forest regression for online capacity estimation of lithium-ion batteries," *Appl. energy*, vol. 232, pp.197-210, Dec. 2018.

[64] X Li, C Yuan, Z Wang, "Multi-time-scale framework for prognostic health condition of lithium battery using modified Gaussian process regression and nonlinear regression," *J. Power Sources*, vol. 467, pp. 228358, Aug. 2020.

[65] A. Géron, "Hands-on machine learning with scikit-learn and tensorflow," *O'Reilly Media*, 2017.

[66] J. Yu, "State of health prediction of lithium-ion batteries: Multiscale logic regression and Gaussian process regression ensemble," *Reliab. Eng. Syst. Saf.*, vol. 174, pp. 82-95, June 2018.

[67] Y. Cheng, D. Song, Z. Wang, C. Lu, and N. Zerhouni, "An ensemble prognostic method for lithium-ion battery capacity estimation based on time-varying weight allocation," *Appl. Energy*, vol. 266, pp.114817, May 2020.

[68] D. Liu, W. Xie, H. Liao, and Y. Peng, "An integrated probabilistic approach to lithium-ion battery remaining useful life estimation," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 3, pp. 660–670, Mar. 2015.

[69] B. Gou, Y. Xu and X. Feng, "An ensemble learning-based data-driven method for online state-of-health estimation of lithium-ion batteries," *IEEE Trans. Transp.*, vol. 7, no. 2, pp. 422-436, June 2021.

[70] J. Meng, L. Cai, D.I. Stroe, J. Ma, G. Luo, and R. Teodorescu, "An optimized ensemble learning framework for lithium-ion battery state of health estimation in energy storage system," *Energy*, vol. 206, pp. 118140, Sep. 2020.

[71] .S. Khaleghi, Y. Firouz, M. Berecibar, J. Van Mierlo, and P. Van Den Bossche, "Ensemble gradient boosted tree for SoH estimation based on diagnostic features," *Energies*, vol. 13, no. 5, pp. 1262, Mar. 2020.

[72] A. Guarino, W. Zamboni and E. Monmasson, "A comparison of ensemble machine learning techniques for the estimate of residual capacity of li-ion batteries," in *IEEE 29th International Symposium on Industrial Electronics (ISIE)*, Delft, Netherlands, pp. 1307-1312, 2020.

[73] W. Liu and Y. Xu, "A data-driven method for online health estimation of Li-ion batteries with a novel energy-based health indicator," *IEEE Trans. Energy Convers.*, vol. 35, no. 3, pp. 1715-1718, Sep. 2020.

[74] D. Liu, W. Xie, H. Liao and Y. Peng, "An integrated probabilistic approach to lithium-ion battery remaining useful life estimation," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 3, pp. 660-670, Mar. 2015.

[75] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489-501, Dec. 2006.

[76] Y. H. Pao, G. H. Park, and D. J. Sobajic, "Learning and generalization characteristics of the random vector Functional-link net," *Neurocomputing*, vol. 6, no. 2 pp. 163-180, Apr. 1994.

APPENDIX A. PARAMETERS OF THE TESTED BATTERIES

TABLE A.1. The datasheet of the tested $LiFePO_4$ battery.

Item	Value
Chemistry	$LiFePO_4$/C
Type	cylindrical
Dimensions	Ø 26×65 mm
Weight	76 g
Nominal capacity	2.5 Ah
Nominal voltage	3.3 V
Maximum voltage	3.6 V
Cut-off voltage	2.0 V
Maximum continuous charge current	10 A
Maximum continuous discharge current	50 A
Operating temperature	−30 °C to 55 °C
Storage temperature	−40 °C to 60 °C

APPENDIX B. EXPERIMENTAL SETUP

The Maccor battery test station, as shown in Fig. B.1 is used to perform cyclic aging test with the frequency regulation profile. The Memmert UNP 500 oven, as shown in Fig. B.2 is used to perform calendar aging test with five test conditions (TCs) consisting of SOC (i.e., 10%, 50% and 90%) and temperature (i.e., 40 °C, 47.5 °C and 55 °C). The FuelCon battery test station, as shown in Fig. B.1 is used to perform the reference measurements during for both cyclic aging and calendar aging analysis.

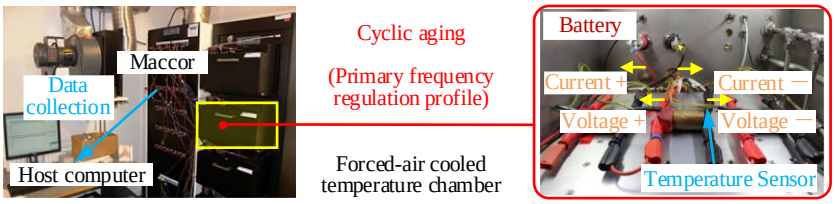


Fig. B.1. Experimental setup for cyclic aging test.

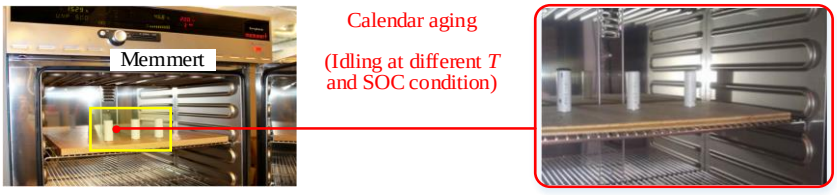


Fig. B.2. Experimental setup for calendar aging test.

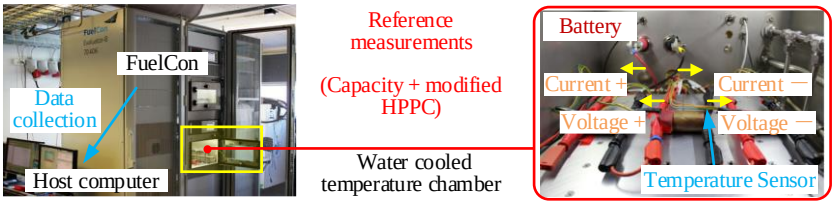


Fig. B.3. Experimental setup for reference measurement during cyclic aging and calendar aging.

APPENDIX C. ACCELERATE AGING TESTS

C.1. CYCLIC AGING TEST

The whole cyclic aging test profile consists an aging test and the reference measurements, the latter of which includes periodic capacity test and hybrid pulse power characterization (HPPC) test, as shown in Fig. C.1. The climatic chamber is set to 25 °C for all the processes comprised in cyclic aging. Batteries are aged with a one-week mission profile from the energy storage system providing primary frequency regulation to the grid. During this period, the battery SOC varied from 10% and 90%, as shown in Fig. C.2. In order to describe this mission profile in more detail, the SOC variation, current, and voltage response in the zoomed part are also provided.

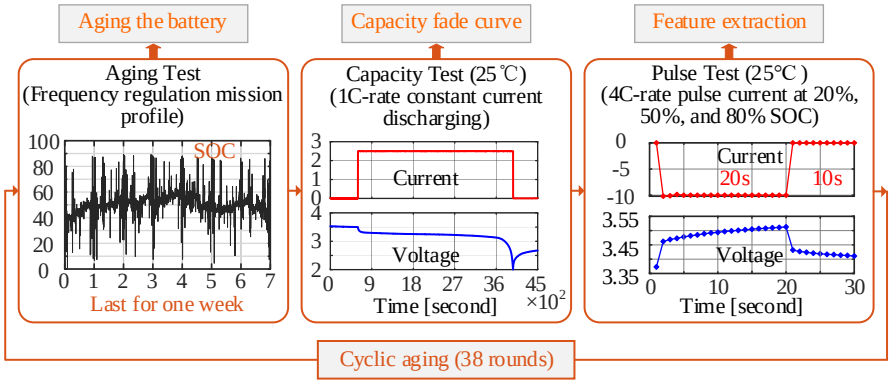


Fig. C.1. Flowchart of the cyclic aging and reference tests procedure. Source: [40, J3].

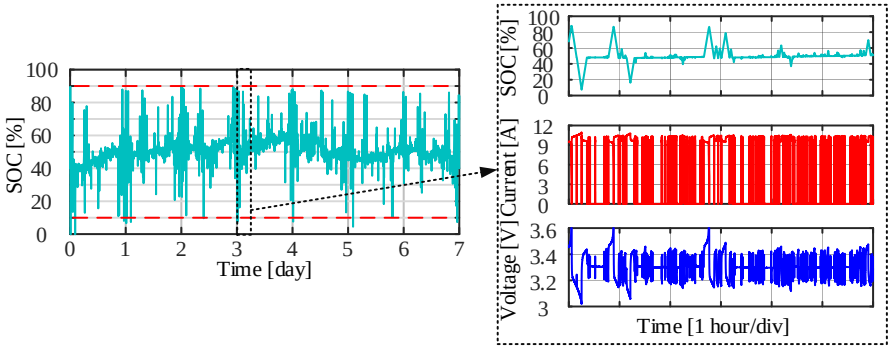


Fig. C.2. SOC variation of the tested batteries when performing the aging mission profile, and the current and voltage responses in the zoomed part. Source: [40, J3].

C.2. CALENDAR AGING TEST

Like the cyclic aging test, the calendar test profile also consists an aging test and the reference measurements, as shown in Fig. C.3. The temperature is adjusted to different values when the aging tests are performed. And the temperature of the water-cooled chamber is controlled at 25 °C during the reference measurements. Considering the storage time, SOC, and temperature as stress factors, the test matrix is designed on $LiFePO_4$ batteries, as shown in Fig. C.4. Fifteen batteries are aged at five TCs consisting of SOC (i.e., 10%, 50% and 90%) and temperature (i.e., 40 °C, 47.5 °C and 55 °C). The aging tests are stopped after every month for quantify the performance parameters of the batteries. The entire tests last for 43 months for Case 1, Case 2, and Case 3, and 27 months for Case 4 and Case 5.

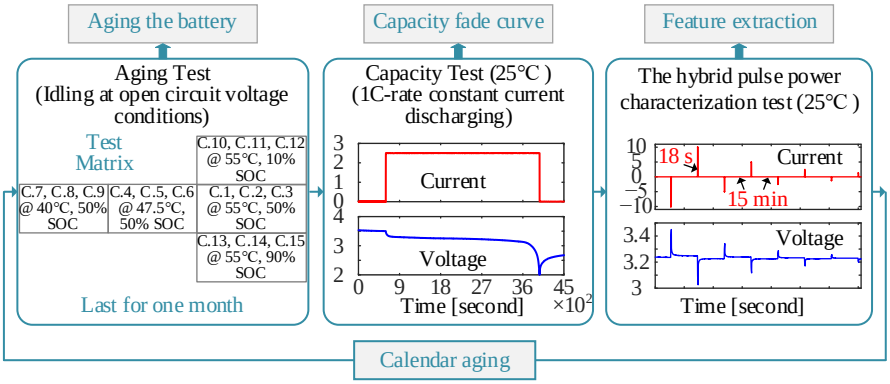


Fig. C.3. Flowchart of the calendar aging and reference tests procedure. Source: [40, J2].

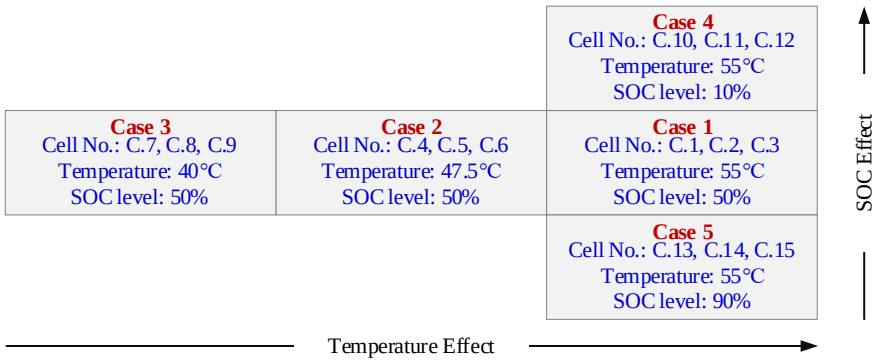


Fig. C.4. The test matrix of calendar aging (possible interaction between stress factors is not considered). Source: [40, J2].

C.3. REFERENCE MEASUREMENTS

1) Capacity test: The capacity test is carried out under different C-rate. Batteries are first charged with a 1C-rate constant current (CC) until the voltage reaches 3.6 V. Then the voltage is kept at 3.6 V until the current equals 0.1 A where batteries are considered fully charged (CC-CV charging). After 15 minutes of relaxation for achieving electrochemical stability, the current battery capacity is measured following a 1C-rate constant current discharging procedure (CC discharging). After relaxing the battery for one hour, the capacity test is repeated with a 4C-rate current. During both charging and discharging, the battery data is sampled with one second.

2) HPPC test: A modified version of HPPC test is designed, the details are presented in Fig. C.5 and Fig. C.6. Starting from the fully discharged state, the battery is charged with a constant current of 1C-rate until it reached to 20%, 50%, and 80% SOC, respectively. Each time the specified SOC level is reached, the pulse test profile composed with a consecutive current pulse of four C-rates (i.e., 4C-rate, 2C-rate, 1C-rate, and 0.5C-rate) is performed for feature extraction. Each pulse lasts for 20s for both charging and discharging conditions, with 15 minutes of relaxation time in between.

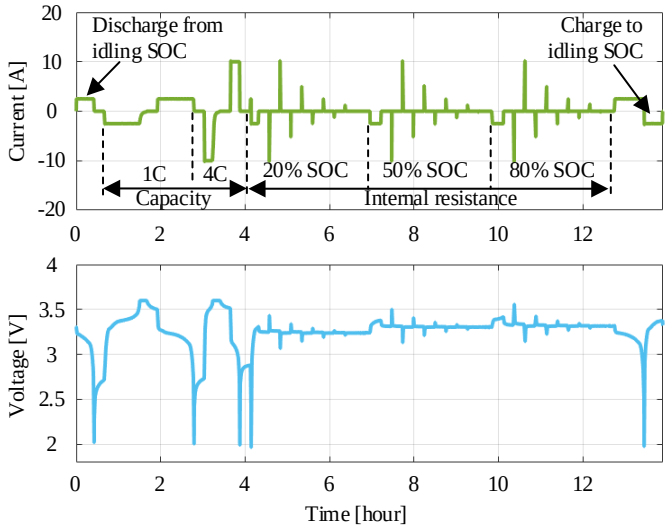


Fig. C.5. Current (a) and voltage response (b) during the reference measurements. Source: [40, J2].

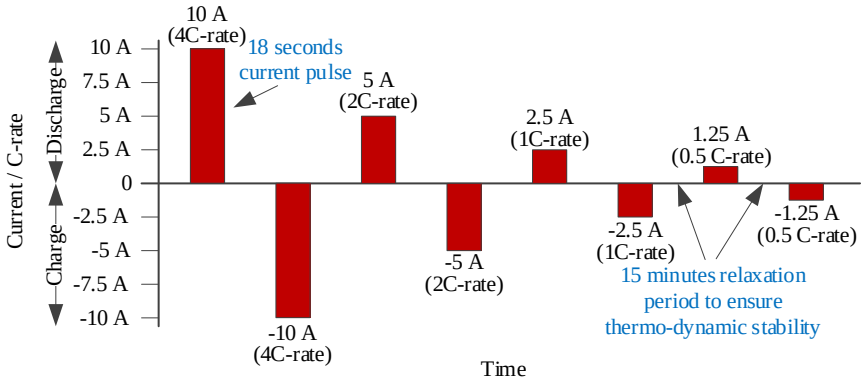


Fig. C.6. The modified HPPC test profile used to extract SOH features of the $LiFePO_4$ battery cells at 20%, 50% and 80% SOC. Source: [40, J2].

APPENDIX D. SELECTED MACHINE LEARNING ALGORITHMS

D.1. NONLINEAR REGRESSION

Algorithm 1: Nonlinear regression

Require: y_i, $\bar{y}_i$, and $\hat{y}_i$: the real, average, and the estimated capacity fade values, respectively

Require: w_j : the estimated value of *j*th weight

Require: α: The learning rate

input: A training data set of N battery capacity fade data pairs $D_N = \{(\mathbf{x}_i, y_i), i = 1, 2, \ldots, N\}$, weight w, and learning rate α

initialization: initialize w_j randomly

for $i = 1, 2, \ldots, N$ **do**

Calculate the sum of squared errors between the model and the output

► $E_w = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$

for $j = 0, 1, \ldots, d$ **do**

Calculate the gradient of the weight ► $\frac{\partial E_w}{\partial w_j} = 0$

Update the parameters iteratively by the gradient descent method to minimize E_w until it converges to the desired local minimum

► $w_j = w_j - \alpha \frac{\partial E_w}{\partial w_j}$

end for

end for

Output: the regression function $f(t, w)$

Fig. D.1. The nonlinear regression algorithm. Source: [J2].

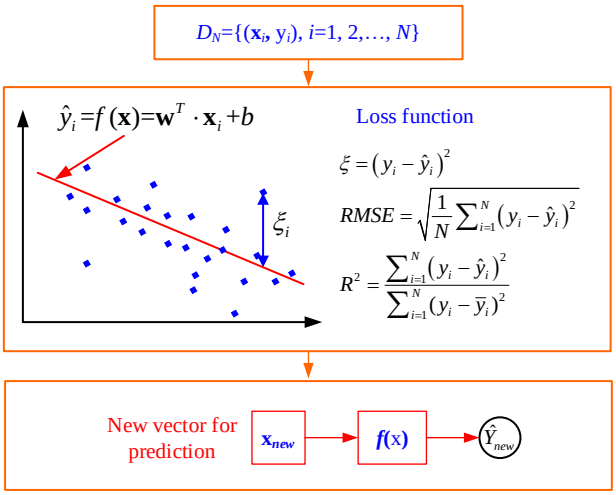


Fig. D.2. The illustration of nonlinear regression. Source: [J1].

D.2. SUPPORT VECTOR MACHINE

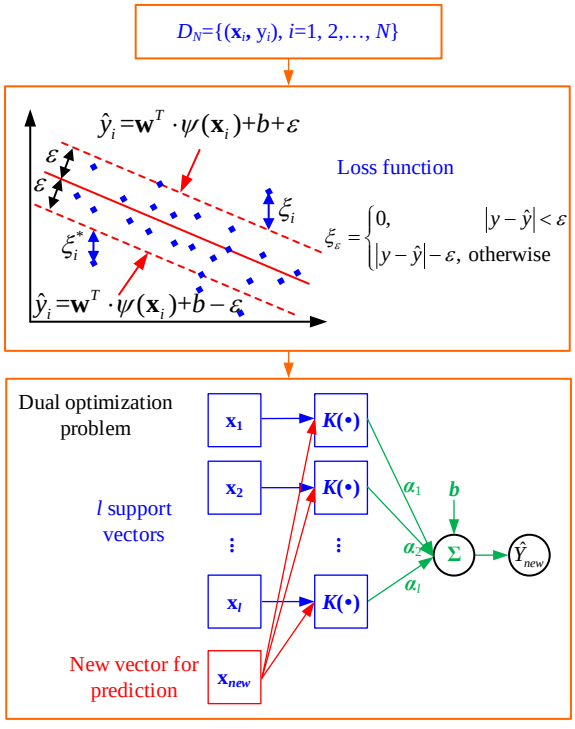


Fig. D.3. The illustration of support vector machine. Source: [J1].

Algorithm 2: Support vector machine for regression

Require: $\mathbf{x}_i$: the feature vector

Require: y_i and $\hat{y}_i$: the real and the estimated SOH value, respectively

Require: N: the sample number

Require: $\mathbf{w}$ and b: the coefficients matrix that needs to be optimized

Require: $\psi(\cdot)$: the target mapping

Require: d and $\tilde{d}$: the dimension of original vector space and the new feature space, respectively

Require: ξ_i and ξ_i^*: slack variables

Require: C: a positive constant regulating the penalty

Require: α_i, α_i, β_i, β_i^* : Lagrange multipliers

Require: $K(\mathbf{x}_i, \mathbf{x})$: the kernel function

input: A training data set of N battery SOH data pairs $D_N = \{(\mathbf{x}_i, y_i), i = 1, 2, \ldots, N\}$, training inputs $\mathbf{x}$

initialization: initialize $\mathbf{w}$ randomly

Calculate the ε-insensitive loss function

► $\xi_\varepsilon\left(\hat{y}_i, y_i\right) = \begin{cases} 0, & \left|y_i - \hat{y}_i\right| < \varepsilon \\ \left|y_i - \hat{y}_i\right| - \varepsilon, & \text{otherwise} \end{cases}, \ \forall i \in \{1, 2, \ldots, N\}$

$$\min_{\substack{\mathbf{w}\in R^d \\ \xi_i,\, \xi_i^* \in R^N}} \ \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_i^N\left(\xi_i + \xi_i^*\right)$$

Create the primal SVM optimization ► $s.t. \begin{cases} y_i - \mathbf{w}^T \cdot \psi(\mathbf{x}_i) - b \le \varepsilon + \xi_i \\ \mathbf{w}^T \cdot \psi(\mathbf{x}_i) + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases}$

Create the min-max problem by introducing Lagrange multipliers α_i, α_i, β_i, β_i^*

Create the max-min problem by satisfying the Karushe-Kuhne-Tucker (KKT) conditions, i.e., $\nabla_{\mathbf{w}} L = 0$, $\nabla_b L = 0$, $\nabla_{\xi_i} L = 0$, $\nabla_{\xi_i^*} L = 0$

► $\mathbf{w} = \sum_{i=1}^n\left(\alpha_i - \alpha_i^*\right)\psi(\mathbf{x}_i)$, $\sum_{i=1}^n\left(\alpha_i - \alpha_i^*\right) = 0$, $\begin{cases} C = \alpha_i + \xi_i \\ C = \alpha_i^* + \xi_i^* \end{cases}$

Compute $\mathbf{w}$ and b from α ► $b = y_i - \sum_{i=1}^N\left(\alpha_i - \alpha_i^*\right)\psi(\mathbf{x}_i)^T \cdot \psi(\mathbf{x}_i)$

for examples i where $0 < \alpha_i, \alpha_i^* < C$

Output: the regression function $f(\mathbf{x}) = \mathbf{w}^T \cdot \psi(\mathbf{x}) + b = \sum_{i=1}^N (\alpha_i^* - \alpha_i)^T \cdot K(\mathbf{x}_i, \mathbf{x}) + b$

Fig. D.4. The support vector machine algorithm. Source: [J1].

D.3. GAUSSIAN PROGRESS REGRESSION

Algorithm 3: Gaussian progress regression

Require: **x**: the input matrix
Require: $m(\mathbf{x})$ and $k(\mathbf{x}, \mathbf{x}')$: the mean and covariance functions, respectively
Require: $\boldsymbol{y}$ and $\boldsymbol{y}^*$: the real and the estimated SOH value, respectively
Require: *Gaussian process* $f(\mathbf{x}) \sim N(m(x), k(x, x'))$

input: A training data set of N battery SOH data pairs $D_N = \{(\mathbf{x}_i, y_i), i = 1, 2, \ldots, N\}$, training inputs $\mathbf{x}$, $m(\mathbf{x})$, $k(\mathbf{x}, \mathbf{x}')$
Calculate the prior distribution ► $\mathbf{y} \sim N(0, K_f(x, x') + \sigma^2 \mathbf{I})$

Calculate the hyper parameters σ and f by the maximum likelihood method
Obtain the joint prior distribution of $\boldsymbol{y}$ and predicted value $\boldsymbol{y}^*$

► $$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}^* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K_f(x, x) + \sigma^2 \mathbf{I} & K_f(x, x^*) \\ K_f(x, x^*)^T & K_f(x^*, x^*) \end{bmatrix}\right)$$

Output: the posterior distribution for the a given input

► $p(\mathbf{y}^* \mid \mathbf{x}, \mathbf{y}, \mathbf{x}^*) = N\left(\bar{\mathbf{y}}^*, \mathrm{cov}(\mathbf{y}^*)\right)$

Fig. D.5. The Gaussian progress regression algorithm. Source: [C4].

D.4. EXTREME LEARNING MACHINE

Algorithm 4: Extreme learning machine

Require: **X**: the input matrix
Require: **Y**: the output vector given input matrix **X**
Require: **W**: the input weight matrix
Require: $\mathbf{W_0}$: the bias of hidden nodes
Require: **β**: the output weight vector
Require: g(·):the activation function
input: N battery capacity fade data pairs $D_N = \{(t_i, y_i), i = 1, 2, \ldots, N\}$, weight w, and learning rate α
initialization: initialize **W** randomly

Calculate the hidden layer output matrix **H** ► $\mathbf{H} = g\left(\mathbf{W} \cdot \mathbf{X}^{\mathbf{T}} + \mathbf{b}\right)$

Calculate the Moore-Penrose inverse of **H** ► $\mathbf{H}^+ = \left(\mathbf{H}^T \mathbf{H}\right)^{-1} \mathbf{H}^T$

Output: the output weight vector ► $\boldsymbol{\beta} = \mathbf{H}^+ \cdot \mathbf{Y}$

Fig. D.6. The extreme learning machine algorithm. Source: [C5].

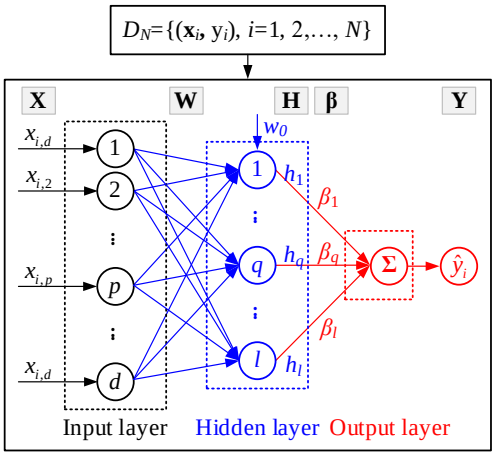


Fig. D.7. The structure of ELM. Source: [C5].

APPENDIX E. ENTROPY ALGORITHMS

E.1. APPROXIMATE ENTROPY ALGORITHM

Algorithm 5: Approximate entropy

Require: N: the total number of data
Require: r: the tolerance for accepting matrices
Require: m: the dimension of vectors
input: a given series $\{v(1), v(2), \ldots, v(N)\}$
initialization: initialize r as a positive real number, m as a positive integer
Loop:
for $i, j = 1, 2, \ldots, N\text{-}m+1$ **do**
Generate vectors $V_m(i)$ and $V_m(j)$ ► $V_m(i) = [v(i), v(i+1), \ldots, v(i+m-1)]$, $V_m(j) = [v(j), v(j+1), \ldots, v(j+m-1)]$
end for
for $i, j = 1, 2, \ldots, N\text{-}m+1$, $k = 0, 1, \ldots, m\text{-}1$ **do**
Calculate the absolute distance between vector $V_m(i)$ and $V_m(j)$
► $d_{ij}^m = \max\{|v(i+k) - v(j+k)|\}$
Calculate the similarity degree D_{ij}^m
► $D_{ij}^m(r) = \begin{cases} 1, & d_{ij} \le r \\ 0, & d_{ij} > r \end{cases}$
Calculate W^m ► the sum of D_{ij}^m
end for
for $i = 1, 2, \ldots, N\text{-}m+1$,
Calculated the conditional probability $C_i^m(r)$ ► $C_i^m(r) = \frac{1}{N-m+1} W^m(i)$
Calculate the probability of matching points
► $\Phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \log C_i^m(r)$
end for
***m* = *m* + 1 and repeat loop only once**
Output: the approximate entropy (AE) of the input series
► $\mathrm{AE}(m,r,N) = \lim_{N\to\infty} \left[\Phi^m(r) - \Phi^{m+1}(r)\right]$
Because the length of the data is always limited, the approximate entropy can be estimated by the statistic ► $\mathrm{AE}(m,r,N) = \Phi^m(r) - \Phi^{m+1}(r)$

Fig. E.1. The approximate entropy algorithm. Source: [J3].

E.2. SAMPLE ENTROPY ALGORITHM

Algorithm 6: Sample entropy

Require: N: the total number of data
Require: r: the tolerance for accepting matrices
Require: m: the dimension of vectors
input: a given series $\{v(1), v(2),\ldots, v(N)\}$
initialization: initialize r as a positive real number, m as a positive integer
Loop:
for $i, j = 1, 2, \ldots, N\text{-}m+1$ **do**
Generate vectors $V_m(i)$ and $V_m(j)$ ► $V_m(i) = [v(i), v(i+1), \ldots, v(i+m-1)]$, $V_m(j) = [v(j), v(j+1), \ldots, v(j+m-1)]$
end for
for $i, j = 1, 2, \ldots, N\text{-}m+1$, $k = 0, 1, \ldots, m\text{-}1$ **do**
Calculate the absolute distance between vector $V_m(i)$ and $V_m(j)$
► $d_{ij}^m = \max\{|v(i+k) - v(j+k)|\}$
Calculate the similarity degree $D_{ij}{}^m$ ► $D_{ij}^m(r) = \begin{cases} 1, & d_{ij} \le r \\ 0, & d_{ij} > r \end{cases}$
Calculate W^m ► the sum of $D_{ij}{}^m$
end for
***m* = *m* + 1 and repeat loop only once**, then calculate W^{m+1}
for $i = 1, 2, \ldots, N\text{-}m$ **do**
Calculated the conditional probability $B_i{}^m(r)$ and $A_i{}^m(r)$
► $B_i^m(r) = \frac{1}{N-m-1} W^m(i)$
► $A_i^m(r) = \frac{1}{N-m-1} W^{m+1}(i)$
Calculate the probability of matching points $B^m(r)$ and $A^m(r)$
► $B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r)$
► $A^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(r)$
end for

Output: the sample entropy (SE) of the input series ► $\mathrm{SE}(m, r, N) = -\ln\left[\frac{A^m(r)}{B^m(r)}\right]$

Fig. E.2. The sample entropy algorithm. Source: [J3].

E.3. FUZZY ENTROPY ALGORITHM

Algorithm 7: Fuzzy entropy

Require: N: the total number of data
Require: r: the tolerance for accepting matrices
Require: m: the dimension of vectors
input: a given series $\{v(1), v(2),\ldots, v(N)\}$
initialization: initialize r as a positive real number, m as a positive integer
Loop:
for $i, j = 1, 2, \ldots, N\text{-}m+1$ **do**
Generate vectors $V_m(i)$ and $V_m(j)$ ► $V_m(i) = [v(i), v(i+1), \ldots, v(i+m-1)]$, $V_m(j) = [v(j), v(j+1), \ldots, v(j+m-1)]$
for $k = 0, 1, \ldots, m\text{-}1$ **do**
Calculate the baseline of the vector $V_m(i)$ ► $v_0(i) = \frac{1}{m}\sum_{k=0}^{m-1} v(i+k)$
end for
Remove the baseline from the generated vectors
end for
for $i, j = 1, 2, \ldots, N\text{-}m+1, k = 0, 1, \ldots, m\text{-}1$ **do**
Calculate the fuzzy distance between vector $V_m(i)$ and $V_m(j)$
► $d_{ij}^m = \max\{|v(i+k) - v(j+k)|\}$
Calculate the similarity degree $D_{ij}{}^m$ ► $D_{ij}^m(r) = \exp(-\log(2) \times (d_{ij}^m / r)^2$
Calculate the similarity degree W^m ► the sum of $D_{ij}{}^m$
end for
***m* = *m* + 1 and repeat loop only once**, then calculate W^{m+1}
for $i = 1, 2, \ldots, N\text{-}m$ **do**
Calculated the conditional probability $B_i{}^m(r)$ and $A_i{}^m(r)$
► $B_i^m(r) = \frac{1}{N-m-1}\sum_{i,j=1}^{N-m+1} d_{ij}^m$, $A_i^m(r) = \frac{1}{N-m-1}\sum_{i,j=1}^{N-m+1} d_{ij}^{m+1}$
Calculate the probability of matching points $B^m(r)$ and $A^m(r)$
► $B^m(r) = \frac{1}{N-m}\sum_{i=1}^{N-m} B_i^m(r)$, $A^m(r) = \frac{1}{N-m}\sum_{i=1}^{N-m} A_i^m(r)$
end for

Output: the fuzzy entropy (FE) of the input series ► $\mathrm{FE}(m,r,N) = -\ln\left[\frac{A^m(r)}{B^m(r)}\right]$

Fig. E.3. The fuzzy entropy algorithm. Source: [J4].

E.4. COARSE-GRAINED PROCEDURE FOR MULTISCALE ENTROPY CALCULATION

Algorithm 8: Coarse-grained procedure

Require: τ: windows of length
Input: time series $\{v(1), v(2), \ldots, v(N)\}$, τ
Divide the input series into non-overlapping windows of length τ
Calculate the average value of the data points inside each window
Obtain the consecutive coarse-grained time series

► $y(j) = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} v(i), \quad 1 \le j \le N/\tau$

Output: the consecutive coarse-grained time series $\{y(1), y(2), \ldots, y(N/\tau)\}$

Fig. E.4. The coarse-grained procedure for multiscale entropy calculation. Source: [J3].

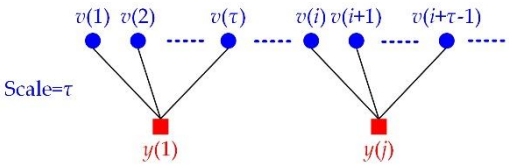


Fig. E.5. Schematic diagram of the coarse-graining process. Source: [J3].

E.5. ENTROPY PARAMETER SELECTION METHOD

Algorithm 9: Minimization of the maximum approximate/sample/fuzzy entropy relative error

Require: CP: the conditional probability obtained during AE, SE, or FE calculation
Require: σ_{CP}: the standard deviation of CP
Require: $\sigma_{g(CP)}$: the standard approximation of $g(CP)$, and $\sigma_{g(CP)} \cong |g'(CP)|\sigma_{CP}$
initialization: $g(CP) = -\log(CP)$, $g(SE) = -\log(-\log(CP))$
Calculate the relative errors of the SE estimates

► $\sigma_{g(\mathrm{SE})} \cong |g'(SE)|\sigma_{CP} = \frac{\sigma_{CP}}{-\log(CP)CP}$

Calculate the relative errors of the CP estimates

► $\sigma_{g(CP)} \cong |g'(CP)|\sigma_{CP} = \frac{\sigma_{CP}}{CP}$

Select r to minimize the quantity $\max\left(\frac{\sigma_{CP}}{-\log(CP)CP}, \frac{\sigma_{CP}}{CP}\right)$

Output: the optimal selected r

Fig. E.6. The entropy parameters selection method. Source: [J4].

APPENDIX F. NOISE SUPPRESION ALGORITHMS

F.1. EMPIRICAL MODE DECOMPOSITION

An IMF is a function that satisfies two the following conditions: i) In the whole data set, the number of extrema and the number of zero crossings must either equal or differ at most by one; ii) At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero [C3].

Algorithm 10: Empirical mode decomposition

Require: *nn*: the number of iterations
Require: c_{nn}: the component
Require: $h_{nn:}$ the residue
Require: N_{EMD}: the total number of IMF
input: Original signal $v(t)$
repeat
Identify the extrema of the input.
Find the upper and lower envelopes $e_{max}(t)$ and $e_{min}(t)$ by cubic spline interpolating.
Calculate the average curve $h(t)$ between the two envelopes ►
$h_1(t) = \frac{1}{2}\left(e_{max}(t) + e_{max}(t)\right)$.
Extract the remainder signal ► $c_1(t) = v(t) - h_1(t)$.
Until any of the following predetermined criteria are met
► Criteria 1: the component, c_{nn}, or the residue, h_{nn}, becomes so small that it is less than the predetermined value of substantial consequence
► Criteria 2: the residue, h_{nn}, becomes a monotonic function from which no more IMF can be extracted
During the process, $h(t)$ is regarded as the IMF and the $h(t)$ in the last step is kept as the residue
Output: a series IMF and the residue ► $\begin{cases} h_2(t) = h_1(t) - c_2(t) \\ \vdots \\ h_{nn}(t) = h_{nn-1}(t) - c_{nn}(t) \end{cases}$, and the decomposed original signal is described as $v(t) = \sum_{n=1}^{N_{EMD}} IMF_n(t) + r(t)$

Fig. F.1. The empirical mode decomposition algorithm. Source: [C3].

F.2. MOVING AVERAGE METHOD

Algorithm 11: Moving average

Require: $2\times l+1$: the window width
Require: x, y: the original data and the smoothed data, respectively
Require: N: the length of the original series
input: a given series $x = \{x_1, x_2, \ldots, x_N\}$, window length $2\times l+1$
for $i = 1, 2, \ldots, N$ **do**
Determine the central point x_i
Calculate the smoothed value ▶ $y_i = \begin{cases} \frac{1}{i}\sum_{k=1}^{i} x_k, & i-l<1 \\ \frac{1}{n-i+1}\sum_{k=i}^{N} x_k, & i+l>N \\ \frac{1}{2l+1}\sum_{k=i-l}^{i+l} x_k, & \text{others} \end{cases}$
end for
Output: a smoothed data series $y = \{y_1, y_2, \ldots, y_N\}$

Fig. F.2. The Moving average algorithm. Source: [C3].

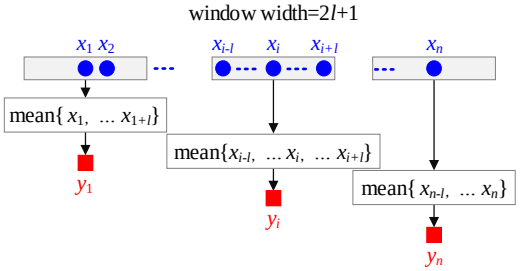


Fig. F.3. Diagram of the moving average method. Source: [C4].

F.3. MOVING MEDIAN METHOD

Algorithm 12: Moving median

Require: $2\times l+1$: the window width
Require: x, y: the original data and the smoothed data, respectively
Require: N: the length of the original series
input: a given series $x = \{x_1, x_2, \ldots, x_N\}$, window length $2\times l+1$
for $i = 1, 2, \ldots, N$ **do**
Determine the central point x_i
Calculate the smoothed value ▶ $y_i = \begin{cases} \text{median}\{x_1 \cdots x_i\}, & i-l>1 \\ \text{median}\{x_i \cdots x_N\}, & i+l>N \\ \text{median}\{x_{i-l} \cdots x_{i+l}\}, & \text{others} \end{cases}$

end for
Output: a smoothed data series $y = \{y_1, y_2, \ldots, y_N\}$

Fig. F.4. The Moving median algorithm. Source: [C3].

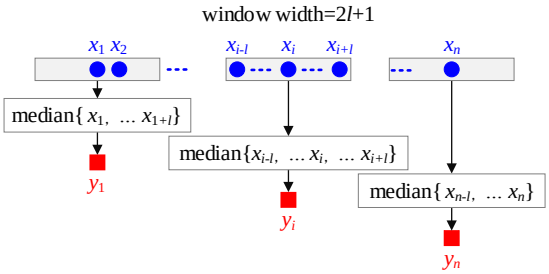


Fig. F.5. Diagram of the moving median method. Source: [C4].

F.4. GAUSSIAN FILTER

Algorithm 13: Gaussian filter

Require: $2\times l+1$: the window width
Require: x, y: the original data and the smoothed data, respectively
Require: N: the length of the original series
Require: σ: the variance of the original series
input: a given series $x = \{x_1, x_2, \ldots, x_N\}$, window length $2\times l+1$, variance σ
for $i = 1, 2, \ldots, N$ **do**

Determine the window span $[i-l, i+l]$, and generate the local feature series $\{x_{i-l}, \ldots, x_i, \ldots, x_{i+l}\}$

Calculate the Gaussian weight ω_i ► $\omega_i = \frac{1}{2\pi\sigma}\exp\left(-\frac{x_i^2}{2\sigma^2}\right)$

Normalize the weight of the central data x_i ► $\omega_i = \frac{\omega_i}{\sum_{k=i-l}^{i+l}\omega_k}$

Calculate the smoothed value as the weighted sum of row data x_i

► $y_i = \sum_{k=i-l}^{i+l} x_k\omega_k$

end for
Output: a smoothed data series $y = \{y_1, y_2, \ldots, y_N\}$

Fig. F.6. The Gaussian filter. Source: [C3].

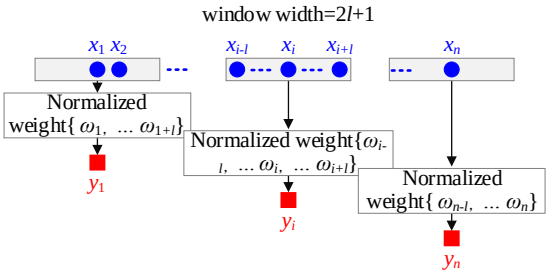


Fig. F.7. Diagram of the Gaussian filter. Source: [C4].

F.5. SAUTZKY-GOLAY FILTER

Algorithm 14: Sautzky-Golay filter

Require: $2\times l+1$: the window width
Require: x, y: the original data and the smoothed data, respectively
Require: N: the length of the original series
input: a given series $x = \{x_1, x_2,\ldots, x_N\}$, window length $2\times l+1$, variance σ
for $i = 1, 2, \ldots, N$ **do**
Determine the window span $[i-l, i+l]$, and generate the local feature series $\{x_{i-l}, \ldots, x_i, \ldots, x_{i+l}\}$
Perform the quadratic polynomial fitting on the data in the span using the linear least squares regression
Calculate the smoothed value y_i by the regression model
end for
Output: a smoothed data series $y = \{y_1, y_2, \ldots, y_N\}$

Fig. F.8. The Sautzky-Golay filter. Source: [C3].

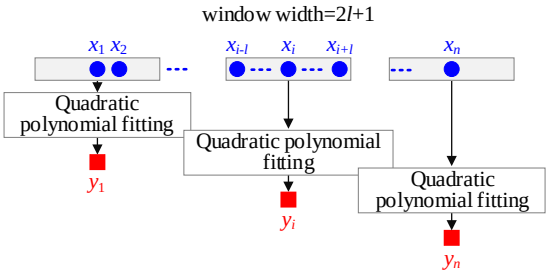


Fig. F.9. Diagram of the Sautzky-Golay filter. Source: [C4].

F.6. LOCALLY WEIGHTED SCATTERPLOT SMOOTHING

Algorithm 15: Locally weighted scatterplot smoothing

Require: $2\times l+1$: the window width
Require: x, y: the original data and the smoothed data, respectively
Require: N: the length of the original series
Require: σ: the variance of the original series
Require: x_i and x_j: the central data and the its neighbor, respectively
Require: dd_j: the maximum distance between x_i and x_j
input: a given series $x = \{x_1, x_2, \ldots, x_N\}$, window length $2\times l+1$, variance σ
for $i = 1, 2, \ldots, N$ **do**

Determine the window span $[i-l, i+l]$, and generate the local feature series $\{x_{i-l}, \ldots, x_i, \ldots, x_{i+l}\}$

Calculate the regression weights ω_j for each data in the window span by the tricube weight function

► $$\omega_j = \begin{cases} \left(1-\left(\dfrac{x_i - x_j}{dd_j}\right)^3\right)^3, & \left|\dfrac{x_i - x_j}{dd_j}\right| < 1 \\ 0, & \left|\dfrac{x_i - x_j}{dd_j}\right| \geq 1 \end{cases}$$

Perform the first-degree polynomial fitting on the data in the span using the weighted least-squares regression ► $S = \sum \omega_j \left(a + bx_j - y_j\right)^2$

Perform the quadratic polynomial fitting on the data in the span using the linear least squares regression

Calculate the smoothed value y_i by the weighted regression model
end for
Output: a smoothed data series $y = \{y_1, y_2, \ldots, y_N\}$

Fig. F.10. The locally weighted scatterplot smoothing. Source: [C3].

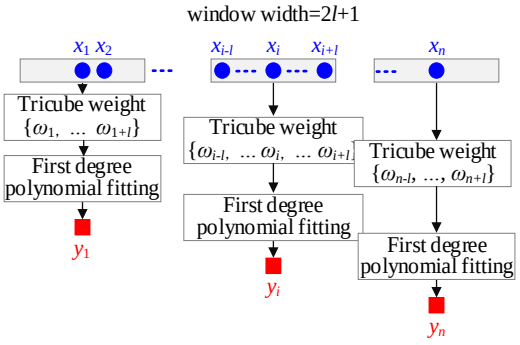


Fig. F.11. Diagram of locally weighted scatterplot smoothing. Source: [C4].

F.7. ROBUST LOCALLY WEIGHTED SCATTERPLOT SMOOTHING

Algorithm 16: Robust locally weighted scatterplot smoothing

Require: $2\times l+1$: the window width
Require: x, y: the original data and the smoothed data, respectively
Require: N: the length of the original series
Require: σ: the variance of the original series
Require: x_i and x_j: the central data and the its neighbor, respectively
Require: s: the median of the residuals in the span $[i\text{-}l, i+l]$
input: a given series $x = \{x_1, x_2, \ldots, x_N\}$, window length $2\times l+1$, variance σ
for $i = 1, 2, \ldots, N$ **do**

Determine the window span $[i\text{-}l, i+l]$, and generate the local feature series $\{x_{i-l}, \ldots, x_i, \ldots, x_{i+l}\}$

Calculate the regression weights ω_j for each data in the window span by the 'bisquare' function

► $$\omega_j = \begin{cases} \left(1-\left(\dfrac{x_i - x_j}{6s}\right)^2\right)^2, & |x_i - x_j| < 6s \\ 0, & |x_i - x_j| \geq 6s \end{cases}$$

Perform the first-degree polynomial fitting on the data in the span using the weighted least-squares regression ► $S = \sum \omega_j \left(a + bx_j - y_j\right)^2$

Perform the quadratic polynomial fitting on the data in the span using the linear least squares regression

Calculate the smoothed value y_i by the weighted regression model

end for
Output: a smoothed data series $y = \{y_1, y_2, \ldots, y_N\}$

Fig. F.12. The robust locally weighted scatterplot smoothing. Source: [C3].

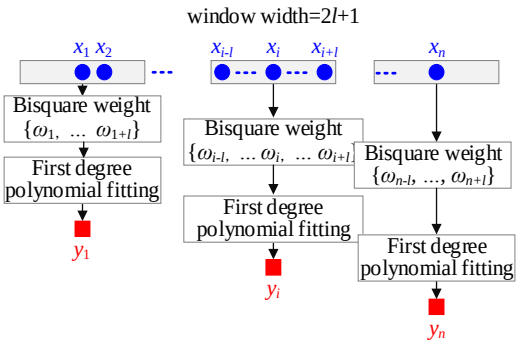


Fig. F.13. Diagram of the robust locally weighted scatterplot smoothing. Source: [C4].

PART II. PAPERS

JOURNAL PAPER I

[J1] A Review of Non-Probabilistic Machine Learning-Based State of Health Estimation Techniques for Li-Ion Battery

X. Sui, S. He, S.-B. Vilsen, J. Meng, R. Teodorescu, and D.-I. Stroe

The paper has been published in
Appl. Energy, vol. 300, pp. 117346, Oct. 2021.
DOI: 10.1016/j.apenergy.2021.117346.

JOURNAL PAPER II

[J2] The Degradation Behavior of $LiFePO_4$/C Batteries During Long-Term Calendar Aging

X. Sui, M. Świerczyński, R. Teodorescu, and D.-I. Stroe

The paper has been published in
Energies, vol. 14, no. 6, pp. 1732, Mar. 2021.
DOI: 10.3390/en14061732.

JOURNAL PAPER III

[J3] The Effect of Voltage Dataset Selection on the Accuracy of Entropy-Based Capacity Estimation Methods for Lithium-Ion Batteries

X. Sui, D.-I. Stroe, S. He, X. Huang, J. Meng, and R. Teodorescu

The paper has been published in
Appl. Sci., vol. 9, no. 19, pp. 4170, Oct. 2019.
DOI: 10.3390/app9194170.

JOURNAL PAPER IV

[J4] Fuzzy Entropy-Based State of Health Estimation for Li-Ion Batteries

X. Sui, S. He, J. Meng, R. Teodorescu, and D.-I. Stroe

The paper has been published in
IEEE Trans. Emerg. Sel. Topics Power Electron., vol. 9, no. 4, pp. 5125-5137, Aug. 2021. DOI: 10.1109/JESTPE.2020.3047004.

CONFERENCE PAPER I

[C1] State of Health Estimation for Lithium-Ion Battery Using Fuzzy Entropy and Support Vector Machine

X. Sui, S. He, D.-I. Stroe, and R. Teodorescu

The paper has been published in the
Proc. IEEE ECCE Asia, pp. 1417-1422, Dec. 2020.
DOI: 10.1109/IPEMC-ECCEAsia48364.2020.9368182.

CONFERENCE PAPER II

[C2] Fuzzy Entropy-Based State of Health Estimation of $LiFePO_4$ Batteries Considering Temperature Variation

X. Sui, S. He, J. Meng, R. Teodorescu, and D.-I. Stroe

The paper has been published in the
Proc. IEEE ECCE, pp. 4401-4406, Oct. 2020.
DOI: 10.1109/ECCE44975.2020.9236267.

CONFERENCE PAPER III

[C3] Lithium-Ion Battery State of Health Estimation Using Empirical Mode Decomposition Sample Entropy and Support Vector Machine

X. Sui, S. He, D.-I. Stroe, and R. Teodorescu

The paper has been published in the
Proc. IEEE APEC, pp. 3424-3429, Mar. 2020.
DOI: 10.1109/APEC39645.2020.9124327.

CONFERENCE PAPER IV

[C4] Data Smoothing in Fuzzy Entropy-Based Battery State of Health Estimation

X. Sui, S. He, X. Huang, R. Teodorescu, and D.-I. Stroe

The paper has been published in the
Proc. IEEE IECON, pp. 1779-1784, Oct. 2020.
DOI: 10.1109/IECON43393.2020.9255281.

CONFERENCE PAPER V

[C5] Fast and Robust Estimation of Lithium-Ion Batteries State of Health Using Ensemble Learning

X. Sui, S. He, S.-B Vilsen, R. Teodorescu, and D.-I. Stroe

The paper has been published in the
Proc. IEEE ECCE, pp. 1779-1784, Nov. 2021.
10.1109 / ECCE47101.2021.9595113

ISSN (online): 2446-1636
ISBN (online): 978-87-7573-972-1 AALBORG UNIVERSITY PRESS