Research on key issues of health evaluation based on data-driven method for lithium-ion batteries

by

Jiwei Wang

Supervisor: Professor Dr Abuliti Abudula

A Dissertation Presented to

the Graduate School of Science and Technology

of Hirosaki University in partial fulfillment of the requirement for

the degree of Doctor of Philosophy in Safety Science and Technology

HIROSAKI UNIVERSITY

2022

ABSTRACT

In recent years, to resolve the global energy crisis and environmental pollution problems, major economies and automobile manufacturers in the world have strengthened the research and development of electric vehicles. As one of the most important systems of electric vehicles, a battery management system that ensures the accurate estimation of the key states of traction batteries is the key to achieve the safe and reliable operation of the batteries. The aging process of lithium-ion batteries is a dynamic coupling process affected by various factors, such as temperature and current rate. The internal aging mechanism is very complex and the electrochemical parameters, such as the property of battery materials, are not easy to be obtained, which greatly increases the difficulty of establishing the mechanism model. To improve the state of health (SOH) estimation accuracy for lithium-ion batteries, the following studies on data-driven methods has been carried out in this dissertation.

1. **Analysis and test of the aging characteristics of lithium-ion batteries.** A battery aging test platform is built to explore the influence of different factors on battery aging. A lithium-ion battery aging test procedure is designed, including aging tests at different ambient temperatures, charging rates, and depths of discharge. Based on the battery cycling aging test and the battery calendar aging test, statistical analysis and comparative research on the aging characteristics of lithium-ion batteries are carried out. Two empirical aging models are constructed to calculate the capacity loss of the batteries.

2. **Estimation of the SOH of lithium-ion battery cells based on the improved Gaussian Process Regression (GPR) model.** The single kernel function based on the GPR model is not accurate in estimating the SOH of lithium-ion batteries because it cannot accurately capture the phenomenon of overall capacity decay and local capacity

recovery. Herein, a method for estimating the SOH of lithium-ion batteries based on the modified GPR combined with charging and discharging features is proposed. First, the changes in battery voltage and temperature curves among different aging cycles are analyzed in detail, and health indicators (HIs) that can effectively represent the health status of the battery are proposed. Then, the Pearson correlation analysis method is used to quantify the correlation between the HIs and SOH, and three HIs with strong correlation are employed in this study. Next, a novel compound kernel function is proposed for the battery SOH estimation, and different pairs of mean function and kernel function chosen from four mean functions and sixteen kernel functions are used to construct the GPR models, and their estimation accuracy is compared subsequently. Finally, four different batteries with various initial health levels from the NASA battery dataset are used to verify the performance of the proposed method. Experiments show that its estimated mean-absolute-error (MAE) and root-mean-square-error (RMSE) are only 1.7%, and 2.41%, respectively. Compared with a single kernel function, the GPR model based on a composite kernel function is more suitable for capturing the battery aging characteristics of various trends and can achieve an accurate estimation of the battery SOH.

3. **Battery pack health estimation based on early data.** For the estimation of the SOH of the battery pack, it takes a lot of time and economic cost to complete the full-life aging test of the battery pack. Aiming at solving above problems, a method for predicting the future health of the battery pack is proposed, which involves the aging data of the battery cells over their entire life cycles and the early cycling data of the battery pack. Firstly, the HIs are extracted from the experimental data, and high correlations between the extracted HIs and the capacity are verified by the Pearson correlation analysis method. To predict the future health of the battery pack based on

the HIs, degradation models of HIs are constructed by using an exponential function, long short-term memory network, and their weighted fusion. The future HIs of the battery pack are predicted according to the fusion degradation model. Then, based on the GPR algorithm and the battery pack data, a data-driven model is constructed to predict the health of the battery pack. Finally, the proposed method is validated by a series-connected battery pack with fifteen 100 Ah lithium iron phosphate battery cells. The MAE and RMSE of the health prediction of the battery pack are 7.17% and 7.81%, respectively, indicating that the proposed method has satisfactory accuracy. Compared with the single feature decay model, the fusion feature decay model can predict the future HIs of the battery pack with more accuracy, which contributes to the satisfactory prediction accuracy of battery pack health based on the GPR model.

Keywords: electric vehicle; lithium-ion battery; health indicator; state of health; data-driven; Gaussian process regression

ACKNOWLEDGEMENTS

Sincere thanks to all the experts who worked so hard to review the dissertation!

Thank you all very much.

   Jiwei Wang

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Introduction

1.1 Research background and significance

1.1.1 Background

With modern society's continuous progress and rapid development, human beings rely on energy considerably. According to the study of *World Energy Statistics Yearbook 2021* [1], the global fossil energy consumption in 2020 occupies 83.1% of the total disposable energy consumption, among which oil, coal, and natural gas occupy 31.2%, 27.2%, and 24.7% of the total energy structure, respectively. The excessive consumption of fossil fuels has caused problems such as accelerated depletion of fossil fuel energy, increased greenhouse effect, and atmospheric pollution. More than 100 countries and regions worldwide have reached a broad consensus to address the global energy crisis and environmental pollution and proposed implementing the "zero-carbon" or "carbon neutral" goal. By March 2022, *the Energy & Climate Intelligence Unit* has calculated the progress of the world's major countries and regions toward zero emissions, including two countries that have achieved "carbon neutrality," 12 countries have enacted legislation, and 32 countries have published policy documents. Some countries and regions have participated in the "carbon neutral" global program, as shown in Table 1.1. In addition, 61 countries and regions are still in the proposal or discussion stage.

Table 1.1. Many countries and regions have globally participated in the "Carbon Neutral" initiative.

Type of commitment	Specific countries and regions (planning time)	amount
Implemented	Suriname, Bhutan	2
In law	Germany (2045), Sweden (2045), Japan (2050), France (2050), UK (2050), Korea (2050), Canada (2050), Spain (2050), Ireland (2050), Denmark (2050), Hungary (2050), New Zealand (2050), EU (2050)	13
In policy document	Maldives (2030), Finland (2035), Antigua and Barbuda (2040), United States (2050), Italy (2050), Australia (2050), Belgium (2050), Romania (2050), Austria (2050), Chile (2050), Portugal (2050), Greece (2050), Ecuador (2050), Panama (2050), Croatia (2050), Lithuania (2050), Costa Rica (2050), Slovenia (2050), Uruguay (2050), Luxembourg (2050), Latvia (2050), Laos (2050), Malta (2050), Fiji (2050), Belize (2050), the Marshallese (2050), Monaco (2050), Turkey (2050), China (2060), Ukraine (2060), Sri Lanka (2060)	32
Declaration /pledge	Brazil (2050), Thailand (2050), Argentina (2050), Malaysia (2050), Vietnam (2050), Colombia (2050), South Africa (2050), United Arab Emirates (2050), Kazakhstan (2050), Israel (2050), Estonia (2050), Cape Verde (2050), Andorra (2050), Russia (2060), Saudi Arabia (2060), Nigeria (2060), Bahrain (2060), India (2070)	18

Data Sources: ECIU

Statista research shows that the transportation systems account for 17% of global greenhouse gas emissions and is the fastest-growing source of greenhouse gas emissions in the world [2]. With the further development of economic globalization, the scale of future human and logistics flows will continue to increase for a considerable period, and hundreds of millions of fossil fuel vehicles will generate a large amount of greenhouse gases into the atmosphere every year and accelerate the climate crisis. To effectively solve the problem and reduce the reliance on traditional fossil fuels, the development of new energy vehicles has reached a world consensus. China, America, Japan, Germany, and other countries are developing rapidly in the research and

development of new energy vehicles. Major countries and regions (cities) around the world have proposed plans to ban the sale of fuel vehicles, as shown in Table 1.2.

Table 1.2. Time of fuel vehicle prohibition plan in different countries (regions/cities) worldwide.

Country/City	proposed time	Presenting method	Implement-ation time	Prohibited range
Norway	2016	State planning	2025	Gasoline/diesel vehicles
Netherlands	2016	Motions	2035	Gasoline/diesel passenger cars
Paris, Madrid, Athens, Mexico	2016	Mayor signs action agreement	2025	Diesel vehicles
California, USA	2018	Government Decree	2029	Fuel Bus
Germany	2016	Motions	2030	Internal combustion locomotive
France	2017	Oral statements by government officials	2040	Gasoline/diesel vehicles
UK	2017	Oral statements by government officials	2040	Gasoline/diesel vehicles
	2018	Transportation Sector Strategy		
	2020	Government documents	2030	
Scotland, UK	2016	Government documents	2032	Gasoline/diesel vehicles
India	2016	Oral statements by government officials	2030	Gasoline/diesel vehicles
Taiwan, China	2016	Government Action Program	2040	Gasoline/diesel vehicles
Hainan, China	2016	Government Planning	2030	Gasoline/diesel vehicles
Ireland	2016	Oral statements by government officials	2030	Gasoline/diesel vehicles
Rome, Italy	2016	Oral statements by government officials	2024	Diesel vehicles
Japan	2020	Government documents	2035	Gasoline/diesel vehicles

1.1.2 Electric vehicles

The automobile industry is undergoing a big revolution. The internal combustion

engine (ICE) system will be upgraded by electric vehicles (EVs) in the future, even though it has ruled the automobile industry for a century. From 2010 to 2021, global EV growth is shown in Figure 1.1. Although the number of EVs has grown rapidly over the past seven years, the total number is still relatively small. In the latter five years, the number of EVs grew much more than before. Growth has been particularly impressive over recent three years, even as the global *COVID-19* pandemic shrank the market for conventional cars. In particular, the sales of electric vehicles are more than double to 6.6 million in 2021. All of the net growth in global vehicle sales last year came primarily from EVs.
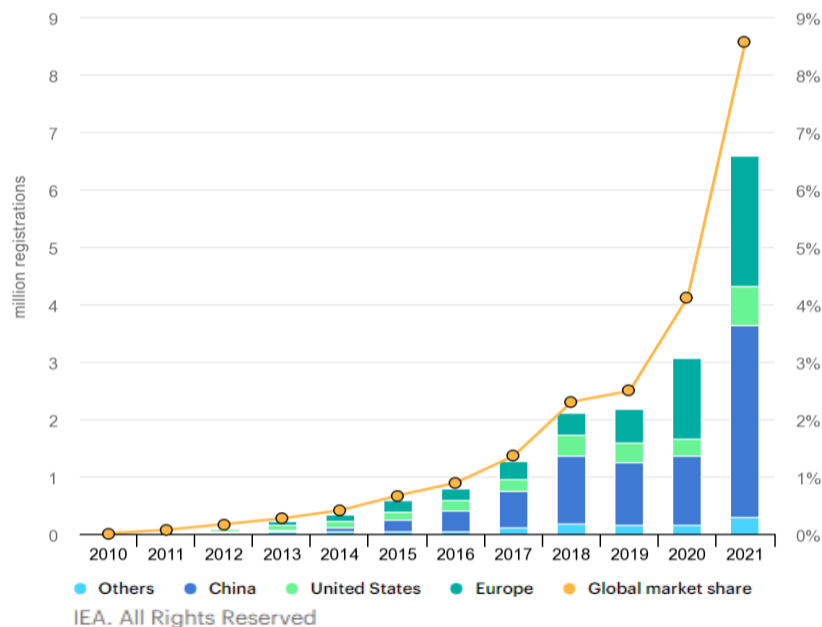
Figure 1.1. Global sales and sales market share of electric cars, 2010-2021.

Compared with traditional fuel vehicles, EVs have become an important part of transportation electrification due to their better economy and environmental friendliness. According to different sources of driving power, EVs can be divided into four categories [3]: battery electric vehicles (BEVs), hybrid electric vehicles (HEVs), plug-in hybrid electric vehicles (PHEVs), and fuel cell electric vehicles (FCEVs), as shown in Figure 1.2.

Figure 1.2. Representative vehicles of four different types of EVs.

(a) BEV © Tesla; (b) HEV ©BYD; (c) PHEV ©VW; (d) FCEV © Toyota

(a) BEV

A BEV means that the driving power comes entirely from the electric motor. The electric energy driving the motor comes from the on-board rechargeable battery or other electric energy storage device. Compared with fuel cars and other types of EVs, BEVs have the highest energy efficiency, cheap electricity, and less cost. In addition, BEVs can be charged by using the low valley of electricity at night, which has the function of regulating the peak and valley loads of the grid system and improving the grid efficiency.

(b) HEV

The HEVs are new energy hybrid vehicles that combine a traditional ICE with an electric propulsion system. HEVs have two relatively independent propulsion systems. Through the power compensation of the motor drive system and the power adjustment of "peak-shaving and valley-filling," the engine can frequently work in a high-efficiency and low-emission area. Compared with traditional fuel vehicles, HEVs improve energy conversion efficiency and reduce fuel consumption.

(c) PHEV

PHEV is one type of hybrid vehicle that can be recharged by plugging in. It has

pure electric and hybrid modes, which can be considered as the fusion of EVs and HEVs. Compared with HEVs, this type of vehicles have larger capacity power battery packs, more powerful electric motor drive systems, and smaller displacement engines.

(d) FCEV

An FCEV uses electrical energy generated by an electrochemical reaction in a fuel cell with a catalyst between a fuel, such as a hydrogen and oxygen in the air. In contrast to typical EVs, FCEVs are powered by an onboard fuel cell unit. This type of vehicles have the advantages of energy saving, environmental friendliness, high efficiency, and long overload capacity. FCEVs have become one of the most promising development directions. However, there are still problems such as high cost and complicated auxiliary equipment that need to be further solved.

The comparison of the advantages and disadvantages of the above four types of electric vehicles is shown in Table 1.3 [4].

Table 1.3. The comparison of advantages and disadvantages of four different types of EVs.

Evaluation Indicators	BEV	HEV	PHEV	FCEV
Drive mode	motor	ICE + motor	ICE + motor	motor
Carbon emission	zero	lower	lower	near zero
Noise	lower	low	low	lower
Cost	normal	higher	higher	highest
Range	normal	farther	farther	normal
Energy System	electrical energy	oil/electrical energy	oil/electrical energy	hydrogen
Typical company-model	Tesla-Model Y	VW-Tiguan	BYD-Tang	Toyota-Mirai

1.1.3 Power battery technology

Power battery, as one of the core technologies of EVs, is the primary power source of EVs. The driving range, acceleration performance, and braking energy recovery rate of EVs is inseparable related to the performance of the power battery, which are reflected in the strong hill-climbing ability, acceleration ability, and range of EVs. The

main function of the power battery is significantly different for different types of EVs. For BEVs, the power battery is the only power source, mainly used to drive EVs for longer mileage and to supply the starting, acceleration, and climbing of EVs with high discharge current, which requires higher specific energy of the power battery. For the HEVs, the power battery mainly plays an auxiliary role in supplementing the power required for acceleration and hill-climbing so that HEVs have better acceleration and hill-climbing performance. The power capability of batteries is also very important [5]. With the continuous progress of power battery technology, EVs' safety, economy, and power have obvious advantages compared with traditional cars.

(1) The performance of power battery

Power battery in the EV driving process should meet the following requirements [6] [7]:

High energy density. Higher energy density can effectively reduce the mass and volume of batteries, enabling lighter weight and more extended cruising range for EVs.

High power density. The high power density of batteries can provide enough instantaneous power and improve the power performance of the EV.

Long cycle life. The long cycle life of the battery can effectively reduce the cost of using EVs and replacement frequency.

Excellent charge and discharge characteristics. The ideal charging and discharging characteristics of the battery can shorten the charging time and improve the usage performance. When the car working in braking mode, it gets higher braking energy recovery efficiency. To avoid battery over-charging and over-discharging, which can extend the battery's service life to some extent. In addition, the better discharge characteristics have a strong continuous power supply capability.

Good battery consistency. Good battery consistency refers to the small variability

of the performance, which can reduce the adverse consequences of significant differences in battery performance during battery pack use and rapid aggravation.

(2) Power battery type

Power batteries in the field of EVs can be divided into lead-acid batteries, nickel metal hydride batteries, lithium-ion batteries, and fuel cells.

The lead-acid battery is invented by French scientist Gaston Planté in 1859 and is the first commercially available rechargeable battery. The basic composition of this battery is a positive electrode plate (PbO_2), negative electrode plate (Pb), and electrolyte (H_2SO_4 aqueous solution), and its internal chemical energy and direct current (DC) energy are converted to each other, in other words, the battery can be reused by charging and discharging. Lead-acid batteries are mainly used as ignition devices for car starting. Due to their low energy density, short cycle life, irreversible environmental pollution caused by a large amount of lead emissions during production and processing, and recycling and other shortcomings, they are not suitable for the power supply of modern electric vehicle.

Nickel-metal hydride battery is a new green battery developed in the 1990s. It has the advantages of high energy, long life, low pollution, and good low-temperature characteristics, and its all-around performance is better than that of lead-acid batteries. Therefore, NiMH batteries are used in batch for HEVs under Toyota, Honda, Ford, and other large automobile brands. However, NiMH batteries have not eliminated the memory effect. There are also new problems, such as severe heating during charging, which may cause safety accidents and poor performance in high-current charging.

Lithium-ion battery is a kind of secondary battery that mainly relies on the movement of lithium ions between the positive and negative electrodes. It mainly comprises a positive electrode, negative electrode, electrolyte, separator, etc.[8]. Since

1991, Sony Japan launched commercial lithium-ion batteries, and lithium-ion battery technology has been developed rapidly. Lithium-ion batteries have higher energy density and power density than NiMH batteries. In addition, Lithium-ion battery owns long cycle life, low self-discharge rate, high operating voltage and have other excellent performance.

The fuel cell is a device that converts the chemical energy of a fuel directly into electrical energy through electrode reaction, which is the fourth type of power generation technology after hydroelectric, thermal, and nuclear power generation. As a new type of power generation technology, the reaction process of fuel cells does not involve combustion. So this kind of battery has the remarkable characteristics of high efficiency and cleanliness. Based on the above advantages, the research and development of fuel cell technology are highly valued by governments, which will become the clean and efficient power generation technology in the 21st century. Today, fuel cell developers Ballard Power Systems, international fuel cell companies, and automobile manufacturers DaimlerChrysler, Ford Motor Company, Toyota Motor Corporation, and Volkswagen have joined hands to develop fuel cells and fuel cell vehicles.

A comparison of the performance of four types of power batteries currently applicable to EVs is in Table 1.4 [9, 10]. Lead-acid batteries are inexpensive and relatively abundant in resources, but their specific energy and power density are low. Compared with lead-acid batteries, the performance of NiMH batteries has been improved obviously, but their charging and discharging efficiency and specific energy efficiency are still not satisfactory. Although the fuel cell is one of the most environmentally-friendly batteries, its cost has to be reduced, and further improving the performance of the fuel cell is still the focus of future research. Through comprehensive comparison,

the lithium-ion battery has the best all-around performance. At present, the lithium-ion battery has been widely used in BEVs and hybrid vehicles.

Table 1.4. Comparison of performance parameters of four different power batteries.

Battery Type	Specific energy density(Wh/Kg)	Specific power density(Wh/L)	Specific power(W/Kg)	Lifetime （cycle）	Cost (USD/Kwh)
Lead-acid battery	30~50	60~100	200~400	400~600	120~150
NiMH battery	60~120	100~300	200~400	600~1200	150~300
Lithium-ion battery	120~200	140~280	200~450	1000~2000	150~180
Fuel Cell	300~400	50~200	30~1000	5k~10k (h)	200~240

(3) Battery performance indicators

As the power source of EVs, the battery plays an essential role in EVs, and the actual effect of the battery is mainly evaluated through its indicators of the battery. The performance indicators for assessing the battery are voltage, capacity, internal resistance, energy, output efficiency, self-discharge rate, service life, etc. [11].

Voltage refers to the potential difference between the positive and negative electrodes and is one of the critical parameters used to characterize the performance and state of the battery.

The amount of electricity that can be released from the battery under certain discharge conditions is called the capacity of the battery, which can be calculated by the product of discharge current and discharge time. Battery capacity can be divided into actual capacity and rated capacity. The actual capacity is equal to the product of discharge current and discharge time. Its value is less than the theoretical value. The parameter reflects the size of the battery's actual storage capacity, which is proportional to the range of EVs.

The internal resistance of the battery refers to the assistance of the current flowing

through the battery. The higher the internal resistance of a battery, the more energy it consumes and the lower the efficiency of its use. As the number of times the battery is charged and discharged increases, the internal resistance of the battery will increase to different degrees due to the consumption of electrolytes and the reduction of chemical activity inside the battery.

The battery's energy refers to the electrical energy that the battery can output under certain discharge conditions. The battery's specific energy is a comprehensive index to reflect the quality level of the battery and affect the overall quality and range of the electric vehicle. Energy density refers to the electric energy output per unit volume of the battery.

The battery's power refers to the energy output per unit time under certain discharge conditions, which can evaluate the electric vehicle's acceleration performance and hill-climbing ability. The power output per unit mass of the battery becomes the specific power, and the power output per unit volume is called power density.

Self-discharge rate refers to the rate of capacity decline of the battery during storage, i.e., the battery without load is the rate of capacity loss due to its discharge.

SOC is the ratio of the remaining power to the battery's rated capacity under a specific discharge rate, and the relative amount of the percentage of SOC is usually used to indicate the changing state of charge in the battery.

SOH refers to the percentage of the maximum available capacity of the battery to the rated capacity; the higher the value, the better the health of the battery.

Battery lifetime includes calendar life and cycle life, and the remaining useful life (RUL) can be obtained by subtracting the used life from the calibrated life.

1.1.4 Battery management technology

Battery Management System (BMS) is a device used to safely monitor and effectively manage the battery pack and improve battery usage efficiency. For EVs, its main task is to ensure the safety and reliability of the power battery system, provide the status information required for vehicle control and energy management, and take effective intervention measures for the power battery system when abnormalities occur. Through the system's effective control of battery pack charging and discharging, it can achieve increased range, extended service life, reduced operating costs, and ensure the safety and reliability of power battery applications. The main functions of BMS include battery state detection, battery state analysis, battery safety protection, energy control management, and battery information management, as shown in Figure 1.3 [12, 13].
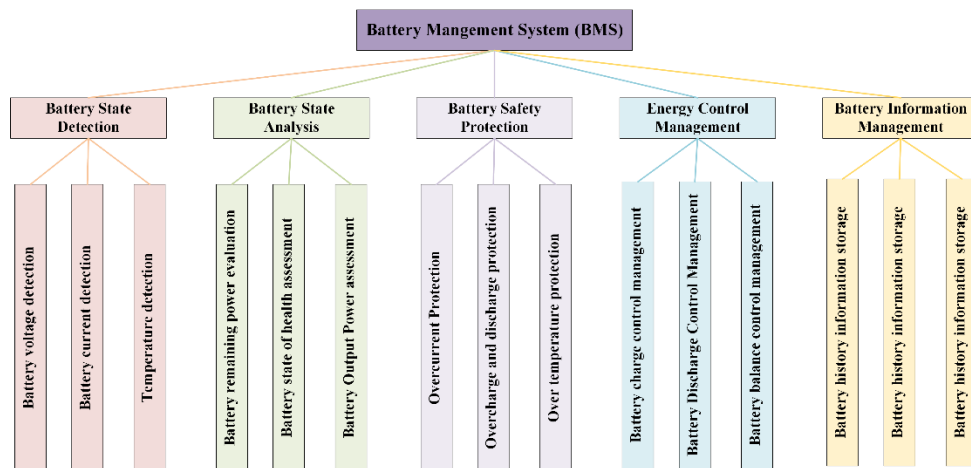
Figure 1.3. Basic functions of the battery management system.

(1) Battery state detection

Detecting physical quantities such as voltage, current, and temperature is the most basic function of BMS, so collecting accurate and reliable information is very important.

(2) Battery state analysis

Battery state analysis includes state of charge (SOC), state of health (SOH), and

state of power (SOP). Among them, SOC is the percentage of available capacity of the battery to the rated capacity. SOH reflects the aging degree of the battery. It is usually used to assess the battery's health status by the declining capacity or increasing internal resistance. SOP indicates the power that the battery can provide or absorb for the vehicle drivetrain in an instantaneous time.

(3) Battery safety protection

Battery safety protection mainly refers to the safety control of the battery, mainly including over-current protection, over-charge, and over-discharge protection, over-temperature protection. For example, lithium batteries are prone to permanent damage under the working conditions of overcharge and over-discharge, resulting in capacity decay or even fire and explosion, so the safety protection function of BMS is essential.

(4) Energy control management

Energy control management includes charging, discharging, and equalization control management. As the electric vehicle driving conditions are very complex, emergency acceleration, braking, up and down hills, and other driving operations of the random trigger will cause complex and variable dynamic load. Battery charging and discharging management provides real-time optimal control of voltage, current, and other parameters during battery charging and discharging, which can bring out the more excellent performance of the power battery pack. Battery balance control management refers to the practical measures to reduce the adverse effects caused by inconsistent batteries, optimize the battery pack's overall discharge efficiency and extend the battery pack's overall life.

(5) Battery information management

The battery management system needs to integrate several functional modules and reasonably coordinate the communication between the modules. The BMS needs to

process the operating data of the power battery. The above data is used to notify the driver through the meter, is transmitted to other controllers or the vehicle cloud through the communication network, and is stored as the intermediate historical data of the battery management system.

1.2 Battery modeling theory

To monitor the operation status of lithium-ion batteries in real-time, establishing a more accurate battery model has become a key link to achieving battery health management. Since the power battery is a very complex electrochemical system with many factors and uncertainties, its mathematical modeling is a multidisciplinary intersection problem and has become a hot spot. Currently, battery models mainly include the electrochemical mechanism, equivalent circuit, and data-driven models.

1.2.1 Electrochemical model

The electrochemical model describes the ion transport and electrochemical reaction processes inside the battery using complex and mutually coupled nonlinear partial differential equations. The model can not only predict the variation of the external voltage of the cell more accurately, but also model the microscopic parameters such as electrolyte over potential, electrode current density, solid-phase ion concentration, and the rate of internal side reactions [14]. It has obvious advantages in the research, prediction and evaluation of electrochemical reaction/coupling process, but the model involves many nonlinear equations, and the calculation is complicated. Reducing the computational load and developing simplified electrochemical models have become essential research directions. Currently, the two commonly used simplified electrochemical models are the pseudo-two-dimensional (P2D) model and the single-particle (SP) model.

(1) P2D model

In the mid-1990s, M. Doyle et al. at the University of California, Berkeley, created the P2D model with porous electrodes and solution theory, laying the foundation for developing the electrochemical mechanism model. The model assumes equal current density on the battery poles and believes that the battery can be approximated as multiple identical structures (sandwich structures) inside the battery. The whole battery is simplified as a sandwich structure. The schematic diagram of the P2D model of the lithium-ion battery, as shown in Figure 1.4 [15], mainly consists of a copper collector, negative electrode coating, diaphragm, positive electrode coating, and aluminum collector.

Meanwhile, the model simplifies the active battery material as a uniform sphere and uses Fick's diffusion law to describe the radial diffusion of lithium-ion particles to produce another dimension. P2D model can accurately describe the internal processes such as the Faraday effect, lithium-ion diffusion within the active particles, electrochemical reactions on the active particle surface, lithium-ion diffusion and migration in the electrolyte, and Ohm's law in the anode and diaphragm [15, 16].
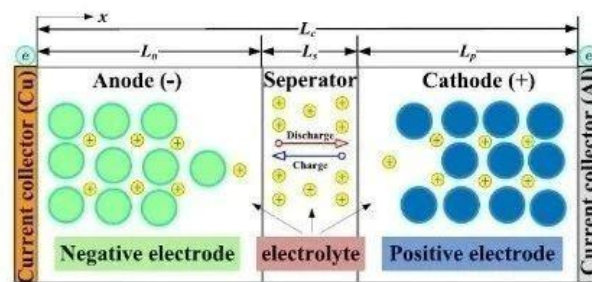
Figure 1.4. Schematic diagram of the P2D model.

(2) SP model

To reduce the complexity of the mechanism model, White et al. proposed a single-particle model based on the P2D model, and the structure is shown in Figure 1.5 [17]. This schematic is based on the simplification of liquid-phase mass transfer and the

assumption of uniformity of current density, and the simulation of the overall cell behavior is achieved by replacing the whole electrode with a single active substance particle.

In the P2D model, the lithium-ion diffusion process integrates the ion diffusion in the solid phase active material and the ion transport in the electrolyte phase. However, the solid-phase diffusion time factor is usually several orders of magnitude higher than the liquid-phase diffusion time factor [18]. The effect of the electrolyte on the cell behavior is simplified and equated to the liquid-phase diffusion resistance in the SP model to reduce the number of solution domains and improve computational efficiency [19].
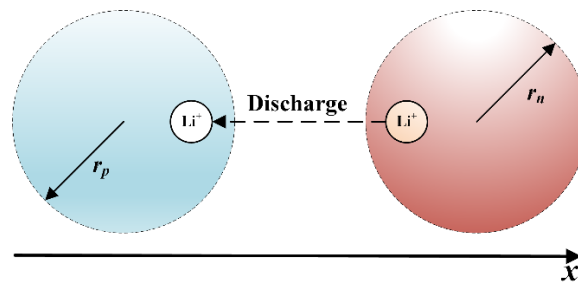
Figure 1.5. Schematic diagram of the SP model for lithium-ion battery.

With the continuous development of model simplification and numerical calculation methods, electrochemical mechanism models have received more and more attention and applications in the field of lithium-ion battery simulation and design in recent years. However, in the process of model construction, it is difficult to obtain material property parameters such as electrical conductivity and diffusion coefficient. Therefore, it is still difficult to transfer to the existing BMS platform for online applications.

1.2.2 Equivalent circuit model

The equivalent circuit model uses circuit elements such as resistors, capacitors,

and voltage sources to form a circuit network based on the physical characteristics to describe the static and dynamic characteristics of the battery. The equivalent circuit model has the characteristics of small calculation, easy parameter identification, simple model, etc. In recent years, the study of battery state estimation based on the equivalent circuit model has become one of the hot spots in academic research, and the commonly used equivalent circuit models include the R_{int} model, Thenenin model, PNGV model, and 2RC network model [20].

(1) R_{int} model

R_{int} model is the simplest equivalent circuit model for the battery, as shown in Figure 1.6, where the voltage source V_{OC} represents the open-circuit voltage, the ohm internal resistance R_0 of the battery can be calculated from the current ratio of the voltage across the battery, I is the load current, and V_b is the terminal voltage.
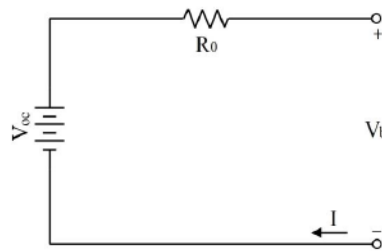
Figure 1.6. Schematic diagram of R_{int} model.

The output voltage versus input current for the R_{int} model is shown in Equation (1.1).

$$V_{b} = V_{OC} - IR_{o} \tag{1.1}$$

The R_{int} model is used as the basic equivalent circuit model, which has the advantage of a simple model equivalent circuit and easy parameter identification, but it cannot accurately simulate the transient characteristics of the battery. Especially under high current pulses, the polarization effect is ignored. Under complex conditions such as short circuits, overload, or transient high power output, the model accuracy is poor.

Therefore, the R_{int} model is only suitable for simulation environments with small current variations but lays the foundation for building a more accurate higher-order equivalent circuit model.

(2) Thevenin model

The Thevenin model is shown in Figure 1.7, which mainly consists of an open-circuit voltage source V_{OC}, an RC parallel network characterizing the battery polarization phenomenon, and an ohm internal resistance R_0. Compared with the R_{int} model, this model considers both the transient and steady-state characteristics of the battery, where the open-circuit voltage source is used to simulate the steady-state characteristics of the battery, while the RC parallel network and the ohm resistance R_0 are used to simulate the transient characteristics of the battery.
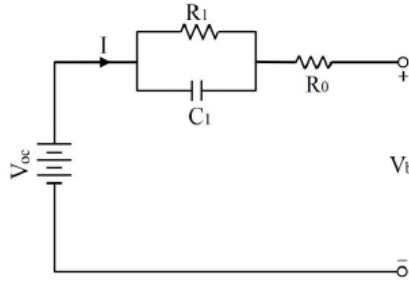
Figure 1.7. Schematic diagram of Thevenin model.

According to Kirchhoff's law, the relationship between voltage and current for the Thevenin model is shown in Equation (1.2).

$$\begin{cases} \dfrac{dV_1}{dt} = \dfrac{I}{C_1} - \dfrac{V_1}{R_1 C_1} \\ V_b = V_{OC} - V_1 - IR_o \end{cases} \tag{1.2}$$

Thevenin model is based on the R_{int} model with the addition of RC parallel networks characterizing the effects of capacitance and resistance generated during the electrolysis of the battery, which can describe the dynamic characteristics of the battery. Due to the moderate accuracy and computational complexity of the model, it is suitable

for modeling the battery characteristics.

(3) PNGV model

The PNGV model is the standard equivalent circuit model for PNGV as proposed by the United States Council for Automotive Research (USCAR) in *the PNGV Battery Test Manual* published in 2001. This model adds a series equivalent capacitor C_b to the Thevenin equivalent circuit model to describe the open-circuit voltage change with time accumulation of load current, as shown in Figure 1.8. Where V_{oc} is the open-circuit voltage, R_0 is the ohm internal resistance, and R_1 and C_1 represent the polarized internal resistance and polarized capacitance of the battery, respectively.
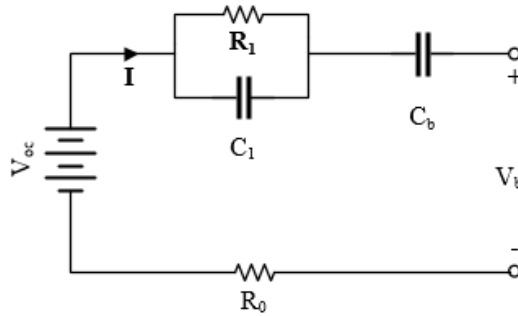
Figure 1.8. Schematic diagram of PNGV model.

According to Kirchhoff's law, the PNGV model corresponds to the expression shown in Equation (1.3).

$$V_{b} = V_{OC} - \frac{1}{C_b}(\int Idt) - V_1 - IR_o \tag{1.3}$$

The PNGV model has a higher computational complexity than the Thevenin model, and the relative accuracy is improved to some extent.

(4) 2RC network model

The 2RC network model adds an *RC* network to the Thevenin model for representing the diffusion effect of the cell, as shown in Figure 1.9. Where V_{oc} denotes the open-circuit voltage, R_0 denotes the ohm internal resistance, R_1 and C_1 denote the polarization internal resistance and polarization capacitance of the cell, respectively,

and R_2 and C_2 denote the diffusion internal resistance and diffusion capacitance of the cell, respectively.
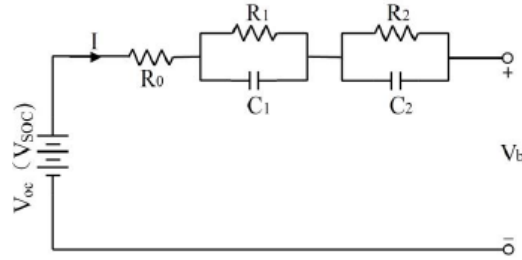
Figure 1.9. Schematic diagram of the second-order equivalent circuit model.

According to Kirchhoff's law, the PNGV model corresponds to the expression shown in Equation (1.4).

$$V_b = V_{OC} - R_0 I - V_1 - V_2 \tag{1.4}$$

The model can simulate the dynamic performance of the battery more accurately. However, due to the addition of an RC network, the model parameters of the battery and the dimensionality of the state space equations are increased, so the computational complexity and computational time of the model are better than those of the Thevenin model.

1.2.3 Data-driven model

Data-driven model is one of the methods to realize the health state estimation of the Lithium-ion battery, which is analyzed from the perspective of the external characteristics of the Lithium-ion battery and uses the current, voltage, and other data in the charge/discharge characteristic curve to build the battery health state estimation model. The method does not need to consider the complex electrochemical processes inside the battery. The main methods to achieve battery health state estimation are based on data-driven models including support vector machine (SVM), relevance vector machine (RVM), Gaussian process regression (GPR), and artificial neural network

(ANN).

  (1) SVM

SVM is a modeling method based on statistical learning theory that maps the input space to a high-dimensional feature space through a nonlinear transformation to obtain a nonlinear relationship between the input and output variables [21, 22]. Today, it has become an essential tool in machine learning and data mining, and the algorithm has gained widespread applications in text classification, handwriting recognition, image classification, bioinformatics, etc. The SVM model that is usually used for regression estimation is shown in Equation (1.5) [23].

$$f(x) = \omega\varphi(x) + b, x \in R^m, b \in R \tag{1.5}$$

where the input space is denoted by R^m while satisfying x belongs to R^m and $f(x)$ belongs to R. With the help of $\varphi(x)$, x can be mapped into the feature space, $\varphi(x)$ represents the nonlinear mapping, and w, b denote the weights and intercepts of the hyperplane, respectively.

To apply SVM to solve the regression problem, the insensitive loss function is introduced, while the SVM model requires the minimization of the ε-insensitive loss function, as in Equation (1.6).

$$L(f(x), y) = \begin{cases} 0 & , |y - f(x)| \le \varepsilon \\ |y - f(x)| - \varepsilon, & \text{otherwise} \end{cases} \tag{1.6}$$

where $f(x)$ denotes the predicted value, y denotes the observed value, and ε denotes the allowable error between the true value and the estimated value. The loss function of Equation (1.6) defines an ε-tube centered on $f(x)$, 2ε-insensitive loss function is shown in Figure (1.10). If the difference between the observed and predicted values of the sample are less than ε, the loss is equal to 0, the higher the accuracy of the regression model, while for samples outside the ε-band region, they are considered by slackening

the variables ξ, ξ^* and a non-negative penalty coefficient C.
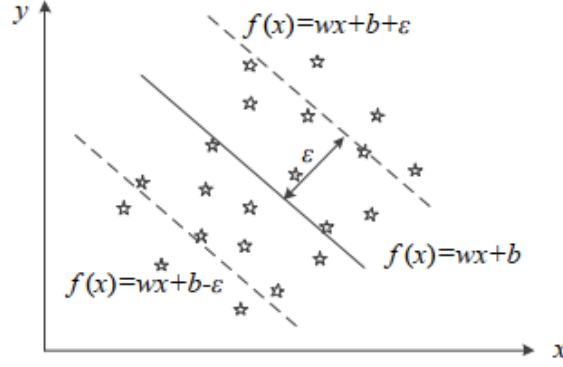
Figure 1.10. Principle of the insensitive loss function.

Minimizing the ε-insensitive loss is equivalent to solving the optimization problem for

w and b. The problem of regression estimation using a standard SVM can be expressed

as an optimization problem as in Equation (1.7) [24].

$$\begin{cases} \min\limits_{w,b,\xi} \dfrac{1}{2}\|w\|^2 + C\sum\limits_{i=1}^{l}(\xi_i + \xi_i^*) \\ \quad y_i - w\varphi(x_i) - b \le \varepsilon + \xi_i \\ s.t. \quad w\varphi(x_i) + b - y_i \le \varepsilon + \xi_i^*, \quad i = 1,2,...,l \\ \quad \xi_i, \xi_i^* \ge 0 \end{cases} \tag{1.7}$$

where C denotes the penalty factor to balance the mean square error and the

generalizability of the model over the test data; ε denotes the radius of the pipeline; and

ξ_i, ξ_i^* denotes the slack variables corresponding to sample x_i.

For the optimization problem of Equation (1.7), the Lagrange Function (1.8) is

introduced.

$$L(w,b,\xi_i,\xi_i^*,\alpha_i,\alpha_i^*,\beta_i,\beta_i^*) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) + \sum_{i=1}^{l}\alpha_i[y_i - w\varphi(x_i) - b - \varepsilon - \xi_i]$$
$$+ \sum_{i=1}^{l}\alpha_i^*[y_i - w\varphi(x_i) - b - \varepsilon - \xi_i^*] - \sum_{i=1}^{l}\beta_i\xi_i - \sum_{i=1}^{l}\beta_i^*\xi_i^* \tag{1.8}$$

where $\alpha_i, \alpha_i^*, \beta_i, \beta_i^*$ is the Lagrange multiplier.

According to the Karush-Kuhn-Tucker (KKT) condition, the problem of Equation

(1.7) can be transformed into a pairwise optimization problem, as in Equation (1.9).

$$\begin{cases} \max_{\alpha,\alpha^*} & \sum_{i=1}^{l} y_i(a_i - a_i^*) - \varepsilon \sum_{i=1}^{l} y_i(a_i + a_i^*) - \frac{1}{2}\sum_{i,j}^{l}(a_j - a_j^*)(a_i - a_i^*)K(x_i,x) \\ & \sum_{i=1}^{l}(a_i - a_i^*) = 0, \\ s.t. & a_i \geq 0, \\ & a_i^* \leq C, \end{cases} \tag{1.9}$$

By solving Equation (1.8), the regression function is obtained as in Equation (1.10).

$$f(x) = \sum_{i=1}^{l} \alpha_i k(x_i,x) + b \tag{1.10}$$

where $k(\cdot,\cdot)$ denotes the kernel function that satisfies the Mercer condition.

The main thought of SVM is to solve the problem. First, the original data set is mapped to a high-dimensional space by introducing some kernel function, which makes the data linearly separable in the high-dimensional space. However, SVM still has some shortcomings, such as the prediction results do not contain probabilistic information to determine the uncertainty of the prediction results; the kernel function $k(x, x_i)$ must satisfy the Mercer condition. The above shortcomings become a bottleneck for SVM to improve the accuracy of results in the problem-solving process, which restricts the application scope of SVM and its practical application in engineering.

(2) RVM

To address the shortcomings of SVM, Michael E. Tipping proposed RVM, a machine learning method based on Bayesian theory, Markov property, auto-correlation decision prior, and maximum likelihood theory [25]. This method has high sparsity and probability-based learning characteristics, which overcomes the SVM kernel function and must satisfy the Mercer condition, the basic principle of RVM is described below [26].

Assuming that the training sample $D = \{(x_1,t_1),(x_2,t_2),...,((x_N,t_N))\}$, N is the

23

number of samples, the regression model of RVM is as in Equation (1.11).

$$y(x; \boldsymbol{w}) = \sum_{i=1}^{N} w_i k(x, x_i) + \varepsilon_i \qquad (1.11)$$

Where $k(x,x_i)$ is the kernel function, $w=\{w_i\}_{i=1}^{N}$ represents the different weights, and ε_i is the Gaussian white noise $\varepsilon_i \sim N(0,\sigma^2)$ that satisfies the normal distribution, and the probability distribution of the target value can be derived, as in Equation (1.12).

$$p(t_i \mid x) = N(t_i \mid y(x_i; \boldsymbol{w}), \sigma^2) \qquad (1.12)$$

Where $N(\cdot)$ denotes a Gaussian distribution function based on an expectation of $y(x_i; w)$, variance σ^2 over t_i, and it is reasonable to assume that the targets t_i are mutually independent random variables, and the likelihood function of the entire training set is as in Equation (1.13).

$$p(\boldsymbol{t} \mid \boldsymbol{w}, \sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp\{-\frac{1}{2\sigma^2}\|\boldsymbol{t} - \boldsymbol{\phi}\boldsymbol{w}\|^2\} \qquad (1.13)$$

where $\boldsymbol{w} = (w_0, ..., w_N)^T$ is the weight vector, $\boldsymbol{\phi} = [\phi(x_1), \phi(x_2), ..., \phi(x_N)]^T$ is an $N \times (N+1)$ kernel matrix, $\boldsymbol{\phi}(x_n) = (1, K(x_n, x_1), ..., K(x_n, x_N))^T$, where $K(x_n, x_i)$ is the kernel function. w, σ^2 can be obtained by performing the maximum likelihood estimation of Equation (1.13).

In RVM, to improve the generalization ability, a hyper-parameter vector $\boldsymbol{\alpha}$ is introduced, and a Gauss prior distribution with zero mean is assigned to the weight vector w as in Equation (1.14).

$$p(\boldsymbol{w} \mid \boldsymbol{\alpha}) = \prod_{i=0}^{N} N(w_i \mid 0, \alpha^{-1}) \qquad (1.14)$$

Suppose the sample to be tested is t_*, and the distribution of the corresponding predicted values is as in Equation (1.15).

$$p(t_* \mid t) = \int p(t_* \mid \boldsymbol{w}, \boldsymbol{\alpha}, \sigma^2) p(\boldsymbol{w}, \boldsymbol{\alpha}, \sigma^2 \mid t) dw d\alpha d\sigma^2 \qquad (1.15)$$

The derivation of Equation (1.15) from Markov's property and Bayes' theorem leads to Equation (1.16).

$$p(\boldsymbol{w} \,|\, t, \boldsymbol{\alpha}, \sigma^2) = (2\pi)^{-\frac{N+1}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(w-\mu)^T \Sigma^{-1}(w-\mu)\right\} \qquad (1.16)$$

where the covariance matrix $\Sigma = (\sigma^2 \phi^T \phi + A)^{-1}$, the mean of the weights $\mu = \sigma^{-2} \Sigma \phi^T t$, and the diagonal array $A = diag(\alpha_0, \alpha_1, ..., \alpha_N)$.

The partial differential of Equation (1.16) is found and made zero. The hyper-parameter and variance update Equations are obtained as in Equation (1.17) (1.18).

$$\alpha_i^{new} = \frac{\gamma_i}{\mu_i^2} \qquad (1.17)$$

$$(\sigma^2)^{new} = \frac{\|t - \phi\mu\|^2}{N - \Sigma_i \gamma_i} \qquad (1.18)$$

where $\gamma_i = 1 - \alpha_i \Sigma_{ii}$, Σ_{ii} are the *i-th* diagonal elements. The hyper-parameters $\boldsymbol{\alpha}$ and σ^2 estimates can be obtained through continuous iterations, while the weighted posterior mean $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$ are continuously updated until the convergence requirement is satisfied (the hyper-parameter values stop changing or the maximum number of iterations is reached), thus obtaining the optimal values of the variables $\boldsymbol{\alpha}_{MP}$ and σ_{MP}^2.

In the iterative process, most of the hyper-parameters α_i will tend to infinity, and the corresponding weights will slowly converge to zero. Therefore, some of the kernel functions with zero weights will be removed from the model and will not have any effect on the prediction results. Therefore, the correlation vector machine has a more sparse performance than the support vector machine, and the other α_i will steadily converge to a finite value, and the corresponding x_i is called the correlation vector.

Finally, the expectation is y_* according to Equation (1.15), and the noise

variance is σ_*^2.

$$y_* = \mu^T \phi(x_*) \tag{1.19}$$

$$\sigma_*^2 = \sigma_{MP}^2 + \phi(x_*)^T \sum \phi(x_*) \tag{1.20}$$

(3) GPR

GPR is a nonlinear modeling method in the field of machine learning. It has a rigorous statistical learning theoretical foundation and can effectively handle the identification of high-dimensional nonlinear complex systems [27]. Since Gaussian process regression is a model based on statistical learning, its properties are completely determined by its mean function *m(x)* and covariance function *k(x,x')*, assuming a random process *f(x)*, the Gaussian process can be defined as Equation (1.21).

$$f(x) \sim GP(m(x), k(x, x')) \tag{1.21}$$

where the mean value function *m(x)* represents the expected value *E(f(x))* of the function *f(x)* at the input point *x*. In general, the mean function specifies the central tendency of the potential function, but in most practical applications, there is usually a lack of a priori knowledge of the mean function *m(x)*, which is usually set zero in consideration of symmetry. The covariance matrix contains information about the shape and structure of the potential function, which is determined by the covariance function and its corresponding parameters. The covariance function *k(x, x')* is a measure of the confidence level of the mean function *m(x)* and is denoted by *E[(f(x) - m(x))(f(x) - m(x))]*.

For the regression problem, the model can be expressed as an Equation (1.22).

$$y = f(x) + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2) \tag{1.22}$$

where *f(x)* is the value of the function corresponding to the input vector *x*, ε is Gaussian white noise with mean taken as zero and variance σ^2, and *y* denotes the measured value.

From this, the prior distribution of y is obtained as in Equation (1.23).

$$y \sim N(0, K(X,X) + \sigma^2 I) \tag{1.23}$$

According to the Gaussian process definition, the joint prior distribution of the observed value y and the predicted value y_* is as in Equation (1.24).

$$\begin{bmatrix} y \\ y_* \end{bmatrix} \sim \left(0, \begin{bmatrix} K(X,X) + \sigma^2 I & K(x',X) \\ K(X,x') & \mathbf{k}(x',x') \end{bmatrix} \right) \tag{1.24}$$

where $K(X, X) = K_{n*n} = (k_{ij})$ is the $n \times n$ order positive definite and symmetric covariance matrix, $K(X, x) = K(X, x)^T$ denotes the $n \times 1$ order covariance matrix between the test sample input x' and the training sample input x and $K(x', x')$ is the covariance scalar of the test input.

According to the Bayesian formula, the posterior probability can be calculated to obey the Gaussian distribution as in Equation (1.25).

$$y_* \mid X, y, x' \sim N(\bar{y}_*, \mathrm{cov}(y_*)) \tag{1.25}$$

where the predicted mean and corresponding variance are Equation (1.26) and (1.27), respectively.

$$\bar{y}_* = K(X, \mathbf{x}')(K(X,X)) + \sigma^2 I)^{-1} y \tag{1.26}$$

$$\mathrm{cov}(y_*) = \mathbf{k}(, \mathbf{x}', \mathbf{x}') - (K(X, \mathbf{x}'))[K(X,X)] + \sigma^2 I)^{-1} y \tag{1.27}$$

$\bar{y}_*$ is the predicted mean of the GPR model at the test point, and the confidence interval

of the prediction result is determined by the variance of the covariance matrix $\mathrm{cov}(y_*)$.

  (4) ANN

ANN is an information processing system established by human beings based on understanding the operation law of neural networks in the brain, which can imitate the structure of neural networks in the brain to achieve specific functions [28]. Since ANN

is an adaptive nonlinear dynamic system formed by many linear or nonlinear neurons following a certain method of interconnection, the learning process of the neural network is to continuously adjust the connection weights among the neurons according to the input patterns, so as to form a memory of the input patterns. Therefore, ANN has the advantages of high speed of information processing, high fault tolerance, nonlinear operation, and self-learning. The neuron model with n input data is shown in Figure 1.11. The model mainly includes input data, weights and deviations, activation function, and output data.
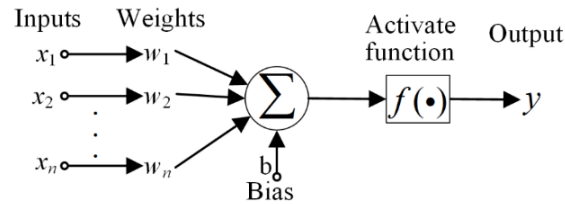
Figure 1.11. Schematic diagram of the neuron model.

In the neuron model, $X=(x_1,x_2,...,x_n)$ is the input of the neuron; $W=(w_1,w_2,...,w_n)$ is the weight value corresponding to each input; b is the bias value of the neuron; $f(\bullet)$ is the activation function of the neuron and y is the output of the neuron, as shown in Equation (1.28).

$$y = f(\bullet) = \sum_{i=1}^{n} w_i x_i + b \qquad (1.28)$$

ANN can be divided into convolutional neural networks and recurrent neural networks according to their principles and structures, and this paper introduces recurrent neural networks as an example. A recurrent neural network (RNN) is a neural network model proposed by Saratha Sathasivam in 1982 for processing sequential problems [29]. In this neural network, the hidden layer neurons receive information not only from the input layer but also from the hidden layer at the previous moment, and the basic structure of the RNN network is shown in Figure 1.12.
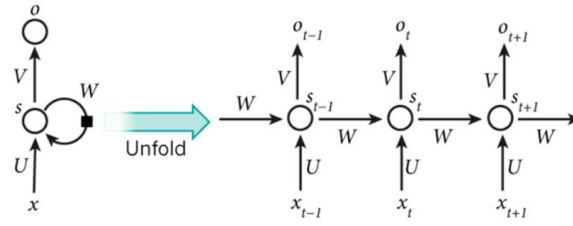
Figure 1.12. Schematic diagram of RNN structure.

Where x_t denotes the input of the *t-th* sequence and s_t is the state of the *t-th* step of the hidden layer, which is the memory unit of the network. s_t is calculated based on the output of the current input layer and the state of the hidden layer of the previous step. It can be considered that the hidden layer state s_t is the memory unit of the network; s_t contains the hidden layer states of all previous steps. o_t is the output of the *t-th* step, and the output o_t of the output layer is only related to s_t of the current step. In the RNN, each layer shares parameters *U*, *V*, and *W*, which greatly reduces the number of parameters to be learned in the network and increases the training speed. RNN input-to-hidden connections are parameterized by the weight matrix *U*, the hidden-to-hidden cyclic connections are parameterized by the weight matrix *W* and the hidden-to-output connections are parameterized by the weight matrix *V*. The left side of Figure (1.12) represents the RNN drawn using cyclic connections, and the right side of Figure (1.12) represents the computational plot of the expansion.

Based on the analysis and comparison of the above three different models, it can be seen that the electrochemical mechanism model has a clear physical meaning of parameters and has the advantages of high accuracy and good reliability. However, the model parameters are difficult to obtain and the calculation time is long, so it cannot be matched with the existing BMS hardware environment for online applications. The equivalent circuit model, as a semi-empirical semi-mechanical model, can meet the accuracy application requirements with limited computational resources and has lower empirical dependence and better scalability and reliability [15]. The data-driven

approach is based on learning and analyzing the historical data of the research object with the help of data-driven theories and techniques. The raw data are transformed into relevant information and behavioral models for battery state estimation and prediction. Especially for nonlinear problems with obvious advantages, the use of data-driven models for battery modeling can better establish the relationship between battery input and output.

1.3 Research content and structure

1.3.1 Main research content

In this dissertation, based on the problems related to the health status of lithium-ion batteries, three targeted research points are summarized. Firstly, battery aging experiments are built to analyze the effects of temperature, rate, and state of charge on battery aging, and based on analyzing the battery aging mechanism, relevant characteristics of battery aging are summarized. Secondly, based on the improved Gaussian process regression, a novel composite kernel function is proposed to capture the overall decay trend and local regeneration characteristics of battery aging, and a battery single cell health state estimation model is established by combining three different groups of HIs. Finally, three new sets of HIs are proposed based on the whole-life data of battery cells and the early-life data of battery packs, and a decay model of the fused characteristics of battery pack HIs is constructed; combined with Gaussian process regression, the estimation of the battery pack health state is realized. The main relevant studies are as follows:

（1）**Analysis and test of the aging characteristics of lithium-ion batteries.** Given the influence of different environmental factors on battery aging, a battery aging test platform is built. A lithium-ion battery aging test procedure is designed, including

aging tests at different temperatures, different charging rates, and different depths of discharge. Based on the battery cycle aging test and the battery calendar aging test, statistical analysis and comparative research on the aging characteristics of lithium-ion batteries are carried out.

（2）**Estimation of the SOH of lithium-ion battery cells based on the improved GPR model.** The single kernel function based on the GPR model is not accurate in estimating the SOH of the lithium-ion battery, and cannot accurately capture the phenomenon of overall capacity decay and local capacity recovery. A method for estimating the SOH of lithium-ion batteries based on modified GPR combined with charging and discharging features is proposed. First, the changes in battery voltage and temperature curves among different aging cycles are analyzed in detail, and health indicators (HIs) that can effectively represent the health status of the battery are proposed. Then, the Pearson correlation analysis method is used to quantify the correlation between HIs and SOH, and three HIs with strong correlation are employed in this paper. Next, a novel compound kernel function is proposed for battery SOH estimation, and different pairs of mean function and kernel function chosen from four mean functions and sixteen kernel functions are used to construct GPR models, and their estimation accuracy is compared subsequently. Finally, four different batteries with various initial health conditions from the NASA battery dataset are used to verify the performance of the proposed method. Experiments show that its estimated mean-absolute-error (MAE) and root-mean-square-error (RMSE) are only 1.7%, and 2.41%, respectively. The results show that compared with a single kernel function, the GPR model based on a composite kernel function is more suitable for capturing the battery aging characteristics of various trends and can achieve an accurate estimation of the battery SOH.

（3） **Battery pack health estimation based on early data.** For the estimation of the SOH of the battery pack, it takes a lot of time and economic cost to complete the full-life aging test of the battery pack. Aiming at the above problems, a method for predicting the future health of the battery pack is proposed, which uses the aging data of the battery cells over their entire life cycles and the early cycling data of the battery pack. Firstly, HIs are extracted from the experimental data, and high correlations between the extracted HIs and the capacity are verified by the Pearson correlation analysis method. To predict the future health of the battery pack based on the HIs, degradation models of HIs are constructed by using an exponential function, long short-term memory network, and their weighted fusion. The future HIs of the battery pack are predicted according to the fusion degradation model. Then, based on the GPR algorithm and the battery pack data, a data-driven model is constructed to predict the health of the battery pack. Finally, the proposed method is validated with a series-connected battery pack with fifteen 100 Ah lithium iron phosphate battery cells. The MAE and RMSE of the health prediction of the battery pack are 7.17% and 7.81%, respectively, indicating that the proposed method has satisfactory accuracy. The results show that compared with the single feature decay model, the fusion feature decay model can predict the future HIs of a battery pack with more accuracy, which contributes to the satisfactory prediction accuracy of battery pack health based on the GPR model.

1.3.2 Chapter arrangement

Around the research content and main work of this dissertation, the content of each chapter is arranged as follows.

Chapter 1. First, the global background and significance of this research are outlined, and EVs are introduced as an important part of future green transportation.

The characteristics of four different types of EVs are analyzed and compared. The key role played by battery management technology in the field of EVs is focused on. Then, the main research results on battery modeling theory are reviewed and the advantages and disadvantages of the three main battery models are compared.

Chapter 2. The aging characteristics of lithium-ion batteries are verified and analyzed through experiments. Firstly, the battery aging experimental platform is built, and the cycle aging experiment and calendar aging experiment is conducted on Lithium-ion batteries to verify the effects of different temperature, rate, SOC, and other indexes on battery aging. Secondly, the common methods for extracting battery characteristic factors are analyzed and compared.

Chapter 3. Estimation of the SOH of lithium-ion battery cells based on the improved GPR model. The single kernel function based on the GPR model is not accurate in estimating the SOH of a lithium-ion battery, and cannot accurately capture the phenomenon of overall capacity decay and local capacity recovery. A method for estimating the SOH of lithium-ion batteries based on modified GPR combined with charging and discharging features is proposed.

Chapter 4. Battery pack health estimation based on early data. For the estimation of the SOH of the battery pack, it takes a lot of time and economic cost to complete the full-life aging test of the battery pack. Aiming at the above problems, a method for predicting the future health of the battery pack is proposed, which uses the aging data of the battery cells over their entire life cycles and the early cycling data of the battery pack.

Chapter 5. The research work and related results in the full paper are summarized, and the research related to the health status of lithium-ion batteries prospects.

References

[1] Pimonenko, T.V., O.V. Liulov, and N.Y. Letunovska, Circular and carbon-free economy: road map for harmonizing national and European standards for energy market. 2021.

[2] Kwan, S.C., Hashim, J.H. A review on co-benefits of mass public transportation in climate change mitigation. *Sustainable Cities and Society*. 2016. **22**: p. 11-18.

[3] Zhang Lei, Research on modeling and state estimation algorithms for supercapacitors for electric vehicles. 2016, PhD thesis, Beijing Institute of Technology.

[4] Guo Xiangwei, Research on battery charge state estimation and equalization technology for electric vehicles. 2016, PhD thesis, South China University of Technology.

[5] Ma, Y., *XIN NENG YUAN QI CHE DONG LI DIAN CHI JI SHU*. 2end ed. 2020: Peking University Press.

[6] Liu, Shulin, Study on modeling and state estimation of lithium-ion power battery for electric vehicles. 2017, PhD thesis, Shandong University.

[7] By R. Xiong, Core algorithms for power battery management systems. 2018: Beijing: Machinery Industry Press.

[8] (USA) by Davide Andria; translated by Li Jianlin et al. Large-scale lithium-ion battery management systems. 2016: Beijing: Machinery Industry Press.

[9] Li Quan, Research on key technologies of lithium-ion power battery management system. 2017, PhD thesis, Hunan University.

[10] Zhang Chi, Research on the aging characteristics and state estimation algorithm of automotive power lithium-ion batteries. 2019, PhD thesis, Wuhan University of Technology.

[11] Wang, Shuai, Research on data-driven method for predicting the remaining life of lithium-ion batteries. 2017, PhD thesis, Harbin Institute of Technology.

[12] Tan Xiaojun, Electric Vehicle Power Battery Management System Design. 2011: Guangzhou: Sun Yat-sen University Press.

[13] Yulong Shao, Research on key technologies and construction of hardware-in-the-loop test system for electric vehicle BMS. 2018, PhD thesis, Jilin University.

[14] Jokar, A., et al., Review of simplified Pseudo-two-Dimensional models of lithium-ion batteries. *Journal of Power Sources.* 2016. **327**: p. 44-55.

[15] Shen, Chia-Ni, Study on model development and charge state estimation method for lithium-ion batteries. 2017, PhD thesis, Shanghai Jiaotong University.

[16] Kim, U.S., C.B. Shin, and C.-S. Kim, Modeling for the scale-up of a lithium-ion polymer battery. *Journal of Power Sources.* 2009. **189**(1): p. 841-846.

[17] Haran, B.S., B.N. Popov, and R.E. White, Determination of the hydrogen diffusion coefficient in metal hydrides by impedance spectroscopy. *Journal of Power Sources.* 1998. **75**(1): p. 56-63.

[18] Zheng, G., B.N. Popov, and R.E. White, Electrochemical determination of the diffusion coefficient of hydrogen through an LaNi4. 25Al0. 75 electrode in alkaline aqueous solution. *Journal of the Electrochemical Society.* 1995. **142**(8): p. 2695.

[19] Luo, W., et al., A new extension of physics-based single particle model for higher charge–discharge rates. *Journal of Power Sources.* 2013. **241**: p. 295-310.

[20] Shang, Yunlong, Design and Implementation of Optimization of Vehicle Li-ion Power Battery State Estimation and Equalization Management System. 2017, PhD thesis, Shandong University.

[21] Zhu Fa, A study of some problems in support vector machines. 2019, PhD thesis,Nanjing University of Technology.

[22] Deng, Z., et al., Maximum available capacity and energy estimation based on support vector machine regression for lithium-ion battery. *Energy Procedia.*2017. **107**: p. 68-75.

[23] Vapnik, V., The nature of statistical learning theory. 1999: Springer science & business media.

[24] Wang, F.N., Research on Robustness Models and Algorithms for Support Vector Machines. 2015, PhD thesis,China Agricultural University.

[25] Tipping, M.E., Sparse Bayesian learning and the relevance vector machine. *Journal of machine learning research*.2001. **1**(Jun): p. 211-244.

[26] Wu, Jing, et al., Study on soft measurement modeling based on dynamic multicore correlation vector machine. *Journal of Chemical Engineering*. 2019. 70(04): p. 1472-1484.

[27] Rasmussen, C. and C. Williams, GaussianProcessesforMachineLearning. AdaptiveComputation and Machine Learning. 2006, The MIT Press, Cambridge.

[28] Haykin, S.; Ye, S. W.; Shi, Z. S., Principles of neural networks. 2004: Beijing: Machinery Industry Press.

[29] Zeyu Zheng, Bowen Liang, Siyu Gu, TensorFlow Hands-on Google Deep Learning Framework 2nd ed. 2018: Beijing: Electronic Industry Press.

Chapter 2. Lithium-ion battery aging characteristics

In 1985, Aki Yoshino et al. discovered that carbon material could be used as the anode material of lithium rechargeable batteries. They proposed a carbon material with a graphite structure instead of a lithium metal anode by using lithium and transition metal composite oxide as the battery's cathode material[1]. Since the first commercialization of lithium-ion batteries by Sony Corporation in Japan in 1991, lithium-ion batteries have been widely used in many fields such as electric vehicles, mobile electronic devices, and large-scale energy storage due to their excellent performance in terms of energy density, power density, cycle life, and environmental friendliness[2, 3]. In recognition of their outstanding contributions to the field of lithium-ion batteries, Akira Yoshino et al. were awarded the Nobel Prize in Chemistry in 2019.

However, the battery performance will inevitably decay when using lithium-ion batteries repeatedly. When the battery's state of health no longer meets the actual requirements, the battery needs to be replaced promptly.   Since battery cost is a high percentage for many applications, extending and predicting battery life can significantly increase its economic value. So, it is necessary to quantify lithium-ion battery aging, determine the degradation mechanism, and establish aging models for life diagnosis and prediction.

2.1 Introduction of this chapter

A lithium-ion battery is a dynamic and time-varying electrochemical system with a nonlinear and complex internal reaction mechanism. The main manifestation is that the specific performance of lithium-ion batteries shows a gradual deviation from normal indicators, including capacity decay or diving, shortened service life, deterioration of multiplier performance, and increased self-discharge rate [4]. To ensure the stability of

the battery system, different types of Lithium-ion batteries must work in a specific voltage window interval. The upper voltage limit of the battery is determined by the upper limit of the electrochemical window of the organic electrolyte. The lower voltage is determined by the reduction potential characteristics of the cathode material itself, and the actual working voltage of the battery beyond the voltage window is usually called over-charge or over-discharge. When the battery is overcharged, the organic electrolyte solvent is oxidized and decomposed on the cathode surface, resulting in the electrolyte deviating from the optimal state and the ionic conductivity decreasing. When the battery is over-discharged, it easily leads to the reduction of transition metal cations in the cathode material and the destruction of the original lattice structure, resulting in battery capacity decay [5].

2.1.1 Lithium-ion battery working principle

Lithium-ion battery is a kind of secondary battery in which lithium-ions can go back and forth between the positive and negative electrodes and de-embed, which is named "rocking chair battery". It is mainly composed of a positive electrode, negative electrode, diaphragm, and electrolyte, and its working principle is shown in Figure 2.1 [6]. When charging, the lithium atoms in the positive electrode ionize into lithium ions and electrons that are removed from the lattice of the positive electrode material. The lithium ions pass through the electrolyte and diaphragm under the action of the applied electric field and are embedded in the lattice of the negative electrode material, which is reduced to lithium atoms to form a compound of carbon inserted into the laminar structure of the negative electrode graphite. The more lithium ions embedded in the negative electrode material, the higher the charging capacity. When discharging, it is the opposite of the charging process.
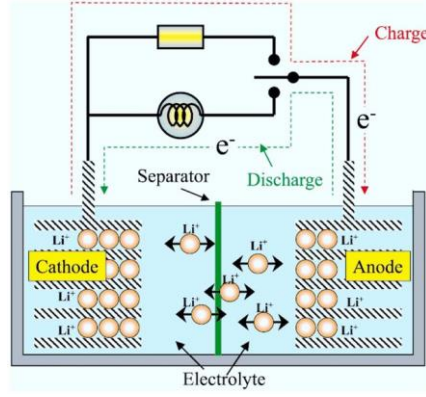
Figure 2.1. Schematic diagram of Lithium-ion battery operation.

During the charging and discharging process of a lithium-ion battery, the de-embedding of lithium ions usually causes the change of crystal layer spacing without destroying crystal structure, and the chemical structure of positive and negative electrode materials is unchanged. The battery's actual capacity is determined by the lithium-ion content of the round-trip de-embedding and the potential difference between the positive and negative electrodes in the de-lithium and embedded lithium. During the first few charge and discharge cycles of the battery, the electrode surface will form a passivation film due to the decomposition of the electrolyte, which is known as a solid electrolyte interphase (SEI) film. Taking the widely used lithium iron phosphate ion battery as an example, the electrode reactions occurring during charging and discharging can be expressed as Equation (2.1-2.2) and the total reaction as Equation (2.3) [7].

$$\text{Positive reaction：} \quad LiFePO_4 \underset{discharge}{\overset{charge}{\rightleftharpoons}} Li_{1-x}FePO_4 + xLi^+ + xe^- \qquad (2.1)$$

$$\text{Negative reaction：} \quad xLi^+ + xe^- + 6C \underset{discharge}{\overset{charge}{\rightleftharpoons}} Li_xC_6 \qquad (2.2)$$

$$\text{Total reaction：} \quad LiFePO_4 + 6C \underset{discharge}{\overset{charge}{\rightleftharpoons}} Li_{1-x}FePO_4 + LixC_6 \qquad (2.3)$$
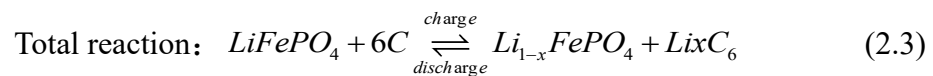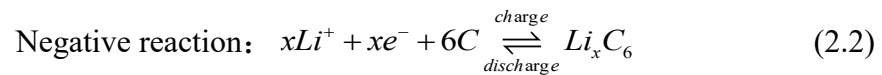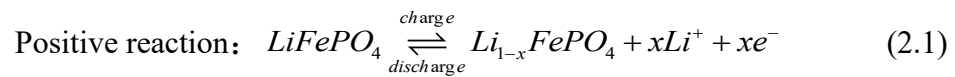
2.1.2 Lithium-ion battery characteristics

Due to the differences in the positive and negative electrode materials, electrolyte materials, and production processes of lithium-ion batteries, the batteries exhibit different performances, resulting in a wide variety of lithium-ion batteries. The different of battery shape structure can be divided into square batteries, cylindrical batteries, and soft pack batteries. The electrolyte material used can be divided into lithium-ion batteries and lithium polymer batteries. According to the cathode material, they can be divided into Lithium Cobalt Battery (LCO), Lithium Manganese Oxide Battery (LMO), Ternary Material Battery (NCM), and Lithium Iron Phosphate Batteries (LFP). A brief introduction and comparative analysis of the above four batteries of different materials as follow.

The cathode material of LCO battery is lithium cobalt oxide, which is generally available in two types: layered structure and spinel structure, which is unstable and has poor cycling performance. The preparation methods of lithium cobalt oxide include the solid-phase method, spray drying method, sol-gel method, etc. The advantage of using lithium cobalt oxide as a cathode material is that it is relatively easy to prepare and has a stable structure, but its cost is relatively high due to the scarcity of cobalt resources.

The cathode material of LMO battery is lithium manganese oxide, which is more price advantageous, and the manganese has non-toxicity and less pollution. Manganese oxides mainly have tunnel structure and layer structure, and tunnel structure manganese oxides include MnO_2 and its derivatives, which are mainly used in lithium primary batteries, and the cycle performance of the battery is not ideal when used as lithium-ion battery cathode material. While the lithium manganese oxide with layered structure mainly includes orthogonal $LiMnO_2$, $LiMnO_3$, and their lithiated derivative. Among them, the preparation methods of orthogonal $LiMnO_2$ include the sol-gel method,

hydrothermal method, solid-phase reaction method, and ion-exchange method. The de-lithium capacity of this class of lithium manganese oxide is high in the voltage range of 3.0~4.5V, but the structure is not stable enough, resulting in poor cycling performance.

Ternary material batteries, modified doping of orthogonal $LiMnO_2$ can effectively improve its cycling ability, such as doping with Al, Cr, Co, Ni, and other elements. Among them, doping with Co and Ni can obtain ternary material electrodes. The preparation methods of ternary materials mainly include the sol-gel method, solid-phase reaction method, simple combustion method, co-precipitation method, and spray pyrolysis method. The electrochemical properties of the ternary materials obtained by different preparation methods will also be different, but the safety performance of this type of battery needs to be further improved.

The cathode material of LFP battery is lithium iron phosphate, which does not contain cobalt, nickel, and other precious metal elements, the raw material price is relatively low, and phosphorus, iron, lithium, and other elements are abundant in the earth's crust. In addition, LFP stack density, multiplier performance, low-temperature performance, and other issues have been gradually resolved in recent years through technical means, significantly broadening the application of lithium-ion batteries. Currently, the commercialized lithium iron phosphate batteries have excellent safety performance, long cycle life, and good stability in high temperature and high heat environments, which has led to the widespread application of lithium iron phosphate batteries in energy storage and electric vehicles [8, 9].

A comparison of the performance of common lithium-ion batteries is shown in Table 2.1 [10-12].

Table 2.1. Comparison of the performance of common types of lithium-ion batteries.

Battery	LCO	LMO	NCM	LFP
Chemical formula	$LiCoO_2$	$LiMnO_4$	$LiNi_xCo_yMn_{(1-x-y)}O_2$	$LiFePO_4$
Operating voltage(V)	3.6	3.7	3.7	3.3
Crystal structure	layered rock salt	spinel	layered rock salt	olivine
Energy density (Wh/kg)	150~200	100~150	150~220	100~130
Cycle life (cycle)	500~1000	300~700	1000~2000	>2000
Thermal runaway temperature (℃)	180	265	210	310
Security features	normal	good	better	best

2.2 Lithium-ion battery aging mechanism

Battery failure and gradual performance degradation result from complex, interrelated phenomena that depend on battery chemistry, the environment, and the actual operating conditions [13]. The aging mechanism and external characteristics of lithium-ion batteries are shown in Figure 2.2. Due to the numerous external factors affecting battery aging, such as environmental and mechanical stress, and the complex electrochemical system inside the battery involving multiple physical fields such as electric-thermal-mechanical, the performance of the battery decays continuously during long-term use and storage. Generally, each aging mechanism results from the mutual interaction of two or more external factors. For example, the presence of SEI film plays an inhibitory role in the formation of new SEI films. In contrast, the cracks formed after the decomposition or breakage of the old SEI film will further promote the generation of new SEI films. In addition, the exfoliation of graphite leads to the reduction of contact area between collector and electrode materials and the increase of contact impedance, resulting in the decay of cell power characteristics. In summary, the complex aging mechanism manifests as an aging pattern of active lithium loss and loss of positive and negative active materials, which is a fundamental cause of battery power

degradation and capacity decay [14, 15].
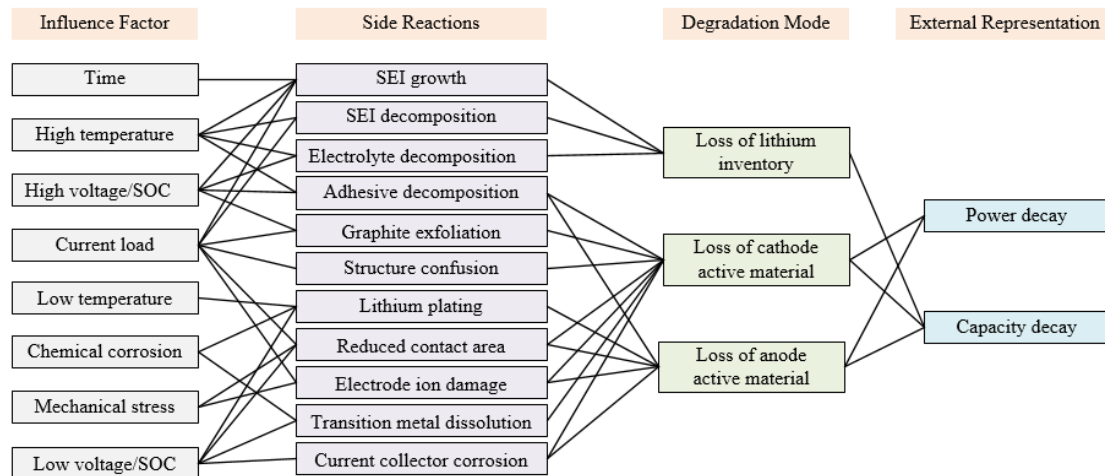
Figure 2.2. The aging mechanism of lithium-ion batteries and its effect on external characteristics.

2.2.1 Battery aging type

According to the different application conditions of aging, the battery aging type can be divided into calendar aging and cycle aging [16].

(1) Calendar aging

Calendar aging refers to the self-discharge reaction of the battery during the storage process, resulting in the gradual decay of its performance index. When the lithium-ion battery works normally, an SEI film is generated on the contact surface of the electrode and electrolyte, whose main component is the product of electrolyte reduction at the negative electrode. This film has a specific protective effect on the stability of the battery's internal structure, and its thickness, composition, and structure deteriorate during the aging process of the battery, becoming the leading cause of battery power decay and capacity decline [14]. Existing studies show that the calendar aging of lithium-ion batteries is related to the growth of the thickness of the SEI film and the gradual decrease of the number of lithium-ions in the cycle.

Considering the calendar aging of Lithium-ion batteries is related to the state of

charge (SOC) and storage temperature of Lithium-ion batteries, and the higher the state of charge and storage temperature, the faster the capacity of Lithium-ion batteries decreases. To effectively slow down the calendar aging of Lithium-ion batteries, it is usually necessary to turn down the charge state of Lithium-ion batteries and store them at a lower temperature when storing them.

(2) Cycle aging

Unlike calendar aging, lithium-ion batteries are accompanied by irreversible capacity loss during cycling, and their internal phenomena, such as performance decay and component aging are called cyclic aging [17, 18]. Battery cyclic aging is mainly caused by fatigue fracture caused by lattice strain cycling of the active material. For LFP particles, the strain field distributed along the phase interface will lead to lattice dislocation or fracture, and the increase of defects will significantly reduce the diffusion rate of lithium ions, leading lead to the degradation of battery performance. The lattice fracture will also lead to the disassociation of the active material from the conductive agent and the inability to obtain the electrons required for the electrochemical reaction, ultimately leading to the loss of available capacity. For negative graphite electrodes, the lattice sheets are maintained by weak van der Waals forces between the layers, and lithium-ion de-embedding tends to cause distortion and destruction of the lattice, which continuously drives electrolyte decomposition and deposition on the newly exposed graphite surface to form SEI films, eventually causing battery aging. Many factors affect cyclic aging, mainly the ambient temperature and charge/discharge rate. The higher the temperature, the greater the charge/discharge rate will be the capacity decay rate.

2.2.2 Battery aging external factors

The external factors affecting battery aging mainly include temperature, charge/discharge rate, depth of discharge (DOD) and SOC. This section analyzes the chemical reaction mechanism of battery decay under the influence of a single external factor's influence [12].

(1) Temperature

The ambient temperature plays a crucial role in the internal chemical reaction rate of the battery. When the temperature rises, the impact on battery aging occurs mainly at the positive electrode, with less impact on the negative electrode. While in a low-temperature environment, lithium ions are neither easy to complete the deem bedding. In contrast, the liquid material will lead to increased viscosity or even solidification under low-temperature conditions, which eventually makes crossing the diaphragm difficult. In addition, the mass transfer process of charged ions and the speed of electrochemical reactions is also reduced, which ultimately leads to the reduction of lithium metal ions, lithium deposition, and the crystal structure of the active material is easily damaged. Therefore, in low-temperature conditions, resulting in increased resistance of lithium-ion batteries, the actual capacity will rapidly decline, ultimately leading to rapid battery aging.

(2) Charging and discharging rate

The charge/discharge rate is another major factor that affects battery aging. When the battery exceeds the rated charge and discharge rate, the thermal effect of the current increases the temperature of the battery itself, accelerates the electrochemical reaction rate inside the battery, and gradually intensifies the side reactions of aging. A high current will further aggravate the polarization phenomenon inside the battery. In addition, a high current brings a large number of lithium ions embedded in the cathode

material, causing an impact on the stability of the material. When discharging at a high current, there is also the stability problem of heating and de-embedding of cathode material. At this time, too many lithium-ions are transported to the negative electrode, making the deposition of lithium monomers occur, leading to loss of capacity. The high current destroys the structural stability of the battery electrode material and causes the cathode material to react with the electrolyte. This further causes corrosion of the cathode and the deposition of lithium salts on the cathode surface, resulting in the thickening of the SEI film and loss of active material, which in turn increases the internal resistance of the lithium battery and decreases the actual available capacity.

(3) DOD

DOD refers to the ratio of the capacity released during the discharge process to the battery's nominal capacity, and the value is negatively correlated with the number of charge/discharge cycles. When a lithium-ion battery is discharged deeply for a long time, a large number of active lithium ions and electrolytes inside the battery are consumed, which reduces the ionic conductivity and leads to an increase in electrolyte impedance and electrode interface charge transfer impedance. The larger the DOD value, the fewer remaining cycles. Therefore, the battery should be charged in time to prevent deep discharge and accelerate battery aging.

(4) SOC

SOC is one of the main causes of battery aging, and the mean SOC within the cycle interval is positively correlated with the battery capacity decay rate. In addition, the greater the difference between the upper and lower limits of the cycle interval, the faster the capacity decay. From this point of view, the full charge and discharge of the battery will accelerate the aging of the battery, so when using it, shallow charge and discharge will help delay the aging of the battery. If the battery needs to be stored for a

long time, try to avoid the battery in a fully charged state or high SOC.

2.2.3 Internal mechanism of battery aging

As the number of charges and discharge cycles of Lithium-ion batteries continues to increase, their performance and lifetime gradually decay. Common causes of lithium battery aging include structural changes of active material, the phase change of active material, cracking or breaking of active particles, transition metal leaching, volume expansion, solid electrolyte interface overgrowth, SEI decomposition, etc. [19]. As its internal aging mechanism is very complex as shown in Figure 2.3 [4, 14], the main causes of lithium-ion battery aging can be divided into three categories: 1) loss of lithium inventory (LLI) [20]; 2) loss of positive active material (LAM$_p$) [14]; and 3) loss of negative active material (LAM$_n$) [21, 22]. Along with the capacity decay of Lithium-ion batteries, there is often an increase in the internal resistance of the battery and the consumption of electrolytes.
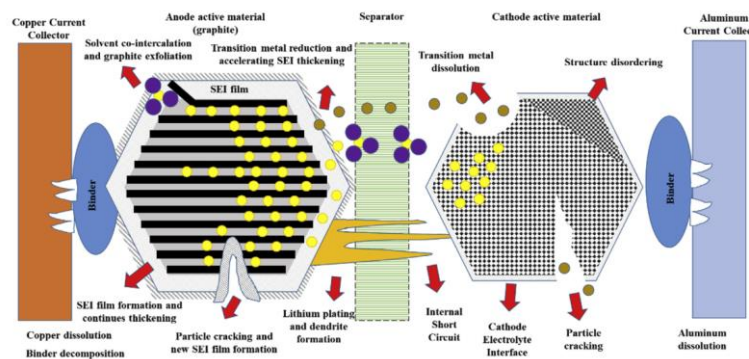
Figure 2.3. Schematic diagram of the battery aging mechanism.

(1) LLI

The active lithium ions inside the battery are continuously detached and embedded between the positive and negative electrodes to complete the charging and discharging of the lithium-ion battery. As the amount of cathode material is closely related to the active lithium inside the battery, when the limited amount of lithium ions in the battery

is decreasing, the battery capacity is also reduced accordingly. During the initial cycling phase, the growth rate of SEI film is relatively fast, and then the growth rate is relatively flat. The stability of the SEI film has an important impact on the amount of active lithium inside the cell during the long-term continuous cycling process, accompanied by irreversible reactions and loss of active lithium. Existing studies show that the loss of reactive lithium becomes the main cause of battery capacity degradation [23-25].

(2) LAM_p

The cathode material is one of the critical factors in determining the lifetime of lithium batteries. The dissolution of transition metals in the anode will not only lead to the loss of anode material, but the dissolved transition metals will also pass through the diaphragm and precipitate on the surface of the anode, accelerating the formation of the negative SEI film. For LFP batteries, the graphite system cells have faster capacity decay in high-temperature cycles, mainly due to the dissolution of Fe^{3+} in the anode and deposition on the anode surface, which further undergoes reduction and generates Fe metal particles, leading to increased anode polarization [26, 27]. The volume expansion of the LFP material during the charging and discharging process is 6.77%, so the LFP material exhibits a very excellent cycle life. However, LFP is also accompanied by partial dissolution of Fe elements during cycling and precipitation on the anode surface, catalytic electrolyte decomposition causing an increase in cell impedance and loss of active Li, which eventually leads to capacity decay and power degradation.

(3) LAM_n

The reactions that occur during the aging and decay of the negative electrode material are mainly the rupture/regeneration of SEI film and solvent molecule co-embedding. In the pre-cycle after cell synthesis, the SEI film is generated on the surface of the active particles. It leads to an increase in the internal resistance of the negative

electrode. The rupture and regeneration of SEI film lead to the loss of active material and the consumption of available lithium ions and electrolytes, increasing the battery's internal resistance. Under low-temperature charging or high-rate charging, lithium metal is easily precipitated from the cathode surface. The precipitated lithium metal is very active and reacts with the electrolyte, causing the loss of available lithium ions and increased internal resistance. Similar to the positive electrode, the negative electrode collector and binder will also decompose and corrode during use, eventually leading to capacity degradation and power degradation.

The above three internal environmental factors interact and influence, resulting in the decline of the available capacity of lithium-ion batteries. Of course, the internal environment of lithium-ion batteries is not closed and will also be affected by the external environment. An unsuitable external climate will exacerbate the changes in the internal environment of the lithium-ion battery, which will accelerate the decline in available capacity. In order o slow down the decline of the available capacity of Lithium-ion batteries, effective battery management strategies are needed to reduce the external environment's influence on the battery's internal environment [28].

2.2.4 External characterization of battery aging

For most power batteries, the capacity decay shows non-linearity, as shown in Figure 2.4. Usually, the capacity decay characteristics are divided into three stages during the whole life cycle of the battery. In the first stage, mainly because the SEI film formed at the beginning is unstable, the electrolyte will continue to decompose on the surface of the negative electrode, so the consumption of active lithium and the growth of SEI film become the main factors of rapid capacity decay in this stage. In the second stage, the thickness of SEI film increases significantly, the decomposition rate of

electrolyte on the negative electrode surface decreases greatly, and the decay rate of battery capacity decreases significantly, while some other side reactions eventually lead to a flat trend of the capacity decay curve. In the third stage, the pores inside the cathode at the end of life are filled with a large amount of SEI film, causing the precipitation of metallic lithium on the cathode surface, which accelerates the loss of active lithium in the battery and leads to the rapid decay of battery capacity [14].
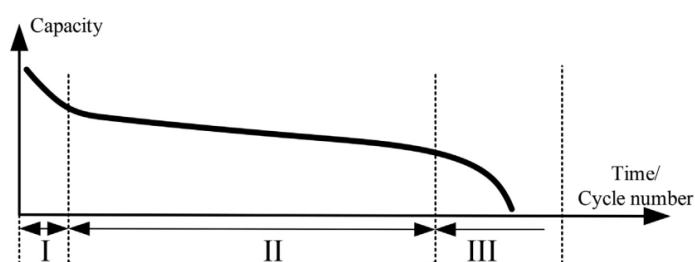
Figure 2.4. Capacity decay characteristic curve of Lithium-ion battery.

2.3 Battery aging analysis method

The study of the aging mechanism of lithium-ion batteries involves various fields, such as electrode material science, electrochemistry, health management, etc. The research methods are mainly divided into three types: in situ online analysis method, disintegration physical and chemical analysis method, and external characterization analysis method [29-31].

2.3.1 In-situ online analysis method

The in-situ online analysis method uses in-situ analysis equipment to monitor the changes in internal physical quantities of the battery online during the cycling process. It analyzes battery aging by studying the evolution of the internal physical quantities of the battery during the aging process. In-situ testing techniques mainly include in-situ X-ray diffraction analysis (XRD) technology, neutron wire technology, preparation of transparent model cells to observe online the aging mechanism, and evolution of the

battery during charging and discharging with the help of an optical microscope, scanning electron microscopy (SEM), or Transmission Electron Microscopy (TEM), etc.

2.3.2 Disassembly physical and chemical analysis method

The disassembled physical-chemical analysis method disassembles the aged battery sample under a specific environment to obtain the battery's internal components and obtain the battery's aging characteristics by utilizing material analysis or reassembling the buckling. This analysis method mainly includes SEM, energy-dispersive X-ray spectroscopy (EDS), etc. It is used to investigate the changes of elements or structures in the active material, diaphragm, and other battery components before and after aging or reassemble the positive and negative electrode layers into a buckle half-cell to investigate the aging mechanism of the battery.

2.3.3 External characteristics analysis method

The external characteristics analysis method extracts the aging characteristics of the battery by comparing the changes of exterior features, such as charge/discharge curves and impedance spectra, before and after aging and doing appropriate processing on these external characteristics. This analysis method mainly includes incremental capacity analysis (ICA), differential voltage (DVA), discharge curve fitting, and electrochemical impedance spectrum methods. The existing methods to analyze the aging mechanism of batteries can be divided into three categories: the external characterization method, the disassembly analysis method, and the in-situ online analysis method. The above three battery aging analysis methods and their corresponding advantages and disadvantages are shown in Table 2.2 [29].

Table 2.2. Comparison of advantages and disadvantages of three types of battery aging mechanism analysis methods.

Analysis Method	Main Methods	Advantages	Disadvantages
In-situ testing techniques	In-situ XRD, neutron line techniques, model cells	Non-destructive, real-time observation, clear and intuitive for study subjects	Precise and expensive experimental equipment, scarce experimental time, difficult equipment construction and sample preparation
Disintegration materialization	Deductive power, SEM. TEM, EDS	Direct observation of internal changes in the battery and verification of the cause of aging	Damage to the object of study, the ability to investigate only a specific state of aging, the need for a reference sample
External Characteristics	ICA、DVA、EIS、R-Q	Non-destructive to the object of study, and can investigate the evolution of the aging process	Based on hypothetical speculation, a certain combination of disintegration materialization analysis is required

Among the above three battery aging analysis methods, the in situ observation technique requires high experimental equipment and sample requirements. It is difficult to guarantee the testing time, so it is less applied in battery aging analysis. Although the disintegration physical analysis method can directly observe the internal changes of the battery and verify the cause of aging, it will cause damage to the research object and is often used in the laboratory environment. The external characterization method has a wide range of application scenarios because it can achieve damage-free analysis of the battery. Therefore, researchers in different fields can use appropriate schemes to analyze the battery aging mechanism according to the actual needs.

2.4 Battery health indicators

Health indicators (HIs) extracted and selected are essential data pre-processing steps for data-driven estimation and significantly impact estimation performance. This chapter divides the battery HIs into performing direct measurements and indirect calculations [32].

2.4.1 Measurement-based HIs

The variables that BMS can measure include current, voltage, temperature, and time. Therefore, a simple method is to extract the HIs from these measured variables, which researchers have widely studied. The measured variable-based HIs can be divided into four categories: voltage curve-based, time-based, temperature-based, and others. Existing HI extraction methods are based on measured variables, as shown in Table 2.3.

Table 2.3. Existing HIs extraction methods are based on measured variables.

Variable Type	Health indicators	Operating conditions
Voltage curve-based	voltage variation during the equal time (VVET)	Charge and discharge
	the slope of the voltage curve (SV)	Charge and discharge
	initial voltage (IV)	Charge and discharge
	cut-off points (COP)	Charge and discharge
	voltage distance (VD)	Charge and discharge
Time-based	the time interval during equal voltage increase (TEVI)	Charge
	the time interval of equal voltage decrease (TEVD)	Discharge
	the time interval during the equal current decrease (TECD)	Charge
	time of constant current charge (TCCC)	Charge
	time of the constant voltage charge (TCVC)	Charge
	the time interval of equal temperature increase (TETI)	Charge
Temperature-based	maximum temperature (MAT)	Charge and discharge
	location of maximum temperature(LMAT)	Charge and discharge
	mean temperature (MET)	Charge and discharge
	temperature increase during the equal time (TIET)	Charge and discharge
Others	cycle number (CM)	Charge and discharge
	the current decrease during equal charge time	Charge
	the ratio of CC time to CV time (RCV)	Charge

2.4.2 Computation-based HIs

The health information reflected by measured variables is limited. The calculated variables are used to extract HIs that reflect more details about the SOH. This method

transforms the measured variables, and then the HIs are extracted from the transformed curves. The mainstream transformations include incremental capacity (IC) analysis, differential voltage (DV) analysis, and differential temperature (DT) analysis [33, 34]. Existing HIs extraction methods are based on calculated variables, as shown in Table 2.4.

Table 2.4. Existing HIs extraction methods are based on calculated variables.

Variable Type	Health indicators	Operating conditions
IC-based	incremental capacity peak value (ICP)	Charge / Discharge
	incremental capacity valley value (ICV)	Charge / Discharge
	incremental capacity peak location (ICPL)	Charge / Discharge
	incremental capacity valley location (ICVL)	Charge / Discharge
	incremental capacity peak area (ICPA)	Charge / Discharge
	distance between two peaks (DBTP)	Charge / Discharge
DV-based	differential voltage peak (DVP)	Charge / Discharge
	differential voltage valley (DVV)	Charge / Discharge
	differential voltage peak location (DVPL)	Charge / Discharge
	differential voltage valley location (DVVL)	Charge / Discharge
	distance between two valleys (DBTV)	Charge / Discharge
DT-based	differential temperature peak (DTP)	Charge / Discharge
	differential temperature peak location (DTPL)	Charge / Discharge
Others	sample entropy (SE)	Discharge
	parameters of the equivalent circuit model (PECM)	Charge
	frequency analysis (FA)	Charge and discharge
	mechanical parameter (MP)	Charge and discharge
	curve fitting coefficients (CFC)	Charge

2.5 Battery aging experiment

The analysis of the whole life cycle external characteristics of the power battery and the study of the aging model is inseparable from the support of the battery life cycle data. Accurate and systematic aging experimental data simulating the realistic environment is essential in the battery lifetime evaluation system. Therefore, the design

of a reliable battery cycle aging test and calendar aging test, and highly accurate equipment to establish the whole life cycle database are the prerequisites for battery aging model research. As shown in Figure 2.5, the experimental platform is set up for the battery aging test. It includes a battery tester, a data logger, a thermal chamber, an electrochemical workstation, a computer, and some batteries. The platform is used to verify the effects of different temperatures, mean SOC, DOD, charge-discharge rate, storage SOC, and storage time on battery aging.
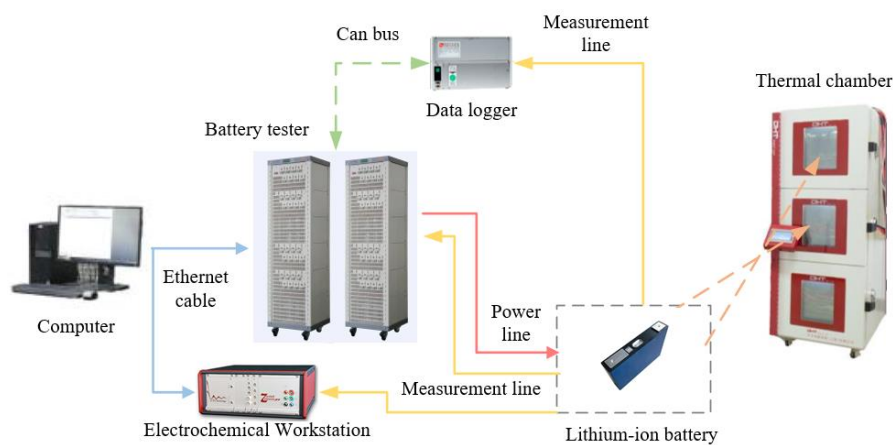
Figure 2.5. The platform for the battery aging experiment.

2.5.1 Lithium-ion battery cycle lifetime experimental research

(1) Test purpose

In this test, to investigate the factors influencing the cyclic aging rate of NCM622 lithium battery, a test platform for a single-cell battery test is built, and a cyclic aging test scheme considering different temperatures, mean SOC in cyclic intervals, DOD and charging rate is designed. This test platform provides the basis for studying the battery cyclic aging model.

(2) Experimental steps

The cyclic aging test is mainly considered in four dimensions: temperature, charging rate, DOD, and mean SOC of the cyclic interval. The cyclic aging test scheme

is shown in Table 2.5, with a discharge rate of 1C and no resting between charging and discharging.

Table 2.5. A cyclic aging test program.

NO.	Test conditions			Test termination conditions	sample size
	SOC interval /%	Charge rate/C	temperature/℃		
1	0~30	1	25	the capacity decay reaches 20%	2
2	45~55				2
3	70~100				2
4	35~65				2
5	10~90				2
6	0~100				2
7	10~90	1.2			2
8	10~90	1.5			2
9	0~30	1	45		2
10	45~55				2
11	70~100				2
12	35~65				2
13	10~90				2
14	0~100				2
Note: C=50Ah, upper/lower limit voltage: 4.2V/2.75V, the battery has a cycle life of 1500 at 1C 100% DOD at 25°C.					

A cycle aging test is conducted monthly to retest the battery characteristics data.

(a) Actual battery capacity: discharge with C/3 constant current until the lower cut-off voltage reaches 2.75V, take C/3 constant current charge to 4.2V, charge with constant voltage until the current is less than C/20, then discharge with C/3 constant current until the lower cut-off voltage is 2.75V, repeat two times.

(b) Low current discharge curve: charge with C/3 current to the cut-off voltage, leave it for 1h, and discharge with C/20 current to 2.75V.

(c) Hybrid Pulse Power Characterization (HPPC) test for all cells at different SOC levels at 25℃ ambient temperature: The cells were fully charged and gradually

discharged to 80%, 60%, 50%, 40%, 20% SOC state for HPPC test and 80%, 50%, 20% SOC state for EIS test. Before the test is set aside for 2h, the frequency range of EIS is selected as 100 mHz~5kHz, constant potential mode, and the amplitude is 2mV. HPPC test is discharged with 2C current for 18s, set aside for 40s, then charged with 1C for 10s. then set aside for 40s. then charged with 2C current for 18s, set aside for 40s, discharged with 1C current for 10s, and finally set aside for 40s.

(3) Calculation of experimental results and analytical discussion
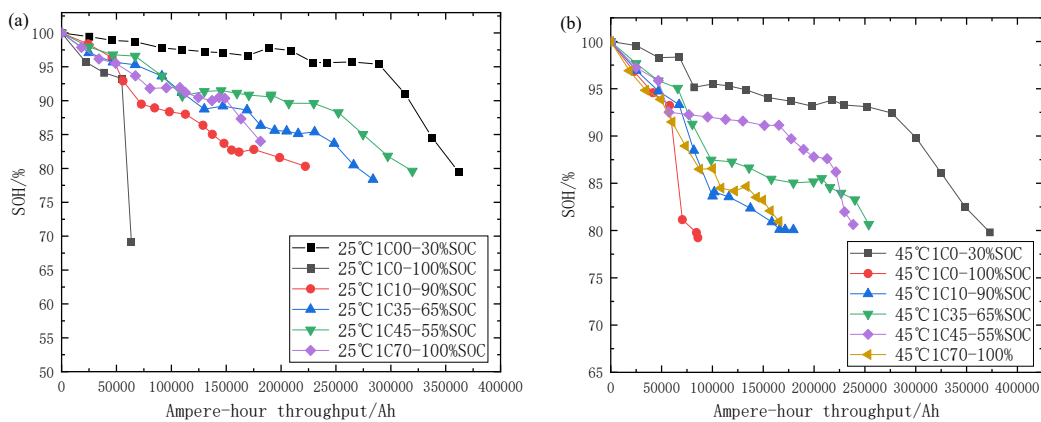
Figure 2.6. Aging curves for different cycle intervals and temperatures.

(a) 25℃ under different DOD; (b) 45℃ under different DOD.

As shown in Figure 2.6, at different temperatures, full charge and discharge for the 0% to 100% SOC interval intensifies the battery aging process, reaching the 80% life cutoff at approximately 650 cycles. 0% to 30% SOC interval cycles have the slowest aging rate, indicating that the smaller the mean SOC of the cycle interval, the slower the battery decline.
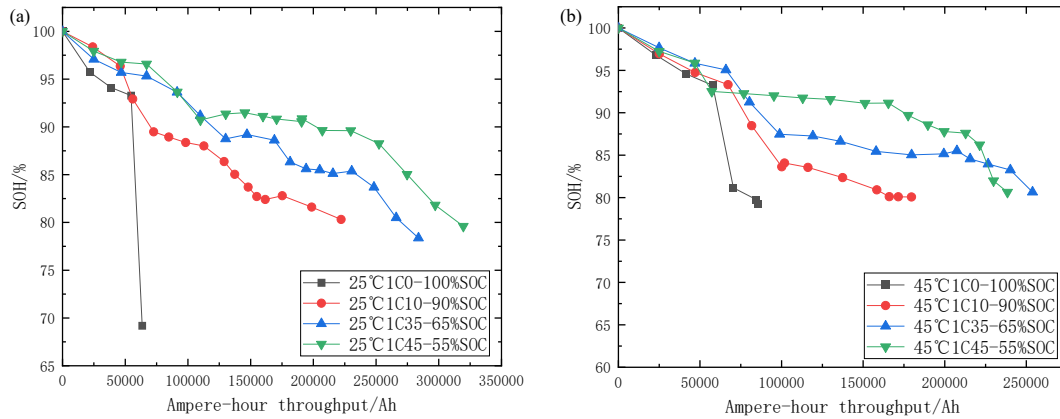
Figure 2.7. Aging curves for same mean SOC and different temperatures.

(a) 25℃ under same mean SOC; (b) 45℃ under same mean SOC.

As shown in Figure 2.7, the larger the DOD, the faster the battery aging rate under the same temperature conditions, mean SOC of the cycling interval, and charging rate. Comparing the 0%~100% SOC and 10%~90% SOC cycling intervals, full charge and discharge will significantly accelerate the aging process, and cycling between 10%~90% SOC will slow the battery aging.
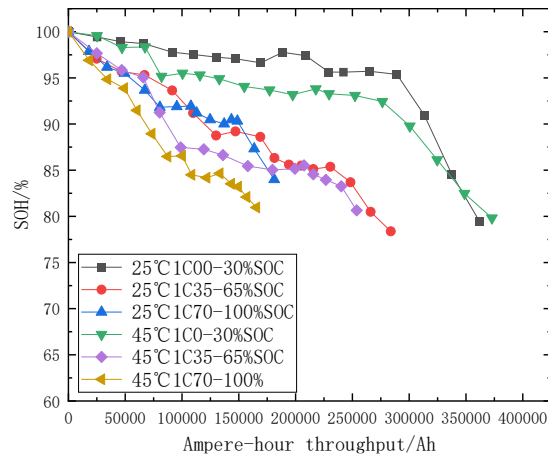
Figure 2.8. Aging curves for different temperatures at 30% DOD.

As shown in Figure 2.8, the slowest cyclic aging rate was observed between 0% to 30% SOC at 30% DOD and 25°C. At 30% DOD and 45°C, the cyclic aging rate is the fastest between 70%~100% SOC, indicating that the smaller the mean SOC of the cyclic interval, the slower the battery decay. Meanwhile, the decay rate will be accelerated with the increase in temperature.
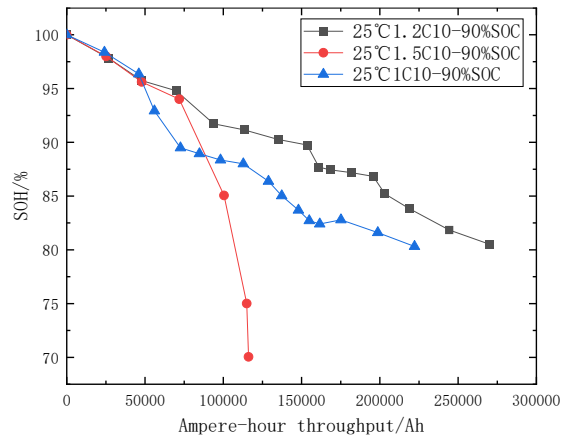
Figure 2.9. Aging curves for different charging rates at 25°C.

Figure 2.9 shows that the aging rate is the fastest at 1.5C charging rate, while the battery aging rate is the slowest at 1.2C charging rate, indicating that this type of battery is suitable for charging between 1C and 1.5C charging rate.
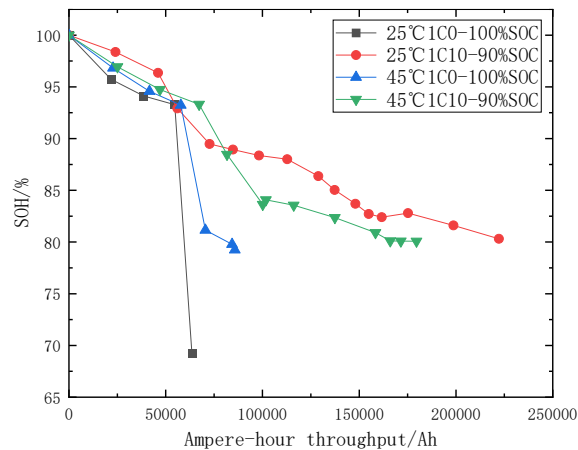
Figure 2.10. Aging curves for same mean SOC at different temperatures.

As seen in Figure 2.10, the aging rate is exacerbated by full charge and discharge at different temperatures. The battery decay rate is faster at 0% to 100% SOC at 25°C than at 10% to 90% SOC at 45°C, indicating that large interval cycling has a more significant effect on the battery aging rate than temperature.

2.5.2 Lithium-ion battery calendar lifetime experimental research

(1) Test purpose

A lithium battery test platform was built for this test to explore the factors that affect the calendar aging rate of NCM622 lithium batteries, and a calendar aging test scheme was designed considering different storage temperatures, storage SOCs, and storage times. The test platform establishes the calendar aging data of the NCM622 lithium battery for one year, which provides a basis for studying the calendar aging model.

(2) Test steps

The calendar aging test mainly considers temperature, storage SOC, and storage time. After the battery is discharged at C/3 to the cut-off voltage, it is charged at C/3 to the corresponding SOC. The calendar aging test is shown in Table 2.6.

Table 2.6. Storage aging test protocol.

NO.	SOC status/%	temperature/℃	Test termination conditions	sample size
1	100			2
2	95			2
3	90	25		2
4	70			2
5	50			2
6	20		Storage for 12 months	2
7	100			2
8	95			2
9	90	45		2
10	70			2
11	50			2
12	20			2
Note C=50Ah, upper/lower voltage limit: 4.2V /2.75V.				

The battery characteristic data is re-tested once a month in the calendar aging test.

(a) Actual battery capacity: discharge at C/3 constant current until the lower cut-off voltage is 2.75V. Take C/3 constant current charge to 4.2V and constant voltage charge until the current is less than C/20. Then C/3 constant current is discharged until the lower cut-off voltage of 2.75V. Repeat the above operation two times.

(b) Small current discharge curve: charge with C/3 current to cut-off voltage, let

stand for 1h, and discharge with C/20 current to 2.75V.

(c) Perform HPPC tests on all cells at different SOC levels at an ambient temperature of 25°C: gradually discharge the batteries to 80%, 60%, 50%, 40%, and 20% SOC after fully charged, and perform HPPC tests at SOC. EIS test is performed under 80%, 50%, and 20% SOC states. Before the test, it was put on hold for 2h, the frequency range of EIS was selected as 100 mHz~5kHz, the potentiostatic mode was selected, and the amplitude was 2mV. During the HPPC test, first discharge with 2C current for 18s, put it on hold for 40s, then charge it with 1C for 10s, and then put it on hold for 40s. Charge with 2C for 18s, put on hold for 40s, then discharge with 1C for 10s, and finally put on hold for 40s.

(3) Calculation and analysis of experimental results

As can be seen from Figure 2.11, at different temperatures, the aging rate of the battery stored at 20% SOC is the slowest, followed by storage at 50% SOC, and storage at greater than 50% SOC will accelerate battery degradation, but the difference is not It is not obvious, so such batteries are suitable for storage under low SOC conditions.
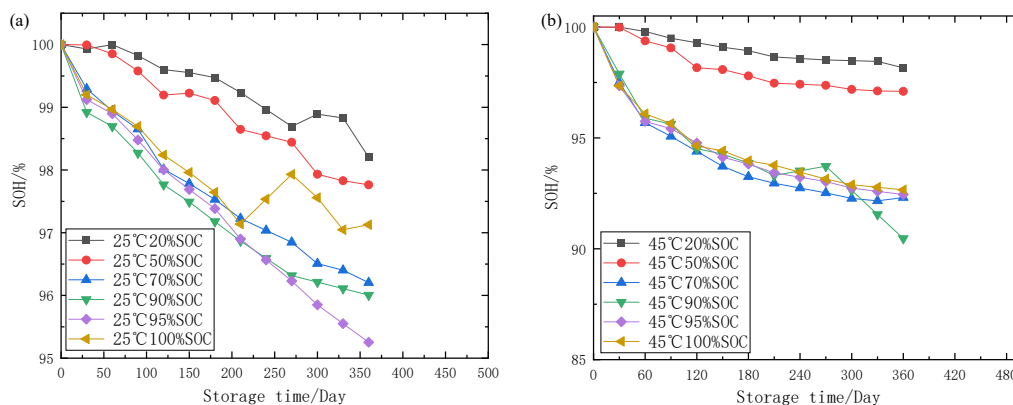
Figure 2.11. Battery aging curves are stored at different SOC and temperature.

(a) 25℃ under different SOC intervals; (b) 45℃ under different SOC intervals.

As shown in Figure 2.12, the aging rate of the battery stored at 45°C is significantly greater than that stored at 25°C. This phenomenon is more apparent   when stored at

61

high SOC, indicating that such batteries are suitable for storage at room temperature.
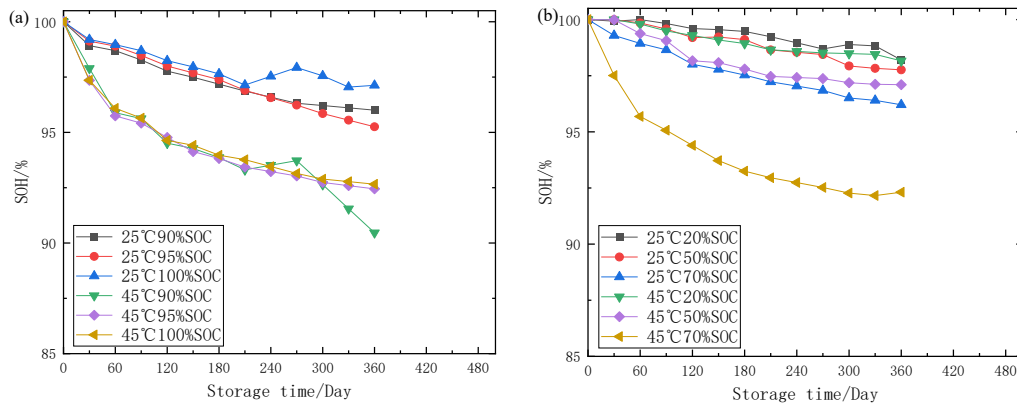
Figure 2.12. Battery aging curves at different temperatures.

(a) aging curves under high SOC; (b) aging curves under normal SOC.

2.6 Chapter summary

This section mainly analyzes and verifies the aging mechanism of lithium-ion batteries. Firstly, the working principle of the Lithium-ion battery is introduced, and the characteristics of four common Lithium-ion batteries are analyzed and compared. Secondly, the internal and external factors affecting the aging of Lithium-ion batteries are analyzed in detail and the battery aging mechanism analysis method is introduced, while the health indicators characterizing the battery aging are investigated and studied. Finally, lithium-ion battery cycle aging experiment is designed to verify the effects of temperature, mean SOC of cycle interval, DOD, and charge rate on battery aging, in which temperature, mean SOC of cycle interval, and charge/discharge rate show a positive correlation with battery aging cycle. Meanwhile, a lithium-ion battery calendar aging experiment was designed to verify the effects of temperature, storage SOC, and storage time on battery calendar aging, and the results showed that temperature, storage SOC, and storage time were positively correlated with battery calendar aging.

References

[1] Zhang, Z.J. and P. Ramadass, Lithium-ion battery systems and technology, in *Batteries for Sustainability*. 2013, Springer. p. 319-357.

[2] Scrosati, B. and J. Garche, Lithium batteries: Status, prospects and future. *Journal of power sources*. 2010. **195**(9): p. 2419-2430.

[3] Tarascon, J.-M. and M. Armand, Issues and challenges facing rechargeable lithium batteries, in *Materials for sustainable energy: a collection of peer-reviewed research and review articles from Nature Publishing Group*. 2011, World Scientific. p. 171-179.

[4] Hu, X., et al., Battery lifetime prognostics.*Joule*. 2020. **4**(2): p. 310-346.

[5] Bai, P., et al. Review of theoretical research on aging failure mechanisms and optimal management methods for lithium-ion batteries. *2011 Annual Meeting of the Chinese Society of Automotive Engineering*. 2011. Beijing, China.

[6] Huang, P., et al., The combustion behavior of large scale lithium titanate battery. *Scientific reports*. 2015. **5**(1): p. 1-12.

[7] By R. Xiong, Core algorithms for power battery management systems. 2018: Beijing: Machinery Industry Press.

[8] Liu, Xingtao, Study on the state estimation strategy of power lithium batteries for electric vehicles. 2014, PhD thesis, University of Science and Technology of China.

[9] Liang Guangchuan, Zong Jiyue, Cui Xuxuan, editors, Lithium iron phosphate cathode materials for lithium-ion batteries. 2013: Beijing: Science Press.

[10] Wang YJ, Modeling, state estimation and management strategies of powered lithium batteries. 2017, PhD thesis,University of Science and Technology of China.

[11] Zhang Chi, Research on the aging characteristics and state estimation algorithm of automotive power lithium-ion batteries. 2019, PhD thesis,Wuhan University of Technology.

[12] Chen, Yuan, State estimation and lifetime modeling of lithium-ion batteries. 2021, PhD thesis, Hefei University of Technology.

[13] Palacín, M.R. and A. de Guibert, Why do batteries fail? *Science*.2016. **351**(6273): p. 1253292.

[14] Han, X., et al., A review on the key issues of the lithium ion battery degradation among the whole life cycle. *ETransportation 2019; 1: 100005*.

[15] Tian Jiaqiang, Health assessment and fault diagnosis of energy storage lithium battery system. 2021, PhD thesis, University of Science and Technology of China.

[16] Agubra, V. and J.J.M. Fergus, *Lithium ion battery anode aging mechanisms.* 2013. **6**(4): p. 1310-1325.

[17] Kang, Yongzhe, Research on capacity estimation and fault diagnosis method for lithium-ion battery pack. 2021, PhD thesis, Shandong University.

[18] Berecibar, M., et al., Critical review of state of health estimation methods of Li-ion batteries for real applications. *Renewable and Sustainable Energy Reviews.* 2016. **56**: p. 572-587.

[19] Zhao Yunwu, Study on aging analysis and side reaction modeling method for lithium-ion batteries. 2019, Master Thesis, Harbin Institute of Technology.

[20] Zheng Huiyuan, Research on capacity degradation mechanism and suppression method of lithium-ion batteries. 2017, PhD thesis, Suzhou University.

[21] Vetter, J., et al., Ageing mechanisms in lithium-ion batteries.*Journal of power sources*. 2005. **147**(1-2): p. 269-281.

[22] Ren, Dongsheng, et al., Progress of research on the whole life cycle safety evolution of lithium-ion batteries. *Energy Storage Science and Technology*. 2018. 7(06): p. 957-966.

[23] Dubarry, M., C. Truchot, and B.Y. Liaw, Cell degradation in commercial LiFePO4 cells with high-power and high-energy designs.*Journal of Power Sources.*2014. **258**:

p. 408-419.

[24] Dubarry, M., C. Truchot, and B.Y. Liaw, Synthesize battery degradation modes via a diagnostic and prognostic model. *Journal of Power Sources.* 2012. **219**: p. 204-216.

[25] Dubarry, M. and B.Y. Liaw, Identify capacity fading mechanism in a commercial LiFePO4 cell. *Journal of Power Sources.* 2009. **194**(1): p. 541-549.

[26] Yang Peng, Study of capacity decay of lithium-ion batteries. 2013, Master Thesis, Shanghai Jiao Tong University.

[27] Amine, K., J. Liu, and I. Belharouak, High-temperature storage and cycling of C-LiFePO4/graphite Li-ion cells. *Electrochemistry communications.* 2005. **7**(7): p. 669-673.

[28] Yu B., Study on the degradation modeling and remaining life prediction method for lithium-ion batteries. 2020, PhD thesis, National University of Defense Technology.

[29] Liu, Z.X., Study on aging distribution and state estimation of lithium-ion batteries. 2019, PhD thesis,Tsinghua University.

[30] Song, Y., et al., Advances in lithium-ion battery aging research. *Power Technology.* 2018. 42(10): p. 1578-1581.

[31] Yuan-Yuan Li, Research on Lithium-ion Battery Health Management Based on Data-Driven Modeling. 2021, PhD thesis, University of Electronic Science and Technology.

[32] Hu, X., et al., Battery health prediction using fusion-based feature selection and machine learning. *IEEE Transactions on Transportation Electrification.* 2020. **7**(2): p. 382-398.

[33] Li, Y., et al., Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review. *Renewable and sustainable energy reviews.* 2019. **113**: p. 109254.

[34] Xiong, R., L. Li, and J. Tian, Towards a smarter battery management system: A

critical review on battery state of health monitoring methods. *Journal of Power Sources*. 2018. **405**: p. 18-29.

Chapter 3. State of Health Estimation based on modified Gaussian Process Regression for Lithium-ion Batteries

3.1 Introduction of this chapter

Lithium-ion batteries (LIBs) have been commercially produced by Sony for the first time for 31 years, which have been widely used in consumer electronics, electric vehicles, and micro-rids due to their low and falling costs, low self-discharge, high energy densities, and long lifetimes [1-3]. The battery management system (BMS) plays an irreplaceable role in ensuring the safety and durability of LIBs, especially for electric vehicles [4]. However, with the increase of cyclic charging and discharging operations or storage time, battery performance gradually declines, resulting in a decrease in battery capacity and an increase in internal resistance. Battery state of health (SOH) has been used to evaluate the aging status of the battery in actual operation. The current energy storage capacity of the battery or the power supply capacity compared with a fresh battery can be used to quantify SOH [5]. SOH plays a key role in ensuring the timely maintenance of the battery system and preventing safety accidents [6]. Ideally, the SOH of the battery when it leaves the factory can be considered 100%, and it decreases with the increase of the number of cycles and reaches zero at the end of life (EOL) of the battery. Therefore, accurate SOH can help us correctly judge the aging degree of the battery and provide guidance for the rational use of the battery [7]. In the present research, the definition of SOH is generally divided into two types: internal resistance definition method and capacity definition method. Wherein the second definition is more widely used, usually defined as the ratio of the current maximum capacity to the rated capacity of the battery [8]. According to the definition of battery SOH, considering that the actual capacity of the battery is less than 70% of the rated capacity, many of its performance deteriorations have increased. Therefore, 70%～80%

battery SOH is regarded as EOL[9, 10]. As far as we know, LIBs are non-linear electrochemical systems, and their internal electrochemical reactions are coupled with each other, and they are affected by the mechanical-electrical-thermal coupling, which makes it difficult to accurately describe their aging process, which makes SOH estimation a huge challenge.

To cope with the above-mentioned problems, in recent years, researchers have proposed a variety of battery SOH estimation methods. Especially with the large-scale increase in the number of electric vehicles, research on battery SOH estimation and remaining useful life (RUL) prediction has received unprecedented attention [11]. The battery SOH estimation approaches can be roughly divided into three categories according to the principle and structure, specifically: empirical approaches, model-based approaches, and data-driven approaches [12, 13]. Empirical approaches is mainly through calendar test or cycle test, and the data obtained is used to describe the simple characteristics of battery aging, and its accuracy and reliability are not high enough [9]. However, their accuracy is limited because the battery degradation is generally affected by some unknown stressing factors. In addition to empirical methods, model-based approaches mainly use differential equations or differential equations to simplify battery electrochemical systems to quantify battery health over the entire battery life, such as equivalent circuit models (ECMs) [14-16], and electrochemical models (EMs) [17, 18]. Generally, EMs with partial differential equations (PDE) has a relatively large computational cost, and it is difficult to apply to vehicle embedded systems for long-term prediction [19]. Compared with the previous two methods, the data-driven method has attracted widespread attention in recent years due to its flexibility and model-free characteristics [20].

Recently, machine learning as an important data-driven method has been rapidly developed and widely used. Artificial neural network (ANN) [21], support vector

machine (SVM) [22], Gaussian process regression (GPR) [23] are important components of machine learning. Since GPR is a probability model that can quantify the level of uncertainty through confidence intervals, it has obvious advantages compared with the other two non-probilistic models of machine learning. In addition, because GPR is constructed based on Bayesian theory, the prediction results can be clearly explained in a probabilistic way [24]. Furthermore, the structure of the GPR is especially simple, as its performance is mainly decided by a mean function and a covariance function [19]. In most of the current literature, the mean function in the estimation methods based on GPR mainly includes zero mean [25, 26], constant mean [27], linear mean, and quadratic polynomial (QP) mean [28, 29]. The covariance function (also called kernel function) of GPR is generally single, such as squared exponential (SE) kernel [26, 28, 30-32], Periodic (Per) kernel [28, 31], and Matern kernel [31-33]. Among them, the zero mean function combined SE kernel function is widely used to realize SOH estimation [9]. Due to the irreversible nature of the aging of the battery during operation, the battery exhibits the aging characteristics of continuous capacity degradation and small-scale capacity regeneration during the full life cycle. A single kernel function based on the GPR model is hard to meet the accuracy requirement of SOH estimation for batteries with different aging trends because of its limited ability to capture the battery aging process. To solve this problem, the double squared exponential (DSE) function [28] and the period function combined with the standard SE function [30] are proposed to achieve high prediction performance. More attention needs to be paid to composite kernel function to further improve the estimation accuracy.

Although the GPR has been successfully used to estimate battery SOH with acceptable accuracy, it is also valuable to research the method to further improve its accuracy. In this chapter, a method for estimating the SOH of lithium-ion batteries based on modified GPR combined with charging and discharging features is proposed. First,

the changes in battery voltage and temperature curves among different aging cycles are analyzed in detail, and HIs that can effectively represent the health status of the battery is proposed. Then, the Pearson correlation analysis method is used to quantify the correlation between HIs and SOH, and three HIs with strong correlation are adopted in this article. Next, a novel compound kernel function is proposed for battery SOH estimation, and different pairs of mean function and kernel function chosen from four mean functions and sixteen kernel functions are used to construct GPR models, and their estimation accuracy is compared subsequently. Finally, four different batteries with various initial health conditions from the NASA battery dataset are used to verify the performance of the proposed method. Three indicators, MAE, RMSE, and 95% confidence interval (CI), are used to evaluate the calculation results of each pair of different mean and kernel functions. The proposed SOH estimation method mainly includes three parts: Data Acquisition, Estimation Algorithm, and Error Analysis. The specific flowchart is shown in Figure 3.1.
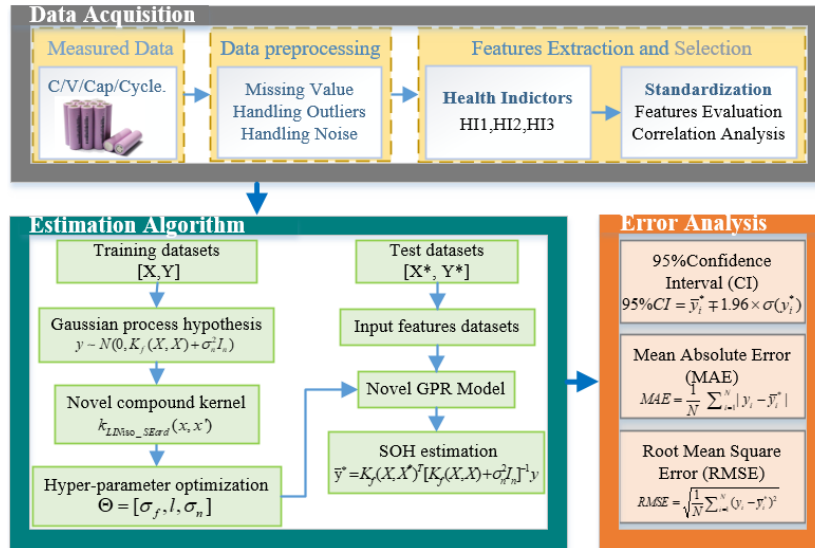
Figure 3.1. The framework of the GPR-based model for battery SOH estimation.

The main contributions of this chapter include the following three points:

(1) Three health indicators (HIs) closely related to battery SOH are extracted from the charging and discharging curves. The time corresponds to the three HIs of the same

charging voltage interval, the same charging voltage interval, and the same discharging temperature interval, respectively.

(2) A novel composite kernel function with automatic relevance detection is proposed for the GPR model to improve the accuracy of battery health estimation. This model can effectively capture the two different phenomena presented by battery capacity decay.

(3) For the first time, a comparative study of the four mean functions and sixteen kernel functions based on the GPR model is carried out. The best pair of mean function and kernel function is determined for battery SOH estimation.

The remainder of this chapter is as follows. The aging process and experimental data are analyzed in Section 3.2. Section 3.3, HIs related to SOH are selected by the Pearson correlation analysis method. Section 3.4 describes the basic principles of GPR and proposes a new compound kernel function for accurate SOH estimation. The experimental results, verification, and discussion are in Section 3.5. The conclusion is arranged at the end of this article.

3.2 Battery dataset

The battery dataset of NASA Ames Prognostics Center of Excellence has been widely concerned and applied worldwide due to its good data integrity and early publication time. This chapter uses the battery data numbered NO.5, NO.6, NO.7, and NO.18 from this public dataset [34]. The rated capacity and voltage of batteries are all 2Ah and 3.7V, respectively. In the cyclic aging stage, each battery undergoes three different test conditions, including charging, discharging, and impedance measurement at room temperature (24 °C), as shown in Figure 3.2(a). First, charge all batteries with a current of 1.5A until the voltage reaches 4.2V, and then switch to the constant voltage

(CV) charging mode and keep the voltage constant until the charging current drops to 20mA. The above process can make the batteries fully charged and is usually called the constant current and constant voltage (CC-CV) charging process. During the constant current (CC) discharging process, batteries are discharged with 2A CC until the voltage of the four batteries drop to the discharge cut-off voltages of 2.7V, 2.5V, 2.2V, and 2.5V, respectively. Electrochemical impedance spectroscopy (EIS) is conducted in each aging cycle with a frequency range of 0.1Hz and 5kHz. With the repeated cycling of the battery, the battery will gradually decline until the available capacity reaches 70% of the rated capacity and end the experiment. The detailed description of the experiment is shown in Table 3.1.

Table 3.1. Detailed description of experiments for four batteries.

Battery ID	Rated Capacity (Ah)	Rated Voltage (V)	Voltage upper/lower limit(V)	Charge /Discharge current magnitude(A)	Temperature(℃)
NO.5	2	3.7	4.2/2.7	1.5/2	24
NO.6	2	3.7	4.2/2.5	1.5/2	24
NO.7	2	3.7	4.2/2.2	1.5/2	24
NO.18	2	3.7	4.2/2.5	1.5/2	24

Figure 3.2(a) presents the corresponding changes in voltage, current, and temperature under each cycle test. It notes that the temperature reaches its maximum at the end of discharge. While the capacity degradation is shown in Figure 3.2(b), in which a long-term downward trend with the increase of cycle number is evident, some local capacity regeneration is observed during the aging process. The degradation trends of the four batteries are relatively similar. Still, the degradation speeds show apparent difference and the corresponding cycles when the four batteries reach 70% of the rated capacity are also different.
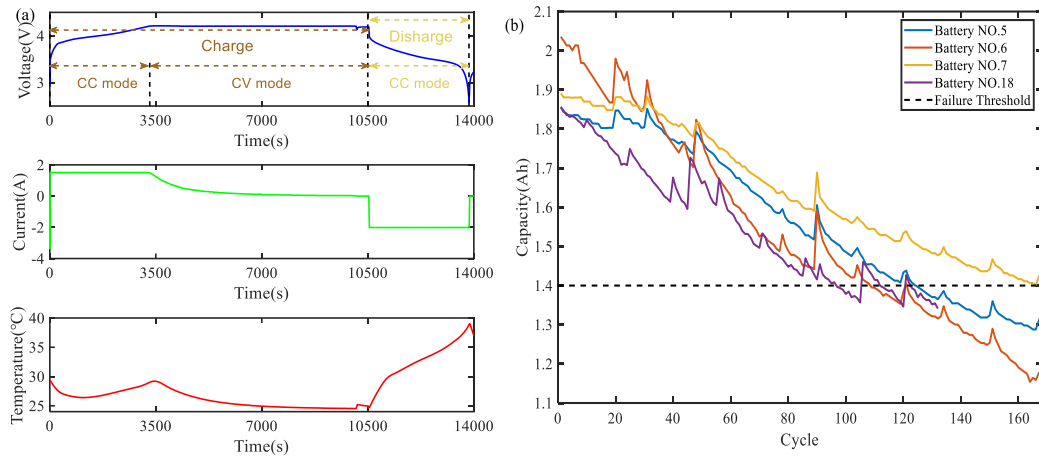
Figure 3.2. Battery aging cycle principle and capacity degradation characteristics.

(a) aging cycle test for voltage, current, and temperature; (b) aging curves of different batteries.

3.3 Features extraction and selection

Based on machine learning algorithms, extracting data related to labels as output is necessary to establish a supervised data-driven model. The data quality directly affects the training effect and estimation accuracy of machine learning, so the extraction and selection of features are vital.

3.3.1 Features extraction

During the normal operation of the battery, battery parameters such as current, voltage, temperature, and time can be directly measured by BMS. Only a tiny amount of data can be used to analyze the battery aging process, even though the large amount of data collected by BMS and screening high-value data from massive data has become a vital issue that needs to be solved. As one of the effective methods to solve the above problems, extracting features from considerable data has been widely used in data-driven methods. Data preprocessing is a critical step of features for GPR estimation and determines the reliability of the estimation method to a certain extent. Features dramatically reduces data redundancy and plays an essential role in improving the

efficiency of raw data [6]. Researchers have conducted a lot of in-depth research on different methods of extracting features from experimental battery measurement or historical data, such as voltage, time, temperature [23].

As the number of battery cycles increases, the electrolyte inside the battery is consumed, and the electrode activity decreases, resulting in an increase in resistance and a decrease in capacity [35]. The voltage and temperature curves of NO.5 battery under different charge and discharge cyclic conditions are shown in Figure 3.3. The experimental results show that these different cycle curves can reflect the aging process of the battery from multiple angles. These charging-discharging curves shift to the left with the cycle gradually increasing. In the CC-CV charging and CC discharging modes, three features are extracted as features to represent the change process of battery health and performance, named HI1, HI2, and HI3.

HI1: The time corresponds to the same charge voltage interval. Due to the irreversibility of the electrochemical reaction inside the battery, the battery capacity continuously reduces as its repeated use. With the gradual decrease of the active electrolyte, the electrode activity continues to weaken, and the available capacity of the battery and the charging capacity during the same voltage range gradually decrease, especially between 3.8V and 4.1V. Considering the positive correlation between the above two parameters, HI1 can be used as a feature factor to characterize the battery aging process.

HI2: The time corresponds to the same discharge voltage interval. As the internal resistance of the battery becomes larger and larger during the aging process, the aging is further aggravated, resulting in the gradual shortening of the time from 3.8V to 2.8V, and the shorter discharge time. Therefore, the time corresponding to the discharge voltage of the CC operating mode can be used as an important indicator to quantify

battery health.

HI3: The time corresponds to the same discharge temperature change. Due to the positive correlation between the discharge temperature and the internal resistance, the internal resistance of the battery gradually increases as the number of cycles increases. Therefore, when the same energy is released, the time required for the battery to rise by 9°C from the lowest temperature is gradually shortened, so HI3 can be used to characterize the change process of the battery's entire life.
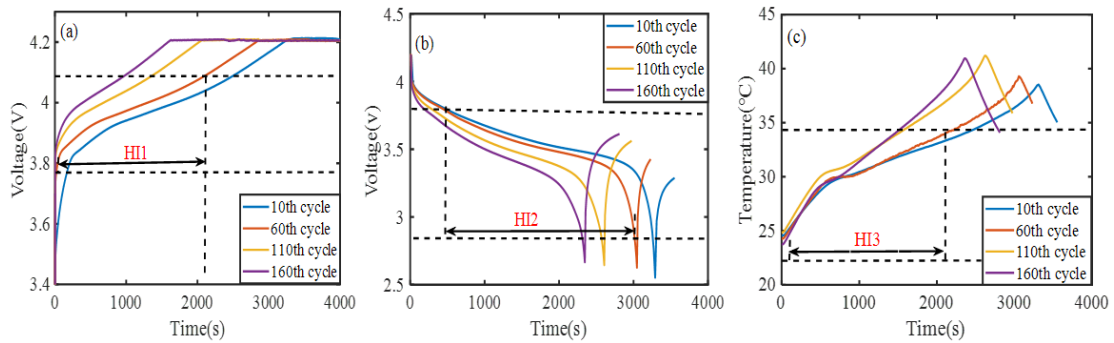
Figure 3.3. Charging and discharging properties of NO.5 battery under different cycles. (a) voltage aging curve (charging); (b) voltage aging curve (discharging); (c) temperature aging curve (discharging).

3.3.2 Features selection

Features selection is an important preprocessing process in the early stage of battery SOH estimation based on the GPR model. The purpose of HI selection is to give priority to the features closely related to battery health and ignore the less important ones. That is, removing unimportant and redundant features not only helps to reduce the calculation cost but also helps to obtain more accurate and reliable results [36]. In the process of studying battery aging, the characteristics of battery aging from different perspectives are proposed, such as based on voltage, current, and temperature [6]. Therefore, it is very important to ensure that HIs that have a strong correlation with the battery SOH are used as the training data set of the compound kernel function based on

the GPR model. To quantify the correlation between the selected HI and SOH, it is necessary to adopt a method that can effectively measure the correlation between HI and SOH. The high-value features not only retain valuable data and eliminate a large amount of irrelevant data, reduce the calculation cost, but also improve the reliability and efficiency of battery SOH estimation [37]. Since the HIs extracted in this paper are characterized by a normal distribution with a good linear relationship with the labels, Pearson correlation analysis (PCA) was used to analyze the correlation between HI and SOH, and the PCA expression is shown in Equation (3.1) [38]:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{3.1}$$

Where the value of r_{xy} ranges from -1 to 1. The closer the absolute value of the correlation coefficient is to 1, the better the relevant feature value of the label. Accordingly, the absolute value of the correlation coefficient is closer to 0, indicating a poorer correlation. In this chapter, three HIs are selected as comparison sequences, and SOH is used as the reference sequence. Through the PCA algorithm, we can get the correlation between the characteristics of the battery and the SOH, as shown in Table 3.2, which shows that there is a strong correlation between SOH and HIs. Therefore, they are used as the input data of the GPR model to complete the battery SOH estimation.

Table 3.2. Pearson correlation coefficient between HIs and SOH.

Battery ID	HI1	HI2	HI3
NO.5	0.9947	0.9958	0.9918
NO.6	0.9875	0.9958	0.9838
NO.7	0.9878	0.9871	0.9774
NO.18	0.9745	0.9954	0.9898

3.4 Methodology

The GPR model is based on Bayesian framework and has the advantages of flexibility, probability, and non-parameter, which leads to a lot of research on battery SOH estimation. The basic principles of the GPR model are described first in Section 3.4.1. Then the mean function and kernel function that plays a key role in Gaussian Process (GP) are analyzed in Section 3.4.2. According to the corresponding analysis results, a novel composite kernel function is proposed in Section 3.4.3.

3.4.1 Basic principles of GPR model

The GPR model is a probabilistic modeling framework that applying Bayesian inference by combining machine learning and statistics [27]. GP can be expressed as $f(x) \sim GP(m(x), k_f(x,x'))$. Its properties are completely determined by $m(x)$ and $k_f(x,x')$ as shown in Equation (3.2).

$$\begin{cases} m(x) = E(f(x)) \\ k_f(x,x') = E[(f(x) - m(x))(f(x') - m(x'))] \end{cases} \tag{3.2}$$

where $m(x)$ is the mean function and $k_f(x,x')$ is the covariance function. The zero mean function and the squared exponent (SE) covariance function (also known as the RBF kernel) have become the most commonly used combination of the GPR model [31]. As for practical application, the observed value y can be expressed by function $f(x)$ with Gaussian noise as shown in Equation (3.3).

$$y = f(x) + \varepsilon \tag{3.3}$$

where ε is the white noise that obeys the Gaussian distribution, which plays a key role in the prior distribution of GP. The specific Equation , as shown in Equation (3.4) - (3.5).

$$y \sim N(0, K_f(X,X) + \sigma_n^2 I_n) \tag{3.4}$$

$$K_f(X,X) = (\sigma_f^2 e^{-\frac{(x_i - x_j)^2}{2l^2}})_{n \times n} \tag{3.5}$$

There is a positive correlation between the distances of the corresponding x_i and x_j, σ_f^2 and l represent signal variance and length scale, respectively. The maximum likelihood method is one of the most effective methods to learn the hyper-parameter set $\Theta = [\sigma_f, l, \sigma_n]$ through the training data set. The corresponding negative logarithmic marginal likelihood (NLML) expression is as Equation (3.6) follows:

$$L = logp(y \mid X, \Theta) = -\frac{1}{2}y^T[K_f(X,X)+\sigma_n^2 I_n]^{-1}y - \frac{1}{2}log(det(K_f(X,X)+\sigma_n^2 I_n)) - \frac{nlog2\pi}{2} \quad (3.6)$$

The conjugate gradient method is widely used to find the optimal solution. The basic principle is to obtain the maximum values of objective function by deriving the log-likelihood function, as shown in Equation (3.7).

$$\frac{\partial}{\partial \Theta_i} logp(y \mid X, \Theta) = \frac{1}{2}[\alpha\alpha^T - (K_f(X,X)+\sigma_n^2 I_n)^{-1}\frac{\partial}{\partial \Theta_i}(K_f(X,X)+\sigma_n^2 I_n)] \quad (3.7)$$

in which $\alpha = (K_f(X,X)+\sigma_n^2 I_n)^{-1}y$. The variables of the new training set X^* all obey the Gaussian distribution. The prediction of the test sample is achieved through the conditional probability distribution function $p(y^* \mid X, y, X^*)$. Therefore, the joint prior distribution of the observed value y and the predicted value y^* are expressed as Equation (3.8).

$$\begin{bmatrix} y \\ y^* \end{bmatrix} \sim N\left(0, \begin{pmatrix} (K_f(X,X)+\sigma_n^2 I_n & K_f(X,X^*) \\ K_f(X,X^*)^T & K_f(X^*,X^*) \end{pmatrix}\right) \quad (3.8)$$

The GP prediction equation is expressed by Equation (3.9).

$$p\left(y^* \mid X, y, X^*\right) = N\left(y^* \mid \overrightarrow{y}^*, \sigma^2(y^*)\right) \quad (3.9)$$

Where $\overline{y}^*$ and $\sigma^2(y^*)$ are given as Equation (3.10) [39].

$$\begin{cases} \overline{y}^* = K_f(X,X^*)^T[K_f(X,X)+\sigma_n^2 I_n]^{-1}y \\ \sigma^2(y^*) = K_f(X^*,X^*) - K_f(X,X^*)^T[K_f(X,X)+\sigma_n^2 I_n]^{-1}K_f(X,X^*)) \end{cases} \quad (3.10)$$

Where $\overline{y}^*$ and $\sigma^2(y^*)$ are the predicted mean and covariance, respectively. In addition, to evaluate the predictive performance of the GPR model, this paper uses three key indicators, including 95% confidence interval (CI), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). The 95% CI of the GPR model based on the

compound kernel function represents the uncertainty of the estimation results as shown in Equation (3.11). MAE and RMSE shown in Equation (3.12) are used to describe the deviation between the predicted value and the actual test value.

$$95\%CI = \bar{y}_i^* \mp 1.96 \times \sigma(y_i^*) \tag{3.11}$$

$$\begin{cases} MAE = \dfrac{1}{N} \sum_{i=1}^{N} | y_i - \bar{y}_i^* | \\ RMSE = \sqrt{\dfrac{1}{N} \sum_{i=1}^{N} (y_i - \bar{y}_i^*)^2} \end{cases} \tag{3.12}$$

Here, y_i and y_i^* represent actual and estimated values, respectively.

3.4.2 Mean function

The properties of the GPR model are completely determined by the mean function and the covariance function. The mean function has a significant impact on the data far away from the field of prediction, so a proper choice of the mean function is very important [39]. Since different mean functions have a greater impact on the long-term prediction (training input and test input are farther), the commonly used mean functions include: zero mean function [27], constant mean function [40], linear mean function [28, 29], and QP mean function [29] . The long-term prediction will eventually return to the above-mentioned mean function. The expressions of the four mean functions as listed in Table 3.3. This chapter combines the above four mean functions with different kernel functions to construct GPR models.

Table 3.3. Expressions of four different mean functions [41].

NO.	Mean Function	Expression
1	Zero[29]	0
2	Constant[40]	C
3	Linear[28, 29]	$AX + B$
4	Quadratic Polynomial [28]	$AX^2 + BX + C$

3.4.3 Covariance function

Covariance function is another important component that affects the accuracy of battery SOH estimation. Figure 3.2(b) shows that the battery capacity has experienced a long-term downward trend, while there are occasional, obvious discontinuities and rising steps. However, a GPR model based on a single SE kernel function commonly used to describe the battery aging trend is not enough to capture different trends well. The entire degradation trend and local regeneration phenomenon of LIBs should be considered to achieve accurate estimation. For the above battery aging process, a compound covariance function can be built by combining different single kernel functions to describe the complex problem [27, 31].

In this paper, the local regeneration phenomenon in the entire aging process is approximated by the linear covariance function with isotropic distance metric (LINiso), while the long-term capacity decay is described by the squared exponent (SE) covariance function with automatic relevance determination (SEard). Therefore, a compound covariance function with better battery aging capturing ability can be constructed by combining the LINiso covariance function and the SEard covariance function, and can be expressed in Equation (13).

$$k_{LINiso\_SEard}(x,x') = x^T l^{-2} x' + \sigma_f^2 exp(-\frac{1}{2}(x-x')^T \Lambda^{-2}(x-x')) \qquad (3.13)$$

Where $\Lambda = diag(l)$ is a diagonal matrix corresponding to the dimension of the input space, and σ_f^2 is the signal variance. The length scale l is also called the ARD parameter, which is used to determine the change rate of the function in the relevant input space. In addition to the proposed compound kernel function, other compound kernel functions can also be formed by combining different single ones. Finally, a total of sixteen different kernel functions including single and compound kernel functions are

introduced, and their formulas are listed in Table 3.4. The performance of the above

kernel functions on battery SOH estimation will be compared in next section.

Table 3.4. Formulas of sixteen different kernel functions [41].

NO.	Kernel Function	Formula
1	linear covariance function (LIN)	$k_{LIN}(x,x') = x^T I x'$
2	linear covariance function with ARD (LINard)	$k_{LINard}(x,x') = x^T \Lambda^{-2} x' \qquad \Lambda = diag(\lambda)$
3	linear covariance function with isotropic distance measure (LINiso)	$k_{LINiso}(x,x') = x^T l^{-2} x'$
4	diagonal squared exponential (SEiso)	$k_{SEiso}(x,x') = \sigma_f^2 exp(-((x-x')^T(x-x'))/(2l^2))$
5	ARD squared exponential (SEard)	$k_{SEard}(x,x') = \sigma_f^2 exp(-2^{-1}(x-x')^T \Lambda^{-2}(x-x'))$
6	Maternard_5/2 [27]	$k_{M52}(x,x') = \sigma_{M52}^2 exp(1 + (\sqrt{5}(x-x'))/\Lambda_{M52} + 5(x-x')/(3\Lambda_{M52}^2)) exp(-(\sqrt{5}(x-x'))/\Lambda_{M52})$
7	SEiso + SEiso	$k_{SEiso\_SEiso}(x,x') = k_{SEiso\_1}(x,x') + k_{SEiso\_2}(x,x')$
8	SEard + SEard	$k_{SEard\_SEard}(x,x') = k_{SEard\_1}(x,x') + k_{SEard\_2}(x,x')$
9	SEiso + SEard	$k_{SEiso\_SEard}(x,x') = k_{SEiso}(x,x') + k_{SEard}(x,x')$
10	SEiso + Periodic [25]	$k_{SEiso\_per}(x,x') = k_{SEiso}(x,x') + k_{per}(x,x')$ $k_{per}(x,x') = \sigma_{Per}^2 exp(-(2l_{Per}^{-2} sin^2(\omega(x-x')/2\pi)))$
11	SEard + Periodic	$k_{SEard\_per}(x,x') = k_{SEard}(x,x') + k_{per}(x,x')$
12	LIN + SEiso	$k_{LIN\_SEiso}(x,x') = k_{LIN}(x,x') + k_{SEiso}(x,x')$
13	LINiso + SEiso	$k_{LINiso\_SEiso}(x,x') = k_{LINiso}(x,x') + k_{SEiso}(x,x')$
14	LINiso + SEard	$k_{LINiso\_SEard}(x,x') = k_{LINiso}(x,x') + k_{SEard}(x,x')$
15	LINard + SEiso	$k_{LINard\_SEiso}(x,x') = k_{LINard}(x,x') + k_{SEiso}(x,x')$
16	LINard + SEard	$k_{LINard\_SEard}(x,x') = k_{LINard}(x,x') + k_{SEard}(x,x')$

3.5 Results and discussion

The proposed GPR model based on a novel composite kernel function is analyzed

and verified for battery SOH estimation in this section. The effect of different mean and

kernel functions on estimation results is compared in section 3.5.1. Four batteries are

used to analyze the generalization and effectiveness of the GPR models in section 3.5.2.

In sections 3.5.3 and 3.5.4, the robustness of the proposed method is verified for batteries with various initial health.

3.5.1 Estimation results for different models

In order to verify the estimation accuracy of the four different mean functions and the new composite kernel function (LINiso-SEard), taking battery NO.5 as an example, its battery health estimation results and the corresponding results of relative errors are shown in Figure 3.4. The first 52 cycles of NO.5 battery data are used as the training set, and the remaining data are used as the test set. Among them, Figure 3.4 (a), (c), (g), (e) is the estimation results of the new composite kernel function based on four different mean functions. The black dots represent the actual cycle corresponding to the SOH, the dark blue solid line represents the estimated SOH value, and the light blue area is the 95% confidence interval. Accuracy is better than the QP mean function but worse than the zero mean function. Figure 3.4(b), (d), (f), (h) represent the relative errors corresponding to the four mean functions. With the continuous increase of the cycle, the relative errors have two characteristics. The relative errors were less than 1% in the first 52 cycles and gradually increased in the subsequent cycles. The maximum relative error corresponding to the first three mean functions is less than 6%, and the fourth mean value is as high as 20%. In contrast, the relative error corresponding to the first mean function is less than 2% in the first 100 cycles, which is better than the other three mean functions. Therefore, the GPR model constructed with zero mean and the new composite kernel function (LINiso-SEard) has satisfactory estimation accuracy.
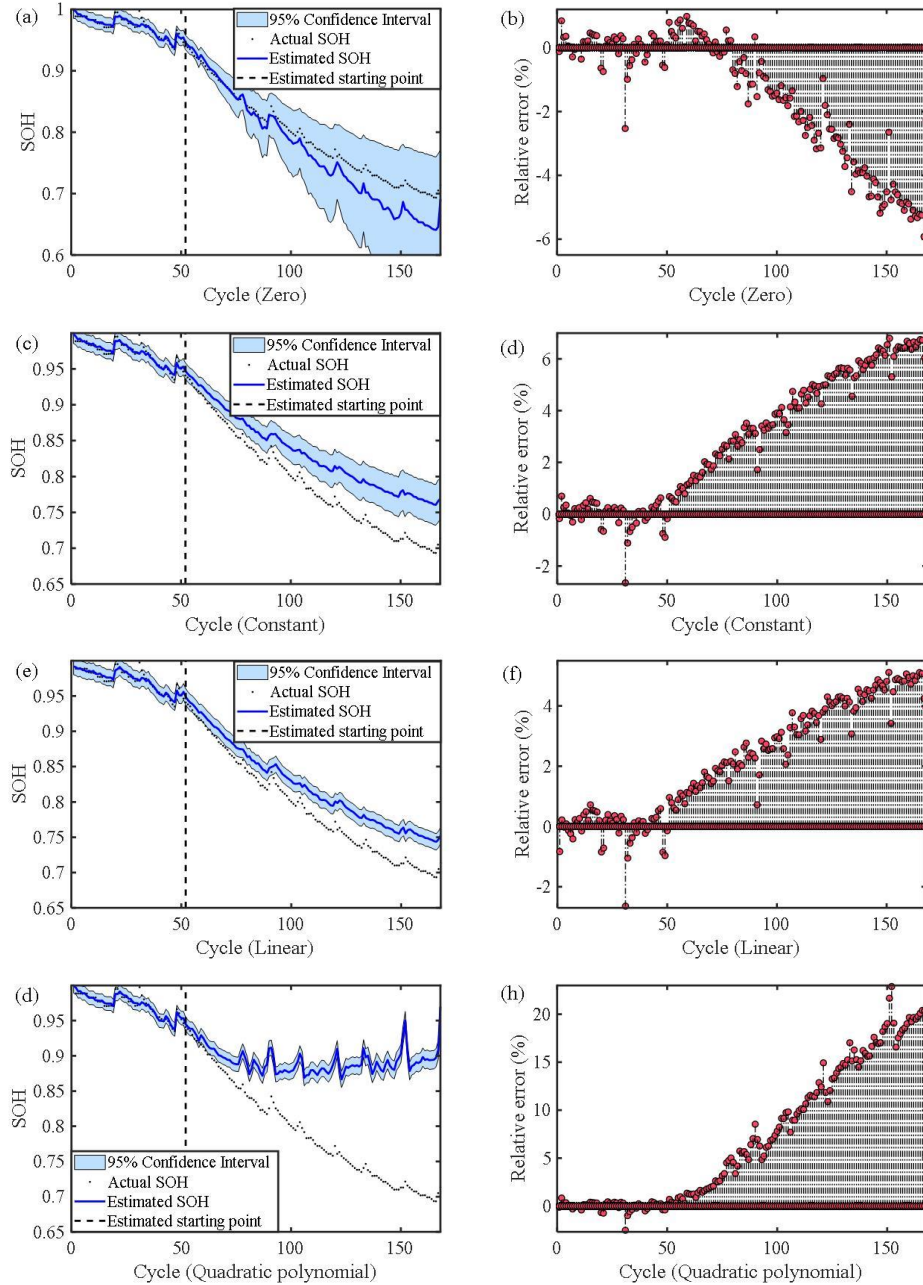
Figure 3.4. Battery SOH estimation results based on different mean functions.

(a) and (b) results of zero mean; (c) and (d) results of constant mean; (e) and (f) results of linear

mean; (g) and (h) results of QP mean.

To verify the generalization ability of the proposed method, four batteries (NO.5,

NO.6, NO.7, and NO.18) are used to verify and compare the MAE and RMSE based

on the four mean functions. The results as shown in Table 3.5. The MAE and RMSE

corresponding to the QP mean function is much larger than those of the other three. The

MAE and RMSE value corresponding to the NO.6 battery is the largest, 15.69% and 20.71%, respectively. The results show that the GPR estimation model constructed by the QP mean function and the LINiso-SEard kernel function cannot simultaneously capture the global decay trend and the local capacity regeneration phenomenon characterized by battery capacity decay. The MAE and RMSE corresponding to the zero-mean function are better than those of the other three. Even though the RMSE of the NO.6 battery corresponding to the constant-mean function is slightly larger than the zero-mean function. The corresponding errors of the constant mean function and the linear mean function are between that of zero and QP mean functions, and their MAE and RMSE are very close. For different GPR models, the zero mean function is the most widely used for battery health estimation [27], which is also employed in this paper.

Table 3.5. The statistical errors of SOH estimation for four mean functions (four batteries).

| NO. | Mean Function | Estimation Error (%) | | | | | | | |
| | | NO.5 | | NO.6 | | NO.7 | | NO.18 | |
		MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
1	Zero	1.70	2.41	1.58	1.96	0.59	0.79	0.43	0.53
2	Constant	2.98	3.78	1.69	1.81	0.91	1.14	0.69	0.86
3	Linear	2.29	2.83	1.83	1.95	0.81	1.03	0.70	0.87
4	QP	7.24	10.16	15.69	20.71	6.12	8.78	1.47	2.01

Then, to further verify the generalization ability of the proposed method, four batteries (NO.5, NO.6, NO.7, and NO.18) are used to verify and compare based on sixteen kernel functions corresponding to MAE and RMSE, the results as shown in Table 3.6. The sixteen kernel functions (refer to Table 3.4 for details) can be divided into two categories: single kernel function (former six) and compound kernel function (later ten).In a single kernel function, the estimation results of the first three kernel functions all have better estimation accuracy. It can be seen from Table 3.6 that the SEiso kernel function has the larger error among all kernel functions, then the SEiso

kernel function has the largest MAE and RMSE errors, reaching 15.84% and 22.12%, respectively. Therefore, the commonly used single SEiso kernel function cannot accurately capture battery aging information. Besides, the composite kernel function obtained by adding the two SEiso does not improve the accuracy, and the four compound kernel functions (NO.8,9,10,11) also have poor estimation accuracy. In contrast, the five combinations of the linear and SE kernel function (NO.12,13,14,15,16) have relatively good estimation accuracy. The MAE and RMSE of the LINiso-SEard kernel function are 0.3% and 0.39%, respectively, which are better than the single LINiso and SEard. It indicates that the SEard kernel function can effectively improve accuracy. The above results verify that the composite kernel function formed by combing LINiso and SEard kernel functions can accurately capture the trends of battery aging.

Table 3.6. The statistical errors of SOH estimation for sixteen different kernel functions.

NO.	Kernel Function	Estimation Error (%)							
		NO.5		NO.6		NO.7		NO.18	
		MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
1	$k_{LIN}(x,x')$	1.51	1.79	0.51	0.69	1.64	2	0.52	0.64
2	$k_{LINard}(x,x')$	1.27	1.5	0.55	0.71	1.62	1.97	0.57	0.7
3	$k_{LINiso}(x,x')$	1.6	1.9	0.51	0.68	1.67	2.04	0.52	0.64
4	$k_{SEiso}(x,x')$	7.32	10.24	3.69	4.75	10.95	15.09	0.93	1.32
5	$k_{SEard}(x,x')$	4.39	6.26	0.75	0.96	15.84	22.12	0.49	0.65
6	$k_{M52}(x,x')$	2.57	3.44	4.31	5.86	7.47	10.16	1.7	2.32
7	$k_{SEiso\_SEiso}(x,x')$	7.32	10.24	3.69	4.75	10.95	15.09	0.93	1.32
8	$k_{SEard\_SEard}(x,x')$	4.39	6.26	0.96	1.28	0.83	1.14	0.49	0.64
9	$k_{SEiso\_SEard}(x,x')$	4.37	6.24	1.03	1.31	4.56	6.02	0.41	0.53
10	$k_{SEiso\_per}(x,x')$	4.38	6.25	1.21	1.63	4.64	5.84	0.58	0.76
11	$k_{SEard\_per}(x,x')$	4.39	6.26	1.21	1.63	4.64	5.85	0.58	0.76
12	$k_{LIN\_SEiso}(x,x')$	1.49	1.84	0.33	0.47	1.41	1.8	0.31	0.42
13	$k_{LINiso\_SEiso}(x,x')$	2.82	3.79	0.33	0.48	1.53	1.96	0.3	0.42
14	$k_{LINiso\_SEard}(x,x')$	0.42	0.55	0.28	0.45	1.26	1.61	0.3	0.39

| 15 | $k_{LINard\_SEiso}(x,x')$ | 1.46 | 1.83 | 0.43 | 0.56 | 1.3 | 1.67 | 0.44 | 0.56 |
| 16 | $k_{LINard\_SEard}(x,x')$ | 1.46 | 1.83 | 0.82 | 0.94 | 1.87 | 2.31 | 0.58 | 0.71 |

3.5.2 Estimation results for different batteries

In this section, the batteries numbered No.5, No. 6, No.7, and No.18 are used to verify the generalization ability of the proposed method. Figure 3.5 presents the SOH estimation results and relative errors of the four batteries and the corresponding statistical errors are shown in Figure 3.6. The first 52 cycles of the battery are used as the training set, and the remaining are used as the test set. The confidence interval is closely related to the confidence level of the prediction results, 95% confidence interval (blue shaded area) is also provided in Figure 3.5 to quantify the uncertainty of the battery SOH estimation. The narrower the blue area, the higher the degree of credibility of the prediction model estimation results. According to the Figure 3.5 (a), (c), (g), (e), the feasibility of the SOH estimation results of the other three batteries is better than NO.5. The relative errors of the four batteries are shown in Figure 3.5(b), (d), (f), (h). It can be seen that the maximum relative error of the NO.5 battery is 5.92%, which is greater than the other three batteries. The maximum relative error of NO.18 battery accuracy is less than 1.42%. As the first two batteries continue aging, the relative error also increases. However, Figure 3.5(f), (h) show different results, and their errors change not obviously. The NO.18 battery has a higher accuracy, part of the reasons are that it enters the cut-off state earlier than the other three batteries. Although the estimated accuracy of each cell is different, they are all less than 1.8% in MAE. The above results show that the proposed method has good generalization ability for different batteries.
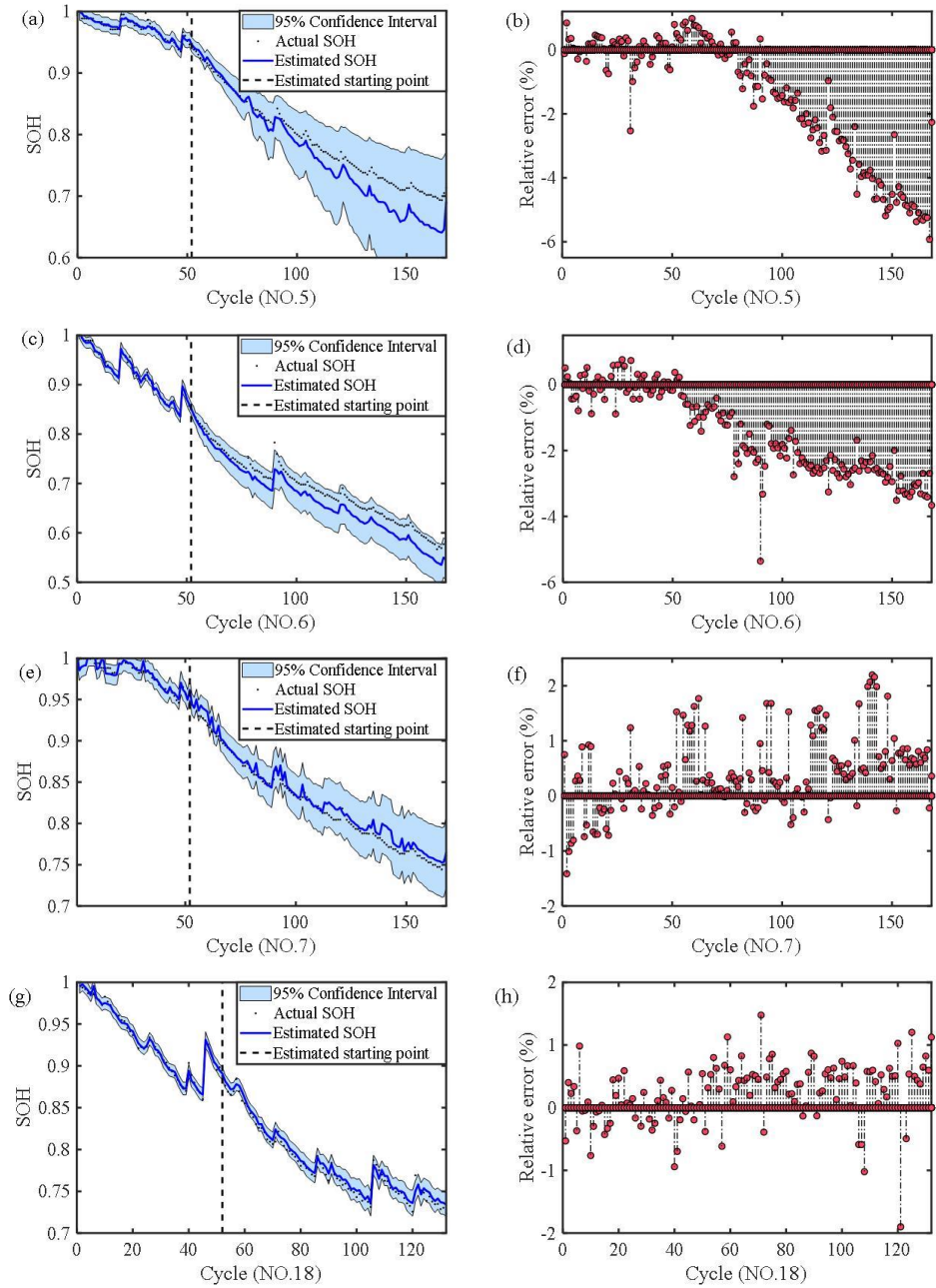
Figure 3.5. SOH estimation results for different batteries.

(a) and (b) results of NO.5; (c) and (d) results of NO.6; (e) and (f) results of NO.7; (g) and (h)

results of NO.18.
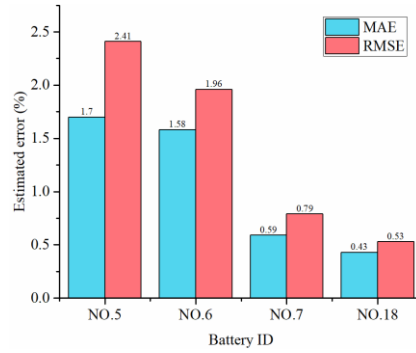
Figure 3.6. The statistical errors of SOH estimation for four batteries.

3.5.3 Estimation results with different starting points

In order to further illustrate the robustness and effectiveness of the proposed method throughout the battery lifespan cycle, the data of the NO.5 battery with different starting points is used to test the proposed method. The SOH estimation results with different starting points are shown in Figure 3.7. As indicated in Figure 3.7(a), (c), (g), (e), the first 40 cycles, the first 70 cycles, the first 100 cycles, and the first 120 cycles are selected as the training data, respectively. It can be observed that as increase of the training set, the accuracy continues to improve, and the region of the confidence interval gradually becomes narrower. This indicates that its reliability gradually improves. The relative error of each cycle corresponding to different training sets is shown in Figure 3.7(b), (d), (f), (h). As the training data increases, the relative error gradually decreases. The above results indicate that the proposed method has a strong robustness and high accuracy under training sets with different estimated starting points.
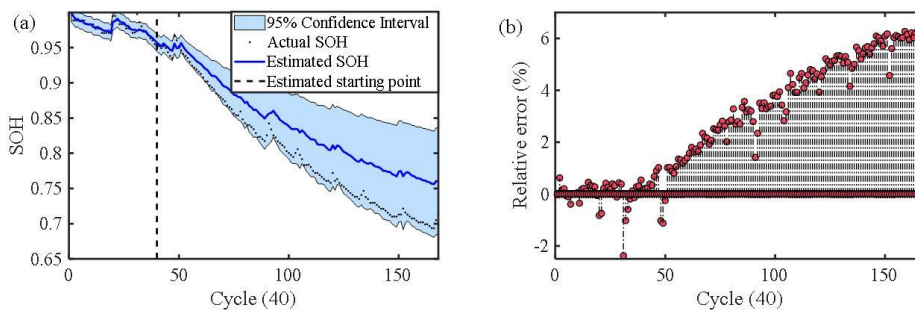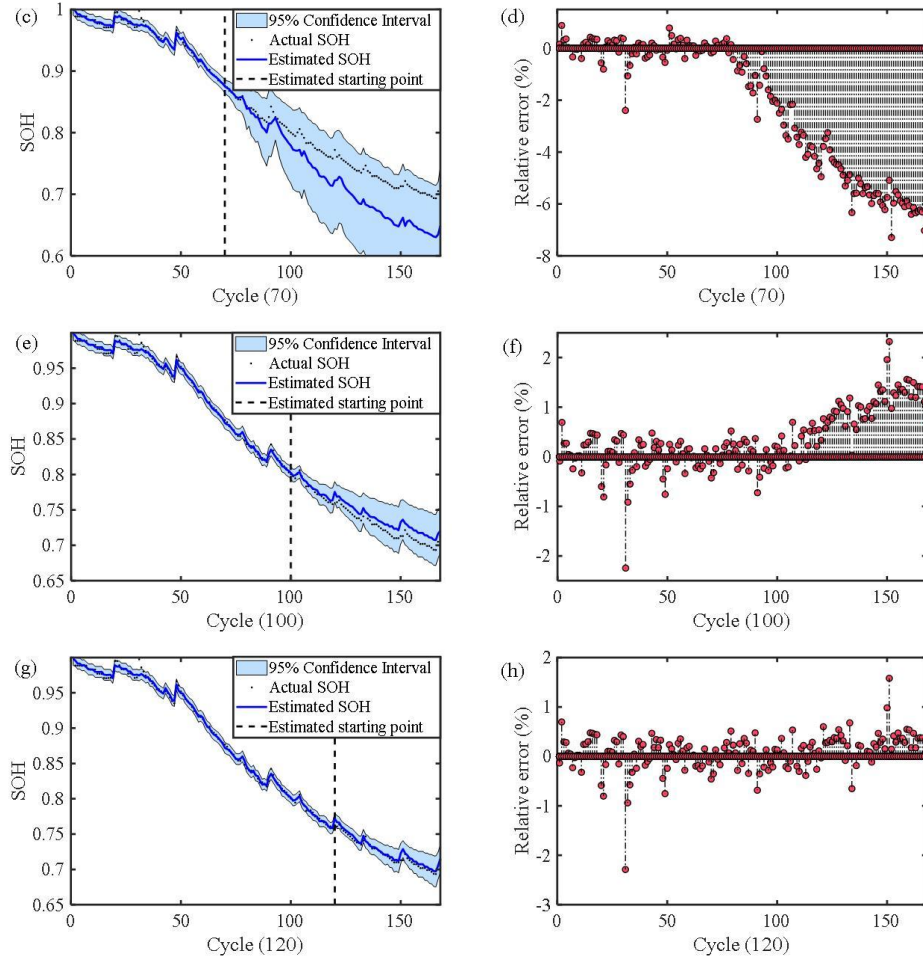
Figure 3.7. Battery SOH estimation results with different starting points.

(a) and (b) results of 40 cycles; (c) and (d) results of 70 cycles; (e) and (f) results of 100 cycles;

(g) and (h) results of 120 cycles

3.5.4 Estimation results by using multiple batteries aging information

In order to analyze the relationship between the aging data of multiple batteries and a single battery. This part uses the aging data of any three of the four known batteries as the training set to estimate the lifespan cycle of the fourth battery. The three batteries numbered NO.5, NO.6, and NO.7 are used as the training set as shown in Figure 3.8, which shows that it has high training accuracy. The relative error is mostly less than 1%, and the maximum relative error is 1.96 %. The NO.18 battery is used as the test set, and its estimation results are shown in Figure 3.9. It can be seen that high accuracy

is obtained, and most of the relative errors are less than 2% and the maximum relative error is 2.94%.

Further, any three of the four batteries are used as the training set, and the fourth battery is used as the test set. The statistical errors of SOH estimation are shown in Figure 3.10 (including the other three cases). The accuracy of NO.18 as test set is the highest, its MAE and RMSE are 1.04% and 1.27%, respectively. The worst accuracy happens on NO.6, with the above two errors up to 3.52% and 3.61%, respectively. The reason for the above phenomenon may be that NO.18 has the shortest SOH variation range, while NO.6 has the longest SOH variation range. When using NO.6 as the test set, the training set cannot cover the SOH range of the test set, which leads to larger errors inevitably. In general, by using multiple battery aging data as training sets, high-precision SOH estimation models can be obtained.
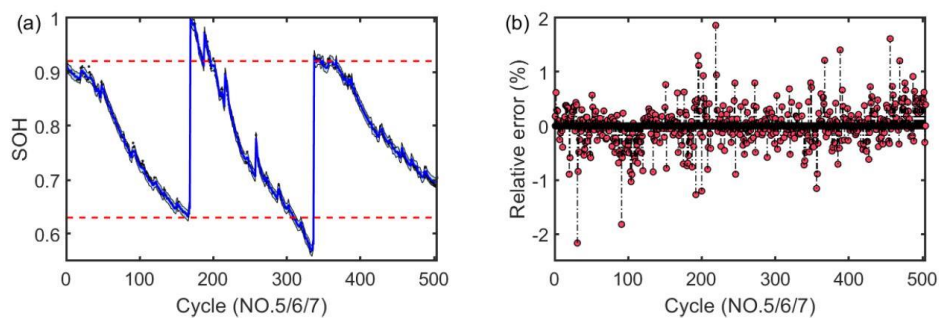
Figure 3.8. The SOH training result with three batteries.

(a) NO.5, NO.6, and NO.7 as the training set; (b) training set cyclic relative error.
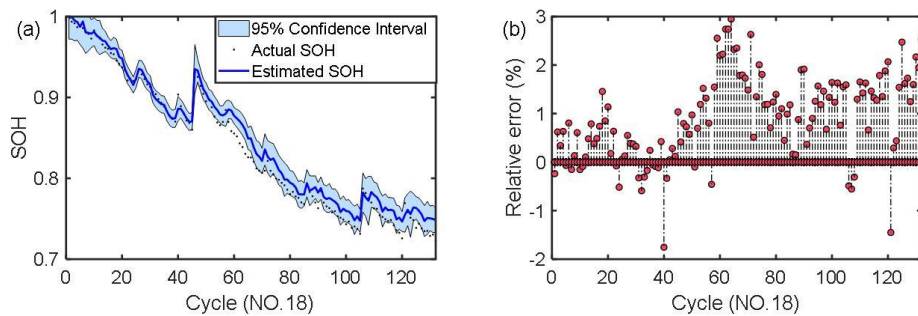
Figure 3.9. The SOH estimation results with NO.18 battery.

(a) lifespan cycle estimation result; (b) test set cyclic relative error.
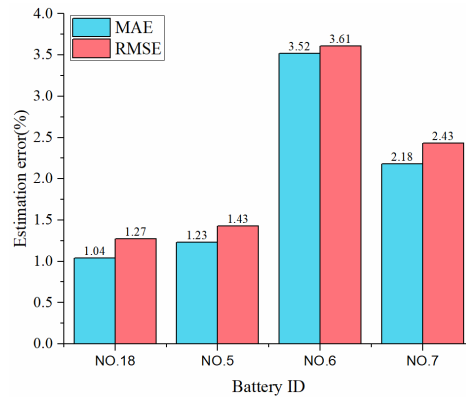
Figure 3.10. The statistical errors of SOH estimation for any three batteries as training set.

3.6 Chapter summary

In this chapter, a battery SOH estimation method based on features extraction and a modified GPR model is presented. Based on the analysis of measured battery current, voltage, and temperature, three effective features are extracted from the charge and discharge processes. Due to the different characteristics of the battery aging process, a single kernel function cannot be used to achieve a satisfactory estimation accuracy. A novel composite kernel function by combining the LINiso and SEard kernel functions is proposed, which can effectively describe the local regeneration phenomenon and the long-term degradation trend of battery health.

Experiments on four batteries of the same material from NASA show that the proposed method is suitable for estimating battery with different initial starting points and suitable for other batteries with the same starting points. The MAE and RMSE of SOH estimation are only 1.7% and 2.41% for NO.5 battery, respectively, and are lower than 1.8% and 2.5% for all batteries. By using multiple battery data as a training set, a high-precision model can be established to estimate battery SOH over its whole life cycle.

Due to the difference of different batteries during charging and discharging, the limitations of this method may not be suitable for other types of batteries. In order to solve the shortcomings of this method, we will focus on its improvement to ensure the applicability for different kinds of battery cells and battery packs in the future. In addition, data-driven battery SOH estimation will play an essential role in battery health management and secondary utilization guidance. More advanced algorithms will be proposed by combining big data technology.

References

[1] X. Hu, L. Xu, X. Lin, M.J.J. Pecht, Battery lifetime prognostics, *Joule*, 4 (**2020**) 310-346.

[2] K.A. Severson, P.M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M.H. Chen, M. Aykol, P.K. Herring, D. Fraggedakis, M.Z. Bazant, S.J. Harris, W.C. Chueh, R.D. Braatz, Data-driven prediction of battery cycle life before capacity degradation, *Nat. Energy*, 4 (**2019**) 383-391.

[3] Y. Che, Z. Deng, X. Lin, L. Hu, X. Hu, Predictive battery health management with transfer learning and online model correction, *IEEE Transactions on Vehicular Technology*, 70 (**2021**) 1269-1277.

[4] B. Jiang, H.F. Dai, X.Z. Wei, T.J. Xu, Joint estimation of lithium-ion battery state of charge and capacity within an adaptive variable multi-timescale framework considering current measurement offset, *Applied Energy*, 253 (**2019**).

[5] Y. Li, K. Liu, A.M. Foley, A. Zülke, M. Berecibar, E. Nanini-Maury, J. Van Mierlo, H.E. Hoster, Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review,*Renewable and sustainable energy reviews*, 113 (**2019**) 109254.

[6] Y. Li, D.-I. Stroe, Y. Cheng, H. Sheng, X. Sui, R. Teodorescu, On the feature selection for battery state of health estimation based on charging–discharging profiles, *J. Energy Storage*, 33 (**2021**).

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

[8] T. Okoshi, K. Yamada, T. Hirasawa, A.J. Emori, Battery condition monitoring (BCM) technologies about lead–acid batteries, *J. Power Sources*, 158 (**2006**) 874-878.

[9] X. Li, Z. Wang, J. Yan, Prognostic health condition for lithium battery using the partial incremental capacity and Gaussian process regression, *J. Power Sources*, 421 (**2019**) 56-67.

[10] Z. Deng, X. Lin, J. Cai, X. Hu, Battery health estimation with degradation pattern recognition and transfer learning, *J. Power Sources*, 525 (**2022**).

[11] M.-F. Ge, Y. Liu, X. Jiang, J. Liu, A review on state of health estimations and remaining useful life prognostics of lithium-ion batteries, *Measurement*, 174 (**2021**).

[12] A. Jokar, B. Rajabloo, M. Desilets, M. Lacroix, Review of simplified Pseudo-two-Dimensional models of lithium-ion batteries, *J. Power Sources*, 327 (**2016**) 44-55.

[13] Y. Wang, J. Tian, Z. Sun, L. Wang, R. Xu, M. Li, Z. Chen, A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems, *Renewable and Sustainable Energy Reviews*, 131 (**2020**).

[14] M. Bahramipanah, D. Torregrossa, R. Cherkaoui, M. Paolone, Enhanced equivalent electrical circuit model of lithium-based batteries accounting for charge redistribution, state-of-health, and temperature effects, *IEEE Transactions on Transportation Electrification,* 3 (**2017**) 589-599.

[15] X. Hu, F. Feng, K. Liu, L. Zhang, J. Xie, B. Liu, State estimation for advanced battery management: Key challenges and future trends, *Renewable and Sustainable Energy Reviews*, 114 (**2019**) 109334.

[16] X. Hu, H. Yuan, C. Zou, Z. Li, L. Zhang, Co-estimation of state of charge and state of health for lithium-ion batteries based on fractional-order calculus, *IEEE Transactions on Vehicular Technology,* 67 (**2018**) 10319-10329.

[17] R. Xiong, L. Li, Z. Li, Q. Yu, H.Mu, An electrochemical model based degradation state identification method of Lithium-ion battery for all-climate electric vehicles

application, *Applied energy* , 219 (**2018**) 264-275.

[18]Z. Deng, X. Hu, X. Lin, L. Xu, J. Li, W. Guo, A Reduced-Order Electrochemical Model for All-Solid-State Batteries, *IEEE Trans. Transp. Electrif.*, 7 (**2021**) 464-473.

[19]K. Liu, X. Hu, Z. Wei, Y. Li, Y. Jiang, Modified Gaussian Process Regression Models for Cyclic Capacity Prediction of Lithium-Ion Batteries, *IEEE Trans. Transp. Electrif.*, 5 (**2019**) 1225-1236.

[20]Z. Deng, X. Hu, X. Lin, L. Xu, Y. Che, L. Hu, General Discharge Voltage Information Enabled Health Evaluation for Lithium-Ion Batteries, *IEEE/ASME Trans. Mechatron.*, (**2020**) 1-1.

[21]Z. Xia, J.A.A. Qahouq, Adaptive and fast state of health estimation method for lithium-ion batteries using online complex impedance and artificial neural network, in: 2019 IEEE Applied Power Electronics Conference and Exposition (APEC), IEEE, **2019**, pp. 3361-3365.

[22]X. Feng, C. Weng, X. He, X. Han, L. Lu, D. Ren, M. Ouyang, Online state-of-health estimation for Li-ion battery using partial charging segment based on support vector machine, *IEEE Transactions on Vehicular Technology*, 68 (2019) 8583-8592.

[23]X. Hu, Y. Che, X. Lin, S. Onori, Battery health prediction using fusion-based feature selection and machine learning, *IEEE Transactions on Transportation Electrification*, (**2020**) 1-1.

[24]L. Li, P. Wang, K.H. Chao, Y. Zhou, Y. Xie, Remaining Useful Life Prediction for Lithium-Ion Batteries Based on Gaussian Processes Mixture, *PLoS One*, 11 (**2016**) e0163004.

[25]J. Liu, Z.J.I.A. Chen, Remaining useful life prediction of lithium-ion batteries

based on health indicator and Gaussian process regression model,*IEEE Access*, 7 (**2019**) 39474-39484.

[26] X. Li, C. Yuan, X. Li, Z.J.E. Wang, State of health estimation for Li-Ion battery using incremental capacity analysis and Gaussian process regression, *Energy*, 190 (**2020**) 116467.

[27] C. KI Williams, Gaussian processes formachine learning, Taylor & Francis Group, **2006**.

[28] D. Liu, J. Pang, J. Zhou, Y. Peng, M. Pecht, Prognostics for state of health estimation of lithium-ion batteries based on combination Gaussian process functional regression, *Microelectron. Reliab.*, 53 (**2013**) 832-839.

[29] Jia, J. Liang, Y. Shi, J. Wen, X. Pang, J. Zeng, SOH and RUL Prediction of Lithium-Ion Batteries Based on Gaussian Process Regression with Indirect Health Indicators, *Energies*, 13 (**2020**).

[30] D. Yang, X. Zhang, R. Pan, Y. Wang, Z.J.J.o.P.S. Chen, A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve, *J. Power Sources*, 384 (**2018**) 387-395.

[31] R.R. Richardson, M.A. Osborne, D.A. Howey, Gaussian process regression for forecasting battery state of health, *J. Power Sources*, 357 (**2017**) 209-219.

[32] K. Liu, Y. Li, X. Hu, M. Lucu, W.D. Widanage, Gaussian Process Regression With Automatic Relevance Determination Kernel for Calendar Aging Prediction of Lithium-Ion Batteries, *IEEE Trans. Ind. Inf.*, 16 (**2020**) 3767-3777.

[33] R.K. Pandit, D. Infield, Comparative analysis of Gaussian Process power curve models based on different stationary covariance functions for the purpose of improving model accuracy, Renewable *Energy*, 140 (**2019**) 190-202.

[34] B. Saha, K. Goebel, S. Poll, J. Christophersen, measurement, Prognostics methods

for battery health monitoring using a Bayesian framework, *IEEE Transactions on instrumentation and measurement,* 58 (**2008**) 291-296.

[35] X. Han, L. Lu, Y. Zheng, X. Feng, Z. Li, J. Li, M. Ouyang, A review on the key issues of the lithium ion battery degradation among the whole life cycle, *eTransportation*, 1 (**2019**).

[36] S. Greenbank, D.A.J.a.p.a. Howey, Automated feature selection for data-driven models of rapid battery capacity fade and end of life, *arXiv e-prints,* (**2021**): arXiv-2101.

[37] W. Daelemans, V. Hoste, F. De Meulder, B. Naudts, Combined optimization of feature selection and algorithm parameters in machine learning of language, in: European Conference on Machine Learning, Springer, 2003, pp. 84-95.

[38] P.J.B. Sedgwick, Pearson's correlation coefficient, 345 (**2012**).

[39] S. Roberts, M. Osborne, M. Ebden, S. Reece, N. Gibson, S. Aigrain, Gaussian processes for time-series modelling, *Philos Trans A Math Phys Eng Sci*, 371 (**2013**) 20110550.

[40] S.A. Aye, P.S. Heyns, An integrated Gaussian process regression for prediction of remaining useful life of slow speed bearings based on acoustic emission, *Mechanical Systems and Signal Processing*, 84 (**2017**) 485-498.

[41] C.E. Rasmussen, H.J.T.D. Nickisch, The gpml toolbox version 4.0, (**2016**).

Chapter 4. Early Prognostics of Lithium-ion Battery Pack Health

4.1 Introduction of this chapter

Lithium-ion batteries (LIBs) have been widely used in portable electronics, electric vehicles, and grid-side energy storage systems because of their high energy density, no memory effect, low self-discharge current, long lifecycle, wide temperature range, and other advantages [1–3]. In LIBs, as a complex electro-thermal coupled, time-varying nonlinear electrochemical system, the increasing number of charge–discharge cycles or long storage times cause the loss of active materials inside the battery and the precipitation of lithium ion, which eventually leads to the aging of the lithium-ion battery (e.g., increased internal resistance or reduced capacity [4,5]). The probability of potential safety problems increases dramatically in the later stages of aging, so the accurate prediction of battery health during regular operation plays a vital role in eliminating the battery life anxiety of energy storage, providing maintenance strategies, and avoiding safety incidents. As an essential indicator to characterize the health and life of a battery, the end of life (EOL) is defined as 80% state of health (SOH), where the ratio of current maximum available capacity to the rated capacity of the battery is defined as the SOH [6].

In recent years, research related to battery health has received much attention from scholars, who have obtained a series of research results [5,7–9]. Usually, the battery SOH prediction is performed through the battery management system by combining models with related algorithms based on collected key parameters (e.g., voltage, current, temperature, time, etc.). The research approaches in battery health are mainly divided into model-based methods and data-driven approaches. Model-based approaches describe the internal dynamics of the battery at different scales through mathematical

models (e.g., empirical models [8,10], equivalent circuit models [11–13], and electrochemical models [14–16]), while determining the balance between the complexity of the battery model and the prediction accuracy is still a complex problem that needs to be further addressed. In contrast, data-driven approaches based on machine learning algorithms are dedicated to mining the close relationship between battery aging data and the state of health in the laboratory or under actual operating conditions, without the need to construct mathematical models. Currently used machine learning algorithms mainly include: artificial neural network (ANN) [17], support vector machines (SVM) [18], relevance vector machines (RVM) [19], and Gaussian process regression (GPR) [5,20–22]. Among these, the first three algorithms have a common disadvantage that can easily lead to overfitting [23]. However, GPR is generally computationally efficient, flexible, and easy to implement, though it lacks higher robustness scores. Thus, this paper uses GPR to capture the potential coupling between HIs and battery capacity for achieving battery health prediction since it uses a statistical machine learning process and shows better accuracy and uncertainty expression.

At present, the demand for a long lifetime of battery packs in the field of energy storage is becoming more and more prominent. However, the lifetime test of battery packs takes up much time and much expense. Due to the inconsistency between cells and a higher increase in temperature in the battery pack, the battery presents totally different degradation characteristics compared with its individual cells, and it usually has a much shorter lifetime. To make full use of the aging data of battery cells and to reduce battery pack aging test time, this chapter proposes a method for predicting the future health of the battery pack using the aging data of the battery cells along with the entire lifecycle and early cycling data of the battery pack. An exponential function, a long short-term memory network (LSTM) model, and their weighted fusion are employed to construct the degradation models of HIs, so that the future HIs can be

predicted. Then, by combining the early cycling data of the battery pack with the GPR algorithm, a data-driven model is constructed to achieve the health prediction of the battery pack. The main contributions of this paper include:

(1) An HI fusion degradation model is established to capture the global decay and the local variation of battery HIs simultaneously. An exponential degradation model is fitted to capture the global decay of HIs, while the LSTM degradation model is constructed to imitate the local variation of HIs. By weighting the exponential-based model and LSTM-based model, a fusion degradation model can be created, which can inherit the advantages of these two models.

(2) An early prognostic method of battery pack health is proposed. Based on the early cycling data of the battery pack and the fusion degradation model of HIs, the future HIs of the battery pack can be obtained. Taking the HIs as the inputs of the GPR algorithm, data-driven models can be constructed to predict the future health of the battery pack.

(3) Three health prediction models based on the GPR algorithm are constructed for comparison, including the exponential function-based (EXP-GPR) model, an LSTM-based (LSTM-GPR) model, and the weighted fusion-based (EXP-LSTM-GPR) model. The results show that the fused degradation model has better accuracy.

The remaining of the paper is organized as follows: Section 4.2 introduces the cell and aging experiments; Section 4.3 describes the principle of cell HI extraction and correlation analysis; Section 4.4 presents the main methods used in this paper; Section 4.5 presents the results and discussion, and finally the conclusion of this chapter.

4.2 Aging experiments

This section selects the battery cells under different operating conditions (sixteen battery cells divided into six groups, as shown in Table 4.1) and the battery pack under

the same operating conditions (35 °C_0.5C-0.5C) to complete the aging experiment. In order to analyze the effects of temperature and current on battery aging, battery cells under four different temperatures and three different currents are used for the experiments; to improve the reliability of the experiments, two or three cells are set up in the same experimental environment for each group of cells. The battery cells with a lithium iron phosphate (LiFePO$_4$) cathode have 100 Ah-rated capacity. The battery pack is made up of 15 cells of the same type connected in a series (six groups). The voltage, current, time, and temperature data of the continuous cycle can be collected directly by the charge/discharge tester. The tested battery cells and battery packs are charged and discharged under constant current and constant voltage (CC-CV) conditions. The temperature of the battery pack is set to 35 °C, and the charge/discharge current is 0.5 C. Four types of temperature settings are set for the single-cell experiment: 25, 35, 45, and 55 °C. Three types of charge current settings are set: 0.3, 0.5, and 1C. Finally, two kinds of discharge current settings are set: 0.5C and 1C. The sampling time interval is 30 s for the battery pack and 10 s for the battery cells. The resting time is set to 10 min, as shown in Table 4.1.

Table 4.1. Battery and aging test parameters.

Experimental Setup Conditions	Cell #1	Cell #2	Cell #3	Cell #4	Cell #5	Cell #6	Pack
Temperature (℃)	25	25	35	35	45	55	35
Charge and discharge policies	0.5C		0.3C	0.5C			0.5C
		1C/1C			1C/1C	1C/1C	
(CC-CV/CC)	/0.5C		/1C	/0.5C			/0.5C
Actual temperature range (℃)	25/29	25/31	35/44	36/40	45/52	54/64	36/47

The platform for the aging experiment of the battery as show in Figure 4.1(a), which includes a battery tester, a data logger, a thermal chamber, a computer, a series-connected battery pack, and sixteen battery cells. In the battery aging test experiment, a complete cycle of the charging and discharging process of Cell #4 as show in Figure

4.1(b), where the solid red line indicates the charging and discharging current. The solid blue line indicates the charging and discharging voltage. In the charging phase, the battery is first charged with a constant current (CC) of 0.5C, and when the voltage reaches the upper cutoff voltage, it switches to constant voltage (CV) charging mode until the charging current drops to C/20, which ends the charging process. In the discharging stage, the battery is discharged with constant current discharge at 0.5C until the discharge voltage reaches the lower cutoff voltage; then, the discharging process is finished. The above charging and discharging steps are repeated until the battery capacity decays to 80% of its rated capacity, which marks the end of the experiment. The battery capacity aging curve is shown in Figure 4.2, where Figure 4.2(a) is the capacity decay curve of Cell #4 alone, and Figure 4.2(b) is the capacity decay curve of the battery pack. The battery cell offers different degrees of capacity regeneration in the decay process, among which there are three apparent fluctuations due to the long resting time. The comparison between the battery cell's aging curve and the battery pack's aging curve shows that the number of cycles of the pack is significantly less than the number of cycles of the battery cell when it reaches 80% SOH.
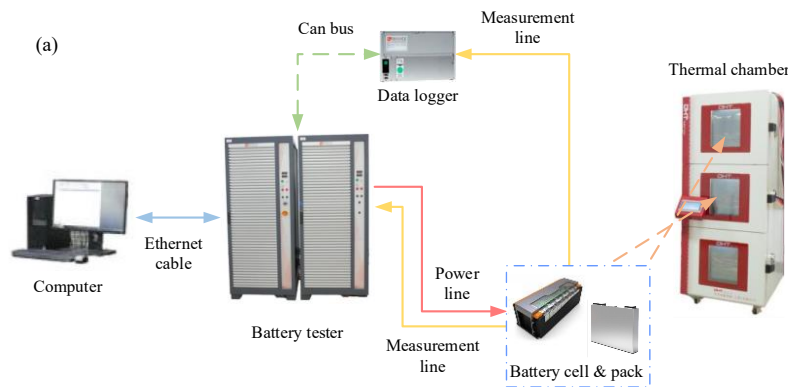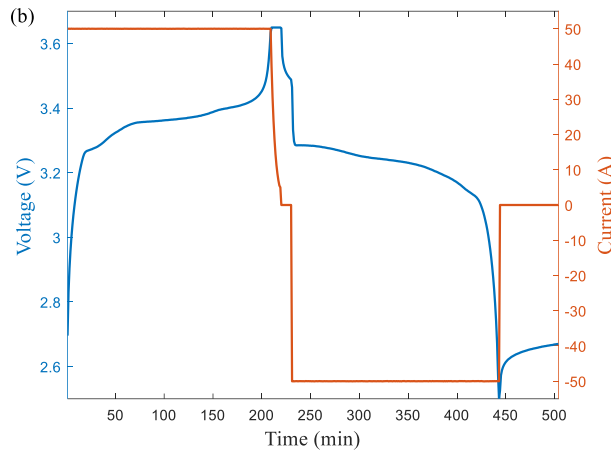
Figure 4.1. Aging test design for battery.

(a) platform for aging experiment of battery; (b) complete cycle charge and discharge curves.
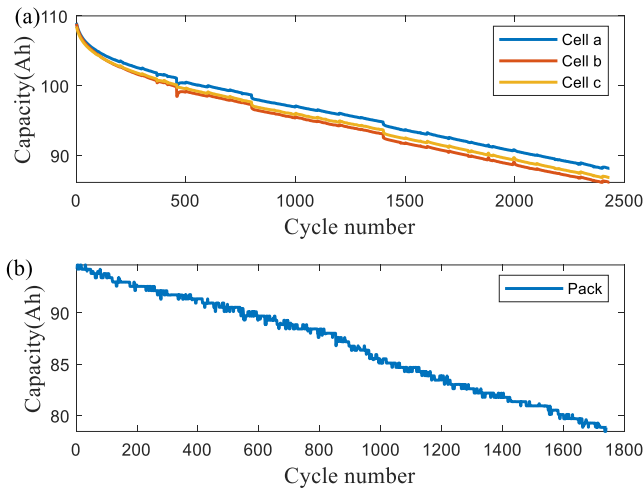
Figure 4.2. Battery aging and capacity degradation curves.

(a) battery cell aging curves; (b) battery pack aging curves.

4.3 HIs extraction and selection

Batteries generate large amounts of data in aging tests, and for data-driven battery health prediction methods, extracting HIs directly or indirectly from battery aging data is an essential step toward eliminating redundant data and improving computational efficiency. The extraction principle of extracting three sets of HIs from battery aging data is outlined in section 4.3.1. The correlation analysis of HIs is performed by the Pearson correlation analysis method in section 4.3.2.

4.3.1 HIs extraction

This section presents the extraction principles of the three HIs. In our previous work [24–26], two HIs, the standard deviation Q sequence (*stdQ*) and the standard deviation of different Q sequence (*stddQ*), have been verified to have excellent performance as the basis for the research of battery pack health prediction in this paper. The incremental capacity (IC) analysis method, as a non-destructive and effective means, has been widely used to evaluate the health status of LIBs. Since the *IC_peak* can capture the electrochemical processes occurring inside the battery and is closely related to the capacity decay [27], it is used as the third HI extracted in this paper.

Figure 4.3(a) represents the *Q-V* discharge curve of the battery under two different aging cycles (100th cycle and 1000th cycle), with the increasing number of cycles releasing the same amount of electricity at a decreasing voltage value. Firstly, a constant discharge voltage interval is selected, and the discharge voltage fragment V_{seg} is shown in Equation (4.1):

$$V_{\text{seg}} = \left[V_{\text{max}}, V_{\text{max}} - \Delta V, V_{\text{max}} - 2 * \Delta V, \cdots, V_{\text{min}} \right] \tag{4.1}$$

where ΔV is the voltage interval that can be obtained through the ampere–time integration of the same voltage interval corresponding to the capacity sequence *Q(v)*, as shown in Equation (4.2):

$$Q(V) = \left[Q_1, Q_2, \cdots, Q_{N_P} \right] \tag{4.2}$$

The difference between the capacities of two different cycles can be noted as $\Delta Q_{\text{c2-c1}}(v)$. The standard deviation of the capacity sequence is indicated as *Q(v)*, and the capacity sequence of different cycle differences are indicated as *stdQ* and *stddQ*, respectively.

As shown in Figure 4.3(b), the increasing number of cycles leads to the increase in the amount of lithium ion loss inside the cell, causing the peak of the IC curve to decrease gradually. Its decay characteristics describe the aging process of the cell. The expression of the IC curve as show in Equation (4.3) [20]:

$$(f^{-1})' = \frac{dQ}{dV} = \frac{I \cdot dt}{dV} = I \cdot \frac{dt}{dV} \tag{4.3}$$

Where $V_{min} = 3.1v, V_{max} = 3.4v, N_p = 50, \Delta V = 6mv, dv = 70mv.$
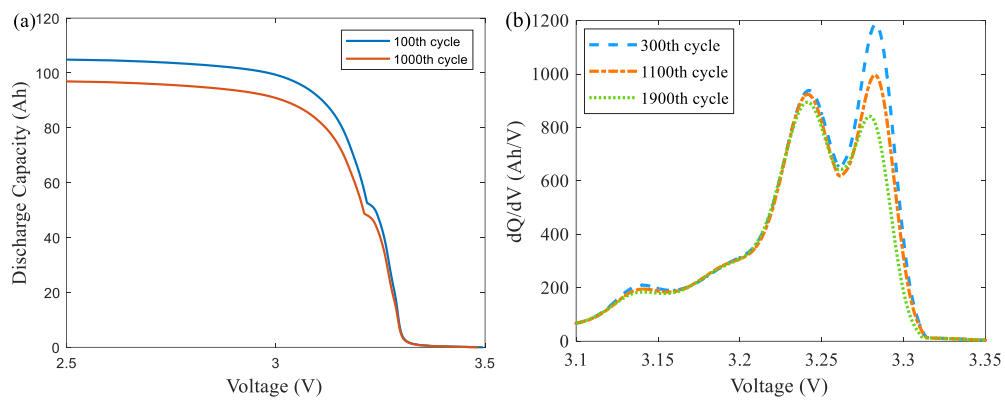
Figure 4.3. Battery aging characteristics curves (Cell #4).

(a) *Q-V* discharge curves for different cycles; (b) IC curves for different cycles.

4.3.2 HIs selection

According to the principle of HIs extraction in Section 4.3.1, the capacity and three sets of HIs decay curves of the battery cells and battery packs can be obtained. Figure 4.4(a) shows the aging decay curve corresponding to the single cell with a rated capacity of 100Ah. With the increasing number of cycles, the first 100 cycles show exponentially decreasing characteristics. From the 101st cycle to the end of life, the capacity showed a linear decay trend, and three capacity regeneration phenomena appeared locally. Figure 4(b), (c), (d) show the decay curves of *stdQ*, *stddQ*, and *IC_peak*, respectively. After normalizing the extracted three groups of HIs, their decay curves are consistent

with the decay curves of the monomer capacity, indicating that the proposed three groups of HIs have a closely related aging characteristic with the capacity.
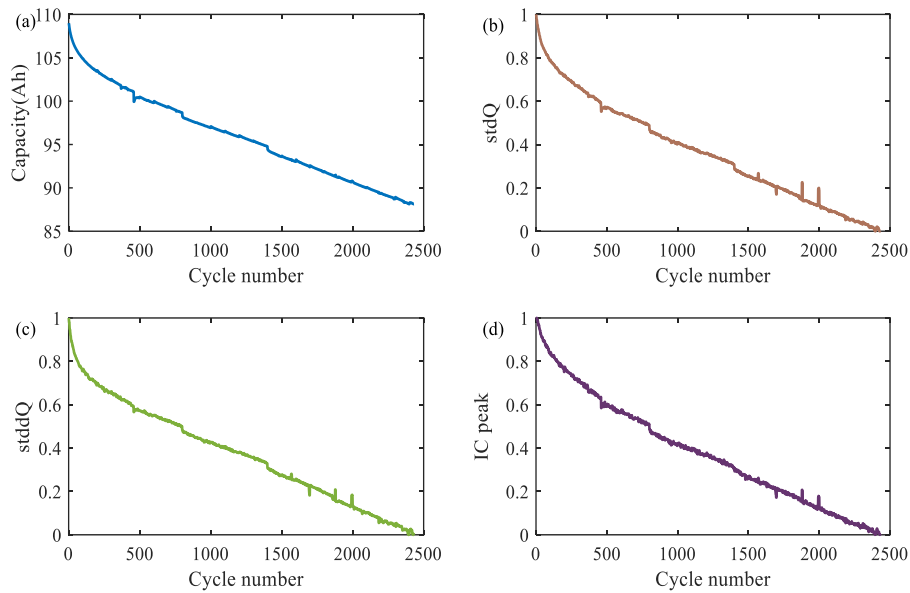
Figure 4.4. Aging curves of monomer capacity and HIs (Cell #4).

(a) capacity aging curve; (b) stdQ aging curve; (c) stddQ aging curve; (d) IC peak aging curve.

Figure 4.5 shows the decay curves of the pack capacity and three HIs of the fifteen cells. Figure 4.5(a) shows the decay curves of the pack capacity; after the abnormal data are eliminated, the remaining capacity shows a linear monotonic decreasing trend with slight local fluctuations. In Figure 4.5(b), (c), the decay curves of *stdQ* and *stddQ* are shown, respectively. There are two apparent fluctuations and opposite directions in two cluster decay curves, indicating that anomalies occurred in two regions during the aging experiments. Figure 4.5(d) shows that *IC_peak* has the same aging characteristics as the capacity decay curve. Although there are two fluctuations in the same direction as the fluctuations in Figure 4.5(b), the changes are minor.
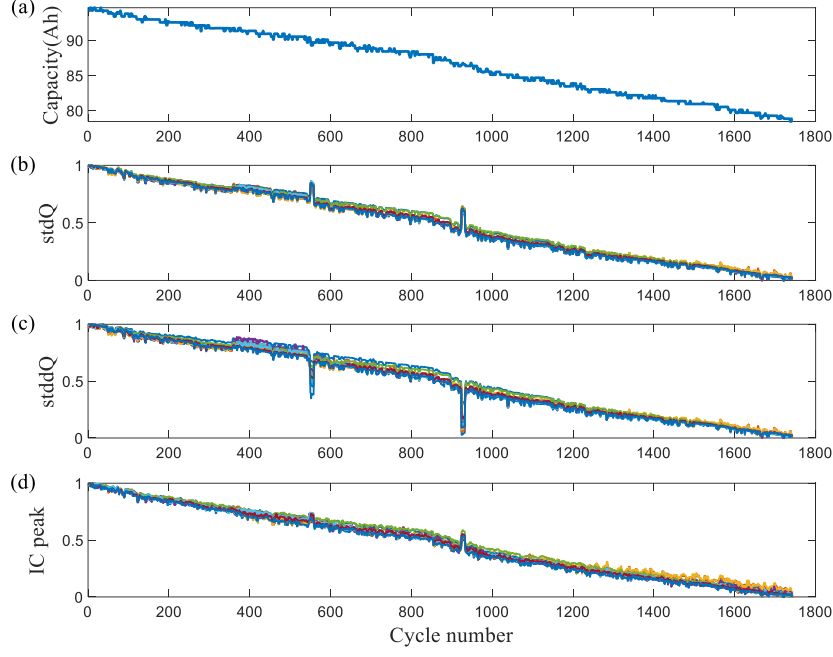
Figure 4.5. Battery pack capacity and HI decay curves.

(a) capacity aging curve; (b) stdQ aging curves; (c) stddQ aging curves; (d) IC peak aging curves.

The above analysis shows that the HIs of cells and the battery pack shows a decay characteristic consistent with the overall capacity. To further quantify the correlation degree between the HIs and the capacity, the quantitative analysis is carried out in this paper using the Pearson correlation analysis method, as shown in Equation (4.4) [28]:

$$\rho_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{4.4}$$

where x_i and y_i are the sample observations of x and y, respectively, and $\bar{x}$ and $\bar{y}$ are the mean of the sample values of variables x and y, respectively.

The HI correlation coefficients between cells, pack, and the corresponding capacity can be obtained from Equation (4.4), and the results are shown in Figure 4.6. As can be seen from the figure, the three sets of HIs used for the same operating condition battery cell and pack all reach 0.99. The *stdQ* HIs contain both battery voltage and discharge capacity curve information. The *stdQ* reflects the uneven battery discharge energy with the voltage, revealing the essential factors to achieve high-

accuracy prediction using this feature. It indicates that the HIs strongly correlate with the battery capacity, and the extracted three sets of HIs that meet the requirements are used to establish the HI degradation model in the next stage.
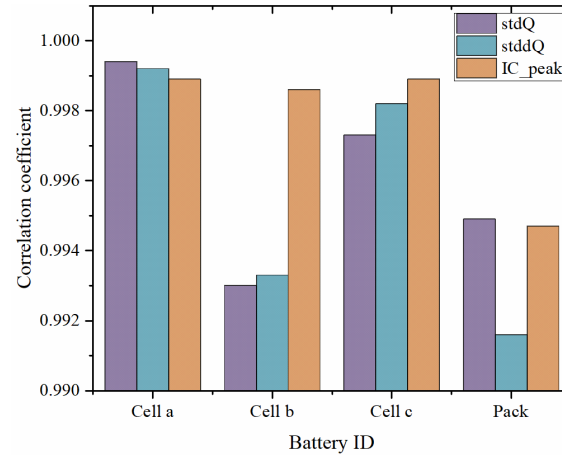
Figure 4.6. Statistical values of the correlation coefficients of the HIs of single cells (Cell #4) and battery packs in the same operating condition.

4.4 Methodology

In this section, a method for battery pack health prognostics is proposed. A brief description of the overall prediction process is given in Section 4.4.1, the degradation model used to predict the future HIs of the pack is introduced in Section 4.4.2, and the pack health estimation model based on the GPR algorithm is presented in Section 4.4.3.

4.4.1 Battery pack health prognostics

The proposed scheme in this chapter consists of three sections and the flowchart of the proposed scheme for battery pack health prognostics as show in Figure 4.7. First section: data acquisition. First, the measured data of the battery with noise and abnormalities are pre-processed, including filling, deletion, noise reduction, smoothing, and normalization. Then, the critical data, such as voltage, current, time, etc., are screened out from the battery cells and battery pack data. Second section: the model

construction process. First, three sets of HIs—*stdQ*, *stddQ*, and *IC_peak*—are extracted from the Q-V partial discharge curve based on the principle of HI extraction. Second, the correlation between the HIs and capacity is assessed based on the Pearson correlation analysis method. Next, the exponential degradation model and the LSTM degradation model are established based on the cell's entire life aging data. Then, the first 10% of the battery pack HIs are used to fine-tune the above two degradation models to obtain the battery pack HI exponential degradation model and LSTM degradation model, and the two models are fused according to different weights to obtain the battery pack HI fusion degradation model. Finally, the battery pack capacity estimation model is constructed based on the GPR algorithm combined with the early battery pack HIs and capacity. Third section: battery pack health prediction. First, the predicted value of the battery pack's future cycling HIs can be obtained by fine-tuning the HI fusion degradation model. Then, the predicted value is used as the input of the battery pack capacity estimation model, which can realize the battery pack health prediction.
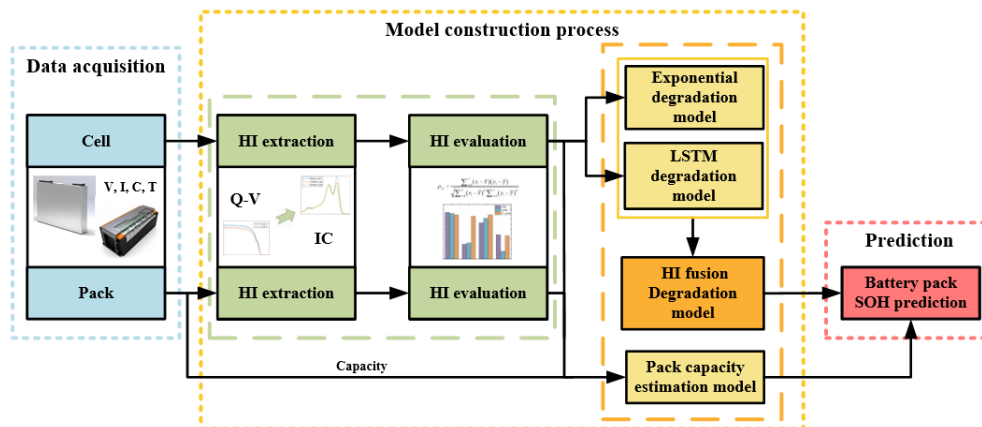

Figure 4.7. Flowchart of the proposed scheme for battery pack health prognostics.

4.4.2 HIs Degradation model

A HI fusion degradation model is proposed for predicting the HIs of the future cycles of the battery pack in this section. The main components of the fusion model are

also explicitly described: exponential (EXP) model and LSTM model.

(1) EXP model

In this section, the battery pack HI degradation model is established based on the battery pack full-life HIs through the double exponential empirical formula. The early battery pack HIs are used to realize the correction of the battery cell HI degradation model. A degradation model of battery cell HIs is established to recursively predict future cycle HIs, based on a data-driven approach to achieve the prediction of battery pack health.

By analyzing the aging characteristics of the battery, it is seen that the exponential model is closely related to the decay characteristics of the extracted HIs and can accurately track the global aging trend of the battery HIs, and the exponential model is shown in Equation (4.5):

$$y = ae^{bx} + ce^{dx} \tag{1.5}$$

where a, b, c, and d denote the model parameters to be determined, x is the number of battery cycles, and y is the HI.

Taking Cell #4 as an example, the HI EXP degradation model is constructed, and it includes two steps. Firstly, the model parameters are initialized as para0 = [a, b, c, d] = [0.02, −0.00701, 0.968, −0.9155×10^{-5}]. Then, the updated model parameters (para1 = [a, b, c, d] = [0.024, 2.7219×10^{-4}, 1.0416, −8.4452×10^{-5}]) are obtained using least-squares identification to track the real degradation curve of the cell.

(2) LSTM model

LSTM is a kind of specialized recurrent neural network (RNN) for avoiding gradient vanishing and exploding problems [29]. The network structure of LSTM, as shown in Figure 8, mainly consists of three gates and two memory states [6] (e.g., forgetting gate f_t, input gate i_t, output gate O_t, long memory C_t, and short memory h_t).

The forgetting gate f_t, which is used to calculate the degree of forgetting of the information, is processed by the sigmoid function and takes a value between zero and one, where one means all retained and zero means all forgotten. The input gate it, used to calculate the information saved to the state unit, consists of two parts, it as the amount of current input information to save to the unit state, and $\tilde{c}_t$ as the new information generated by the current input to add to the unit state, both of which generate a new memory state. As a result, the current moment of the unit state consists of the product of the forgetting gate input and the previous moment state plus the product of the two parts of the input gate, that is, C_t. The output gate O_t is used to calculate the extent to which the information is output at the current moment. In the previous hidden state h_{t-1}, the current input x_t is passed to the sigmoid function. The updated cell state is passed to the *tanh* function, and the *tanh* function is multiplied with the sigmoid function output to determine the information that the hidden state should carry and what the hidden state should use as the output. Then, the new cell state and the new hidden state are transferred to the next time step, and the corresponding expression of the LSTM structure can be shown by Equation (4.6) [30–33]:
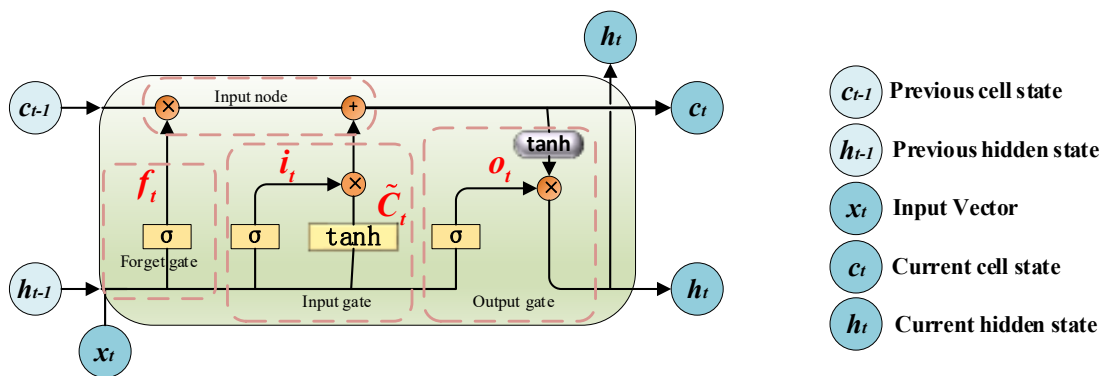

Figure 4.8. The network structure and operations LSTM.

$$\begin{cases} f_t = \sigma\left(w_f \cdot [h_{t-1}, x_t]\right) + b_f \\ i_t = \sigma\left(w_i \cdot [h_{t-1}, x_t]\right) + b_i \\ \tilde{C}_t = \tanh\left(w_c \cdot [h_{t-1}, x_t]\right) + b_c \\ C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \\ o_t = \sigma\left(w_o \cdot [h_{t-1}, x_t]\right) + b_o \\ h_t = o_t * \tanh\left(C_t\right) \end{cases} \qquad (4.6)$$

where σ and *tanh* are the sigmoid and hyperbolic activation functions, respectively, *w* and *b* are different weight matrices and bias matrices, respectively. C_{t-1} and C_t are the previous moment cell state and current cell state, respectively, and C_t is the updated cell state. h_{t-1} and h_t, are the previous hidden state and the current hidden state. C_{t-1}, h_{t-1}, and x_t are the inputs of LSTM, while C_t and h_t are the outputs.

Based on the above introduction of the three gates and two memory states in the LSTM network, the HI degradation model is established as follows. Firstly, the battery cell HIs and pack HIs as inputs to the LSTM network are rearranged in the data format. Secondly, the neural network (including the LSTM network layer and the fully connected layer) is constructed, the input single-unit HIs are trained for the network, and the trained network parameters are drawn from the battery single-unit HI degradation model. Finally, the long short-term memory network is frozen, and the early cyclic HIs of separate cells in the battery pack are input into the neural network. The fully connected layer in the network is trained again, a new fully connect layer is trained, and the new neural network is the LSTM degradation model of the HIs of every single cell in the battery pack.

(3) Fusion degradation model of HIs

Since the HI EXP degradation model only captures the global aging characteristic of the battery HIs but ignores the influence of local fluctuations on the results, it will lead to the loss of crucial information. The LSTM degradation model has the advantage of capturing local changes. However, with continuous recursion, the errors gradually

accumulate, leading to too many errors to meet the accuracy requirements. To solve the problem of simultaneously capturing the global aging characteristic and local fluctuations of the HIs, a fusion degradation model in which different weights combine the exponential degradation model and the LSTM degradation model is proposed. The fusion HI degradation model (FHDM) as show by Equation (4.7):

$$FHDM = \alpha * LSTM + (1-\alpha) * EXP \tag{4.7}$$

where a is the weight of the LSTM degradation model, taking values between 0 and 1. LSTM is the long short-term memory neural network degradation model, and EXP is the exponential degradation model.

4.4.3 GPR theory

The Gaussian process regression (GPR) algorithm is a machine learning method based on Bayesian theory. The advantages of being flexible, nonparametric, and able to integrate uncertainty expressions are widely valued in battery health prediction. This section provides a brief introduction to the GPR algorithm [34].

A formula for any GPR problem as shown in Equation (4.8):

$$y = f(x) + \varepsilon \quad \varepsilon \sim N(0, \sigma^2) \tag{4.8}$$

where y is the observed value containing the noise, ε is a Gaussian white noise that satisfies the mean value of zero, and $f(x)$ is a function that obeys the Gaussian probability distribution, as shown in Equation (4.9):

$$f(x) \sim GP(m(x), k_f(x, x')) \tag{4.9}$$

where $m(x)$ is the mean function, usually taken to zero, k_f is the kernel function, used to characterize the distance or similarity between the two points of the input quantity, and the commonly used covariance function is the squared exponential (SE) covariance function, as shown in Equation (4.10) [35]:

$$K_f = \sigma_f^2 \exp(-\frac{1}{2}(x-x')l^{-1}(x-x')) \qquad (4.10)$$

where the signal covariance σ_f^2 is the output amplitude and l is the characteristic length scale.

For the observed value y that obeys a Gaussian distribution and considers Gaussian white noise, the prior distribution is as shown in Equation (4.11):

$$y \sim N(0, K_f(X,X) + \sigma_n^2 I_n) \qquad (4.12)$$

The set of parameters $\Theta = [\sigma_f, l, \sigma_n]$ in Equation (4.10) is the hyper-parameter, and the optimal solution of the hyper-parameter is obtained by establishing the negative logarithmic marginal likelihood function (NLML) to find the partial derivatives of the hyper-parameter, and then using the conjugate gradient method to minimize the partial derivatives, as shown in Equation (4.12):

$$L = logp(y|X,\Theta) = -\frac{1}{2}y^T[K_f(X,X)+\sigma_n^2 I_n]^{-1}y - \frac{1}{2}log(det(K_f(X,X)+\sigma_n^2 I_n)) - \frac{nlog2\pi}{2} \qquad (4.12)$$

The joint Gaussian distribution of the observed value y and the predicted value y* as shown in Equation (4.13):

$$\begin{bmatrix} y \\ y^* \end{bmatrix} \sim N\left(0, \begin{bmatrix} (K_f(X,X)+\sigma_n^2 I_n) & K_f(X,X^*) \\ K_f(X,X^*)^T & K_f(X^*,X^*) \end{bmatrix}\right) \qquad (4.33)$$

According to Bayesian theory, the posterior distribution as show in Equation (4.14):

$$p\left(y^*|X,y,X^*\right) = N\left(y^*|\bar{y}^*, \sigma^2(y^*)\right) \qquad (4.14)$$

where $\bar{y}^*$ is the predicted mean, and $\sigma^2(y^*)$ is the predicted covariance, as given in Equation (4.15):

$$\begin{cases} \bar{y}^* = K_f(X,X^*)^T[K_f(X,X)+\sigma_n^2 I_n]^{-1}y \\ \sigma^2(y^*) = K_f(X^*,X^*) - K_f(X,X^*)^T[K_f(X,X)+\sigma_n^2 I_n]^{-1}K_f(X,X^*)) \end{cases} \qquad (4.15)$$

To quantitatively evaluate the actual effect of the proposed scheme, three indicators—the 95% confidence interval (CI), the mean absolute error (MAE), and the root mean square error (RMSE)—are used to evaluate the prediction performance of the GPR algorithm, as shown in Equation (4.16):

$$\left\{ \begin{array}{l} 95\%CI = \bar{y}^* \mp 1.96 \times \sigma(y^*) \\ RMSE = \sqrt{\dfrac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y}_i^*)^2} \\ MAE = \dfrac{1}{N}\sum_{i=1}^{N}|y_i - \bar{y}_i^*| \end{array} \right. \tag{4.16}$$

where 95%CI is the confidence interval, $\bar{y}^*$ and y^* are the predicted value and variance, respectively. N is the size of the testing set, y_i is the actual value, and $\bar{y}_i^*$ is the estimated value.

4.5. Results and discussion

In this section, the experimental aging data are used to evaluate battery pack health prognosis methods. The single degradation and fused degradation models are validated based on the HIs by using the experimental data of battery aging under the same and different operating conditions in section 4.5.1. The effect of the battery pack's health prediction model is verified by the future cycling the HIs predicted by the single degradation model and the fused degradation model, and the prediction results are quantitatively analyzed and discussed in section 4.5.2.

4.5.1 HIs prediction

This section describes two types of HI degradation models: the single degradation models (EXP degradation model and LSTM degradation model) and the fusion degradation model. Firstly, the prediction accuracy of the single degradation model is

verified by using three sets of HIs under the same and different operating conditions, and the applicability range of each single degradation model is compared. Secondly, the prediction accuracy of the fusion degradation model is verified using the HIs under the same condition and the different operating conditions, and the results are analyzed and discussed accordingly.

(1) Single degradation model

The single degradation model includes the EXP degradation model and the LSTM degradation model. For the EXP degradation model under the same operating condition (charge and discharge current is 0.5C and temperature is 35°C), the corresponding decay curves before and after the correction of the HI degradation model for Cell #4 under the same operating condition are shown in Figure 4.9. Figure 4.9(a1), (b1), (c1) shows the decay curves of the three HIs of Cell #4 under the same operating condition (before correction), where the blue line is the actual value and the red line is the estimated value of the EXP degradation model based on the least-squares method. Using the battery pack early 10% HIs to correct the HI EXP degradation model, the model of the battery pack corresponding to the three sets of HI decay curves (after correction) are shown in Figure 4.9(a2), (b2), (c2), where the dotted line is the actual value and the solid line is the fitted value. The MAE and RMSE between the actual values and the predicted values for the EXP degradation model of the battery pack under the same operating condition are 0.034, 0.021, 0.006, and 0.2745, 0.2746, 0.3191, respectively. The aging characteristic of individual HIs in the battery pack is in good agreement with the global aging characteristic of the predicted values, indicating that the EXP degradation model can capture the global HI decay tendency with high accuracy, but cannot capture the local HI variation trend.
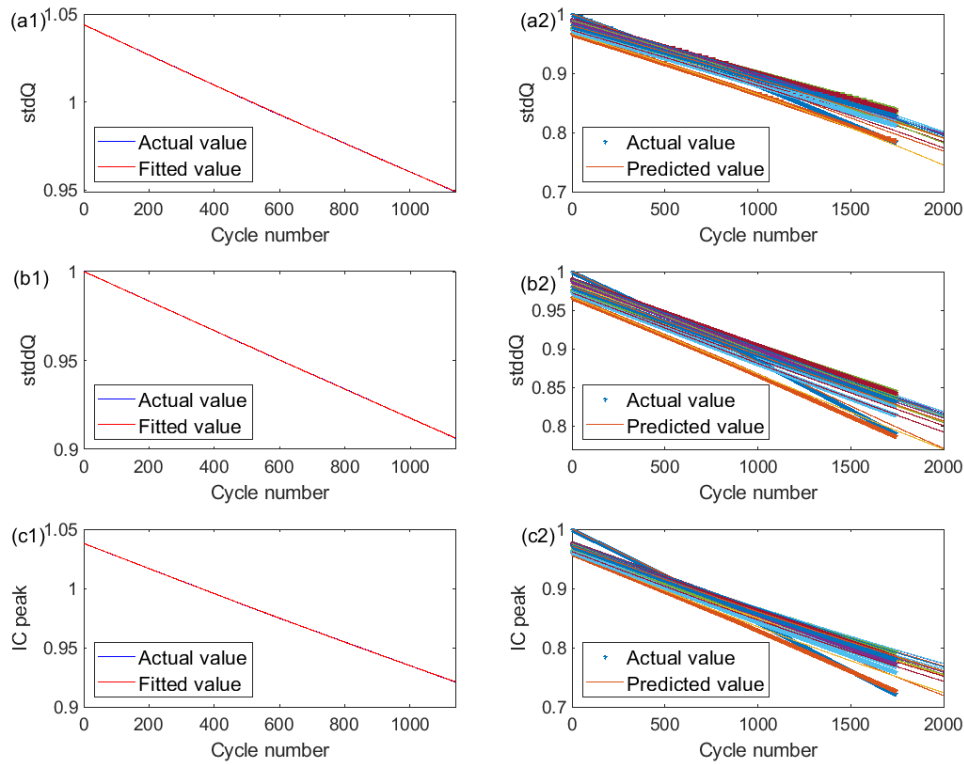
Figure 4.9. Decay curve of the HI EXP degradation model before and after 10% pack correction. Before correction: (a1) *stdQ*; (b1) *stddQ*; (c1) *IC peak*. After correction: (a2) *stdQ*; (b2) *stddQ*; (c2) *IC peak*.

For the LSTM degradation model under the same operating condition, the three sets of the HI decay curves corresponding to Cell #4 with the fitted curves are shown in Figure 4.10(a1), (b1), (c1) (before correction), where the blue line is the actual value and the red line is the predicted value of the LSTM degradation model. Figure 4.10(a2), (b2), (c2) shows the decay curves of the three HIs corresponding to the battery pack with the predicted values (after correction). Due to the gradual accumulation of errors in the recursive rolling process of the LSTM model, although the deviation of the predicted values from the actual values is slight for the first 900 cycles, the errors increase rapidly from the 901st cycle to the 1742nd cycle as the cycles continue to accumulate. Then, the LSTM degradation model experiences the error accumulation effect under the same operating condition, and the error is too large to correct at the later stage. The MAE and RMSE between the actual values and the predicted values

for the LSTM degradation model of the battery pack under the same operating condition
are 0.0523, 0.0401, 0.11 and 0.1624, 0.1933, 0.0939, respectively. The results show
that the local aging characteristics of the pack HIs are in good agreement with the
predicted values, indicating that the LSTM degradation model can capture the ageing
characteristic of local HIs. Nevertheless, the ability to capture the aging characteristic
of global HIs is relatively poor due to the disadvantages of cumulative errors.

Figure 4.10. Decay curves of the LSTM degradation model with HIs before and after 10% pack
correction.

Before correction: (a1) *stdQ*; (b1) *stddQ*; (c1) *IC peak*. After correction: (a2) *stdQ*; (b2) *stddQ*;
(c2) *IC peak*.

For the LSTM degradation model under the different operating conditions, the
corresponding three sets of HI aging curves and fitted curves (before correction) for six
groups of battery cells under different operating conditions are shown in Figure
4.11(a1), (b1), (c1), where the blue line is the actual value. The red line is the predicted
value of the LSTM degradation model. The three sets of HI decay curves of the battery

pack are shown in Figure 4.11(a2), (b2), (c2), where the dotted line is the actual decay curve and the solid line is the LSTM degradation model prediction curve (after correction). Compared with the same condition decay curves in Figure 4.10(a2), (b2), (c2), the corresponding prediction accuracy of the LSTM degradation model prediction values under the different operating conditions is improved, where the improvement of *stddQ* and *IC peak* is prominent. The MAE and RMSE between actual values and the predicted values for the LSTM degradation model of the battery pack under the different operating conditions are 0.0203, 0.128, 0.197 and 0.2417, 0.2543, 0.2855, respectively. By comparing this with the LSTM degradation model under the same operating condition, the result shows that the consistent local and global aging characteristics of the pack HI aging trends with the predicted values are significantly improved. The results indicate that the type, quantity, and quality of the battery aging data under the different operating conditions play crucial roles in enhancing the ability of the LSTM degradation model to capture the local and global HI aging trends.
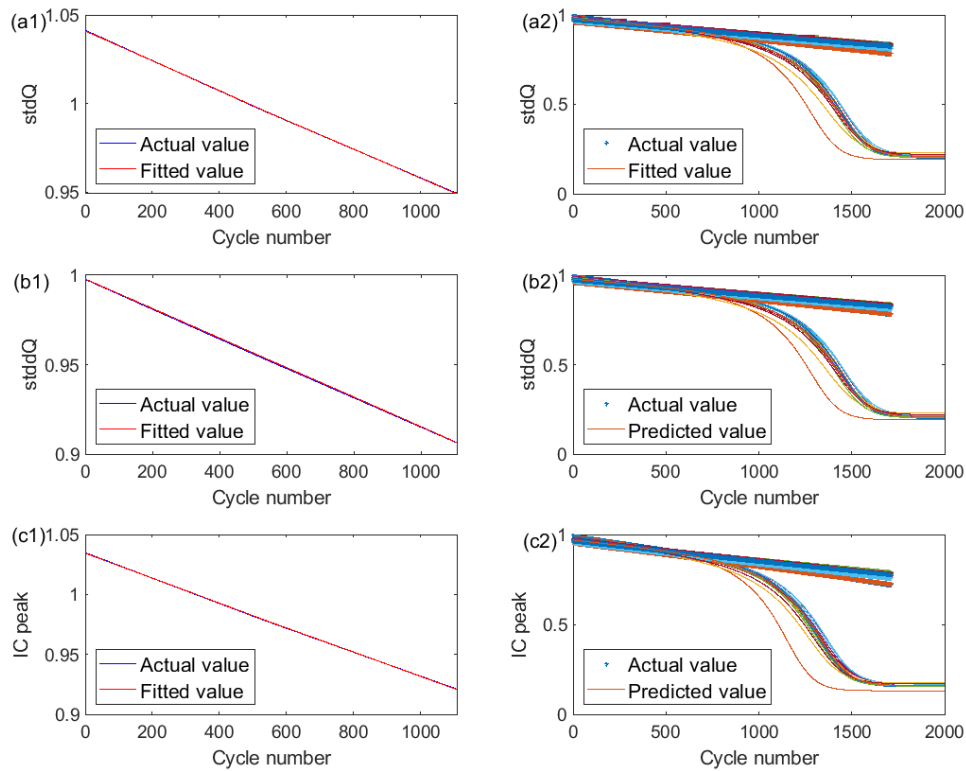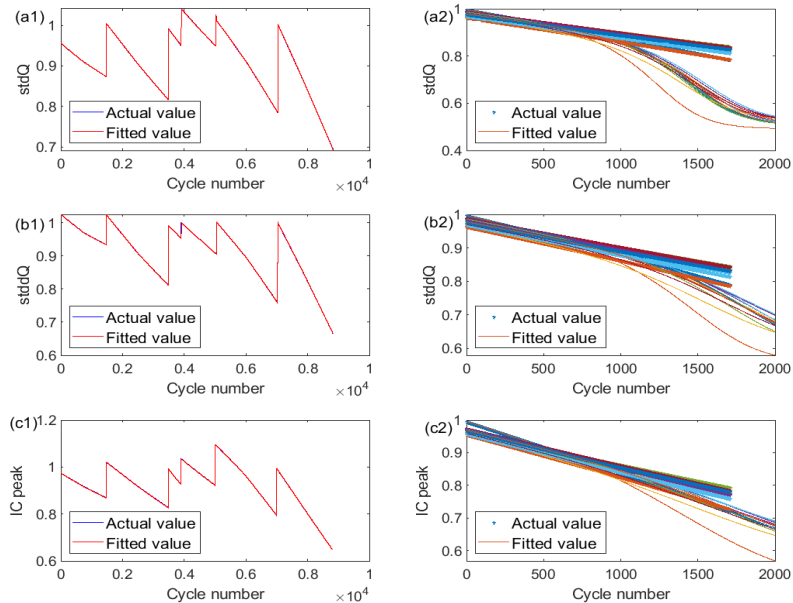

Figure 4.11. Decay curves of the LSTM model of the HIs before and after 10% pack correction.

Before correction: (a1) *stdQ*; (b1) *stddQ*; (c1) *IC peak*. After correction: (a2) *stdQ*; (b2) *stddQ*;

(c2) *IC peak*.

119

(2) Fusion degradation model

The single HI EXP degradation model cannot capture the local variation of the HI decay characteristics. In contrast, the single HI LSTM degradation model cannot capture the global variation of the HI decay characteristics. The battery aging data of the same and different operating conditions are used here to verify the proposed fusion degradation model to solve the above shortcomings.

The actual decay curves and the predicted decay curves are based on the fusion degradation model for the three HIs of the battery pack under the same operating condition, as shown in Figure 4.12(a), (b), (c). The dotted lines are the actual values of the HIs, and the solid lines are the predicted values of the HIs. The MAE and RMSE of the three HIs of the battery pack are 0.0132, 0.0097, 0.0215 and 0.256, 0.2604, 0.2823, respectively. The overall accuracy of MAE and RMSE of the fusion degradation model is better than the single HI EXP degradation model. However, the predicted value of *IC peak* is much lower than the single HI EXP degradation model. The MAE of the fusion degradation model is better than the LSTM degradation model. The RMSE is between the predicted values of the single EXP degradation model and the predicted values of the single LSTM degradation model.

The decay curves of the three sets of HIs and the predicted decay curves based on the fusion model are shown in Figure 4.13(a), (b), (c), where the dotted lines are the actual values of the HIs and the solid lines are the predicted values of the HIs under the different operating conditions. The MAE and RMSE of the fusion degradation model for the three sets of HIs are 0.0068, 0.042, 0.0034 and 0.2682, 0.2707, 0.3126, respectively. The results show that the fusion degradation model under the different operating conditions has better accuracy and ability to capture the HIs than the single HI EXP degradation model and single LSTM degradation model.
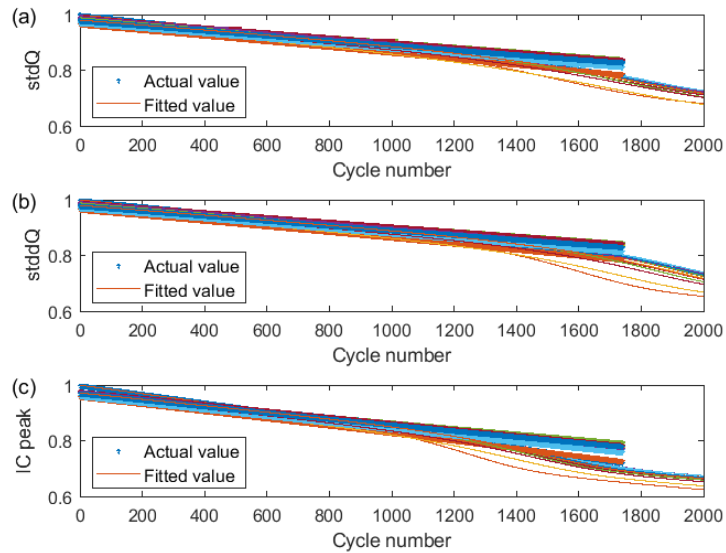
Figure 4.12. Decay curves of HI fusion model after 10% pack correction (weight coefficient of

LSTM degradation model is 0.2) (same operating condition).

(a) *stdQ*; (b) *stddQ*; (c) *IC peak*.
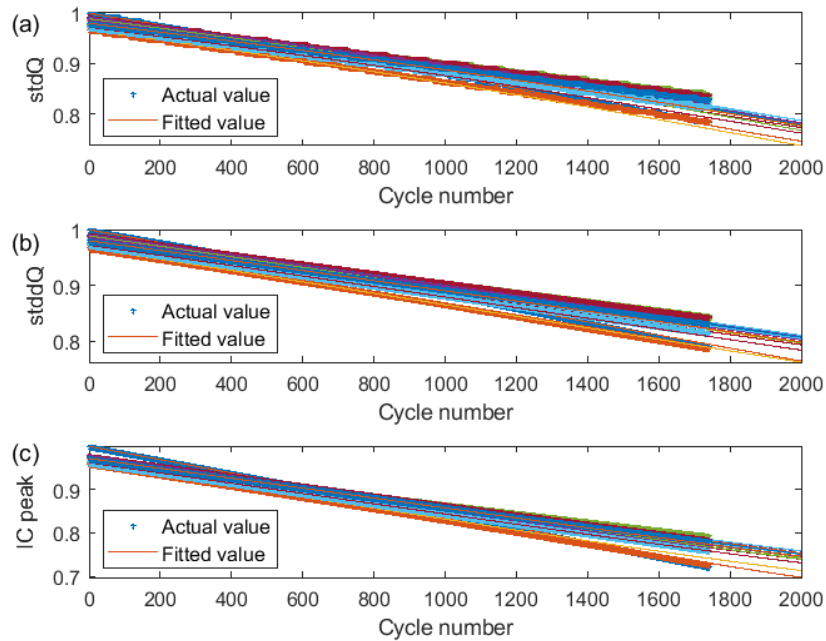

Figure 4.13. Decay curves of HI fusion model after 10% pack correction (different operating

conditions).

(a) *stdQ*; (b) *stddQ*; (c) *IC peak*.

The prediction accuracies correspond to the single degradation model and the fused

degradation model under the different operating conditions, as shown in Figure 4.14.

Figure 4.14(a) represents the mean absolute error (MAE) corresponding to the three HIs under different degradation models. By comparing the MAE under the three degradation models, the results show that the prediction accuracy of the fusion degradation model is higher than that of the exponential degradation model and the LSTM degradation model, in which the *stdQ* and *stddQ* accuracy of the single degradation models (EXP degradation model and LSTM degradation model) are substantially improved. Figure 4.14(b) represents the root mean square error (RMSE) corresponding to the three HIs under different degradation models. The RMSE corresponding to the three sets of HIs of the fusion degradation model has errors between the exponential degradation model and the LSTM degradation model. The results show that the fusion degradation model can capture the global aging trend of the HIs in the EXP degradation model and eliminate the shortcoming of excessive cumulative errors in the recursive process of the LSTM degradation model, and thus has satisfactory estimation accuracy.
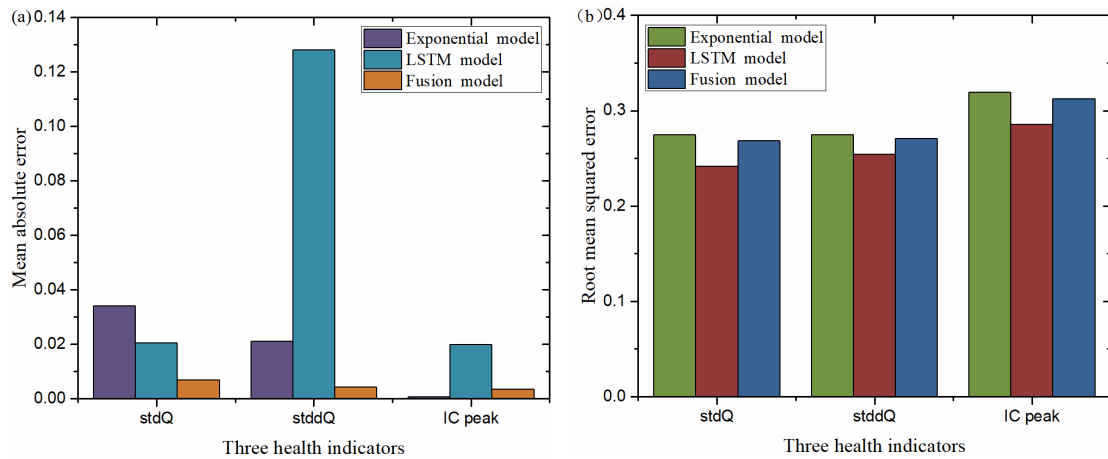

Figure 4.14. Error statistics of three different HI degradation models.

(a) three HIs MAE statistic; (b) three HIs RMSE statistic.

4.5.2 Battery pack SOH prediction

In this section, the prediction accuracy of the battery pack health prediction method is verified based on the predicted value of the HI fusion decay model under the same operating condition and the different operating conditions, combined with the battery pack capacity estimation model. Among them, the same operating condition means that both battery cells and the battery pack work at 35 ℃ _0.5C-0.5C; different operating conditions mean that battery cells work in six different environments (see Table 4.1 for details) and battery pack works at 35 ℃ _0.5C-0.5C.

(1) Prediction results using the same operating condition

Under the same operating condition, the battery pack cycle life early 10% HIs and capacity are used as the training set of the battery pack capacity estimation model. On the basis of the EXP degradation model, the LSTM degradation model, and the fusion degradation model, used to predict the future HIs of the battery pack, the predicted battery pack state of health is analyzed against the actual value to calculate the error. Figure 4.15 shows the corresponding battery pack state of health prediction results under the three groups of models: the HI exponential degradation model-GPR capacity prediction model (EXP-GPR), the HI exponential degradation model-HI LSTM degradation model-GPR capacity prediction model (EXP-LSTM-GPR), and the HI LSTM degradation model-GPR capacity prediction model (LSTM-GPR). Figure 4.15(a1), (b1), (c1) shows the prediction error of the battery pack for each of the three cases. The solid red line indicates the battery state of the health cutoff threshold. The red dotted line is the actual value, the solid blue line is the estimated value, and the green shaded area is the 95% confidence interval used to describe the uncertainty of the prediction result. Figure 4.15(a2), (b2), (c2) shows the relative errors under the three groups of models EXP-GPR, EXP-LST M-GPR, and LSTM-GPR, respectively.

Among them, Figure 4.15(a2) shows that the relative error corresponding to the first 406 cycles is negative, while the relative error from the 407th cycle to the 850th cycle is positive and exponentially increasing. Then, the error increases rapidly from the 851st cycle to the 1742nd cycle, and the maximum relative error is only 1.7%. Figure 4.15(b2) indicates that the relative errors of the fusion model are all negative when the weights of the single degradation model are 0.5. The relative error is less than −0.4% for the first 600 cycles, and the relative error shows exponential growth from the 601st cycle to the 1200th cycle. The error increases linearly and rapidly from the 1201st cycle, with the maximum relative error reaching −15%. Comparing Figure 4.15c with Figure 4.15b, the result shows that the two relative errors have the same trend, but the maximum relative error of the latter is about twice as large as that of the former.
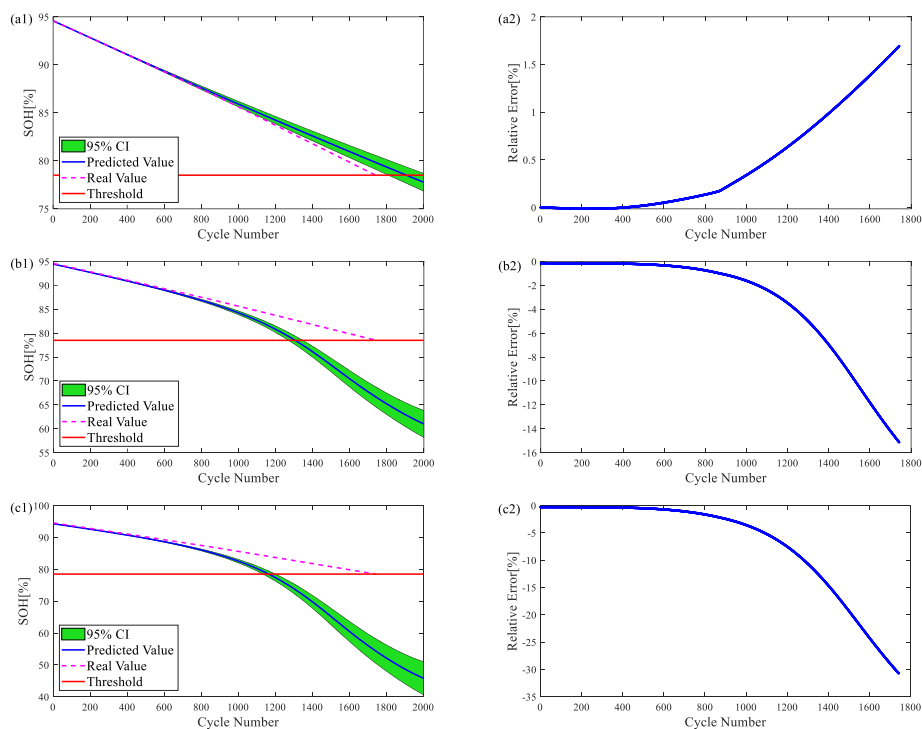

Figure 4.15. Battery pack SOH prediction results and relative error (same operating condition). EXP-GPR prediction model: (a1) SOH prediction results; (a2) relative error. EXP-LSTM-GPR prediction model: (b1) SOH prediction results; (b2) relative error. LSTM-GPR prediction model: (c1) SOH prediction results; (c2) relative error.

Under the same operating condition, the EXP degradation model and LSTM degradation model are fused based on different weights, and the HI fusion degradation model thus predicts the future cycle HIs of the battery pack, which is combined with the battery pack capacity estimation model to achieve a state of health prediction. The errors under the different weights are shown in Figure 4.16, where the horizontal coordinates indicate the weight of the LSTM degradation model in the fusion model and the vertical coordinates indicate the corresponding errors under different weights. The results show that the fusion degradation model can track the capacity regeneration phenomenon with different degrees of accuracy due to its inability to cover the range of parameter variation within the battery pack. The prediction accuracy based on the battery pack health state prediction model needs to be optimized.
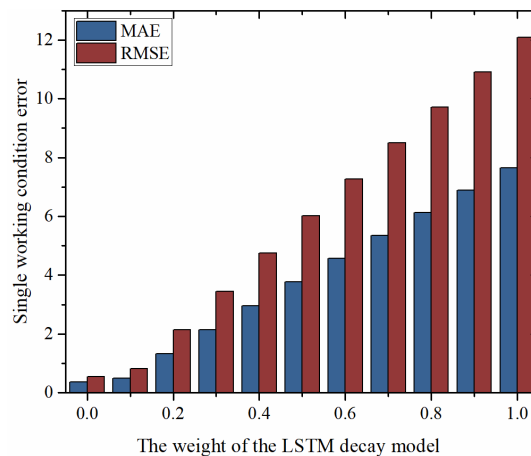
Figure 4.16. Error statistics of fusion degradation model with different weights (same operating condition).

(2) Prediction results using the different operating conditions

Under the different operating conditions, using the HI degradation model to predict the future HIs, the prediction results will verify the prediction accuracy of the battery pack capacity estimation model and the corresponding prediction results and relative errors, as shown in Figure 4.17. Figure 4.17(a) shows the fusion degradation model under the different operating conditions and the LSTM degradation model with a

weight coefficient of zero predicted for the battery pack. The relative error of the early cycle is negative, the relative error of the mid-cycles is positive and shows an exponential growth trend, and the relative error of the later cycle increases rapidly, while the maximum relative error is only 1.7%. Figure 4.17c shows that the predicted value is consistently smaller than the actual value as the number of cycles increases. The relative error is less than −0.5% and remains stable for the first 600 cycles and increases linearly and rapidly from the 601st cycle to the 1742nd cycle, with the maximum relative error being less than -4%. Compared with Figures 4.17a and c, Figure 4.17(b) shows the prediction HIs of the fusion model when the weights of the single degradation model are 0.5, and the SOH prediction accuracy based on the fusion degradation model is higher than the others. The first 880 cycles are stable within −0.35%, and the error gradually increases after the 880th cycle, with a maximum relative error of 0.87%.
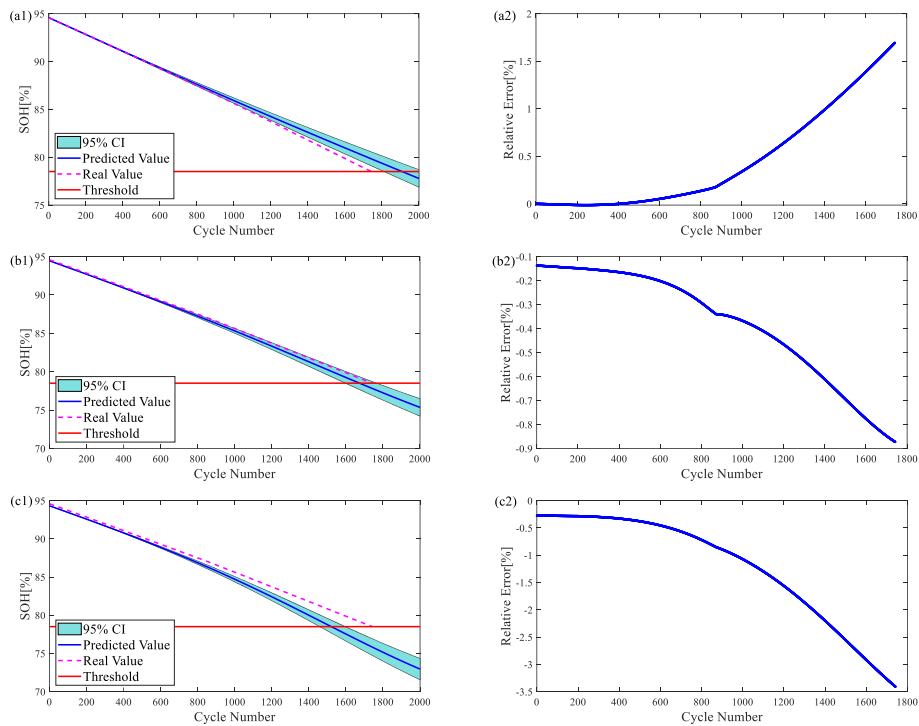
Figure 4.17. Battery pack SOH prediction results and relative error (different operating conditions).

EXP-GPR prediction model: (a1) SOH prediction results; (a2) relative error. EXP-LSTM-GPR prediction model: (b1) SOH prediction results; (b2) relative error. LSTM-GPR prediction model: (c1) SOH prediction results; (c2) relative error.

Under the different operating conditions, the EXP degradation model and the LSTM degradation model are constructed as fusion models with different weights to predict the future cycle HIs of the battery pack, and the battery pack capacity estimation model is validated based on the predicted HIs. The errors correspond to the different weights of the fusion degradation model, as shown in Figure 4.18, where the horizontal coordinates indicate the weights of the LSTM degradation model in the fusion degradation model, and the vertical coordinates indicate the errors corresponding to the different weights. With the increasing weight of the LSTM model, both MAE and RMSE show a 'V' shape, and both errors of MAE and RMSE reach the maximum when the LSTM degradation model is adopted separately. The highest prediction accuracy is obtained by fusing the EXP degradation model and the LSTM degradation model with 0.7 and 0.3 weights under the different operating conditions, and the corresponding MAE and RMSE are 7.17% and 7.81%, respectively. The results show that the fused degradation model can achieve satisfactory prediction accuracy when combined with the battery pack capacity estimation model for SOH prediction.
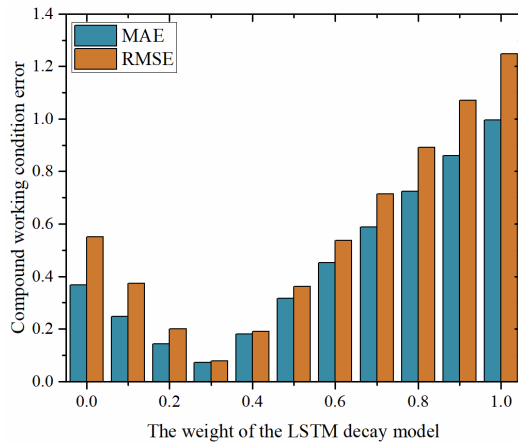
Figure 4.18. Statistical errors corresponding to different weights of the fusion degradation model (different operating conditions).

4.6 Chapter summary

This chapter proposes a new method to predict the health of a battery pack based on its early cycling data and the complete aging data of battery cells. Firstly, three sets of HIs are extracted from the experimental data of battery cells and the battery pack. The correlations between the HIs and capacity are verified by the Pearson correlation analysis method. The results show that their correlation coefficients are greater than 0.99, indicating the HIs are highly related to the battery capacity. To predict the future HIs of battery cells, an exponential degradation model is fitted to capture the global decay of HIs, while a LSTM degradation model is constructed to imitate the local variation of HIs. A fusion degradation model can be created by weighting the exponential-based model and LSTM-based model, which can inherit the advantages of these two models. Then, an early prognosis method for battery pack health is proposed. Based on the early cycling data of the battery pack and the fusion degradation model of HIs, the future HIs of the battery pack can be obtained. Taking the HIs as the inputs of the GPR algorithm, data-driven models can be constructed to predict the future health of the battery pack. Finally, three health prediction models based on the GPR algorithm

are constructed for comparison, including the exponential function-based (EXP-GPR) model, LSTM-based (LSTM-GPR) model, and their weighted fusion-based (EXP-LSTM-GPR) model. A comparison of the errors of the three different models for two different operating conditions is shown in Table 4.2. The results show that the fused degradation model has better accuracy under the different operating conditions, and its MAE and RMSE are 7.17% and 7.81%, respectively.

Table 4.2. Comparison of the errors of different models under different working conditions.

Operating Conditions	Accuracy	Model		
		EXP-GPR	LSTM-GPR	EXP-LSTM-GPR
Same operating	MAE	0.3681	7.6437	2.1477
conditions	RMSE	0.5517	12.0944	3.4502
Different operating	MAE	0.3681	0.9951	0.0717
conditions	RMSE	0.5517	1.2476	0.0781

The method proposed in this paper is of excellent engineering utility for the rapid development of battery packs and for evaluating their performance indexes. It can save more than 50% of the aging experiment time and labor costs, make full use of the existing test data to predict the life of unknown battery cells and packs, and facilitate the development and selection cycle. The whole-lifecycle prediction is completed by using the data of a small number of cycles, which can improve the accuracy of the life prediction of battery cells and battery packs and broaden the application scope to solve the problem of life prediction under different temperatures, different currents, and different cell models. However, the proposed scheme in this paper only validates the same type of battery cells and battery packs. This method can be applied to different types of batteries and capacity diving problems within the scope of principle analysis, but further validation is needed, which is also our next problem to solve.

References

[1] Hu, X.; Xu, L.; Lin, X.; Pecht, M. Battery Lifetime Prognostics. *Joule* **2020**, *4*, 310–346.

[2] Severson, K.A.; Attia, P.M.; Jin, N.; Perkins, N.; Jiang, B.; Yang, Z.; Chen, M.H.; Aykol, M.; Herring, P.K.; Fraggedakis, D.; et al. Data-driven prediction of battery cycle life before capacity degradation. *Nat. Energy* **2019**, *4*, 383–391.

[3] Deng, Z.; Hu, X.; Lin, X.; Kim, Y.; Li, J. Sensitivity Analysis and Joint Estimation of Parameters and States for All-Solid-State Batteries. *IEEE Trans. Transp. Electrif.* **2021**, *7*, 1314–1323.

[4] Zhang, Y.; Wang, C.-Y.; Tang, X. Cycling degradation of an automotive LiFePO$_4$ lithium-ion battery. *J. Power Sources* **2011**, *196*, 1513–1520.

[5] Shu, X.; Shen, S.; Shen, J.; Zhang, Y.; Li, G.; Chen, Z.; Liu, Y. State of health prediction of lithium-ion batteries based on machine learning: Advances and perspectives. *iScience* **2021**, *24*, 103265.

[6] Li, W.; Sengupta, N.; Dechent, P.; Howey, D.; Annaswamy, A.; Sauer, D.U. Online capacity estimation of lithium-ion batteries with deep long short-term memory networks. *J. Power Sources* **2021**, *482*, 228863.

[7] Yang, S.; Zhang, C.; Jiang, J.; Zhang, W.; Zhang, L.; Wang, Y. Review on state-of-health of lithium-ion batteries: Characterizations, estimations and applications. *J. Clean. Prod.* **2021**, *314*, 128015.

[8] Li, Y.; Liu, K.; Foley, A.; Zülke, A.; Berecibar, M.; Nanini-Maury, E.; Van Mierlo, J.; Hoster, H.E. Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review. *Renew. Sustain. Energy Rev.* **2019**, *113*, 109254.

[9] Ge, M.-F.; Liu, Y.; Jiang, X.; Liu, J. A review on state of health estimations and remaining useful life prognostics of lithium-ion batteries. *Measurement* **2021**, *174*, 109057.

[10] Schmalstieg, J.; Käbitz, S.; Ecker, M.; Sauer, D.U. A holistic aging model for Li (NiMnCo) O_2 based 18650 lithium-ion batteries. *J. Power Sources* **2014**, *257*, 325–334.

[11] Hu, X.; Li, S.; Peng, H. A comparative study of equivalent circuit models for Li-ion batteries. *J. Power Sources* **2012**, *198*, 359–367.

[12] Lai, X.; Gao, W.; Zheng, Y.; Ouyang, M.; Li, J.; Han, X.; Zhou, L. A comparative study of global optimization methods for parameter identification of different equivalent circuit models for Li-ion batteries. *Electrochim. Acta* **2019**, *295*, 1057–1066.

[13] Hu, X.; Yuan, H.; Zou, C.; Li, Z.; Zhang, L. Co-Estimation of State of Charge and State of Health for Lithium-Ion Batteries Based on Fractional-Order Calculus. *IEEE Trans. Veh. Technol.* **2018**, *67*, 10319–10329.

[14] Jokar, A.; Rajabloo, B.; Désilets, M.; Lacroix, M. Review of simplified Pseudo-two-Dimensional models of lithium-ion batteries. *J. Power Sources* **2016**, *327*, 44–55.

[15] Galeotti, M.; Cinà, L.; Giammanco, C.; Cordiner, S.; Di Carlo, A. Performance analysis and SOH (state of health) evaluation of lithium polymer batteries through electrochemical impedance spectroscopy. *Energy* **2015**, *89*, 678–686.

[16] Xu, L.; Lin, X.; Xie, Y.; Hu, X. Enabling high-fidelity electrochemical P2D modeling of lithium-ion batteries via fast and non-destructive parameter identification. *Energy Storage Mater.* **2021**, *45*, 952–968.

[17] Lipu, M.S.H.; Hannan, M.; Hussain, A.; Hoque, M.; Ker, P.J.; Saad, M.; Ayob, A. 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.* **2018**, *205*, 115–133.

[18] Klass, V.; Behm, M.; Lindbergh, G. A support vector machine-based state-of-health estimation method for lithium-ion batteries under electric vehicle operation. *J. Power Sources* **2014**, *270*, 262–272.

[19] Li, H.; Pan, D.; Chen, C.L.P. Intelligent Prognostics for Battery Health Monitoring Using the Mean Entropy and Relevance Vector Machine. *IEEE Trans. Syst. Man Cybern. Syst.* **2014**, *44*, 851–862.

[20] Li, X.; Wang, Z.; Yan, J. Prognostic health condition for lithium battery using the partial incremental capacity and Gaussian process regression. *J. Power Sources* **2019**, *421*, 56–67.

[21] Che, Y.; Deng, Z.; Hu, X. Battery pack state of health estimation with general health indicators and modified gaussian process regression. In Proceedings of the 34th International Electric Vehicle Symposium and Exhibition (EVS34), Nanjing, China, 25–28 June 2021.

[22] Deng, Z.; Hu, X.; Li, P.; Lin, X.; Bian, X. Data-Driven Battery State of Health Estimation Based on Random Partial Charging Data. *IEEE Trans. Power Electron.* **2021**, *37*, 5021–5031.

[23] Hu, X.; Che, Y.; Lin, X.; Deng, Z. Health Prognosis for Electric Vehicle Battery Packs: A Data-Driven Approach. *IEEE/ASME Trans. Mechatron.* **2020**, *25*, 2622–2632.

[24] Deng, Z.; Hu, X.; Lin, X.; Xu, L.; Che, Y.; Hu, L. General Discharge Voltage Information Enabled Health Evaluation for Lithium-Ion Batteries. *IEEE/ASME Trans. Mechatron.* **2020**, *26*, 1295–1306.

[25] Che, Y.; Deng, Z.; Li, P.; Tang, X.; Khosravinia, K.; Lin, X.; Hu, X. State of health prognostics for series battery packs: A universal deep learning method. *Energy* **2021**, *238*, 121857.

[26] Yunhong, C.; Zhongwei, D.; Tang, X.; Xianke, L.; Xianghong, N.; Hu, X. Lifetime and Aging Degradation Prognostics for Lithium-ion Battery Packs Based on a Cell to Pack Method. *Chin. J. Mech. Eng.* **2022**, *35*, 1–16.

[27] Weng, C.; Feng, X.; Sun, J.; Peng, H. State-of-health monitoring of lithium-ion battery modules and packs via incremental capacity peak tracking. *Appl. Energy* **2016**, *180*, 360–368.

[28] Zhou, Y.; Huang, M.; Chen, Y.; Tao, Y. A novel health indicator for on-line lithium-ion batteries remaining useful life prediction. *J. Power Sources* **2016**, *321*, 1–10.

[29] Wu, Y.; Xue, Q.; Shen, J.; Lei, Z.; Chen, Z.; Liu, Y. State of Health Estimation for Lithium-Ion Batteries Based on Healthy Features and Long Short-Term Memory. *IEEE Access* **2020**, *8*, 28533–28547.

[30] Mitrovic, T.; Xue, B.; Li, X. (Eds.) *AI 2018: Advances in Artificial Intelligence: 31st Australasian Joint Conference, Wellington, New Zealand, 11–14 December 2018, Proceedings*; Springer: Berlin/Heidelberg, Germany, 2018.

[31] Lipton, Z.C.; Berkowitz, J.; Elkan, C. A critical review of recurrent neural networks for sequence learning. *arXiv* **2015**, arXiv:1506.00019.

[32] Che, Y.; Deng, Z.; Lin, X.; Hu, L.; Hu, X. Predictive Battery Health Management With Transfer Learning and Online Model Correction. *IEEE Trans. Veh. Technol.* **2021**, *70*, 1269–1277.

[33] Deng, Z.; Lin, X.; Cai, J.; Hu, X. Battery health estimation with degradation pattern recognition and transfer learning. *J. Power Sources* **2022**, *525*, 231027.

[34] Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2008; Volume 2.

[35] Liu, D.; Pang, J.; Zhou, J.; Peng, Y.; Pecht, M. Prognostics for state of health estimation of lithium-ion batteries based on combination Gaussian process functional regression. *Microelectron. Reliab.* **2013**, *53*, 832–839.

Chapter 5. Conclusion and Outlook

5.1 Conclusion

This dissertation takes lithium-ion battery as the research object, and firstly designs for lithium-ion battery aging test, and analyzes the lifetime influencing factors and aging characteristics of lithium-ion power battery singles through aging experimental data. Secondly, a battery health state estimation method is established based on the data-driven approach to achieve health state estimation for lithium-ion battery cells and pack. The specific research results are as follows.

（1）**Analysis and test of the aging characteristics of lithium-ion batteries.** Given the influence of different environmental factors on battery aging, a battery aging test platform is built. A lithium-ion battery aging test procedure is designed, including aging tests at different ambient temperatures, different charging rates, and different depths of discharge. Based on the battery cycle aging test and the battery calendar aging test, statistical analysis and comparative research on the aging characteristics of lithium-ion batteries are carried out.

（2）**Estimation of the state of health (SOH) of lithium-ion battery cells based on the improved Gaussian Process Regression (GPR) model.** The single kernel function based on the GPR model is not accurate in estimating the SOH of the lithium-ion battery, and cannot accurately capture the phenomenon of overall capacity decay and local capacity recovery. A method for estimating the SOH of lithium-ion batteries based on modified GPR combined with charging and discharging features is proposed. First, the changes in battery voltage and temperature curves among different aging cycles are analyzed in detail, and health indicators (HIs) that can effectively represent the health status of the battery are proposed. Then, the Pearson correlation analysis

method is used to quantify the correlation between HIs and SOH, and three HIs with strong correlation are employed in this paper. Next, a novel compound kernel function is proposed for battery SOH estimation, and different pairs of mean function and kernel function chosen from four mean functions and sixteen kernel functions are used to construct GPR models, and their estimation accuracy is compared subsequently. Finally, four different batteries with various initial health conditions from the NASA battery dataset are used to verify the performance of the proposed method. Experiments show that its estimated mean-absolute-error (MAE) and root-mean-square-error (RMSE) are only 1.7%, and 2.41%, respectively. The results show that compared with a single kernel function, the GPR model based on a composite kernel function is more suitable for capturing the battery aging characteristics of various trends and can achieve an accurate estimation of the battery SOH.

（3）**Battery pack health estimation based on early data.** For the estimation of the SOH of the battery pack, it takes a lot of time and economic cost to complete the full-life aging test of the battery pack. Aiming at the above problems, a method for predicting the future health of the battery pack is proposed, which uses the aging data of the battery cells over their entire life cycles and the early cycling data of the battery pack. Firstly, HIs are extracted from the experimental data, and high correlations between the extracted HIs and the capacity are verified by the Pearson correlation analysis method. To predict the future health of the battery pack based on the HIs, degradation models of HIs are constructed by using an exponential function, long short-term memory network, and their weighted fusion. The future HIs of the battery pack are predicted according to the fusion degradation model. Then, based on the GPR algorithm and the battery pack data, a data-driven model is constructed to predict the health of the battery pack. Finally, the proposed method is validated with a series-

connected battery pack with fifteen 100 Ah lithium iron phosphate battery cells. The MAE and RMSE of the health prediction of the battery pack are 7.17% and 7.81%, respectively, indicating that the proposed method has satisfactory accuracy. The results show that compared with the single feature decay model, the fusion feature decay model can predict the future HIs of a battery pack with more accuracy, which contributes to the satisfactory prediction accuracy of battery pack health based on the GPR model.

5.2 Outlook

This dissertation addresses the deficiencies of existing research on lithium-ion battery health estimation. It focuses on the health state prediction of lithium-ion battery singles and lithium-ion battery packs, based on a data-driven approach, for different operating conditions. However, given the authors' current limited knowledge, there are still many issues to be further addressed in the research on lithium-ion battery health state prediction, and subsequent work will be carried out in the following aspects.

(1) Enhancing data-driven interpretability. The mainstay of the data-driven approach is the training and prediction of the model with the help of machine learning methods. However, the machine learning method is a black-box model without sufficient knowledge about its internal optimization. Therefore, while obtaining satisfactory prediction results, more thinking about the reasons that produce excellent model results is needed. This method will help optimize and find a model with better generalization ability.

(2) Integration of data-driven methods with electrochemical mechanism modeling methods. In building the health state prediction model of Lithium-ion batteries, the data-driven approach starts from the data characteristics without considering the specific model construction. In contrast, the mechanical model

approach completes the typical model building with the help of Lithium-ion battery parameters and then realizes the study of the battery aging phenomenon.

(3) Creation of actual lithium-ion battery operating conditions data. In the actual operation of lithium-ion batteries, there are various operating conditions, but the current research mainly stays in laboratory conditions. To establish a lithium-ion battery health state prediction model with stronger generalization performance, it is necessary to consider a variety of operating conditions and collect richer experimental data and actual operating conditions data, such as electric vehicles and renewable energy smart grids.

(4) Joint estimation of multiple states at different time scales. The existing research mainly focuses on single-state estimation or prediction, but the joint estimation and prediction of multiple states with a strong coupling of electric-thermal-mechanical-aging inside the actual battery, such as the joint estimation of different time scales of SOC, SOE, SOH, SOP, SOT, SOS, etc., will be a crucial technology to be overcome in the next-generation intelligent battery system.

List of publications and presentations

Publications

(1) **Jiwei Wang**, Zhongwei Deng, Tao Yu , Akihiro Yoshida , Lijun Xu, Guoqing Guan, and Abuliti Abudula. " State of health estimation based on modified Gaussian process regression for lithium-ion batteries ", Journal of Energy Storage.51(2022):104512.

(2) **Jiwei Wang**, Zhongwei Deng, Kaile Peng, Xinchen Deng, Lijun Xu, Guoqing Guan, and Abuliti Abudula. "Early Prognostics of Lithium-Ion Battery Pack Health", Sustainability, 14(4) (2022)2313.

(3) Nutthaphak Kitiphatpiboon, Suchada Sirisomboonchai, Meng Chen, Shasha Li, Xiumin Li, **Jiwei Wang**, Xiaogang Hao, Abuliti Abudula, Guoqing Guan. " Facile fabrication of O vacancy rich CuVOx nanobelt@NiO nanosheet array for hydrogen evolution reaction ", Electrochimica Acta.405(2022):139623.

(4) Wang, Jiajia, Xiyan Yue, Xin Cao, Zhao Liu, Amar M. Patil, **Jiwei Wang**, Xiaogang Hao, Abuliti Abudula, and Guoqing Guan. "Metal organic frameworks derived CoS2/NiS2 heterostructure toward high-performance sodium storage anode materials", Chemical Engineering Journal 431 (2022): 134091.

(5) XU Lijun, WANG Weiqing, DUAN Youlian, **WANG Jiwei**. " Optimized method of photovoltaic generator-water electrolyser direct coupling through DC/DC converter ", Chinese Journal of Power Sources.42(2018): 1668-1671

(6) HU Bing，XU Lijun，HE Shan，SU Xin, **WANG Jiwei**, " Researching progress of hydrogen production by PEM water electrolysis under the goal of carbon emission peaking and carbon neutrality ", Chemical Industry and Engineering Progress.(DOI：10.16085/j.issn.1000-6613.2021-2464)

(7) **Jiwei Wang**, Yunhong Che, Zhongwei Deng, Kaile Peng, Guoqing Guan, and Abuliti Abudula. "Lifetime prognostics of lithium-ion battery pack based on its early cycling data and complete degradation information of battery cells." In 2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA), IEEE, 2021.

(8) Zhao Liu, Jiajia Wang, Xiyan Yu, Zhengkun Xie, Hongxin You, **Jiwei Wang**, Abuliti Abudula, Guoqing Guan,"A flexible Li2SnO3-coupled PEO-based single-ion conducting composite solid-state electrolyte for highly-stable Li metal batteries",Journal of Alloys and Compounds, 911(5)(2022)165138.

(9) Zhao Liu, Jiajia Wang, Xiyan Yu, Zhengkun Xie, Hongxin You, **Jiwei Wang**, Abuliti Abudula, Guoqing Guan,"Foldable nano-Li2MnO3 integrated composite polymer solid electrolyte for all-solid-state Li metal batteries with stable interface", Journal of Colloid and Interface Science,621(2022)232-240.

(10) Peng Liu, Tao Lei, Qian Xiang, Zexuan Wang, **Jiwei Wang**. "Animal Fur Recognition Algorithm Based on Feature Fusion Network", Journal of Multimedia Information System. 9(1)(2022)1-10.

Conference presentations

(1) 2020 年 08 月 05-06 日 第 29 回日本エネルギ−学会大会（口頭発表）

(2) 2020 年 09 月 24-26 日 化学工学会第 51 回秋季大会（口頭発表）

(3) 2020 年 09 月 26-27 日 令和 2 年度化学系学協会東北大会（口頭発表）

(4) 2021 年 03 月 22-24 日 電気化学第 88 会大会（口頭発表）

(5) 2021 年 08 月 04-05 日 第 30 回日本エネルギ−学会大会（口頭発表）

(6) 2021 年 12 月 17-19 日 2021 IEEE 2nd International Conference on Information

Technology, Big Data and Artificial Intelligence (ICIBA) （口頭発表）

(7) 2022 年 08 月 04-05 日 第 31 回日本エネルギ−学会大会（口頭発表）