



**Equivalent Circuit Model-based Research on SOC Estimation
Method of Lithium Battery**

by

Tong Wu

Supervisor: Dr Gareth Jones

Project submitted as part of the requirements for the award of
MSc Software Engineering

September 2022

Declaration and Form of Consent

DECLARATION

I, ...Tong Wu..... declare that I am the sole author of this Project; that all references cited have been consulted; that I have conducted all work of which this is a record, and that the finished work lies within the prescribed word limits.

This work has not previously been accepted as part of any other degree submission.

Signed: Tong Wu

Date: September 2022

FORM OF CONSENT

I _____ Tong Wu _____ hereby consent that my Project, submitted in candidature for the M.Sc. Applied Computing, if successful, may be made available for inter-library loan or photocopying (subject to the law of copyright), and that the title and abstract may be made available to outside organisations.

Signed _____ Tong Wu _____

Date _____ September 2022_____

Where necessary replace M.Sc. Applied Computing with your named award.

ABSTRACT

In response to the deteriorating environment and the scarcity of oil resources, the purely electric vehicle industry is growing at a rapid pace worldwide. As an important part of purely electric vehicles, research into lithium-ion battery packs for purely electric vehicles is supported by many companies. The State of Charge (SOC) of a battery is an important area of research for power Li-ion battery packs and is a direct indicator of the amount of charge remaining in the battery. Accurate estimation of battery SOC is important in battery research not only to prevent overcharge and overdischarge, but also to improve the overall performance of the battery. The main work of this paper is as follows:

Firstly, the battery characteristics are used as an entry point to analyse its operating mechanism, to compare and study the advantages and disadvantages of various battery models, to model the battery rationally and to lay the foundation for further battery SOC estimation. The analysis focuses on the effects of different temperatures and different capacity decay cases on the battery performance.

Secondly, the advantages and disadvantages of the equivalent circuit-based model and the support vector regression algorithm (SVR) are combined to compensate for the error in the estimation results based on the equivalent circuit-based model through quadratic estimation on the one hand, and on the other hand, to address the problem of increasing error in the case of battery SOC capacity below 20%, extremely low temperature conditions and a large degree of battery capacity decay, the model reconstruction method is used to re-perform the model training. Finally, the algorithm is tested off-line at different temperatures and with different degrees of capacity degradation to verify the superiority of the algorithm.

Finally, after proposing the SOC estimation method based on the equivalent circuit model, in order to make the SOC estimation method truly applied and verified in practice, a set of front and back-end separated battery management cloud platform is finally built, which enables the uploading of various data of the battery management system through the Bluetooth transmission module, and can analyse and process these

data, so as to display the status of the battery management system in real time and be used to understand the real-time status of the battery system.

Keywords: SOC SVR Equivalent circuit model Cloud Platform

TABLE OF CONTENTS

1 INTRODUCTION	1
2 LITERATURE REVIEW.....	3
2.1 Background to the development of power batteries.....	3
2.2 Current status of research on power batteries.....	4
2.3 Current status of power batteries SOC estimation research.....	5
2.4 Current status of research on battery management cloud platforms.....	11
3 POWER battery characterisation and modelling theory.....	13
3.1 Power battery characterization.....	13
3.2 Power Battery Modelling Theory.....	16
4 Support vector machine regression model prediction based on the DP model.....	23
4.1 SVR Algorithm.....	23
4.2 DP model-SVR combination algorithm implementation.....	28
4.3 Analysis of experimental results.....	52
4.4 Model Summary.....	55
5 Design of a cloud platform for power battery management.....	56
5.1 Functional analysis.....	56
5.2 Development environment for the Power Battery Management Cloud Platform..	56
5.3 Development model for a power battery management cloud platform.....	56
5.4 Construction of a cloud platform for power battery management.....	57
5.5 Cloud Platform Functionality Showcase.....	58
6 Full summary and REFLECTION.....	62
6.1 Full summary.....	62
6.2 Reflection.....	62

1. Introduction

Fossil energy is used all the time in human life, however, it is a non-renewable resource, not unlimited in quantity, and the amount of resources is getting smaller and smaller over time, and one day it will be consumed. The use of oil and gas as the main fossil energy sources is not only a resource constraint, but also puts enormous pressure on the natural environment. In order to solve the problem of resource shortage and environmental pollution, all industries need to make fundamental changes to their energy sources and look into the development of clean and renewable energy sources. As an important sector in the development of society, the automotive industry has made a significant contribution to social development, but it has also brought a series of problems. Throughout its history, the automotive industry has been dominated by petroleum-based vehicles, but the non-renewable and polluting nature of its main energy source has inevitably led to a shift in the direction of a more environmentally friendly vehicle that can alleviate energy shortages. Purely electric vehicles are a good option in this direction, contributing to environmentally friendly changes and to the sustainable development of society. Different countries are becoming aware of this change in direction and are investing more money in research and development to optimise the solution to the problem. The most commonly used power source for pure electric vehicles is the lithium-ion battery, whose low self-discharge rate, high energy density, recyclability, long life and environmental friendliness make it stand out from other batteries. The BMS is the regulatory system in the battery pack and has the most profound impact on battery performance. During the operation of an electric vehicle, the BMS monitors the internal state of the battery in real time, prevents abnormalities in the charging and discharging state and performs regular intelligent maintenance to protect the battery, thereby increasing its service life. The State of Charge (SOC) is an important parameter of the BMS and indicates the amount of capacity remaining. Accurate estimation of the battery SOC can prevent over-discharge or over-charge of the battery, help protect the battery from explosion or fire and improve battery performance. However, in the actual operation of electric vehicles, SOC cannot be

directly measured due to the limitations of the equipment, and the variability of the environment and the complexity of real-time operating conditions make it more difficult to accurately estimate SOC. The study of Li-ion power battery modelling and charge state estimation can effectively increase the battery life and improve the performance of electric vehicles, and achieve real-time accurate estimation of Li-ion power battery SOC, which is of practical significance to promote the application of Li-ion power batteries in electric vehicles.

With the development of IoT technology, 5G technology and cloud technology, new vitality has been injected into the development of battery management systems. The development of the traditional battery management system's backend management is to build the system framework, planning and design functions, programming and development independently, this development method has long development cycle, large workload, poor system stability, high programming ability requirements for developers, how to use the current technology to simplify the development of the backend management system, rapid development of the required BMS backend management system is the current need to solve one of the One of the problems that needs to be solved is how to make use of current technology to simplify the development of back office management systems and quickly develop the required BMS.

Based on the current pain points of the battery management system's backend management development, the built cloud platform includes the storage of data, remote monitoring, abnormal alarm, data playback, data presentation, etc. The various layers cooperate with each other to provide a good management system for users. This solution allows for the rapid, efficient and low cost development of the required BMS back office management system.

2. Literature Review

2.1 Background to the development of power batteries

As a rechargeable battery, scientific work on the power batteries dates back to the nineteenth century. In the 1860s, the French scientist Gaston Plante developed the world's first rechargeable battery, the lead-acid battery, and in 1899, the Swedish scientist Waldemar Jungner^[5] developed another nickel-metal rechargeable battery, the nickel-cadmium battery. The nickel-cadmium battery was the first battery to use an alkaline electrolyte, and its group of electrodes consisted mainly of metallic nickel and metallic cadmium immersed in a solution of potassium hydroxide. In the same year, another type of nickel-metal battery, the nickel-iron battery, was invented, but it was not used on a large scale because it was difficult to seal and inefficient, and because it was found to have a short life and to leak easily after commercialisation. The development of the entire battery industry. Many major breakthroughs were made in materials, structural design and manufacturing processes, which led to rapid improvements in battery performance. To date, battery technology has been widely used in transportation, electronic equipment, aerospace, defence and military applications^[8].

In the initial development stage the performance of new energy vehicles is far from the level of traditional fuel vehicles. Lead-acid batteries are more mature in technology because of their long development time, their energy density is relatively high, their performance is stable and safe. However, the energy density of gasoline can reach more than 10,000 W-h/kg, so lead batteries must have a lot of space and mass to achieve the same range as gasoline, which means that cars equipped with lead batteries have to be bigger and heavier than traditional fuel cars^[12]. In addition, lead batteries emit huge amounts of lead-containing chemical gases during processing, production and recycling, causing serious air pollution, which is contrary to the purpose of developing new energy vehicles. Nickel-cadmium batteries have improved energy density to 50 W-h/kg and increased cycle times to 500, offering better performance. Again, because metallic cadmium is irreversible in terms of soil contamination, it is not widely used. The nickel-

metal hydride battery that emerged in the same period was a less polluting battery with an energy density of 60 to 80 W-h/kg and good low temperature characteristics, both of which have higher overall qualities than the two categories mentioned above. Since 1990, NiMH batteries have been widely used in the hybrid models of many major car brands. However, NiMH batteries were also not widely favoured because of their charging heat and other problems. Lithium-ion batteries, on the other hand, have a high specific energy and power density, allowing them to store more energy for the same volume and mass. With these advantages, the mass of the battery pack can be reduced by 40 to 50 per cent and the volume reduced by 20 to 30 per cent^[15]. At the same time, compared to the above three types of batteries, they release fewer atmospheric pollutants, and the materials used in their manufacture are relatively environmentally friendly, energy-saving and environmentally friendly, in line with the original intention of the development of electric vehicles. Based on these advantages, lithium-ion batteries are the preferred choice for electric vehicle power batteries.

2.2 Current status of research on power batteries

Power batteries are difficult to break through in many technologies because of their complex internal physicochemical reactions, which involve multiple disciplines and fields. However, with the rapid development of the electric vehicle industry in recent years, many of the technical bottlenecks of power batteries are being solved one by one. In the early stages of development, lithium-ion batteries had high manufacturing costs due to the high price of the cathode material LiCoO_2 ^[20], but through several years of development, LiMn_2O_4 and LiFePO_4 were used as alternatives, which greatly reduced the cost and gradually improved the discharge performance with the development of battery management technology. Li-ion batteries are therefore widely favoured. Currently on the market widely used in electric vehicles, the main types of batteries are ternary lithium batteries and lithium iron phosphate batteries, such as Chery Automobile and Geely Automobile and other brands are using ternary lithium batteries. Current research on lithium-ion batteries covers a wide range of disciplines and fields, including materials such as cathode and anode materials, electrolyte composition,

polymer diaphragms, etc., and battery management systems such as the collection of terminal voltage and current information of the entire battery pack and individual cells, monitoring of status, safety protection, charge and discharge control, energy control, equalisation management, etc.

2.3 Current status of power batteries SOC estimation research

2.3.1 Current status of research on SOC estimation methods

The role and importance of the SOC of an electric vehicle's on-board power battery is comparable to that of a traditional vehicle's remaining fuel level display, and accurate SOC estimation is extremely critical for the entire BMS of an electric vehicle. As the main power source of the vehicle, accurate SOC estimation can ensure the safety of the power battery during use. However, due to the complexity of the internal structure of the power battery and the diversity of its chemical properties, as well as the complex and variable operating conditions of the on-board battery, the combination of multiple factors makes accurate SOC estimation difficult. Common SOC estimation methods can be divided into four categories: characterisation-based parameters, ampere-time integration, model-based and data-driven^[21].

1. Characterisation-based parameters

Characterisation parameters, also known as characteristic parameters, are used to establish a mapping relationship between the characterisation parameters and the battery SOC for the battery SOC based on offline data obtained from standard experiments. Once the relationship has been established, the values of the characterisation parameters of the battery are measured or calculated during the charging and discharging process, and the SOC of the battery at this moment is determined based on the previously established one-to-one correspondence. The characterisation parameters involved for power batteries include residual capacity, impedance, OCV, etc.

As the discharge current of a new energy vehicle is not constant under operating conditions, it is difficult to determine the residual capacity through ampere-time integration, making this method a good predictor only under standard test conditions.

The impedance spectrum method requires the precise determination of the internal impedance of the battery during charging and discharging, which is used to calibrate the mapping between the state of charge and the impedance of the battery, and finally to determine the SOC value under the current impedance value by looking up the table. Compared to the above, OCV only needs to measure the current open circuit voltage and is therefore relatively easy to implement and stable. However, an accurate OCV can only be obtained after a few hours of resting a battery that has just been charged and discharged, so in practice it is often necessary to use it in conjunction with an identification method.

2. Ampere-time integration

The Anschluss method of integration is estimated by means of the SOC definition:

$$z(t) = z(t_0) - \frac{\int_{t_0}^t \eta_i i_L(\tau) d\tau}{C_{max}} \quad (1)$$

Where, $z(t)$ refers to the estimated SOC value at that t moment, $z(t_0)$ refers to the initial value of power batteries SOC, η_i indicates the power batteries charging and discharging Coulomb efficiency, which is usually 1 under discharging conditions and 0.98~1 under charging conditions; $i_L(\tau)$ refers to the power batteries charging and discharging current at τ moment; C_{max} refers to the maximum available capacity of the power battery under current conditions.

The classical and easy-to-use method of integration still suffers from a number of shortcomings:

- (1) The problem of accurate acquisition of initial values of power batteries SOC.
- (2) The problem of accurate acquisition of charge and discharge circuits. Current sensors are susceptible to interference from noise, temperature drift and other, etc. unavoidable factors, which reduce the acquisition accuracy. On the other hand, errors in the current cause cumulative errors in the integration calculation, and the rounding calculation of the controller also has an impact on the final result.

(3) Battery performance degradation issues. Battery ageing can lead to a reduction in static battery capacity.

These three shortcomings further reduce the reliability of the method. In order to reduce the errors and minimise the limitations of the above factors, they are often addressed by periodically updating the initial calibration. This is why the method is often used in combination with other methods of updating initial values.

3. Model-based

The model-based approach refers to the use of reliable performance models to simulate the physical entity of the battery to characterise its internal reactions, commonly used models include electrochemical models, equivalent circuit models, fractional order models etc. Based on the determined equivalent model of the power battery, a suitable filtering method is used as the algorithm for the state estimation stage to compensate for errors in the ampere-time integration and to correct for SOC.

The process is as follows:

- (1) A priori estimation: the current SOC estimate is calculated using the Anschluss integral.
- (2) Model parameter acquisition: Determine the value of the model parameter at this moment in time by looking up the table based on the mapping between the battery characterisation parameters and the SOC.
- (3) Obtain the model end voltage: Calculate the end voltage at this moment based on the mapping between the end voltage and the parameters.
- (4) End voltage error acquisition: Error calculation based on the model end voltage from the previous step and the actual measured voltage.
- (5) The error from the previous step is used to correct the SOC estimate and apply it to the input value at the next moment.

As can be seen from the process, the model-based approach forms a closed loop by continuously correcting the initial value of the ATS integration and in doing so improves the robustness of the algorithm. The accuracy of the SOC prediction of the

model-based approach is therefore dependent on the accuracy of the model built and the effectiveness of the initial value correction process, i.e. the state estimation algorithm, which can be affected by inaccurate model accuracy and inappropriate state estimation algorithms. On the other hand, the initial value calibration by the filtering algorithm can cause large fluctuations in the estimated value if the correction is too large, and how to determine the magnitude of the correction needs to be determined according to the actual situation.

4. Data-driven

The data-driven approach means that the model is trained by relying on sufficient and reliable offline data. The input data includes the charge/discharge current, charge/discharge voltage, temperature and internal resistance of the battery, while the output data includes the SOC and measurement error of the battery. The process is as follows:

- (1) Pre-processing of offline data is the processing of training data through data cleaning, data normalisation and data chunking, so that the training data meets the input and output requirements of the model.
- (2) The model structure is determined by the amount of data in the training set. The model is then trained using the training data set, usually with a validation set validation result error or number of training rounds as a condition.
- (3) The model is tested by bringing in the trained model from the previous step using the test set and using the output error to determine if the model's prediction accuracy meets the requirements, if it does then the training is deemed complete, otherwise return to the second step to redefine the model structure.

The data-driven approach excels at solving problems with non-linear relationships like power battery SOC estimation. It does not need to consider the specific relationship between input and output quantities and interference factors, but it needs to consider whether the amount of data selected is sufficiently large and reliable and whether the selected test data correctly reflects the battery characteristics. However, the acquisition

of large amounts of reliable data is more difficult to achieve under practical working conditions, so the selection of a suitable data-driven method is also a key consideration. The advantages and disadvantages of the above four types of methods are as follows:

Table 1 Evaluation of four types of SOC estimation methods

Methods	Advantages	Disadvantages	Accuracy	Robustness
Based on characterisation parameters	Simple and easy to implement Low calculation costs	Vulnerable to, environmental factors Requires periodic calibration Requires precision measuring instruments	Bad	Good
ATS Points	Simple and easy to implement Low calculation costs	Dependence on SOC initial values Open-loop algorithm requires periodic correction Vulnerable to current drift, noise, aging, etc.	General	Bad
Model-based	High calculation accuracy Closed loop feedback control Strong self-adaptation	Needs to be supported by an accurate model High calculation costs Requires high initial value accuracy	Excellent	Good
Data driven	High estimation accuracy Adept at solving non-linear problems	High algorithm complexity High dependence on training set	Excellent	General

2.3.2 SOC estimation difficulties

As explained above, the performance of the model-based approach depends primarily on the accuracy of the determined model and the state estimation algorithm. In the state estimation part the Kalman filter type of algorithm is a widely used algorithm, which is based on the principle of treating the output error of the circuit model as noise and taking a recursive approach to filtering the system noise. However, initially the simple Kalman filtering algorithm was only valid for linear systems, so in order to extend the algorithm to non-linear systems, the Extended Kalman Filter (EKF) algorithm was later proposed, which was extended to non-linear systems by Taylor series expansion, but the effectiveness of this method relies on the exact initial value.

The data-driven approach uses a large amount of experimental data for model training and has a very good fitting capability, does not require consideration of the internal construction of the power battery and the objective mapping relationships between the various feature quantities, and is theoretically applicable to SOC estimation for any type of battery. The most typical method is the neural network algorithm, but the problem of overfitting the model needs to be considered. On the other hand, the structure of the models trained by neural networks is more complex and computationally intensive, making them difficult to implement in practical engineering.

From the above analysis, it is clear that the main difficulties in the current SOC estimation are:

(1) Construct a comprehensive equivalent model. Although the electrochemical model can accurately represent the particle changes and electrode reactions inside the battery through a series of partial differential equations and algebraic equations, the model is too complex and computationally intensive, requiring a high level of computing power from the BMS. The construction of an equivalent circuit model based on characterisation and characteristics parameters does not allow the actual state of the battery to be represented chemically, and the internal structure of the battery is not as accurate as it could be if only simple circuit components were used. The accuracy of the established circuit model will inevitably have a direct impact on the accuracy of the

SOC estimation of the power battery, so it is important to establish a suitable and reliable battery model. important direction.

(2) Considering the complex working conditions of batteries in the actual operating process, and the interference of environmental noise in the operating current, and the influence of battery terminal voltage by ambient temperature and battery capacity decay, the SOC estimation algorithms discussed according to the above contents of this paper all have different degrees of limitations in the face of the complex operating conditions of batteries. Therefore, it has become a research priority for researchers to explore the reasonable matching relationship between the model and the algorithm, so as to improve the traditional algorithm and derive an optimisation algorithm with more universality, reliability, accuracy and practicality.

2.4 Current status of research on battery management cloud platforms

For the research and design of battery management systems, in recent years, Li et al. have proposed a flexible and reliable battery management approach based on battery big data platforms and information physical system (CPS) technology, developed a function extraction model based on deep learning algorithms and a function-oriented battery modelling approach to improve the accuracy of cloud-based battery models. Advances in battery technology have accelerated the transition of the battery management system (BMS) from a pure monitoring unit to a multifunctional integrated unit. To implement effective control of the BMS, it will be accurate and effective to use more comprehensive and faster battery models that reflect the behaviour of the battery charge level to the vehicle. The literature ^[25] reviewed the research and development of multi-physical field model simulation and multi-functional integrated BMS technology, and proposed some key model-based techniques to implement a collaborative green energy vehicle smart BMS. the literature ^[26] analysed state estimation methods from key indicators of the BMS, and made a study of the state of charge (SOC), state of energy (SOE), state of power (SOP), state of health (SOH), remaining useful life (RUL)

and state of equilibrium (SOB) estimation methods are systematically reviewed and discussed.

The current battery management system's background management still has the following problems:

- (1) The state estimation accuracy of the battery management system is not high and the various parameters of the battery system cannot be queried in real time.
- (2) High cost of historical data storage for the whole life cycle of the battery system.
- (3) High cost of data transmission to the backend management system using wireless transmission technology, as well as long development cycle and high development cost of the backend management system.

Advantages of the cloud platform in the field of lithium battery energy storage management:

- (1) Li-ion battery management systems are highly dependent on long-term historical data and use historical data to make ageing predictions for batteries, while the cloud platform has strong computing and data storage capabilities and is able to calculate and analyse big data.
- (2) The cloud platform is able to continuously and accurately monitor lithium battery parameters, online condition estimation, fault prediction and diagnosis.

3. Power battery characterisation and modelling theory

3.1 Power battery characterization

3.1.1 Temperature characteristics of the battery

1. The effect of temperature on power battery capacity

We know that temperature affects the physical and chemical reaction rates of particles. The intuitive effect inside the battery is that the rate of movement of the active particles inside the battery decreases and the rate of chemical reaction decreases. According to laboratory data from the National Laboratory for Electric Vehicles at Beijing University of Technology, the test curves for the maximum available capacity of an 18650 type lithium iron phosphate battery with a rated capacity of 27A-h under ambient conditions of -10°C, 0°C, 25°C and 40°C are as follows:

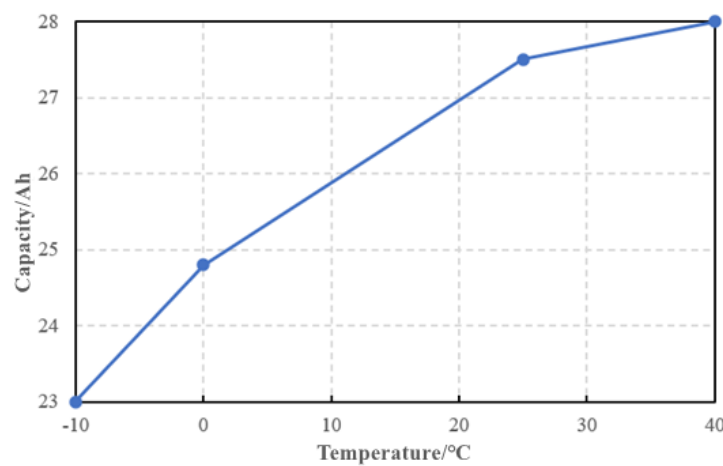


Figure 1 Maximum available capacity of the power battery at different temperatures

As can be seen from the graph, the relationship between the capacity of the power battery and the temperature is positively correlated in the range of -10 to 40°C, i.e. the higher the temperature the greater the discharge capacity. The maximum capacity at -10°C is nearly 23 Ah and the maximum available capacity at 40°C is nearly 28 Ah, with a difference of up to 5Ah. This is sufficient to show that the effect of temperature on battery characteristics is not negligible. The main reason for this effect is that the activity of the battery material decreases as the temperature decreases, and the migration rate of the active particles slows down as the temperature decreases, thus affecting the charge and discharge performance of the battery. The end voltage affects the charging and discharging efficiency of the battery. Therefore, under the same model and the same experimental conditions, when the temperature is higher, the end voltage

of the power battery is relatively high, the discharge efficiency is high and the discharge time is long; when the temperature is lower, the battery will reach the cut-off voltage earlier.

3. Effect of temperature on open circuit voltage

The OCV-SOC mapping relationship for power batteries is a key part of the model-based algorithm. At different temperatures, the curve cannot be stabilized due to the change of internal particles and thus the change of internal resistance of the battery, which is influenced by the temperature. According to the laboratory data from the National Laboratory of Electric Vehicles of Beijing University of Technology, the OCV-SOC curves for an 18650 LiFePO₄ battery with a rated capacity of 27A-h at -10°C, 25°C and 45°C are shown in the figure, from which it can be found that the effect of temperature on the open circuit voltage is negligible in the SOC 10% to 100% band, but in the SOC below 10% The difference between the curves in the 10% and 100% SOC range is negligible. The internal resistance of the battery increases at low temperatures, resulting in a shorter discharge time to reach the cut-off voltage, which in turn causes the problem of inadequate battery discharge, the effect of which on the graph is a reduction in the rate of voltage drop.

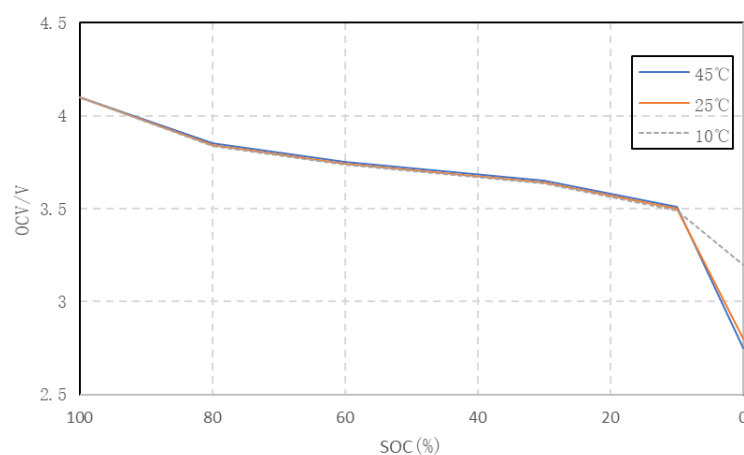


Figure 2 OCV-SOC curves at different temperatures

3.1.2 Battery capacity decay characteristics

1. Capacity decay in relation to the number of cycles

Figure 3 shows the maximum available capacity of the battery under different ageing conditions. It can be seen from the graph that the maximum available capacity and the

number of battery cycles are negatively correlated. The main reason for this phenomenon is that as the number of cycles increases, the battery tends to age, which is directly reflected in the increase in internal resistance. The concentration of active particles in an ageing battery decreases, resulting in a loss of active material inside the battery, which is often seen in batteries that have been in use for a long time, resulting in a depression in capacity. At the same time, impurities in the electrolyte increase, affecting the exchange of lithium ions and free metal particles in the electrolyte will adhere to the electrodes, resulting in an increase in transfer resistance, which in turn reduces the discharge capacity of the battery.

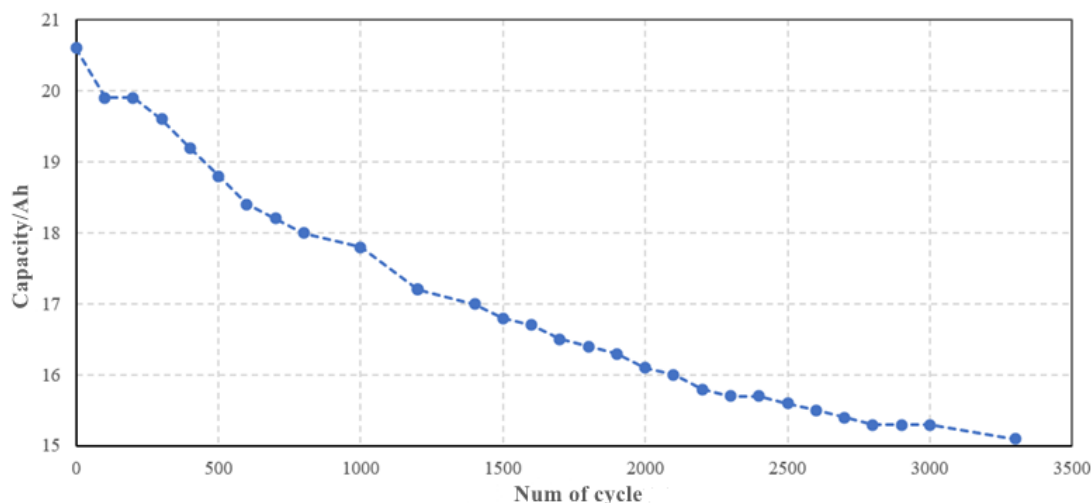


Figure 3 Maximum usable capacity of power batteries in different aging states

2. Effect of capacity decay on open circuit voltage

According to the laboratory data of the National Laboratory of Electric Vehicles of Beijing University of Technology, the OCV-SOC curve of an 18650 LiFePO₄ battery with a rated capacity of 27Ah, Figure 4 shows the OCV-SOC curve under different capacity decay, from the figure below it can be seen that with the battery in the SOC less than 10% stage, the influence of the number of battery cycles on the open circuit voltage is relatively large, the more cycles, the faster the curve drops. The higher the number of cycles, the faster the curve decreases. Under the same conditions of discharge, the more severe the capacity decay, the faster the OCV of the battery falls. As with low temperature conditions, their time to reach the as of voltage is correspondingly shorter.

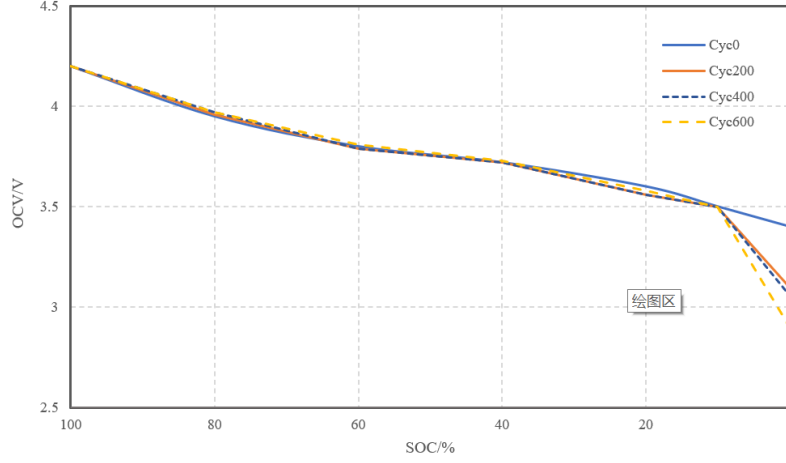


Figure 4 OCV-SOC curves for different ageing conditions

3.2 Power Battery Modelling Theory

3.2.1 Electrochemical modelling

Electrochemical modelling is a direct response to the internal structure and properties of a battery and has a higher degree of accuracy because of its ability to accurately describe the various physical and chemical reactions within the battery. However, because of its ability to accurately describe the internal chemical structure of a battery, the electrochemical model is complex to build, requiring the integration of multiple physical and chemical knowledge, and is computationally intensive and difficult to analyse.

The most representative model in electrochemical modelling is the pseudo two-dimensional (P2D) model. The main principles of which are briefly described as follows:

(1) Diffusion of lithium ions in the liquid phase

The liquid phase refers to the electrolyte in which the lithium ion transfer process is expressed using Fick's second law:

$$\varepsilon_e \frac{\partial a_e}{\partial t} = \frac{\partial}{\partial x} (G_e^{\text{eff}} \frac{\partial a_e}{\partial x}) + b(1-t_+^0)j_r \quad (2)$$

In the above equation, $\frac{\partial}{\partial x} (G_e^{\text{eff}} \frac{\partial a_e}{\partial x})$ represents the effect of diffusion on the lithium ion concentration in the liquid phase, $b(1-t_+^0)j_r$ represents the effect of migration on the lithium ion concentration in the liquid phase, and its value is usually small and can be approximated as 0. ε_e is the liquid phase volume fraction, a_e is the lithium ion concentration in the liquid phase, x is the direction of the electrode plate thickness, G_e^{eff} is the lithium ion liquid phase effective diffusion coefficient, and b is the surface area per unit volume of electrode particles.

(2) Diffusion of lithium ions in the solid phase

The solid phase refers to the interior of the material particles and, unlike in the liquid phase, the coordinate system established in the solid phase is spherical and its diffusion equation in the solid phase is as follows:

$$\frac{\partial c_e}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} (D_s r^2 \frac{\partial c_s}{\partial r}) = D_s \left(\frac{2}{r} \frac{\partial c_s}{\partial r} + \frac{\partial^2 c_s}{\partial r^2} \right) \quad (3)$$

where c_s is the solid phase lithium ion concentration, r is the radius direction of the solid spherical particle and D_s is the lithium ion solid phase diffusion coefficient. The conservation equations are solved in the r -dimensions of the spherical particles - pseudo two-dimensional, hence the model is called the P2D model.

(3) Liquid phase Ohm's law

$$k^{eff} \frac{\partial \varphi_e}{\partial x} = - \frac{2RTk^{eff}}{F} (t_+^0 - 1) \frac{\partial \ln c_e}{\partial x} - i_e \quad (4)$$

where k^{eff} is the liquid phase ionic conductivity, φ_e is the liquid phase potential energy of the battery, R is the ideal gas constant, T is the battery temperature, F is the Faraday constant and i_e is the liquid phase current density.

(4) Solid Phase Ohm's Law

$$\sigma^{eff} \frac{\partial \varphi_s}{\partial x} = -i_s \quad (5)$$

where σ^{eff} is the effective diffusive conductivity of the solid phase, φ_s is the solid phase potential energy of the Li-ion battery, and i_s is the solid phase current density.

(5) Equation of conservation of electric charge

According to the law of conservation of charge, the battery charge and discharge current is equal to the sum of the liquid phase current and the solid phase current inside the power battery, i.e:

$$i_s + i_e = i \quad (6)$$

where i is calculated by the formula:

$$i = \frac{i_L}{S} \quad (7)$$

In the above equation, i_L is the external current and S is the effective area of the electrode. i_s denotes the solid phase current, which is calculated as follows and described by Faraday's law:

$$i_s = nFj_r \quad (8)$$

where $n=1$, is the lithium ion charge number and F is the Faraday constant.

3.2.2 Equivalent circuit model

The equivalent circuit model of a battery consists of circuit components such as resistors, capacitors and voltage sources, which are combined in different ways to form a circuit network to describe the external characteristics of the battery. The model is relatively simple in structure, small in calculation and has good real-time performance. Figure 5 shows a typical equivalent model consisting of an n th order RC network.

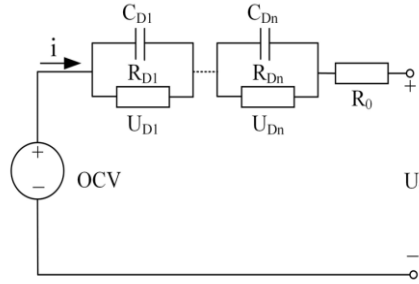


Figure 5 Circuit structure of the n -RC model

The dynamic equations of the n -RC circuit are shown below:

$$\begin{cases} U_t = U_{OC} - R_0 i - U_{D1} - \dots - U_{Dn} \\ i = \frac{U_{D1}}{R_{D1}} + C_{D1} \frac{dU_{D1}}{dt} = \dots = \frac{U_{Dn}}{R_{Dn}} + C_{Dn} \frac{dU_{Dn}}{dt} \end{cases} \quad (9)$$

where U_{OC} denotes open circuit voltage, i denotes load current, R_0 denotes internal resistance, the direction of load current in Figure 5 is positive, R_{Dn} denotes polarisation resistor, C_{Dn} denotes polarisation capacitor, and U_{Dn} denotes the voltage values across the polarisation resistor and polarisation capacitor.

The RC circuit model with multiple RC networks has a high degree of accuracy, as the RC network can simulate many internal reactions that affect the characteristics of the battery, such as polarisation and diffusion reactions, and the higher the order of the circuit model, the closer the simulation results will be to the real value. The higher the

order of the model, the closer the simulation results will be to the true value, but on the other hand, the higher the order of the model, the higher the complexity of the circuit, and the higher the computational process and computational effort, so that when choosing a model, the accuracy of the SOC estimation is very much related to the accuracy of the equivalent circuit model, and the higher the accuracy of the model, the higher the computational accuracy, but at the same time the computational effort will increase, resulting in a significant reduction in the computational speed. Therefore, a balance needs to be found between model accuracy and computational complexity.

Commonly used equivalent models are as follows:

(1) Rint Model

The Rint model is shown in Figure 6:

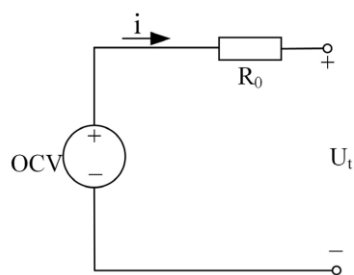


Figure 6 Rint model circuit structure

U_{oc} denotes the open circuit voltage of the power battery and the resistance R_0 denotes the ohmic internal resistance of the battery. Output voltage versus input current for the Rint model:

$$U_t = U_{oc} - iR_0 \quad (10)$$

(2) Thevenin Model

The circuit structure of the Thevenin model is shown in Figure 7:

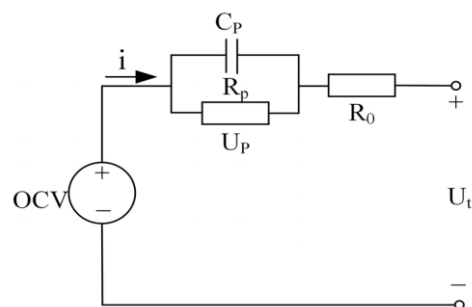


Figure 7 Thevenin model circuit structure

In the diagram, R_p represents the polarisation resistor and C_p represents the polarisation capacitor, and U_p is the voltage of the parallel network of R_p and C_p . For the RC parallel network when first energised, the voltage across the capacitor is 0, and as charging proceeds the voltage across the capacitor gradually rises to be opposite the resistor terminals, so its circuit equation is:

$$\begin{cases} \dot{U}_p = \frac{i}{C_p} - \frac{U_p}{R_p C_p} \\ U_p = U_{OC} - U_p - iR_0 \end{cases} \quad (11)$$

The accuracy of the circuit is improved by the use of an RC parallel network to describe the polarisation response within the battery.

(3) DP Model

The circuit structure of the DP model is shown in figure 8:

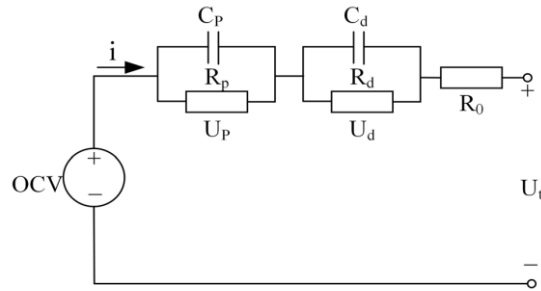


Figure 8 DP model structure

The model uses two parallel RC networks to represent the polarisation characteristics and the diffusion characteristics inside the battery for the battery model equivalence. Where R_p represents the polarisation resistance and C_p represents the polarisation capacitance and R_d represents the diffusion resistance and C_d represents the diffusion capacitance. The equations of the model are:

$$\begin{cases} \dot{U}_p = \frac{i}{C_p} - \frac{U_p}{R_p C_p} \\ \dot{U}_d = \frac{i}{C_d} - \frac{U_d}{R_d C_d} \\ U_t = U_{OC} - U_p - U_d - iR_0 \end{cases} \quad (12)$$

3.2.3 Fractional order model

The fractional-order model is proposed to improve the accuracy of equivalent circuits for purely capacitive-resistive components. We have experimentally found that the frequency response characteristics of a bilayer are different from those of a pure capacitor, and this inconsistency can be compensated for by constant phase angle components. The circuit model is shown in Figure 9.

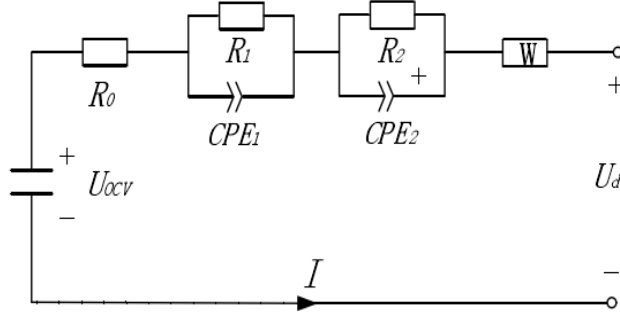


Figure 9 Fractional order model structure

Similar to the equivalent circuit model, the transfer function expression is:

$$G(s) = \frac{U_t(s) - U_{OC}(s)}{i_L(s)} = -(R_i + \frac{R_{ct}}{1 + R_{ct}Ys^\alpha})$$

$$= -\frac{(R_i + R_{ct}) + R_{ct}R_iYs^\alpha}{1 + R_{ct}R_iYs^\alpha} \quad (13)$$

That is:

$$U_t(s) - U_{OC}(s) + R_{ct}Ys^\alpha[U_t(s) - U_{OC}(s)]$$

$$= -(R_{ct} + R_i)i_L(s) - R_{ct}R_iYs^\alpha i_L(s) \quad (14)$$

An inverse Laplace transform of the above equation gives, at the k th sampling time point:

$$U_t(k) = -(R_{ct} + R)i_L(k) - R_{ct}YD^\alpha i_L(k) + U_{OC}(k)$$

$$+ R_{ct}YD^\alpha [U_{OC}(k) - U_t(k)] \quad (15)$$

Electrochemical models can describe the internal chemistry of a power battery with high accuracy but are computationally intensive. The equivalent circuit model uses ideal electrical components to build a simulated circuit model to describe the internal physical characteristics of the battery, and multiple RC parallel connections to reflect the polarization reaction of the battery. The fractional-order model is designed to

compensate for the frequency response characteristics of the double layer that pure capacitive components do not have, and is compensated for by constant phase angle components, which is highly accurate but also computationally intensive, and is commonly used for SOH estimation. Therefore, the equivalent circuit model is chosen for estimating SOC in this paper.

4. Support vector machine regression model prediction based on the DP model

4.1 SVR Algorithm

4.1.1 SVR algorithm and derivation

SVR creates a "spacing band" on either side of the linear function, so that no loss is calculated for all samples falling within the band, and only those outside the band are included in the loss function. The model is then optimised by minimising the width of the band and the total loss. Since the model needs to drop some edge points for minimizing the interval band, the so the relaxation variables ξ_i and ξ_i^* , representing the distances between the upper and lower edge points of figure 10 and the middle real line (y-axis projection distances rather than Euclidean distances are calculated directly as the difference on the y-axis between the points and the projection of the points on the real line.) are introduced .

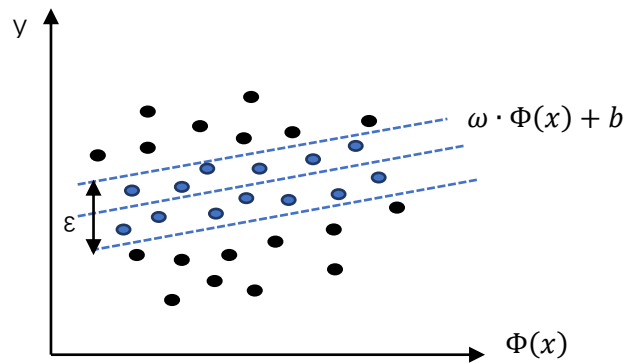


Figure 10 Classified hyperplane

SVR objective function is:

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (16)$$

$$\text{s. t. } \begin{cases} y_i - \omega x - b \leq \epsilon + \xi_i \\ \omega x + b - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (17)$$

Where ξ_i and ξ_i^* are:

$$\begin{cases} \xi_i = y_i - (\omega x + b + \varepsilon), y_i > \omega x + b + \varepsilon \\ \xi_i = 0, \text{ otherwise} \end{cases} \quad (18)$$

$$\begin{cases} \xi_i^* = (\omega x + b - \varepsilon) - y_i, y_i < \omega x + b - \varepsilon \\ \xi_i^* = 0, \text{ otherwise} \end{cases} \quad (19)$$

To minimise the objective function, construct the Lagrangian function according to the constraints:

$$\begin{aligned} L = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^1 (\xi_i + \xi_i^*) - \sum_{i=1}^1 \alpha_i (\varepsilon + \xi_i - y_i + \omega x + b) - \sum_{i=1}^1 \alpha_i^* (\varepsilon + \xi_i^* + \\ y_i - \omega x - b) - \sum_{i=1}^1 (\eta_i \xi_i + \eta_i^* \xi_i^*) \quad (20) \\ \text{s. t. } \alpha_i, \alpha_i^*, \eta_i, \eta_i^* \geq 0 \end{aligned}$$

The original question can be transformed into:

$$\min_{\omega, b} \max_{\alpha_i^{(*)}, \eta_i^{(*)}} L(\omega, b, \xi_i, \xi_i^*, \alpha_i, \alpha_i^*, \eta_i, \eta_i^*) \quad (21)$$

Pairs of the original question:

$$\max_{\alpha_i^{(*)}, \eta_i^{(*)}} \min_{\omega, b} L(\omega, b, \xi_i, \xi_i^*, \alpha_i, \alpha_i^*, \eta_i, \eta_i^*) \quad (22)$$

Find the partial derivatives of ω , b , ξ_i, ξ_i^* , respectively, and let the partial derivative be 0:

$$\begin{cases} \frac{\partial L}{\partial \omega} = \omega - \sum_{i=1}^1 (\alpha_i^* - \alpha_i) x_i = 0 \\ \frac{\partial L}{\partial b} = \sum_{i=1}^1 (\alpha_i^* - \alpha_i) = 0 \\ \frac{\partial L}{\partial \xi_i} = C - \alpha_i - \eta_i \\ \frac{\partial L}{\partial \xi_i^*} = C - \alpha_i^* - \eta_i^* \end{cases} \quad (23)$$

Therefore, if x_i is a support vector, i.e. $\alpha_i \neq 0$ or $\alpha_i^* \neq 0$, then one of the following two equations $\xi_i = y_i - (\omega x + b + \varepsilon)$ or $\xi_i^* = (\omega x + b - \varepsilon) - y_i$ holds. This indicates that the support vector falls outside the ε -pipeline region. For samples that fall within the ε -pipeline region, we determine that these samples have no effect on the model assumptions, based on $\alpha_i = \alpha_i^* = 0$. The determination of the parameter ε determines sparsity and accuracy: when the value of ε is large, we obtain a relatively sparse solution, but at the same time it leads to a less accurate solution. Setting the right value of ε therefore has a significant impact on the effectiveness of the algorithm.

4.1.2 Improved SVR algorithm

1. Optimisation of parameter C and kernel parameters

The penalty factor C and the kernel parameters of the general SVR are determined before training. The penalty factor C determines the maximum error tolerance of the training data, so we want C to be as large as possible, but too high a value of C will lead to overfitting of the training model, and the prediction effect will be reduced, thus affecting the generalisation of the training model. The kernel parameter γ determines the complexity of mapping the sample data to a higher dimensional space, and when the mapped dimension is too high it will also lead to overfitting problems, and too low it will fail to meet the linearity requirement and lead to underfitting.

We therefore want to determine the optimal solution for these two parameters by parameter search before training the model, and the grey wolf optimisation algorithm can help us solve this problem. The basic idea of the grey wolf algorithm is that we assume that the parameter to be optimised is infinite, and then determine a rule to move the parameter towards the target value until it reaches a set error threshold or a set number of iterations, which is the process of finding the optimal parameter. The basic idea is similar to the collective hunting of the grey wolf population: the wolves in the pack are divided into four ranks, α , β , δ , ω wolves in descending order, and during the hunting process, the leading wolf, the wolf that tracks the prey and the last wolf that hunts the prey are determined according to the rank. When the hunt is finally successful, the α wolf is the optimal solution of the parameter we want to find the best.

The specific process is as follows:

(1) Classification into classes

The whole wolf pack is divided into four ranks according to the degree of adaptation from high to low, defined as α , β , δ , ω wolves. Among them, α , β , δ , ω are in the leadership level and can lead wolves of lower rank than themselves. Each iteration will replace a new leader by selecting the best wolves, so the leadership is dynamic and

is constantly updated to ensure that the leadership of the population does not decline and that prey can be caught on every hunt.

(2) Tracking prey

Once the ranking is completed, the prey tracking process can begin. All wolves in the pack are first located and distances are calculated before the hunting work is carried out. During the hunting process, the positions of α , β , δ , ω change, where the distances between individual grey wolves are:

$$D = |C \cdot X_p(t) - X(t)| \quad (24)$$

Updated by:

$$X(t+1) = X_p(t) - A \cdot D \quad (25)$$

where X_p is the position of the prey, t is the number of iterations, X is the current position of the grey wolf, and A is a random variable ($A = 2a \cdot r_2 - a, r_2 \in [0,1]$). It determines whether the grey wolf population is the prey search range; when $A > 1$, the grey wolf population will expand the prey search range, and when $A < 1$, the grey wolf population will narrow the search range. C denotes the perturbation of the prey ($C = 2r_1, r_1 \in [0,1]$), the main parameter controlling the parameter A is a , which we call the convergence factor and is calculated as follows:

$$a = 2 - 2\left(\frac{t}{\max}\right) \quad (26)$$

where t denotes the current number of iterations and $\max$ denotes the maximum number of iterations.

(3) The Hunt

When prey is found by the leadership, the highest ranking α wolf will lead the β and δ wolves to direct the whole pack, guiding them to surround the prey and capture it successfully. When prey is found, α , β , δ wolves use their distance from the prey to guide the other grey wolves to update their direction of movement and step length to keep approaching the target position in the following way:

$$\begin{aligned}
D_\alpha &= |C_1 X_\alpha(t) - X(t)| \\
D_\beta &= |C_2 X_\beta(t) - X(t)| \\
D_\delta &= |C_3 X_\delta(t) - X(t)|
\end{aligned}
\tag{27}$$

The direction of movement and step size of $X(t+1) = X_p(t) - A \cdot D$ are as follows:

$$\begin{aligned}
X_1 &= X_\alpha - A_1 \cdot D_\alpha \\
X_2 &= X_\beta - A_2 \cdot D_\beta \\
X_3 &= X_\delta - A_3 \cdot D_\delta \\
X(t+1) &= \frac{X_1 + X_2 + X_3}{3}
\end{aligned}
\tag{28}$$

$$X(t+1) = \frac{X_1 + X_2 + X_3}{3} \tag{29}$$

Eq. 22 determines the final position of the ω wolf and that A and C are the random and adaptive vectors respectively.

The flow chart of the algorithm implementation is as follows:

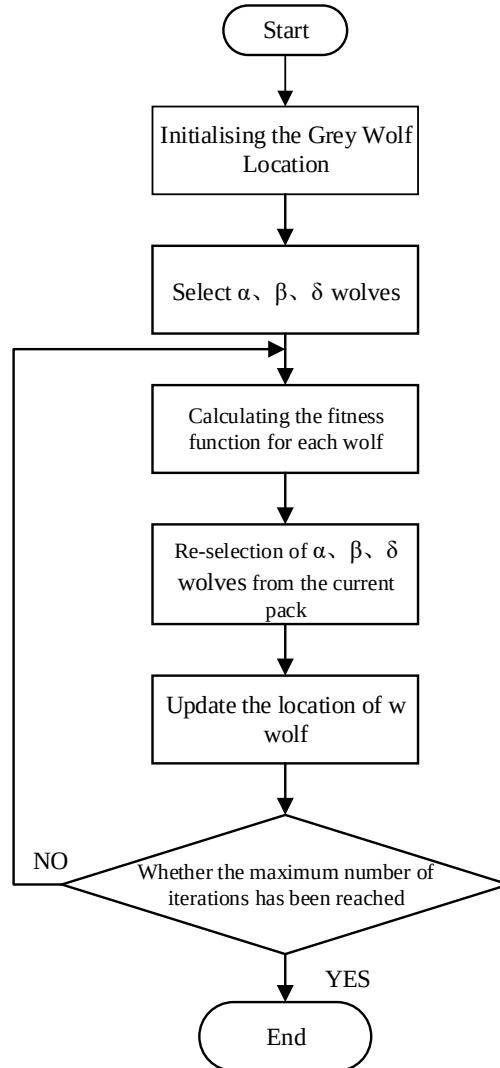


Figure 11 Flow chart of the Grey Wolf algorithm

4.2 DP model-SVR combination algorithm implementation

4.2.1 Experimental data sources

1. Off-line test data

(1) Open circuit voltage test

The main purpose of the open-circuit voltage test is to obtain data on the battery at different temperatures and under different capacity decay conditions to perform the performance test of the algorithm. The specific steps of the open-circuit voltage test are as follows.

- (i) Conduct a discharge experiment on the battery with a constant current and stop the discharge when the cut-off voltage is reached. Leave it for 3h and measure its voltage value.
- (ii) Charge the battery with a constant current of 2700mA, the cut-off condition is that the charging current drops to the cut-off current, then leave it for 10min.
- (iii) Discharge the battery at a constant current of 1890mA, the cut-off condition is that the battery voltage drops to the cut-off voltage and then is left for 10min.
- (iv) Skip steps (i) to (ii) for 600 cycles.

The charge and discharge voltage and current curves for one cycle (excluding resting) are shown below:

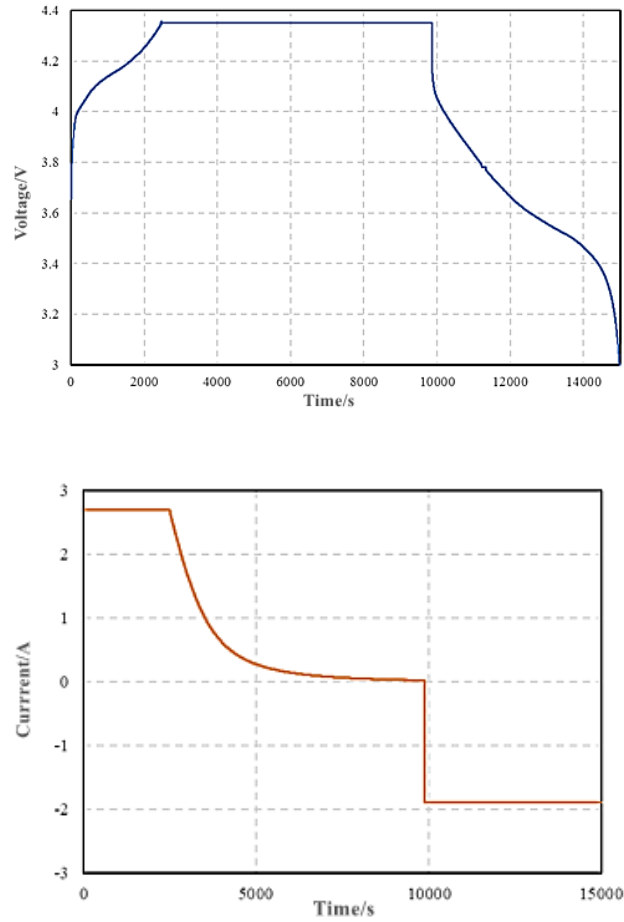


Figure 12 Charge and discharge voltage and current curves during one cycle (excluding resting)

(2) Hybrid Power Pulse (HPPC) Characterisation Testing

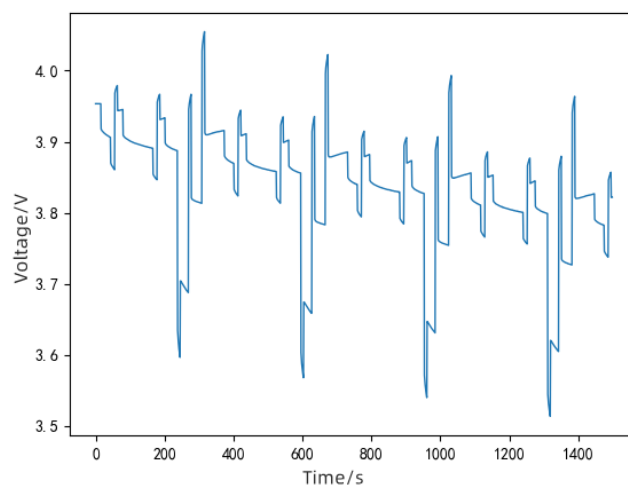
According to power battery modelling theory, experimental test data from hybrid pulse dynamics are often used for off-line identification of model parameters.

The HPPC test procedure is specified as follows:

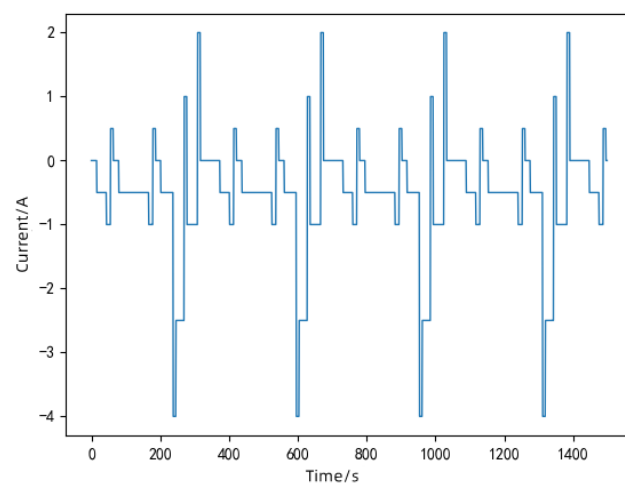
- (i) First, the experimental battery is charged using a standard current.
- (ii) The battery is fully charged and left to stand for 4 hours.
- (iii) Using the HPPC current, the power battery is discharged at a constant current for a fixed period of time and then left to stand for 1 hour.
- (iv) Repeat the operation set in step (iii) until the power battery reaches its discharge cut-off voltage.

2. Dynamic working data

Due to the limitations of laboratory experimental conditions, it is not possible to conduct dynamic working conditions experiments on power batteries, so in order to verify the estimation effect of the algorithm in this paper, two main types of dynamic working conditions data are used, namely dynamic stress test data (DST) and US Federal Urban Driving Conditions data (FUDS). Both types of data are derived from experimental testing of 18650 type batteries. The Dynamic Stress Test (DST) is designed to simulate the stress conditions experienced by a vehicle during acceleration and deceleration, cornering, driving on rough roads, etc. Figure 12 shows the dynamic stress test voltage data and current data. Figure 13 shows the voltage data and current data for the US federal city driving conditions.

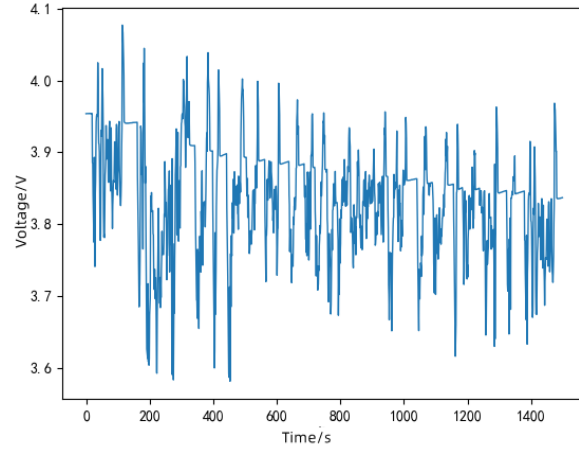


(a) DST test voltage

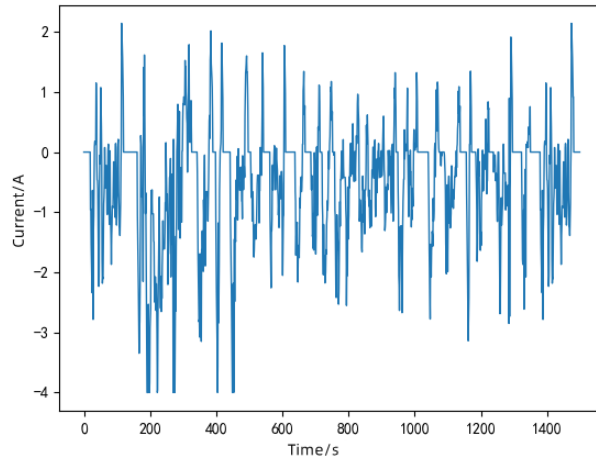


(b) DST test current

Figure 12 Dynamic stress test data



(a) FDUS test voltage



(b) FUDS test current

Figure 13 US Federal Urban Driving Service Test Data

4.2.2 Experimental data processing

There are a number of clustering methods for pre-processing experimental data, and the basic but efficient K-means clustering method is chosen to save unnecessary computational costs. The basic principle is to first select an initial set of n samples as clustering centres, then assign the samples to the originally determined clustering centres according to the principle of minimum distance (error sum of squares criterion) to obtain the initial clusters, and then use the distance as an evaluation index of similarity, so as to determine whether the classification is a composite requirement, and if not, to modify the classification, and so on and so forth until the requirements are met. The specific implementation of the K-means algorithm is as follows:

(1) Choose n initial clustering centres $l_1(1), l_2(1), \dots, l_n(1)$. Take the initial n samples as clustering centres.

(2) The samples of the sample set are assigned to the n clustering centres according to the principle of minimum distance. The minimum distance principle means that, if $\|x - l_j(m+1)\| < \|x - l_i(m+1)\|, i, j = 1, 2, \dots, k, i \neq j$, then $x \in f_j(m)$, $f_j(m)$ denotes the cluster domain represented by the jth cluster centre during this iteration.

(3) The clustering centres calculated from step (2), i.e.:

$$z_i(m+1) = \frac{1}{N_i} \sum_{x \in f_i(m)} x, i = 1, 2, \dots, k \quad (30)$$

where N_i is the number of samples in the i-th clustering domain $f_i(m)$. Its mean vector is used as the new clustering centre such that the error sum of squares criterion function:

$$J = \sum_{x \in f_i(m)} \|x - z_j(m+1)\|^2, i = 1, 2, \dots, k \quad (31)$$

reaches minimum value.

(4) If $Z_i(m+1) = Z_i(m)$, the algorithm converges, the calculation is complete; otherwise return to step (2) for the next iteration.

K-mean clustering of the voltage values of the five groups of batteries according to the above steps gave the following results.

Table 2 Batteries' voltages and k-means clustering voltages

Battery 1	Battery 2	Battery 3	Battery 4	Battery 5	K average vaule
4.1984	4.1994	4.1994	4.1991	4.1992	4.1991
4.1986	4.1987	4.1987	4.1984	4.1989	4.1987
4.1983	4.1984	4.1984	4.1986	4.1987	4.1985
4.1980	4.1984	4.1981	4.1980	4.1984	4.1982
4.1978	4.1981	4.1981	4.1981	4.1984	4.1982

4.2.3 Model construction and parameter search

1. Model construction

The model first needs to be discretized and the transfer function of the n-RC model is

$$G(s) = \frac{U_t(s) - U_{OC}(s)}{i_L(s)} = -(R_i + \frac{R_{D1}}{1 + R_{D1}C_{D1}s} + \frac{R_{D2}}{1 + R_{D2}C_{D2}s} + \dots + \frac{R_{Dn}}{1 + R_{Dn}C_{Dn}s}) \quad (32)$$

Let $E_L(s) = U_t(s) - U_{OC}(s)$, then:

$$E_L(s) = -i_L(s)(R_i + \frac{R_{D1}}{1 + R_{D1}C_{D1}s} + \frac{R_{D2}}{1 + R_{D2}C_{D2}s} + \dots + \frac{R_{Dn}}{1 + R_{Dn}C_{Dn}s}) \quad (33)$$

Mapping relations from the s-plane to the z-plane:

$$s = \frac{2}{\Delta t} \frac{1 - z^{-1}}{1 + z^{-1}} \quad (34)$$

where Δt is sampling interval.

The equation based on the z-plane is:

$$G(z^{-1}) = \frac{c_{n+1} + \dots + c_{2n+1}z^{-2n+1}}{1 - c_1z^{-1} - \dots - c_nz^{-n}} \quad (35)$$

where is the custom model coefficient.

Eq. 34 can be expressed in discrete time as:

$$E_{L,k} = c_1 E_{L,k-1} + \dots + c_n E_{L,k-n} + c_{n+1} i_{L,k-n} \quad (36)$$

Define $U_{OC,k}$ at moment k as a function of the SOC value z_k , the temperature value T_k and the aging state $A_{ge,k}$ as:

$$U_{OC,k} = f(z_k, T_k, A_{ge,k}) \quad (37)$$

Then the derivative of $U_{OC,k}$ with respect to time can be expressed as:

$$\frac{dU_{OC}}{dt} = \frac{\partial U_{OC}}{\partial z} \frac{dz}{dt} + \frac{\partial U_{OC}}{\partial T} \frac{dT}{dt} + \frac{\partial U_{OC}}{\partial A_{ge}} \frac{dA_{ge}}{dt} \quad (38)$$

We assume that a small change in the maximum battery capacity per unit sampling interval Δt has no effect on SOC, i.e. $dz/dt \approx 0$. Then the effect of power battery charge variation on SOC over Δt is:

$$\frac{dz}{dt} = \frac{\eta_i i_L \times \Delta t}{C_{\max} \times 3600} \quad (39)$$

Assuming that the operating current of the power battery is fixed at 1C, the current maximum available capacity is C_{max} , the sampling interval Δt is 1s, and the charge/discharge multiplier efficiency of the power battery η_i is 1, the above equation can be calculated as:

$$\frac{dz}{dt} = \frac{1 \times C_{max} \times 1}{C_{max} \times 3600} = \frac{1}{3600} < 0.03\% \quad (40)$$

The capacity decay of a power battery is a long-term process, so we determined that the battery did not have an ageing problem during the unit sampling time. Therefore, equation 40 can be simplified as:

$$\frac{dU_{OC}}{dt} = 0 \quad (41)$$

In the discrete time domain:

$$\Delta U_{OC,k} = U_{OC,k} - U_{OC,k-1} \approx 0 \quad (42)$$

Then:

$$U_{t,k} = (1 - \sum_{i=1}^n c_i) U_{OC,k} + c_1 U_{t,k-1} + c_2 U_{t,k-2} + \dots + c_n U_{t,k-n} + c_{n+1} i_{L,k} + c_{n+2} i_{L,k-1} + \dots + c_{2n+1} i_{L,k-n} \quad (43)$$

Define the data and parameter matrices of the system as:

$$\begin{cases} \Phi_{n,k} = [1 & U_{t,k-1} & \dots & U_{t,k-n} & i_{L,k-1} & i_{L,k-2} & \dots & i_{L,k-n}] \\ \Theta_{n,k} = [(1 - \sum_{i=1}^n c_i) U_{OC,k} & c_1 & c_2 & c_3 & \dots & c_{2n+1}]^T \end{cases} \quad (44)$$

Then:

$$y_k = \Phi_{n,k} \Theta_{n,k} \quad (45)$$

Based on the results of the parameter identification, the OCV can be resolved as:

$$U_{OC,k} = \frac{U_{t,k}}{1 - \sum_{i=1}^n c_i} \quad (46)$$

2. Off-line parameter identification

The main function of off-line parameter identification is to determine the variation of the individual resistor-capacitance values of the RC equivalent model during the charging and discharging process.

The steps to achieve this are as follows:

(1) Determine the parameter update region

Determine the range of SOC (x, y) from the experimental data.

(2) Power battery model parameter identification

Step 1: Calculate the interval of the SOC involved in this parameter identification. The specific calculation equation is as follows:

$$[z_s + (m-1) \times 0.005z_s + (m-1) \times 0.005 + \partial] (47)$$

Step 2: The parameters of the different order equivalent models are identified according to the established RC equivalent circuit model.

Step 3: The third step is to determine the reliability of the parameter identification results according to R^2 . When $R^2 > a$, we consider that the result meets the requirements and the result is retained; when $R^2 < a$, the result does not meet the requirements and we need to add more data to re-identify.

(3) Model parameter evaluation and optimisation

The accuracy of the model is judged by the voltage results output from the circuit simulation. An error threshold b is set and if the error does not exceed the threshold, the model is considered to be accurate, if the error exceeds the threshold, the model is not considered to be accurate.

(4) Model order determination and system parameter update

The determination of the model order is based on the AIC, and since the model order n is much smaller than the data length L_D and the mean value of the squared residuals $\hat{s}_k^2 < 1V^2$ under the optimal parameters, the AIC criterion can be simplified as:

$$AIC = 2 \log \hat{s}_k^2 + 2n (48)$$

where $\hat{s}_k^2$ is calculated as follows:

$$\hat{s}_k^2 = \frac{1}{L_D} \sum_{k=1}^{L_D} (y_k - \hat{y}_k)^2 \quad (49)$$

where $\hat{y}_k$ is the estimated terminal voltage of the model.

3. Parameter identification results

Figure 14 and 15 show the simulation results of power battery model parameter identification and open circuit voltage identification results respectively.

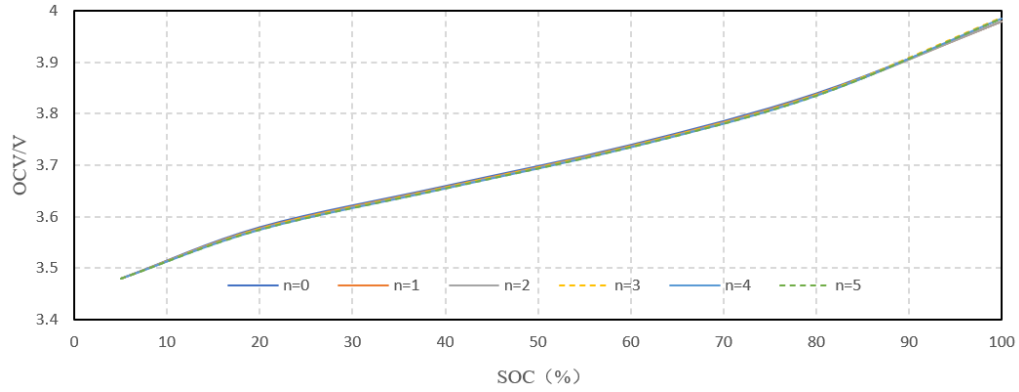


Figure 14 Open circuit voltage identification results

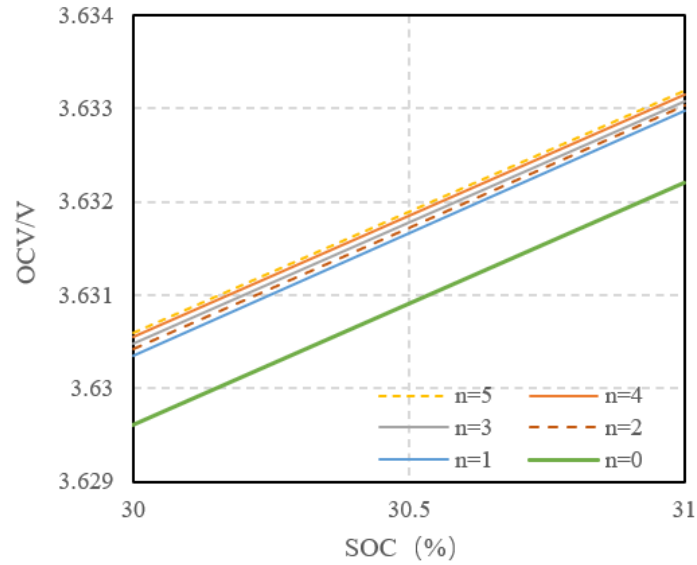


Figure 15 Detail view of identification results

In order to evaluate each order model, a suitable order model is selected. The analysis of the RC equivalent circuit is carried out in terms of four aspects: evaluation coefficients, maximum error, evaluation error and fitting calculation time.

Table 3 Analysis of the absolute value of the voltage prediction error

Number of RC networks	R ²	Maximum error/mV	Average error/mV	Time/s
n=0	0.926	65.786	1.701	0.61
n=1	0.991	19.002	0.116	1.12
n=2	0.992	18.508	0.114	4.56
n=3	0.992	18.131	0.112	10.13
n=4	0.993	19.115	0.112	15.32
n=5	0.991	19.079	0.111	27.21

From Figures 14 and 15 it can be obtained that in the range 0-5, the order of the RC network has very little effect on the estimation results. However, the maximum error in the prediction results of the Rint model reaches nearly 66mV, which is too different from the other orders to meet expectations. However, the higher the order, the more the estimation error is not significantly reduced, but rather the computational cost increases leading to an increase in computation time.

Table 4 AIC evaluation results

Number of RC networks	n=1	n=2	n=3	n=4	n=5
AIC	2.73	10.46	14.68	17.79	20.53

Combining prediction accuracy and model evaluation, n=2, i.e. the DP circuit, was chosen as the equivalent model.

4. Status estimation

The performance of the model-based approach is dependent on the model correctness and the state estimation algorithm to quantify it, according to the previous section. The extended Kalman filter algorithm is employed in this paper to handle state estimation.

For a non-linear discrete system with $f(x_{k-1}, u_{k-1})$ as the equation of the system state function and $h(x_k, u_k)$ as the function of the system observation equation, the general form of the observation equation of the state equation is:

$$\begin{cases} x_k = f(x_{k-1}, u_{k-1}) + \omega_{k-1} \\ y_k = h(x_k, u_k) + v_k \end{cases} \quad (50).$$

Where x represents the system's state, u its input, y its output, ω_{k-1} its white noise with covariance Q_k , v_k its sequence of measured white noise with mean 0 and covariance R_k , and ω and v are independent of one another.

At each moment, linearize $f(x_k, u_k)$ and $h(x_k, u_k)$ with a first-order Taylor expansion, i.e.

$$\begin{cases} f(x_k, u_k) \approx f(\hat{x}_k, u_k) + \left. \frac{\partial f(x_k, u_k)}{\partial x_k} \right|_{x_k=\hat{x}_k} (x_k - \hat{x}_k) \\ h(x_k, u_k) \approx h(\hat{x}_k, u_k) + \left. \frac{\partial h(x_k, u_k)}{\partial x_k} \right|_{x_k=\hat{x}_k} (x_k - \hat{x}_k) \end{cases} \quad (51)$$

where $\hat{x}_k$ is the estimated value of x_k . Define $A_k \triangleq \left. \frac{\partial f(x_k, u_k)}{\partial x_k} \right|_{x_k=\hat{x}_k}$, $C_k \triangleq \left. \frac{\partial h(x_k, u_k)}{\partial x_k} \right|_{x_k=\hat{x}_k}$. Bringing equation 50 into equation 51 yields the linearised system equation of state and the observation. The equations are:

$$\begin{cases} x_k \approx A_{k-1}x_{k-1} + [f(\hat{x}_{k-1}, u_{k-1}) - A_{k-1}\hat{x}_{k-1}] + \omega_{k-1} \\ y_k \approx C_k x_k + [h(\hat{x}_k, u_k) - C_k \hat{x}_k] + v_k \end{cases} \quad (52)$$

Based on the above system equations, the estimation process is established:

- (i) Initialisation: Sets the initial value of the observer: x_0, P_0, Q_0, R_0 .
- (ii) A priori estimation-prediction: state from time $(k)^-$ to time $(k)^+$: the system state prediction and error covariance prediction update equations are as follows:

$$\begin{cases} \hat{x}_k^- = f(\hat{x}_{k-1}, u_{k-1}) \\ P_k^- = A_{k-1}P_{k-1}A_{k-1}^T + Q_{k-1} \end{cases} \quad (53)$$

- (iii) Posterior estimation-correction: state from time $(k)^-$ to time $(k)^+$: This step corrects the state estimate and covariance estimate with the measurement y_k at moment k . The estimates are expressed as $\hat{x}_k^+$ and P_k^+ , respectively, and the measurement update equation is as follows:

$$\text{New interest matrix: } e_k = y_k - h(\hat{x}_k^-, u_k)$$

Kalman gain matrix: $K_k = P_k^- C_k^T (C_k P_k^- C_k^T)^{-1}$

$$H_k = \frac{1}{M} \sum_{i=k-M+1}^k e_i e_i^T$$

Adaptive noise covariance matching: $R_k = H_k - C_k P_k^- C_k^T$

$$Q_k = K_k H_k K_k^T$$

System status correction: $\hat{x}_k^+ = \hat{x}_k^- + K_k e_k$

Error covariance correction: $P_k^+ = (I - K_k C_k) P_k^-$

(iv) Time scale update: the state and covariance matrix at moment $(k)^+$ is used as output to prepare the state estimate at moment $k+1$.

5. Error compensation

The data driven method has a higher accuracy of the results but does not take into account the impact of temperature and capacity decay on the SOC estimation results. The Kalman filter based method treats the output error as noise to compensate for the error in the initial value of the ATS integration, but the final accuracy is not ideal. In order to further reduce the SOC estimation error by using more precise terminal voltages and currents, the benefits of these two methods can be combined. The output data from the model-based method can then be used as the training data for the SVR for a second estimation. As inputs, the filtered current, end voltage, and SOC values are chosen. As outputs, the second projected SOC and error are chosen. Finally, the validation data is utilized to check the model's accuracy. The exact process is shown in Figure 16.

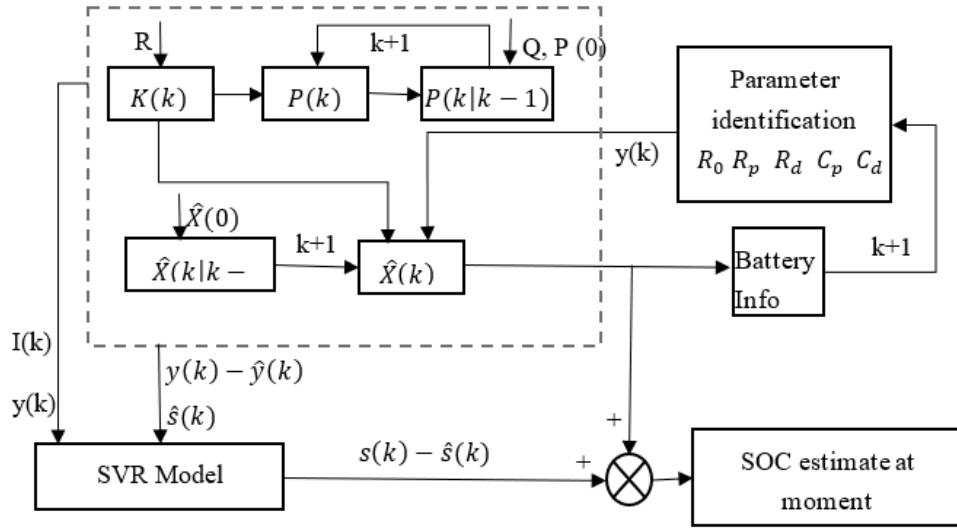


Figure 16 Error compensation flow chart

4.2.4 Model reconstruction based on prediction errors

Based on the changes in battery performance under very low temperature and deep ageing conditions discussed previously. At very low temperatures, the activity of the active particles inside the battery decreases, the movement of the transport particles slows down, the chemical reaction rate decreases, the capacity loss of the power battery increases and the maximum available capacity decreases. As the number of battery cycles increases, the number of active particles decreases and the electrode active material melts, resulting in increased impurities in the electrolyte and increased impedance within the battery, further leading to a decrease in battery performance. The effect of all of the above issues plus the polarisation of the battery on the model is that the data distribution of the original data set used for training is significantly different from the current data distribution, and the effect of the support vector set regression algorithm is highly dependent on the distribution of the data set, so if the original trained model is used to test the new data it will reduce the prediction effect. Therefore, the model needs to be re-tuned, and a model reconstruction based on the prediction error is proposed to solve the problem of increasing error due to the change in the distribution of the data set.

Model refactoring requires an event to trigger it, when the error in the model prediction exceeds a tolerable upper limit. When the prediction error exceeds a threshold, it means that the training model at this point is no longer applicable to the data distribution at this moment in time, and the model needs to be retrained using data containing the dataset at this moment in time to improve the accuracy of the prediction. The effects of temperature and capacity decay on battery characteristics discussed in Chapter 2 reveal that the larger part of the data variation exists at the moment when the battery SOC drops below 20%. Therefore, the prediction reliability of the model is high at the stage where the SOC is greater than 20% and decreases at the stage where the SOC is less than 20%, at which point it is necessary to retrain the model by including data where the SOC is less than 20%. In practice, we consider that when the prediction error is greater than 0.03 then the prediction error is considered too large and the model needs to be reconstructed.

4.2.5 Model reliability validation

1. Introduction to Simscape

Simscape is a modeling and simulation platform enhanced by MATLAB, which is based on a variety of physical domains, including mechanical, thermal fields, physical signals and more. It provides users with base modules and components for each physical domain, which can be experimented with to build some complex components or devices. Because Simscape components use physical connections to each other, the models are matched to the structure of the system being developed. Features are as follows.

- (1) Simscape can be used to model and simulate a variety of physical domain systems, such as electrical and mechanical, but only a single physical domain system can be built, and multiple physical domain crossovers are not supported.
- (2) Includes a library of physical modules and a library of mathematical operations, which can be used to build and develop custom modules.
- (3) Can use Simscape language for custom component and module design, and can expand physical modelling components, domains and libraries.

(4) Physical units for parameters and variables are provided and unit conversions are automatically performed.

(5) Models for simulation include modules provided by relevant physical modelling tools without the need to purchase these modelling tools.

(6) C code generation is supported.

Main features:

(1) Physical system modelling

System modelling using Simscape is built in the same way as the actual physical model. Simscape uses a non-causal modelling approach for model construction: the individual physical domain modules are equivalent to the corresponding physical components, such as operational amplifiers, current generators, AC power supplies, etc. The connections between modules are equivalent to physical connections, i.e. physical parameter transfer. This building model reflects the physical structure of the system rather than the mathematical relationships between the modules of the system, so that the model built is highly compatible with the schematic. In addition Simscape can automatically construct differential algebraic equations (DAE) that describe the physical properties of the system based on the model built.

(2) Component library

The Simscape component library includes a variety of basic physics modules in the physical domain that can be used. The main focus here is on resistive components and physical signals. Some basic electronic circuit components, such as resistors, capacitors and inductors, and some more complex components such as amplifiers and transformers. SimElectronics™. The physical signals in the Simscape model require units to be determined when setting up the module parameters and Simscape automatically converts the units when solving the physical model. By automatically converting units, Simscape can perform mathematical operations on physical signals without considering units, increasing the efficiency of the operations while retaining the ability to describe the physical relationships of the system.

(3) Simscape language

The Simscape language is based on the MATLAB programming language and allows the programmer to customise components according to the MATLAB programming language rules. The customisation includes component parameters, ports, equations of state, dynamics equations, etc. The custom components allow for dynamic acquisition and updating of parameters. Once a standardised custom component has been created, it is possible to adjust different parameters to achieve different representations of the same component, so there is no need to repeat the definition.

The creation of custom components allows for the extension of the relevant physical domain and the generation of C-code for Simulink models containing custom components. By using the Simscape language, it is possible to customise the various influences on the main physical characteristics of the physical components created. This allows a trade-off between model fidelity and simulation speed.

2. Simscape model building

Figure 17 shows the Simscape circuit simulation model with the custom physical components connected by blue lines and the modules under Power System connected by black lines. The data for the custom modules needs to be loaded in the workspace in advance and the units set in advance. It is important to note that the Simscape model must be constructed with the Solver Configuration module, whose output can be connected to any part of the circuit. In addition, unlike Simulink modelling, the Electrical Reference module must be added to the circuit, the main purpose of which is to establish a reference point. The PS-Simulink Converter is used to convert the physical model signal into the simulated model signal.

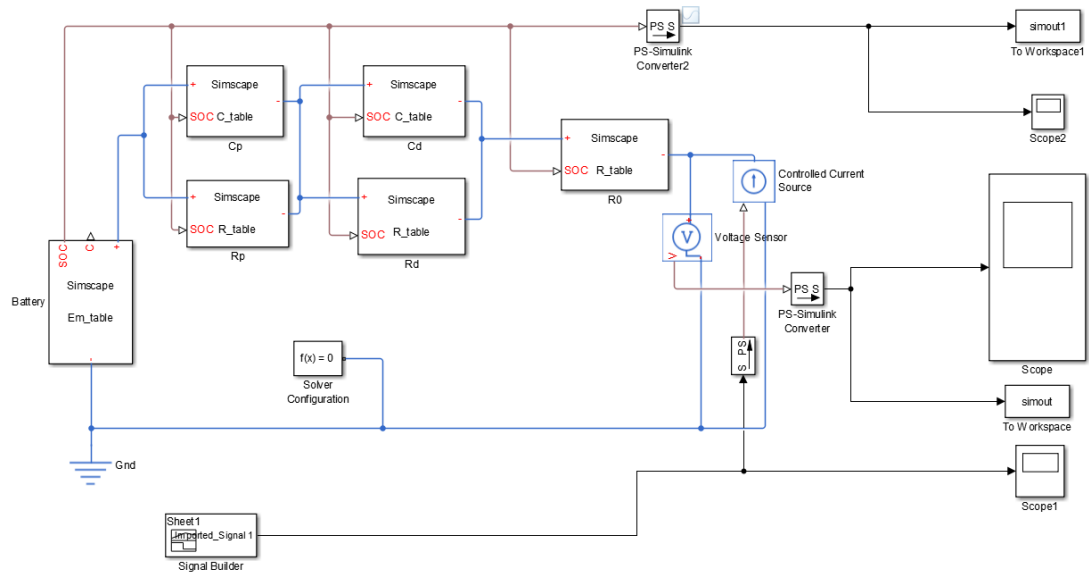


Figure 17 Simscape circuit simulation model

The main components of which custom creation are shown below.

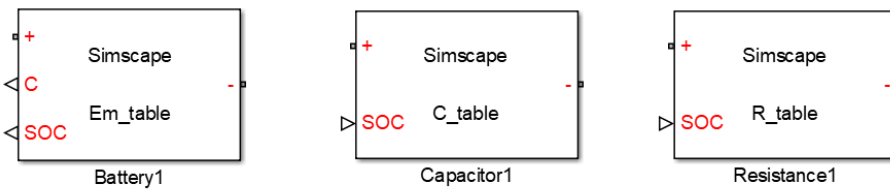


Figure 18 Created circuit components

Some of the codes are as follows:

nodes

p = foundation.electrical.electrical; % +:left

n = foundation.electrical.electrical; % -:right

end

equations

v == p.v - n.v;

% Capacity table lookup

C == tablelookup(SOC_Table,C_Table,SOC,...

interpolation=linear,extrapolation=nearest)

SOC == 1 - Qe/C;

```

% Electrical equation by table lookup

v == tablelookup(SOC_Table,Em_Table,SOC,...
interpolation=linear,extrapolation=nearest)

let

R = tablelookup(SOC_Table,R_Table,SOC,...
interpolation=linear,extrapolation=nearest)

in

v == i*R;

let

C = tablelookup(SOC_Table,C_Table,SOC,...
interpolation=linear,extrapolation=nearest)

in

% Electrical equation

i == C * v.der;

```

where equations represent the kinetic equations of the components. The values for each parameter of the component are obtained by looking up the table.

3. Model parameter estimation

The aim of parameter estimation is to obtain the values of the various capacitive and resistive components in the equivalent circuit to simulate the variation of the characteristic quantities of the battery during charging and discharging. The more accurate the values obtained for the components, the more accurate the model will be.

Firstly, in the circuit model, R_0 , R_p , R_d , C_p , C_d are not fixed and the values of these resistors and capacitors change dynamically during the charging and discharging process. Therefore, this paper uses Parameter Estimation to automatically optimise the parameters for the Simscape model of the battery until the simulation output matches the experimental results exactly. The current used in the simulation is the HPPC compound pulse current, as shown in Figure 19 for the test signal current used in the simulation.

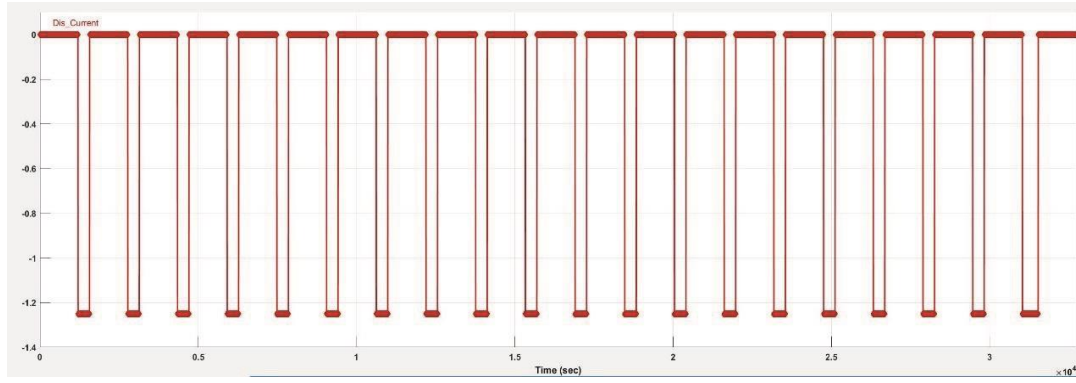


Figure 19 Current signal generators

Figure 20 shows the parameter estimation process. First, the HPPC composite pulse current is input, and the initial parameters are input according to the battery model to determine the choice of parameter estimation results through the set simulation output voltage and the collected voltage error threshold. When the error is less than the determined threshold, we believe that the parameter estimation result is more accurate. Otherwise, it is brought back into the model for calculation until the result error reaches the expected value.

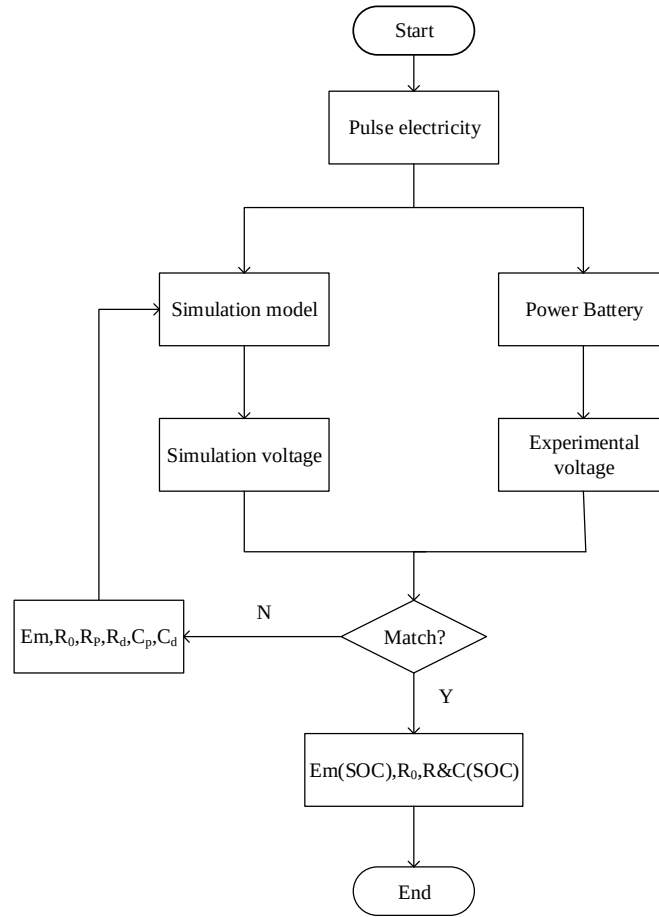


Figure 20 Flowchart of parameter estimation

4. Model Accuracy Verification

Figure 21 shows the result of using only the dual-polarization equivalent circuit algorithm. Due to the polarization reaction at the end of the discharge, the error greatly increases, and the maximum voltage error reaches 0.215V.

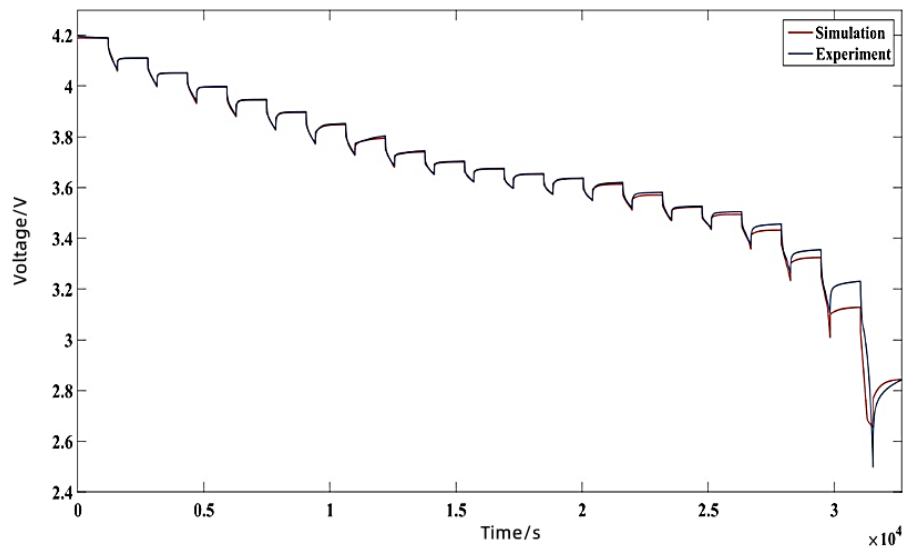


Figure 21 Results of the model-based algorithm voltage fitting

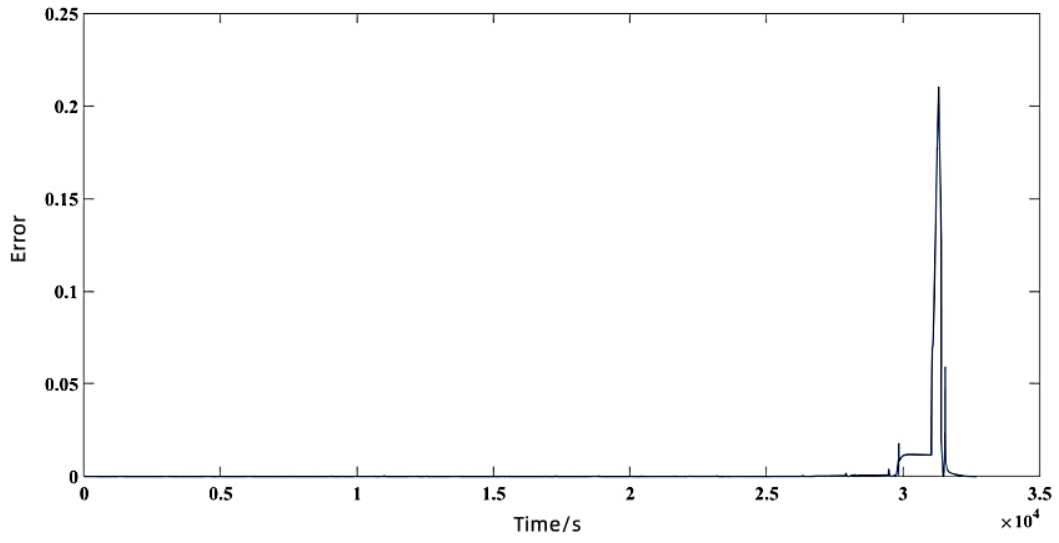


Figure 22 Voltage fitting error

Figure 23 shows the results of the voltage output after the improved algorithm, and it can be seen that the estimation accuracy has improved, and although there is still a large error in the polarisation response phase, the error value is greatly reduced to less than 0.06V.

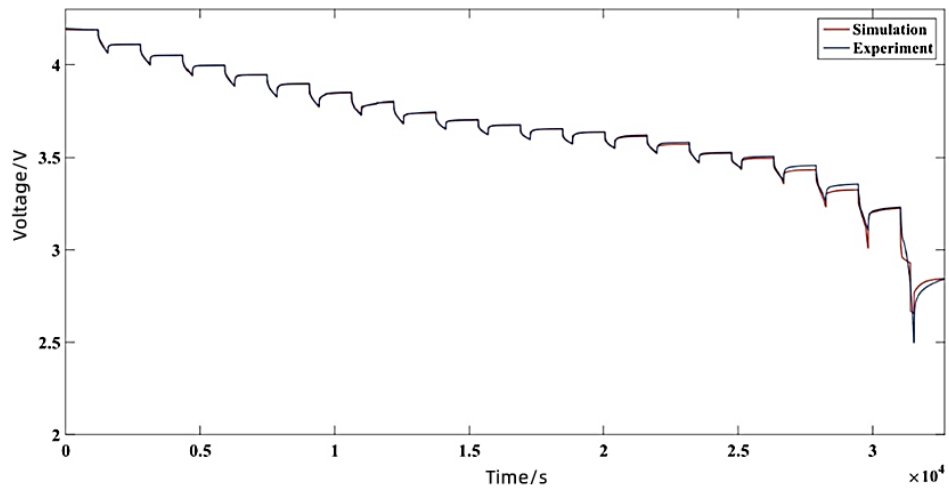


Figure 23 Improved algorithm voltage fitting results

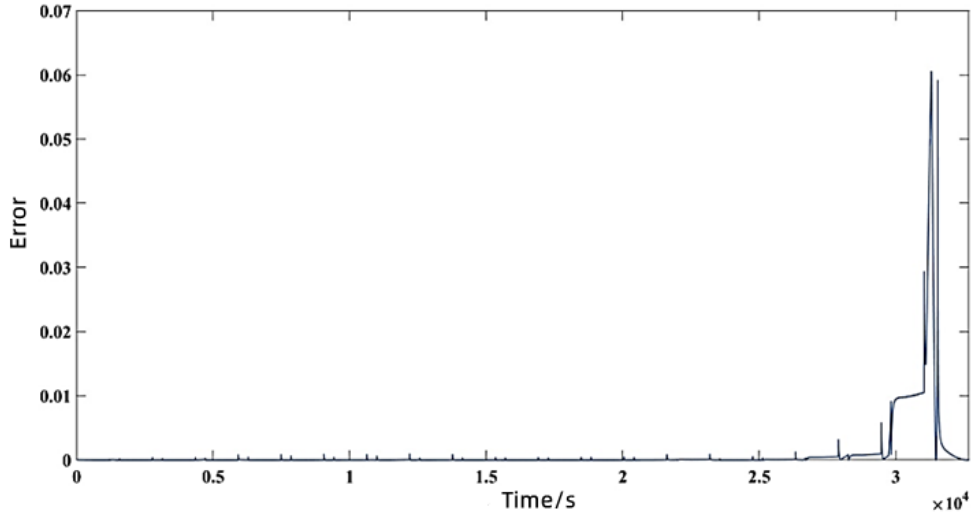


Figure 24 Voltage fitting error

The results of the parameter estimation for each of these resistive elements are shown in Table 5:

Table 5 Estimation results for each resistance parameter of the Simscape model

SOC	R_o/Ω	R_d/Ω	R_p/Ω
1.00	0.024109	0.0039032	0.0095040
0.95	0.024048	0.0037528	0.0120000
0.90	0.024903	0.0037992	0.0156243
0.85	0.024903	0.0051960	0.0237440
0.80	0.024903	0.0065664	0.0240560
0.75	0.024414	0.0005040	0.0079888
0.70	0.024598	0.0064430	0.0108880
0.65	0.024598	0.0063120	0.0108880
0.60	0.023987	0.0063928	0.0114480
0.55	0.023438	0.0058440	0.0125680
0.50	0.023560	0.0073248	0.0146960
0.45	0.023682	0.0089840	0.0176560
0.40	0.023926	0.0072150	0.0148240
0.35	0.023926	0.0073600	0.0158720
0.30	0.024292	0.0066136	0.0098160
0.25	0.025452	0.0061272	0.0178320
0.20	0.027405	0.0146000	0.0211360
0.15	0.030213	0.0176160	0.0234640
0.10	0.036133	0.0264160	0.0321520
0.05	0.057872	0.0692800	0.0724240

The results of the parameter estimates for capacitive components and voltages are shown in Table 6:

Table 6 Estimation results for each capacitance and voltage parameter of the Simscape model

<i>SOC</i>	<i>Em/V</i>	<i>C_d/F</i>	<i>C_p/F</i>
1.00	4.1903560	2808.726181	29945.28620
0.95	4.1117720	4961.812421	25200.37200
0.90	4.0536350	3569.943803	27726.57450
0.85	3.9991610	1848.093042	12622.13612
0.80	3.9488060	987.7796588	13244.09711
0.75	3.9002820	335.7490711	25047.56659
0.70	3.8535900	821.3614337	80281.04335
0.65	3.8046080	1524.302673	69406.68626
0.60	3.7463190	2218.590686	47388.19008
0.55	3.7054250	1422.749459	25238.70146
0.50	3.6771950	979.8766662	21005.71584
0.45	3.6561380	689.1767751	19953.55686
0.40	3.6392000	407.4298830	19866.43281
0.35	3.6213470	807.7051684	18019.15323
0.30	3.5821320	1463.224940	40617.35941
0.25	3.5281150	1953.667154	25039.25527
0.20	3.5061420	691.8722364	19336.67676
0.15	3.4574650	655.1402678	19757.92704
0.10	3.3564510	473.5162193	15044.16522
0.05	3.2320900	130.9718193	8610.405390

The results of the combined algorithm and the traditional EKF algorithm for SOC estimation were compared. From Figure 25 and Figure 26, it can be seen that the accuracy of the prediction of the algorithm in this paper has improved significantly compared to the EKF algorithm, with the maximum error not exceeding 0.015.

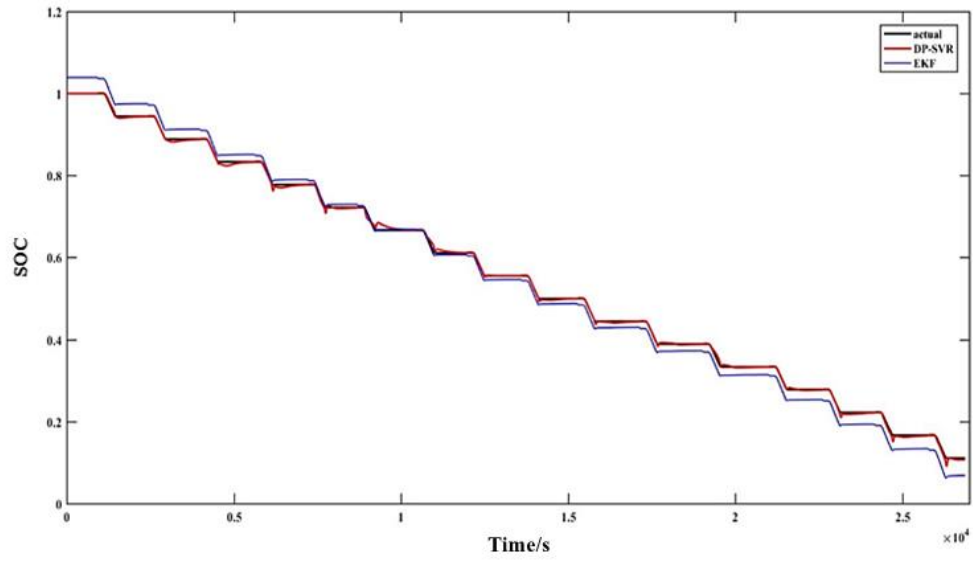


Figure 25 Comparison of the combined algorithm SOC estimation results with the EKF algorithm

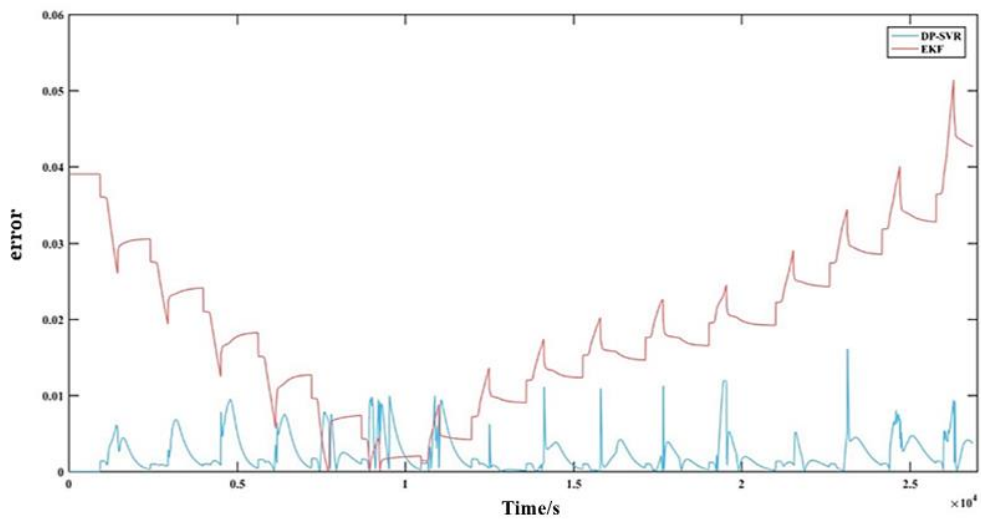


Figure 26 Two algorithms for estimating error

The algorithm error results are shown in Table 7:

Table 7 Comparison of the errors of the two algorithms

Algorithms	Mean absolute error	Root Mean Square Error
EKF	0.0193	0.0227
DP-SVR	0.0062	0.0073

From the error comparison results, it can be seen that the method used in this paper reduces the average absolute error by about 0.013 and the root mean square error by about 0.015 compared to the extended Kalman filter algorithm, indicating that the accuracy of the algorithm in this paper has been significantly improved compared to the EKF.

4.3 Analysis of experimental results

4.3.1 Predicted effects at different ambient temperatures

The algorithm is evaluated using experimental test data of Li-ion batteries under varied temperature circumstances in order to confirm the adaptability of the suggested approach at various temperatures. The one-to-one correspondence between the SOC to be estimated and the input data established in this section is $R = \{(x_1, SOC_1), (x_2, SOC_2), \dots, (x_n, SOC_n)\}$, where SOC_k is the actual value of SOC at time k ; x is the input vector at time k . In this method is defined as $x = [V(k), I(k), T(k)]$, where $V(k)$, $I(k)$ and $T(k)$ are the voltage, current and temperature values at moment k respectively. The test results are shown in the following Figure 27.

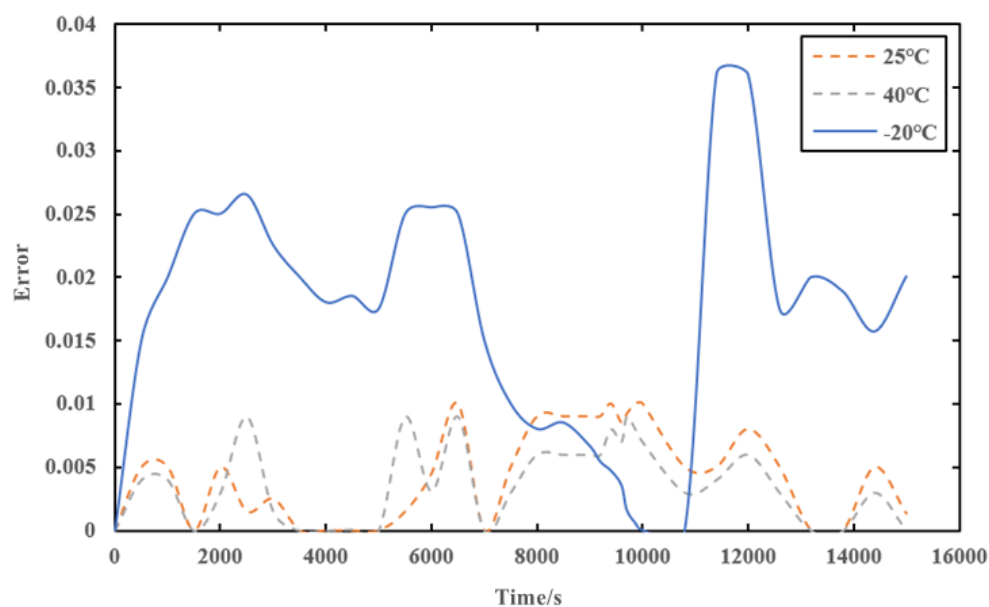


Figure 27 Prediction error curves for three temperature conditions

The results of the prediction errors at the three temperatures are shown in Table 8.

Table 8 Comparison of power battery SOC prediction errors at 25°C, 40°C and -20°C

Temperature	Average error	Root mean square error
25°C	0.0046	0.0060
40°C	0.0039	0.0050
-20°C	0.0155	0.0184

A comparison of the power battery SOC prediction errors at 25°C, 40°C and -20°C shows that the average error in the predictions of the algorithms in this paper is within 0.005 at both 25°C and 40°C ambient temperatures, but reaches 0.015 at very low temperatures which is consistent with the theoretical situation discussed above.

From the experimental results, it can be seen that the SOC estimation at -20°C shows a larger error. The reason for this is the reduced activity of the particles inside the battery at very low temperatures, which results in slower migration rates and lower chemical reaction rates, leading to higher internal resistance. Therefore, low temperature environments are not suitable for the use of power batteries as a power source for electric vehicles.

4.3.2 Predicted effects under different capacity decay scenarios

The model was trained and validated using experimentally measured ageing battery test data based on the SOC estimation results for multiple cycle batteries at 25°C temperature point as shown in Figure 28 in order to verify the SOC estimation performance of the proposed method in this paper for power batteries after multiple cycles. For testing, batteries with 200 cycles and 600 cycles were chosen.

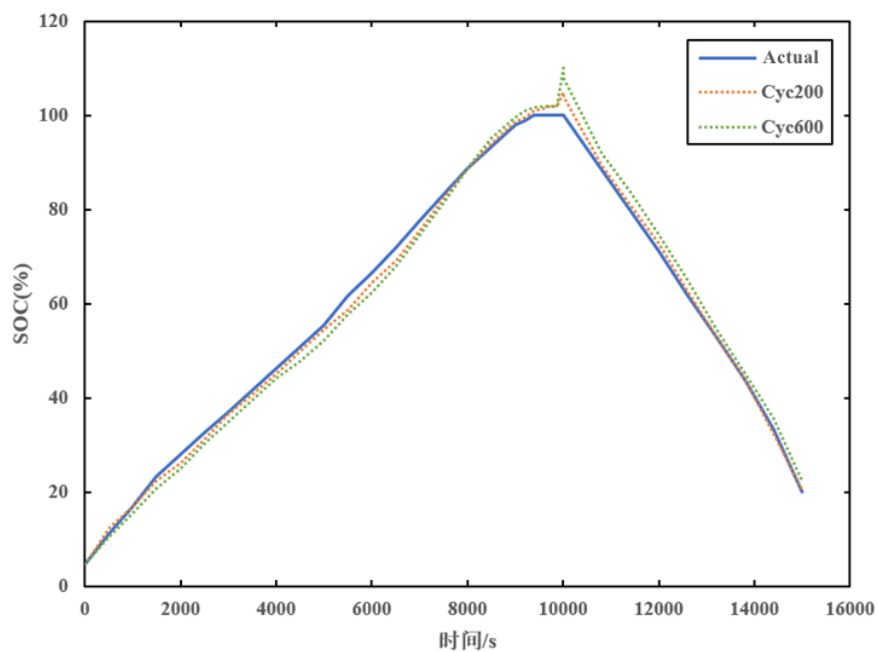


Figure 28 SOC static test curve after 200 and 600 cycles

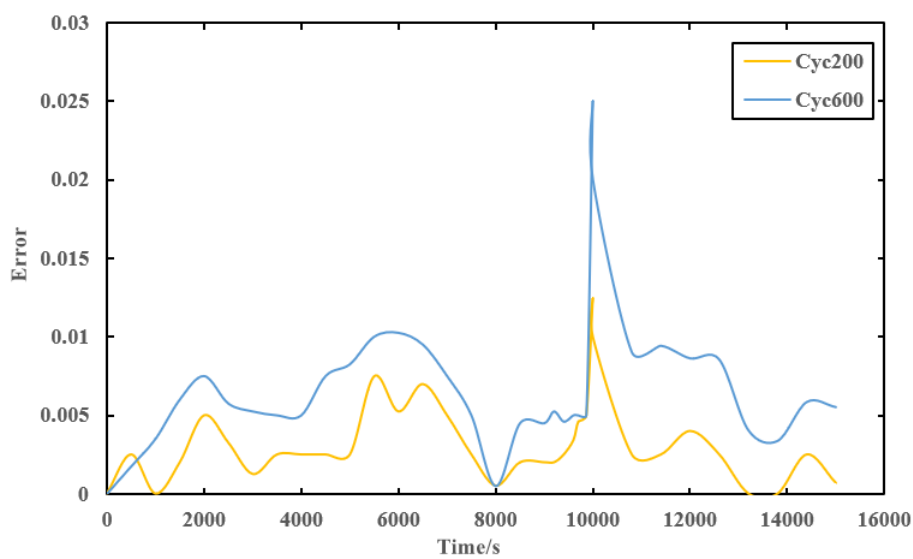


Figure 29 Error comparison curve

The results of the error comparison are shown in Table 9:

Table 9 Comparison of power battery SOC prediction errors for 200 and 600 cycles

Number of cycles	Average error	Root mean square error
200	0.0035	0.0037
600	0.0073	0.0066

The algorithm estimates the battery's SOC after several cycles with a maximum error of no more than 0.01 as shown in Figures 28 and 29. For the reasons outlined in the discussion above, Figure 28 shows that the highest inaccuracy happens when the battery is depleted after it has finished charging and is at rest. After several cycles, the battery's internal resistance increases, lowering its discharge capacity. Additionally, the concentration of active particles inside the battery decreases, and the amount of impurities in the electrolyte increases, causing a passivation film to form on the electrode's surface. This causes the battery to reach the cut-off voltage more quickly, which causes a decline in battery performance.

4.4 Model Summary

For effective battery management, an accurate evaluation of the battery SOC state is crucial. This chapter studies the estimating approach that combines the data-driven and equivalent circuit model algorithms on the basis of the research of battery modeling and the AEKF algorithm. The chapter proposes a combined algorithm based on DP-SVR by fully utilizing the excellent characteristics of the data-driven algorithm in terms of state prediction accuracy, robustness, and generalization in order to address the issue that the filtering-only algorithm cannot consider multiple influencing factors in SOC estimation and to further improve the battery SOC estimation performance based on the equivalent model algorithm. The combined algorithm enhances the estimation precision and applicability of the EKF algorithm, fully combines the benefits of the filtering algorithm and the data-driven algorithm, and further optimizes the filtering results by training the SVR model using the output data of the circuit model, based on which the SVR model output error compensation is built. The test results for various temperatures and various capacity decay data demonstrate that the proposed DP-SVR algorithm can significantly improve the comprehensive prediction of battery SOC compared to the EKF algorithm, with the maximum mean square error of 0.5% and the average error below 1%, which enhances prediction accuracy, generalization, and robustness of SOC estimation generally.

5. Design of a cloud platform for power battery management

5.1 Functional analysis

The background management system based on the cloud platform is an important part of the whole battery management system to achieve data management, so the background management system of this design should include battery status monitoring, real-time monitoring, battery system abnormal alarm, log information and other functions.

Battery status monitoring and control: used to monitor all kinds of data of the battery system.

Real-time monitoring: mainly used to monitor the change curve of SOC and power of the battery.

Data analysis functions: including user historical usage data and usage logs.

Abnormal alarm: send alarm information to users or management personnel in real time when the monitored data is in abnormal condition, so that

The relevant personnel to deal with abnormal conditions in a timely manner.

5.2 Development environment for the Power Battery Management Cloud Platform

This cloud platform adopts a front-end and back-end development model, mainly consisting of a web front-end page, a database and a scripting language for dynamic page development, which is deployed on the cloud server after the page and back-end are built. According to the specific situation of this design, Vue is chosen as the development framework for the front-end, and MySQL is chosen as the database for this design according to the characteristics of the moderate amount of data in the back-end, which is a lightweight open source database suitable for application in small and medium-sized projects; PHP is chosen as the scripting language that can be open source and cross-platform. The combination of PHP + MySQL is highly efficient, easy to install in a Windows environment, efficient in execution and mature in application.

5.3 Development model for a power battery management cloud platform

There are two main development models for Web applications, C/S and B/S. C/S is a client-side/server-side application, which needs to run independently, and B/S is a

browser-side/server-side application, which installs the specific information and procedures on the server side, and the user only needs to enter the URL in the browser to read the front-end display data on the server side. It is simple to maintain, easy to use, short development cycle, safe and reliable, etc. Based on this, this design uses the B/S development model for the development of web applications, where multiple users can access the web server at the same time to obtain the corresponding data information.

5.4 Construction of a cloud platform for power battery management

Web application building mainly includes front-end and back-end development framework design, database design, BMS device access, cloud server deployment, the following will introduce the implementation process of each module.

(1) Front-end and back-end development framework design

The front-end of this design uses the framework of Vue and is developed in the environment of Web Storm. Currently the more popular front-end development frameworks are Vue, React and so on, considering the ease of development of the author equivalent to two frameworks, this design uses the Vue framework.

The back end of this design uses the open source PHP scripting language to develop in the PHP Storm environment. The current mainstream PHP development frameworks are ThinkPHP, Yii, CodeIgniter, Lavarel, etc. Considering the scale of the project, the compatibility of the system and the ease of development, this design uses the ThinkPHP framework with an object-oriented development structure and the MVC model, which can simulate the implementation of the Struts tag library and support MySQL, MSSQL, Sqlite, MySQL, MSSQL, Sqlite, Sqlite, Sqlite, Sqlite, Sqlite, Sqlite, Sqlite, Sqlite and Sqlite. MSSQL, Sqlite, PostgreSQL, Oracle and other databases, with the advantages of simple development, high efficiency, easy expansion, rich documentation and examples, framework compatibility, etc., very suitable for the development of small and medium-sized projects.

(2) Database design

The database is the long-term memory of the Web database application, storing information in and being able to access information from the database by accepting and

processing information typed by the user in the browser window. The design uses a MySQL database to store and manage data, using the database management tool Phpmyadmin to create and insert relevant fields, SQL statements to search for data, and data files to import and export to the database. PHP includes specific statements to send commands to the MySQL server, allowing easy database operations between PHP and MySQL.

(3) Access to the BMS and deployment of the cloud server

The access to the BMS and the battery management platform is carried out through the Bluetooth transmission module. The data transmission time is set to 1s and the data is transmitted through the TCP/IP protocol.

The cloud server selected for this battery management cloud platform is AliCloud server with 1 core (vCPU) 2 GiB CPU & memory and CentOS 7.9 64-bit operating system.

5.5 Cloud Platform Functionality Showcase

1. Battery Info

Battery Info focuses on basic battery information, with two sub-pages showing basic battery status and individual battery information respectively.

(1) Battery State

This page displays the basic battery status information, including SOC, voltage, current, cycle count, BMS production date, number of NTC, equalisation status, etc. The function page is shown in Figure 30.

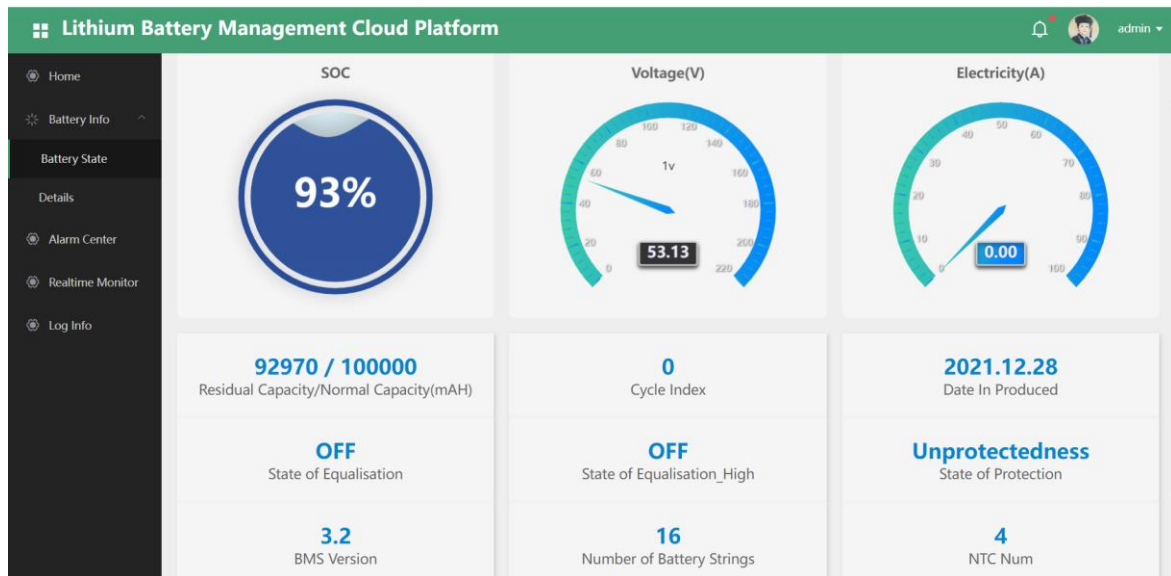


Figure 30 Battery State page

(2) Details

This page mainly displays the individual battery information, including the number of battery strings and the voltage of each string, as well as the maximum voltage, minimum voltage and maximum voltage difference in a single battery. The page is displayed as shown in Figure 31:

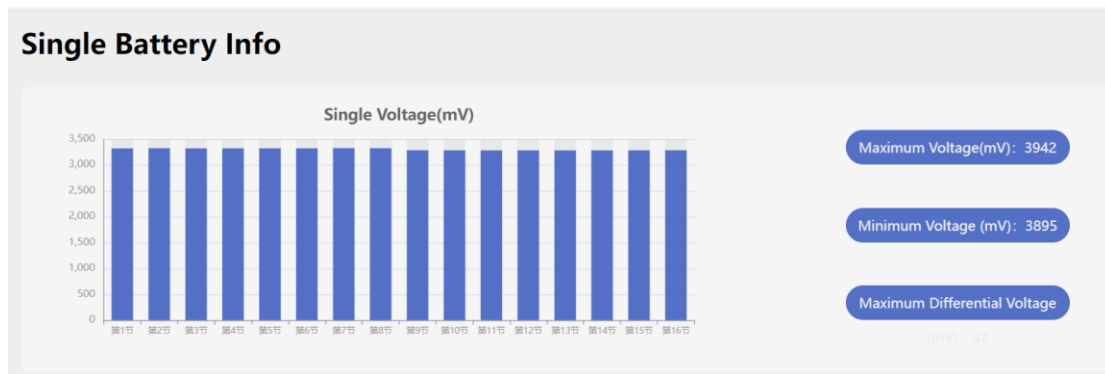


Figure 31 Details page

2. Alarm Center

This page shows the real-time battery protection status, if the protection occurs, an alarm will be generated to remind the user to check the battery to reduce the accident. The page displays the functions as shown in Figure 32 :

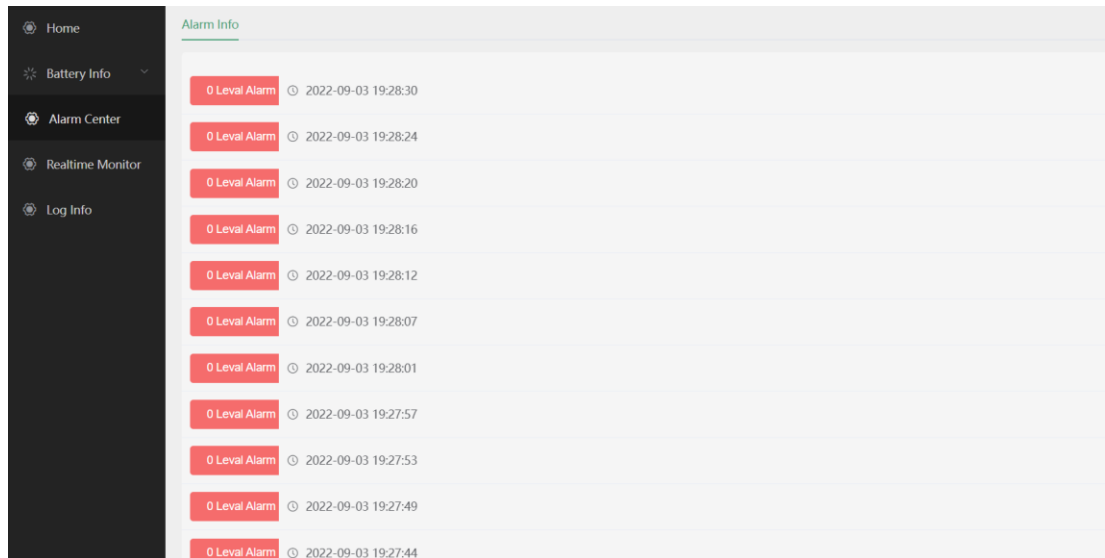


Figure 32 Alarm Center page

3. Realtime Center

This page mainly shows the real-time changes in SOC and power, and will show the SOC and power data over the past two hours in the form of line graphs, so that users can monitor them in real time. The function of the page is shown in Figure 33 (at this point the battery is idle, so the current is 0 and the SOC changes are small).

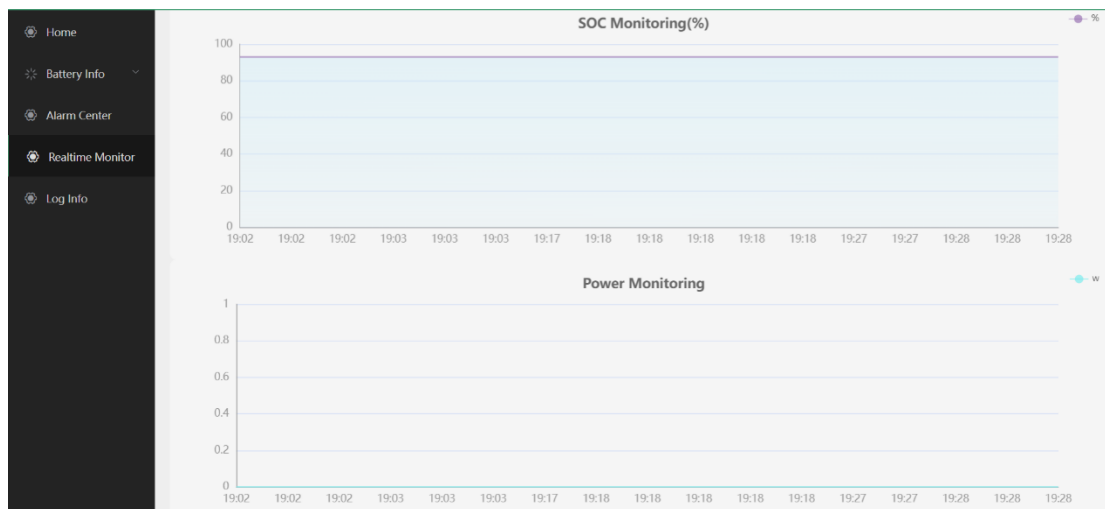


Figure 33 Realtime Center page

4. Log Info

This page is the log information display interface, mainly to extract the current, voltage, cycle times and other information for the past two hours, so that users can easily check the battery status. The page functions are displayed in Figure 34.

Home

Battery Info

Alarm Center

Realtime Monitor

Log Info

Voltage

Voltage

Cycle Times

SOC

Electricity

	Time	Data
	2022-09-03 19:28:30	53.13
	2022-09-03 19:28:24	53.13
12023	2022-09-03 19:28:20	53.13
12022	2022-09-03 19:28:16	53.13
12021	2022-09-03 19:28:12	53.13
12020	2022-09-03 19:28:07	53.13
12019	2022-09-03 19:28:01	53.13

Home

Battery Info

Alarm Center

Realtime Monitor

Log Info

ID	Time	Data
12025	2022-09-03 19:28:30	53.13
12024	2022-09-03 19:28:24	53.13
12023	2022-09-03 19:28:20	53.13
12022	2022-09-03 19:28:16	53.13
12021	2022-09-03 19:28:12	53.13
12020	2022-09-03 19:28:07	53.13
12019	2022-09-03 19:28:01	53.13
12018	2022-09-03 19:27:57	53.13
12017	2022-09-03 19:27:53	53.13
12016	2022-09-03 19:27:49	53.13
12015	2022-09-03 19:27:44	53.13
12014	2022-09-03 19:27:39	53.13
12013	2022-09-03 19:27:35	53.13
12012	2022-09-03 19:27:31	53.13

Figure 34 Log Info page

6. Full summary and reflection

6.1 Full summary

The SOC of a power battery is defined as the current remaining capacity of the battery and therefore the accuracy of the SOC estimate affects the efficiency of the battery's stored energy usage. Accurate SOC estimation can prevent overcharging and over-discharging of the battery, which in turn can extend the battery's operating time. However, in practical engineering applications, the SOC of a battery is affected by the operating environment, with factors such as temperature and capacity degradation affecting battery characteristics and making accurate SOC estimation more difficult.

This paper focuses on the effects of temperature and battery ageing on the open circuit voltage and maximum discharge capacity of batteries, and proposes improvements based on the equivalent model approach for these two problems. The data-driven and equivalent model approaches are combined to increase the accuracy and generalisability of the algorithm by complementing each other's strengths.

In order to apply the improved SOC estimation method, a cloud platform was built and deployed on a cloud server. The SOC of this cloud platform is calculated using the algorithm proposed in this paper as a method to monitor the changes in battery status and complete the health monitoring of the battery.

6.2 Reflection

Although the research on power SOC estimation in this paper has yielded preliminary results, there are still unresolved issues that can be further investigated:

- (1) This paper is based on the SOC estimation of power Li-ion single cell batteries, this paper does not investigate the problem of battery power balance and consistency after the formation of single cell batteries, which is one of the directions that can be explored in depth later. Moreover, only offline data is used for testing, and the accuracy of the algorithm under complex working conditions is not considered.
- (2) The power battery management cloud platform built in this paper is based on one BMS device, and different devices will be added for diversified management to achieve

the universality of the platform, while enhancing the platform's capacity for concurrency and building a stable and efficient cloud platform.

References

- [1] LIPU M S H, HANNAN M A, HUSSAIN A, et al. A review of state of health and remaining useful life estimation methods for lithium-ion battery in electric vehicles: Challenges and recommendations[J]. Journal Of Cleaner Production, 2018,205:115-133.
- [2] QIU X H, WU W X, WANG S F. Remaining useful life prediction of lithium-Ion battery based on improved cuckoo search particle filter and a novel state of charge estimation method[J].Journal of Power Sources, 2020,450: 227700.
- [3] BI J, ZHANG T, YU H Y. State-of-health estimation of lithium-ion battery packs in electric vehicles based on genetic resampling particle filter[J]. Applied Energy, 2016,186:558-568.
- [4] VIDAL C, MALYSZ P, KOLLMEYER P, et al. Machine learning applied to electrified vehicle battery state of charge and state of health estimation: state-of-the-art[J]. IEEE ACCESS, 2020,8:52796-52814.
- [5] TAN X J,ZHAN D,LYU P X, et al. Online state-of-health estimation of lithium-ionbattery based on dynamic parameter identification at multi timescale and support vector regression[J]. Journal of Power Sources, 2020,22233.
- [6] WEI J W, DONG G Z, CHEN Z H. Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression[J]. IEEE Transactions on Industrial Electronics, 2018,65(7): 5634-5643.
- [7] MENG J H, CAI L, LUO G Z. Lithium-ion battery state of health estimation with short-term current pulse test and support vector machine[J]. Microelectronics Reliability, 2018,88-90:1216-1220.
- [8] ZHOU D, YIN H T, FU P. Prognostics for state of health of lithium-ion batteries based onGaussian process regression[J]. Mathematical Problems in Engineering, 2018,2018.
- [9] KODJO S R M, KDDAHECH A, DUMUR D, et al. State-of-health estimators coupled to a random forest approach for lithium-ion battery aging factor ranking[J].Journal of Power Sources, 2020,229154.
- [10] PHATTARA K, NITA Y. Data-driven prognostic model of li-ion battery with deeplearning algorithm[C]. //Annual Reliability and Maintainability Symposium (RAMS). Orlando,2019:1-6.
- [11] DAI H D, ZHAO G C, LIN M Q. A novel estimation method for the state of health of lithium-Ion battery using prior knowledge-based neural network and Markovchain[J]. IEEE Transactions on Industrial Electronics,2020,66(10):7706-7716.
- [12] KIM J, YU J, KIM M, et al.Estimation of li-ion battery state of health ba

- sed onmultilayer perceptron: as an EV application[J]. IFAC PapersOnLine, 2018,51 (28):392-397.
- [13] XIA Z Y, QAHOUQ A, JABER A. Adaptive and fast state of health estimation method for lithium-ion batteries using online complex impedance and artificial neural network[C]. //Annual IEEE Applied Power Electronics Conference and Exposition (APEC), 2019:3361-3365.
 - [14] VENUGOPAL P, VIGNESWARAN T. State-of-health estimation of li-ion batteries in electric vehicle using IndRNN under variable load condition[J]. Energies, 2019,12(22): 4338.
 - [15] YOU G W, PARK S, OH D. Diagnosis of electric vehicle batteries using recurrent neural networks[J]. IEEE Transactions on Industrial Electronics, 2017,64 (6):4885-4893.
 - [16] LI P H, ZHANG Z J, XIONG Q Y, et al. State-of-health estimation and remaining useful life prediction for the lithium-ion battery based on a variant long short-term memory neural network[J]. Journal of Power Sources, 2020,459:228069.
 - [17] ZHOU Y T, HUANG Y N, PANG J B, et al. Remaining useful life prediction for supercapacitor based on long short-term memory neural network[J]. Journal of Power Source, 2019,440:227149.
 - [18] TAN Y D, ZHAO G C. Transfer learning with long short-term memory network for state-of-health prediction of lithium-ion batteries[J]. IEEE Transactions on Industrial Electronics, 2020,67(10):8723-8731.
 - [19] CROCIONI G, PAU D, DELORME J M, et al. Li-ion batteries parameter estimation with tiny neural networks embedded on intelligent IoT microcontrollers[J]. IEEE ACCESS, 2020,8:122135-122146.
 - [20] ZHOU D H, LI Z Y, ZHU J L. State of health monitoring and remaining useful life prediction of lithium-ion batteries based on temporal convolutional network[J].IEEE ACCESS, 2020,8:53307-53320.
 - [21] SHEN S, MOHAMMADKAZEM S, CHEN X Y. A deep learning method for online capacity estimation of lithium-ion batteries[J]. Journal of Energy Storage, 2019,25:100817.
 - [22] CHEN Z, XUE Q, XIAO E X. State of health estimation for lithium-ion batteries based on fusion of autoregressive moving average model and Elman neural network[J].IEEE ACCESS, 2019,7:102662-102678.
 - [23] He H, Xiong R, Guo H. Online estimation of model parameters and state-of-charge of LiFe PO₄ batteries in electric vehicles[J].Applied Energy,2012,89(1):413-420.
 - [24] Li S, Zhao P. Big data driven vehicle battery management method: A novel cyber-physical system perspective[J]. Journal of Energy Storage, 2021, 33: 102064.

- [25] Shen M, Gao Q. A review on battery management system from the modeling efforts to its multiapplication and integration[J]. International Journal of Energy Research, 2019, 43(10): 5042-5075.
- [26] Wang Y, Tian J, Sun Z, et al. A Comprehensive Review of Battery Modeling and State Estimation Approaches for Advanced Battery Management Systems[J]. Renewable and Sustainable Energy Reviews, 2020.