Batteries & Supercaps

Research Article
doi.org/10.1002/batt.202500620

Chemistry Europe
European Chemical Societies Publishing

www.batteries-supercaps.org

Data-Driven State of Health for Lithium-Ion Batteries: Feature Engineering, Estimation Approaches, and Future Directions

Zhiqiang Lyu, Xiang'en Li, Zhirui Jin, Hao Wang, Yuan Chen,* and Longxing Wu*

Amidst the prevailing trends of electrification, intelligence, and connectivity, the convergence of new energy vehicles and big data herald transformative opportunities for China's automotive industry. In this context, the estimation of the state of health (SOH) of lithium-ion batteries assumes critical importance for ensuring the safe and efficient operation of electric vehicles. This paper presents a comprehensive exposition of data-driven methodologies for SOH estimation of lithium-ion batteries. It begins with an overview of commonly accepted definitions and the key factors influencing battery SOH, followed by an in-depth exploration of feature engineering—including the processes of feature extraction and selection. The discussion then delves into data-driven approaches to battery SOH management, encompassing both nonprobabilistic and probabilistic models. While numerous methods exist for estimating SOH, each possesses distinct strengths and limitations. Looking ahead, data-driven SOH estimation is poised to evolve toward the integration of multisource data fusion, enhancement through small-sample and transfer learning techniques, incorporation with physical modeling, and expansion across domains. These advancements are anticipated to significantly enhance the precision and dependability of SOH estimation and catalyze the broader deployment of battery technologies across various sectors.

1. Introduction

Lithium-ion batteries hold a central and indispensable role within the energy architecture of contemporary society, with their applications spanning a multitude of critical domains—from portable electronic devices to electric vehicles (EVs) and expansive energy storage infrastructures.[1,2] In these application scenarios, the precise estimation of battery state of health (SOH) serves as a critical cornerstone for realizing efficient battery management and safeguarding the system's safe and stable operation.[3,4] For all categories of devices and systems dependent on lithium-ion batteries, an accurate assessment of the SOH holds profound significance. In the realm of EVs, precise SOH estimation is intrinsically linked to driving range, overall performance, and operational safety. Drivers must be able to plan their journeys judiciously, guided by reliable SOH data, to avert issues such as becoming stranded enroute due to inadequate battery charge.[5–7] Meanwhile, automobile manufacturers and battery suppliers can refine battery management strategies grounded in SOH data to prolong battery lifespan and curtail operational costs. Within energy storage systems, SOH estimation serves as a vital pillar in securing a dependable power supply and enhancing the overall efficiency of energy utilization.[8,9] By precisely discerning the battery SOH, system operators can devise more scientifically grounded charging and discharging strategies, mitigate power fluctuations stemming from battery degradation or malfunction, and enhance the overall stability and reliability of the energy storage system.

The conventional approach to estimating battery SOH primarily relies on constructing physicochemical models to simulate, with precision, the intricate electrochemical processes occurring within the battery. However, the aging of batteries entails a labyrinth of chemical reactions and physical transformations, influenced by a multitude of factors including temperature, charge–discharge rates, and environmental conditions.[10] For instance, excessively low temperatures can diminish the rate of chemical reactions within the battery, leading to a noticeable reduction in capacity.[11,12] Conversely, high charge and discharge rates intensify battery polarization[13,14] and significantly curtail its lifespan. These factors are intricately interwoven, rendering precise physicochemical modeling a formidable challenge, as it struggles to faithfully capture the battery's health under real-world operating conditions. Such models often demand extensive prior knowledge, intricate parameter configurations, and stringent experimental environments, making them ill-suited to accommodate the complex and dynamic nature of practical applications—ultimately undermining the accuracy and reliability of the resulting estimations.

Z. Lyu, X. Li, Z. Jin, H. Wang
School of Internet
Anhui University
Hefei 230039, China

Y. Chen
School of Artificial Intelligence
Anhui University
Hefei 230039, China
E-mail: cumtjiangsucy@126.com

L. Wu
College of Intelligent Manufacturing
Anhui Science and Technology University
Chuzhou 233100, China
E-mail: batterywu@163.com

In this era marked by the swift advancement of information technology, data-driven approaches are garnering growing prominence in the realm of battery SOH estimation. In contrast to conventional techniques that hinge upon intricate chemical mechanisms within the battery, data-driven approaches infer SOH by uncovering latent patterns within vast volumes of real-world operational data.[15] With the aid of modern sensors capable of capturing extensive information—such as voltage, current, and temperature—this paradigm fully leverages empirical data to render more comprehensive and precise assessments of battery health. Moreover, it adeptly manages complex and variable operational conditions that traditional models often struggle to address, exhibiting strong adaptability and generalization capabilities. This opens an innovative pathway toward resolving the persistent challenge of accurate SOH estimation. This paper centers its investigation on data-centric SOH estimation techniques, with primary emphasis on two dimensions: feature extraction and selection, and algorithmic design. Through a rigorous analysis of pertinent technologies, coupled with the distinct attributes of practical battery operation data, it aspires to propose forward-thinking solutions that offer robust technical support for the widespread deployment of lithium-ion batteries across diverse sectors. The paper presents a thorough review of data-driven SOH estimation approaches, explores their evolution in feature engineering and model development, and delves into prevailing challenges and prospective directions. Crucially, this review moves beyond a siloed discussion of features and models. It establishes a coherent framework that links the inherent properties of cathode materials to the extraction of degradation-sensitive features, which in turn, informs the selection and optimization of data-driven models. This material-property-model paradigm is crucial for developing reliable, universal, and interpretable battery health assessment methods tailored to specific battery chemistries. It aims to furnish meaningful insights and references to advance both research and application within this vital field.

2. Battery SOH Definition and Influencing Factors

2.1. Battery SOH Definition

Battery SOH, as a pivotal metric that precisely reflects both the current performance and the degree of aging of a battery, holds critical importance in the assessment of battery lifespan. It not only encapsulates the performance degradation over time but also serves as a vital indicator for estimating the battery's Remaining Useful Life (RUL), offering essential guidance for maintenance, replacement, and the optimization of associated systems. According to IEEE Standard 1188–1996, when a battery's capacity diminishes to 80% of its original value, the performance of the corresponding device may experience a marked decline, potentially leading to failures and operational downtime. Hence, to ensure the uninterrupted functionality of equipment, timely replacement of aging batteries reaching this threshold is imperative. Presently, battery SOH is commonly defined through two approaches: one based on capacity, as shown in (Equation 1), and the other grounded in internal resistance, as shown in (Equation 2) . It is worth emphasizing that, although capacity and internal resistance serve as intuitive and easily operable indicators for assessing battery SOH, their real-time estimation proves challenging in practical applications. This difficulty stems from the intricate electrochemical reactions occurring within the battery, coupled with the susceptibility of the measurement process to noise interference—factors that inevitably introduce bias into the estimation results. Consequently, there exists a pressing demand for the development of novel technologies capable of predicting battery SOH online, characterized by high precision and swift responsiveness.

$$\mathrm{SOH} = \frac{C_i}{C_0} \times 100\% \tag{1}$$

$$\mathrm{SOH} = \frac{R_e - R_i}{R_e - R_0} \times 100\% \tag{2}$$

where C_i and C_0, respectively, indicate the current actual capacity and the initial rated capacity of the battery, R_e, R_i, and R_0, respectively, indicate the internal resistance of the battery when it reaches the end-of-life state, the internal resistance of the battery in the current state, and the internal resistance of the battery in the initial state. These two definitions reflect different aspects of battery degradation. The capacity-based SOH is widely adopted due to its direct correlation with the energy storage capability of the battery, which is critical for applications such as EVs where driving range is a key concern. On the other hand, the internal resistance-based SOH is particularly useful for evaluating power capability and instantaneous performance, especially in high-rate applications like acceleration or regenerative braking. It is noteworthy that both definitions have been standardized in industry practice, such as in IEEE Standard 1188–1996, and are typically used in conjunction with each other. However, real-time estimation of these parameters remains challenging due to the complex electrochemical dynamics within the battery and the susceptibility of measurements to noise and operational variability. Consequently, there is a growing interest in developing data-driven methods that can indirectly infer SOH from readily measurable operational data, such as voltage, current, and temperature profiles, thereby overcoming the limitations of direct measurement-based approaches. Therefore, in data-driven SOH estimation, the focus shifts from direct measurement to the extraction of informative health features (HFs) that are correlated with capacity fade or resistance increase, which will be discussed in the following sections.

2.2. Factors Affecting SOH of Lithium-Ion Batteries

Battery degradation is an inevitable and intrinsic process of aging. The SOH of lithium-ion batteries is influenced by a multitude of interrelated factors that collectively shape their performance and longevity. Among these, temperature emerges as a particularly influential external variable. In cold environments, the sluggish diffusion of lithium ions and significant loss of Li^+ result in heightened

Research Article
doi.org/10.1002/batt.202500620
Chemistry Europe
European Chemical Societies Publishing

internal resistance and accelerated capacity fade, a phenomenon exemplified by the notable decline in EV range during winter. Conversely, elevated temperatures induce irreversible chemical reactions, diminish the quantity of active materials, promote excessive growth of the solid electrolyte interphase (SEI) layer, and expedite battery aging—manifesting, for instance, in the rapid capacity loss of electronic devices in hotter climates. Moreover, parameters such as charge and discharge cut-off voltages, current rates, and depth of discharge also play pivotal roles in determining battery health. Elevated charging cutoff voltages, high-rate charging currents, and excessive discharging all serve to accelerate the degradation of battery health, with the habitual use of fast-charging technologies standing as a quintessential example of this phenomenon. Internally, the formation of the SEI film depletes lithium ions, thereby diminishing capacity, while progressive degradation heightens the risk of internal short circuits, subtly yet persistently altering the battery's intrinsic properties throughout its charge–discharge cycles. The accumulation of lithium ions may give rise to lithium dendrites—needle-like structures capable of piercing the separator and inducing internal short circuits, posing grave risks to both safety and SOH particularly in batteries that have undergone severe aging. Additionally, volumetric changes and mechanical stresses during cycling compromise the integrity of the electrode structure, diminishing the contact surface between active materials and the electrolyte, leading to increased internal resistance and a corresponding decline in capacity. The dissolution of active materials and electrolytes diminishes the quantity of reactive constituents and disrupts ion transport, respectively—culminating in a decline in battery capacity, a rise in internal resistance, and an overall deterioration in performance. In essence, these internal and external influences are intricately interconnected, collectively shaping the battery's aging trajectory and the evolving trend of SOH. The key factors impacting the SOH of lithium-ion batteries are summarized as shown in **Figure 1**.

2.3. Battery Type Differences (Based on Cathode Materials)

The SOH degradation of lithium-ion batteries essentially refers to the gradual deterioration of the electrochemical activity and

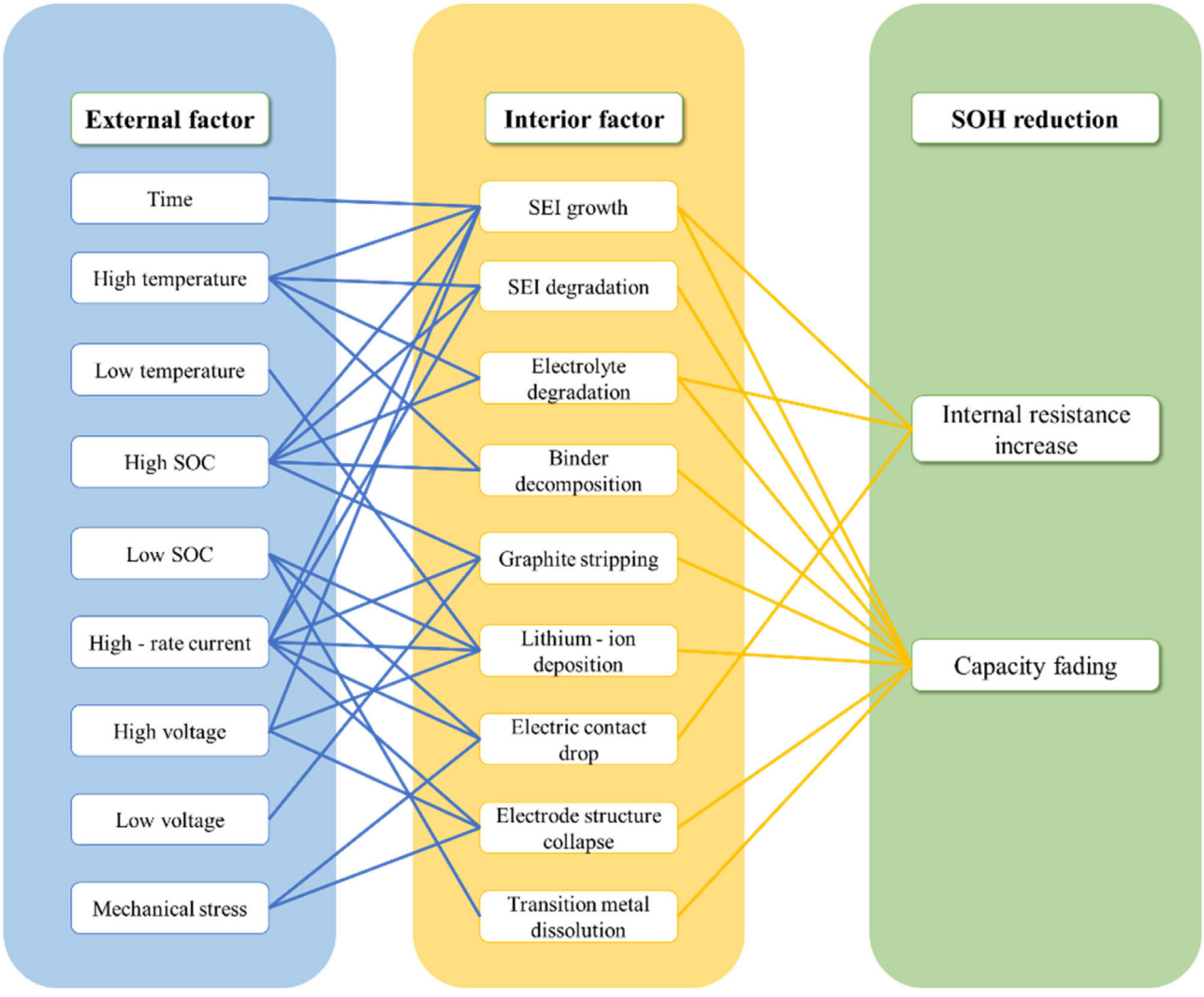

Figure 1. Internal and external factors that affect battery SOH.

structural stability of cathode materials during cycling. Due to differences in crystal structure, elemental reactivity, and reaction mechanisms among different cathode materials, their aging pathways and SOH degradation laws vary significantly. This not only directly determines the extraction direction of HFs in feature engineering but also affects the adaptation logic of data-driven models. Based on the experimental characteristics and literature conclusions of mainstream cathode materials, this section systematically analyzes the differences in their aging mechanisms and the impacts on SOH estimation.

2.3.1. $LiCoO_2$ (Lithium Cobalt Oxide) Batteries

The SOH degradation of $LiCoO_2$ batteries is primarily driven by two core mechanisms: layered structure failure and transition metal dissolution. $LiCoO_2$ adopts an α-$NaFeO_2$-type layered crystal structure, which enables the free intercalation and deintercalation of Li^+ between layers. However, inherent limitations in its structural stability and elemental chemical activity lead to significant aging issues. Specifically, when $LiCoO_2$ is charged to above 4.2 V, the massive deintercalation of Li^+ causes the interlayer spacing of CoO_2 to shrink from 0.46 to 0.42 nm. Under long-term cycling, lattice distortion accumulates continuously, eventually triggering the collapse of the layered structure. This process reduces the Li^+ diffusion coefficient from an initial $10^{-10}\,cm^2\,s^{-1}$ to below $10^{-12}\,cm^2\,s^{-1}$, directly manifesting as continuous capacity fading—with the capacity retention rate generally dropping below 60% after 500 cycles.[16] Under operating conditions with even higher voltages, Co^{3+} is prone to oxidation to Co^{4+} and dissolution into the electrolyte. The dissolved Co ions migrate to and deposit on the surface of the graphite anode, disrupting the ionic selectivity of the SEI layer. As a result, the battery internal resistance increases by 80%–100% of its initial value after 300 cycles.[17,18] Additionally, $LiCoO_2$ has an electronic conductivity of only $10^{-3}\,S\,cm^{-1}$. High-rate charging exacerbates the polarization effect, causing the charging voltage plateau to be reached earlier. The duration of the constant current (CC) phase is shortened by 30%–40% compared to 0.5C charging, which further accelerates the failure of active materials[19] and greatly limits its application in high-rate scenarios. Key factors influencing battery SOH estimation: for lithium cobalt oxide batteries, feature engineering should focus on capturing anomalous variations within high voltage ranges. For instance, the rate of change in charging time within the 4.2–4.3 V range reflects the degree of polarization, while the voltage drop slope during discharge in the 1.5–2.0 V range is closely associated with increased internal resistance due to cobalt ion dissolution.

2.3.2. NCM (Lithium Nickel Cobalt Manganese Oxide) Batteries

NCM batteries have become the mainstream choice for power batteries in new energy vehicles due to their high specific capacity. However, the high chemical activity of their Ni element significantly complicates the aging mechanism, with core degradation issues focusing on three aspects: first, the loss of active sites caused by Ni^{2+} oxidation—during cycling, Ni^{2+} in NCM is easily oxidized to Ni^{3+}/Ni^{4+}; some oxidized Ni ions detach from the lattice to form inactive NiO_2, directly reducing the number of active sites on the cathode. Additionally, the strong oxidizing property of Ni^{4+} catalyzes the decomposition of the electrolyte, generating by-products such as CO_2 and HF, which further damage the electrode–electrolyte interface stability.[20,21] Second, particle microcracks induced by volume change—the volume change rate of NCM during charge–discharge reaches 4%–6%. Under long-term cycling, microcracks accumulate inside the cathode particles, disrupting the transport paths of Li^+ and electrons, and ultimately manifesting as "sudden capacity drop". For example, after 200 cycles of an NCM811 battery, the capacity retention rate may plummet from 90% to 75%, and the subsequent degradation rate accelerates significantly.[22,23] Third, environmental sensitivity amplified by poor thermal stability—high-nickel NCM is prone to thermal decomposition in environments above 45 °C; the released oxygen reacts with the electrolyte, causing abnormal thickening of the SEI layer. This leads to a 20%–30% increase in capacity degradation amplitude at low temperatures compared to room temperature.[24,25] The critical impact of such complex aging characteristics on SOH estimation lies in the fact that feature engineering for NCM batteries must focus on capturing the abrupt changes that occur during later cycles. For instance, in differential voltage analysis (DVA) curves, characteristic peak voltages may shift suddenly—for example, from 3.8 to 3.7 V—which directly reflects interrupted transport paths caused by particle microcracks. Similarly, a sharp decline exceeding 15% in the peak area of incremental capacity analysis (ICA) curves is often strongly correlated with active site loss resulting from nickel ion oxidation. Through such sudden change features, the core aging stages and degradation roots of NCM batteries can be accurately identified.

2.3.3. LFP (Lithium Iron Phosphate) Batteries

LFP batteries owe their exceptional structural stability to their olivine crystal structure, where strong covalent bonds (P—O bonds) within the structure provide robust support. Their aging mechanism is relatively mild, with core issues focusing on two aspects: SEI layer growth and lithium-ion inventory loss. From the perspective of structural characteristics, LFP exhibits a mere 1%–2% volume change rate during charge–discharge cycles, far lower than that of layered cathode materials such as NCM. The olivine structure effectively inhibits particle cracking, resulting in more gradual capacity fading—even after 1000 cycles, the capacity retention rate can still reach over 80%.[26,27] In terms of interfacial behavior and ion dynamics, although the bulk structure of LFP is stable, the electrolyte continues to decompose on the anode surface during long-term cycling, forming an SEI layer. The gradual thickening of this SEI layer increases the migration resistance of Li^+, manifesting as a slow rise in internal resistance, with a typical growth rate of less than 50% after 1000 cycles.[28,29] Meanwhile, LFP has a relatively low Li^+ deintercalation potential. During cycling, some Li^+ ions are trapped by the SEI layer or form irreversible lithium compounds, leading to a reduction in the total amount of Li^+ available for electrochemical reactions—this is also the core cause of its

capacity fading.[30,31] For the estimation of SOH, this slow and gradual aging characteristic implies that feature extraction should focus on long-term time-series features. Examples include the gradual extension of charging time in the CC phase and the slow shift in peak positions of the incremental capacity (IC) curve. These indicators, which change gradually with cycling, enable more accurate reflection of the battery's aging degree.

3. Feature Engineering

The primary driver behind the aging of lithium-ion batteries lies in the occurrence of electrochemical side reactions during charge and discharge cycles. These undesirable reactions lead to the gradual depletion of active materials at both the anode and cathode, as well as the consumption of lithium ions and electrolyte, ultimately resulting in diminished capacity and a decline in power output. However, these degradation mechanisms are not readily discernible from the raw charge–discharge data of the battery. Thus, within data-driven methodologies, a pivotal task is the extraction of meaningful and effective HFs capable of capturing both the battery's performance state and its degree of aging. The charge and discharge profiles of lithium-ion batteries are rich repositories of valuable diagnostic information.[32,33] Through the application of feature engineering, a range of critical parameters—such as charging plateau voltage, charging duration, discharging plateau voltage, and discharge capacity—can be precisely extracted. However, raw battery charge–discharge datasets are frequently marred by noise and redundancy, which can significantly undermine the accuracy of performance assessment. Fortunately, preprocessing techniques such as data cleansing and normalization within the feature engineering framework can effectively eliminate such noise, including anomalous data points resulting from sensor measurement deviations.[34]

Aging datasets serve as a cornerstone in feature engineering, offering crucial support for a profound comprehension of battery performance degradation patterns and the precise evaluation of SOH. However, as battery technology continues to advance, the longevity of batteries has significantly improved, rendering aging tests increasingly time-consuming and costly. In light of this, numerous research institutions have opted to make their battery aging datasets publicly accessible, thereby facilitating the validation and benchmarking of SOH prediction methodologies for the broader research community.

The NASA datasets,[35,36] stand as a renowned and widely utilized public resource within the new energy sector, encompassing data on solid-state batteries, lithium-ion batteries, and more. Curated primarily by NASA's Ames Research Center and the Prognostics Center of Excellence, these datasets offer comprehensive records of battery charging, discharging, and impedance testing under diverse temperature conditions. They encompass a rich array of parameters, including current, voltage, temperature, discharge capacity, and electrochemical impedance spectroscopy (EIS) data. For instance, applications such as battery SOH estimation and RUL prediction are commonly explored.[37,38] The experimental conditions are well-defined—for example, temperature is maintained at a stable 24 °C; charging is performed using a CC of 1.5A until the battery voltage reaches 4.2 V, followed by constant voltage (CV) charging until the current tapers to 20mA. Discharging is executed under a CC of 2A until the battery voltage descends to a predetermined threshold. The original data format is mat, must be preprocessed to extract key parameters such as capacity, current, and voltage.

The CALCE open datasets,[39,40] originating primarily from the Center for Advanced Life Cycle Engineering at the University of Maryland, are extensively utilized in the domain of battery research, offering robust support for a wide range of studies and practical applications.[41–43] It encompasses operational data for a variety of battery types, including lithium-ion batteries. Notably, the CS2 series of standardized lithium-ion battery datasets stands out as particularly representative, meticulously documenting battery performance under rigorously controlled charge cycle tests. Each CS2 battery follows an identical charging protocol—specifically, the conventional CC/CV method—wherein a CC of 0.5C is applied until the voltage attains 4.2 V, after which the voltage is maintained at 4.2 V until the charging current declines below 0.05 A.

The Stanford-MIT dataset, a collaborative effort between Stanford University, the Massachusetts Institute of Technology, and the Toyota Research Institute, stands as the most extensive publicly available dataset of its kind. It draws upon a vast array of battery charge–discharge experiments, encompassing over 1,000 cycles across numerous battery samples subjected to five distinct charging and discharging protocols. The dataset meticulously chronicles the evolution of electrical parameters—such as voltage, current, and temperature—alongside critical metrics like charge and discharge capacity and internal resistance throughout the cycling process. These datasets offer a multifaceted perspective and robust support for investigating battery performance, assessing SOH, and uncovering the underlying mechanisms of aging. The most recent Stanford-MIT collection comprises two datasets,[44,45] one of which encompasses 124 commercial lithium-ion batteries subjected to failure through repeated cycling under fast—charging conditions. These lithium iron phosphate/graphite cells, produced by A123 Systems, are cycled within a horizontal cylindrical fixture connected to a 30-channel Arbin LBT potentiostat, housed in a forced-convection thermal chamber maintained at 48 °C. Each battery possesses a nominal capacity of 1.1 Ah and a rated voltage of 3.3 V. The second dataset likewise utilizes A123's APR18650M1A lithium iron phosphate power cells for aging experiments; however, its focus lies in optimizing fast-charging strategies for lithium-ion batteries. Consequently, every cell in this dataset undergoes charging through one of 224 distinct six-step, 10 min fast-charging protocols.

3.1. Feature Extraction

When constructing a data-driven battery SOH estimation model, data feature extraction is crucial. The quality of data features directly affects the accuracy and efficiency of the model. High-quality features can improve the model's ability to capture changes in battery status, reduce redundant information, reduce the complexity of model training, and make the model more accurate and

efficient in assessing battery SOH. Many researchers have mined data feature related to battery aging based on charge and discharge data and regarded them as HFs. It can be extracted from voltage data, current data and temperature data during charging and discharging. Common extraction methods include feature extraction based on charge and discharge stage, feature extraction based on feature curve, and another feature extraction method based on model parameters.

In the development of data-driven models for estimating battery SOH, the extraction of data features plays a pivotal role. The caliber of these features directly influences the model's precision and computational efficiency. High-quality features enhance the model's sensitivity to variations in battery condition, diminish redundant information, lower the complexity of training procedures, and ultimately render SOH assessment more accurate and efficient.[46–48] A multitude of researchers have explored aging-related data features derived from charging and discharging behavior, designating them as HFs. These can be extracted from voltage, current, and temperature profiles recorded during operation. Common approaches include stage-based feature extraction from charge and discharge phases,[49,50] curve-based feature extraction techniques,[51] as well as methodologies centered on model parameter estimation.

3.1.1. HF Based on Charge and Discharge Curves

During the charging phase, the CC–CV charging process is closely related to the battery SOH. As illustrated in **Figure 2**, the precharging stage begins with a gentle current and a gradually increasing voltage. This phase primarily serves to awaken the battery gently, minimizing the risk of damage that might result from a sudden inrush of current—akin to rousing a slumbering cell with delicate care. In the CC phase, the current is held steady until the voltage ascends to its designated threshold, upon which the process transitions into the CV stage. There, the voltage is maintained at its rated level while the charging current tapers down to a predefined minimum. As the battery ages, the duration of the CC stage progressively diminishes, and the voltage curve steepens—an indication of rising internal resistance within the cell.

Drawing upon this law of variation, metrics such as total charging duration, the temporal proportion of the CC and CV phases, and the integral area within specific voltage intervals can serve as salient characteristic parameters indicative of battery health. Tian et al.[52] distilled seven distinct features from battery charging signals—namely voltage, current, and time—including the temporal ratios of the CC and CV phases, average voltage and duration across different stages, and the mean current during the CV phase. These features encapsulate meaningful representations of the battery SOH, offering both interpretability and physical significance. A random forest regression model was employed to quantify the relative importance of each feature, guided by the purity index and the corresponding loss function. The purity index serves as a metric to evaluate the homogeneity of decision tree nodes, during the construction of a random forest, features that contribute more significantly to enhancing node purity are deemed more influential. The loss function, on the other hand, quantifies the discrepancy between the model's predictions and the actual values. Features that play a greater role in minimizing the loss function are considered to hold higher importance. Through this approach, the relative significance of each feature within the model can be systematically quantified.

In the constant discharge stage, lithium inventory loss occurs during battery discharge, which is closely related to battery aging. After analyzing the trend of CS2−37 battery discharge voltage, it can be found that with the increase of the number of discharge cycles, the time required for complete discharge will gradually decrease, and the rate of voltage drop will gradually accelerate. Usually, researchers choose the time of different stages in the discharge process as the aging characteristic parameters, which help to dig into the original data information affecting the battery SOH, and then more accurately assess the SOH of the battery degradation is closely related to the discharge process. Discharge is a normal working process of the battery, but frequent or unreasonable discharge will accelerate battery degradation. In theory, the hidden degradation trend of constant current CC phase voltage

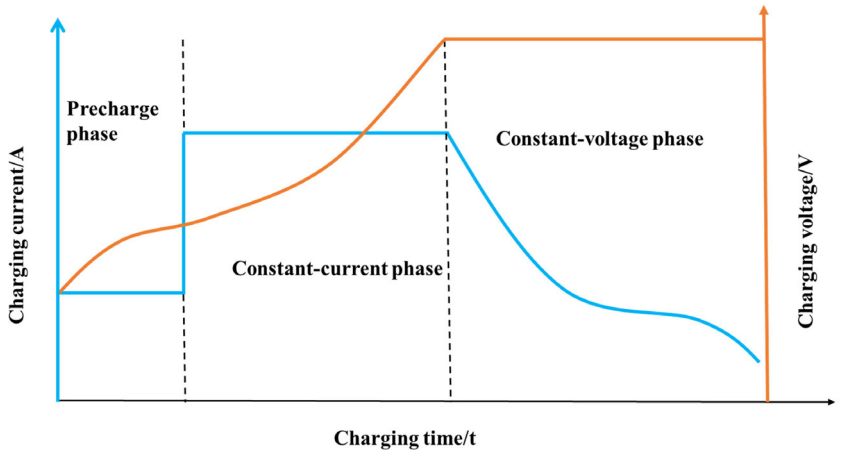

Figure 2. Charge curve of CC and CV stage.

can be used to extract the HF parameters, which has a key value for the evaluation of battery status and other related research and practice.

During the CC discharge phase, the loss of lithium inventory becomes evident, a phenomenon intrinsically tied to battery aging. An analysis of the discharge voltage trend for the CS2—37 battery reveals that, as the number of discharge cycles increases, the time required for a complete discharge gradually diminishes, while the rate of voltage decline correspondingly accelerates. Researchers often extract the durations of various discharge phases as characteristic aging parameters, which serve to unearth critical insights embedded within the raw data that influence the battery SOH. These parameters, in turn, enable a more nuanced and precise evaluation of SOH.[13,53,54] Battery degradation is intimately connected to the discharge process: although discharging is an essential operational function, excessive or suboptimal discharge conditions can hasten deterioration. Theoretically, the latent degradation patterns observed in the CC phase voltage may be harnessed to extract HFs parameters—elements of paramount importance for assessing battery status and informing both research and practical applications.

3.1.2. HF Based on IC Curves

The ICA method stands as a widely adopted, nondestructive in situ approach for investigating battery aging mechanisms. By conducting aging evaluations, performance assessments, and pulse tests, the IC of the battery is calculated, from which the IC curve is subsequently derived, as illustrated in **Figure 3**. This curve encapsulates the phase transitions of active materials during the charging and discharging processes, with its distinctive peaks bearing a strong correlation to the progression of battery aging.[55–58]

For instance, in the analysis of IC curves from CS2—37 batteries across various charging cycles, Gaussian filtering is first employed to eliminate noise, after which parameters such as the position, gradient, and area of characteristic peaks are extracted. The evolution of these parameters serves as an effective indicator of aging severity and the battery SOH. While ICA is extensively utilized, the IC curve remains highly sensitive to the initial state of charge (SOC) and the charge–discharge rates, often resulting in deformation and displacement of features that compromise the precision of SOH estimation. Wang et al.[59] introduced a filtering technique grounded in the finite time differentiator (FTD) to process high-frequency sampled data, yielding smoother IC curves. A series of experiments were meticulously designed and conducted under varying initial SOC and charge–discharge rates to investigate their impact on HFs. Consequently, a SOH estimation approach, adaptable to diverse initial SOC and dynamic operating conditions, was developed to enhance both accuracy and robustness. Wang employed FTD filtering to accommodate dynamic charge–discharge rates and mitigate distortion in IC curve features caused by variations in operating conditions, whereas He et al.[60] broadened its application to suppress noise in high-frequency sampled data. Compared with traditional methods such as Kalman filtering, Gaussian filtering, and moving average filtering, the FTD method exhibits superior noise suppression performance when processing high-frequency signals: on one hand, it can more thoroughly filter out random noise introduced by sensor errors in high-frequency sampling, making the characteristic peaks on the IC curve more clearly distinguishable; on the other hand, its computational efficiency is comparable to that of traditional filtering methods, while it further improves curve smoothness and avoids the loss of feature information caused by over-filtering. Nevertheless, the selection of filtering parameters remains inherently subjective, introducing ambiguity into the IC curve derivation. Furthermore, throughout the evolution of the IC curve, fitting errors are inescapable. Given the complexity of real-world measurement conditions, accurately quantifying the extent of measurement noise within these error margins proves challenging. Li et al.[61] transformed the traditional Incremental capacity–voltage (IC–V) curve into an incremental capacity–time (IC–T) curve, extracting HFs through the lens of time-series analysis. This transformation effectively circumvents issues arising from voltage measurement inaccuracies and numerical approximation errors, enabling the identification of novel health indicators from a temporal perspective.

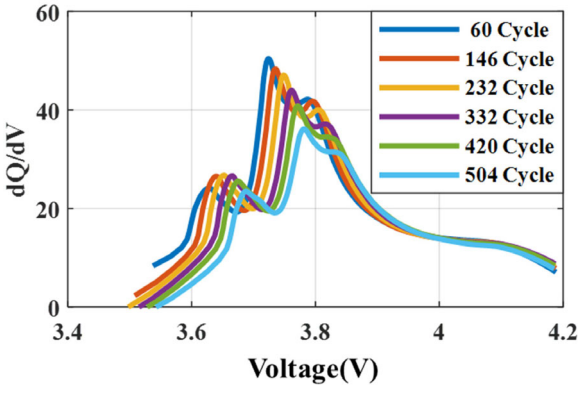

Figure 3. IC curves with aging.

3.1.3. HF Based on Differential Voltage Curves

The DVA method, akin to the ICA method, is widely employed for battery safety diagnostics and the prediction of SOH.[62–64] Distinctively, the DVA approach places greater emphasis on subtle variations in voltage, uncovering the underlying electrochemical reaction mechanisms within the battery by examining the differential relationship between voltage and capacity throughout the charging and discharging processes.[65,66] The DV curves of CS2—37 batteries across various charging stages were meticulously processed and examined. It was observed that electrolyte decomposition, the depletion of active materials, and the escalation of internal resistance collectively led to shifts in the positions of characteristic peaks and valleys, with the overall curve trending toward lower capacities. By scrutinizing parameters such as the axis coordinates, area, and time required to reach these peaks and valleys, critical insights into battery aging can be uncovered, enabling more effective management of the battery SOH throughout its lifecycle. Xia et al.[67] introduced a novel feature set—encompassing

peak height, peak position, peak area, and Q2—extracted from the IC/DV curve reconstructed via a second-order RC model voltage. The Pearson correlation coefficients between these features and SOH exceeded 0.9, surpassing those obtained through conventional filtering techniques. These refined indicators offer a more holistic and precise reflection of battery aging, yielding more representative features for both SOH and RUL prediction, thereby enhancing the overall accuracy and robustness of prognostic models. although both ICA and DVA extract HFs through the differentiation of voltage–capacity curves, there are significant differences in their sensitivity to battery aging mechanisms, computational complexity, and SOH estimation accuracy. These differences are strongly correlated with battery cathode materials and operating conditions. In terms of accuracy, for ICA: the mean absolute error (MAE) of SOH estimation is 1.2%−3.5% for $LiCoO_2$ batteries in the NASA dataset;[55] in the CALCE dataset for NCM batteries, due to the sudden capacity drop in the late cycling stage, the MAE increases to 4.0%.[59] For DVA: the MAE is 0.8%−2.8% for LFP batteries in the CALCE dataset;[65] in LFP batteries after 1000 cycles, since aging is dominated by the slow growth of the SEI layer, the MAE can be as low as 0.5%.[67]Regarding computational cost, ICA requires high-frequency voltage sampling and complex denoising processes such as Gaussian filtering and FTD.[60] Under the same hardware conditions, the single-cycle data processing time of ICA is 30%−50% longer than that of DVA, and the single-cycle data volume of ICA is ≈2–3 times that of DVA. In contrast, DVA only requires low-frequency sampling and can achieve denoising through moving average filtering;[66] its single-cycle data processing time is usually less than 100 ms, with a small data volume, making it suitable for resource-constrained embedded devices. In terms of anti-interference capability, ICA is sensitive to charging rates: when charged at 1C, the IC peak shifts by up to 0.05 V, leading to a 20%–30% increase in SOH estimation error.[59] However, the error fluctuation of ICA is less than 1% within the temperature range of 25–45 °C,[55] showing good adaptability to high-temperature operating conditions. DVA has high tolerance to rate changes, with an error fluctuation of less than 0.8% within the range of 0.5C–2C,[65] but it is sensitive to temperature: at a low temperature of −10 °C, the depth deviation of the DV valley reaches 15%, resulting in a 10%–15% increase in SOH estimation error,[64] and its performance degrades significantly under low-temperature operating conditions. For applicable battery types, ICA is preferentially suitable for batteries with obvious voltage plateaus, such as $LiCoO_2$ and high-nickel NCM batteries, and can distinguish multistep redox reactions through multiple characteristic peaks. DVA is preferentially suitable for batteries with flat voltage plateaus, such as LFP batteries, and can accurately capture the internal resistance increase caused by SEI layer growth through the depth of the characteristic valley.

3.1.4. HF Based on Model Parameters

In battery SOH assessment and HF extraction, electrochemical model[68,69] and equivalent circuit model[70,71] are two mainstream modeling methods, each with its own advantages. Equivalent circuit method is to equivalent the battery to a circuit model composed of resistor, capacitor, inductor and other circuit components, common Rint model, Thevenin model[72] and PNGV model. The Rint model is shown in **Figure 4**a, which is composed of a CV source and a resistor in series. Thevenin model is derived from Rint model, in order to simulate the nonlinear behavior inherent in charge and discharge process, RC network is introduced. The Thevenin model is often referred to as the first-order RC model, as shown in Figure 4b. The PNGV model is a nonlinear low-order model. Based on the Thevenin model, an additional capacitor C_b is connected in series to describe the changes in the open circuit voltage of the battery caused by the accumulation of load current over time. The PNGV model is shown in Figure 4c.

Compared to the simplistic internal resistance model and the first-order RC (Thevenin) equivalent, the PNGV model offers a

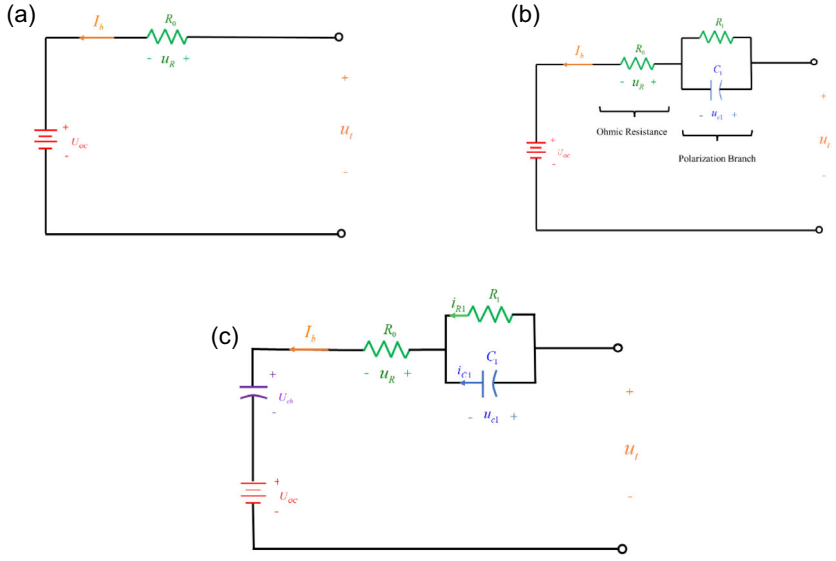

Figure 4. Equivalent circuit model. a) Rint mode; b) Thevenin model; c) PNGV model.

more nuanced depiction of the battery's output characteristics. Zhu et al.[73] proposed an integer variable-order RC equivalent circuit model that dynamically adapts its structural order in accordance with the distribution features of the battery's voltage plateau. In regions where the open-circuit voltage varies rapidly, a high-order model is employed, while in stable plateau regions, a lower-order model suffices—allowing for a more precise representation of both the dynamic and static behavior of the battery under diverse operating conditions. This model facilitates the acquisition of terminal and polarization voltages across varying SOC and current ratios. These dynamic parameters serve as real-time indicators of SOH evolution, enabling the timely detection of latent battery issues and offering a robust mechanism for continuous SOH monitoring. Amir et al.[74] developed a 2-RC equivalent circuit model, treating the resistance and capacitance parameters as critical diagnostic features. The evolution of these parameters mirrors the underlying electrochemical and physical transformations within the battery, rendering them closely indicative of the battery SOH. Their interdependencies were elucidated through the application of Kirchhoff's voltage law. Furthermore, the Arrhenius equation was incorporated to account for temperature as a dynamic factor, thereby expanding the dimensionality of the feature space and capturing its influence on reaction kinetics, battery performance, and lifespan. These enriched features were employed in a nonlinear least squares curve fitting framework to estimate SOH, with their efficacy substantiated through empirical validation on the UMD CALCE and NASA Ames datasets.

The electrochemical model is based on the internal electrochemical principle of the battery, and can describe the complex physical and chemical phenomena of the battery during the charge and discharge process, such as the embedding and removal of lithium-ions between the positive and negative electrodes, charge transfer, diffusion process, etc.[75] Common electrochemical models include button cell model, single particle model, pseudo-2D model and so on, as shown in **Figure 5**.

The negative electrode material is Li_xC_6, which is indicated by the blue area and sphere in the figure. The current collector is a copper current collector. The positive electrode material is Li_yMO_2, which is indicated by the green area and sphere in the figure. The current collector is an aluminum current collector. During discharge, the battery is connected to the load to supply power externally. Electrons flow from the negative electrode through the load to the positive electrode, while lithium ions are released from the negative electrode and migrate to the positive electrode through the electrolyte and separator. During charging, the external power source charges the battery. Electrons flow into the negative electrode, while lithium ions escape from the positive electrode and are intercalated into the negative electrode through the electrolyte and separator. By estimating the parameters of the electrochemical model, the characteristic parameters reflecting the internal state of the battery can be obtained. For instance, intrinsic parameters such as internal resistance, electrode reaction rate constants, and diffusion coefficients may serve as vital indicators. Moreover, state variables derived from electrochemical models—such as electrode potential and the spatial distribution of lithium-ion

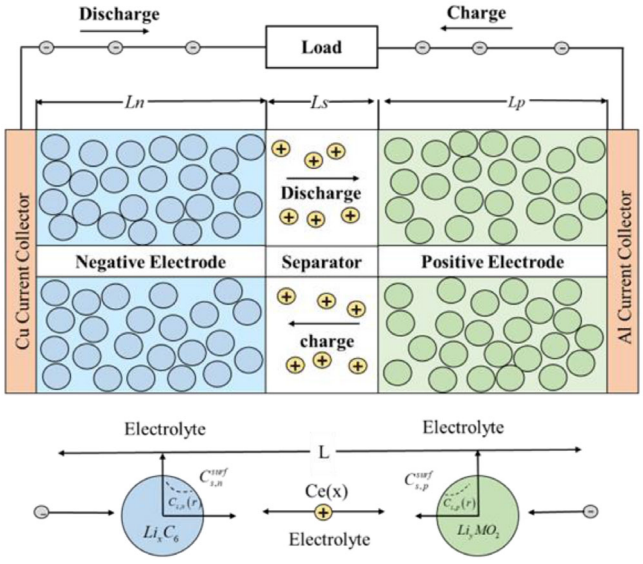

Figure 5. The working principle of the electrochemical model.

concentration—can likewise be harnessed as defining features for the evaluation of battery SOH.

For instance, by observing fluctuations in electrode potential throughout the charging and discharging cycles, one can gain insight into the internal electrochemical reactions and infer the battery's SOH. Alhammadi et al.[76] employed a multiphysical model to resolve the conservation of mass, momentum, and charge within the battery, thereby deriving internal descriptors such as equilibrium potential, velocity profiles, and current distribution. A modified Randles equivalent circuit model was then utilized to fit the EIS data. Parameters within this circuit—such as series Ohmic resistance and charge transfer resistance—embody distinct physical interpretations and can be treated as characteristic features of the battery. By adjusting the properties of the electrodes, separator, and electrolyte, the resulting EIS spectra and equivalent circuit parameters were examined to extract the influence of these components on the battery's intrinsic attributes. Zhang et al.[77] established an electrochemical model of solid-state batteries to obtain characteristic variables such as lithium-ion concentration C_{Li^+} in the solid electrolyte and lithium concentration C_{Lis} in the cathode. In the process of model analysis, features that have an important impact on battery state estimation were selected by comparing different models. For example, it was found that models containing lithium-ion generation/recombination reaction terms were more accurate in predicting electrolyte lithium-ion concentration and voltage. By simplifying the model, a simplified voltage function with the lithium-ion concentration at the boundary of the electrolyte domain and the lithium concentration at the cathode as the main characteristics is determined.

3.2. Feature Selection

Through the foregoing analysis, it becomes evident that data-driven battery SOH estimation approaches encompass a vast array of potential HFs. However, an overabundance of HFs can

significantly elevate the computational burden of model training, while low-quality features may introduce noise, thereby impairing the model's learning performance. Moreover, the applicability of certain HFs is often limited to specific operational scenarios. For instance, features extracted from the CV charging phase are irrelevant for batteries lacking such conditions, while characteristics like the discharge capacity–voltage curve or average voltage decay—derived from constant discharge data—prove ineffective when applied to NASA's random walk aging datasets, which involve dynamic discharge profiles. Consequently, it is imperative to selectively identify accurate and robust HFs tailored to the intended application context. Analytical methods rooted in correlation coefficients and dimensionality reduction offer promising strategies for effectively filtering and refining these health indicators.

3.2.1. Feature Screening Based on Correlation Coefficient

Common analytical techniques grounded in correlation coefficients include the Pearson correlation coefficient,[78–82] with its computational formulations outlined in (Equation 3) through (Equation 5), as well as the Spearman rank correlation coefficient,[83,84] the Kendall correlation coefficient,[85,86] and grey relational analysis (GRA).[87–89] The Pearson method necessitates that the variables be continuous and typically presumes a normal—or approximately normal—distribution. In contrast, the Spearman coefficient imposes no stringent assumptions regarding data distribution and is applicable to both continuous and ordinal variables. These approaches quantitatively assess the strength of association between variables, thereby facilitating the identification of features most intimately correlated with battery SOH.

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\overline{y})^2}} \tag{3}$$

Among

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{4}$$

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{5}$$

where r_{xy}represents the correlation coefficient between variable x and variable y, n represents the length of the health vector, x_i is a HF value indexed by i, $\overline{x}$ is the mean value of the HF, y_i is the actual capacity of the battery corresponding to the HF x_i, and $\overline{y}$ is the mean value of the battery capacity. Pearson correlation coefficient is an indicator used in statistics to measure the strength and direction of the linear relationship between two variables.[90] Its value is between -1 and 1, when r is 1, it means that the two variables are completely positively correlated, when r is -1, it means that the two variables are completely negatively correlated, and when r is 0, it means that the two variables are not linearly correlated. When r is close to 1 or -1, it indicates that there is a positive linear correlation between the two variables. The closer r is to 1 or -1, the stronger the linear correlation degree is. As shown in **Figure 6**, the color bar on the right side marks the correlation coefficient range from -1 to 1. Yellow represents a positive correlation, and the brighter the color, the stronger the positive correlation; the green color represents negative correlation, and the darker the color, the stronger the negative correlation.

By widely collecting battery data in various stages of use, including key characteristic parameters such as voltage, current, internal resistance, and cycle number in detail, and accurately recording the corresponding battery SOH value, a complete data set is constructed. Then the Pearson correlation coefficient formula is used to calculate the combination of each characteristic parameter and SOH in the data set. According to the Pearson correlation coefficient obtained by the calculation, the correlation degree between each characteristic parameter and SOH is deeply analyzed, so as to determine the effect of each parameter on the SOH of battery. Chen et al.[91] used the Pearson correlation coefficient method to select six strongly correlated features from the differential temperature curve and combined with the gated cycle unit neural network to predict the SOH of battery. Bao et al.[92] innovatively proposed a global-local embedded module to learn the original information of battery voltage, current, and temperature through parallel convolutional flows of two different depths, and effectively establish the mapping relationship between battery charge/discharge curve and SOH by using Pearson correlation coefficient and other methods. This method achieves higher accuracy and robustness in SOH estimation.

GRA quantifies the degree of similarity and association between a reference sequence and a comparative sequence by computing their respective correlation coefficients and correlation degrees.[93] At its core, the method assesses the closeness of relationships by evaluating the geometric resemblance between their corresponding curves: the greater the similarity in shape, the stronger the inferred correlation. conversely, diminished similarity implies weaker correlation. The procedure begins by preprocessing the data, mapping raw values to a normalized range of (0,1) and subsequently determining both the correlation coefficient and overall correlation degree between the reference and each comparative sequence. This enables a thorough analysis of inter-factor relationships. The mathematical expressions for the correlation coefficient and correlation degree are presented in (Equation 6) and (7), respectively.

$$\zeta_i(k) = \frac{\min_i \min_k |x_0(k)-x_i(k)| + \rho \max_i \max_k |x_0(k)-x_i(k)|}{|x_0(k)-x_i(k)| + \rho \max_i \max_k |x_0(k)-x_i(k)|} \tag{6}$$

$$r_i = \frac{1}{n}\sum_{k=1}^{n} \zeta_i(k) \tag{7}$$

where $x_0(k)$ is the value of the reference sequence at the k moment, $x_i(k)$ is the value of the comparison sequence i at the k moment, and ρ is the resolution coefficient, the value range is 0–1, usually 0.5, its core function is to adjust the significance of differences

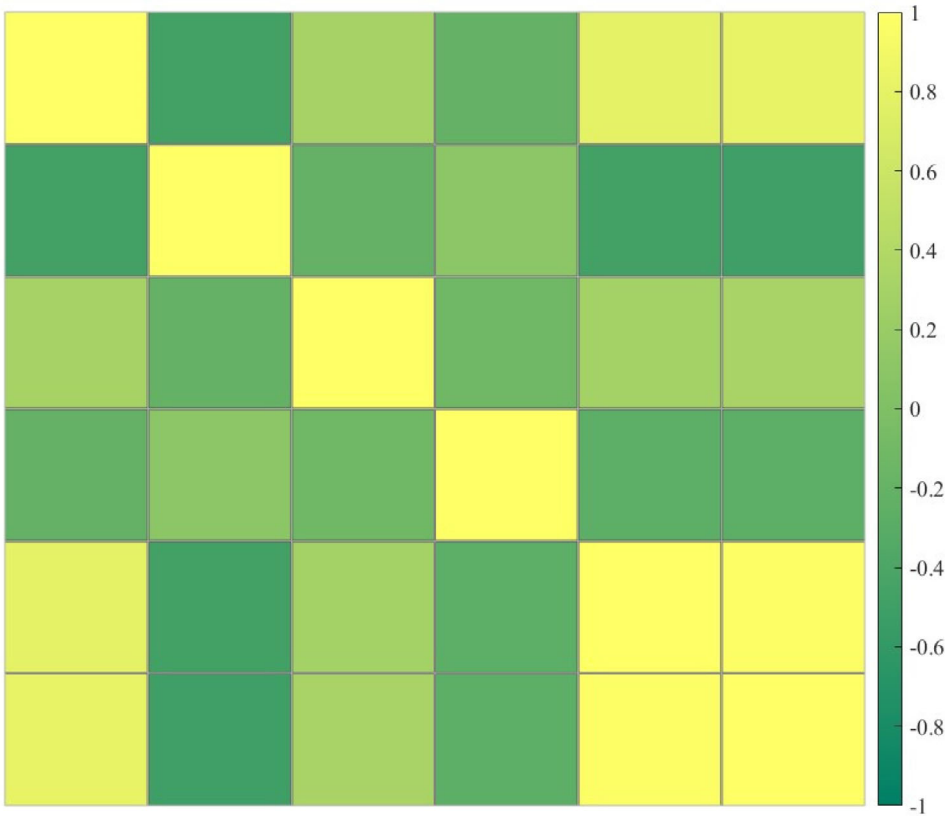

Figure 6. Pearson correlation thermogram.

between correlation coefficients, preventing data noise from amplifying minor differences or masking the true differences in the correlation between features and SOH. r_i represents the correlation degree between the comparison sequence i and the reference sequence, and n is the number of data. By calculating the arithmetic average of the correlation coefficient at each moment, the correlation degree that comprehensively reflects the correlation degree of the two sequences is obtained. In the assessment of battery SOH, the SOH of the battery is taken as a reference sequence, and various characteristic parameters of the battery are taken as a comparison sequence. By analyzing the correlation degree between these characteristic parameters and SOH, the key characteristic parameters that have a greater impact on SOH are determined. Li et al.[94] proposed a battery SOH estimation method combining ICA and entropy weight method (EWM). Firstly, IC curves were obtained by interpolation method, health indicators were extracted from some IC curves for GRA, and the importance of indicators was evaluated by EWM. Finally, the grey correlation degree of reference sequence and comparison sequence was calculated to evaluate battery SOH. The experiment verified that the maximum estimated error of this method was less than 4%.

3.2.2. Feature Screening Based on Data Dimensionality Reduction

In the evaluation of battery SOH, feature selection techniques grounded in data dimensionality reduction serve as powerful tools to eliminate redundant information, thereby enhancing both the efficiency and precision of predictive models. Among the most commonly employed methods are principal component analysis (PCA),[95,96] recursive feature elimination (RFE),[97] and other related approaches. Given that battery datasets are often characterized by high dimensionality and significant redundancy, RFE proves particularly effective in isolating the most influential features pertinent to SOH, thus streamlining the dataset and optimizing model performance. The key features selected by RFE can provide more valuable input to the model, thus improving the performance of the model in SOH estimation. PCA is an unsupervised linear dimensionality reduction algorithm,[98] whose goal is to transform the original high-dimensional data into a new set of low-dimensional data, namely the principal component, by linear transformation. These principal components are linear combinations of the original data and are orthogonal to each other, i.e. the correlation between them is zero. The SOH of a battery is affected by many factors, involving a large number of parameters, such as capacity, internal resistance, open circuit voltage and other related parameters. PCA can reduce the dimensionality of these high-dimensional data, simplifying the data structure while preserving the main information. Banguero et al.[99] conducted PCA on 26 parameters associated with battery capacity, internal resistance, and open-circuit voltage, ultimately preserving the first five principal components, which accounted for 80.25% of the total variance. This dimensionality reduction rendered the complex dataset more tractable, while simultaneously illuminating the most influential

features impacting battery SOH. Wu et al.[100] applied PCA to optimize six aging indicators derived from the durations of CC and CV phases in lithium-ion battery charging curves. The first two principal components demonstrated a notably high contribution to the aging characteristics across batteries, thereby significantly reducing input dimensionality without compromising key informational content. In a similar vein, Chen et al.[101] developed an algorithm to estimate battery capacity SOH using an open-source fast-charging dataset. By transforming time-domain data into histograms and leveraging PCA, the number of features fed into the feedforward neural network (FNN) was reduced by 80%, effectively eliminating redundant information and enhancing model efficiency. Yang et al.[102] applied PCA to process the real and imaginary part data of impedance at each frequency point, aiming to extract key features or quantify data uncertainty and 95% confidence ellipses are constructed to quantify uncertainty, thereby realizing the separation of random errors and systematic errors. The analysis reveals that a sampling window slightly longer than the signal length can improve the consistency in the medium and low frequency ranges; at low temperatures, the minor axis of the confidence ellipse should be used to evaluate uncertainty; an increase in the amplitude of the excitation current leads to a decrease in uncertainty. Additionally, the measurement uncertainties of both pseudo-random binary sequence and discrete-interval binary sequence are determined by the energy at each frequency point. These findings provide a basis for the optimization of binary sequences and the improvement of EIS measurement accuracy.

4. Data-Driven Approaches to Battery SOH Estimation

Battery SOH is influenced by a multitude of intricate factors, including internal electrochemical dynamics, external environmental conditions, and operational parameters. Traditional estimation approaches grounded in physical modeling often fall short of comprehensively and precisely capturing these complexities. In contrast, data-driven methodologies are capable of unlocking the latent value embedded within vast volumes of real-world operational data, which encapsulate the battery's behavior across varying temperatures, charge–discharge rates, and usage frequencies. As illustrated in **Figure 7**, a taxonomy of SOH prediction strategies for lithium-ion batteries is delineated. With ongoing advancements in battery technology, novel materials and structural configurations continue to emerge. The principal merit of data-driven methodologies lies in their ability to circumvent the need for mastering the intricate physicochemical models unique to each battery cell. Instead, they enable the construction of estimation models rooted in empirical usage data from specific cells. This model-agnostic characteristic has garnered widespread attention from researchers both domestically and internationally. Data-driven techniques are broadly categorized into two classes: nonprobabilistic models and probabilistic models.[103–106] With regard to SOH estimation, nonprobabilistic models infer battery health based on predefined rules or explicit functional relationships. Among these, support vector machine (SVM),[107–109] artificial neural network (ANN),[110,111] and their associated topologies are extensively applied across the literature. In contrast, probabilistic models explicitly account for the uncertainties inherent in battery SOH prediction. By leveraging probability distributions to characterize battery health, these models offer a nuanced depiction of potential SOH outcomes along with their corresponding likelihoods. Representative approaches within this category include Gaussian process regression (GPR)[112,113] and relevance vector machines (RVM),[114] etc.

4.1. Nonprobability Model

4.1.1. ANN

An ANN is a computational paradigm inspired by the architecture of the human brain's nervous system,[115–118] as illustrated in **Figure 8**, and serves as a pivotal tool in the estimation of battery SOH. Initially, diverse forms of data are introduced into the input layer, after which they are propagated forward through the network. At each neuron, the weighted sum of inputs is computed, augmented by a bias term, and subsequently transformed via an activation function. This output then serves as the input for the subsequent layer, continuing through the network until reaching the final output layer. ANN are adept at capturing the complex nonlinear relationships between battery SOH and a multitude of interdependent variables—an endeavor where conventional linear models often fall short. In scenarios involving dynamic battery behavior under varying operational conditions, such as rapid charging/discharging or fluctuating temperatures, tailored ANN architectures—like recurrent neural network (RNN) and their derivatives—can effectively adapt to and model these variations, thereby delivering precise and reliable SOH estimation. In the field of battery SOH estimation, the method based on neural network has achieved a more detailed classification. Among them, the mainstream methods can be roughly classified into five categories: FNN,[119,120] whose structure is relatively simple, and the signal is one-way transmitted from the input layer to the output layer, showing certain advantages in processing some simple mapping relationships. RNN,[121] with unique memory function, can effectively use the time sequence information in the sequence data, which is very consistent with the dynamic change characteristics of data during the battery operation. convolutional neural network (CNN)[122,123] is good at extracting local features of data. If the battery data is properly structured, key information can be accurately captured with the help of its convolutional layer and pooling layer. Deep neural network (DNN), with its complex multilayer architecture, can deeply explore the hidden law of battery data and achieve a more accurate estimate of SOH. And hybrid neural network (HNN),[124] by integrating the advantages of different types of neural networks and integrating multiple network structures and algorithms, to show greater adaptability and flexibility in dealing with complex and variable battery SOH estimation tasks. Although traditional SVM and GPR can take advantage of some features related to battery SOH, there are limitations in the utilization of raw data, and it is difficult to fully tap the rich information

Research Article
doi.org/10.1002/batt.202500620
Chemistry Europe
European Chemical Societies Publishing

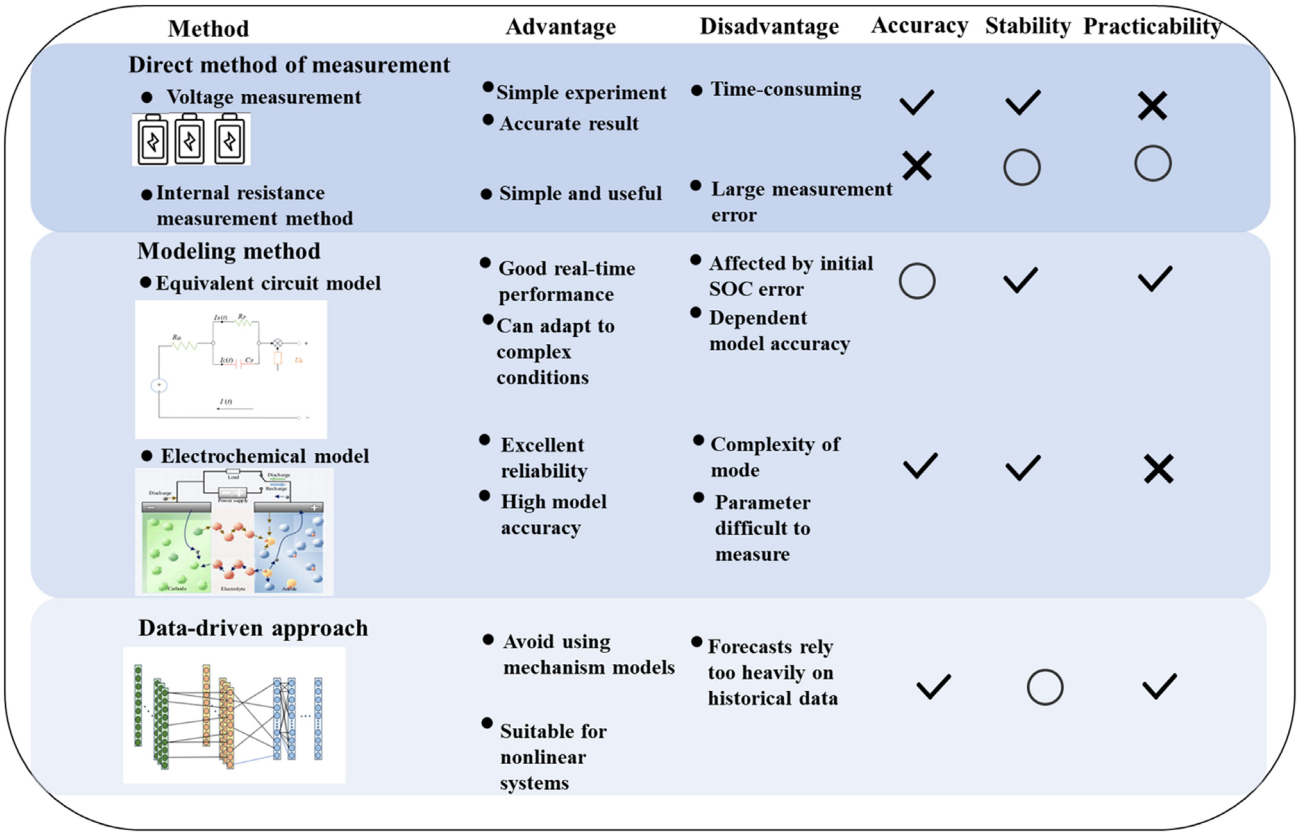

Figure 7. Classification of SOH prediction methods for lithium-ion batteries.

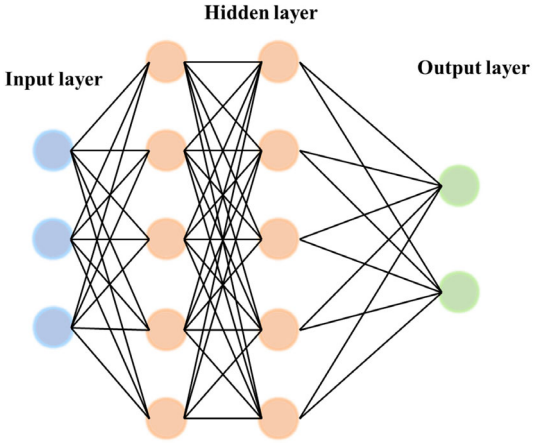

Figure 8. Basic structure diagram of ANN.

contained in it. Although deep learning has the advantage of automatic learning, it exposes many defects in the process of data preprocessing and feature fusion, and it is difficult to realize the efficient use of data and the organic integration of features. Lu et al.[125] constructed a CNN architecture based on feature fusion, aiming to estimate the SOH of battery based on voltage measurement data of part of its cycle. Different from the traditional approach of directly using voltage data as input, this method first extracts three key feature sequences, covering volume–voltage curves and their differentiation relative to voltage and cycle life, respectively, in order to dig more valuable information from the dimension of cycle and cycle, thereby improving the estimation accuracy. When fuzing multisource data, existing methods based on multidimensional features fail to make effective use of the internal relationships between the data and fuze the data before input to the model, which makes it difficult for the model to learn these relationships and affects the performance of SOH estimation. And in practical application, the data of battery charging and discharging are susceptible to electromagnetic interference to produce noise, and the noise type and level are different in different running stages. Early fusion data is easy to cause information interference and noise propagation, but the existing methods are not suitable for this, and the performance of the model will decline in the noisy environment. Therefore, Chen et al.[126] proposed an attention-mechanism-based multi-feature fusion framework (AM-MFF) to achieve more accurate, robust and interpretable SOH estimation of lithium-ion batteries by means of independent CNN module to extract features, attention-mechanism to capture relationships among features, interpretation of input contributions, and data enhancement to resist noise. The experimental results fully prove the advantages of the framework compared with traditional methods under different conditions, and provide a more effective technical solution for lithium-ion battery management.

From the perspective of the trade-off between model complexity and real-time applicability, different ANN architectures exhibit

significant differences: FNNs have a small number of parameters[119] and a single-sample inference time of only 5–10 ms, making them suitable for resource-constrained scenarios such as consumer electronics. However, their simple "input–output" one-way propagation structure fails to capture the time-series dependencies of battery data, resulting in a MAE of 2.5%–3.8% for SOH estimation in the NASA dataset,[120] which indicates insufficient accuracy. RNNs enhance time-series modeling capabilities through memory units, reducing the MAE to 1.8%–2.2% in the CALCE dataset.[121] Nevertheless, their number of parameters increases to 10^5–10^6, and the inference time extends to 15–25 ms. Additionally, multilayer stacking of RNNs is prone to gradient vanishing, leading to a 15%–20% accuracy degradation in the late cycling stage of batteries. Deep learning models, such as CNN-long short-term memory network (LSTM),[127] combine convolutional operations for local feature extraction with LSTM for capturing temporal dependencies. These models achieve a mean absolute error between 0.8% and 1.2% on the NASA dataset. However, their parameter counts surges to the range of 10^6–10^7, and the inference time exceeds 30 milliseconds—significantly surpassing the 50 ms real-time threshold required for onboard battery management systems (BMS). Furthermore, these models require a substantial amount of data spanning 200–500 cycles,[128] which limits their generalization capability in real-world EV scenarios characterized by random charging and fragmented data. A trade-off phenomenon is commonly observed where improving accuracy leads to increased complexity and reduced real-time responsiveness. This phenomenon highlights a key limitation of artificial neural network architectures in engineering applications, and simultaneously achieving both high accuracy and real-time performance still remains a significant challenge to this day.

4.1.2. SVM

SVM is a supervised machine learning algorithm mainly used for classification problems.[109,129] SVM, as a deterministic algorithm, has been widely used in battery health diagnosis and prediction. It has many obvious advantages: its estimation accuracy is considerable, and it can analyze and judge the data accurately. The results show good robustness to outliers, even if there are some outliers with strong interference in the data, there will be no significant deviation to the overall diagnosis and prediction results. SVM, as a deterministic algorithm, has been widely used in battery SOH diagnosis and prediction. It has many obvious advantages: its estimation accuracy is considerable, and it can analyze and judge the data accurately. The results show good robustness to outliers, even if there are some outliers with strong interference in the data, there will be no significant deviation to the overall diagnosis and prediction results. In EVs applications, battery capacity and resistance degrade over time and use, affecting vehicle range and power, and traditional laboratory standard tests cannot be directly used in the vehicle. Therefore, Klass et al.[130] proposed a SOH estimation method based on SVM for battery models and virtual standard performance tests. Building a data-driven model using variables such as battery current, voltage, and temperature during EVs operation eliminates the need for extensive up-front lab work, battery specific knowledge, and additional equipment. For capacity and resistance, two key SOH indicators, they created different types of data sets for SVM training, including temperature-independent and temperature-dependent models, to adapt to different application scenarios. Although the existing data-driven SOH estimation methods have advantages, the method based on time domain signal cannot provide detailed information inside the battery, and the data collection is complex, high energy consumption, and the estimation is inaccurate. Moreover, when intelligent algorithms are used to optimize the parameters of the model, traditional algorithms such as grey wolf optimization (GWO) algorithm have linear convergence factors, which are difficult to balance the global and local search capabilities, easy to fall into local optimality, and do not consider the strict social rank of wolves when determining the final position, which affects the accuracy and speed of battery SOH estimation Therefore, Liu et al.[129] constructed a novel data-driven model that combines kernel support vector machine (KSVM) and nonlinear GWO (NGWO). In this model, the key hyperparameters of KSVM are optimized by NGWO, and the local optimization dilemma is successfully overcome by introducing new nonlinear convergence factors and dynamic weighting scheme to improve the traditional GWO, and the performance SOH estimation is greatly enhanced. SVM can also be used for regression problems, called support vector regression (SVR),[131–133] the core idea of which is to find an optimal hyperplane that separates different classes of data points as much as possible. The SVR algorithm is used to constrain the optimization problem shown in (Equation 8), and satisfies (Equation 9).

$$\min \frac{1}{2}\omega^T\omega + C\sum_{i=1}^{1}(\xi_i + \xi_i^*) \tag{8}$$

$$\begin{cases} y_i - (\boldsymbol{\omega}^{\mathrm{T}}\mathbf{X}_i + b) \le \varepsilon + \xi_i^* \\ s.t. \begin{cases} (\boldsymbol{\omega}^{\mathrm{T}}\mathbf{X}_i + b) - y_i \le \varepsilon + \xi_i \\ \xi_i,\ \xi_i^* \ge 0,\ \ i = 1, \ldots, I \end{cases} \end{cases} \tag{9}$$

where $\boldsymbol{\omega}$ represents the model parameter, y_i is the actual value, $\boldsymbol{X}_i$ is the eigenvector, b is the biased term, ξ_i and ξ_i^* is the relaxation variable and must be non-negative. SVR is very effective in dealing with high-dimensional data and nonlinear regression problems, which has attracted the attention of many scholars at home and abroad. Wei et al.[134] took the lead in constructing a battery health assessment model based on SVR. They used battery capacity as the state variable, selected the capacity-related coefficient B and the charging capacity in the CV stage as input variables to establish an SVR-based capacity degradation model, and further expanded it to form a joint prediction framework for both battery SOH and RUL, laying a foundation for the application of SVR in the field of battery health assessment. Building on this, Li et al.[135] addressed the issue that the key parameters of the SVR model are susceptible to initial settings and difficult to achieve global optimization. They proposed a lithium-ion battery SOH estimation method that combines the improved Antlion Optimization Algorithm with SVR. By using improved Antlion Optimization Algorithm to perform global

optimization on the key parameters of the SVR model, they effectively made up for the limitations of the SVR model established by Wei et al.[132] in parameter adjustment, and further improved the performance of the SVR model in SOH estimation tasks.

4.1.3. LSTM

LSTM,[136–138] as a special type of RNN family, it is designed to overcome two major problems that traditional RNN faces when dealing with long series data, namely, gradient disappearance and gradient explosion. These two problems have seriously restricted the ability of traditional RNN to learn and capture long-term dependance on data for a long time. LSTM successfully solves these difficulties with its unique design architecture, enabling the model to learn and accurately capture long-term dependencies in long series data more efficiently.[139] The core structure of LSTM is the memory unit, which connects the whole network and can realize the efficient transfer of information between different time steps. There are three important gating structures on the memory unit, namely input gate, forgetting gate and output gate, as shown in **Figure 9**. The calculation process of LSTM is shown in (Equation 10).

$$\begin{cases} f_t = \sigma\left(B_f + \overline{W}_f x_t + V_f h_{t-1}\right) \\ i_t = \sigma\left(B_i + \overline{W}_i x_t + V_i h_{t-1}\right) \\ C_t^* = \tanh\left(B_C + \overline{W}_C x_t + V_C h_{t-1}\right) \\ C_t = C_{t-1} f_t + i_t C_t^* \\ O_t = \sigma\left(B_o + \overline{W}_o x_t + V_O h_{t-1}\right) \\ h_t = O_t \tanh(C_t) \end{cases} \tag{10}$$

where *f*, *i*, *O*, and *C* represent the memory door, input door, output door, and storage unit in turn. *B* indicates bias, *W* represents the weight of x_t the input, *V* indicates the weight of the hidden state h_{t-1} at the last time. σ is sigmoid activation function, can limit the output value in the range of 0–1, tanh is a hyperbolic tangent function.

Multilayer LSTM networks are prone to degradation, which affects model performance. Therefore, Xu et al.[127] proposed an improved CNN-LSTM model (CNN-LSTM-SKIP), the model integrates the exceptional local feature extraction capability of CNNs with the dynamic modeling advantage of LSTM in capturing temporal dependencies. Furthermore, drawing on the cascading concept of dense networks, it incorporates skip connections to enhance information flow and gradient propagation. Mitigate degradation of multilayer LSTM The model structure includes 1D convolution layer, pooling layer, LSTM layer and fully connected layer, which improves information transmission efficiency through skip connection. Although the data-driven method can directly use data for prediction, it has strict requirements on data. In practice, due to the uncertainty of household charging behavior, it is difficult to obtain complete charge–discharge cycle data and voltage, current and other data in a specific range, which limits this method. Due to the strong randomness of battery charging curve, it is difficult to extract aging information from random charging curve segment. Therefore, Zhang et al.[128] proposed a multi-layer LSTM algorithm with attention mechanism, which can make use of charging curve segments of arbitrary length and relative position, and is more suitable for actual working conditions. In data processing, the charging stage is divided first, and then the original data is segmented and normalized. With the improved LSTM algorithm as the base learner, several base learners focus on different parts of the charging curve respectively, and use BP neural network as a meta-learner to integrate the output of the base learner to obtain the final SOH estimate. Experiments show that the estimation error of this method is only about 1% when only the charging curve

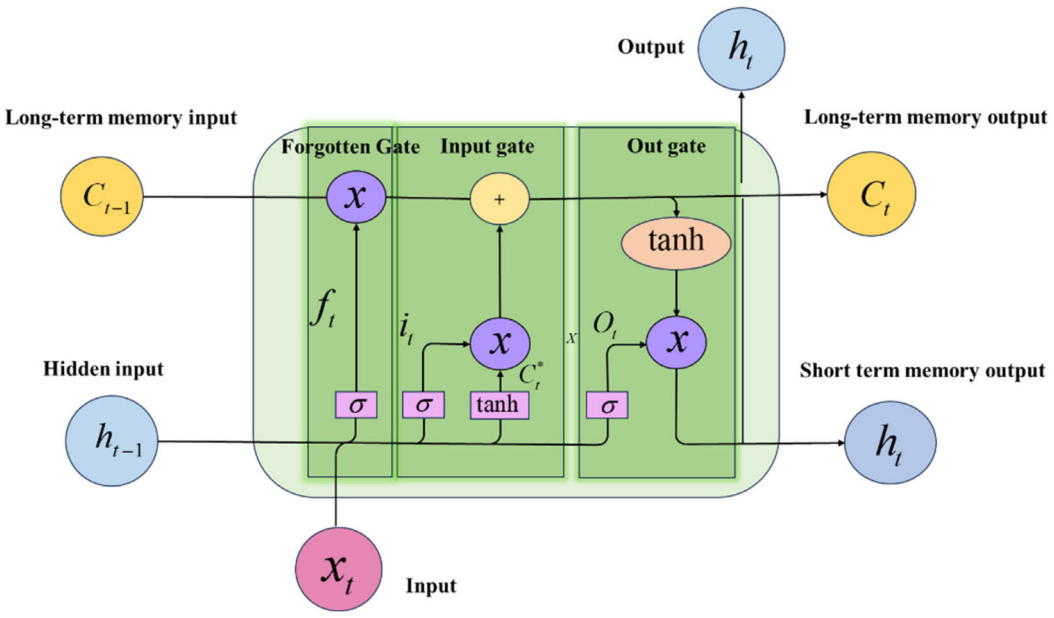

Figure 9. LSTM basic structure.

section is used. Tian et al.[97] proposed the RFE-DBiLSTM model. Based on the RFE method of SVM, this model evaluates the importance of features through iterative training, gradually eliminates features that have little impact on the target variable, screens out key features, thereby reducing the computational complexity of the model and improving its prediction performance. By stacking multilayer bidirectional LSTM units, the DBiLSTM component can capture forward and reverse time-series information simultaneously, effectively addressing the nonlinear and dynamic changes in battery data, and enhancing the ability to process complex time-series data and capture long-term dependencies. However, traditional methods have limitations in handling complex battery aging processes, especially in noisy environments. To address this issue, Tao et al.[140] made further improvements: they extended the hidden layer structure of the LSTM network and separated the gating weight parameters with positive and negative correlations to reduce the risk of overfitting. Meanwhile, they proposed a new network weight update algorithm that combines the Extended Kalman Filter (EKF) and Gradient Descent. By leveraging the inherent noise suppression capability of EKF, this algorithm effectively improves the robustness of the model, making it more suitable for complex and noisy battery operating conditions.

4.1.4. Transformer

The transformer model is an advanced sequence-to-sequence encoder–decoder architecture in the field of deep learning.[141,142] It breaks with tradition and does not use a circular structure, but uses a multihead attention mechanism to accurately capture complex connections between input sequences. This innovative design enables the Transformer model to generalize to a wide range of data sets of different sizes. Compared with traditional models, Transformer model has obvious advantages in processing long series data and has excellent parallel processing capability, which greatly improves processing efficiency.[57,143] This advantage has enabled it to achieve remarkable results in many fields such as natural language processing and computer vision. SOH estimation of lithium-ion batteries is critical to the reliability and safety of BMS. Traditional methods have the problems of time consuming, high cost or insufficient generalization ability. Gu et al.[144] combined the advantages of CNN and transformer. CNN extracts local features in layers through convolution operation, and transformer aggregates global features through self-attention mechanism. The combination of the two can capture both local and global information of battery data. Chen et al.[145] took the lead in exploring the application of transformer architecture in battery SOH estimation, targeting the limitation that the original vision transformer (VIT)—initially designed for computer vision tasks—cannot be directly applied to SOH estimation. They proposed a modified regression vision transformer (RVIT) network: first, by calculating the Pearson correlation coefficient between HFs and SOH, they screened out sampling points with high correlation to SOH, then grouped these points and determined the optimal number of sampling points through analyzing training speed and precision. To address VIT's incompatibility with battery data formats, a dimension conversion layer was added to RVIT to transform 1D battery data into 3D data, enabling multi-to-one or one-to-one SOH estimation; additionally, a trainable regression token was introduced to collect SOH-related information, which was then input into the transformer encoder combined with positional encoding, and SOH prediction was finally realized via the MLP regression layer. This work verified Transformer's potential in high-precision SOH estimation, as RVIT achieved excellent performance on NASA and CALCE datasets with prediction curves closer to the true values, outperforming other conventional methods. Building on Chen exploration of transformer in SOH estimation, Zhang et al.[58] further optimized the transformer architecture to address the demand for multistep SOH prediction. They proposed an improved transformer that integrates dilated causal convolution (DCC) with multihead attention mechanisms, reconstructing a new encoder–decoder structure specifically for extracting battery SOH features. By framing battery SOH prediction as a sequence-to-sequence mapping problem, this improved architecture enables multistep prediction. Comparative experiments showed that, compared with the basic transformer, temporal convolutional network (TCN), LSTM, and CNN-LSTM, Zhang et al.[58] modified transformer not only maintained high accuracy but also exhibited better generalization ability in multistep SOH prediction—effectively expanding the application scope of transformer-based methods in battery health assessment and it fills the gap in dynamic health trend prediction left by Chen et al.[145]

4.2. Probabilistic Model

4.2.1. GPR

GPR is a random process in which any combination of finite random variables obeys the joint normal distribution.[146,147] In the regression problem, the function to be predicted is assumed to be a Gaussian process. This is determined by specifying a mean function (such as a zero-mean function) and a covariance function (kernel function). Common covariance functions are the radial basis function (RBF), which describes the correlation between different input points, and its parameters affect the smoothness of the function. When the observed data is available, the posterior distribution of the function under the given data can be calculated according to Bayes' theorem. The posterior distribution is still a Gaussian process, but the mean and covariance are updated according to the observed data. PR does not need to make assumptions about functional forms and can be flexible to you and various data. GPR can naturally provide uncertainty in SOH estimates.[148] In battery management, this is very important. For example, when the uncertainty is high, it indicates that the current battery usage conditions may be beyond the range covered by the training data, or there may be an abnormal condition in the battery, which can alert the user or management system to conduct a deeper examination. The relationship between the SOH of a battery and the conditions of use is often nonlinear. G As a nonparametric model, GPR does not need to make assumptions about this relationship, and can fit complex nonlinear relationships well.[112,149] For example, the SOH attenuation law of the battery may be different at different temperatures

and charge and discharge ratios, and GPR can effectively capture these complex changes. Qian et al.[113] took the lead in addressing the issues of poor data separability and heteroscedasticity caused by outliers in traditional GPR when processing high-dimensional data. They proposed a weighted GPR (WGPR) model based on an improved local outlier factor (LOF). By introducing the EWM to enhance the LOF algorithm, this model not only solved the problem of ineffective data differentiation in high-dimensional spaces but also assigned weights to data samples, thereby eliminating heteroscedasticity induced by outliers and laying a foundation for the application of GPR in complex data scenarios. Building on this, Buchanan et al.[150] further overcame the reliance of traditional GPR on battery usage history and proposed a CNN-GPR model that combines CNN with GPR. In this model, CNN is responsible for extracting key features from battery data, while GPR performs SOH estimation based on these features. This model can not only handle random partial charging events commonly encountered in practical scenarios but also provide confidence intervals for prediction results—a feature that is particularly crucial for industries such as microgrids and transportation, where batteries are often not fully charged during use. Additionally, the output of the predicted error range enables the BMS to operate more accurately under partial charging conditions, making up for the limitation of the model in Qian et al.,[113] which failed to consider the randomness of actual charging. Xiong et al.,[151] on the other hand, focused on the problem of data loss in random charging scenarios and proposed an SOH estimation method for lithium-ion batteries based on GPR and LSTM network. Specifically, LSTM excels at handling dynamic dependencies in time-series data, while GPR can provide reliable probabilistic predictions when data is incomplete. The combination of the two enables the GPR-LSTM method to effectively estimate the SOH of lithium-ion batteries in random charging scenarios and maintain high accuracy even when some HFs are missing. This further enhances the applicability of GPR in complex real-world operating environments and forms a complete research chain, covering key aspects such as data preprocessing optimization, operational condition adaptability improvement, and enhanced robustness against data loss.

4.2.2. RVM

RVM is a sparse probability model based on Bayesian learning theory,[152,153] which can be used for classification and regression. It is similar to support vector machine in that it uses kernel function to map data to high-dimensional feature space and transform low-dimensional nonlinear relationship into high-dimensional linear relationship to deal with nonlinear problems. However, correlation vector machines have unique advantages that are different from support vector machines. RVM can accurately screen out weakly correlated sample points, reduce data scale, and improve processing efficiency.[114] When constructing kernel functions, it does not need to set penalty factors, and the expression shown in (Equation 11) can flexibly create various forms of kernel functions, expanding the application scope of the model and improving the adaptability.

$$kx - x_c = \exp\left\{\frac{-x - x_c^2}{2b^2}\right\} \tag{11}$$

where x represents any point in the prediction space, x_c is the center point of the kernel function, usually a fixed point or reference point, b is the bandwidth parameter, which controls the width or smoothness of the kernel function. In particular, compared with SVM, the solution obtained by RVM is sparser. This not only reduces the complexity of model calculations and reduces the need for storage space, but also significantly improves the flexibility and resilience of models in complex data environments. Therefore, RVM has shown great potential and value in many machine learning application scenarios. To address the issue that the single kernel function of traditional RVM cannot accurately fit the battery capacity degradation law, Lyu et al.[154] proposed a hybrid kernel relevance vector machine (HKRVM). By fuzing global kernel functions and local kernel functions, and leveraging a weight assignment mechanism to give full play to the advantages of different kernel functions, HKRVM can more accurately characterize the battery capacity degradation law. Meanwhile, to optimize the weights and kernel parameters in HKRVM, the genetic GWO (GGWO) algorithm was introduced. This algorithm enhances the global search capability of the model and significantly improves its optimization performance, making the model more efficient and accurate when dealing with battery capacity-related issues. Ultimately, it realizes the synchronous prediction of battery SOH and RUL during the charging process, laying a foundation for the application of RVM in battery health management. Building on this, Chen et al.[155] further overcame the limitations of traditional RVM, such as insufficient accuracy and generalization ability in battery SOH estimation, and proposed an optimized RVM method based on the dynamic integrated Bat Algorithm (BA-RVM). A dynamic ensemble learning strategy was introduced: experimental data of different batteries were divided into multiple training subsets, which were used to train multiple BA-RVM sub-models, respectively. At the same time, a dynamic weight update method was proposed to update the weights of the sub-models in real time according to the similarity between the sample battery and the battery under test. This enables the model to more accurately reflect the actual state of the battery, thereby achieving efficient online estimation of battery capacity. Its effectiveness has been fully verified using NASA battery data, making up for the shortcomings of the model proposed by Lyu et al.,[154] which failed to consider the impact of individual battery differences on estimation results. Chen et al.,[156] on the other hand, focused on the problems of battery capacity regeneration and randomness in single prediction, and proposed a hybrid algorithm BLS-RVM based on BLS and RVM. By integrating the strong feature learning ability of BLS with the probabilistic prediction advantage of RVM, and combining the EMD algorithm to denoise experimental data, the above problems were effectively solved Meanwhile, validation across multiple datasets demonstrates that the model exhibits high accuracy and strong generalization capability, significantly enhancing the applicability of stochastic voltage models under complex battery operating conditions. This effort further establishes a comprehensive research

framework encompassing kernel function design, optimization strategies, data preprocessing, and model integration.

4.2.3. Probabilistic Neural Network (PNN)

PNN is a kind of feedforward neural network based on Bayesian decision theory,[157] which performs nonparametric probability density estimation through the Parzen window method and is particularly suitable for fast pattern classification tasks. This network adopts a four-layer structure: The input layer receives feature vectors, the pattern layer measures the similarity between the input and the training samples using RBFs (usually Gaussian kernels), the summation layer integrates the probability density distributions of various categories, and finally the output layer implements classification decisions based on the maximum a posteriori probability criterion. PNN has advantages such as high training efficiency (training can be completed with a single forward computation), support for online learning, and progressive convergence to Bayesian optimal solutions. However, it also has limitations such as high storage overhead (all training samples need to be retained) and sensitive kernel function parameters, making it more suitable for classification scenarios of small and medium-sized datasets, such as industrial fault diagnosis and medical image analysis. Lin et al.[158] estimated the SOH of lithium-ion batteries using PNN. By conducting life cycle tests of CC/CV charging and CV discharging on 110 lithium-cobalt batteries, the electrical characteristics of charging and discharging as well as the life cycle test data were obtained to estimate the SOH. The research finds that the CC charging time, the instantaneous voltage drop at the beginning of discharge and the open-circuit voltage are important characteristics for estimating SOH. Li et al.[159] proposed a flexible estimation framework based on PNNs for the capacity estimation problem of lithium-ion batteries in multidimensional charging scenarios of EVs. By constructing the CLPNN model including the CNN-LSTM and the probabilistic regression layer, the spatio-temporal features are extracted from the charging data at short time intervals and the capacity mean value and uncertainty estimation are output. Two result fusion strategies, namely pruning the mean and uncertainty screening, are introduced to optimize the estimation accuracy of multiple scenarios. For the first time, probabilistic layer transfer learning was adopted to solve the problem of scarce data in the target domain. Crossdomain knowledge transfer was achieved by fixing the CNN-LSTM layer and fine-tuning the probabilistic layer through the pretrained model of the source domain.

4.3. Overview of the Estimation Results from Different Algorithms

Traditional physics-based models often struggle to deliver precise estimations due to the intricate interplay of internal electrochemical phenomena and external environmental variables. In contrast, data-driven approaches demonstrate superior adaptability and flexibility by uncovering latent patterns embedded within real-world operational datasets. Nonprobabilistic models—such as ANN, SVM, LSTM networks, and transformer architectures—evaluate battery SOH through explicit functional mappings or heuristic rules, excelling particularly in the interpretation of nonlinear, high-dimensional time-series data. Probabilistic models—such as GPR and RVM—quantify the inherent uncertainty in SOH estimation through probability distributions, thereby delivering more resilient and nuanced predictions. Moreover, this study highlights the growing significance of hybrid frameworks that synergistically integrate the strengths of multiple modeling paradigms alongside novel enhancements such as attention mechanisms and dynamic weight optimization. These innovations collectively contribute to elevated accuracy, robustness, and generalization capabilities. An overview of the diverse methodologies for SOH estimation is provided in the referenced literature and illustrated in **Figure 10**. Some data-driven methods are shown in **Table 1**, some evaluation indexes are added, such as the RMSE (root mean square error), MAPE (mean absolute percentage error), MAE (mean absolute error). Nonetheless, no single approach proves universally optimal; each comes with its own set of advantages and limitations. Consequently, the practical resolution of SOH estimation challenges necessitates a holistic and judicious selection and fine-tuning of algorithms and parameters tailored to specific application contexts.

5. Expectation

In the current era of rapid technological development, data-driven methods have brought new opportunities for battery SOH estimation. In the future, its development in this field is expected to make progress in the following key directions, as shown in **Figure 11**.

Multisource data fusion and in-depth mining. Most contemporary data-driven models continue to depend heavily on conventional parameters such as battery voltage, current, and temperature—readily available but inherently limited. Looking ahead, a broader spectrum of data sources is poised to be integrated, including acoustic and optical signals, as well as insights into the microstructural composition of battery materials. By synthesizing these diverse modalities, a more holistic and nuanced representation of the battery's internal state can be achieved, thereby significantly enhancing the precision and dependability of SOH estimations.

Reinforcement of few-shot learning and transfer learning. In real-world application scenarios, acquiring extensive high-quality battery data proves to be a formidable challenge, often demanding substantial investments of time and resources. Consequently, the emergence of few-shot learning has underscored its strategic advantages. By refining algorithmic structures and learning mechanisms, such models can adeptly discern the nuanced aging behaviors of batteries and accurately capture the evolving patterns of SOH, even when limited data is available. Moreover, transfer learning is poised to gain broader adoption, enabling models to swiftly adapt to novel battery deployment environments by leveraging insights gleaned from analogous systems or related domains—thereby mitigating the dependency on large-scale datasets.

Integration with physical models. By integrating data-driven methodologies with physics-based battery models, the

Research Article
doi.org/10.1002/batt.202500620
Chemistry Europe
European Chemical Societies Publishing

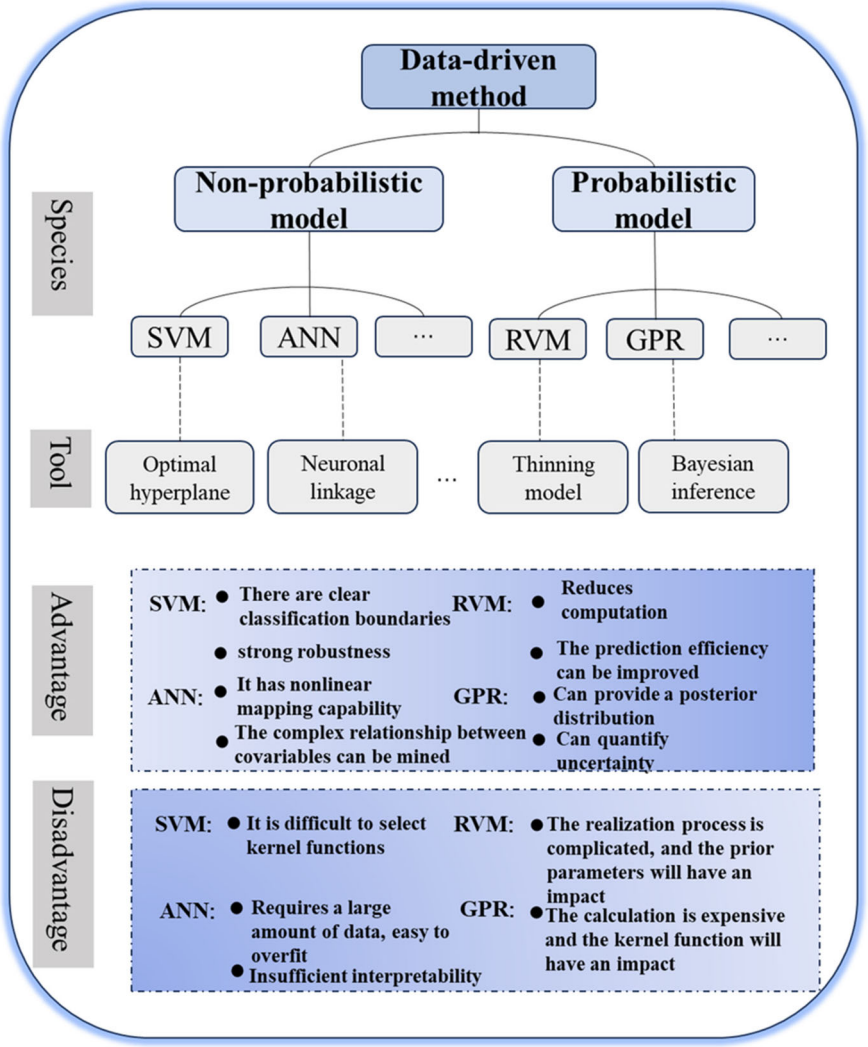

Figure 10. Advantages and disadvantages of several data-driven approaches.

complementary strengths of both paradigms can be maximally harnessed. While physical models elucidate the underlying electrochemical mechanisms within the battery, offering a rigorous theoretical foundation, data-driven approaches excel in compensating for the limitations of these models under intricate and variable operating conditions. Moreover, they enhance the model's adaptability and predictive fidelity. This synergistic fusion paves the way for a more holistic, robust, and precise estimation of battery SOH. To better bridge theoretical proposals with practical applications, a concrete and feasible implementation pathway is the development of a physics-informed hybrid neural network framework. Such a framework can be specifically applied to critical issues such as predicting the growth kinetics of the solid electrolyte interface. For instance, classical physical equations describing SEI growth—such as diffusion-reaction equations—can be embedded into the objective function of the neural network in the form of regularization terms or constraints, thereby leveraging physical laws to guide the training process of the data-driven model. Meanwhile, key electrochemical parameters within the battery, including lithium-ion diffusion coefficients and reaction rate constants, can serve as physically interpretable outputs of the network, replacing traditional black-box features. Such a framework integrates physics-based knowledge with data-driven modeling, effectively combines the mechanistic interpretability of physical models with the powerful nonlinear fitting capability of neural networks, offering a clear implementation path toward building highly accurate, interpretable, and data-efficient models.

Crossdomain application expansion. The applicability of data-driven battery SOH estimation extends well beyond the realms of EVs and energy storage systems, reaching into diverse sectors such as aerospace and consumer electronics. In these domains, battery performance and safety are of equal paramount importance. By facilitating crossdomain expansion, data-driven methodologies offer unified technical support for battery management across varied industries, thereby catalyzing transformative advancements in battery technologies on multiple fronts.

Looking ahead, as battery big data platforms become increasingly sophisticated and AI algorithms undergo continual

Research Article
doi.org/10.1002/batt.202500620
Chemistry Europe
European Chemical Societies Publishing

Table 1. Some data-driven methods.

Methodology		References	Error indicator	Advantage	Challenge
Nonprobabilistic model	CNN	[125]	MAE $<$ 0.28% MAPE $<$ 0.32%	High data efficiency	Poor interpretability
	KSVM	[129]	RMSE $\approx$ 0.72% MAE $\approx$ 0.50%	Strong robustness	Difficulty in generalizing across multiple temperatures and limited data
	LSTM	[128]	MAPE $\approx$ 0.93%	Strong temporal feature extraction	A single model struggles to adapt to multiple charging stages
	Transformer	[144]	MAE $\approx$ 0.27% RMSE $\approx$ 0.34%	The prediction accuracy is extremely high	High computational complexity
Probability model	WGPR	[113]	MAE $\leq$ 0.72% RMSE $\leq$ 0.92%	High computational efficiency	The ability to handle high-dimensional features needs to be improved
	HKRVM	[154]	RMSE $\approx$ 0.95%	Intelligent optimization of weights and parameters	High computational complexity
	PNN	[158]	MAE $\approx$ 0.39%	High training efficiency	Generalized limitations

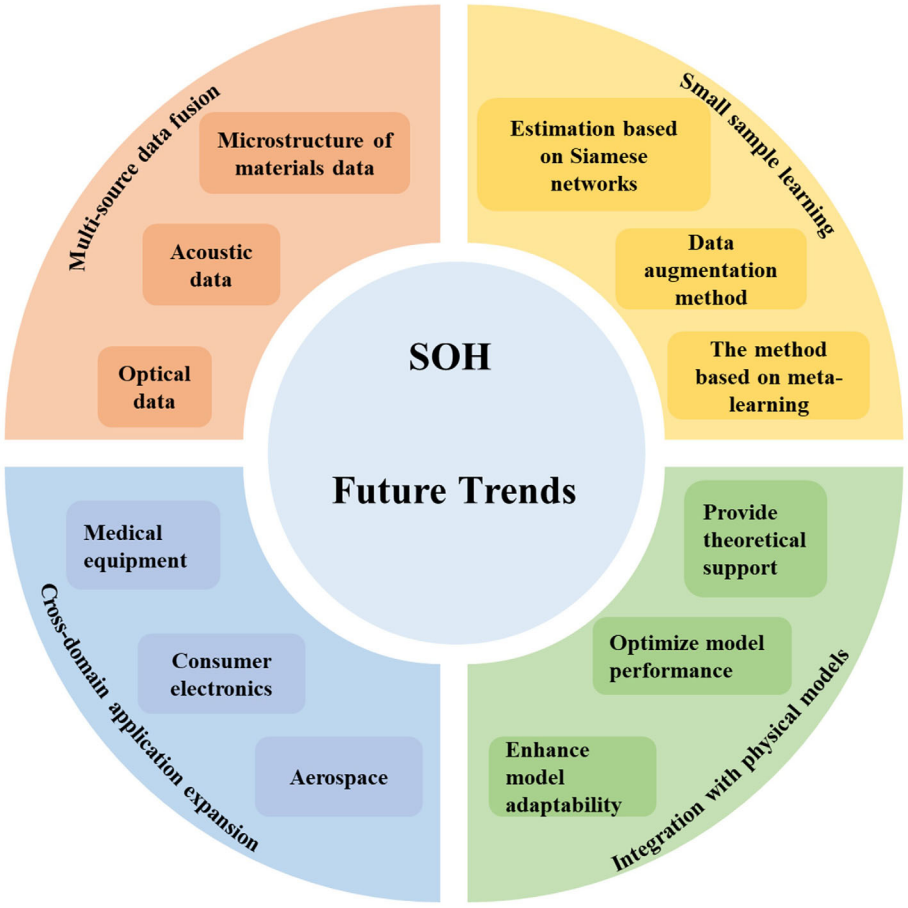

Figure 11. Future trends in battery SOH estimation.

refinement, data-driven approaches will more adeptly address the challenges of SOH prediction under complex operating conditions. These advancements will contribute significantly to enhancing battery safety, prolonging service life, and optimizing second-life applications—thereby offering robust support for the achievement of carbon neutrality objectives.

6. Conclusion

As the core component throughout the entire operational lifecycle of EVs, the accurate evaluation of the SOH of lithium-ion batteries directly determines vehicle safety, operational stability, and overall energy efficiency. Focusing on lithium-ion batteries,

this paper first systematically elaborates on two mainstream SOH definitions based on capacity fade and internal resistance increase and then conducts an in-depth exploration of the research progress in data-driven methods for SOH estimation.

This study provides a comprehensive examination of key factors influencing battery SOH estimation, feature engineering techniques, and data-driven methodologies through establishing a coherent feature-model matching logic, thereby offering a reliable and systematic reference framework: first, understanding the inherent degradation mechanisms of cathode materials; second, designing feature extraction strategies sensitive to these mechanisms; and finally, selecting the model most suitable for learning from the resulting feature sets. This shift from a generic, trial-and-error approach to a tailored, physics-informed methodology represents a significant step forward for the field.

It is also necessary to objectively acknowledge that while the current research has made progress in establishing a technical system for data-driven SOH estimation, it still has limitations in terms of the disconnection between laboratory scenarios and real-world operating conditions. For instance, it relies heavily on data from single cells under controlled laboratory conditions, failing to fully address the needs of engineering applications—such as the consistent degradation of battery packs, crossdomain adaptation of multiple cathode materials, and multimodal data fusion. This also exposes the pain point in the transformation from laboratory technology to engineering application. To bridge this gap, future research should focus on overcoming three key challenges: diversifying data scenarios beyond single-cell experiments, enhancing material generalization capabilities, and developing robust methods for practical multisource data fusion. Ultimately, this will provide critical technical support for optimizing the full life-cycle management of lithium-ion batteries and achieving carbon neutrality goals.

Existing research on data-driven SOH estimation, while having established a solid technical framework, is limited by its primary reliance on controlled laboratory data. This creates a notable disconnect between research findings and real-world engineering applications, which require addressing complex issues such as battery pack degradation, crossdomain adaptation for various cathode materials, and multimodal data fusion. To bridge this gap, future research should focus on overcoming three key challenges: diversifying data scenarios beyond single-cell experiments, enhancing material generalization capabilities, and developing robust methods for practical multisource data fusion. Such advancements are crucial for optimizing the full life-cycle management of lithium-ion batteries and supporting global carbon neutrality initiatives.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (grant no. 62303007, 62503002), and the Doctoral Research Start-up Funding (grant no. S020318015/028) of Anhui University.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Zhiqiang Lyu: conceptualization (lead); writing—review & editing (lead). **Xiang'en Li**: methodology (lead); writing—original draft (lead). **Zhirui Jin**: investigation (lead); visualization (lead). **Hao Wang**: software (lead); validation (lead). **Yuan Chen**: conceptualization (supporting); supervision (lead); writing—review & editing (equal). **Longxing Wu**: investigation (supporting); writing—review & editing (supporting).

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Keywords: battery management · data-driven approaches · feature engineering · lithium-ion batteries · state of health estimation

[1] S. Liu, M. Bai, X. Tang, W. Wu, M. Zhang, H. Wang, W. Zhao, *Chem. Eng. J.* **2021**, *417*, 128997.
[2] S. Zhang, R. Cao, X. Pu, A. Zhao, W. Chen, C. Song, Y. Fang, *J. Energy Chem.* **2024**, *92*, 162.
[3] X. Pan, A. Chojnacka, F. Béguin, *Energy Storage Mater.* **2021**, *40*, 22.
[4] L. Kong, Y. Luo, S. Fang, T. Niu, G. Chen, L. Yang, R. Liao, *Green Energy Intell. Transp.* **2025**, *4*, 100192.
[5] M. Kapetanović, A. Núñez, N. van Oort, R. M. P. Goverde, *Energy Convers. Manage.* **2024**, *303*, 118202.
[6] B. Pang, H. Zhu, Y. Tong, Z. Dong, *J. Energy Storage* **2024**, *101*, 113963.
[7] S. Pushparaj, M. Sanjai, R. Gopalakrishnan, R. Radhamani, *2024 9th Int. Conf. on Communication and Electronics Systems (ICCES), Coimbatore, India* **2024**, pp. 323–330, https://doi.org/10.1109/ICCES63552.2024.10859535.
[8] R. R. Kumar, C. Bharatiraja, K. Udhayakumar, S. Devakirubakaran, K. S. Sekar, L. Mihet-Popa, *IEEE Access* **2023**, *11*, 105761.
[9] Y. Pan, J. Meng, S. Zhang, W. Liu, J. Fu, Q. Peng, B. Xiao, Z. Song, *J. Energy Storage* **2025**, *127*, 117147.
[10] M. Wang, S. Wu, Y. Chen, W. Luan, *Appl. Energy* **2025**, *378*, 124909.
[11] Optimization of objective temperature for battery heating at low temperatures to improve driving range and battery lifetime of battery electric vehicles.
[12] C. Wang, N. Cui, Z. Cui, H. Yuan, C. Zhang, *J. Energy Storage* **2022**, *53*, 105075.
[13] Z. He, X. Ji, H. Ding, Y. Li, G. Xu, Y. Dai, H. Lou, *J. Power Sources* **2025**, *631*, 236235.
[14] M. Leng, K. Liu, Y. Gao, H. Chen, P. Yan, X. Liu, *J. Energy Storage* **2024**, *97*, 112682.
[15] Y. Xie, S. Wang, G. Zhang, P. Takyi-Aninakwa, C. Fernandez, F. Blaabjerg, *J. Energy Chem.* **2024**, *97*, 630.
[16] The Atomic and Electronic Structure Changes Upon Delithiation of LiCoO2: From First Principles Calculations.
[17] M. Itagaki, N. Kobari, S. Yotsuda, K. Watanabe, S. Kinoshita, M. Ue, *J. Power Sources* **2005**, *148*, 78.
[18] Cobalt in high-energy-density layered cathode materials for lithium ion batteries.
[19] W. Yang, W. Chen, H. Zou, J. Lai, X. Zeng, Y. Zhang, X. Zeng, K. Ding, S. Zhang, L. Ma, Z. Li, Q. Zheng, *Angew. Chem., Int. Ed.* **2025**, *64*, e202424353.
[20] Z. Li, H. Yi, X. Li, P. Gao, Y. Zhu, *ACS Appl. Mater. Interfaces* **2024**, *16*, 28537.
[21] M. Jiang, D. L. Danilov, R. A. Eichel, P. H. L. Notten, *Adv. Energy Mater.* **2021**, *11*, 2103005.

[22] Y. Jiang, F. Guo, L. Qiu, T. Liu, Y. Hu, W. Yang, Y. Liu, Y. Sun, Z. Wu, Y. Song, X. Guo, *ACS Appl. Mater. Interfaces* **2023**, *15*, 35072.
[23] A. Patel, M. L. Rasche, S. Mallick, S. Kim, M. Jiang, M. P. Paranthaman, H. Lopez, R. B. Gupta *Energy Adv.* **2025**, https://doi.org/10.1039/d5ya00032g.
[24] X. Fan, Z. Zhang, G. Mao, Y. Tong, K. Lin, H. Tong, W. Wei, Q. Tian, X. Guo, *Rare Met.* **2023**, *42*, 2993.
[25] J. Yan, H. Huang, J. Tong, W. Li, X. Liu, H. Zhang, H. Huang, W. Zhou, *Interdiscip. Mater.* **2022**, *1*, 330.
[26] C. Hu, M. Geng, H. Yang, M. Fan, Z. Sun, R. Yu, B. Wei, *Coatings.* **2024**, *14*, 832.
[27] H. Sharifi, B. Mosallanejad, M. Mohammadzad, S. M. Hosseini-Hosseinabad, S. Ramakrishna, *Ionics* **2022**, *28*, 213.
[28] S. F. Schuster, T. Bach, E. Fleder, J. Müller, M. Brand, G. Sextl, A. Jossen, *J. Energy Storage* **2015**, *1*, 44.
[29] E. Sarasketa-Zabala, F. Aguesse, I. Villarreal, L. M. Rodriguez-Martinez, C. M. López, P. Kubiak, *J. Phys. Chem. C* **2015**, *119*, 896.
[30] M. Kassem, C. Delacourt, *J. Power Sources* **2013**, *235*, 159.
[31] L. Wang, J. Qiu, X. Wang, L. Chen, G. Cao, J. Wang, H. Zhang, X. He, *Escience* **2022**, *2*, 125.
[32] J. Sun, X. Liu, X. Li, S. Chen, S. Xing, Y. Guo, *Energy* **2025**, *315*, 134472.
[33] X. Qiang, W. Liu, Z. Lyu, H. Ruan, X. Li, *Green Energy Intell. Transp.* **2024**, *3*, 100169.
[34] Q. Li, W. Xue, *J. Energy Storage* **2025**, *112*, 115453.
[35] Adaptation of an Electrochemistry-based Li-Ion Battery Model to Account for Deterioration Observed Under Randomized Use.
[36] https://www.nasa.gov/intelligent-systems-division/discovery-and-systems-health/pcoe/pcoe-data-set-repository/ (accessed: October 2024).
[37] Y. Chen, D. Liu, X. Ding, H. Jiang, *Adv. Eng. Inf.* **2024**, *59*, 102316.
[38] H. Dai, J. Wang, Y. Huang, Y. Lai, L. Zhu, *Renewable Energy* **2024**, *222*, 119907.
[39] https://calce.umd.edu/battery-data (accessed: October 2024).
[40] Y. Xing, E. W. M. Ma, K. Tsui, M. Pecht, *Microelectron. Reliab.* **2013**, *53*, 811.
[41] J. Feng, F. Cai, H. Li, K. Huang, H. Yin, *Process Saf. Environ.* **2023**, *180*, 601.
[42] H. Pang, K. Chen, Y. Geng, L. Wu, F. Wang, J. Liu, *Energy* **2024**, *293*, 130555.
[43] S. Saxena, C. Hendricks, M. Pecht, *J. Power Sources* **2016**, *327*, 394.
[44] https://www.nature.com/articles/s41586-020-1994-5 (accessed: November 2024).
[45] https://www.nature.com/articles/s41560-019-0356-8 (accessed: November 2024).
[46] J. Du, C. Zhang, S. Li, L. Zhang, W. Zhang, *Energy* **2024**, *297*, 131276.
[47] X. Li, D. Yu, S. B. Vilsen, D. I. Stroe, *J. Energy Chem.* **2024**, *92*, 591.
[48] B. Zhao, W. Zhang, Y. Zhang, C. Zhang, C. Zhang, J. Zhang, *Appl. Energy* **2024**, *358*, 122325.
[49] M. Dubarry, C. Truchot, B. Y. Liaw, *J. Power Sources* **2012**, *219*, 204.
[50] S. Li, C. Zhang, J. Du, L. Zhang, Y. Jiang, *Appl. Energy* **2025**, *377*, 124634.
[51] A. Sonthalia, J. S. Femilda Josephin, E. G. Varuvel, A. Chinnathambi, T. Subramanian, F. Kiani, *Energy* **2025**, *317*, 134569.
[52] J. Tian, J. Zhang, H. Luo, C. Huang, M. Chow, Y. Jiang, S. Yin, *2024 IEEE 33rd Int. Symposium on Industrial Electronics (ISIE), Ulsan, Republic of Korea* **2024**, pp. 1–6, https://doi.org/10.1109/ISIE54533.2024.10595821.
[53] B. Ospina Agudelo, W. Zamboni, F. Postiglione, E. Monmasson, *Energy* **2023**, *263*, 125637.
[54] K. Q. Zhou, Y. Qin, C. Yuen, *J. Energy Storage* **2024**, *100*, 113502.
[55] Y. Li, M. Abdel-Monem, R. Gopalakrishnan, M. Berecibar, E. Nanini-Maury, N. Omar, P. Van den Bossche, J. Vab Mierlo*J. Power Sources* **2018**, *373*, 40.
[56] Y. Li, M. Abdel-Monema, R. Gopalakrishnan, M. Berecibar, E. Nanini-Maury, N. Omar, P. Van den Bossche, J. Vab Mierlo, *J. Power Sources* **2018**, *393*, 230.
[57] Z. Xu, Z. Chen, L. Yang, S. Zhang, *Appl. Soft Comput.* **2024**, *165*, 112072.
[58] Y. Zhang, Z. Li, L. Kong, H. Xu, H. Shen, M. Chen, *J. Energy Storage* **2025**, *105*, 114538.
[59] G. Wang, N. Cui, C. Li, Z. Cui, H. Yuan, *J. Energy Storage* **2023**, *73*, 109010.
[60] J. He, X. Bian, L. Liu, Z. Wei, F. Yan, *J. Energy Storage* **2020**, *29*, 101400.
[61] K. Li, N. Xie, *Energy* **2024**, *303*, 131888.
[62] Y. Liu, Y. Liu, H. Shen, L. Ding, *J. Energy Storage* **2025**, *110*, 115347.
[63] R. Sun, J. Chen, B. Li, C. Piao, *Energy* **2025**, *319*, 134756.
[64] X. Zhang, X. Gao, L. Duan, Q. Gong, Y. Wang, X. Ao, *Appl. Energy* **2025**, *377*, 124404.
[65] M. Berecibar, M. Garmendia, I. Gandiaga, J. Crego, I. Villarreal, *Energy* **2016**, *103*, 784.
[66] M. Spielbauer, J. Soellner, P. Berg, K. Koch, P. Keil, C. Rosenmüller, O. Bohlen, A. Jossen, *J. Energy Storage* **2022**, *55*, 105564.
[67] F. Xia, K. Wang, J. Chen, *J. Energy Storage* **2023**, *64*, 107161.
[68] F. Guo, L. D. Couto, G. Mulder, K. Trad, G. Hu, O. Capron, K. Haghverdi, *J. Energy Storage* **2024**, *101*, 113850.
[69] D. Qin, L. Li, H. Wu, J. Chen, J. Yang, *J. Energy Storage* **2024**, *97*, 112921.
[70] Fifth-order resistance-capacitance-based optimal equivalent circuit model of lithium-ion batteries with improved transient search optimization algorithm.
[71] C. Wang, M. Yang, X. Wang, Z. Xiong, F. Qian, C. Deng, C. Yu, Z. Zhang, X. Guo, *J. Energy Storage* **2025**, *110*, 115346.
[72] M. A. A. Mohamed, T. F. Yu, G. Ramsden, J. Marco, T. Grandjean, *J. Energy Storage* **2025**, *113*, 115581.
[73] G. Zhu, M. Hu, C. Qiu, K. Deng, *Energy* **2025**, *322*, 135694.
[74] S. Amir, M. Gulzar, M. O. Tarar, I. H. Naqvi, N. A. Zaffar, M. G. Pecht, *IEEE Access* **2022**, *10*, 18279.
[75] M. Bao, D. Liu, Y. Wu, Z. Wang, J. Yang, L. Lan, Q. Ru, *Electrochim. Acta* **2024**, *494*, 144449.
[76] A. Alhammadi, A. Fetyan, R. Agung Susantyoko, I. Mustafa, M. O. Bamgbopa, *J. Energy Chem.* **2025**, *102*, 329.
[77] D. Zhang, S. Tang, L. D. Couto, Z. Wei, V. Viswanathan, *Automatica* **2025**, *173*, 112092.
[78] L. Fang, S. Liu, F. Cheng, *Comput. Electr. Eng.* **2024**, *118*, 109481.
[79] H. Gong, Y. Li, J. Zhang, B. Zhang, X. Wang, *Eng. Appl. Artif. Intell.* **2024**, *131*, 107865.
[80] Z. Li, Y. Yang, L. Li, D. Wang, *J. Energy Storage* **2023**, *60*, 106584.
[81] L. Zhang, N. Wu, S. Han, J. Xiao, S. Ruan, *2023 IEEE Int. Conf. on Energy Internet (ICEI), Shenyang, China* **2023**, pp. 340–345, https://doi.org/10.1109/ICEI60179.2023.00071.
[82] X. Chen, H. Yang, Y. Yang, H. Li, Y. Chen, C. Zhang, L. Wu, L. Yan, *Green Energy Intell. Transp.* **2025**, 100339, https://doi.org/10.1016/j.geits.2025.100339
[83] J. Jiang, X. Zhang, Z. Yuan, *Expert Syst. Appl.* **2024**, *249*, 123633.
[84] X. Li, Y. Ma, Q. Zhou, X. Zhang, *Appl. Soft Comput.* **2024**, *167*, 112253.
[85] I. Couso, O. Strauss, H. Saulnier, *Fuzzy Sets Syst.* **2018**, *343*, 50.
[86] A. Marsick, H. André, I. Khelf, Q. Leclère, J. Antoni, *Mech. Syst. Signal Process.* **2024**, *216*, 111486.
[87] W. Qu, W. Shen, J. Liu, *J. Energy Storage* **2021**, *42*, 103102.
[88] C. Yang, X. Xia, B. Peng, Z. Wang, H. Zhang, E. Liang, *Energy* **2024**, *311*, 133360.
[89] K. Chen, J. Li, K. Liu, C. Bai, J. Zhu, G. Gao, G. Wu, S. Laghrouche, *Green Energy Intell. Transp.* **2024**, *3*, 100151.
[90] Correlation based-graph neural network for health prognosis ofnon-fully charged and discharged lithium-ion batteries.
[91] Z. Chen, H. Zhao, Y. Zhang, S. Shen, J. Shen, Y. Liu, *J. Power Sources* **2022**, *521*, 230892.
[92] Z. Bao, J. Nie, H. Lin, J. Jiang, Z. He, M. Gao, *Energy* **2023**, *282*, 128306.
[93] Z. Xu, N. Xie, H. Diao, *Energy* **2023**, *283*, 129167.
[94] X. Li, Z. Wang, L. Zhang, C. Zou, D. D. Dorrell, *J. Power Sources* **2019**, *410–411*, 106.
[95] Y. Cheng, C. Lu, T. Li, L. Tao, *Energy* **2015**, *90*, 1983.
[96] L. Xuan, L. Qian, J. Chen, X. Bai, B. Wu, *IEEE Access* **2020**, *8*, 164693.
[97] F. Tian, S. Chen, X. Ji, J. Xu, M. Yang, R. Xiong, *Int. J. Electrochem. Sci.* **2025**, *20*, 100891.
[98] M. Schmid, C. Endisch, *J. Energy Storage* **2022**, *53*, 104815.
[99] E. Banguero, A. Correcher, Á. Pérez-Navarro, E. García, A. Aristizabal, *Renewable Energy* **2020**, *146*, 2438.
[100] M. Wu, Y. Zhong, J. Wu, Y. Wang, L. Wang, *Energy* **2023**, *283*, 129061.
[101] J. Chen, P. Kollmeyer, F. Chiang, A. Emadi, *2023 IEEE Transportation Electrification Conf. & Expo (ITEC), Detroit, MI, USA* **2023**, pp. 1–6, https://doi.org/10.1109/ITEC55900.2023.10187012.
[102] K. Yang, Z. Feng, Y. Zhu, Y. Zhang, B. Feng, Z. Song, J. Meng, *J. Energy Storage* **2025**, *131*, 117617.
[103] Y. Li, S. Wang, L. Chen, C. Qi, C. Fernandez, *Energy* **2023**, *282*, 128776.
[104] F. Meng, F. Yang, J. Yang, M. Xie, *Reliab. Eng. Syst. Saf.* **2023**, *237*, 109361.
[105] A. Tang, Y. Jiang, Y. Nie, Q. Yu, W. Shen, M. G. Pecht, *Energy* **2023**, *279*, 128137.
[106] C. Zhang, L. Luo, Z. Yang, S. Zhao, Y. He, X. Wang, H. Wang, *Green Energy Intell. Transp.* **2023**, *2*, 100108.
[107] An SVM-Based Health Classifier for Offline Li-Ion Batteries by Using EIS Technology.
[108] Online data-driven battery life prediction and quick classification based on partial charging data within 10 min.
[109] A. Nuhic, T. Terzimehic, T. Soczka-Guth, M. Buchholz, K. Dietmayer, *J. Power Sources* **2013**, *239*, 680.
[110] W. Li, A. Li, A. C. Yin Yuen, Q. Chen, T. B. Yuan Chen, I. M. De Cachinho Cordeiro, P. Lin, *Appl. Therm. Eng.* **2025**, *259*, 124874.
[111] C. Mehta, A. V. Sant, S. P. Optimized, *Green Energy Intell. Transp.* **2024**, *3*, 100175.

[112] F. Li, Y. Min, Y. Zhang, Y. Zhang, H. Zuo, F. Bai, *Reliab. Eng. Syst. Saf.* **2024**, *242*, 109787.
[113] C. Qian, N. He, L. He, R. Li, F. Cheng, *Eng. Appl. Artif. Intell.* **2024**, *138*, 109314.
[114] C. Wang, X. Zhang, X. Yun, X. Meng, X. Fan, *Energy* **2023**, *285*, 129466.
[115] L. Driscoll, S. de la Torre, J. A. Gomez-Ruiz, *J. Energy Storage* **2022**, *50*, 104584.
[116] Y. Luo, K. Lu, *J. Energy Storage* **2022**, *52*, 105062.
[117] S. Zhang, B. Zhai, X. Guo, K. Wang, N. Peng, X. Zhang, *J. Energy Storage* **2019**, *26*, 100951.
[118] Y. Zou, Z. Lin, D. Li, Z. Liu, *J. Energy Storage* **2023**, *73*, 109069.
[119] C. Chen, R. Xiong, R. Yang, W. Shen, F. Sun, *J. Cleaner Prod.* **2019**, *234*, 1153.
[120] R. Guo, W. Shen, *Energy* **2022**, *254*, 124270.
[121] S. Ansari, A. Ayob, M. S. Hossain Lipu, A. Hussain, M. H. Md Saad, *Expert Syst. Appl.* **2024**, *238*, 121904.
[122] S. Wang, P. Wang, L. Wang, K. Li, H. Xie, F. Jiang, *J. Energy Storage* **2025**, *107*, 114952.
[123] Z. Yuan, T. Tian, F. Hao, G. Li, R. Tang, X. Liu, *J. Power Sources* **2024**, *609*, 234697.
[124] M. Kurucan, M. Özbaltan, Z. Yetgin, A. Alkaya, *Renewable Sustainable Energy Rev.* **2024**, *192*, 114262.
[125] Z. Lu, Z. Fei, B. Wang, F. Yang, *Energy* **2024**, *288*, 129690.
[126] S. Chen, J. Liu, H. Yuan, Y. Tao, F. Xu, L. Yang, *Appl. Energy* **2025**, *381*, 125116.
[127] H. Xu, L. Wu, S. Xiong, W. Li, A. Garg, L. Gao, *Energy* **2023**, *276*, 127585.
[128] L. Zhang, J. Zhang, T. Gao, L. Lyu, L. Wang, W. Shi, L. Jiang, G. Cai, *J. Energy Storage* **2023**, *74*, 109370.
[129] S. Liu, L. Fang, X. Zhao, S. Wang, C. Hu, F. Gu, A. Ball, *J. Energy Storage* **2024**, *102*, 114052.
[130] V. Klass, M. Behm, G. Lindbergh, *J. Power Sources* **2014**, *270*, 262.
[131] S. M. Jameel, J. M. Altmemi, A. A. Oglah, M. A. Abbas, A. H. Sabry, *J. Energy Storage* **2024**, *104*, 114554.
[132] Y. Zhang, Y. Liu, J. Wang, T. Zhang, *Energy* **2022**, *239*, 121986.
[133] T. Zhu, S. Wang, Y. Fan, N. Hai, Q. Huang, C. Fernandez, *Energy* **2024**, *306*, 132464.
[134] J. Wei, G. Dong, Z. Chen, *IEEE Trans. Ind. Electron.* **2018**, *65*, 5634.
[135] Q. Li, D. Li, K. Zhao, L. Wang, K. Wang, *J. Energy Storage* **2022**, *50*, 104215.
[136] L. Guo, H. He, Y. Ren, R. Li, B. Jiang, J. Gong, *Eng. Appl. Artif. Intell.* **2024**, *127*, 107317.
[137] J. Kim, D. Han, P. Lee, J. Kim, *Etransportation* **2023**, *18*, 100293.
[138] M. Lin, J. Wu, J. Meng, W. Wang, J. Wu, *Energy* **2023**, *268*, 126706.
[139] S. Kim, Y. Y. Choi, K. J. Kim, J. Choi, *J. Energy Storage* **2021**, *41*, 102893.
[140] J. Tao, S. Wang, W. Cao, Y. Cui, C. Fernandez, J. M. Guerrero, *Energy* **2024**, *312*, 133541.
[141] N. H. Paulson, J. Kubal, S. J. Babinec, *Cell Rep. Phys. Sci.* **2024**, *5*, 101928.
[142] J. Zhao, X. Han, Y. Wu, Z. Wang, A. F. Burke, *J. Energy Chem.* **2025**, *102*, 463.
[143] Z. Li, X. Zhang, W. Gao, *Energy* **2024**, *311*, 133418.
[144] X. Gu, K. W. See, P. Li, K. Shan, Y. Wang, L. Zhao, K. C. Lim, N. Zhang, *Energy* **2023**, *262*, 125501.
[145] L. Chen, S. Xie, A. M. Lopes, X. Bao, *Int. J. Electr. Power* **2023**, *152*, 109233.
[146] X. Li, C. Yuan, X. Li, Z. Wang, *Energy* **2020**, *190*, 116467.
[147] Y. Yang, S. Chen, T. Chen, L. Huang, *J. Energy Storage* **2023**, *61*, 106797.
[148] Y. Zhou, G. Dong, Q. Tan, X. Han, C. Chen, J. Wei, *Energy* **2023**, *262*, 125514.
[149] S. Pohlmann, A. Mashayekh, F. Stroebl, D. Karnehm, M. Kuder, A. Neve, T. Weyh, *J. Energy Storage* **2024**, *88*, 111649.
[150] S. Buchanan, C. Crawford, *J. Energy Storage* **2024**, *76*, 109799.
[151] X. Xiong, Y. Wang, K. Li, Z. Chen, *J. Energy Storage* **2023**, *72*, 108390.
[152] W. Guo, M. He, *Appl. Soft Comput.* **2022**, *124*, 108967.
[153] S. Jia, B. Ma, W. Guo, Z. S. Li, *J. Manuf. Syst.* **2021**, *61*, 773.
[154] Z. Lyu, G. Wang, R. Gao, *Energy* **2022**, *251*, 123852.
[155] Z. Chen, S. Zhang, N. Shi, F. Li, Y. Wang, J. Cui, *Appl. Soft Comput.* **2022**, *129*, 109615.
[156] Z. Chen, N. Shi, Y. Ji, M. Niu, Y. Wang, *Energy* **2021**, *234*, 121269.
[157] Y. Che, Y. Zheng, Y. Wu, X. Sui, P. Bharadwaj, D. Stroe, Y. Yang, X. Hu, R. Teodorescu, *Appl. Energy* **2022**, *323*, 119663.
[158] H. Lin, T. Liang, S. Chen, *IEEE Trans. Ind. Inf.* **2013**, *9*, 679.
[159] Q. Li, J. Zhong, J. Du, Y. Yi, J. Tian, Y. Li, C. Lai, T. Lu, J. Xie, *Energy* **2024**, *294*, 130881.

Manuscript received: August 15, 2025
Revised manuscript received: September 18, 2025
Version of record online: