

Multi-modal framework for battery state of health evaluation using open-source electric vehicle data

Received: 6 May 2024

Accepted: 21 January 2025

Published online: 29 January 2025

Hongao Liu^{1,2,5}, Chang Li^{1,5}, Xiaosong Hu²✉, Jinwen Li², Kai Zhang³, Yang Xie¹, Ranglei Wu¹ & Ziyou Song⁴✉

Accurate, practical, and robust evaluation of the battery state of health is crucial to the efficient and reliable operation of electric vehicles. However, the limited availability of large-scale, high-quality field data hinders the development of the battery management system for state of health estimation, lifetime prediction, and fault detection in various applications. In this work, to gain insights into underlying factors limiting battery management system performance in real-world vehicles, we analyze the operational data of 300 diverse electric vehicles over three years to understand the disparities between field data and laboratory battery test data and their effect on state of health estimation. Furthermore, we propose a deep learning-based multi-modal framework to effectively leverage historical vehicle data for efficient, accurate, and cost-effective state of health estimation. The proposed paradigm exhibits considerable potential for numerous applications in state estimation and diagnostics in multi-sensor systems. Furthermore, we make the field data of these electric vehicles publicly available aiming to promote further research on the development of effective and reliable battery management systems for real-world vehicles.

Lithium-ion batteries, with their high energy densities and long life-spans, have emerged as a promising energy source, particularly in electric vehicles (EVs)^{1,2}, which are becoming increasingly popular. However, the electrochemical mechanism of lithium batteries leads to irreversible aging, which results in performance degradation in terms of capacity reduction and internal resistance increase. The aging of lithium-ion batteries can be attributed to the solid electrolyte interface growth, loss of active materials, and lithium plating^{3,4}. In an EV battery system that consists of substantial individual cells, cell-to-cell variation results in rapid degradation because of the barrel effect⁵. The state of health (SOH) is widely used as a metric to measure battery degradation. SOH is defined as the current internal resistance of the battery compared with that of a new battery. SOH is also defined as the ratio of

the current capacity to the initial capacity⁶. Accurate SOH estimation is critical to the development of the next-generation battery management system (BMS) to achieve superior safety and performance in EVs⁷. Battery aging reduces the available capacity, which limits the single-charge cruise range of EVs⁸. In practice, a battery reaches the end of its life when its SOH decreases to 80% or 70%. During the warranty period, manufacturers are responsible for repairing or replacing batteries. Therefore, the accurate and efficient estimation of SOH is a critical topic of research for EV manufacturers and consumers.

High-quality, detailed, and informative data are essential for battery SOH estimation and lifetime forecasting. Battery aging tests are lengthy and require hours of charging, resting, and discharging to induce degradation. Hundreds of charging and discharging cycles are

¹State Key Laboratory of Intelligent Vehicle Safety Technology, Chongqing, China. ²College of Mechanical and Vehicle Engineering, Chongqing University, Chongqing, China. ³School of Energy and Power Engineering, Chongqing University, Chongqing, China. ⁴Department of Mechanical Engineering, National University of Singapore, Singapore, Singapore. ⁵These authors contributed equally: Hongao Liu, Chang Li. ✉e-mail: xiaosonghu@ieee.org; ziyou@nus.edu.sg

required to accurately delineate degradation patterns and characteristics. Consequently, the availability of open-source battery datasets is critical for advancing battery research. Many publicly accessible battery testing datasets have significantly contributed to progress in this field. The CALCE team⁹ has provided extensive data from battery cycling tests under various loads and temperatures, encompassing 18,650 cells, A123 systems, and CS2 batteries, among others. Researchers from the University of Oxford¹⁰ have enriched the research community with several datasets, encompassing three distinct sets focused on battery degradation. Zhang et al.¹¹ compiled a comprehensive dataset featuring over 20,000 electrochemical impedance spectroscopy (EIS) spectra from commercial batteries across various SOH, state of charge (SOC), and temperatures. Severson et al.¹² contributed a dataset encompassing 124 commercial lithium-ion cells cycling under rapidly changing conditions. Li et al.¹³ proposed a time-series dataset comprising 48 cells aged under identical conditions and provided valuable insights into battery aging processes. Lu et al.¹⁴ and Weng et al.¹⁵ have curated datasets for researchers. Also, sponsored by the U.S. Department of Energy Office of Electricity, a regularly updated battery archive¹⁶ for the public is accessible. These repositories provide battery testing data under various conditions for use in battery degradation studies.

In addition to datasets obtained from laboratory settings, field data are crucial to advancing battery management studies in real-world environments. On the one hand, battery is a nonlinear, time-varying system, so limited lab testing will only generate a partial understanding of batteries. These real-world operational conditions may generate some extrapolation of battery characteristics obtained from lab testing, thereby generating significant errors in real-world BMS algorithms. On the other hand, constrained by experimental settings, battery tests in laboratories are typically conducted at the individual cell level, whereas the battery systems in EVs are complex assemblies of multiple cells connected in series and parallel and exhibit considerable individual cell inconsistencies. A multiplicity of signal sources and a complex set of factors exist. Therefore, field data are crucial for supporting BMS research. Researchers, automotive companies, and third-party platforms have collaborated to provide field data available to the public. Pozzato et al.¹⁷ obtained a dataset from an EV in use for over a year to provide considerable empirical evidence. Crucially, performance indicators closely linked to temperature variations were derived using this field data to improve our understanding of the environmental effect on battery health. Zhang et al.¹⁸ obtained a comprehensive battery field dataset from 347 EVs and used it to investigate the feasibility of detecting battery faults accurately through deep learning techniques. Deng et al.¹⁹ provided a dataset that includes the charging records of 20 EVs for more than 25 months. These publicly available datasets aid battery management research, encompassing health evaluation, lifetime prediction, and fault detection, among other areas.

Apart from the data basis, SOH estimation methods have been devised^{20–23}. Data-driven machine learning methods exhibit considerable potential to solve the large-scale SOH estimation problem^{24–26}. Features for battery health evaluation indicate the input of the machine learning models, which can be acquired from multiple sources, such as EIS analysis^{25,27}, incremental capacity/differential voltage (IC/DV) analysis²⁸, differential thermal voltammetry analysis²⁹, and vehicle operating status¹⁷. Furthermore, numerous machine learning algorithms^{30–32}, such as Gaussian process regression (GPR)³³, feedforward neural network (FNN)³⁴, convolutional neural network (CNN)^{14,35}, recurrent neural networks (RNN)³⁶, and long-short-term memory network (LSTM, a modified RNN method)^{19,37,38} have been developed for rapid SOH estimation and prediction. In these algorithms, deep neural networks are constructed to realize automatic feature extraction. For example, Jiang et al.³⁹ proposed an extreme learning machine-based method for SOH estimation and validated its performance on the

public dataset from NASA⁴⁰, CALCE⁹, and MIT¹². However, due to the lack of available data, this method has not been validated on field data. In addition, the model primarily investigated SOH estimation of battery cells using a series of point features as inputs, while failing to consider the complex characteristics of real-world battery modules/packs, such as inconsistencies among cells. Wang et al.⁴¹ proposed a physics-informed neural network for accurate estimation of battery SOH. The results indicated that features extracted from the current and voltage data during the constant current-constant voltage process before the battery is fully charged held promise for accurate SOH estimation. However, it is important to note that the actual charging process of EVs typically does not include constant voltage charging, and there were also remarkable uncertainties in constant current charging conditions, such as current fluctuations and measurement noises in practice. Pozzato et al.¹⁷ analyzed field data from EVs' battery packs and extracted three performance indicators that were easy to calculate and closely linked to battery performance. The analysis of the field data was very detailed, but the method for SOH estimation in real-world scenarios was not provided. Qi et al.⁴² extracted labeled capacity and multiple health features from two EVs, and a CNN-BiGRU (Bidirectional Gated Recurrent Unit)-based model was established to build the relationship between capacities and features. This model attempted to address the issue of SOH estimation based on field data. However, it merely used traditional point features as inputs without fully considering multi-modal information. Additionally, due to data limitations, the model has not been extensively validated on large-scale EVs with diverse driver behaviors. It should be noted that a significant challenge in applying machine learning to real-world EVs is the limited generalizability of algorithms across large numbers of vehicles, as different operational conditions produce various features that may significantly affect algorithm performance. Thus, it remains unclear from existing studies whether machine learning methods can consistently deliver reliable battery monitoring across a wide range of EVs operating under diverse conditions. This is a critical gap in the field, which this study aims to address comprehensively. Generally, regular machine learning methods and FNN exhibit excellent performance with point features that are acquired according to expertise and experience, and CNN is used in two-dimensional (2D) picture processing. LSTM is widely adopted in one-dimensional (1D) sequence feature extraction and sequence prediction, such as SOC estimation⁴³. The gating mechanism of LSTM enables it to learn long-term data dependencies and selectively retain useful information while ignoring irrelevant details. As a result, LSTM can effectively address the gradient vanishing problem and enhance generalization capabilities. However, the gating mechanism in LSTM also increases the number of network parameters, leading to issues such as being prone to overfitting, slow inference speed, and gradient explosion. Therefore, a simple model renders addressing the complex multi-source information obtained by EV BMS. Limited studies have focused on SOH estimation for operating EVs, especially under diverse, random driving conditions. Previous studies were limited by the volume and quality of field data. Furthermore, the model performance of existing methods is restricted by a single-modal input.

In this work, to address these challenges, we investigate the historical data from 300 diverse EVs monitored over three years (Fig. 1a). Based on the field data, we analyze the operational characteristics of real-world EVs and summarize the bottlenecks in transferring theoretical methods of SOH estimation to engineering applications. Furthermore, we detail factors influencing the accuracy of algorithms and introduce a multi-modal deep learning framework to accurately estimate the SOH of EV batteries using operational data. The promising outcomes highlight the potential of the proposed framework to revolutionize EV battery testing. The proposed method can considerably reduce time, expenditures, and energy and promote the development of intelligent BMS and cloud-based monitoring

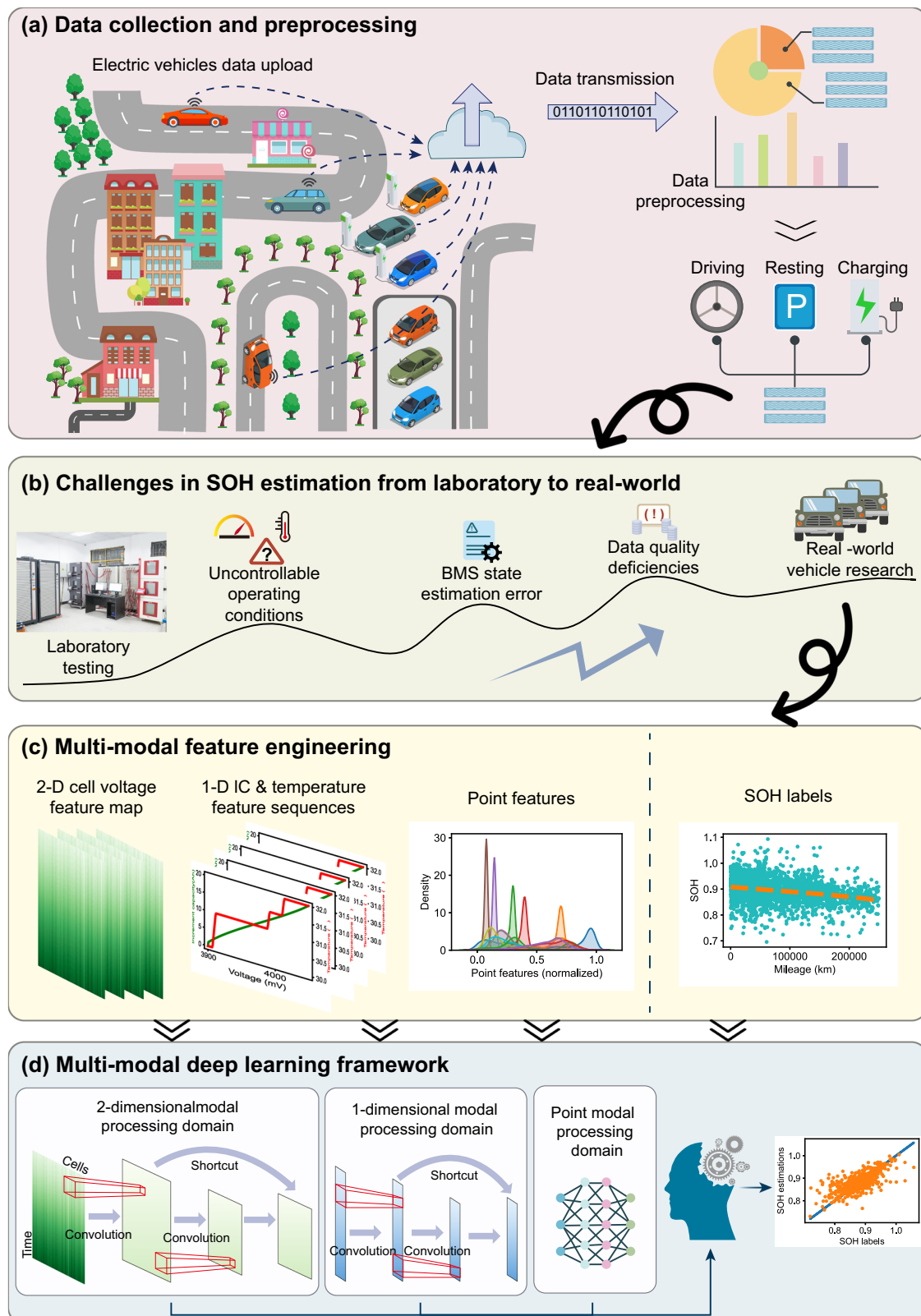


Fig. 1 | State of health (SOH) evaluation for electric vehicles (EVs) based on historical field data. **a** Raw operating data are uploaded to a cloud server from EVs. After data preprocessing, raw data are formulated into comma-separated values files. **b** Challenges in real-world EV SOH estimation. **c** Multi-modal feature engineering. Each SOH label is associated with one two-dimensional (2D) cell voltage

feature map, two one-dimensional (1D) feature sequences related to charge capacity and temperature, and fifteen feature points. **d** Multi-modal SOH estimation framework. The multi-modal features are used as the input of the framework and consist of a 2D processing domain, a 1D processing domain, and a point feature processing domain.

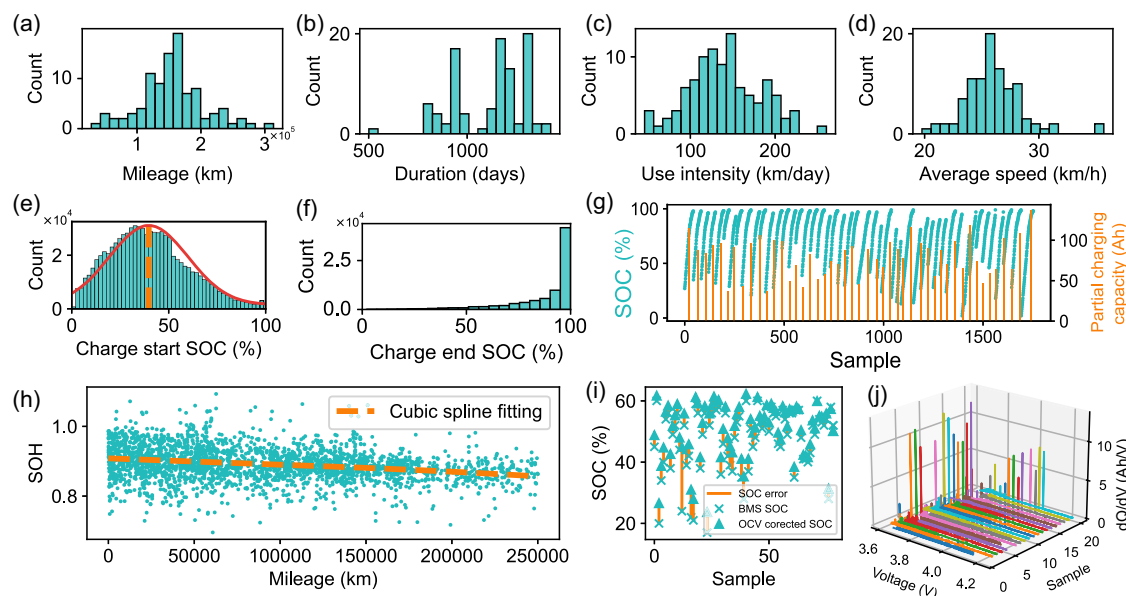


Fig. 2 | EV field data description. **a** Distribution of mileage. **b** Distribution of duration. **c** Distribution of use intensity quantified by driving distance per day. **d** Distribution of the average driving speed. **e** Distribution of charging start state of charge (SOC). **f** Distribution of charging end SOC. **g** EV charging example. Charging behaviors are randomly started and ended, with various partial charging capacities.

h State of health (SOH) degradation of the EVs. **i** SOC estimation error of the battery management system (BMS). **j** Incremental capacity analysis for EV charging process. Extracting peak features from each charging process is difficult, and the evolution of peak features with battery aging is not pronounced. Source data are provided as a Source Data file.

platforms to enhance EV utilization. To encourage the development of battery management algorithms, this large volume of high-quality vehicle data has been made publicly available. These data include complete driving and charging records for EVs, providing an abundant and robust data foundation for studying vehicle usage behaviors and battery degradation prediction.

Results

Real-world EV battery data disclosure and analysis

Full charging and discharging tests of batteries are crucial for designing an accurate SOH estimation method. In this test, the maximum available capacities of batteries are obtained by integrating the charging or discharging current over time. However, full charging and discharging do not occur during EV operation. By contrast, the bench test for EVs is time-consuming and requires considerable energy (see Supplementary Note 1 for the cost of offline SOH benchmark for EVs), rendering the method impractical for large-scale deployment and periodic, frequent SOH estimation for SOH tracking. A promising approach is to estimate SOH based on historical big data of EVs under real-world operations collected by BMS. In this study, we obtained operating data from 300 EVs equipped with NCM lithium-ion batteries over 3 years. The battery pack has a rated capacity of 155 Ah and consists of 96 cells connected in series. The operational conditions and usage intensity of this fleet of EVs are very diverse. Specifically, the driving data encompasses random charging and discharging patterns, including high-current discharging during acceleration, low-current discharging during cruising, and temporal charging during regenerative braking. The specific discharging conditions are determined by the driver's demands and the vehicle's driving status, making them inherently random. Furthermore, the charging data features two scenarios: high-rate multi-stage constant current fast charging, with charging rates reaching up to 0.8 C, and low-rate slow charging at ~ 0.15 C. For fast charging conditions, the current rate is controlled jointly by the battery pack temperature, the highest cell voltage, and the power limitation of the charging equipment. Unlike laboratory data, field data lacks idle-time records as the BMS stops recording when the EVs are turned off. These raw data include time stamps,

cumulative mileage, SOC, total current, pack voltage, highest/lowest cell voltage, and cell voltage. These data were generated at a sampling frequency of 0.1 Hz, and the information on the raw field data is presented in Supplementary Note 2.

Figure 2a displays the distribution of the total traveling mileage of vehicles. Most of these vehicles have been driven for more than 10,000 km and have been serviced for more than 800 days, as displayed in Fig. 2b. After prolonged use and considerable mileage accumulation, vehicle batteries exhibit notable degradation, providing numerous samples for battery degradation research. Figure 2c illustrates various driving intensities of this fleet of EVs. Figure 2d displays the given average driving speed. In contrast to the stable cycling conditions and testing environments in laboratories, the diverse driving styles of various users and road conditions result in considerable aging differences among vehicles. This diversity provides researchers with a broad, nondiscriminatory set of samples, and can be used to test the applicability and robustness of the developed algorithms. For the charging behavior, as depicted in Fig. 2e and f, most drivers tend to start charging the vehicles with the SOC under 50% and stop charging until the battery is fully charged. By data cleansing, we obtain the SOH labels of these vehicles with ampere-hour integration and open circuit voltage (OCV)–SOC correction. Figure 2h displays the SOH labels of these vehicles, and a general capacity decay pattern can be noticed by a cubic spline fitting curve. The method for capacity label calculation is elaborated comprehensively in the “Methods” and Supplementary Note 3.

SOH estimation in field scenarios has numerous challenges, such as complex and uncontrollable vehicle operating conditions, low accuracy of BMS state estimation, and data quality deficiencies. By randomly selecting the vehicle as an example, we investigated and revealed that field data characteristics differed from those of laboratory testing and detailed how these variances affect SOH estimation or other battery management tasks.

The complexity of real-world working conditions. The charging and discharging load of EVs almost entirely depends on the user's decision-making, and most users do not deplete their EVs' battery completely.

Therefore, acquiring the capacity labels through a full discharging–full charging approach during usage is not possible. Figure 2g displays the charging SOC intervals of users over a period and the corresponding calculable partial charging capacity. Without a true capacity value to refer to, assessing the level of capacity degradation in the vehicle becomes difficult. Second, in addition to aging degradation, the SOH of a battery is affected by factors such as temperatures and recent charging and discharging conditions, leading to the fluctuation of the calculated battery capacity. The involvement of multiple coupling factors renders conventional empirical aging models inapplicable and poses a challenge to the robustness of SOH estimation methods. Therefore, influencing factors should be comprehensively considered in SOH research on actual vehicles.

SOC estimation error of BMS. The accuracy of the state estimation of the BMS is influenced by various factors, including battery aging, data quality, and estimation algorithms. Among these factors, the estimation error in SOC considerably affects SOH estimation, especially when the vehicle does not undergo a full charging calibration for an extended period. Based on the stationary condition of the vehicle, we compare the differences between the OCV-corrected SOC and the BMS SOC, as displayed in Fig. 2i. The results indicate that with battery aging, the BMS cannot timely update battery capacity information, which results in frequent errors in SOC estimation. This phenomenon poses a challenge to the application of SOH calculation methods based on random local charging segments. To address this issue, SOC is calibrated by taking advantage of the rest period of vehicles to reduce the impact of SOC errors on the extraction of SOH labels. In addition, researchers have proposed a feedforward deep neural network method to achieve accurate and continuous SOC estimation with data noise interference⁴⁴, which is expected to further improve SOH estimation accuracy.

Limited data quality. Data quality has critical importance for the evaluation of battery health in vehicles. Although the data used in this study includes abundant information about the vehicle, field data still has limitations compared with the laboratory testing. The main limitations include data noise interference, missing data issues, and a relatively low sampling frequency. (1) Data noise interference: raw data contain noise due to factors such as lower sensor accuracy, electromagnetic interference, and data transmission defects. For example, during the charging process, the battery cell voltage, which should increase monotonically, may exhibit random fluctuations. In some cases, the battery voltage may suddenly drop to 0, even though the battery itself does not have any actual problems. Potential solutions include replacing the abnormal data through interpolation, monotonic adjustment, or filter-based smoothing methods. (2) Data missing: there are two types of data missing. The first type involves the absence of data associated with specific sensors, where the BMS fails to collect data from a particular sensor. The second type is the entire data frame being missing. Although the battery system is still operating, the corresponding data may not have been collected or recorded. Potential solutions include supplementing the missing data with interpolation as well as fitting methods, or considering the entire segment of data as unreliable and subsequently skipping the calculations for that segment. (3) Lower sampling frequency: due to limitations in BMS performance and data storage capacity, the sampling frequency of field data is considerably lower than that under laboratory conditions. A lower sampling frequency significantly impacts data analysis and calculations under dynamic conditions, such as SOC estimation or capacity calculation based on coulomb counting during driving conditions. In addition, because the BMS fails to capture the battery's response, some calculations that are sensitive to data accuracy, such as IC analysis, become difficult to perform. Potential solutions include increasing the data sampling frequency or prioritizing analyses that are

less sensitive to sampling frequency under limited conditions, such as converting IC analysis to charge capacity sequence analysis.

Furthermore, the charging profiles of EVs are typically limited by various factors, such as the BMS, charging station protocols, voltage levels, and temperatures. Although effective features⁴⁵, such as IC curves⁴⁶ and incremental voltage difference⁴⁷, can be extracted during the charging process, as displayed in Fig. 2j, peak features cannot be extracted from the IC curves of certain charging processes because of the current switching of multi-stage constant current charging. Moreover, peak features do not exhibit clear, regular changes with battery aging because the charging profile varies under various charging conditions. These factors hinder the application of some common laboratory features, reducing the feasibility and accuracy of SOH estimates based on these features.

The limited studies on SOH estimation using field data can be attributed to the aforementioned reasons, and maintaining the estimation accuracy achievable under laboratory conditions is difficult. Under current conditions, we proposed an SOH estimation framework to investigate the capability boundaries of SOH estimation algorithms with field data. Given its performance in multi-modal feature extraction and integration, the proposed framework can be extended to state estimation and diagnostics of multi-sensor systems. Furthermore, the public availability of large volumes of high-value field data can facilitate refined, efficient, and robust battery management.

Extracting multi-modal battery health indicators

Because of the complexity and uncertainty of EV operational conditions as well as the low sampling frequency, feature extraction based on the charging process is a suitable solution. This study proposes a multi-modal feature extraction method based on charging data to extract comprehensive and effective health indicators (HIs) that reflect the SOH of the battery for subsequent estimation.

First, the segment for HI extraction is determined by a fixed charging voltage interval. The charging strategy of these vehicles is multi-stage constant current charging in which the current is controlled by voltage and temperature. In this study, the voltage interval of [3900 mV, 4050 mV] is adopted, as displayed in Fig. 3a. This is the constant current segment that covers the largest voltage interval, ensuring that sufficient HIs can be obtained. This segment requires approximately 100 sampling points (1000 s), and its corresponding SOC interval is 71–84%, making it straightforward to capture specific charging behaviors and extract HIs.

Based on domain knowledge of lithium-ion batteries, it is observed that with battery decay, the battery voltage tends to rise more when the same amount of electricity is charged. In other words, within the same charging voltage range, the capacity charged into the battery will gradually decrease as the battery ages¹². Specifically, the peak and position of the IC curve during the charging process change under the coupled effect of loss of lithium inventory and loss of active material. As the area under the IC curve represents the charge capacity, this area diminishes as the battery ages. Thus, the charge capacity, reflected by the area under the IC curve, steadily decreases with battery degradation⁴⁸. Considering the charge capacity sequence as a function of voltage, a standardized feature sequence that indicates battery degradation can be formed, and this has been widely adopted in SOH estimation. As displayed in Fig. 3b, the charge capacity sequences Q_{HI} are extracted from the fixed charging voltage interval as the first HI sequence.

Temperature is a vital indicator of battery SOH. The usable capacity decreases significantly under cold working environments⁴⁹ on account of reduced ionic conductivity of the electrolyte, leading to slowed reaction kinetics of electrode materials, increased internal resistance, and structural changes in electrode materials. On the other hand, the internal resistance increases with battery degradation, leading to a variation in heat generation during charging, which can be

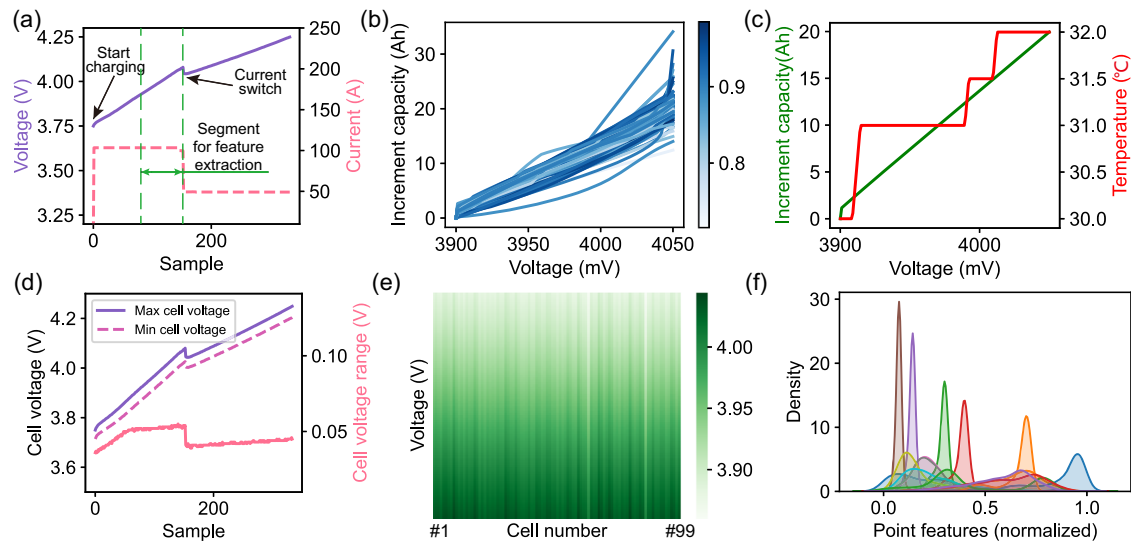


Fig. 3 | Multi-modal health indicators (HIs) extraction. **a** Voltage segment for HI extraction. **b** Charge capacity curves of various SOH samples. **c** HI sequences including temperature sequence and charge capacity sequence. **d** Maximum and

minimum cell voltage and cell voltage range during the charging process. **e** Cell voltage HI map. **f** Distribution of HI points after normalization. Source data are provided as a Source Data file.

captured by the temperature sensors of the battery system. Therefore, we analyze the temperature increment in the voltage interval and simultaneously extract the average temperature sequence of the battery as the second HI sequence. As illustrated in Fig. 3c, the two sequences constitute the one-dimensional modal HIs corresponding to an SOH label.

Second, considering the inconsistency of the cell voltage in the charging process, the voltage data of 96 cells are extracted from raw data. The cell inconsistency issue within a battery pack can be caused by various factors, including manufacturing inconsistency, the design of cooling structures, and balancing strategies, among others. Due to the barrel effect, the capacity of the battery pack is determined by the cell with the highest voltage during charging. Considering the inconsistency among cells, when one cell reaches the voltage limit, the remaining cells are not yet fully charged, thus limiting the charge intake of the battery pack. Furthermore, cell inconsistency leads to local heat accumulation in the cells with higher internal resistance and

the risk of overcharging/discharging, thereby accelerating battery aging. Therefore, it is necessary to consider cell inconsistency when estimating the battery health status⁵⁰. With the interpolation method, a 96×96 two-dimensional pseudo-HI map presented in Fig. 3e is generated. The different cells are depicted from left to right, and the voltage increases are depicted from top to bottom. This pseudo-HI map effectively visualizes cell inconsistency and voltage changes during the charging process, serving as a two-dimensional HI map.

Lastly, as listed in Table 1 and presented in Fig. 3f, a set of HI points is obtained from the vehicle operation data based on engineering intuition, experimental experience, and literature review. Since all energy required for these EVs' operation comes from the battery, there is a strong correlation between the vehicle's cumulative mileage and the battery's degradation state, as shown in Fig. 2h. Therefore, the vehicle's cumulative mileage is selected as the first HI point. In terms of HIs related to battery electrochemical reactions, the average temperature during the charging process is extracted as an HI point from the second HI sequence. Furthermore, various statistical metrics—mean, median, standard deviation, and range—are extracted from Q_{HI} , enriching the model's understanding of battery aging mechanisms. As batteries age, charge capacity curves within a fixed voltage range progressively shift downward, leading to corresponding changes in the mean, median, standard deviation, and range of Q_{HI} as the SOH deteriorates⁵¹. To further characterize cell inconsistencies during charging, statistical metrics such as the mean, median, standard deviation, and range of the cell voltage range sequence (dV) are extracted as Fig. 3d shows. Moreover, charging power is constrained by the maximum cell voltage to prevent overcharging of the battery. Also, charging power is constrained by temperature to avoid thermal runaway at high temperatures and lithium plating during high-rate charging at low temperatures. Since inconsistent charging behaviors can significantly impact the understanding of battery aging characteristics, it is necessary to consider HIs related to charging conditions. Hence, the mean and maximum charging currents, the minimum temperature, and the mean of the minimum temperature sequence during charging were adopted as HI points. In addition, EVs sometimes stop charging at various upper voltages, and a decreasing upper voltage limits the total energy that the battery can store since the battery may not be fully charged. To account for this variability, the charge-ending voltage is adopted as an HI point. These fifteen HI points can effectively reflect battery degradation.

Table 1 | HI points

Index	HI points	Illustration
1	Mileage	Accumulated driving mileage (km)
2	T_{ave}	The average temperature of the battery system during the charging process (°C)
3	Q_{ave}	Average value of Q_{HI} (Ah)
4	Q_{med}	Median value of Q_{HI} (Ah)
5	Q_{std}	Standard deviation of Q_{HI} (Ah)
6	Q_{ran}	Range of Q_{HI} (Ah)
7	dV_{ave}	Average value of dV (mV)
8	dV_{med}	Median value of dV (mV)
9	dV_{std}	Standard deviation of dV (mV)
10	dV_{ran}	Range of dV (mV)
11	C_{max}	Maximum charging current (A)
12	C_{ave}	Average charging current (A)
13	V_{end}	Maximum charge-ending voltage (V)
14	T_{min}	Minimum temperature of the battery system during charging process (°C)
15	T_{min_ave}	Average value of the minimum temperature during the charging process (°C)

Q_{HI} the charge capacity sequence, dV the cell voltage range sequence.

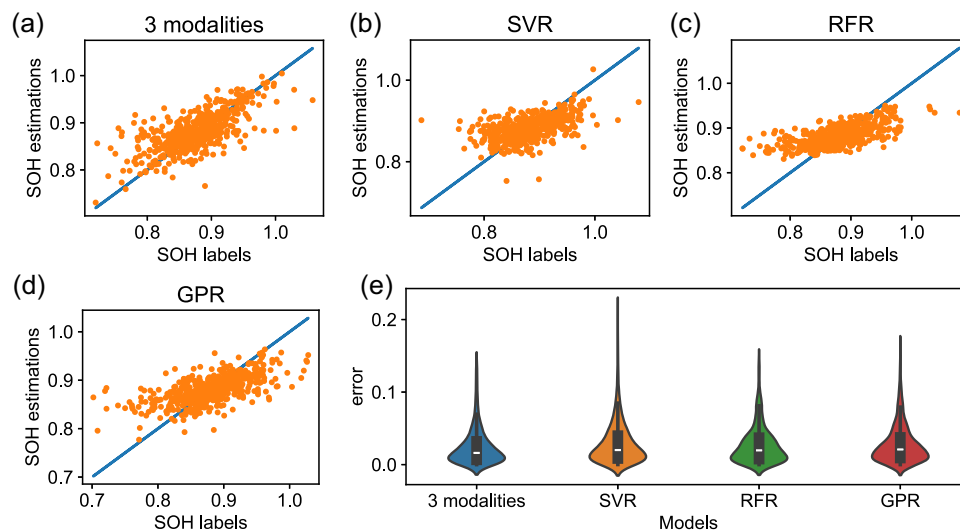


Fig. 4 | SOH estimation result compared with other machine learning (ML) methods. The SOH estimation results based on **a** the multi-modal framework, **b** the support vector machine regression (SVR), **c** the random forest regression (RFR),

and **d** the Gaussian process regression (GPR). **e** The error distribution of various methods. Source data are provided as a Source Data file.

Thus, for each SOH label, we match two HI sequences, one HI map, and fifteen HI points, consisting of a multi-modal HIs-label sample for model training. Typically, the nonlinear relationship between HI points and SOH can be characterized through machine learning methods such as support vector machine regression (SVR), random forest regression (RFR), GPR, and neural network. Although deep learning methods such as CNN and LSTM are suitable for extracting features from 1D and 2D inputs, existing studies do not effectively utilize multi-modal HIs with single-modal models, especially when addressing the SOH estimation problem in a complex scenario with multi-source information. Hence, we propose a multi-modal deep learning-based SOH estimation framework to enable the effective utilization of multi-modal HIs.

Multi-modal deep learning framework

The multi-modal framework, as illustrated in Fig. 1d, comprises a 2D HI map processing domain, a 1D HI sequence processing domain, and an HI point processing domain. In the 1D and 2D processing domains, deep residual convolutional neural networks (ResNets) are used for extracting high-dimensional features, and the HI points domain consists of several fully connected layers with nonlinear activation functions for performing nonlinear regression calculations of HI points. As a variant of CNNs, ResNet integrates shortcut methods to combine input data with post-convolution data, which effectively avoids the vanishing gradient problem and allows network depth increment, thus maximizing high-dimensional feature extraction. Finally, three domains are followed by fully connected layers for SOH regression estimation. The structure of the multi-modal framework is illustrated in Fig. 1d and Supplementary Note 4.

SOH estimation based on the multi-modal deep learning framework

To validate the performance of the proposed multi-modal framework, we randomly split the samples into the training and testing sets. Here, 80% of the samples are adopted as the training set for model tuning, and 20% of the samples are used as testing sets to examine estimation accuracy. In each epoch of the training process, 20% of the training samples are randomly selected as the validation sets to demonstrate the accuracy and improve model applicability. The estimation accuracy may be influenced by the randomness in test data partitioning, model initialization, and parameter iteration optimization. To ensure

the model's consistency and reliability, tests for different models are repeated ten times. To verify the advancement of the proposed framework, model performance is compared with three popular nonlinear regression machine learning (ML) methods that use HI points. The tests are conducted on a laptop with an Intel i7-12700H central processing unit. All neural network-based models are trained on an Nvidia GeForce RTX 3060 graphics processing unit with Python 3.8 as well as Tensorflow 2.10, and other ML methods are supported by Scikit-learn packages. The mean absolute percentage error (MAPE) and root mean squared error (RMSE) were calculated as follows to illustrate the estimation accuracy:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} * 100\% \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

As displayed in Fig. 4, the proposed framework achieves superior estimation accuracy than that of the other three existing methods, whose average estimation MAPE is 2.83%. The SVR model ranks fourth among the three methods, and its average MAPE is 3.34%. The RFR model is the second-best method with a MAPE of 3.10%. However, its estimation performance for lower SOH label samples is weak. The error distribution of the four methods is presented in Fig. 4e. Existing ML methods exhibit a disadvantage in lower SOH estimation, which is crucial for EV battery evaluation after a period of degradation. Therefore, we focused on the estimation result for SOH samples that were lower than 85%. The estimation MAPE for the proposed framework is 4.69%, which is 23.7% lower than the MAPE of the RFR model. The MAPE of the three ML models for low SOH samples are 6.07%, 6.15%, and 6.12%.

To investigate the influence of the network structure and modalities, we add six comparative models. The 2-modalities model removes the input of a 2D HI map and the corresponding domain. The 1-modality model removes the input of the 2D HI map and the 1D HI sequences simultaneously. The CNN model retains all three modalities of HIs as the input but replaces the ResNet structure with regular convolutional networks. In addition, we also replace the network structure of the 1-dimensional modal processing domain with RNN,

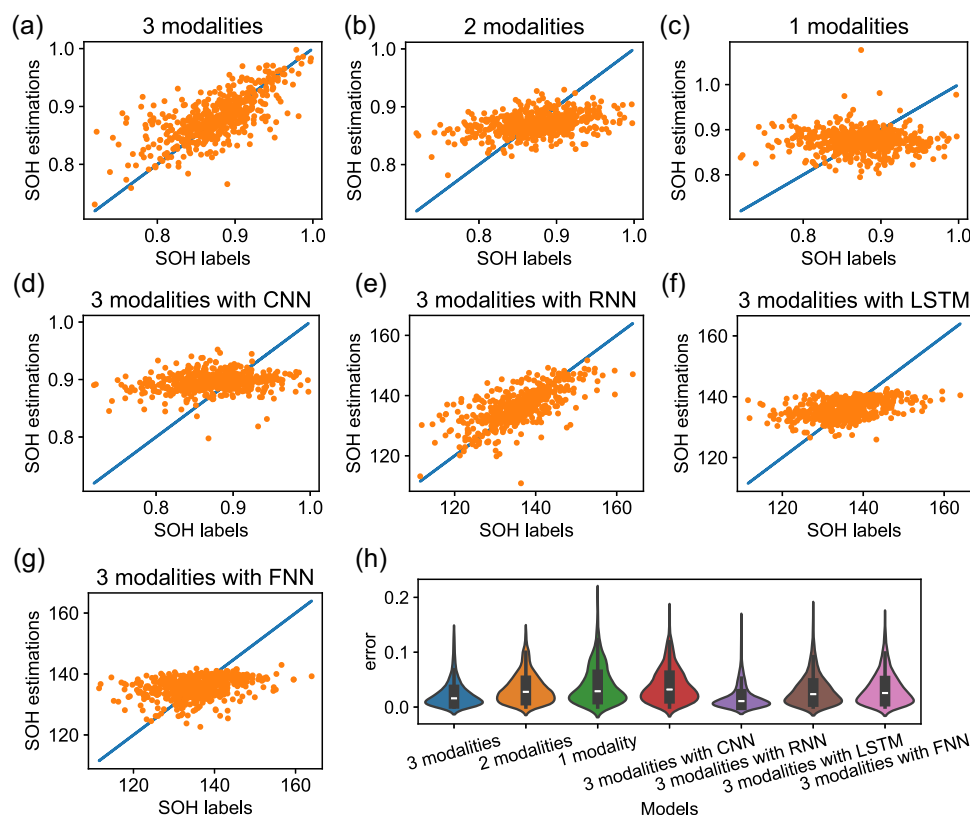


Fig. 5 | SOH estimation results compared with different deep learning methods. SOH estimation results based on **a** 3-modalities framework, **b** 2-modalities framework, **c** 1-modality framework, **d** 3-modalities framework with the convolutional neural network (CNN) structure, **e** 3-modalities framework with the recurrent

neural networks (RNN) structure, **f** 3-modalities framework with the long-short-term memory network (LSTM) structure, and **g** 3-modalities framework with the feedforward neural network (FNN) structure. **h** The error distribution of various methods. Source data are provided as a Source Data file.

LSTM⁵², and feedforward neural network⁵³ (FNN) to verify the impact of network structure on the model. To ensure repeatability and equity, the same learning rate strategy is set for the training processes of all seven models.

Figure 5 displays the testing set estimation results for seven models. The proposed multi-modal model achieves the best estimation accuracy. To evaluate the performance, the MAPE, RMSE, and the maximum estimation error of various models are given in Table 2. By using the 1D HI sequence in the 1-modality model, MAPE is reduced by 18.2%, and RMSE is reduced by 15.2%. Further adding the 2D HI map to the 2-modalities model results in a reduction of 25.7% and 21.3% in MAPE and RMSE, respectively. By contrast, replacing the CNN with ResNet reduces the MAPE and RMSE by 24.7% and 21.1%, respectively. Replacing the 2D processing domain with RNN, LSTM, and FNN may also lead to estimation accuracy decreases. A comparison of the estimation performance of various models is also presented in Table 2. The existing ML methods achieve superior results than that of the deep learning model when using the same inputs that only contain HI points. However, with the participation of multi-modal HIs, the proposed framework exhibits superior accuracy. We used the same framework to conduct SOH evaluation on another batch of vehicles. These vehicles are equipped with LiFePO₄ batteries, and the test results are presented in Supplementary Note 5. The results indicate that the proposed framework can be adapted to different electrochemical systems of batteries and can accurately estimate the SOH of various types of batteries.

Discussion

To develop the SOH evaluation method for real-world EV batteries, we investigate the historical data from 300 EVs monitored over three

years. Based on the field data, we analyze the characteristics of EV operation and summarize the bottlenecks in transferring the theoretical methods of SOH estimation to engineering applications as well as the factors influencing the accuracy of algorithms, these factors include the complex and uncontrollable vehicle operating conditions, low accuracy of the BMS state estimation, unstable charging profiles, and the data quality deficiencies. Furthermore, we investigate a multi-modal deep learning framework to accurately predict the SOH of batteries in EVs leveraging operational data. The approach involves the extraction of multi-modal HIs from a consistent voltage range observed during the charging process of the battery. To use these multi-modal inputs, we pioneer a multi-modal deep learning framework, enriched with 2.9 million trainable parameters, and fortified it by an underlying residual network architecture. By using a dynamic learning rate strategy, the framework achieves remarkably accurate SOH estimations for EV batteries. The MAPE of the SOH estimation results is 2.83%. This result illuminates the potential of the proposed framework for large-scale EV battery evaluation. The method promises not just remarkable reductions in time, economic outlay, and energy consumption but also advancements in intelligent BMS and cloud-based monitoring platforms for improved EV use. Finally, to catalyze advancements in battery management algorithms, we release a substantial collection of high-quality EV battery dataset to the public. This dataset encompasses comprehensive driving and charging logs for EVs to provide a solid dataset pivotal for analyzing vehicle usage patterns and forecasting battery degradation.

While the results of this study are promising, there are certain limitations that warrant further investigation. The proposed multi-modal framework generally provides reliable SOH estimation results. However, unlike the laboratory data, in real-world EVs, sudden changes

Table 2 | SOH estimation performance with various methods

Model	MAPE (%)		RMSE (%)		Max error (%)		Training time (s)
	Average	Std.	Average	Std.	Average	Std.	
3 modalities with ResNet	2.83	0.093	3.26	0.090	17.15	2.186	580
SVR model	3.34	0.114	3.74	0.145	16.557	2.000	0.5
RFR model	3.10	0.264	3.54	0.193	16.00	1.279	10
GPR model	3.11	0.129	3.47	0.139	16.33	1.791	28
2 modalities with ResNet	3.81	0.170	4.14	0.215	30.81	4.459	93
1 modality with ResNet	4.66	0.498	4.88	0.491	28.83	0.942	31
3 modalities with CNN	3.76	0.388	4.13	0.273	26.41	4.120	280
3 modalities with RNN	3.09	0.197	3.89	0.911	34.33	27.78	770
3 modalities with LSTM	3.29	0.395	4.04	0.972	30.34	24.36	977
3 modalities with FNN	3.21	0.204	3.80	0.536	27.51	19.98	900

Std standard deviation.

in battery conditions and data sampling errors can lead to incorrect feature calculations, resulting in abnormal feature inputs. Consequently, the model may generate erroneous results in some corner scenarios. In addition, the dominant aging mechanisms vary among different vehicles. For some, calendar aging is the primary factor, while for others, driving-related degradation is more significant. Furthermore, manufacturing inconsistencies due to mass production can cause inconsistencies among battery cells, leading to markedly different degradation patterns between some vehicles. To mitigate these issues, it is necessary to expand the training data to cover a wider range of scenarios.

Therefore, based on all the above, future research will focus on utilizing more advanced models, such as large language models, once sufficient real-world operational data becomes available, despite this will require time. Moreover, efforts will be made to enhance BMS by integrating more sensors into the battery system, such as individual temperature sensors for each cell and internal temperature and strain sensors inside the cells to improve the battery monitoring performance. Future research will also explore data-driven strategies that incorporate accurate physical information, thereby furnishing models with greater mechanistic understanding and boundary conditions. Due to the framework's ability to handle multi-modal data inputs, more physics-based features, such as information about loss of lithium inventory and loss of active material, can be potentially integrated into the multi-modal framework to improve the estimation performance. Lastly, it will be beneficial to utilize historical data and computational strengths by developing an efficient Cloud BMS, enabling precise monitoring of SOH degradation across extensive vehicle fleets.

Methods

EV data processing

The SOH labels of this study are obtained from the raw EV data with a working condition filtering strategy, SOC correction method, and ampere-hour integral method. First, the rest conditions are extracted from raw data, and determined by the timestamp gap between two coterminous data frames. If the time gap exceeds 1 h, the waking-up moment is set to t_s . After an hour of resting, the terminal voltage approximates the OCV because of the depolarization, and the battery SOC can be calibrated by the maximum cell voltage and the OCV–SOC look-up table. In this study, we use the cell with the maximum voltage to calibrate SOC_s because it typically represents the worst cell in the battery pack. If the calibrated SOC_s is higher than 0.6, this rest segment is skipped to ensure the candidate capacity calculation segment crosses a wide SOC interval to improve SOH calculation accuracy. We continue the search forward until the SOC reaches 1 and set the moment as t_e . The EVs stop charging when the maximum cell voltage

reaches 4.25 V to prevent the worst cell from being overcharged, with the BMS SOC jumping to 1 and the charging current cutting off. At this moment, the battery pack is fully charged and stores the maximum energy. Thus, we obtain the capacity calculation segment (**Seg_{ca}**). Using the ampere-hour integral, the capacity and the SOH label of one segment can be calculated and recorded with **Seg_{ca}**. Detailed information about the SOH label calculation can be found in Supplementary Note 3.

Multi-modal health indicators extraction

Compared to vehicle driving conditions, the charging conditions of EVs are relatively stable. As mentioned, data from the EV's charging segments are utilized when calculating SOH labels. These charging behaviors start at random SOC levels and end with the battery being fully charged. Based on a series of data cleaning and feature extraction methods, HIs are extracted corresponding to the SOH labels.

First, a target voltage sequence **V_{HI}** with a fixed voltage interval for HI extraction is set,

$$V_{HI} = [3900, 3901, \dots, 4050], \quad (3)$$

whose corresponding SOC scope is 71–84%. For a piece of charging process raw data extracted from **Seg_{ca}**, an HI extraction segment can be obtained as follows:

$$Data_{HI} = [Data_{t_1}, Data_{t_2}, \dots, Data_{t_n}], \quad (4)$$

where t_1 and t_n are the moments in which the maximum cell voltage reaches 3900 mV and 4050 mV, respectively. Here, $Data_{t_i}$ is the data frame sampled in t_i . For various charging processes, the length of **Data_{HI}** is different because of diverse charging conditions, different charge starting SOC, various battery health states, and other factors. Subsequently, the HI sequences and planes should be standardized. The EV battery pack adopted in this study contains 96 cells, and their voltages are inconsistent during the charging process. To capture inconsistency, a 96×96 cell voltage feature matrix **V_c** is extracted from **Data_{HI}** through uniform sampling as follows:

$$V_c = \begin{bmatrix} V_{1,1} & V_{1,2} & \dots & V_{1,96} \\ V_{2,1} & \dots & \dots & \dots \\ V_{3,1} & \dots & V_{i,j} & \dots \\ V_{4,1} & \dots & \dots & V_{96,96} \end{bmatrix}, \quad (5)$$

where $V_{i,j}$ represents the i^{th} sampling voltage of the j^{th} cell. Thus, the voltage inconsistency of cells in various charging processes is standardized with the same shape, and **V_c** is adopted as a 2D HI map.

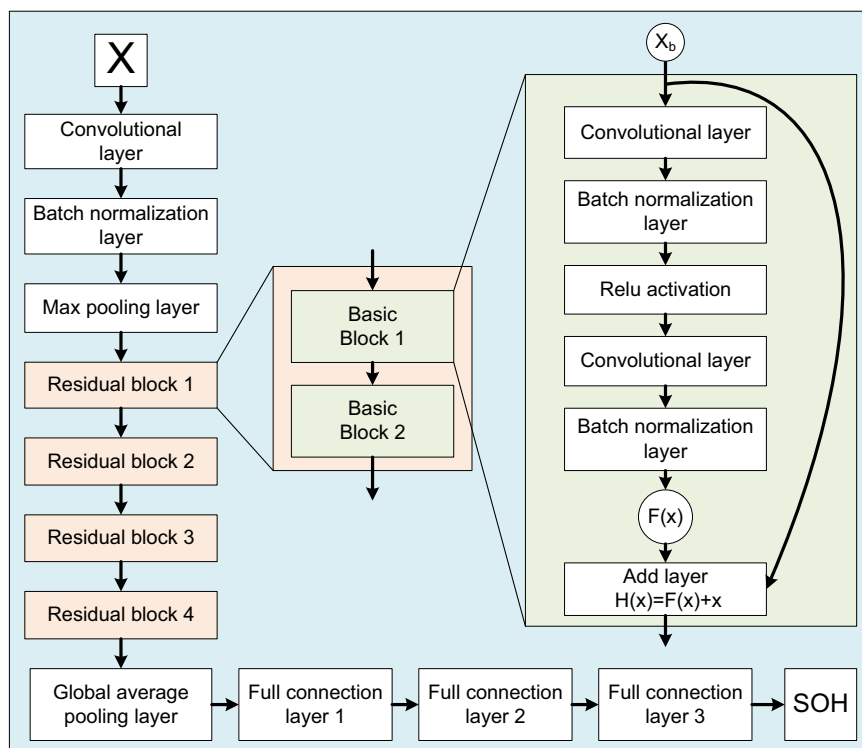


Fig. 6 | ResNet structure diagram for two-dimensional HI map and one-dimensional HI sequence processing domain. The processing domain consists of an input Convolutional layer, a Batch normalization layer, a Max pooling layer, and four Residual blocks. Each Residual block is composed of two Basic Blocks that follow the convolutional residual network architecture. Following this, a global average pooling layer and three fully connected layers are employed to generate the predicted State of Health (SOH).

Second, the maximum cell voltage sequence $V_{\max} = [V_{\max t1}, V_{\max t2}, \dots, V_{\max tn}]$ is extracted from the **Data_{HI}**, and the charge capacity sequence $Q_t = [Q_{t1}, Q_{t2}, \dots, Q_{tn}]$ of each frame is calculated as follows:

$$Q_{ti} = \int_{t1}^{ti} I dt, \quad (6)$$

where **I** refers to the charging current. Because of the variable-length **Data_{HI}**, Q_t should be standardized through interpolation according to **V_{HI}** as follows:

$$\begin{bmatrix} Q_t \\ V_{\max} \end{bmatrix} = \begin{bmatrix} Q_{t1}, Q_{t2}, \dots, Q_{tn} \\ V_{\max t1}, V_{\max t2}, \dots, V_{\max tn} \end{bmatrix} \Rightarrow \begin{bmatrix} Q_{HI} \\ V_{HI} \end{bmatrix} = \begin{bmatrix} Q_1, Q_2, \dots, Q_{151} \\ 3900, 3901, \dots, 4050 \end{bmatrix}. \quad (7)$$

The temperature variation during the charging process is considered. The average temperature sequence is extracted from **Data_{HI}** and is compressed $T_{HI} = [T_1, T_2, \dots, T_{151}]$ by uniform sampling. Thus, **Q_{HI}** and **T_{HI}** are obtained as two 1D HI sequences.

Finally, fifteen HI points are derived from the operating data and the other two modal HIs, considering the vehicle aging conditions, cell inconsistency, and charging behaviors. A detailed illustration is presented in Table 1. With the HIs, a SOH training-testing sample can be defined as follows:

$$\text{Sample} = [\text{output}] | [\text{input}] = [\text{SOH}] \begin{bmatrix} HI_{map} \\ HI_{sequences} \\ HI_{points} \end{bmatrix} = [\text{SOH}] \begin{bmatrix} V_c \\ Q_{HI}, T_{HI} \\ HI_1, HI_2, \dots, HI_{15} \end{bmatrix} \quad (8)$$

Residual network architecture

The multi-modal framework comprises a 2D HI map processing domain, a 1D HI sequence processing domain, and an HI point processing domain. As illustrated in Fig. 6, the 2D HI map processing domain adopts the residual network architecture⁵⁴, which consists of an input convolutional layer, a batch normalization layer, a max pooling layer, and four residual blocks. Here, a global average pooling layer and three full connection layers are adopted to obtain the output of the 2D HI map processing domain. Compared with the regular CNN network, the ResNet mitigates the vanishing gradient problem by adding a shortcut connection that bypasses one or more layers. These shortcut connections enable the gradient to flow easily through the network during training and allow the neural network to go deeper without accuracy degradation.

Each residual block involves two basic blocks, in which the output can be defined as follows:

$$H(x) = F(x) + x, \quad (9)$$

where $H(x)$ is the output of the basic block, and $F(x)$ is the characteristic mapping for input x extracted by several hidden layers. The basic block consists of convolutional layers, batch normalization layers, and rectified linear unit (ReLU) activation functions. For a 2D basic block, given an input $\text{Con}_{in} \in \mathbf{R}^{M \times N}$ and a convolutional filter $\mathbf{w} \in \mathbf{R}^{U \times V}$, the in-between layers can be formulated as follows:

$$\text{Con}_{out}(i, j) = \sum_{u=1}^U \sum_{v=1}^V [w(u, v) \cdot \text{Con}_{in}(i+u-1, j+v-1)] + b, \quad (10)$$

where $\text{Con}_{out}(i, j)$ is the output element in the location (i, j) of the output matrix, and b is the bias of the convolutional layer. The batch

normalization layer can be formulated as follows:

$$BN_{out} = \left(\frac{BN_{in} - \mu}{\sqrt{\sigma^2 + \varepsilon}} \right) * \gamma + \beta, \quad (11)$$

where BN_{out} is the output, BN_{in} is the input, μ and σ are the average value of BN_{in} , ε is a small number for numerical stability, and γ and β are two network trainable parameters. ReLU activation was used to promise the sparsity for the network, mitigate the gradient disappearance, and boost network training. ReLU is one of the most used activation functions and is expressed as follows:

$$y = \max(0, x). \quad (12)$$

Based on a substantial number of comparative trials, the optimal hyperparameters for the model are determined, which is illustrated in Supplementary Note 4.

During training, the warm-up technique is used to allow the model to investigate a wide parameter space in the early stages and achieve global optimization. The learning rate starts at a low initial value and gradually increases to a target learning rate in the first ten epochs (the warm-up period). After the warm-up period, an exponential decline learning rate schedule is adopted to search for optimal parameters. Detailed information about the warm-up method can be found in Supplementary Note 6. After training the model with SOH labels and multi-modal HIs, the structure and parameters of the multi-modal network can be saved. In real-world applications, SOH estimation can be conducted through the following three steps. First, capture the charging behaviors from the historical operational data of EVs and perform data preprocessing, such as outlier handling and removal of low-quality data segments. Second, multi-modal HIs that effectively reflect the battery's aging state can be extracted as long as the charging process completely spans the voltage range [3900 mV, 4050 mV]. Lastly, use the pre-trained model to estimate SOH based on these HIs as inputs.

Data availability

The EV field data are available at the following link: <http://ivstskl.changan.com.cn/?p=2697>. The source data generated in this study have been deposited at <https://github.com/HoraceLiu1010/Multi-modal-SOH-estimation-framework>. Source data are provided with this paper.

Code availability

Code for this work is available at the following link: <https://github.com/HoraceLiu1010/Multi-modal-SOH-estimation-framework>.

References

- Cano, Z. P. et al. Batteries and fuel cells for emerging electric vehicle markets. *Nat. Energy* **3**, 279–289 (2018).
- Lukic, S. M., Cao, J., Bansal, R. C., Rodriguez, F. & Emadi, A. Energy storage systems for automotive applications. *IEEE Trans. Ind. Electron.* **55**, 2258–2267 (2008).
- Hu, X., Xu, L., Lin, X. & Pecht, M. Battery lifetime prognostics. *Joule* **4**, 310–346 (2020).
- Birkel, C. R., Roberts, M. R., McTurk, E., Bruce, P. G. & Howey, D. A. Degradation diagnostics for lithium ion cells. *J. Power Sources* **341**, 373–386 (2017).
- Song, Z. et al. A study of cell-to-cell variation of capacity in parallel-connected lithium-ion battery cells. *ETransportation* **7**, 100091 (2021).
- Han, X. et al. A review on the key issues of the lithium ion battery degradation among the whole life cycle. *ETransportation* **1**, 100005 (2019).
- Onori, S., Spagnol, P., Marano, V., Guezennec, Y. & Rizzoni, G. A new life estimation method for lithium-ion batteries in plug-in hybrid electric vehicles applications. *Int. J. Power Electron.* **4**, 302–319 (2012).
- Moura, S. J., Callaway, D. S., Fathy, H. K. & Stein, J. L. Tradeoffs between battery energy capacity and stochastic optimal power management in plug-in hybrid electric vehicles. *J. Power Sources* **195**, 2979–2988 (2010).
- Pecht, M. CALCE battery data. <https://calce.umd.edu/battery-data> (2015).
- Howey, D. Oxford battery dataset. <https://howey.eng.ox.ac.uk/data-and-code/> (2022).
- Zhang, Y. et al. Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning. *Nat. Commun.* **11**, 1706 (2020).
- Severson, K. A. et al. Data-driven prediction of battery cycle life before capacity degradation. *Nat. Energy* **4**, 383–391 (2019).
- Li, W. et al. One-shot battery degradation trajectory prediction with deep learning. *J. Power Sources* **506**, 230024 (2021).
- Lu, J., Xiong, R., Tian, J., Wang, C. & Sun, F. Deep learning to estimate lithium-ion battery state of health without additional degradation experiments. *Nat. Commun.* **14**, 2760 (2023).
- Weng, A. et al. Predicting the impact of formation protocols on battery lifetime immediately after manufacturing. *Joule* **5**, 2971–2992 (2021).
- Battery Archive. <https://www.batteryarchive.org/index.html> (2024).
- Pozzato, G. et al. Analysis and key findings from real-world electric vehicle field data. *Joule* **7**, 2035–2053 (2023).
- Zhang, J. et al. Realistic fault detection of li-ion battery via dynamical deep learning. *Nat. Commun.* **14**, 5940 (2023).
- Deng, Z. et al. Prognostics of battery capacity based on charging data and data-driven methods for on-road vehicles. *Appl. Energy* **339**, 120954 (2023).
- Waag, W., Fleischer, C. & Sauer, D. U. Critical review of the methods for monitoring of lithium-ion batteries in electric and hybrid vehicles. *J. Power Sources* **258**, 321–339 (2014).
- Plett, G. L. Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 3. State and parameter estimation. *J. Power Sources* **134**, 277–292 (2004).
- Sadabadi, K. K., Jin, X. & Rizzoni, G. Prediction of remaining useful life for a composite electrode lithium ion battery cell using an electrochemical model to estimate the state of health. *J. Power Sources* **481**, 228861 (2021).
- Crawford, A. J., Choi, D., Balducci, P. J., Subramanian, V. R. & Viswanathan, V. V. Lithium-ion battery physics and statistics-based state of health model. *J. Power Sources* **501**, 230032 (2021).
- Che, Y., Hu, X., Lin, X., Guo, J. & Teodorescu, R. Health prognostics for lithium-ion batteries: mechanisms, methods, and prospects. *Energy Environ. Sci.* **16**, 338–371 (2023).
- Kohtz, S., Xu, Y., Zheng, Z. & Wang, P. Physics-informed machine learning model for battery state of health prognostics using partial charging segments. *Mech. Syst. Signal Process.* **172**, 109002 (2022).
- Li, X., Yu, D., Vilsen, S. B. & Store, D. I. The development of machine learning-based remaining useful life prediction for lithium-ion batteries. *J. Energy Chem.* **82**, 103–121 (2023).
- Jones, P. K., Stimming, U. & Lee, A. A. Impedance-based forecasting of lithium-ion battery performance amid uneven usage. *Nat. Commun.* **13**, 4806 (2022).
- Zhao, J. et al. Machine learning for predicting battery capacity for electric vehicles. *eTransportation* **15**, 100214 (2023).
- Wang, Z., Yuan, C. & Li, X. Lithium battery state-of-health estimation via differential thermal voltammetry with Gaussian process regression. *IEEE Trans. Transp. Electr.* **7**, 16–25 (2020).
- Ng, M.-F., Zhao, J., Yan, Q., Conduit, G. J. & Seh, Z. W. Predicting the state of charge and health of batteries using data-driven machine learning. *Nat. Mach. Intell.* **2**, 161–170 (2020).

31. Zhu, J. et al. Data-driven capacity estimation of commercial lithium-ion batteries from voltage relaxation. *Nat. Commun.* **13**, 2261 (2022).
32. Nozarijoubi, Z. & Fathy, H. K. Machine learning for battery systems applications: progress, challenges, and opportunities. *J. Power Sources* **601**, 234272 (2024).
33. Huang, H. et al. An enhanced data-driven model for lithium-ion battery state-of-health estimation with optimized features and prior knowledge. *Automot. Innov.* **5**, 134–145 (2022).
34. Che, Y. et al. State of health prognostics for series battery packs: a universal deep learning method. *Energy* **238**, 121857 (2022).
35. Lee, G., Kwon, D. & Lee, C. A convolutional neural network model for SOH estimation of Li-ion batteries with physical interpretability. *Mech. Syst. Signal Process.* **188**, 110004 (2023).
36. Chaoui, H. & Ibe-Ekeocha, C. C. State of charge and state of health estimation for lithium batteries using recurrent neural networks. *IEEE T Veh. Technol.* **66**, 8773–8783 (2017).
37. Wang, Q. et al. Large-scale field data-based battery aging prediction driven by statistical features and machine learning. *Cell Rep. Phys. Sci.* **4**, 101720 (2023).
38. Chemali, E. & Preindl, M. Neural-network state-of-charge and state of health estimation. In: Google patents, US patent app. 16/688,260 (2020).
39. Jiang, N. et al. Driving behavior-guided battery health monitoring for electric vehicles using extreme learning machine. *Appl. Energy* **364**, 123122 (2024).
40. Saha, B. Battery data set. NASA AMES Prognostics Data Repository. <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/> (2007).
41. Wang, F., Zhai, Z., Zhao, Z., Di, Y. & Chen, X. Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis. *Nat. Commun.* **15**, 4332 (2024).
42. Qi, Q. et al. Battery pack capacity estimation for electric vehicles based on enhanced machine learning and field data. *J. Energy Chem.* **92**, 605–618 (2024).
43. Vidal, C., Malysz, P., Naguib, M., Emadi, A. & Kollmeyer, P. J. Estimating battery state of charge using recurrent and non-recurrent neural networks. *J. Energy Storage* **47**, 103660 (2022).
44. Vidal, C. et al. Robust xev battery state-of-charge estimator design using a feedforward deep neural network. *SAE Int. J. Adv. Curr. Pract. Mobil.* **2**, 2872–2880 (2020).
45. Roman, D., Saxena, S., Robu, V., Pecht, M. & Flynn, D. Machine learning pipeline for battery state-of-health estimation. *Nat. Mach. Intell.* **3**, 447–456 (2021).
46. Mohtat, P., Lee, S., Siegel, J. B. & Stefanopoulou, A. G. Comparison of expansion and voltage differential indicators for battery capacity fade. *J. Power Sources* **518**, 230714 (2022).
47. Naha, A. et al. An incremental voltage difference based technique for online state of health estimation of Li-ion batteries. *Sci. Rep.* **10**, 9526 (2020).
48. Anseán, D. et al. Fast charging technique for high power LiFePO4 batteries: a mechanistic analysis of aging. *J. Power Sources* **321**, 201–209 (2016).
49. Jaguemont, J., Boulon, L. & Dubé, Y. A comprehensive review of lithium-ion batteries used in hybrid and electric vehicles at cold temperatures. *Appl. Energy* **164**, 99–114 (2016).
50. Tian, J. et al. A critical review on inconsistency mechanism, evaluation methods and improvement measures for lithium-ion battery energy storage systems. *Renew. Sustain. Energy Rev.* **189**, 113978 (2024).
51. Deng, Z., Hu, X., Li, P., Lin, X. & Bian, X. Data-driven battery state of health estimation based on random partial charging data. *IEEE Trans. Power Electron.* **37**, 5021–5031 (2021).
52. Chemali, E., Kollmeyer, P. J., Preindl, M., Ahmed, R. & Emadi, A. Long short-term memory networks for accurate state-of-charge estimation of Li-ion batteries. *IEEE Trans. Ind. Electron.* **65**, 6730–6739 (2017).
53. Vidal, C., Malysz, P., Kollmeyer, P. & Emadi, A. Machine learning applied to electrified vehicle battery state of charge and state of health estimation: state-of-the-art. *IEEE Access* **8**, 52796–52814 (2020).
54. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition 770–778* (IEEE, 2016).

Acknowledgements

This work was supported by the National Key Research and Development Program of China under grant number 2022YFB3305403 (X.H.), the project of basic research funds for central universities under grant number 2023CDJQCZX-001 (X.H.) and 2022CDJDX-006 (X.H.), the National Natural Science Foundation of China under grant number U23A20327 (X.H.), the National Natural Science Foundation of China under grant number 72361137006 (X.H.), and the National Natural Science Foundation of China under Grant 52402465 (K.Z.). We also acknowledge support from the Chang'an Automobile Co., Ltd.

Author contributions

H.L. and C.L. conceived the research. H.L., J.L., and K.Z. developed the algorithms. Y.X. and R.W. performed data collection and management. All authors edited and reviewed the manuscript. Z.S. and X.H. supervised the work.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-56485-7>.

Correspondence and requests for materials should be addressed to Xiaosong Hu or Ziyu Song.

Peer review information *Nature Communications* thanks Enhui Liu, Carlos Vidal and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025