



## Survey paper

## An overview of data-driven battery health estimation technology for battery management system

Minzhi Chen <sup>a,1</sup>, Guijun Ma <sup>b,1</sup>, Weibo Liu <sup>c,1</sup>, Nanyin Zeng <sup>d,\*</sup>, Xin Luo <sup>e,\*</sup><sup>a</sup> School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China<sup>b</sup> School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China<sup>c</sup> Department of Computer Science, Brunel University London, Uxbridge, Middlesex UB8 3PH, United Kingdom<sup>d</sup> Department of Instrumental and Electrical Engineering, Xiamen University, Xiamen 361005, China<sup>e</sup> College of Computer and Information Science, Southwest University, Chongqing 400715, China

## ARTICLE INFO

## Article history:

Received 21 November 2022

Revised 20 January 2023

Accepted 13 February 2023

Available online 17 February 2023

Communicated by Zidong Wang

## Keywords:

Battery state of health

Battery management system

Data driven

Data science

Overview

Review

## ABSTRACT

Battery degradation, caused by multiple coupled degradation mechanisms, severely affects the safety and sustainability of a battery management system (BMS). The battery state of health (SOH) is a commonly-adopted metric to evaluate a battery's degradation condition, which should be carefully modeled to facilitate the safety and reliability of a BMS. Recently, owing to the rapid progress of data science-related techniques, data-driven models for battery SOH estimation have attracted great attentions from both academia and industry communities. This paper aims to provide the scientists and engineers with a general overview of data-driven battery SOH estimation technology for BMSs. State-of-the-art models published during 2018–2022 are reviewed with care, including a) feature extraction and selection methods; b) benchmarks, variants and extensions of data-driven SOH estimation models; and c) publicly-available battery SOH datasets. Afterwards, experiments are conducted and analyzed on the Toyota & Stanford-MIT battery SOH datasets for benchmark study. Finally, existing challenges and feature trends are summarized.

© 2023 Elsevier B.V. All rights reserved.

## 1. Introduction

To date, the global energy system is changing rapidly due to the fossil fuel shortage and global warming caused by massive carbon emissions [1–3]. With the exploration of renewable energy such as solar, hydro and wind energy, the energy usage of fossil fuel is on the decline [4,5]. With the characteristics of high power density, high energy efficiency and also the relatively long cycle life, energy-storing batteries have been widely deployed in battery management system (BMS) [6,7]. However, the battery performance is subject to degradation over time due to their long-term operations (which heavily affects the safety and sustainability of BMS), thereby becoming a major defect for real applications [8].

Battery degradation is a main factor affecting battery's performance, which is the consequence of multiple coupled degradation mechanisms [9–11]. The degradation factors (including battery

chemistry, manufacturing, environmental and operating conditions) are cross-dependent [12,13]. Hence, the degradation mechanisms of batteries are highly complicated to characterize in practice.

The battery state of health (SOH) is a commonly-accepted metric to evaluate its degradation level [14–16]. A battery's SOH reflects its current capability to store energy and supply power in contrast to its state at the beginning of its lifecycle [17,18]. By accurately monitoring the operating conditions of batteries, the SOH can realize the early warning and control optimization in a BMS [18–20]. However, the degradation in a battery exists across its entire life and is influenced by many factors e.g., temperature and charge–discharge rates. To investigate the battery degradation rules under these factors calls for the time-consuming cycling tests and characteristic tests [18,21], which results in the economically expensive. Hence, SOH estimation becomes a challenging issue [22,23].

Generally, battery systems are complex. It should be pointed out that the degradation process can be characterized by the capacity fade and power fade (i.e., resistance increase) at a macroscopic scale [18,24]. Conventional SOH estimation models mostly adopt the capacity tests, impedance measurements and other tests

\* Corresponding authors.

E-mail addresses: [chenminzhi001@outlook.com](mailto:chenminzhi001@outlook.com) (M. Chen), [mgi@hust.edu.cn](mailto:mgi@hust.edu.cn) (G. Ma), [Weibo.Liu2@brunel.ac.uk](mailto:Weibo.Liu2@brunel.ac.uk) (W. Liu), [zny@xmu.edu.cn](mailto:zny@xmu.edu.cn) (N. Zeng), [luoxin@swu.edu.cn](mailto:luoxin@swu.edu.cn) (X. Luo).<sup>1</sup> Co-First Authors of this study.

to check the battery health indicators straightforwardly [21,23,25]. For instance, the capacity can be validated by testing the charge transferred through the battery during the charging or discharging process, and the resistance can be obtained by calculating the instantaneous voltage drop during the pulse test [18]. Some well-known measurement-based approaches include the ampere-hour counting method [14], the resistance/impedance method [21], and the electrochemical impedance spectroscopy (EIS) method [17]. Such a method is easy to deploy, however, its application in actual working environment is not reliable enough due to various accidental factors emerging in real world applications [23].

On the other hand, the model-based approaches have been developed over recent years [15,17,21,23,26]. Based on the expert knowledge in mathematics and degradation physics, the model-based approaches (such as Kalman filter (KF) [14,21,27,28] and particle filter (PF) [29–31]) have been designed for battery SOH estimation. Model-based approaches can observe the internal state variables through an iterative mechanism and explain the degradation mechanism between two adjacent cycles [24]. The electrochemical models or equivalent-circuit models could emulate the battery internal dynamic system,

Data-driven models for SOH estimation and remaining useful life (RUL) prediction have been attracting increasing attention in both academia and industry due to the advantage of circumventing complex physical models [24,25,32–36]. It requires a large battery degradation dataset collected by degradation tests incorporating multiple factors affecting battery SOH. The effectiveness of data-driven models also relies on the quality of degradation data. In general, data analysis (DA)-based models [25,34,37] and machine learning (ML)-based models [24,32,35,38–40] are popular data-driven methods for SOH estimation. DA-based models identify features from the differential curves of measured data (such as electrical, thermal or mechanical signals during battery cycling) by fitting analytical functions to the measured data [37]. To do so, the correlations between the battery SOH and the battery's electrical, thermal and mechanical behaviors are constructed. The most frequently used DA-based models include differential voltage (DV)/incremental capacity (IC) analysis, differential mechanical parameter analysis and differential thermal voltammetry (DTV) analysis [34,37].

On the other hand, ML-based models are widely used in data-driven SOH estimation and RUL prediction due to their flexibility

1D-CNN	One-dimensional CNN	GPR	Gaussian Process Regression
2D-CNN	Two-dimensional CNN	GRA	Grey Relational Analysis
AR	Autoregression	GRUN	Gated Recurrent Unit Network
ARD	Automatic Relevance Determination	GWO	Grey Wolf Optimizer
ARIMA	Autoregression Integrated Moving Average	Hybrid-NNs	Hybrid Neural Networks
BMS	Battery Management System	IC	Incremental Capacity
BP	Back-Propagation	KF	Kalman Filter
CC	Constant Current	LightGBM	Light Gradient Boosting Machine
CM	Conventional Models	LR	Linear Regression
CNN	Convolutional Neural Network	LSTM	Long Short-Term Memory
CV	Constant Voltage	ML	Machine Learning
DA	Data Analysis	MMD	Maximum Mean Discrepancy
DBN	Dynamic Bayesian Network	NN	Neural Network
DDA	Deep Domain Adaptation	PCC	Pearson Correlation Coefficient
DKL	Deep Kernel Learning	PF	Particle Filter
DL	Deep Learning	PHM	Prognostics and Health Management
DNN	Deep Neural Network	PSO	Particle Swarm Optimization
DT	Differential Temperature	RBFN	Radial Basis Function Network
DTV	Differential Thermal Voltammetry	RF	Random Forest
DV	Differential Voltage	RNN	Recurrent Neural Network
EIS	Electrochemical Impedance Spectroscopy	RUL	Remaining Useful Life
EL	Ensemble Learning	RVM	Relevance Vector Machine
ELM	Extreme Learning Machine	SOC	State of Charge
Elman-NN	Elman Neural Network	SOH	State of Health
EMD	Empirical Model Decomposition	SVM	Support Vector Machine
FFNN	Feed-Forward Neural Network	SVR	Support Vector Regression
GA	Genetic Algorithm	TCA	Transfer Component Analysis
GAN	Generative Adversarial Network	TL	Transfer Learning
GBT	Gradient Boosting Tree	VAE	Variational Autoencoder
GMM	Gaussian Mixture Model		

which makes it possible to trace the battery degradation process [15,17]. Furthermore, sufficient domain knowledge of the degradation processes leading to failure is highly required to achieve the precise SOH estimation. Nevertheless, due to the battery's complex electrochemical dynamics, the degradation mechanism is hard to grasp [24].

and non-nearly matching ability [24,32]. ML model is a specific data analytical method that learns from data, recognize patterns and make decisions and estimation with minimal human intervention [41–46]. To date, ML-based models are becoming the most popular one for data-driven SOH estimation and RUL prediction, like the neural network (NN)-based models [48,49], Gaussian pro-

cess regression (GPR)-based models [50–53], and random forest (RF)-based models [54,55]. Specifically, ML-based models intend to extract the features affecting battery degradation from the directly measured data such as terminal voltage ( $V$ ), current ( $I$ ), temperature ( $T$ ), and recorded time ( $t$ ). ML-based models learn the battery degradation behaviors during the training process and are thus used to estimate battery SOH.

Nowadays, there are numerous review papers on battery SOH estimations and RUL predictions. Some comprehensive review papers on battery modeling and state estimation approaches have been published in [10,12,56]. In [10,12], a range of estimation approaches of state of charge (SOC), SOH, RUL, state of energy, and state of power have been reviewed and summarized. Recent reviews focusing on the battery SOH estimation and RUL prediction have been presented in [14,15,17–19,21–23], which divides the models of battery SOH estimation and RUL prediction into different categories according to their principles, mathematics and application. Since these reviews discuss broad categories of SOH estimation and/or RUL prediction models, the analysis of data-driven SOH estimation for BMS is not in-deep enough. Here, this paper provides an overview of data-driven battery SOH estimation models applied to BMS, as shown in Fig. 1 and discusses their status, advantages, drawbacks and challenges.

The mainly contributions of this review include:

- 1) **Summary of feature extraction and selection methods.** The feature extraction and selection from the raw data are the foundation for data-driven SOH estimation models. Section III presents the commonly-used feature extraction methods and feature selection methods, including their basic conception and application;
- 2) **Comprehensive survey of data-driven SOH estimation models.** Following the advanced and real-time issues in the filed of battery SOH estimation, the state-of-the-art SOH estimation models mostly published from 2018 to 2022 are reviewed. Section IV provides a detailed introduction on the basic principle, important formulas, and plots of the selected SOH estimation models. Furthermore, the variants and extensions of selected SOH estimation models are introduced, and the benefits and drawbacks of these models are presented in Table 1. The recently developed models from 2020 to 2022 are summarized in Table 2, and relevant applications and novelties are highlighted, which can provide researchers with the most advanced and significant comparison models;
- 3) **Battery datasets and model evaluation metrics.** Battery datasets acquisition is critical for battery degradation research. Section V summarizes the available public battery datasets used in reviewed papers, including the battery number, battery materials, and testing conditions. The resources availability is appended to the Section V. Furthermore, the evaluation metrics of SOH estimation model are represented in the Section V.

- 4) **Benchmark study.** The performances of several state-of-the-art data-driven SOH estimation models are compared through their suitable and representative datasets. The discussion follows based on the comparison results;
- 5) **Challenges and future research.** The existing issues about data-driven SOH estimation model applied in BMS are summarized, and corresponding challenges are discussed. Focusing on the feasible and cost-effective of research directions towards data-driven SOH estimation for BMS, we propose future research proposals based on trends in current research work.

Section II gives the definition of SOH. Section III summarizes the feature extraction and selection methods for an SOH estimation model. Section IV reviews state-of-the-art data-driven SOH estimation models. Section V provides the publicly available battery SOH datasets and evaluation metrics. Section VI reports the benchmark study. Finally, section VII concludes further challenges and future trends regarding data-driven SOH estimation models.

Definition of SOH.

Battery SOH reflects a battery's current degradation level in comparison with its starting point of the lifecycle [21]. SOH quantitatively assesses a battery's health state with one specific performance like capacity [15,21] or internal resistance [57,58]. In an energy storage system where the available energy takes a fundamental role, the capacity is commonly-adopted to describe the battery SOH. In a power-focused application like a hybrid electric vehicle, the internal resistance is often adopted as the SOH metric since an increase in internal resistance can limit the power capacity of systems and decrease the efficiency of the electric vehicle [17]. Generally, the battery is considered at the end of its lifecycle and needs to be replaced when the capacity drops below than 80% of the initial value or the internal resistance doubles [21].

Mathematically, SOH is calculated as the ratio of the current battery capacity or resistance to its initial value [24]. Considering the battery capacity, the SOH can be expressed by [59]:

$$SOH = \frac{C_{current}}{C_{initial}} \times 100\%, \quad (1)$$

where  $C_{current}$  and  $C_{initial}$  denote the current and initial capacity of a battery, respectively. Considering the internal resistance as the metric, the SOH is given by [60]:

$$SOH = \frac{R_{current} - R_{initial}}{R_{initial}} \times 100\%, \quad (2)$$

where  $R_{current}$  and  $R_{initial}$  are the current and initial internal resistance of a battery, respectively.

Note that the internal resistance can be measured by the hybrid pulse power characterization and EIS methods [23,24]. Such a method is highly sensitive to the testing conditions where essential differences can be found in laboratory from real environments [17,18]. Moreover, it is inevitable to terminate power supply when conducting the internal resistance tests, which could lead to

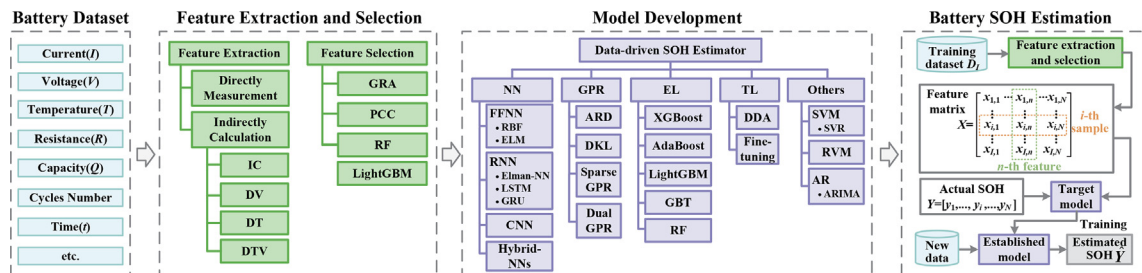
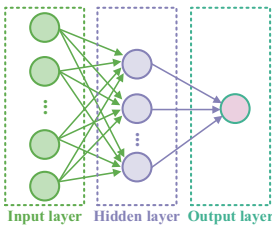
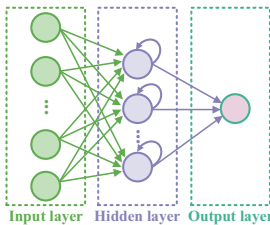
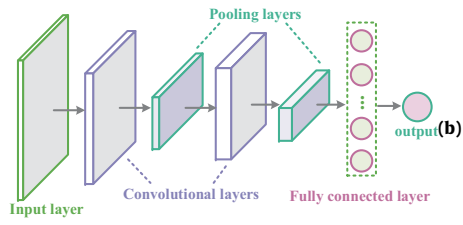
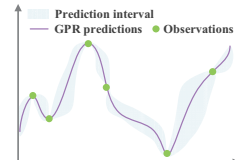
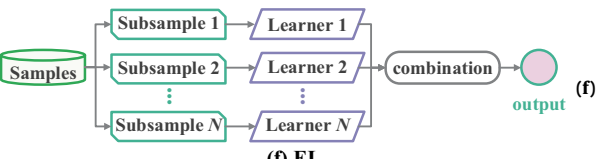


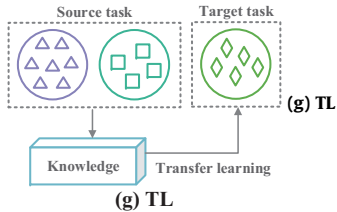
Fig. 1. The workflow of data-driven battery SOH estimation models.

**Table 1**  
Summary of Data-driven SOH Estimation Models.

Models:	Overviews	Ref.		
NN: FFNN#RBFN #ELMRNN#Elman- NN#LSTM #GRUN CNNHybrid-NNs	 <p>(a) <b>FFNN Description:</b> FFNN performs the linear calculation feed to the hidden layer. After the internal training computations, the output results are the estimated SOHs.</p> <p>(a) <b>FFNN</b></p>	 <p>(b) <b>RNN Description:</b> The structure of RNN is driven from the FFNN, but is appended the context units to store historical information. The outputs of hidden layers are fed back as the input of previous layer.</p> <p>(b) <b>RNN</b></p>		
	<p>the linear calculation feed to the hidden layer. After the internal training computations, the output results are the estimated SOHs.</p>		 <p>(b) <b>CNN Description:</b> CNN generally consists of convolutional, pooling, fully connected, and output layers. The convolutional layer is connected with the input data to perform the convolution, mapping the input data onto the features space. The pooling layer shrinks the size of the feature space. Through one or more stacks of convolutional layers and pooling layers, the input data are extracted into the features map.</p> <p>(b) <b>CNN</b></p>	[72,79–82,90-92, 98-101,104-110, 115-118, 120-125,129, 130,132-137]
	<p><b>Pros.</b> Strong ability to express nonlinearity; Capturing the long-term historical information (RNN); Automatically extracting features from raw data (CNN); High estimation accuracy. <b>Cons.</b> Over-fitting;</p>		<p>Poor uncertainty management; Performance depends heavily on the training process which cannot be comprehended; High time cost.</p> <p><b>Description:</b> Deriving from the Bayesian framework, GPR implements the probabilistic estimation with the combination of prior knowledge. It quantifies uncertainty behaviors in dynamic system based on the variance around the estimated mean in Gaussian process. <b>Pros.</b> Provide covariance to generate uncertainty level; Non-parametric; High flexibility. <b>Cons.</b> Performance is heavily affected by kernel functions; High time cost.</p>	
GPR: ARDDKLSparse GPRDual GPR	 <p>(e) <b>GPR</b></p>		[50-53,60, 145-151],	
EL: XGboostAda BoostLight GBMGBTRF	 <p>(f) <b>EL</b></p>	<p>Strong robustness to outliers; Simple to implement; Work well with a large dataset; Able to identify SOH feature importance. <b>Cons.</b> High complexity and computation burden.</p>	[156–111]	

(continued on next page)

Table 1 (continued)

Models:	Overviews	Ref.
TL: DDAFine-tuning	 <p>The diagram illustrates the (g) TL process. It shows a 'Source task' (represented by a dashed box containing two circles with triangles) and a 'Target task' (represented by a dashed box containing two circles with diamonds). Below the source task is a box labeled 'Knowledge' with an arrow pointing to the target task, labeled 'Transfer learning'. The entire process is labeled '(g) TL'.</p>	<p><b>Description:</b> TL can facilitate a qualified leaning model based on limited data. By transferring the knowledge and skills learned from the abundant source domain/tasks to the target domain or new tasks, it compensates the accuracy loss caused by the insufficient training data in the target domain to achieve desired performance. <b>Pros.</b> High generalization performance; Low time cost. <b>Cons.</b> Negative transfer issues.</p> <p>[134,135, 166-170,173]</p>
CM	<p>SVM#SVR</p> <p>RVM</p> <p>AR#ARIMA</p>	<p><b>Description:</b> SVM maps the data into the input space, and transforms the non-linear problems into a linear problem in a high dimensional feature space. In SOH estimation tasks, it adopts a non-linear mapping function to map the input data onto a high feature space, thus implementing the linear regression in this space. <b>Pros.</b> High approximation capability and low time cost on small dataset with high feature dimensions. <b>Cons.</b> Difficulty in kernel function selection; High time cost on large datasets.</p> <p><b>Description:</b> RVM has the same mathematical expression as SVM, but it provides the probabilistic estimation through the Bayesian framework. <b>Pros.</b> It provides the probabilistic SOH estimation. <b>Cons.</b> High time cost on large datasets.</p> <p><b>Description:</b> It is able to facilitate the short-term time-series prediction on a large number of datasets. By combing the observations from the previous time steps linearly, it incorporates previous results into the model to predict the future SOH values. <b>Pros.</b> Ease of implementation; Low time cost. <b>Cons.</b> Low accuracy when modeling the nonlinear degradation of the battery SOH.</p> <p>[24,60,89,118, 124,174,175]</p> <p>[89,141,176]</p> <p>[82,177]</p>



**Table 2**  
State-of-the-art SOH Estimation Models Published in 2020–2022.

Models	Tasks	Novelty	Datasets
<b>Published in 2022:</b>			
GRU-NN [90]	SOH estimation	ELM for temperature variation estimation based on randomly discontinuous short-term charging data; Finite difference for raw DT variation calculation; PCC for high-related features selection; GRUN for SOH estimation.	Oxford
NARX [105]	SOH estimation, RUL prediction	NARX model for offline SOH estimation; An ensemble based on NARX and similarity for online RUL prediction.	38 NMC cells datasets by cycle degradation tests
CNN-AST-LSTM NN [109]	SOH estimation, RUL prediction	1D-CNN and AST-LSTM for features extraction; Bayesian optimization for hyper-parameter adaptation.	NASA
LSTM-GMM-based VAE [86]	RUL prediction	Correlative and monotonic metrics for important features selection; A novel sequence-to-sequence predictive model via a VAE trained by GAN for RUL prediction; A mix-block based on LSTM and a GMM for probabilistic predictions.	Toyota & Stanford-MIT
Multi-feature-based multi-model fusion [145]	SOH estimation	Multiple LR, SVM and GPR are combined for SOH estimation; RF fuses the SOH estimation results.	Oxford
Sparse GPR [150]	SOH estimation	Correlation analysis for high-relative features selection; SparseGPR for SOH estimation.	Sandia
A novel DNNs [156]	RUL prediction	A newly-structured DNN with a special convolutional training strategy for RUL prediction using only one testing cycle.	Toyota & Stanford-MIT
TL-MMD [173]	RUL prediction	A TL model for RUL prediction based on CNN and improved domain adaptation.	Toyota & Stanford-MIT
Deep TL [184]	SOH estimation, RUL prediction	A deep TL model based on CNN and LSTM; CNN for features extraction, LSTM for knowledge capture; Fully-connected modules for capacities map and RUL prediction.	HUST
<b>Published in 2021:</b>			
IMGM [61]	SOH estimation, RUL prediction	Features are extracted from the partial IC curve peak area; Box-Cox transformation method for features optimization; Fuzzy comprehensive evaluation model for SOH estimation; An improved multivariate grey model for RUL prediction.	NASA
BHUMP [54]	SOH estimation	A ML pipeline is designed for SOH estimation by RF, DNNs, Bayesian Ridge Regression, and GPR; The designed pipeline also for automatic feature selection and algorithms calibration based on the segments of charging voltage and current curves.	NANA, CALCE, Oxford
Auto-CNN-LSTM [104]	RUL prediction	Autoencoder for features augment; CNN for mining deep information in features; LSTM for time information extraction in features; DDN for RUL prediction; First-order filter for results smoothing.	NASA
TL [106]	SOH estimation	DNN for complete charging curves estimation; TL for SOH estimation.	NASA, Oxford, CALCE
Hybrid-RBFN [125]	SOH estimation	A hybrid network for SOH estimation by combing the RBFN and AR models; Brownian motion and PF for parameters optimization of NN.	NASA, CALCE
DBN-based battery degradation model [130]	SOH estimation, RUL prediction	An aggregated feature method is designed to characterize the battery degradation dynamics; DBN for features extraction; PF for SOH estimation and RUL prediction.	LiFePO <sub>4</sub> /Graphite LIBs testing data from HKU
DL-based capacity prediction approach [133]	Capacity prediction	A DL model with multiple LSTM layers in both encoder and decoder blocks for estimating the entire capacity degradation trajectory in one shot without iteration or feature extraction.	48 cells datasets by degradation tests
WPD-CNN-HAMB [138]	RUL prediction	A forecasting approach for RUL prediction based on wavelet packet decomposition, two-dimensional CNN (2D-CNN), and adaptive multiple error corrections; Bivariate Dirichlet process mixture for forecasting residues evaluation.	NASA
GPR [142]	SOH estimation	An SOH estimator modeling the internal resistance as a Gaussian Process with Wiener velocity kernel and squared exponential kernel.	1027 solar off-grid cells running 400–760 days
RF [164]	SOH estimation	Two new SOH indicators are designed for degradation assessment; RF for SOH estimation; A degradation factor ranking approach for exploring battery degradation comprehension based on the permutation-based variable importance computation.	Real-life EVs datasets over 7 years.
<b>Published in 2020:</b>			
AST-LSTM NN [110]	SOH estimation, RUL prediction	An AST-LSTM NN model with many-to-one and one-to-on structures for SOH estimation and RUL prediction.	NASA
LS-SVM [124]	SOH estimation	ELM for the future voltage profile prediction based on the short-term charging data; A fixed size least squares-based SVM with a mixed kernel function for SOH estimation; Simulated annealing method for parameters optimization.	3 NCM cells datasets by degradation tests
TL [134]	SOH estimation	A novel TL model is designed for SOH estimation based on the LSTM combined with fully connected layers; Feature expression scoring rule for relevance assessment of multiple prediction tasks.	NASA, CALCE
A novel dual GPR [146]	Battery packs SOH estimation and RUL prediction.	Features extracted from partial charging curves, including the capacity loss, resistance increase, and inconsistency variation are examined; A dual GPR model is built for SOH prediction over the entire cycle life and RUL near the end of life.	Battery pack degradation tests datasets
GPR [147]	RUL prediction	GPR for RUL prediction based on EIS; It automatically determines which spectral features predict degradation.	Cambridge-Faraday
GPR-ARD [148]	Battery calendar degradation prediction	GPR for the underlying mapping among capacity, storage temperature, and SOC; Isotropic kernel function is modified with ARD structure to extract the high relevant input features for improving prediction accuracy and robustness.	9 lithium-ion cells datasets by calendar degradation tests
TL [172]	SOH estimation	A semi-supervised TCA algorithm for the maximal alignments of the distributions of the training and testing domain data in the latent feature space; A kernel ridge regression learner for SOH estimation.	NASA
Semi-parametric degradation model [177]	Battery degradation path prediction, RUL prediction	AR for battery degradation path prediction based on the partial observed path; Confidence intervals from a bootstrap distribution for RUL prediction.	LFP and NMC cells datasets by cycle degradation

expensive costs [19,21]. In contrast, the battery capacity can be directly measured by using the coulomb counting method under the conditions of constant current (CC) charging or discharging [17]. In this case, the capacity is frequently adopted to quantify the battery SOH in practical applications [17].

## 2. Feature extraction and selection

A data-driven SOH estimation model can be considered as a black box with the input of potential features (affecting the battery degradation mechanism) and the output of SOH estimation [24,37]. Note that a data-driven SOH estimation model heavily depends on the size and quality of datasets [25]. Hence, the feature extraction and selection are fundamental and crucial for facilitating accurate SOH estimation [126] [23,24,37,62–65].

### A. Battery SOH feature extraction

The battery SOH features can be further categorized into two types: as shown in Fig. 2:

a) **Directly-measured features** in BMSs such as voltage, current, charging and discharging time, and temperature [37]. Note that the voltage and temperature curves change noticeably during charge–discharge cycles. With the increase of internal resistance and loss of capacity caused by repeated charging and discharging, the voltage slope will increase during the charging process. Consequently, the charging and discharging time for a battery to reach the terminal voltage can be gradually shortened [66]. Furthermore, the temperature will fluctuate during the charging and discharging processes [67]. The temperature values (the maximum/minimum/mean temperature) and frequency also change when a battery ages. The aforementioned features could reflect the characteristics of battery degradation and be considered as the input features for a data-driven SOH estimation model to quantify battery degradation [68].

b) **Indirectly-inferred features** through the incremental and differential analysis based on the voltage and temperature curve [19,23]. Recently, the indirectly calculation methods are proposed to cover the limitations of features that extracted from direct measurement data [25,69]. The commonly-adopted calculation methods include incremental analysis and differential analysis [25,37], both of which are effective tools to analyze the information on the measured variables such as voltage, current, temperature, and time, etc. Based on the curves of measured variables that map the characteristics of battery degradation, the useful features can be efficiently extracted. The most frequently adopted methods are IC [70–72], DV [73], differential temperature (DT) [74], and DTV analysis [75].

IC is calculated from the differentiation of charged or discharged capacity with voltage in a short period. It transforms the voltage plateaus in the charge or discharge curve into the apparent peaks into the IC curve with voltage. Note that each peak in the IC curve reflects a phase equilibrium position in the battery voltage curve [76]. In contrast, DV is defined as the inverse of IC, which dif-

ferentiates the change of voltage with the charged or discharged capacity in a short period [77]. In the DV curve with capacity, the peaks represent the phase transitions in the electrodes hidden in the battery operations, and the valleys represent the plateaus of battery voltage trends [77,78]. Each peak or valley in the IC/DV curve specifically identifies the battery degradation depending on the unique IC/DV peak value [70], position, integrated area [37], and peak position [66]. Therefore, the IC and DV analysis formulate powerful tools to capture the battery degradation trends.

Similar to IC/DV analysis, the DT and DTV analysis introduce the temperature measurements to infer the battery degradation information [73–78]. DT transforms the temperature values into DT curve, which can clearly identify the parse transition and equilibrium position in the temperature curve [78]. It is plotted as the DT value versus capacity calculated by differentiating temperature over a significantly small capacity interval [73]. Note that each peak in a DT curve denotes the rapid temperature increase or decrease tendency [77]. Since the battery temperature is a vital factor affecting the battery degradation, the peak value and positions extracted from a DT curve can be adopted as essential features for SOH estimation [78]. Meanwhile, DTV connects DT and battery voltages for the thermodynamic information regarding the electrode materials [74]. The DTV is obtained by differentiating the temperature to the voltage, and the presented DTV curve plots the DTV values with voltages. Note that each peak in a DTV curve represents a specific phase transition of the negative or positive electrode [74,75].

In an IC/DV/DT/DTV curve, the peak characteristics (e.g., height, width, position, and integrated area) are strongly connected with battery degradation (i.e., capacity loss and internal resistance increase) [73]. Hence, the peaks obtained by above differential analysis methods can be taken as vital features to estimate the potential battery SOH. However, the fitted curves commonly suffer from the non-negligible noises by measurements. The empirical model decomposition (EMD) method [79–82] as well as filter-based methods [72,83–85] (such as moving average filtering [83], Gaussian filtering [84], differential filtering [85] and KF [72]) are commonly adopted for noise depression and outlier detection under such circumstances.

### B. Battery SOH feature selection

To reduce the computational burden as well as improve the estimation precision, feature selection is urgently required to pick up the high-relevant features as the input of the SOH estimation model [25,37,81,86–88]. Recent SOH feature selection methods mostly focus on the correlation coefficient analysis [37], e.g., the grey relational analysis (GRA) [81,89] and Pearson correlation coefficient (PCC) [90–92]. GRA provides a quantitative measurement to the correlation between two or multiple active factors during the system variation. According to the degree of similarity or dissimilarity between the system varying factors, i.e., the grey relation degree, the degree of correlation can be quantified into the [0, 1] interval [81]. Hence, it can be employed to select the features closely connected to the battery SOH. On the other hand, the PCC calculates the linear relationship between the input and SOH estimation accuracy [91], thus helping an SOH estimation model to select features highly relevant to the estimation accuracy. Moreover, Darius et al. [54] adopt the recursive feature elimination and cross-validation to perform automatic feature selection. Wang et al. [93] and Wang et al. [94] adopt the light gradient boosting machine method to fit the feature importance related to SOH.

## 3. Data-driven SOH estimation models

A data-driven SOH estimation model functions on mass data without necessary dependence on the battery internal degradation

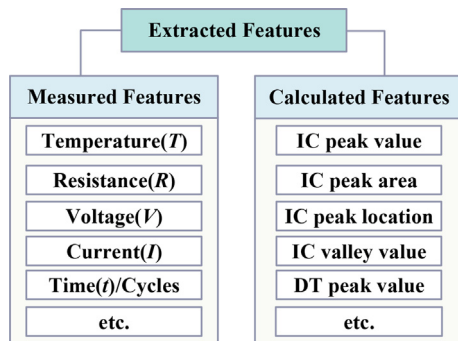


Fig. 2. The extracted features from various features extraction methods.

mechanism (which, however, can be modeled as the expert knowledge to improve a data-driven model's performance). Both directly-measured and indirectly-inferred features can be adopted as its input [24,25,95] to build a reliable relationship between the features and battery SOH, thus achieving precise SOH estimations [33–35]. This section sorts out the data-driven SOH estimation models published during 2018–2022, and classifies them into several broad categories, i.e., NN, GRP, EL, transfer learning (TL), and conventional models (CM). Table 1 summarizes the data-driven SOH estimation models, including their brief plots, advantages, disadvantages and variants. Table 2 lists the state-of-the-art data-driven SOH estimation models published in 2020–2022, which provides researchers with convenience to find the most advanced benchmarks in current research status of battery SOH estimation.

### C. Neural network (NN)

An NN-based learning model possesses excellent ability to address complex non-linear problems [47,48,87,96,97]. In an NN-based battery SOH estimation model, the related features are taken as the input [24], the model parameters are trained to establish the mapping relationship between the input features and battery SOH [60], and the outputs are the estimated battery SOH [98]. According to the type of NN and its popularity in addressing the issue of battery SOH estimation, we focus on the feed-forward neural network (FFNN) [99–103], recurrent neural network (RNN) [104,105] and convolutional neural network (CNN) [106,107]-based battery SOH estimation models. Their general structures are illustrated in Figs. 3–5, respectively. Moreover, researchers are showing increasing interests in hybrid neural networks (hybrid-NNs)-based SOH estimation models [108–113], which will be discussed later.

#### a) Feed-Forward neural network (FFNN)

Given the  $N$ -dimensional features of the  $i$ -th training instance, FFNN takes the features into the input layer as shown in Fig. 3, and then transmits the inputs to the hidden layer as:

$$h_l = g\left(\sum_{n=1}^N x_{(i)n} w_{n,l}^H + b_l\right), l = 1, \dots, L \quad (3)$$

where  $L$  denotes the number of neurons in the hidden layer,  $x_{(i)n}$  is the  $n$ -th input feature,  $w_{n,l}^H$  is the weight that connects the  $n$ -th input neuron and the  $l$ -th hidden neuron,  $b_l$  denotes the bias of the  $l$ -th hidden neuron, and  $g(\cdot)$  denotes the activation function, i.e., Sigmoid, Tanh, rectified linear unit (ReLU), etc. [14,114]. With one hidden layer, the output can be calculated as:

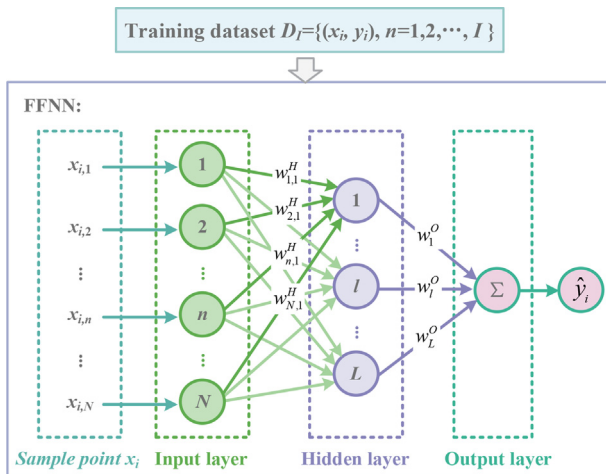


Fig. 3. The structure of FFNN.

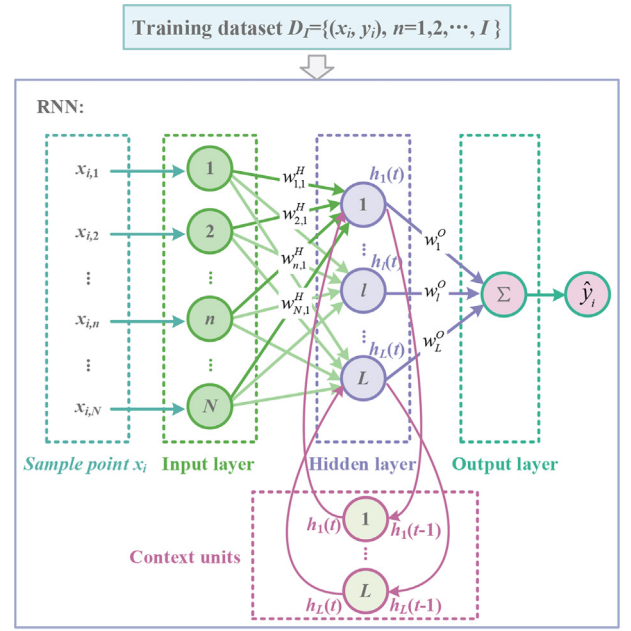


Fig. 4. The structure of Elman-NN.

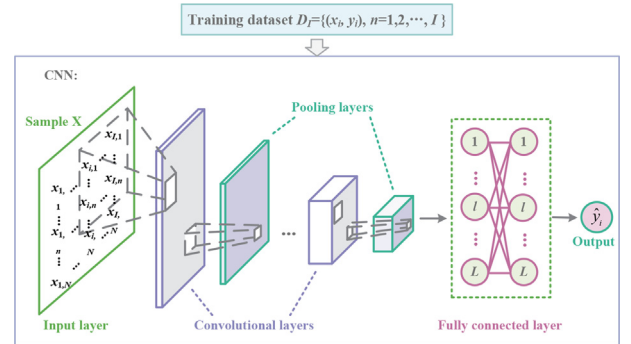


Fig. 5. The structure of CNN.

$$\hat{y}_i = g\left(\sum_{l=1}^L h_l w_l^O\right) \quad (4)$$

where  $h_l$  represents the output of the  $l$ -th hidden neuron,  $w_l^O$  is the output weight of the  $l$ -th hidden neuron. Note that similar inferences can be achieved with multiple hidden layers in an FFNN-based model.

When addressing the issue of battery SOH estimation, an FFNN-based model feeds the features related to the battery SOH degradation into the input layer, and then estimates the battery SOH data with the output results  $\hat{y}_i$  after the internal information transmission. Wu et al. [100] adopt an FFNN to simulate the relationship between SOH and charge curves under different cycle numbers. Considering the non-linearity of battery charge curves, the importance sampling method is applied to the input features before building the FFNN. With the reduced features, the training process is simplified with low computational burden. In order to achieve the best-fitted battery SOH, several optimization algorithms like back-propagation (BP) [91,99,115], genetic algorithm (GA) [116,117] and particle swarm optimization (PSO) [80,118,119] are employed to train an FFNN. Tao et al. [99] adopt GA to optimize the parameters in FFNN on large datasets so as to guarantee the robustness and the estimation accuracy. Meanwhile, an FFNN-



based SOH estimation model is also compatible with prior knowledge [120,121]. For instance, Dai *et al.* [108] build constants on prior knowledge to enhance an FFNN-based SOH estimator. It is able to fit complex nonlinear patterns during battery degradation.

To further accelerate the training process of an FFNN-based SOH estimation model, the extreme learning machine (ELM) [122,123] is adopted. For such a model, the input weights and biases are hypothesized at random instead of being iteratively trained, making the output weights be analytically solvable [123]. Hence, the learning process of ELM costs much less time than a standard FFNN-based model does. Nonetheless, with the less well-adjusted weights and biases, ELM commonly suffers from lower estimation accuracy than a standard FFNN-based model whose parameters are carefully tuned [123].

Chen *et al.* [90] adopt an ELM to estimate the entire temperature trend during the changing process based on the discontinuous and short-term changing data. With it, the multiple-dimensional health-related features are extracted from the entire temperature curves by DT analysis, and six important features are selected by PCC for facilitating battery SOH estimation. Gou *et al.* [92] utilize an ELM to map the relationship between battery SOH-related features and capacity for modeling the battery degradation trend. Considering the high computational efficiency of an ELM model, they further take the ELM as the base learner of an ensemble learning (EL) model for improving the SOH estimation accuracy. Shu *et al.* [124] prove that the ELM is effective in learning the inherent characteristics of voltage variation for enhanced features of battery SOH estimation.

A radial basis function network (RBFN) also outperforms a traditional FFNN in computational efficiency. Compared with an FFNN with multiple hidden layers, an RBFN commonly adopts only one hidden layer and significantly improves the computational efficiency. Lin *et al.* [125] propose an adaptive tunable hybrid RBFN for battery SOH estimation, which achieves adaptively adjustment for the structural parameters in the battery degradation mechanism without the physics-based models. She *et al.* [55] extract related-features through the IC analysis to build an SOH estimation model, which is fused with the slope-bias correction and RBFN. Then, the RBFN is utilized to identify the non-linear relationships.

#### b) Recurrent neural network (RNN)

RNNs are frequently employed in battery SOH estimation, which structure is derived from the conventional FFNN with the appended context units for memorizing the historical information [24,127,128]. It is known that the RNN is effective in capturing the temporal patterns in dynamic SOH data. The Elman neural network (Elman-NN) is a specifically structured RNN [104,105,128], which is frequently adopted in SOH estimation models [82]. As shown in Fig. 4, at the  $t$ -th moment, its context units receive the outputs values of the hidden layer at the  $(t-1)$ -th moment, and then combine them with the input  $h_i(t)$  to achieve the outputs at the  $t$ -th moment. Such principle can be expressed as:

$$\forall l \in \{1, \dots, L\} : \quad h_l(t) = g \left( \sum_{n=1}^N [x_{(i)n}(t)w_{n,l}^H(t) + h_l(t-1)w_l^C(t-1)] + b_l \right), \quad (5)$$

where  $x_{(i)n}$  denotes the input features of the  $n$ -th input neuron at,  $w_{n,l}^H$  is the hidden weight connecting the  $n$ -th input neuron and the  $l$ -th hidden neuron,  $w_l^C$  is the hidden weight connecting the context unit and hidden neuron at the previous moment,  $b_l$  denotes the bias of the  $l$ -th hidden neuron, and  $g(\cdot)$  denotes the activation function, respectively. Similar to the situations in FFNN, the weights and biases in RNN can be optimized by an optimization algorithm like BP [91,115], GA [99,129] and PSO [80,130,131].

Chen *et al.* [82] design an Elman-NN to achieve battery SOH estimation based on the voltage and capacity of battery lifecycle

data. Raw data are processed by the EMD, where the multiple data sequence and related residuals are extracted. The subsequent time series data are modeled by an autoregressive moving average model, and the residuals are modeled by an Elman-NN. Their fused results formulate the estimated SOH data. Khaleghi *et al.* [105] develop a battery prognostics and health management (PHM) model that consists of offline battery SOH estimation and online RUL prediction. The health features extracted from the partial charging voltage curves are fed into an RNN for battery SOH estimation, and the estimated SOH data are utilized as the prediction features for RUL prediction.

To overcome the gradient vanishing and exploding during the process of back propagation in RNN-based SOH estimators [24,25,37], the long short-term memory (LSTM) is adopted [72,132]. LSTM puts three gates, i.e., input, forget and output, into the context unit to control the gradient propagation:

- The input gate selects the new information to be stored in the internal state;
- The forget gate identifies and discards the redundant information that cannot be adopted to determine the estimation results in the next moment; and
- The output gate calculates the outputs and updates the hidden layers.

Li *et al.* [133] propose a model with multiple LSTM layers to estimate the battery degradation trend in one shot without iteration and feature extraction. Cheng *et al.* [79] combine the EMD and LSTM to facilitate SOH estimation and RUL prediction, where the EMD selects irrelevant signals from the SOH data and the LSTM is unitized to predict the RUL. Qu *et al.* [80] propose to incorporate the attention mechanism into LSTM for battery SOH estimation and RUL prediction. They apply incremental learning principle to the achieved SOH estimation model, thus facilitating efficient incorporation of new data. Li *et al.* [110] propose the variant LSTM-NN-based SOH estimator. Such an SOH estimator firstly couples the input and forget gates of the LSTM through a fixed connection to compress the old and new information simultaneously for capturing more beneficial information, and then appends a peephole connection generated from the constant error carousel to the output gate for blocking redundant errors. Considering the LSTM's ability to learn the long-term dependency during the battery degradation, Tan *et al.* [134] and Shu *et al.* [135] combine the LSTM with full connected layers as the base model to establish a TL-based battery SOH estimation model.

A gated recurrent unit network (GRUN) is another kind of gated-based RNN with the simpler structure than an LSTM [136]. GRUN merges the input and forget gates into the update gate by removing the partition between the internal and external states. Thus, the linear relationship between its current state and previous state is directly established, which can be utilized to solve the gradient exploring and eliminating issues. Chen *et al.* [90] propose a GRUN based battery SOH estimation model, where the health features are extracted from DT curves, and closely-related features are selected by PCC as the GRUN inputs. Chen *et al.* [81] build a RNNs consisting of RNN, LSTM, bidirectional LSTM, and GRUN for constructing hybrid SOH estimation model.

#### c) Convolutional neural network (CNN)

A CNN-based battery SOH estimator has strong feature extraction abilities owing to its elaborately designed structure [137,138]. A CNN model with the 2-dimensional input commonly consists of convolutional, pooling, fully-connected and output layers [24]. The data are fed into the input layer and then mapped into the feature space through the filter-sliding-based convolution operations. Sequentially, the pooling condenses the features with a given window for representing higher-level conceptions. Through one or more stacks of convolutional layers and pooling layers, the feature map is successfully built on the input data as

shown in Fig. 5. Finally, the fully connected layer can be trained based on the achieved features to recognize desired patterns from the achieved features.

Considering the CNN's representation ability and LSTM's long-term time series expression ability, Ren *et al.* [104], Li *et al.* [109] and Ma *et al.* [139] propose to combine them for battery SOH estimation. To overcome the limitation of insufficient data, Ren *et al.* [104] apply the autoencoder to augment the data dimension for enhancing the input for CNN and LSTM-based SOH estimator. Li *et al.* [109] combine a 1-D CNN with an LSTM for SOH estimation. To avoid the heavy computational burden of hyper-parameter tuning, they adopt the Bayesian optimization on the priors of the battery data to implement hyper-parameter adaptation. To address the accuracy loss caused by unreliable sliding window size, Ma *et al.* [139] adopt the false nearest neighbors to calculate the required sliding window size for establishing an accurate battery SOH estimator.

Moreover, Tian *et al.* [106] take the partial charging curves as the input for a one-dimensional CNN (1D-CNN) for estimating the entire charging curves, which demonstrate the battery degradation process. In order to exhibit the performance of CNN-based estimator on different types of batteries and operating scenarios, Tian *et al.* employ the TL to estimate the charging curves of new batteries with a trained CNN. The experimental results demonstrate that the modified CNN efficiently improves the estimation accuracy of charging curves.

#### d) Hybrid neural networks (Hybrid-NNs)

The Hybrid-NNs-based battery estimators combining LSTM and CNN are proposed in [104], [109], [139] and [140], one combining ELM and GRUN is proposed in [90], and one combining multiple RNNs is proposed in [81]. The above-mentioned models all adopt the principle of EL to achieve more accurate SOH estimation results. Since the empirical data uncertainty influences the accuracy of an NN-based SOH estimator, Huang *et al.* [86] propose a sequence-to-sequence RUL prediction model based on a variant autoencoder and with the generative adversarial network (GAN) learning scheme. Moreover, such a RUL prediction model further integrates the LSTM and Gaussian mixture model (GMM) for probabilistic prediction. On the other hand, the immeasurable battery capacity data may also affect the SOH estimation accuracy. To address such issue, Lyu *et al.* [141] design an online battery PHM system based on a hybrid model that combines complementary ensemble EMD (CEEMD), NN and relevance vector machine (RVM): the time-series features extracted from the partial charging data are decomposed into high and low frequency signals through the CEEMD model; then the degradation trends of high and low frequency signals are captured and estimated by LSTM and FFNN, respectively. These estimated signals are rebuilt and employed to estimate the degradation trend of features, and finally, the mapping relationship between degradation features and capacity are established through RVM optimized by grey wolf optimizer (GWO) to achieve accurate SOH estimations.

#### D. Gaussian process regression (GPR)

Regression-based battery SOH estimators are also attracting researchers' attentions [50–53]. Deriving from the Bayesian framework, GPR realizes the probabilistic estimation with the prior knowledge combination [142,143]. GPR quantifies uncertainty behaviors in a dynamic system based on the variance estimation in the Gaussian process [60,144]. Since the battery degradation mechanism is influenced by many factors, the GPR model produces the probability distribution of possible battery SOH estimation as follows [25]:

$$\begin{aligned} m(x) &= E(f(x)), \\ k_f(x_i, x_j) &= E[(f(x_i) - m(x_i))(f(x_j) - m(x_j))]; \end{aligned} \quad (6)$$

where the co-variance function  $k(x_i, y_j)$  captures the similarity of different input samples. Owing to its non-parametric and model-free characteristics, GPR can flexibly establish a complex non-linear relationship between the input features and battery degradation trend [60].

Considering the fact that the battery SOH estimation accuracy can be affected by the complex mechanism of battery degradation, Feng *et al.* [50] propose a GPR model to estimate the battery capacity trajectory under multi-battery settings. Related trajectories in history data are employed to enhance the estimation accuracy to target capacity trajectory. They empirically validated that the GPR with the multiple input single output structure obtains better performance than those depending on the single input single output and multiple input multiple output structures. Lin *et al.* [145] propose to fuse GPR, support vector regression (SVR) and multiple linear regression (LR) for battery SOH estimation. The estimated SOH values are further fused by RF to improve the estimation accuracy. Aiming at alleviating the computation and storage costs, Hu *et al.* [146] propose a dual GPR-based SOH estimator which takes advantage of GPR's low training and storage costs.

Zhang *et al.* [147] establish the EIS-capacity GPR model through with a zero mean function and a diagonal covariance function with automatic relevance determination (ARD). Their empirical results indicate that the GPR with the ARD kernel can identify the more irrelevant features from the high-dimensional measurements in the EIS-capacity estimator. Liu *et al.* [148] modify the isotropic kernel function with ARD to improve the accuracy and robustness of an SOH estimator. Ma *et al.* [149] propose an improved GPR model named the deep kernel learning (DKL) model by replacing the traditional kernel function with an NN. Combining with the advantages of an NN and a single squared exponential kernel function, this model is simpler and more robust. To address the random partial charging data, Deng *et al.* [150,151] propose a sparse GPR-based model for battery SOH estimation, where the input features are the random capacity increment sequences extracted from the partial charging processes.

#### E. Ensemble learning (EL)

An EL method refers to the combination and generation of multiple base learners to facilitate a particular learning task [152,153]. By gathering different weights and results from the multiple base learners, EL compensates for the weakness of each single base learner to improve its performance and robustness [24,154,155]. EL methods mostly depend on the dependent or an independent framework. In the dependent framework, base learners are heterogeneous models or algorithms like NN [104,156], support vector machine (SVM) [124] and so on. Boosting is a well-known EL method in the dependent framework, including XGBoost [157,158], AdaBoost [152,159], gradient boosting tree (GBT) [160–162], Light gradient boosting machine (LightGBM) [163], and so on. Song *et al.* [157] and Jafari *et al.* [158] adopt the XGBoost method to build an ensemble for estimating the Lithium-ion battery SOH. Considering the complex nonlinear battery dynamic data, GBT is a powerful SOH estimation model. Yang *et al.* [160] propose a regression GBT algorithm for battery RUL prediction based on extracted features, where the additional trees are constructed by minimizing the prediction residuals from existing base models. Khaleghi *et al.* [161] utilize GBT to estimate the battery SOH in a realistic battery condition. Liu *et al.* [163] adopt lightGBM-based model for RUL prediction owing to its robustness.

On the other hand, homogeneous base models are built on disjoint subsamples in an independent ensemble framework like bootstrapping [24,152]. The RF method adopts the bootstrapping framework for battery SOH estimation owing to its accuracy gain and ease of implementation [54,55,164]. RF consists of multiple individual decision trees as shown in Fig. 6.  $K$  sample subsets are generated from the original SOH dataset  $D_i$  through the bootstrap-

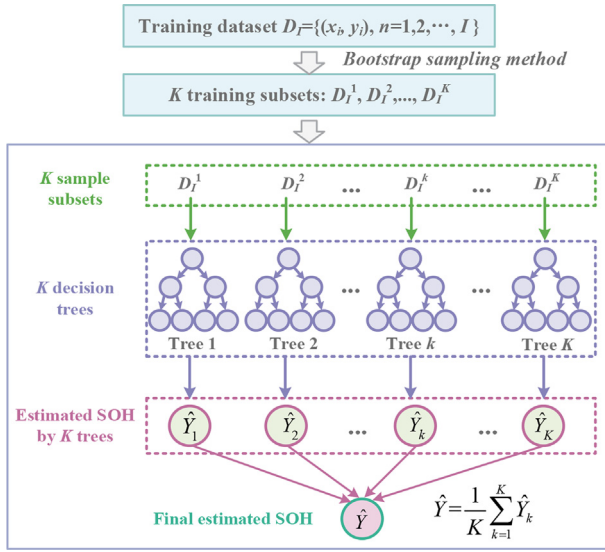


Fig. 6. The structure of RF.

ping as the inputs for multiple decision trees, whose outputs are aggregated to generate the SOH estimation [145,165]. RF decreases the variance that related to the estimated values, thereby alleviating the over-fitting to achieve higher estimation accuracy for unknown battery SOH.

To avert regional training sample's impacts on the model performance, Li *et al.* [159] adopt NN as the base learner to build an EL model with the AdaBoost framework. To address the battery inconsistency in practice, She *et al.* [55] propose an online adaptive EL scheme based on the offline pre-trained based models and online generalized weights. The offline model is established by combing slope-bias correction and RBFN, and the weights are adjusted online according to the new SOH data, thus effectively reducing the data dependency. Inspired by the idea of multiple-model fusion, Lin *et al.* [145] apply the RF method to fusing the battery SOH estimations generated by multiple SOH estimators for further improving the estimation accuracy.

#### F. Transfer learning (TL)

The above-mentioned models, such as NN [99,104], GPR [37,135,149] and EL [55,159], are defined on the precondition that the training and testing data are taken from the same domain. They all require large amount of battery SOH data [135]. When the training data are inadequate, a TL-based approach is able to establish a high-performance SOH estimator from the limited data [166,167]. It transfers the knowledge and skills learned from the abundant source domain to the target domain or new tasks, thus compensating the accuracy loss caused by data insufficiency in the target domain.

Recently, TL and NN are often aggregated for implementing accurate SOH estimators [168], where the fine-tuning strategy is the most commonly-adopted approach. Fine-tuning a pre-trained network model such as CNN on a new dataset can facilitate the knowledge transferring in context of deep learning. Methods have been proposed to fine-tune a) all network parameters, or b) the parameters of the last few layers [135]. The pre-trained model can be also adopted as a fixed feature extractor [168]. The principle of TL with fine-tuning strategy is illustrated in Fig. 7, where the left panel depicts the base model trained by the source datasets, and the right panel depicts the transferring model by setting the target datasets as the input variables. After establishing the base model by pre-training a network on the source data, fine-tuning strategy keeps the base model's most trained parameters, and further feeds the target data into the based model to adjust a small portion of

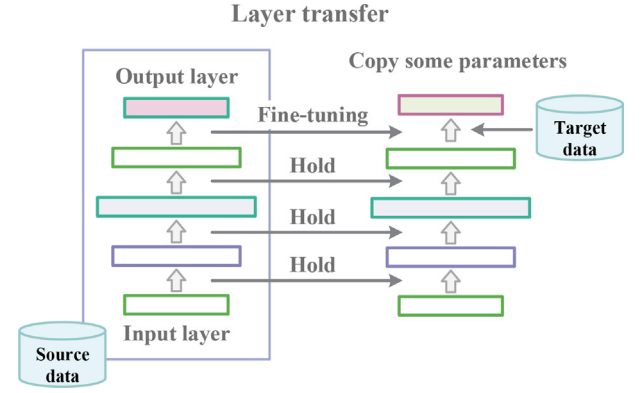


Fig. 7. TL with fine-tuning.

parameters in part of the achieved model's layers [135]. Since only a small portion of parameters in the base model are adjusted according to the target data, the TL with fine-tuning strategy enables high computational efficiency. Moreover, it is able to address SOH-related tasks of different batteries when the target domain data are inadequate with very competitive accuracy [135]. In order to measure the battery SOH inconsistencies among batteries in the same battery pack, Shu *et al.* [135] combine the ideas of LSTM and TL with fine-tuning for battery SOH estimation based on partial charging data. Tan *et al.* [134] design a TL with fine-tuning-based battery SOH estimator, whose base model is an LSTM with fully-connected layers. Tian *et al.* [106] build a TL-based SOH estimator based on a pre-trained deep neural network (DNN) to estimate the new battery degradation trajectory under various working conditions.

The deep domain adaptation (DDA) is also commonly adopted to establish a TL-based SOH estimator [169,170], like maximum mean discrepancy (MMD) [173], Wasserstein distance [171] and transfer component analysis (TCA) [172]. Ma *et al.* [173] propose a TL-based SOH estimator that combines the CNN and MMD, where the CNN is adopted to automatically extract the features from the charging data, and the MMD is employed to decrease the distribution difference between the training and testing datasets.

#### G. Conventional models (CM)

##### a) Support vector Machine (SVM)

SVM is one of the most popular models for classification and regression tasks in complex and non-linear systems [89], which is especially suitable for small datasets [118]. It transforms a low-dimensional non-linear problem into a high-dimensional linear problem through the kernel method [24,60,124] as:

$$y_i(x) = \sum_{n=1}^N \omega_n K(x_i, x_{i,n}) + \varepsilon \quad (7)$$

where  $\omega_n$  represents the weights connecting the feature space to the output,  $K(\cdot)$  is the kernel function and the  $\varepsilon$  is the noise.

Li *et al.* [118] propose a least squares-based SVM for battery SOH estimation, which is optimized by an improved bird swarm algorithm to premature convergence. Shu *et al.* [124] adopt a fixed-size least squares-based SVM to learn the battery degradation trend and implement battery SOH estimation on the ELM-predicted voltage profiles. Furthermore, such model adopts the simulated annealing algorithm to optimize the hyper-parameters. Wei *et al.* [174] establish an SVR-based model to simulate the battery degradation mechanism, whereby the battery SOH estimation model is further achieved via the PF algorithm for estimating degradation parameters. Li *et al.* [175] establish an SVR-based battery degradation model to estimate the battery capacity according to the partially incremental capacity curves.



### b) Relevance vector Machine (RVM)

RVM has the same mathematical expression as that of SVM given in (7), and it provides the probabilistic estimation through the Bayesian framework [176]. RVM's kernel function has high sparsity for distinguishing the margins [25]. Zhao *et al.* [89] combine RVM and grey model to estimate the battery SOH and correct each other's estimation results alternatively, thus reducing the estimation errors. Lyu *et al.* [141] adopt the GWO to optimize the RVM, thereby establish the mapping relationship between degradation features and capacity for battery SOH estimation.

### c) Autoregression (AR)

An autoregression (AR) model can distinguish the short-term time-series from massive data [123,177]. It linearly combines the observations from the previous time steps to predict the SOH in the next time steps [33]. To better fit the non-linear relationship between the SOH features and battery degradation, the autoregression integrated moving average (ARIMA) model is adopted for battery degradation estimation [33,82], which combines the AR and moving average. Chen *et al.* [82] combine the ARIMA and Elman-NN to achieve accurate estimations for battery SOH, where the ARIMA model and Elman-NN model are constructed by training subsequence time-series and residuals, respectively.

Datasets and Metrics.

### H. Available public datasets

This section summarizes several available public datasets published by national prestigious laboratories and institutions. Their details are summarized in Table 3.

**a) NASA battery dataset.** The NASA battery degradation data [178] are generated by testing one type of lithium-ion battery through 3 different operational profiles (i.e., charge, discharge and impedance) at a specific temperature. Charge was conducted in a CC mode at 1.5A till the battery voltage reached 4.2 V, and then continued in a constant voltage (CV) mode till the charge current dropped to 20 mA. Discharge was conducted at 4A till the battery voltage fell to a specific voltage value. Impedance measurement was taken through.

**Code availability:** [54]: <https://doi.org/10.5281/zenodo.4390152>; [106]: <https://github.com/aesabattery/Charging-Curve-Prediction>; [109]: <https://github.com/Lipenghua-CQ/CNN-ASTLSTM>; [133]: <https://git.rwth-aachen.de/isea/battery-degradation-trajectory-prediction>; [147]: <https://github.com/YunweiZhang/ML-identify-battery-degradation>.

**Data availability:** [178] <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/>; [179] <https://web.calce.umd.edu/batteries/data.htm>;

[180] <https://howey.eng.ox.ac.uk/data-and-code/>; [181,182] [https://www.batteryarchive.org/sn\\_study.html](https://www.batteryarchive.org/sn_study.html); [183] <https://data.matr.io/1/>;

[147] <https://doi.org/10.5281/zenodo.3633835>; [184] <https://doi.org/10.17632/nsc7hnsq4s.2>.

an EIS frequency sweep from 0.1 Hz to 5 kHz. Different testing batteries are conducted on the different ambient temperatures and discharge terminal voltages.

**b) CALCE battery dataset.** The CALCE [179] group provides the baseline cycle-life data for public research. The logical starting point for most battery life-cycle testing is to apply a CC/CV charging protocol and then discharge under a CC rate. This dataset provides a useful baseline for comparison against more complex discharge profiles, which can be used for battery SOH estimation and RUL prediction. It includes continuous full and partial cycles curves, storage curves, dynamic-driven curves, open circuit voltage measurements, and impedance measurements.

**c) Oxford battery dataset.** The battery intelligence laboratory at the University of Oxford [180] publishes the battery degradation dataset, which contains battery degradation data from eight lithium-ion pouch batteries. The batteries were all tested in a thermal chamber at 40 centigrade. The batteries were exposed to a CC-CV charging profile, followed by a drive cycle discharging profile that was obtained from the urban Artemis profile. Characterization measurements were taken every 100 cycles.

**d) Sandia battery dataset.** The Sandia national lab [181,182] provides battery cycle degradation data, which are generated by three types of batteries (i.e., lithium-iron-phosphate (LFP) cells from A123 system, lithium-nickel-cobalt-aluminate (NCA) cells from Panasonic and lithium-manganese-cobalt-oxide (NMC) cells from LG Chem) under different temperatures and different discharge currents. Raw data from cycling, abuse and EIS, and the battery information are recorded.

**e) Toyota & Stanford-MIT battery dataset.** The Toyota research institute unites with Stanford and MIT [183] to publish a battery dataset consisting of 124 commercial lithium-ion batteries that were cycled to failure under fast-charging conditions. The tested batteries were cycled in horizontal cylindrical fixtures on a 48-channel Arbin LBT potentiostat in a forced convection temperature chamber set to 30 centigrade. The cells have a nominal capacity of 1.1 Ah and a nominal voltage of 3.3 V. All batteries in this dataset are charged with a one-step or two-step fast-charging policy with the format of 'C1(Q1)-C2', where C1 and C2 are the first and second constant-current steps, and Q1 is the state-of-charge (SOC, %) at which the currents switch, respectively. The second current step ends at 80% SOC, after which the cells charge at 1C CC-CV. The upper and lower cutoff potentials are 3.6 V and 2.0 V, which are consistent with the manufacturer's specifications.

**f) Cambridge & Faraday battery dataset.** The Cambridge University unites with the Faraday Institution [147] to conduct the battery degradation experiments by applying a continuous charge-discharge cycle on 12 45mAh EunicellLR2032 lithium-ion coin batteries. The batteries chemistry is LiCoO<sub>2</sub>/graphite. The batteries are cycled in three climate chambers set to 25, 35 and 45 centigrade, respectively. Each cycle consists of a 1C-rate (45 mA)

**Table 3**  
Publicly-available Battery SOH Datasets.

Battery Datasets	Published Year	Battery (form size chemistry)	Test variables	Data given	Battery counts
NASA [178]	2007	18,650 2Ah	Dhrg, T	Q, IR, V, I, T	34
CALCE [179]	2011	18,650 2.2 Ah LCO	Chrg, Dhrg, T	Q, IR, V, I, T	28
		prismatic 1.1 Ah LCO	Chrg, Dhrg	Q, IR, V, I, T	15
		prismatic 1.35 Ah LCO	Chrg, Dhrg, T	Q, IR, V, I, T	12
		Pouch 1.5 Ah LCO	Chrg, DOD	Q, V, I, T	16
Oxford [180]	2011	18,650 3 Ah NCA/graphite	Chrg, Cal	Q, E, V, I, T	28
Sandia [181,182]	2020	18,650 multiple cells	Dhrg, DOD, T	Q, E, V, I, T	86
Toyota&Stanford-MIT [183]	2021	18,650 1.1 Ah LFP/graphite	Chrg	Q, IR, V, I, T	124
Cambridge&Farada [147]	2020	18,650 1.1 Ah LFP/graphite	Chrg	Q, V, I	12
HUST [184]	2022	18,650 1.1 Ah LFP/graphite	Dhrg	Q, V, I	77

Note that 'Cal' denotes calendar degradation, 'DOD' denotes the depth of discharge, 'Chrg' denotes charge, 'Dhrg' denotes discharge, 'E' denotes 'energy', and 'IR' denotes both internal resistance and impedance.

CC–CV charge up to 4.2 V and a 2C-rate (90 mA) CC discharge down to 3 V. The EIS is measured at nine different stages of charging/discharging during every even-numbered cycle in the frequency range of 0.02 Hz–20 kHz with an excitation current of 5 mA and 15-min open circuit at SOC 0%–100%.

**g) HUST battery dataset.** The Huazhong University of Science and Technology [184] provides a comprehensive dataset consisting of 77 commercial batteries with over 140,000 charge–discharge cycles. The adopted batteries are all LFP/graphite A123 APR18650M1A with 1.1 Ah nominal capacity and 3.3 V nominal voltage. Each battery is tested with a different multi-stage discharge protocol but following the identical fast-charging protocol in two thermostatic chambers at 30 centigrade.

#### B. Evaluation metrics

This section summarizes commonly-adapted evaluation metrics for data-driven SOH estimation models [185].

**a) RMSE.** The root mean squared error describes the deviation between the actual SOHs and their estimations given as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (8)$$

where  $n$  is the cycle number of one battery from initial to fault statuses,  $y_i$  denotes the actual SOH value of testing battery at the  $i$ -th cycle, and  $\hat{y}_i$  denotes its estimation generated by the tested model, respectively.

**b) MAE.** The mean absolute error denotes the absolute difference between the actual SOHs and their estimations formulated as:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (9)$$

Note that small RMSE and MAE stand for high estimation accuracy of a tested SOH estimation model.

**c) MAPE.** The mean absolute percentage error evaluates the ratio of the estimation error to the actual SOHs, which is calculated as:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%. \quad (10)$$

When the MAPE gets closer to 0%, the tested SOH estimation model possesses higher accuracy.

**d)  $R^2$ .** The  $R^2$ , i.e., coefficient of determination, is utilized to measure the fitness degree of the SOH estimations to the actual SOHs of testing battery, which is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad (11)$$

where  $\bar{y}$  denotes the average value of the actual SOHs of one testing battery, i.e.,  $\bar{y} = (1/n) \sum_{i=1}^n y_i$ . The performance of an SOH estimation model is high when its  $R^2$  is close to 1.

## 4. Benchmarks studies

This section presents the benchmark studies, including data preparation, model training strategy and evaluation of several state-of-the-art models, i.e., NN (including LSTM and CNN), GPR (with ARD kernel), EL (RF), and TL (with fine-tuning strategy).

#### A. Data preparation

Ten commercial batteries with cycle lives from 326 to 1266 are selected from the Toyota & Stanford-MIT [123,183] cycle degradation dataset to conduct the experiments. Training dataset with seven batteries under different degradation conditions are utilized for model training, while other three batteries (denoted by B1, B2

and B3) with 45%, 65% and 85% cycle lives are chosen as the testing data. The SOH curves of training batteries and testing batteries are drawn in Fig. 8. From it, we see that different degradation protocols lead to different SOH degradation process versus cycle number.

For each charge–discharge cycle of each battery, the directly measured discharging capacity and temperature, and the calculated DV values are extracted as the input features, and their corresponding SOHs are considered as the labels. The input features and sampling time are fit by a spline function, which are linearly interpolated to a constant length of 1,000 for a uniformly size as the input of the training model.

#### B. Settings

For the model training settings, the seven training batteries are input into each SOH estimation model, and then the pre-trained model estimates the SOH of the remaining three testing batteries. The hyper-parameters like the learning rate and regularization constant are set to be uniform on each tested battery, and the maximum training iteration is set as 1000 uniformly for all tested models.

For the performance comparison of each model, the metrics mentioned in Section V. B are adopted to evaluate the SOH estimation accuracy, and the calculation time (including model training time and the estimation time for each testing battery) is adopted to measure the computation efficiency. Note that the model optimization time (i.e., hyper-parameters tuning time) is not involved in calculation time due to the different optimization process of each tested model. All experiments are conducted on the same bare machine with four Intel Xeon 2.1 GHz and 16-core CPU and 512 GB RAM.

#### C. Comparison results and analysis

Figs. 9–11 illustrate the SOH estimation results of each model on the testing batteries B1–3, respectively. Table 4 summarizes the SOH estimation performance of different models, including the RMSE, MAE, MAPE,  $R^2$ , and time cost on B1–3. From these results, we have the following discoveries:

**a) The degradation protocol is critical for the SOH estimation accuracy.** The selected batteries (including training and testing ones) have different degradation protocols due to the different charging/discharging operations. In the experiments, the involved SOH estimation models are trained on the same training batteries datasets. However, each SOH estimation model represents the different SOH estimation accuracy to the three testing batteries with various cycle lives. For instance, TL, CNN and RF achieve the best SOH estimation accuracy on the testing batteries B1, B2 and B3, respectively. Moreover, CNN achieves the best SOH estimation accuracy on B2 compared with LSTM, GPR, RF, and TL in all accuracy metrics, but has the worst accuracy on B1 and B3. Hence, the accuracy of an SOH estimation model depends on the degradation protocol of one testing battery.

**b) Different SOH estimation models differ from each other vastly in computational cost.** More specifically, the LSTM, CNN, GPR, RF, and TL take 3887.9 s, 492.4 s, 14500.2 s, 1.8 s, and

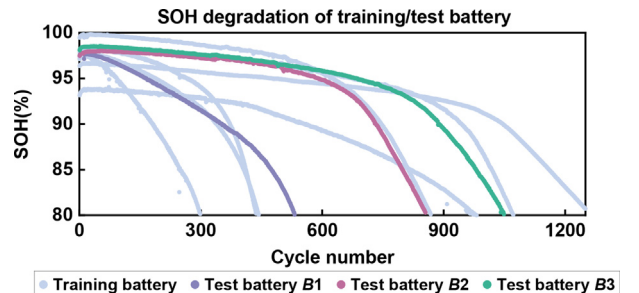


Fig. 8. SOH degradation of training/testing battery.



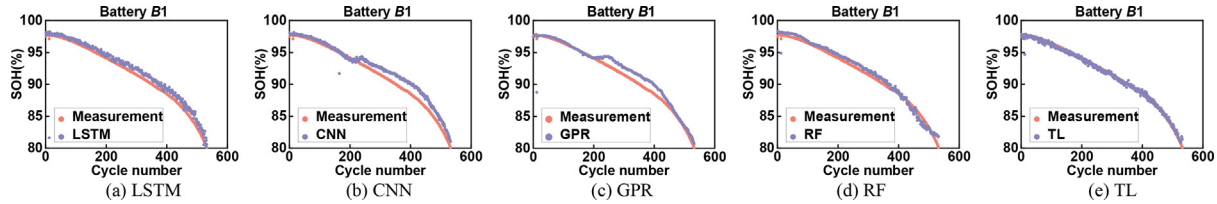


Fig. 9. SOH estimation results of testing battery B1.

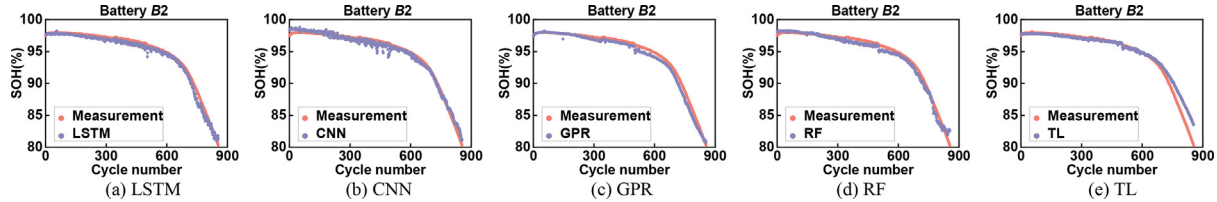


Fig. 10. SOH estimation results of testing battery B2.

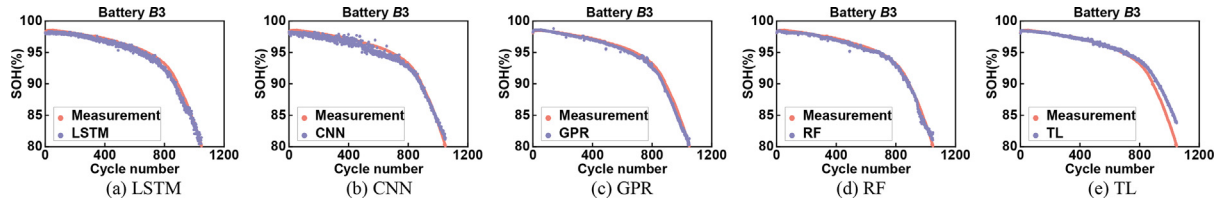


Fig. 11. SOH estimation results of testing battery B3.

Table 4

Performance Comparison of Data-driven SOH Estimation Models.

Testing Battery	Index	LSTM	CNN	GPR	RF	TL
B1	RMSE	0.01005	0.03418	0.00931	0.00594	<b>0.00322</b>
	MAE	0.00697	0.00982	0.00671	0.00525	<b>0.00243</b>
	MAPE	0.77295	1.08991	0.74347	0.57810	<b>0.26692</b>
	$R^2$	0.95315	0.45866	0.95982	0.98365	<b>0.99520</b>
B2	RMSE	0.00552	<b>0.00522</b>	0.00600	0.00566	0.00926
	MAE	0.00442	<b>0.00428</b>	0.00470	0.00465	0.00516
	MAPE	0.47571	<b>0.45734</b>	0.50799	0.50171	0.58527
	$R^2$	0.98394	<b>0.98562</b>	0.98101	0.98313	0.95479
B3	RMSE	0.00608	0.00637	0.00433	<b>0.00352</b>	0.01009
	MAE	0.00505	0.00559	0.00321	<b>0.00261</b>	0.00521
	MAPE	0.54252	0.59050	0.35107	<b>0.28207</b>	0.59762
	$R^2$	0.98148	0.97965	0.99057	<b>0.99378</b>	0.94892
Training Time (s)		3887.9	492.4	14500.2	1.8	120.6
Evaluation Time for B1 (s)		0.91	0.16	0.53	0.01	0.16
Evaluation Time for B2 (s)		0.25	0.12	0.86	0.02	0.11
Evaluation Time for B3 (s)		0.34	0.11	0.96	0.02	0.13

120.6 s for model training, respectively. According to the calculation time records, we find that RF has significantly higher computational efficiency than its peers. Furthermore, although GPR costs much more time than its peers do, it is non-parametric.

## 5. Conclusions

This paper carefully reviews the data-driven battery SOH estimation models published in 2018–2022. The data-driven battery SOH estimation models are categorized into NN, GPR, EL, TL, and CM. The basic concepts and the variants of the models are presented. Representative SOH estimation models published in 2020–2022 along with the available public datasets are summarized. Up to now, various data-driven battery SOH estimators are investigated and published. According to the characteristics of

published battery datasets and the corresponding data-driven SOH estimation models, we conclude that:

a) NASA, CACLE and Oxford datasets take few cells to implement the five-channels testing experiments. It is suitable for NN-based SOH estimation model.

b) Cambridge&Farada dataset only records “Q”, “V” and “I” of 12 cells in the testing experiments. It is a small-simple dataset and is suitable GPR-based SOH estimation model.

c) Sandia, Toyota&Stanford-MIT and HUST datasets include many cells in the testing experiments. It can be used to the TL-based SOH estimation researches.

Nonetheless, challenges remain in the following perspectives:

a) Existing battery degradation data are mostly collected from the laboratory. However, the battery degradation is a complex electrochemistry process, which may vary vastly from the labora-

tory to the real application conditions. Hence, battery SOH monitoring in real applications scenes is a critical task for facilitating reliable and robust battery SOH estimators.

b) Data-driven battery SOH estimation models rely on the quantity and quality of the acquired data. However, the computational burden increases rapidly when the data increase, especially for the complex models like NN-based ones. Efficient SOH estimators for real-time applications are highly desired.

c) Existing battery SOH estimators all rely on dense (or complete) data, while highly-incomplete battery SOH monitoring data are far more frequently encountered in real applications. How to implement accurate and robust battery SOH estimators with highly-incomplete data remains unveiled.

d) The spatio-temporal pattern analysis is required in BMS due to the spatio-temporal continuity of battery SOH. Existing studies partially addresses this issue via the intrinsic model structure like RNN [104,105], LSTM [72,132] or AR [123,177]. However, more comprehensive studies regarding this issue are of great needs.

### CRedit authorship contribution statement

**Minzhi Chen:** Investigation, Methodology, Software. **Guijun Ma:** Investigation, Methodology, Software. **Weibo Liu:** Investigation, Methodology, Software. **Nianyin Zeng:** Visualization, Conceptualization, Methodology, Writing – review & editing. **Xin Luo:** Visualization, Conceptualization, Methodology, Writing – review & editing.

### Data availability

The data sharing links are provided in the manuscript.

### Declaration of Competing Interest

This work was supported in part by the Natural Science Foundation of China under Grants 62073271 and 62272078, the Fundamental Research Funds for the Central Universities of China under Grant 20720220076, and the CAAI-Huawei MindSpore Open Fund under grants CAAIXSJLJ-2021-035A. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### References

- [1] U.S. Department of Energy, Energy Storage Grand Challenge: Energy Storage Market Report, 2020. <https://www.energy.gov/energy-storage-grand-challenge/downloads/energy-storage-market-report-2020>.
- [2] F. Yao, Y. Ding, S. Hong, S. Yang, A survey on evolved LoRa-based communication technologies for emerging internet of things applications, *Int. J. Network Dyn. Intelligence* 1 (1) (2022) 4–19.
- [3] X. Luo, M. Zhou, S. Li, L. Hu, M. Shang, Non-negativity constrained missing data estimation for high-dimensional and sparse matrices from industrial applications, *IEEE Trans. Cybernetics* 50 (5) (2020) 1844–1855.
- [4] United States Environmental Protection Agency, Fact Sheet | Energy Storage (2019) | White Papers | EESI, (n.d.). <https://www.eesi.org/papers/view/energy-storage-2019> (accessed March 9, 2021).
- [5] Z. Li, S. Li, X. Luo, An overview of calibration technology of industrial robots, *IEEE/CAA J. Autom. Sin.* 8 (1) (2021) 23–36.
- [6] Electropaedia, Battery Applications, (n.d.). <https://www.mpoweruk.com/applications.htm> (accessed March 9, 2021).
- [7] Y. Su, H. Cai, J. Huang, The cooperative output regulation by the distributed observer approach, *Int. J. Network Dyn. Intelligence* 1 (1) (2022) 20–35.
- [8] D.I. Stroe, M. Swierczynski, A.I. Stroe, R. Laerke, P.C. Kjaer, R. Teodorescu, Degradation behavior of lithium-ion batteries based on lifetime models and field measured frequency regulation mission profile, *IEEE Trans. Ind. Appl.* 52 (6) (2016) 5009–5018.
- [9] D.I. Stroe, V. Knap, M. Swierczynski, A.I. Stroe, R. Teodorescu, Operation of a grid-connected lithium-ion battery energy storage system for primary frequency regulation: A battery lifetime perspective, *IEEE Trans. Ind. Appl.* 53 (1) (2016) 430–438.
- [10] Y. Wang, J. Tian, Z. Sun, L. Wang, R. Xu, M. Li, and Z. Chen, A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems, *Renewable Sustainable Energy Rev.*, 131, no. 110015, 2020.
- [11] F. Shakiba, M. Shojaei, S. Azizi, M. Zhou, Real-time sensing and fault diagnosis for transmission lines, *Int. J. Network Dyn. Intelligence* 1 (1) (2022) 36–47.
- [12] X. Hu, F. Feng, K. Liu, L. Zhang, J. Xie, and B. Liu, "State estimation for advanced battery management: Key challenges and future trends," *Renewable and Sustainable Energy Reviews*, vol. 114, no. 109334, 2019.
- [13] M. Szankin, A. Kwasniewska, Can AI see bias in X-ray images?, *Int. J. Network Dyn. Intelligence* 1 (1) (2022) 48–64.
- [14] M.H. Lipu, M.A. Hannan, A. Hussain, M.M. Hoque, P.J. Ker, M.M. Saad, A. Ayob, A review of state of health and remaining useful life estimation methods for lithium-ion battery in electric vehicles: Challenges and recommendations, *J. Clean. Prod.* 205 (2018) 115–133.
- [15] H. Tian, P. Qin, K. Li, Z. Zhao, A review of the state of health for lithium-ion batteries: Research status and suggestions, *J. Clean. Prod.*, vol. 261, no. 120813, 2020.
- [16] J. Weddington, G. Niu, R. Chen, W. Yan, B. Zhang, Lithium-ion battery diagnostics and prognostics enhanced with Dempster-Shafer decision fusion, *Neurocomputing* 458 (2021) 440–453.
- [17] S. Jiang, Z. Song, A review on the state of health estimation methods of lead-acid batteries, *J. Power Sources*, vol. 517, no. 230710, 2022.
- [18] S. Yang, C. Zhang, J. Jiang, W. Zhang, L. Zhang, Y. Wang, Review on state-of-health of lithium-ion batteries: Characterizations, estimations and applications, *J. Clean. Product.*, vol. 314, no. 128015, 2021.
- [19] A. Farmann, W. Waag, A. Marongiu, D.U. Sauer, Critical review of on-board capacity estimation techniques for lithium-ion batteries in electric and hybrid electric vehicles, *J. Power Sources* 281 (2015) 114–130.
- [20] G. Zhao, Y. Li, Q. Xu, From emotion AI to cognitive AI, *Int. J. Network Dyn. Intelligence* 1 (1) (2022) 65–72.
- [21] M. Bercibar, I. Gandiaga, I. Villarreal, N. Omar, J. Van Mierlo, P. Van den Bossche, Critical review of state of health estimation methods of Li-ion batteries for real applications, *Renew. Sustain. Energy Rev.* 56 (2016) 572–587.
- [22] S.M. Rezvanianani, Z. Liu, Y. Chen, J. Lee, Review and recent advances in battery health monitoring and prognostics technologies for electric vehicle (EV) safety and mobility, *J. Power Sources* 256 (2014) 110–124.
- [23] R. Xiong, L. Li, J. Tian, Towards a smarter battery management system: A critical review on battery state of health monitoring methods, *J. Power Sources* 405 (2019) 18–29.
- [24] X. Sui, S. He, S. B. Vilsen, J. Meng, R. Teodorescu, D. I. Stroe, A review of non-probabilistic machine learning-based state of health estimation techniques for Lithium-ion battery, *Appl. Energy*, vol. 300, no. 117346, 2021.
- [25] Y. Li, K. Liu, A. M. Foley, A. Zülke, M. Bercibar, E. Nanini-Maury, H. E. Hoster, Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review, *Renewable Sustainable Energy Rev.*, vol. 113, no. 109254, 2019.
- [26] Z. Zhao, P.X. Liu, J. Gao, Model-based fault diagnosis methods for systems with stochastic process—a survey, *Neurocomputing* 513 (2022) 137–152.
- [27] L. Ling, Y. Wei, State-of-charge and state-of-health estimation for lithium-ion batteries based on dual fractional-order extended Kalman filter and online parameter identification, *IEEE Access* 9 (2021) 47588–47602.
- [28] Y. Yuan, X. Luo, M. Shang, Z. Wang, A Kalman-filter-incorporated latent factor analysis model for temporally dynamic sparse data, *IEEE Trans. on Cybernetics* (2022), <https://doi.org/10.1109/TCYB.2022.3185117>.
- [29] G. Dong, Z. Chen, J. Wei, Q. Ling, Battery health prognosis using Brownian motion modeling and particle filtering, *IEEE Trans. Ind. Electr.* 65 (11) (2018) 8646–8655.
- [30] B. Arachchige, S. Perinpanayagam, R. Jaras, Enhanced prognostic model for lithium ion batteries based on particle filter state transition model modification, *Appl. Sci.* 7 (11) (2017) 1172.
- [31] B. Wen, M. Xiao, G. Wang, Z. Yang, J. Li, X. Chen, A fusion prognostic method for remaining useful life prediction based on an extended belief rule base and particle filters, *IEEE Access* 9 (2021) 73377–73391.
- [32] S. Wang, S. Jin, D. Bai, Y. Fan, H. Shi, C. Fernandez, A critical review of improved deep learning methods for the remaining useful life prediction of lithium-ion batteries, *Energy Rep.* 7 (2021) 5562–5574.
- [33] T. Oji, Y. Zhou, S. Ci, F. Kang, X. Chen, X. Liu, Data-driven methods for battery soh estimation: Survey and a critical analysis, *IEEE Access* 9 (2021) 126903–126916.
- [34] E. Vanem, C. B. Salucci, A. Bakdi, Ø. Å. sheim Alnes, Data-driven state of health modelling—A review of state of the art and reflections on applications for maritime battery systems, *J. Energy Storage*, 43, no. 103158, 2021.
- [35] C. Vidal, P. Malysz, P. Kollmeyer, A. Emadi, Machine learning applied to electrified vehicle battery state of charge and state of health estimation: state-of-the-art, *IEEE Access* 8 (2020) 52796–52814.
- [36] L. Chen, Y. Zhang, Y. Zheng, X. Li, X. Zheng, Remaining useful life prediction of lithium-ion battery with optimal input sequence selection and error compensation, *Neurocomputing* 414 (2020) 245–254.
- [37] X. Hu, Y. Che, X. Lin, S. Onori, Battery health prediction using fusion-based feature selection and machine learning, *IEEE Trans. Transport. Electr.* 7 (2) (2020) 382–398.
- [38] X. Li, S. Feng, N. Hou, R. Wang, H. Li, M. Gao, S. Li, Surface microseismic data denoising based on sparse autoencoder and Kalman filter, *Syst. Sci. Control Eng.* 10 (1) (2022) 616–628.

- [39] M. Liang, W. Liu, Y. Wen, H. Yang, Segmentation and weight prediction of grape ear based on SFNet-ResNet18, *Syst. Sci. Control Eng.* 10 (1) (2022) 722–732.
- [40] D. Yang, J. Lu, H. Dong, Z. Hu, Pipeline signal feature extraction method based on multi-feature entropy fusion and local linear embedding, *Syst. Sci. Control Eng.* 10 (1) (2022) 407–416.
- [41] G. Bao, L. Ma, X. Yi, Recent advances on cooperative control of heterogeneous multi-agent systems subject to constraints: A survey, *Syst. Sci. Control Eng.* 10 (1) (2022) 539–551.
- [42] H. Tao, H. Tan, Q. Chen, H. Liu, J. Hu,  $H_\infty$  state estimation for memristive neural networks with randomly occurring DoS attacks, *Syst. Sci. Control Eng.* 10 (1) (2022) 154–165.
- [43] F. Qu, X. Zhao, X. Wang, E. Tian, Probabilistic-constrained distributed fusion filtering for a class of time-varying systems over sensor networks: a torus-event-triggering mechanism, *Int. J. Syst. Sci.* 53 (6) (2022) 1288–1297.
- [44] W. An, P. Zhao, H. Liu, and J. Hu, Distributed multi-step subgradient projection algorithm with adaptive event-triggering protocols: a framework of multiagent systems, *Int. J. Syst. Sci.*, in press, DOI: 10.1080/00207721.2022.2063967.
- [45] X. Li, Q. Song, Y. Liu, F. Alsaadi, Nash equilibrium and bang-bang property for the non-zero-sum differential game of multi-player uncertain systems with Hurwicz criterion, *Int. J. Syst. Sci.* 53 (10) (2022) 2207–2218.
- [46] H. Yu, J. Hu, B. Song, H. Liu, X. Yi, Resilient energy-to-peak filtering for linear parameter-varying systems under random access protocol, *Int. J. Syst. Sci.* 53 (11) (2022) 2421–2436.
- [47] A.K. Jain, J. Mao, K.M. Mohiuddin, Artificial neural networks: A tutorial, *Computer* 29 (3) (1996) 31–44.
- [48] M.H. Hassoun, *Fundamentals of Artificial Neural Networks*, MIT Press, 1995.
- [49] Z. Wang, N. Liu, Y. Guo, Adaptive sliding window LSTM NN based RUL prediction for lithium-ion batteries integrating LTSA feature reconstruction, *Neurocomputing* 466 (2021) 178–189.
- [50] J. Feng, X. Jia, H. Cai, F. Zhu, X. Li, J. Lee, Cross trajectory gaussian process regression model for battery health prediction, *J. Mod. Power Syst. Clean Energy* 9 (5) (2020) 1217–1226.
- [51] M. Lucu, E. Martinez-Laserna, I. Gandiaga, K. Liu, H. Camblong, W.D. Widanage, J. Marco, Data-driven nonparametric Li-ion battery degradation model aiming at learning from real operation data-Part A: Storage operation, *J. Storage Mater.* 30 (101409) (2020) pp.
- [52] K. Liu, X. Hu, Z. Wei, Y. Li, Y. Jiang, Modified Gaussian process regression models for cyclic capacity prediction of lithium-ion batteries, *IEEE Trans. Transport. Electr.* 5 (4) (2019) 1225–1236.
- [53] Z. Wang, J. Ma, L. Zhang, State-of-health estimation for lithium-ion batteries based on the multi-island genetic algorithm and the gaussian process regression, *IEEE Access* 5 (2017) 21286–21295.
- [54] D. Roman, S. Saxena, V. Robu, M. Pecht, D. Flynn, Machine learning pipeline for battery state-of-health estimation, *Nat. Machine Intelligence* 3 (5) (2021) 447–456.
- [55] C. She, Y. Li, C. Zou, T. Wik, Z. Wang, F. Sun, Offline and Online Blended Machine Learning for lithium-ion battery health state estimation, *IEEE Trans. Transport. Electr.* 8 (2) (2022) 1604–1618.
- [56] H. Meng, Y.F. Li, A review on prognostics and health management (PHM) methods of lithium-ion batteries, *Renew. Sustain. Energy Rev.* 116 (109405) (2019) pp.
- [57] M.A. Hoque, P. Nurmi, A. Kumar, S. Varjonen, J. Song, M.G. Pecht, S. Tarkoma, Data driven analysis of lithium-ion battery internal resistance towards reliable state of health prediction, *J. Power Sources* 513 (230519) (2021) pp.
- [58] H. Ji, W. Zhang, X. Pan, M. Hua, Y. Chung, C. Shu, L.J. Zhang, State of health prediction model based on internal resistance, *Int. J. Energy Res.* 44 (8) (2020) 6502–6510.
- [59] X. Tang, C. Zou, K. Yao, J. Lu, Y. Xia, F. Gao, Degradation trajectory prediction for lithium-ion batteries via model migration and Bayesian Monte Carlo method, *Appl. Energy* 254 (113591) (2019) pp.
- [60] H. Rauf, M. Khalid, N. Arshad, Machine learning in state of health and remaining useful life estimation: Theoretical and technological development in battery degradation modelling, *Renew. Sustain. Energy Rev.* 156 (111903) (2022) pp.
- [61] J. Wang, S. Liu, S. Wang, Q. Liu, H. Liu, H. Zhou, J. Tang, Multiple indicators-based health diagnostics and prognostics for energy storage technologies using fuzzy comprehensive evaluation and improved multivariate grey model, *IEEE Trans. on Power Electronics* 36 (11) (2021) 12309–12320.
- [62] N. Yu, R. Yang, M. Huang, Deep common spatial pattern based motor imagery classification with improved objective function, *Int. J. Network Dyn. Intelligence* 1 (1) (2022) 73–84.
- [63] D. Wu, X. Luo, Y. He, M. Zhou, A prediction-sampling-based multilayer-structured latent factor model for accurate representation to high-dimensional and sparse data, *IEEE Trans. Neural Networks Learn. Syst.* (2022), <https://doi.org/10.1109/TNNLS.2022.3200009>.
- [64] X. Luo, M. Chen, H. Wu, Z. Liu, H. Yuan, M. Zhou, Adjusting learning depth in nonnegative latent factorization of tensors for accurately modeling temporal patterns in dynamic QoS data, *IEEE Trans. Autom. Sci. Eng.* 18 (4) (2021) 2142–2155.
- [65] D. Wu, X. Luo, M. Shang, Y. He, G. Wang, M. Zhou, A deep latent factor model for high-dimensional and sparse matrices in recommender systems, *IEEE Trans. Systems, Man, and Cybernetics: Syst.* 51 (7) (2021) 4285–4296.
- [66] J. Yang, B. Xia, W. Huang, Y. Fu, C. Mi, Online state-of-health estimation for lithium-ion batteries using constant-voltage charging current analysis, *Appl. Energy* 212 (2018) 1589–1600.
- [67] Y. Deng, H. Ying, E. Jiaqiang, H. Zhu, K. Wei, J. Chen, G. Liao, Feature parameter extraction and intelligent estimation of the state-of-health of lithium-ion batteries, *Energy* 176 (2019) 91–102.
- [68] R. Xiong, Y. Zhang, J. Wang, H. He, S. Peng, M. Pecht, Lithium-ion battery health prognosis based on a real battery management system used in electric vehicles, *IEEE Trans. Vehicular Technol.* 68 (5) (2019) 4110–4121.
- [69] D. Wu, X. Luo, G. Wang, M. Shang, Y. Yuan, H. Yan, A highly accurate framework for self-labeled semisupervised classification in industrial applications, *IEEE Trans. Ind. Inform.* 14 (3) (2018) 909–920.
- [70] C. Weng, X. Feng, J. Sun, H. Peng, State-of-health monitoring of lithium-ion battery modules and packs via incremental capacity peak tracking, *Appl. Energy* 180 (2016) 360–368.
- [71] D. Anseán, V.M. García, M. González, C. Blanco-Viejo, J.C. Viera, Y.F. Pulido, Y. Sánchez, Lithium-ion battery degradation indicators via incremental capacity analysis, *IEEE Trans. Ind. Appl.* 55 (3) (2019) 2992–3002.
- [72] X. Cui, T. Hu, State of health diagnosis and remaining useful life prediction for lithium-ion battery based on data model fusion method, *IEEE Access* 8 (2020) 207298–207307.
- [73] T. Goh, M. Park, M. Seo, J.G. Kim, S.W. Kim, Capacity estimation algorithm with a second-order differential voltage curve for li-ion batteries with NMC cathodes, *Energy* 135 (2017) 257–268.
- [74] T. Shibagaki, Y. Merla, G.J. Offer, Tracking degradation in lithium iron phosphate batteries using differential thermal voltammetry, *J. Power Sources* 374 (2018) 188–195.
- [75] Y. Merla, B. Wu, V. Yufit, N.P. Brandon, R.F. Martinez-Botas, G.J. Offer, Novel application of differential thermal voltammetry as an in-depth state-of-health diagnosis method for lithium-ion batteries, *J. Power Sources* 307 (2016) 308–319.
- [76] R. Li, J. Hong, H. Zhang, X. Chen, Data-driven battery state of health estimation based on interval capacity for real-world electric vehicles, *Energy* 257 (124771) (2022).
- [77] M. Bercibar, M. Garmendia, I. Gandiaga, J. Crego, I. Villarreal, State of health estimation algorithm of LiFePO<sub>4</sub> battery packs based on differential voltage curves for battery management system application, *Energy* 103 (2016) 784–796.
- [78] L. Wang, C. Pan, L. Liu, Y. Cheng, X. Zhao, On-board state of health estimation of LiFePO<sub>4</sub> battery pack through differential voltage analysis, *Appl. Energy* 168 (2016) 465–472.
- [79] G. Cheng, X. Wang, Y. He, Remaining useful life and state of health prediction for lithium batteries based on empirical mode decomposition and a long and short memory neural network, *Energy* 232 (121022) (2021) pp.
- [80] J. Qu, F. Liu, Y. Ma, J. Fan, A neural-network-based method for rul prediction and soh monitoring of lithium-ion battery, *IEEE Access* 7 (2019) 87178–87191.
- [81] J. Chen, T. Chen, W. Liu, C. Cheng, M. Li, Combining empirical mode decomposition and deep recurrent neural networks for predictive maintenance of lithium-ion battery, *Adv. Eng. Inf.* 50 (101405) (2021) pp.
- [82] Z. Chen, Q. Xue, R. Xiao, Y. Liu, J. Shen, State of health estimation for lithium-ion batteries based on fusion of autoregressive moving average model and elman neural network, *IEEE Access* 7 (2019) 102662–102678.
- [83] Y. Li, M. Abdel-Monem, R. Gopalakrishnan, M. Bercibar, E. Nanini-Maury, N. Omar, J. Van Mierlo, A quick on-line state of health estimation method for Li-ion battery with incremental capacity curves processed by Gaussian filter, *J. Power Sources* 373 (2018) 40–53.
- [84] X. Li, Z. Wang, J. Yan, Prognostic health condition for lithium battery using the partial incremental capacity and Gaussian process regression, *J. Power Sources* 421 (2019) 56–67.
- [85] X. Li, Z. Wang, L. Zhang, C. Zou, D.D. Dorrell, State-of-health estimation for li-ion batteries by combining the incremental capacity analysis method with grey relational analysis, *J. Power Sources* 410–411 (2019) 106–114.
- [86] Y. Huang, Y. Tang, J. VanZwieten, Prognostics with variational autoencoder by generative adversarial learning, *IEEE Trans. Ind. Electr.* 69 (1) (2022) 856–867.
- [87] X. Luo, H. Wu, Z. Li, NeuLFT: A novel approach to nonlinear canonical polyadic decomposition on high-dimensional incomplete tensors, *IEEE Trans. Knowledge Data Eng.* (2022), <https://doi.org/10.1109/TKDE.2022.3176466>.
- [88] D. Wu, Y. He, X. Luo, M. Zhou, A latent factor analysis-based approach to online sparse streaming feature selection, *IEEE Trans. Systems, Man, and Cybernetics, Syst.* 52 (11) (2022) 6744–6758.
- [89] L. Zhao, Y. Wang, J. Cheng, A hybrid method for remaining useful life estimation of lithium-ion battery with regeneration phenomena, *Appl. Sci.* 9 (9) (2019) 1890.
- [90] Z. Chen, H. Zhao, Y. Zhang, S. Shen, J. Shen, Y. Liu, State of health estimation for lithium-ion batteries based on temperature prediction and gated recurrent unit neural network, *J. Power Sources* 521 (230892) (2022) pp.
- [91] Y. Jiang, J. Zhang, L. Xia, Y. Liu, State of health estimation for lithium-ion battery using empirical degradation and error compensation models, *IEEE Access* 8 (2020) 123858–123868.
- [92] B. Gou, Y. Xu, X. Feng, State-of-health estimation and remaining-useful-life prediction for lithium-ion battery using a hybrid data-driven method, *IEEE Trans. Vehicular Technol.* 69 (10) (2020) 10854–10867.



- [93] H. Wang, J. Li, X. Liu, J. Rao, Y. Fan, X. Tan, Online state of health estimation for lithium-ion batteries based on a dual self-attention multivariate time series prediction network, *Energy Rep.* 8 (2022) 8953–8964.
- [94] Y. Wang, P. Kou, J. Fan, N. Zhou, A novel capacity estimation method for li-ion battery cell by applying ensemble learning to extremely sparse significant points, *IEEE Access* 10 (2022) 96427–96441.
- [95] X. Luo, M. Zhou, S. Li, M. Shang, An Inherently nonnegative latent factor model for high-dimensional and sparse matrices from industrial applications, *IEEE Trans. Ind. Inf.* 14 (5) (2018) 2011–2022.
- [96] M. Wang, H. Wang, H. Zheng, A mini review of node centrality metrics in biological networks, *Int. J. Network Dyn. Intelligence* 1 (1) (2022) 99–110.
- [97] L. Jin, X. Zheng, X. Luo, Neural dynamics for distributed collaborative control of manipulators with time delays, *IEEE/CAA J. Autom. Sin.* 9 (5) (2022) 854–863.
- [98] M. Cao, T. Zhang, B. Yu, Y. Liu, A method for interval prediction of satellite battery state of health based on sample entropy, *IEEE Access* 7 (2019) 141549–141561.
- [99] T. Sun, S. Wang, S. Jiang, B. Xu, X. Han, X. Lai, Y. Zheng, A cloud-edge collaborative strategy for capacity prognostic of lithium-ion batteries based on dynamic weight allocation and machine learning, *Energy* 239 (122185) (2022) pp.
- [100] J. Wu, C. Zhang, Z. Chen, An online method for lithium-ion battery remaining useful life estimation using importance sampling and neural networks, *Appl. Energy* 173 (2016) 134–140.
- [101] X. Liang, N. Bao, J. Zhang, A. Garg, S. Wang, Evaluation of battery modules state for electric vehicle using artificial neural network and experimental validation, *Energy Sci. Eng.* 6 (5) (2018) 397–407.
- [102] T. Zhang, D. Zhang, H. Yan, J. Qiu, J. Gao, A new method of data missing estimation with FNN-based tensor heterogeneous ensemble learning for internet of vehicle, *Neurocomputing* 420 (2021) 98–110.
- [103] X. Wang, Y. Sun, D. Ding, Adaptive dynamic programming for networked control systems under communication constraints: a survey of trends and techniques, *Int. J. Network Dyn. Intelligence* 1 (1) (2022) 85–98.
- [104] L. Ren, J. Dong, X. Wang, Z. Meng, L. Zhao, M.J. Deen, A data-driven Auto-CMM-LSTM prediction model for lithium-ion battery remaining useful life, *IEEE Trans. Ind. Inf.* 17 (5) (2021) 3478–3487.
- [105] S. Khaleghi, M.S. Hosen, D. Karimi, H. Behi, S.H. Beheshti, J. Van Mierlo, M. Bercibar, Developing an online data-driven approach for prognostics and health management of lithium-ion batteries, *Appl. Energy* 308 (118348) (2022) pp.
- [106] J. Tian, R. Xiong, W. Shen, J. Lu, X.G. Yang, Deep neural network battery charging curve prediction using 30 points collected in 10 min, *Joule* 5 (6) (2021) 1521–1534.
- [107] E. Chemali, P.J. Kollmeyer, M. Preindl, Y. Fahmy, A. Emadi, Convolutional neural network approach for estimation of li-ion battery state of health from charge profiles, *Energies* 15 (3) (2022) 1185.
- [108] H. Dai, G. Zhao, M. Lin, J. Wu, G. Zheng, A novel estimation method for the state of health of lithium-ion battery using prior knowledge-based neural network and markov chain, *IEEE Trans. Ind. Electr.* 66 (10) (2019) 7706–7716.
- [109] P. Li, Z. Zhang, R. Grosu, Z. Deng, J. Hou, Y. Rong, R. Wu, An end-to-end neural network framework for state-of-health estimation and remaining useful life prediction of electric vehicle lithium batteries, *Renew. Sustain. Energy Rev.* 156 (111843) (2022) pp.
- [110] P. Li, Z. Zhang, Q. Xiong, B. Ding, J. Hou, D. Luo, S. Li, State-of-health estimation and remaining useful life prediction for the lithium-ion battery based on a variant long short term memory neural network, *J. Power Sources* 459 (228069) (2020) pp.
- [111] E. Pinto, E. Nepomuceno, A. Campanharo, Individual-based modelling of animal brucellosis spread with the use of complex networks, *Int. J. Network Dynamics Intelligence* 1 (1) (2022) 120–129.
- [112] M. Chen, C. He, X. Luo, MNL: a highly-efficient model for large-scale dynamic weighted directed network representation, *IEEE Trans. Big Data* (2022), <https://doi.org/10.1109/TBDATA.2022.3218064>.
- [113] L. Hu, S. Yang, X. Luo, M. Zhou, An algorithm of inductively identifying clusters from attributed graphs, *IEEE Trans. Big Data* 8 (2) (2020) 523–534.
- [114] L. Hu, X. Pan, Z. Tang, X. Luo, A fast fuzzy clustering algorithm for complex networks via a generalized momentum method, *IEEE Trans. Fuzzy Systems* 30 (9) (2021) 3473–3485.
- [115] M.S. Hossain Lipu, M.A. Hannan, A. Hussain, M.H.M. Saad, Optimal BP neural network algorithm for state of charge estimation of lithium-ion battery using PSO with PCA feature selection, *J. Renewable Sustainable Energy* 9 (6) (2017).
- [116] X. Shu, G. Li, J. Shen, Z. Lei, Z. Chen, Y. Liu, A uniform estimation framework for state of health of lithium-ion batteries considering feature extraction and parameters optimization, *Energy* 204 (117957) (2020) pp.
- [117] C. Liu, Y. Wang, Z. Chen, Degradation model and cycle life prediction for lithium-ion battery used in hybrid energy storage system, *Energy* 166 (2019) 796–806.
- [118] L. Li, Z. Liu, M. Tseng, A. Chiu, Enhancing the lithium-ion battery life predictability using a hybrid method, *Appl. Soft Comput.* 74 (2019) 110–121.
- [119] J. Chen, X. Luo, M. Zhou, Hierarchical Particle Swarm Optimization-incorporated Latent Factor Analysis for Large-Scale Incomplete Matrices, *IEEE Trans. on Big Data* 8 (6) (2022) 1524–1536.
- [120] F. Han, Q.H. Ling, D.S. Huang, An improved approximation approach incorporating particle swarm optimization and a priori information into neural networks, *Neural Comput. & Applic.* 19 (2010) 255–261.
- [121] S. Ferrari, M. Jensenius, A constrained optimization approach to preserving prior knowledge during incremental training, *IEEE Trans. Neural Network.* 19 (6) (2008) 996–1009.
- [122] Y. Pao, G.H. Park, D.J. Sobajic, Learning and generalization characteristics of the random vector functional-link net, *Neurocomputing* 6 (2) (1994) 163–180.
- [123] K.A. Severson, P.M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, R.D. Braatz, Data-driven prediction of battery cycle life before capacity degradation, *Nat. Energy* 4 (5) (2019) 383–391.
- [124] X. Shu, G. Li, Y. Zhang, J. Shen, Z. Chen, Y. Liu, Online diagnosis of state of health for lithium-ion batteries based on short-term charging profiles, *J. Power Sources* 471 (228478) (2020) pp.
- [125] M. Lin, X. Zeng, J. Wu, State of health estimation of lithium-ion battery based on an adaptive tunable hybrid radial basis function network, *J. Power Sources* 504 (230063) (2021) pp.
- [126] Y. Yuan, Q. He, X. Luo, M. Shang, A multilayered-and-randomized latent factor model for high-dimensional and sparse matrices, *IEEE Trans. On Big Data* 8 (3) (2020) 784–794.
- [127] H. Wu, X. Luo, M. Zhou, Advancing non-negative latent factorization of tensors with diversified regularizations, *IEEE Trans. Services Comp.* 15 (3) (2022) 1334–1341.
- [128] X. Sun, S. Kim, J. Choi, Recurrent neural network-induced Gaussian process, *Neurocomputing* 509 (2022) 75–84.
- [129] X. Han, Z. Wang, Z. Wei, A novel approach for health management online-monitoring of lithium-ion batteries based on model-data fusion, *Appl. Energy* 302 (117511) (2021) pp.
- [130] G. Dong, W. Han, Y. Wang, Dynamic bayesian network-based lithium-ion battery health prognosis for electric vehicles, *IEEE Trans. Ind. Electr.* 68 (11) (2021) 10949–10958.
- [131] M. Shang, Y. Yuan, X. Luo, M. Zhou, An  $\alpha$ - $\beta$ -divergence-generalized recommender for highly accurate predictions of missing user preferences, *IEEE Trans. Cybernetics* 52 (8) (2021) 8006–8018.
- [132] W. Zhang, X. Li, X. Li, Deep learning-based prognostic approach for lithium-ion batteries with adaptive time-series prediction and on-line validation, *Measurement* 164 (108052) (2020) pp.
- [133] W. Li, N. Sengupta, P. Dechent, D. Howey, A. Annaswamy, D.U. Sauer, One-shot battery degradation trajectory prediction with deep learning, *J. Power Sources* 506 (230024) (2021) pp.
- [134] Y. Tan, G. Zhao, Transfer learning with long short-term memory network for state-of-health prediction of lithium-ion batteries, *IEEE Trans. Ind. Electronics* 67 (10) (2019) 8723–8731.
- [135] X. Shu, J. Shen, G. Li, Y. Zhang, Z. Chen, Y. Liu, A flexible state-of-health prediction scheme for lithium-ion battery packs with long short-term memory network and transfer learning, *IEEE Trans. Transport. Electrification* 7 (4) (2021) 2238–2248.
- [136] L. Ungurean, M.V. Micea, G. Carstoiu, Online state of health prediction method for lithium-ion batteries, based on gated recurrent unit neural networks, *Int. J. Energy Res.* 44 (8) (2020) 6767–6777.
- [137] D. Zhou, Z. Li, J. Zhu, H. Zhang, L. Hou, State of health monitoring and remaining useful life prediction of lithium-ion batteries based on temporal convolutional network, *IEEE Access* 8 (2020) 53307–53320.
- [138] P. Ding, X. Liu, H. Li, Z. Huang, K. Zhang, L. Shao, O. Abedinia, Useful life prediction based on wavelet packet decomposition and two-dimensional convolutional neural network for lithium-ion batteries, *Renew. Sustain. Energy Rev.* 148 (111287) (2021) pp.
- [139] G. Ma, Y. Zhang, C. Cheng, B. Zhou, P. Hu, Y. Yuan, Remaining useful life prediction of lithium-ion batteries based on false nearest neighbors and a hybrid neural network, *Appl. Energy* 253 (113626) (2019) pp.
- [140] L. Sanchez, J. Otero, D. Ansean, I. Couso, Health assessment of LFP automotive batteries using a fractional-order neural network, *Neurocomputing* 391 (2020) 345–354.
- [141] Z. Lyu, G. Wang, R. Gao, Li-ion battery prognostic and health management through an indirect hybrid model, *J. Storage Mater.* 42 (102990) (2021) pp.
- [142] A. Aitio, D.A. Howey, Predicting battery end of life from solar off-grid system field data using machine learning, *Joule* 5 (12) (2021) 3204–3220.
- [143] Q. Zhang, Y. Zhou, Recent advances in non-Gaussian stochastic systems control theory and its applications, *Int. J. Network Dyn. Intelligence* 1 (1) (2022) 111–119.
- [144] X. Zheng, X. Deng, State-of-health prediction for lithium-ion batteries with multiple gaussian process regression model, *IEEE Access* 7 (2019) 150383–150394.
- [145] M. Lin, D. Wu, J. Meng, J. Wu, H. Wu, A multi-feature-based multi-model fusion method for state of health estimation of lithium-ion batteries, *J. Power Sources* 518 (230774) (2022) pp.
- [146] X. Hu, Y. Che, X. Lin, Z. Deng, Health prognosis for electric vehicle battery packs: a data-driven approach, *IEEE/ASME Trans. Mechatronics* 25 (6) (2020) 2622–2632.
- [147] Y. Zhang, Q. Tang, Y. Zhang, J. Wang, U. Stimming, A.A. Lee, Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning, *Nat. Commun.* 11 (1) (2020) 1–6, <https://doi.org/10.5281/zenodo.3633835>.
- [148] K. Liu, Y. Li, X. Hu, M. Lucu, W.D. Widanage, Gaussian process regression with automatic relevance determination kernel for calendar degradation prediction of lithium-ion batteries, *IEEE Trans. on Industrial Informatics* 16 (6) (2020) 3767–3777.

- [149] P. Hu, G. Ma, Y. Zhang, C. Cheng, B. Zhou, and Y. Yuan, "State of health estimation for lithium-ion batteries with dynamic time warping and deep kernel learning model," in *IEEE 2020 European Control Conference (ECC)*, pp: 602–607, 2020, Petersburg, Russia..
- [150] Z. Deng, X. Hu, P. Li, X. Lin, X. Bian, Data-driven battery state of health estimation based on random partial charging data, *IEEE Trans. Power Electronics* 37 (5) (2022) 5021–5031.
- [151] Z. Deng, X. Hu, Y. Xie, L. Xu, P. Li, X. Lin, X. Bian, Battery health evaluation using a short random segment of constant current charging, *Iscience* 5 (5) (2020).
- [152] O. Sagi, L. Rokach, Ensemble learning: A survey, *Wiley Interdisciplinary Rev. Data Min. Knowl. Disc.* 8 (4) (2018) e1249.
- [153] G. Ngo, R. Beard, R. Chandra, Evolutionary bagging for ensemble learning, *Neurocomputing* 510 (2022) 1–14.
- [154] K. Xu, H. Yang, C. Zhu, A novel extreme learning machine-based Hammerstein-Wiener model for complex nonlinear industrial processes, *Neurocomputing* 358 (2019) 246–254.
- [155] X. Luo, Z. Liu, M. Shang, J. Lou, M. Zhou, Highly-accurate community detection via pointwise mutual information-incorporated symmetric non-negative matrix factorization, *IEEE Trans. Network Sci. Eng.* 8 (1) (2020) 463–476.
- [156] C. W. Hsu, R. Xiong, N. Chen, J. Li, and N. Tsou, Deep neural network battery life and voltage prediction by using data of one cycle only, *Appl. Energy*, vol. 306, no. 118134, 2022.
- [157] S. Song, C. Fei, H. Xia, Lithium-ion battery SOH estimation based on XGBoost algorithm with accuracy correction, *Energies* 13 (4) (2020) 812.
- [158] S. Jafari, Z. Shahbazi, Y.C. Byun, Lithium-ion battery health prediction on hybrid vehicles using machine learning approach, *Energies* 15 (13) (2022) 4753.
- [159] Y. Li, S. Zhong, Q. Zhong, K. Shi, Lithium-ion battery state of health monitoring based on ensemble learning, *IEEE Access* 7 (2019) 8754–8762.
- [160] F. Yang, D. Wang, F. Xu, Z. Huang, and K. L. Tsui, Life span prediction of lithium-ion batteries based on various extracted features and gradient boosting regression tree model, *J. Power Sources*, vol. 476, no. 228654, 2020.
- [161] S. Khaleghi, Y. Firouz, M. Berecibar, J.V. Mierlo, P.V.D. Bossche, Ensemble gradient boosted tree for SoH estimation based on diagnostic features, *Energies* vol. 13, no. 5 (2020) pp1262.
- [162] M. Huotari, S. Arora, A. Malhi, and K. Främling, Comparing seven methods for state-of-health time series prediction for the lithium-ion battery packs of forklifts, *Appl. Soft Comp.*, vol. 111, no. 107670, 2021.
- [163] H. Liu, Q. Xiao, Z. Jiao, J. Meng, Y. Mu, K. Hou, and H. Jia, LightGBM-based prediction of remaining useful life for electric vehicle battery under driving conditions, in *2020 IEEE Sustainable Power and Energy Conference (ISPEC)*, pp. 2577–2582, 2020, Chengdu, China.
- [164] K. S. Mawonou, A. Eddahech, D. Dumur, D. Beauvois, and E. Godoy, State-of-health estimators coupled to a random forest approach for lithium-ion battery degradation factor ranking, *J. Power Sources*, vol. 484, no. 229154, 2021.
- [165] M. Shang, X. Luo, Z. Liu, J. Chen, Y. Yuan, M. Zhou, Randomized latent factor model for high-dimensional and sparse matrices from industrial applications, *IEEE/CAA J. Automatica Sinica* 6 (1) (2018) 131–141.
- [166] K. Weiss, T.M. Khoshgoftaar, D. Wang, A survey of transfer learning, *J. Big Data* 3 (1) (2016) 1–40.
- [167] C. Li, S. Zhang, Y. Qin, E. Estupinan, A systematic review of deep transfer learning for machinery fault diagnosis, *Neurocomputing* 407 (2020) 121–135.
- [168] M. Long, Y. Cao, J. Wang, M. and Jordan, Learning transferable features with deep adaptation networks, *In Int. Conf. on Machine Learning*, vol. 37, pp. 97–105, 2015.
- [169] G. Wilson, D.J. Cook, A survey of unsupervised deep domain adaptation, *ACM Trans. Intelligent Syst. Technol.* 11 (5) (2020) 1–46.
- [170] W. Mao, J. He, M.J. Zuo, Predicting remaining useful life of rolling bearings based on deep feature representation and transfer learning, *IEEE Trans. Instrum. Measure.* 69 (4) (2020) 1594–1608.
- [171] G. Ma, S. Xu, T. Yang, Z. Du, L. Zhu, H. Ding, Y. Yuan, A transfer learning-based method for personalized state of health estimation of lithium-ion batteries, *IEEE Trans. on Neural Networks and Learning Systems*, DIO. 10.1109/TNNLS.2022.3176925.
- [172] C. Cheng, B. Zhou, G. Ma, D. Wu, Y. Yuan, Wasserstein distance based deep adversarial transfer learning for intelligent fault diagnosis with unlabeled or insufficient labeled data, *Neurocomputing* 409 (2020) 35–45.
- [173] Y. Li, H. Sheng, Y. Cheng, D. I. Stroe, and R. Teodorescu, State-of-health estimation of lithium-ion batteries based on semi-supervised transfer component analysis, *Appl. Energy*, vol. 277, no. 115504, 2020.
- [174] J. Wei, G. Dong, Z. Chen, Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression, *IEEE Trans. Ind. Electronics* 65 (7) (2018) 5634–5643.
- [175] X. Li, C. Yuan, and Z. Wang, State of health estimation for Li-ion battery via partial incremental capacity analysis based on support vector regression, *Energy*, vol. 203, no. 117852, 2020.
- [176] Z. Xue, Y. Zhang, C. Cheng, G. Ma, Remaining useful life prediction of lithium-ion batteries with adaptive unscented kalman filter and optimized support vector regression, *Neurocomputing* 376 (2020) 95–102.
- [177] C. Lin, J. Cabrera, F. Yang, M. Ling, K. L. Tsui, and S. J. Bae, Battery state of health modeling and remaining useful life prediction through time series model, *Appl. Energy*, vol. 275, no. 115338, 2020.
- [178] Prognostics center of excellence - data repository. NASA Ames Progn Res Center. 2007 <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository>.
- [179] CALCE battery research group homepage. 2011. <https://web.calce.umd.edu/batteries/data.htm>.
- [180] Oxford battery team data and code. 2011. <http://howey.eng.ox.ac.uk/data-and-code/>.
- [181] Sandia National Lab. Data for degradation of commercial lithium-ion cells as a function of chemistry and cycling conditions. 2020. [https://www.batteryarchive.org/snL\\_study.html](https://www.batteryarchive.org/snL_study.html).
- [182] H. Barkholtz, A. Fresquez, B. R. Chalamala, and S. R. Ferreira, "A database for comparative electrochemical performance of commercial 18650-format lithium-ion cells," *Journal of The Electrochemical Society*, vol. 164, no.12, 2017.
- [183] Toyota Research Institute. Experimental data platform. 2021. <https://data.matr.io/1/>.
- [184] G. Ma, X. Xu, B. Jiang, C. Cheng, X. Yang, Y. Shen, and Y. Yuan, Real-time personalized health status prediction of lithium-ion batteries using deep transfer learning, *Energy Environ. Sci.*, DOI: 10.1039/D2EE01676A, 2022. <https://doi.org/10.17632/nsc7hnsq4s.2>.
- [185] X. Xu, C. Yu, S. Tang, X. Sun, X. Si, L. Wu, State-of-health estimation for lithium-ion batteries based on wiener process with modeling the relaxation effect, *IEEE Access* 7 (2019) 105186–105201.