Computers and Artificial Intelligence; Vol.3, No.6; 2026
E-ISSN: 3006-3728 P-ISSN: 3006-371X
DOI: https://doi.org/10.70267/cai.v3n6.7892
Published by: Zeus Press

A Review of State-of-Health Estimation and Prediction Methods for Lithium-Ion Traction Batteries in Vehicular Applications: Task Definition, Benchmark Evaluation, and Deployment Challenges

Haozhe Gao*

School of Technology, Beijing Forestry University, Beijing 100083, China

**Corresponding author: Haozhe Gao.*

Abstract

The state of health (SOH) of traction batteries is a common basis for safety warning, residual value assessment, and second-life utilization of electric vehicles, but a gap exists between the laboratory accuracy of existing deep learning methods and their on-board usability. This paper reviews the relevant research progress using on-board usability as the screening criterion. SOH estimation, SOH prediction, and remaining useful life (RUL) prediction are distinguished according to label source and evaluation protocol, and the limits on the comparability of error values are explained. Existing methods are grouped into six categories by modeling mechanism and evaluated category by category using four on-board questions: input availability, cross-domain generalization validation, uncertainty output, and edge-computing resource constraints. Six models are compared on the CALCE CS2 dataset under random splitting, temporal extrapolation, and cell-level holdout protocols. Under random splitting, the mean absolute errors of all six models do not exceed 0.0150. Under temporal extrapolation, they range from 0.0204 to 0.1392, and the R^2 values of three neural networks become negative. Under cell-level holdout, they range from 0.0254 to 0.0378. Under the CALCE CS2 dataset and the settings of this paper, protocol selection changes both the error level and the model ranking.

Keywords

lithium-ion traction battery, state-of-health estimation, state-of-health prediction, deep learning, on-board usability, data splitting protocol

1. Introduction

1.1 Engineering Requirements for On-Board Traction Battery State-of-Health Estimation

New energy vehicles and traction batteries are key components of transport decarbonization [1]. Battery state of health (SOH) determines driving range, charging and discharging capability, and safety limits [2], while practical vehicle requirements focus on degradation warning, residual value pricing, and admission to second-life utilization [3]. On-board conditions differ from laboratory conditions. A battery management system (BMS) collects only voltage, current, and temperature, and a complete constant-voltage charging stage is difficult to obtain [4]. Real-world operating conditions vary with driving behavior and ambient temperature,

© 2026 The Author(s). Authors retain copyright in their articles.

and the aging path changes accordingly [5]. The computing and storage capacities of the main controller are far below those of a training server [4]. These differences mean that laboratory accuracy cannot be used directly as a criterion for on-board usability.

1.2 Coverage and Limitations of Existing Reviews

Zhao et al. [3], Ma et al. [6], and Hu et al. [7] summarized battery and SOH estimation methods. Tang et al. [8], Yao et al. [9], and Luo et al. [10] reviewed estimation and prediction methods. Noura et al. [11] focused on hybrid electric vehicle scenarios, Wang et al. [12] discussed online estimation, and Xie et al. [13] covered whole-life prediction.

The limitations of existing reviews are concentrated at the validation level. Task boundaries are unclear, and state of charge and state of health are often discussed in parallel [14]. The comparability of estimation and prediction errors is not adequately explained. Classification is based on network names, making hybrid architectures difficult to categorize [15]. Performance indicators are mostly extracted from different sources, so the validity of horizontal comparisons is limited [16].

1.3 Work and Contributions of This Paper

This paper addresses four questions. RQ1 concerns the definitions and evaluation protocols of the three SOH tasks. RQ2 asks whether each category of deep learning methods remains valid under on-board constraints. RQ3 asks whether the splitting protocol changes the relative conclusions among methods. RQ4 asks what evidence is still lacking for on-board deployment.

The work of this paper includes distinguishing three types of SOH tasks, evaluating methods using four on-board questions (input availability, cross-domain generalization validation, uncertainty output, and edge-computing resource constraints), and comparing six models on the CALCE CS2 dataset under three protocols.

2. Problem Definition, Data, and Evaluation Protocols

2.1 Definition of SOH and Health Indicators

SOH has two definitions: capacity retention, which is the ratio of current available capacity to initial capacity, and internal resistance growth, which reflects degradation of power capability [2-3, 17]. Health indicators are commonly extracted from charging and discharging curves, including the peak height and peak position of the incremental capacity (IC) curve, differential voltage curve features, the voltage rise rate during constant-current charging, and statistics from partial charging intervals [17-20]. Electrochemical impedance spectroscopy requires dedicated excitation equipment and is difficult to obtain routinely under on-board conditions [21]. Indicators that depend on complete charging and discharging curves are often missing in real vehicles. Indicators from partial charging segments are closer to on-board conditions but are affected by the initial state of charge and temperature [19, 22].

2.2 Distinction Among the Three Tasks

The three tasks differ in label source and evaluation requirements, as shown in Table 1 (corresponding to RQ1). SOH estimation infers the current state from current measurements, and its label is obtained from capacity calibration in the same cycle. SOH prediction extrapolates the SOH trajectory over future cycles [13, 23]. RUL prediction outputs the number of remaining cycles until the end-of-life point [24-25]. Only the latter two must use temporal extrapolation. Because their label sources differ, the errors of estimation and prediction cannot be compared directly [16].

Table 1: Distinction Among SOH Estimation, SOH Prediction, and RUL Prediction

Task	Label source	Task type	Evaluation requirement
SOH estimation	Capacity calibration value from the same cycle	Static regression	Random splitting is only an interpolation baseline; generalization must be validated by cell-level holdout or temporal extrapolation
SOH prediction	Future-cycle SOH sequence	Temporal extrapolation	Temporal extrapolation is required; random shuffling is prohibited

© 2026 The Author(s). Authors retain copyright in their articles.

Task	Label source	Task type	Evaluation requirement
RUL prediction	Remaining number of cycles	Temporal extrapolation	Temporal extrapolation is required; error is measured in cycles

2.3 Public Datasets and On-Board Representativeness

NASA PCoE and CALCE are the most widely used datasets. The former uses constant-current charging and discharging at room temperature, whereas the latter includes multiple rates and temperatures. The experiments in this paper use the CS2 subset (room temperature, approximately 0.5C). The Oxford dataset uses urban-driving current profiles, and MIT-Stanford covers a relatively broad life distribution [26]. Real-vehicle and retired-battery datasets are larger, such as multimodal datasets containing driving and charging records [27].

The limitations of existing datasets include the following: most contain individual cells; coverage of combinations of temperature and operating conditions is limited; and the charging process is controlled and complete, unlike interrupted charging in real vehicles [27-28]. The operating-condition coverage of the datasets is shown in Table 2.

Table 2: Common Public Datasets and Their On-Board Representativeness

Dataset	Cells and scale	Operating-condition and temperature coverage	On-board representativeness
NASA PCoE	18650, 2 Ah	Mainly room temperature; constant-current charging and discharging	Single operating condition; no temperature variation
CALCE CS2 and CX2	18650 and prismatic cells	Multiple rates and temperatures; this paper uses CS2 (room temperature, 0.5C)	Relatively broad coverage, but still individual cells
Oxford Battery Degradation Dataset	Eight pouch cells	Urban-driving current profiles	Load is close to real vehicles; small scale
MIT-Stanford	124 18650 cells	Multiple fast-charging strategies	Broad life distribution; no real-vehicle conditions
Real-vehicle/retired-battery data	Fleet-level scale	Driving and charging records	Closest to on-board conditions; high labeling cost

2.4 Evaluation Metrics and Data Splitting Protocols

Error magnitude is measured by mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE). The first two use the SOH ratio as the unit, MAPE is expressed as a percentage, and goodness of fit is measured by the coefficient of determination (R^2) [13, 29]. The experiments in this paper report MAE, MAPE, and R^2.

Data splitting protocols are divided into three types. Random splitting shuffles all samples before splitting, so adjacent cycles of the same cell may fall on opposite sides of the split. Cell-level holdout uses an entire cell as the unit. Temporal extrapolation uses early cycles for training and later cycles for testing [11, 16]. Random splitting corresponds to interpolation within the same cell and can only serve as a baseline; it does not constitute evidence of generalization performance. Errors under the three protocols are not directly comparable.

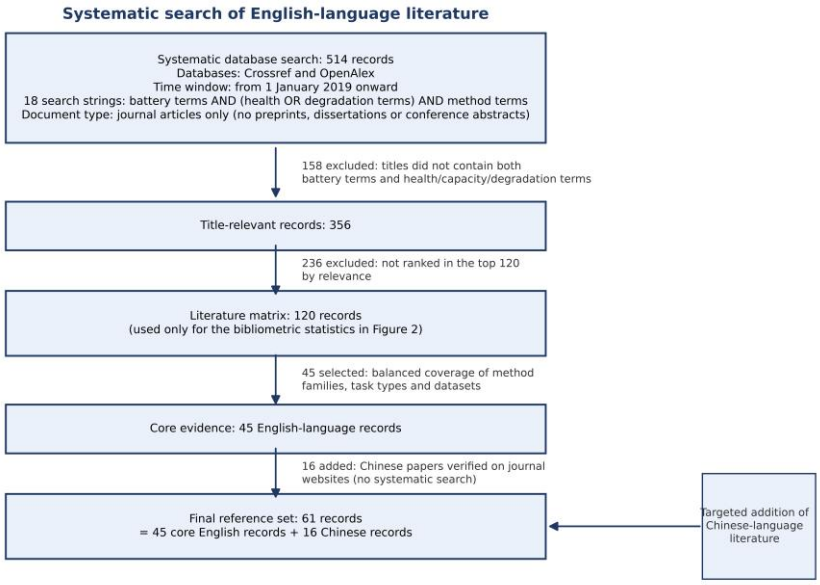
3. Methods and Analysis of On-Board Usability

3.1 Classification Principles and Evidence Base

The evidence was screened according to the process in Figure 1. The number of records and reasons for exclusion at each stage are marked in the figure (corresponding to RQ2).

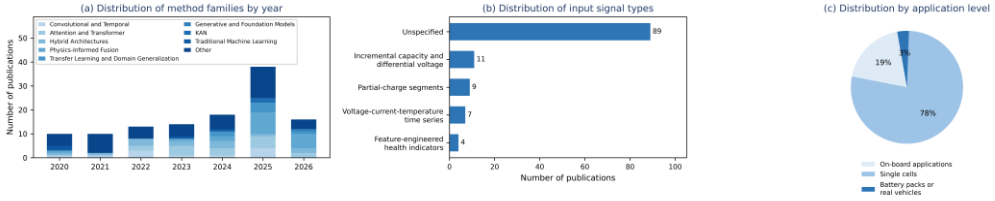
© 2026 The Author(s). Authors retain copyright in their articles.

Figure 1: Literature Screening and Inclusion Process



The English-language portion was obtained through systematic searches of Crossref and OpenAlex from January 1, 2019 onward using 18 search strings. The Chinese-language portion was not searched systematically; 16 records were added after targeted verification on journal websites, and omissions may remain. The final set of 61 records consisted of 45 core English-language records and 16 supplementary Chinese-language records. The title, method family, task definition, input signal, application scenario, and four-question labels of the 120 matrix papers were determined individually according to title-keyword rules; the statistics in Figures 2 and 3 and Section 5 can therefore be checked. Classification follows the main modeling mechanism rather than the network name. A convolutional long short-term memory (LSTM) network consists of local feature extraction and long-range dependency modeling and is therefore placed in the same family as a convolutional gated recurrent network. A new model is classified according to the component of the original system that it replaces; for example, a Kolmogorov-Arnold network replaces the function basis of a multilayer perceptron [30-32].

Figure 2: Bibliometric Characteristics of the Included Literature



3.2 Input Representation and Health Factor Construction

There are three types of input representation: voltage, current, and temperature time series [33-34]; health factors, including IC-curve peak features and equal-voltage-rise charging time [18, 35]; and local curves from charging segments [19-20, 22]. Availability increases in this order. The third type is closest to real-vehicle conditions but is most sensitive to the initial state of charge and temperature and therefore requires explicit normalization [22]. Mode decomposition can suppress noise [36], but the number of modes depends on experience [37].

3.3 Convolutional and Temporal Models

Convolutional networks extract curve shapes through local receptive fields [38]. Temporal convolutional networks enlarge the receptive field through dilated convolution and have been used under multiple-temperature conditions [34]. LSTM networks and gated recurrent units extract temporal dependencies [33, 39],

© 2026 The Author(s). Authors retain copyright in their articles.

while the combination of attention and recurrent structures increases the weights of key time steps [40]. The generalization boundary of these methods is unclear: the model learns the degradation trajectories of the training cells rather than transferable aging laws.

3.4 Attention and Transformer Models

Transformers model long-range dependencies through self-attention [41-42]. Sparse attention and sequence decomposition reduce complexity [43]. Graph convolutional networks can estimate SOH and RUL simultaneously [44]. Their advantage lies in representing long-sequence inputs, at the cost of high parameter counts and memory usage.

3.5 Hybrid Architectures

Hybrid architectures connect local feature extraction with long-range dependency modeling [18, 40, 45-47]. Several studies have reported low errors for such methods under within-cell interpolation settings, but the structures are complex, contain many hyperparameters, and depend on broad training-data coverage for cross-cell generalization [15]. Dual-branch structures also increase parameter count and latency.

3.6 Physics-Informed Fusion

Physics-informed fusion introduces mechanistic knowledge through constraints, loss terms, or network structures [48-52]. Recent studies have combined it with Kolmogorov-Arnold networks [31-32, 53]. Mechanistic constraints limit predictions to physically feasible ranges, and performance outside the training interval is generally better than that of purely data-driven methods. Its limitations are the high cost of parameter identification and the limited transferability of mechanistic terms because battery aging mechanisms vary with chemistry and operating conditions [5].

3.7 Transfer Learning, Meta-Learning, and Domain Generalization

Transfer learning transfers knowledge from a source domain to a target cell through feature alignment or parameter fine-tuning, and the main difficulty is the difference in data distributions. Domain adaptation and domain generalization relax the conditions for using target-domain data [54], while meta-learning addresses small-sample scenarios [55]. These methods respond to cross-domain shortcomings, but their effectiveness depends on the similarity between source and target domains. If both are adjacent cycles from the same group of cells, protocol leakage exaggerates the benefit [16].

3.8 Generative and Foundation Models

A generative pre-trained Transformer can predict degradation trajectories from early cycles [23]. A large language model combined with a multi-time-scale Transformer [56], a Transformer combined with a diffusion model [57], and a prompt-driven large language model [58] have been applied to SOH estimation and prediction. Their advantage lies in knowledge reuse across chemistries and in label-sparse scenarios. Their limitations are inference costs far beyond the BMS budget and a lack of constraints on physical consistency [23, 58].

3.9 Summary of Method Families and Common Problems

Table 3 summarizes the input forms, representative studies, cross-domain validation, and edge-device feasibility of the six method families. Figure 3 counts the evidence for the six families according to the four on-board questions, and each cell reports the number of confirmed records divided by the number of title hits. Coverage was determined in two steps. First, the titles of 120 papers were screened by keyword. Second, the abstracts and experimental settings of the matching papers were checked individually. The criterion was whether a paper provided corresponding experimental evidence: the input had to use incomplete or partial charging data; cross-domain validation had to provide results across cells, datasets, or operating conditions; uncertainty had to output intervals or probabilities and evaluate calibration; and edge deployment had to report measured inference latency, computing resources, or an embedded implementation. There were 38 title hits (36 papers after deduplication), all of which were checked (35 from full texts and one from the abstract).

© 2026 The Author(s). Authors retain copyright in their articles.

Twenty-three instances were confirmed and 15 were excluded. The item-by-item verification records are available from the author.

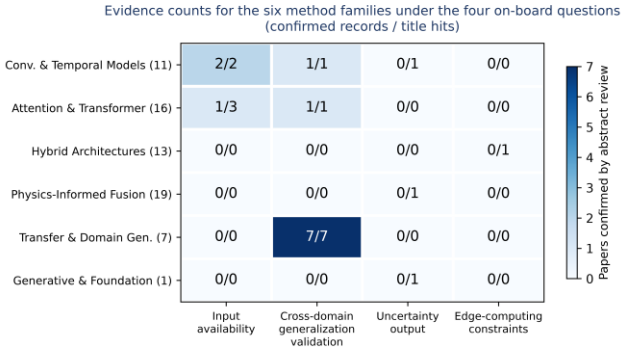
Table 3: Input Forms and On-Board Usability of the Six Method Families

Method family	Representative studies	Input form	Cross-domain validation (qualitative)	Edge-device feasibility (qualitative)
Convolutional and temporal models	[34, 38, 59]	Time series and health factors	Mainly within-cell validation	Low inference cost
Attention and Transformer	[39, 41, 44]	Long-sequence input	Mainly multi-temperature input	Model compression required
Hybrid architectures	[18, 42, 43]	Features and sequences	Limited cross-cell validation	Dual branches; high inference latency
Physics-informed fusion	[32, 48-49, 52]	Mechanistic parameters and curves	Mechanistic constraints favor extrapolation	High online-solution cost
Transfer and domain generalization	[54, 55]	Health factors	Targets cross-domain performance	Moderate inference burden
Generative and foundation models	[23, 56-58]	Sequences and prompts	Validation is at an early stage	Cloud training plus on-board inference

Note: The columns Cross-Domain Validation and Edge-Device Feasibility in Table 3 are qualitative judgments of the representative studies in each row. The counts in Figure 3 are the two-step results of title screening followed by abstract verification; the two use different scopes.

The common problems are as follows. Accuracy improvements do not proceed in step with on-board feasibility: errors within a single cell continue to decrease while deployment difficulty at the edge increases. Six papers with titles involving edge-computing constraints were checked individually, and none reported measured deployment or computing performance [4]. Uncertainty outputs are weak: of five title hits, only two were confirmed to provide interval or probabilistic outputs, while the remaining three were excluded after verification [59-60]. Most studies remain at random splitting within individual cells. The factors that constrain on-board application arise more from evaluation protocols and deployment evidence than from network structure itself.

Figure 3: Evidence Counts for the Six Method Families Under the Four On-Board Questions (Confirmed Records/Title Hits)



4. Controlled Experiments Under Three Data Splitting Protocols

4.1 Experimental Objective

This section compares six models under three splitting protocols using the same dataset and metric definitions (corresponding to RQ3). The reported values are valid only under the settings of this paper.

© 2026 The Author(s). Authors retain copyright in their articles.

4.2 Dataset and Feature Extraction

The experiments use the CALCE CS2 dataset, which contains six 18650 cells, CS2_33 through CS2_38, with a rated capacity of 1.1 Ah and cycling at approximately 0.5C at room temperature. The discharge capacity of a single cycle is calculated from the increments within the cycle, and the SOH label is the ratio of the discharge capacity of that cycle to the discharge capacity of the first cycle.

The input features are taken only from the charging stage. The 12 scalar features include charging duration, initial voltage, final voltage, maximum voltage, voltage mean and standard deviation, the times required to reach 3.8 V, 4.0 V, 4.1 V, and 4.2 V, and the time differences between intervals. A 64-dimensional sequence obtained by resampling the charging curve is also used. After cycles with abnormal capacities and samples with labels outside the range of 0.50 to 1.20 were removed, 4,908 samples remained. Each cell contained 711 to 932 valid cycles, and SOH ranged from 0.50 to 1.12. These features are extracted from controlled, complete charging stages and require a relatively complete charging process. Therefore, the controlled experiments in this Section are used only to test the effect of data splitting protocols on the conclusions, not to demonstrate on-board input availability or deployment feasibility. Partial-charging inputs are discussed in Section 5.

To examine the effect of reduced input information, an input variant with a 4.0 V truncated window was constructed using the same procedure. Only sampling points with charging current greater than 0.05 A and voltage below 4.0 V were retained. The voltage thresholds of the 12-dimensional features were shifted accordingly to 3.6 V, 3.8 V, 3.9 V, and 4.0 V, and the curve was resampled within the same window. Both input types used the same sample-screening rules. The truncated window retained 4,692 samples, 216 fewer than the complete-charging features; the difference came from cycles with fewer than 30 valid sampling points below 4.0 V.

4.3 Three Splitting Protocols and Comparison Models

Random splitting shuffles all samples and divides them at an 8:2 ratio. Test samples lie among training samples, so the model only needs to interpolate. Temporal extrapolation uses the first 70% of cycles of each cell for training and the last 30% for testing. Cell-level holdout uses an entire cell as the test unit, with five cells used for training and one for testing in each fold, rotating through all six cells. The splits under all three protocols are fixed. Random splitting uses seed 42 to generate one 8:2 split; temporal extrapolation is determined by cycle order; and cell-level holdout is determined by cell identifier. Three repetitions refer to model-side random seeds 0, 1, and 2, which affect feature subsampling in the random forest, weight initialization in the shallow neural network and the two deep models, and subsampling of the Gaussian process regression training set.

The comparison includes four conventional models and two deep models. Support vector regression uses a radial basis function kernel, penalty coefficient $C = 10$, kernel coefficient γ set to the scikit-learn 'scale' value, and epsilon-insensitive loss $\varepsilon = 0.005$. Gaussian process regression uses a fixed combination of a constant kernel, a radial basis function kernel (length scale 5.0), and a white-noise kernel (noise level 10^{-3}), without kernel-parameter optimization. Because the computational complexity of exact inference is $O(n^3)$, its training subset is limited to 900 samples. The random forest contains 200 trees. The shallow neural network has two hidden layers with 64 and 32 neurons. It is optimized with Adam using a learning rate of 10^{-3} for a maximum of 1,500 epochs; 10% of the training set is used for internal validation, with early stopping after 40 consecutive epochs without improvement. The one-dimensional convolutional neural network (1D-CNN) contains two convolutional layers followed by adaptive pooling. The LSTM network has a 32-dimensional hidden state. Both are trained for 120 epochs with Adam, a learning rate of 3×10^{-3}, a batch size of 128, and mean squared error loss. In the deep models, the output of the convolutional or recurrent layer is concatenated with the 12-dimensional scalar features and passed to a fully connected output layer. Inputs are standardized using statistics from the training set. The evaluation metrics are MAE, MAPE, and R^2. All hyperparameters take values commonly used in the literature and are determined only from the training set and its internal validation subset; the test set is not used for tuning. No independent validation set is used except for the internal validation of the shallow neural network.

© 2026 The Author(s). Authors retain copyright in their articles.

4.4 Experimental Results

Table 4 presents the performance of the six models under the three protocols. Values are means plus or minus standard deviations. The error bars in Figure 4 use the same definition as Table 4, and R^2 values below -2 are truncated in the display. Under random splitting, MAE ranges from 0.0061 to 0.0150, the mean MAPE over three repetitions does not exceed 1.77%, and all mean R^2 values exceed 0.977. Under temporal extrapolation, MAE rises to 0.0204-0.1392, and the mean R^2 values of the shallow neural network, 1D-CNN, and LSTM fall to -58.0, -1.54, and -6.41, respectively, meaning that their predictions are worse than a baseline using the training-set mean. Under cell-level holdout, MAE is 0.0254-0.0378, mean MAPE is 2.97%-4.35%, and mean R^2 is 0.858-0.940.

Table 4: Performance of Six Models Under Three Data Splitting Protocols (Mean +/- Standard Deviation of Statistical Units)

Model	Random splitting MAE	Temporal extrapolation MAE	Cell-level holdout MAE	Random splitting R^2	Temporal extrapolation R^2	Cell-level holdout R^2
Support vector regression	0.0073±0.0000	0.0435±0.0218	0.0340±0.0127	0.994±0.000	0.727±0.115	0.889±0.092
Gaussian process regression	0.0112±0.0002	0.0204±0.0121	0.0363±0.0148	0.986±0.001	0.935±0.061	0.866±0.128
Random forest	0.0061±0.0000	0.0360±0.0160	0.0281±0.0066	0.992±0.000	0.798±0.149	0.930±0.037
Shallow neural network	0.0111±0.0015	0.1392±0.0797	0.0378±0.0158	0.986±0.003	-58.0±98.2	0.858±0.117
1D-CNN	0.0118±0.0024	0.0434±0.0176	0.0254±0.0080	0.985±0.007	-1.54±1.22	0.940±0.036
LSTM	0.0150±0.0015	0.0566±0.0192	0.0288±0.0057	0.977±0.005	-6.41±4.43	0.928±0.028

Note: Values are the means plus or minus standard deviations of three repetitions, but the statistical unit of the standard deviation differs by protocol. The split under random splitting is fixed, so the standard deviation reflects only three model random seeds. Under temporal extrapolation and cell-level holdout, the values are first averaged over the three seeds and the standard deviation is then calculated across six data units (temporal folds or held-out cells); the two standard deviations are not directly comparable. The cross-seed standard deviations of the latter two protocols are 0.0000-0.0065 and 0.0000-0.0025, respectively; only the shallow neural network reaches 0.0766 under temporal extrapolation. To exclude the effect of sample size, the training subsets of all six models were also limited to 900 samples, and random splitting and cell-level holdout were rerun. MAE was 0.0091-0.0187 under random splitting and 0.0263-0.0649 under cell-level holdout.

The error gap remained under the equal-sample-size comparison. The errors of the models increased by factors of 1.9-4.3, and the model ranking also changed.

Table 4 and Figure 4 show that the protocol changes the relative model ranking. Random forest performs best under random splitting (0.0061), whereas Gaussian process regression performs best under temporal extrapolation (0.0204), and the shallow neural network performs worst (0.1392). Variation caused by changing data units is comparable to that caused by model selection. Under cell-level holdout, the cross-cell standard deviations of the models range from 0.0057 to 0.0158, while the difference between model means is 0.0124. The ratio of the maximum to the minimum six-fold mean MAE is 2.7 (CS2_35 is lowest at 0.0185 and CS2_34 is highest at 0.0492), and this difference arises from cell-to-cell variation.

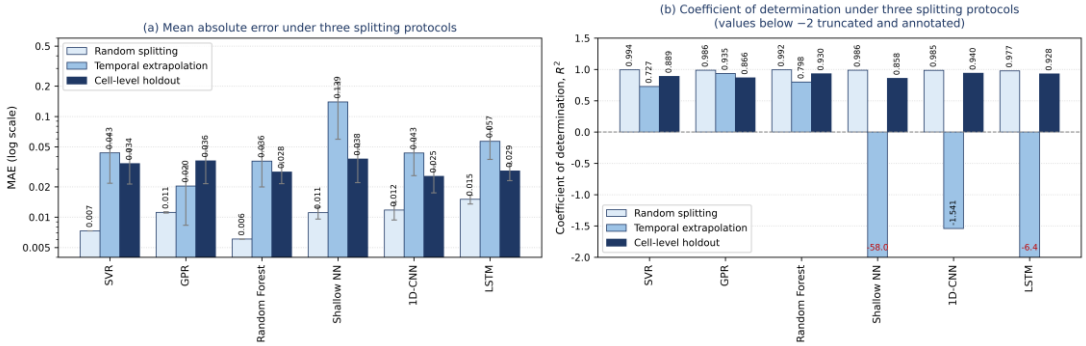
Table 5 compares the 4.0 V truncated window with the complete-charging features. Under random splitting, errors increase comprehensively by factors of 1.3-2.3 (MAE rises from 0.0061-0.0150 to 0.0132-0.0197), indicating that within-cell interpolation accuracy is most sensitive to input completeness. Under cell-level holdout, the factor is 0.96-1.40, with random forest increasing from 0.0281 to 0.0374. Under temporal extrapolation, the factor is 0.54-1.45. The shallow neural network decreases from 0.1392 to 0.0752, indicating that input truncation suppresses this model's instability under extrapolation. The truncated window does not change the relative relationship among protocols; random splitting remains the most optimistic group.

© 2026 The Author(s). Authors retain copyright in their articles.

Table 5: MAE Comparison Between Complete-Charging Features and the 4.0 V Truncated Window (Each Cell Shows Complete → Truncated)

Model	Random splitting	Temporal extrapolation	Cell-level holdout
Support vector regression	0.0073 → 0.0132	0.0435 → 0.0305	0.0340 → 0.0425
Gaussian process regression	0.0112 → 0.0172	0.0204 → 0.0296	0.0363 → 0.0347
Random forest	0.0061 → 0.0138	0.0360→ 0.0375	0.0281 → 0.0374
Shallow neural network	0.0111→ 0.0186	0.1392 → 0.0752	0.0378 → 0.0382
1D-CNN	0.0118 → 0.0197	0.0434 → 0.0421	0.0254 → 0.0355
LSTM	0.0150 → 0.0195	0.0566 → 0.0404	0.0288 → 0.0365

Figure 4: Effects of Three Data Splitting Protocols on Model Performance



4.5 Discussion and Limitations

The difference between random splitting and temporal extrapolation is the difference between interpolation and extrapolation. The difference between temporal extrapolation and cell-level holdout lies in the extrapolation target, and the former shows a more pronounced increase in error. Under random splitting, test samples fall inside the training interval; the fact that training and test samples come from the same source is only one manifestation of this problem.

Therefore, under the CALCE CS2 dataset and the settings of this paper, random splitting reduces the reported error level and changes the model ranking. Its values should not be used as evidence of generalization performance.

The experimental limitations are as follows: the data come from a single dataset and a single rate; model scale and training epochs are simplified to control cost, and no independent validation set is used except for the internal validation of the shallow neural network; temporal extrapolation tests only the same rate; the truncated window still comes from controlled charging and differs from randomly interrupted segments in real vehicles; and repeated splits and repeated holdouts are not performed, so cross-cell variation reflects only the six available cells.

5. Challenges and Research Directions for On-Board Deployment

This Section summarizes common limitations according to the four on-board questions (Table 6, corresponding to RQ4). Literature evidence and experimental evidence are taken from the literature matrix in Section 2 and the experiments in Section 4, respectively.

5.1 Availability of Input Signals

Among the 120 papers included in the statistics, 93 use individual cells as the application scenario, and only nine involve partial charging segments (as judged by title). For input availability, 18 titles matched, of which 12 were confirmed and six were excluded after individual verification. Methods that rely on electrochemical impedance spectroscopy require dedicated excitation equipment [21], and methods that rely on a complete constant-voltage charging stage fail when charging is interrupted [20]. Partial charging segments are the easiest inputs to obtain. A feasible direction is to construct health factors around them and use the initial state of charge and ambient temperature as explicit inputs [19, 22]. The difficulty lies in maintaining stability when segment length and initial conditions vary simultaneously.

© 2026 The Author(s). Authors retain copyright in their articles.

5.2 Generalization Across Cells and Operating Conditions

Based on title screening, only nine papers involve cross-domain transfer or generalization, seven of which belong to the transfer and domain-generalization family. Individual verification confirmed that all nine provide validation across cells, datasets, or operating conditions. The experiments in Section 4 show that errors increase by factors of 1.8-12.5 after switching to temporal extrapolation. Transfer learning, domain generalization, and meta-learning provide feasible paths [54, 55], but their effectiveness depends on the similarity between source and target domains. When the protocol does not strictly separate the two, leakage exaggerates the benefit [16]. A prerequisite is a public benchmark covering multiple chemistries, temperatures, and rates [27].

5.3 Insufficient Data in the Early-Life Stage

Labels are scarce during the early stage of battery use, whereas users and vehicle manufacturers most need early judgments. Generative and foundation models reduce label dependence through pre-training [23, 58], meta-learning has advantages with few samples [55], and the MIT-Stanford dataset covers a relatively broad life distribution [26]. However, early-life windows range from tens to hundreds of cycles across studies, and no conclusion has been reached regarding the trade-off between window length and accuracy.

5.4 Uncertainty and Interpretability

BMS safety warning and residual value assessment require interval estimates. Existing studies provide prediction intervals through randomized methods and Bayesian neural networks [59-60], but the stability of interval coverage under cross-domain conditions has not been validated. Of five title hits, only two were confirmed to output intervals or probabilities. Although the titles of the other three contained ‘interval’ or ‘stochastic,’ these terms actually referred to a time scale or a stochastic configuration network. Attention weights and the physical meanings of health factors lack strict correspondence, making the basis of model judgments difficult to trace. Review studies have also identified interpretability as an unresolved issue [61].

5.5 Edge Computing and Real-Time Performance

The computing and storage budgets of BMS main controllers are limited. Lightweight methods have been developed through pruning, distillation, and fixed-point quantization [4], but individual verification of six papers whose titles involved edge deployment or real-time performance found that none reported measured inference latency, computing-resource usage, or an embedded implementation. Among the 120 papers included in this paper, title-level screening also found no measured comparison of computing performance on a specific BMS platform.

5.6 Research Opportunities and Decision Criteria

The following directions should be prioritized. First, public evaluation benchmarks with protocol labels should be established to cover multiple chemistries, temperatures, and rates, and errors under all three protocols should be reported. Second, condition-adaptive models for partial charging segments should be developed, with initial state of charge and ambient temperature used as explicit inputs, while ensuring that error variation remains acceptable when segment length and initial conditions change. Third, uncertainty should be incorporated into standard evaluation metrics, with both interval coverage and width reported [16, 60].

Table 6: Literature Evidence and Missing Evidence for On-Board Deployment Challenges

Challenge	Evidence in the literature	Evidence still lacking
Input-signal availability	93/120 papers use individual cells; 18 title hits for input, with 12 confirmed	Stability under real charging distributions
Generalization across cells and operating conditions	Nine cross-domain title hits, all nine confirmed; errors in this paper increase by factors of 1.8-12.5	Benchmarks spanning multiple chemistries and temperatures
Insufficient early-life data	Only one paper on generative and foundation models	Trade-off between early-window length and accuracy
Uncertainty and interpretability	Five uncertainty title hits, with two confirmed	Interval coverage under cross-domain conditions
Edge computing and real-time performance	Six edge-computing title hits, with none confirmed after individual verification	Accuracy-computing trade-off for a specific BMS platform

© 2026 The Author(s). Authors retain copyright in their articles.

6. Conclusions

At the task level, this paper distinguishes SOH estimation, SOH prediction, and RUL prediction and explains why high accuracy in an estimation task does not support prediction capability. At the method level, existing approaches are grouped into six categories according to modeling mechanism. A two-step procedure of title screening followed by verification of abstracts and experimental settings was conducted according to the four on-board questions (Figure 3), and all 36 matching papers were checked. All nine papers on cross-domain generalization were confirmed. Twelve of 18 papers on input availability were confirmed, whereas the confirmed counts for uncertainty output and edge-computing constraints were two and zero, respectively, and the confirmed counts were still lower within the six method families. Physics-informed fusion has advantages in cross-condition extrapolation, and generative and foundation models have potential in label-scarce scenarios, but their inference costs exceed the BMS budget.

The comparison under three protocols provides the most direct evidence. The MAEs of all six models do not exceed 0.0150 under random splitting. Under temporal extrapolation, errors increase by factors of 1.8-12.5 and three neural networks produce negative R^2 values. Under cell-level holdout, errors increase by factors of 1.9-4.7. The comparison using the 4.0 V truncated window shows that shortening the input mainly reduces within-cell interpolation accuracy (errors increase by factors of 1.3-2.3) and does not change the relative relationship among protocols. The protocol changes both the error level and the model ranking. Under the CALCE CS2 dataset and the settings of this paper, the influence of protocol selection on the conclusion is no less than that of model selection.

Accordingly, on-board SOH methods should use temporal extrapolation as the minimum validation requirement and report the splitting unit and metric definitions. Uncertainty and edge-resource usage should be standard reporting items. The evidence still lacking for on-board deployment is concentrated on input stability under real charging behavior, interval coverage under cross-domain conditions, and data on the trade-off between accuracy and computing resources for specific BMS platforms.

References

[1] Ouyang, M. (2021). Innovation and development of new energy vehicles for carbon neutrality. Scientific Chinese, (11), 26-31. (in Chinese)

[2] Xiong, R. (2019). Core algorithms of power battery management systems. China Machine Press. (in Chinese)

[3] Zhao, X., Li, M., Yu, Q., Ma, J., & Wang, S. (2023). State estimation of power lithium batteries for electric vehicles: A review. China Journal of Highway and Transport, 36(6), 254-283. https://doi.org/10.19721/j.cnki.1001-7372.2023.06.021 (in Chinese)

[4] Ye, S., Wang, Y., Pang, T., Wang, K., Liu, Q., Fan, G., & Zhang, X. (2025). A review of safety-oriented intelligent management methods for lithium-ion batteries. Drive System Technique, 39(2), 3-14, 22. (in Chinese)

[5] Song, X., Yu, Z., Xiao, Y., Wang, D., Yuan, Q., Li, X., & Zheng, J. (2025). Research progress on the performance changes of lithium-ion batteries after aging. Journal of Jilin University (Engineering and Technology Edition), 55(6), 1817-1833. https://doi.org/10.13229/j.cnki.jdxbgxb.20240763 (in Chinese)

[6] Ma, Z., Li, Y., Zhan, Z., Song, Y., Zhang, Z., & Qiu, Z. (2025). A review of state-of-health estimation methods for lithium-ion batteries. Journal of Chongqing University of Technology (Natural Science), 39(12), 1-10. (in Chinese)

[7] Hu, X., Guo, Y., & Zhang, R. (2022). Review of state-of-health estimation methods for lithium-ion batteries. Journal of Power Supply, 20(1), 126-133. https://doi.org/10.13234/j.issn.2095-2805.2022.1.126 (in Chinese)

[8] Tang, K., Luo, B., Chen, D., Wang, C., Chen, L., Li, F., Cao, Y., & Wang, C. (2025). The State of Health Estimation of Lithium-Ion Batteries: A Review of Health Indicators, Estimation Methods, Development Trends and Challenges. World Electric Vehicle Journal, 16(8), 429. https://doi.org/10.3390/wevj16080429

© 2026 The Author(s). Authors retain copyright in their articles.

[9] Yao, L., Xu, S., Tang, A., Zhou, F., Hou, J., Xiao, Y., & Fu, Z. (2021). A Review of Lithium-Ion Battery State of Health Estimation and Prediction Methods. World Electric Vehicle Journal, 12(3), 113. https://doi.org/10.3390/wevj12030113

[10] Luo, K., Chen, X., Zheng, H., & Shi, Z. (2022). A review of deep learning approach to predicting the state of health and state of charge of lithium-ion batteries. Journal of Energy Chemistry, 74, 159-173. https://doi.org/10.1016/j.jechem.2022.06.049

[11] Noura, N., Boulon, L., & Jemeï, S. (2020). A Review of Battery State of Health Estimation Methods: Hybrid Electric Vehicle Challenges. World Electric Vehicle Journal, 11(4), 66. https://doi.org/10.3390/wevj11040066

[12] Wang, Z., Feng, G., Zhen, D., Gu, F., & Ball, A. (2021). A review on online state of charge and state of health estimation for lithium-ion batteries in electric vehicles. Energy Reports, 7, 5141-5161. https://doi.org/10.1016/j.egyr.2021.08.113

[13] Xie, Y., Wang, S., Zhang, G., Takyi-Aninakwa, P., Fernandez, C., & Blaabjerg, F. (2024). A review of data-driven whole-life state of health prediction for lithium-ion batteries: Data preprocessing, aging characteristics, algorithms, and future challenges. Journal of Energy Chemistry, 97, 630-649. https://doi.org/10.1016/j.jechem.2024.06.017

[14] Yao, J., & Kowal, J. (2025). Towards a smarter battery management system: A critical review on deep learning-based state of charge estimation of lithium-ion batteries. Energy and AI, 21, 100585. https://doi.org/10.1016/j.egyai.2025.100585

[15] Alsuwian, T., Ansari, S., Ammirrul Atiqi Mohd Zainuri, M., Ayob, A., Hussain, A., Hossain Lipu, M. S., Alhawari, A. R. H., Almawgani, A. H. M., Almasabi, S., & Taher Hindi, A. (2024). A review of expert hybrid and co-estimation techniques for SOH and RUL estimation in battery management system with electric vehicle application. Expert Systems with Applications, 246, 123123. https://doi.org/10.1016/j.eswa.2023.123123

[16] Sui, X., He, S., Vilsen, S. B., Meng, J., Teodorescu, R., & Stroe, D. I. (2021). A review of non-probabilistic machine learning-based state of health estimation techniques for Lithium-ion battery. Applied Energy, 300, 117346. https://doi.org/10.1016/j.apenergy.2021.117346

[17] Li, C., Wang, C., Wang, G., Lu, Z., & Ma, C. (2022). Technology of lithium-ion battery state-of-health assessment based on data-driven methods. Electric Power, 55(8), 73-86, 95. https://doi.org/10.11930/j.issn.1004-9649.202202012 (in Chinese)

[18] Dai, Y., & Yu, A. (2022). Combined CNN-LSTM and GRU based health feature parameters for lithium-ion battery SOH estimation. Energy Storage Science and Technology, 11(5), 1641-1649. https://doi.org/10.19799/j.cnki.2095-4239.2021.0623 (in Chinese)

[19] Li, S., Dauletiyarova, A. B., Zhao, X., Zhang, K., & Jiao, Z. (2026). SOH estimation method for lithium-ion batteries based on partial charging voltage features. Automotive Engineering, 48(1), 113-126. https://doi.org/10.19562/j.chinasae.qcgc.2026.01.011 (in Chinese)

[20] Tian, J., Xiong, R., Shen, W., Lu, J., & Sun, F. (2022). Flexible battery state of health and state of charge estimation using partial charging data and deep learning. Energy Storage Materials, 51, 372-381. https://doi.org/10.1016/j.ensm.2022.06.053

[21] Yang, B., Liu, S., Feng, X., Rao, Z., Liu, J., & Wu, W. (2025). Thermal state of lithium-ion batteries: Definition, detection, and estimation. Scientia Sinica Technologica, 55(2), 187-212. https://doi.org/10.1360/sst-2024-0124 (in Chinese)

[22] Deng, Z., Hu, X., Li, P., Lin, X., & Bian, X. (2022). Data-Driven Battery State of Health Estimation Based on Random Partial Charging Data. IEEE Transactions on Power Electronics, 37(5), 5021-5031. https://doi.org/10.1109/tpel.2021.3134701

[23] Hu, J., Fu, P., Wei, Z., Huang, Y., Early, J., Fly, A., & Zhang, Y. (2026). Early prediction of lithium-ion battery degradation with a generative pre-trained transformer. Nature Communications, 17(1), 126. https://doi.org/10.1038/s41467-025-66819-0

© 2026 The Author(s). Authors retain copyright in their articles.

[24] Zhu, Z., Yang, Q., Liu, X., & Gao, D. (2022). Attention-based CNN-BiLSTM for SOH and RUL estimation of lithium-ion batteries. Journal of Algorithms & Computational Technology, 16, 17483026221130598. https://doi.org/10.1177/17483026221130598

[25] Qu, J., Liu, F., Ma, Y., & Fan, J. (2019). A Neural-Network-Based Method for RUL Prediction and SOH Monitoring of Lithium-Ion Battery. IEEE Access, 7, 87178-87191. https://doi.org/10.1109/access.2019.2925468

[26] Severson, K. A., Attia, P. M., Jin, N., Perkins, N., Jiang, B., Yang, Z., Chen, M. H., Aykol, M., Herring, P. K., Fraggedakis, D., Bazant, M. Z., Harris, S. J., Chueh, W. C., & Braatz, R. D. (2019). Data-driven prediction of battery cycle life before capacity degradation. Nature Energy, 4(5), 383-391. https://doi.org/10.1038/s41560-019-0356-8

[27] Liu, H., Li, C., Hu, X., Li, J., Zhang, K., Xie, Y., Wu, R., & Song, Z. (2025). Multi-modal framework for battery state of health evaluation using open-source electric vehicle data. Nature Communications, 16(1), 1137. https://doi.org/10.1038/s41467-025-56485-7

[28] Song, L., Zhang, K., Liang, T., Han, X., & Zhang, Y. (2020). Intelligent state of health estimation for lithium-ion battery pack based on big data analysis. Journal of Energy Storage, 32, 101836. https://doi.org/10.1016/j.est.2020.101836

[29] Gong, Q., Wang, P., & Cheng, Z. (2022). An encoder-decoder model based on deep learning for state of health estimation of lithium-ion battery. Journal of Energy Storage, 46, 103804. https://doi.org/10.1016/j.est.2021.103804

[30] Zhao, X., Dong, Z., Zhou, H., Lin, H., & Ji, X. (2026). A lithium-ion battery state-of-health prediction network integrating Kolmogorov-Arnold representation theory. Acta Physica Sinica, 75(3), 030809. https://doi.org/10.7498/aps.75.20251359 (in Chinese)

[31] Liao, Y., Deng, J., & Hu, Y. (2026). Unified Explicit–Implicit Degradation Dynamics with Physics-Informed Kolmogorov–Arnold Networks for Lithium-Ion Battery SOH Estimation and RUL Prediction Under Different Charging Protocols. Energies, 19(16), 3742. https://doi.org/10.3390/en19163742

[32] Liu, Z., Ye, S., Cui, F., & Ma, Y. (2025). Physical Information-Guided Kolmogorov–Arnold Networks for Battery State of Health Estimation. Energies, 18(22), 5865. https://doi.org/10.3390/en18225865

[33] Liu, C., Li, H., Li, K., Wu, Y., & Lv, B. (2025). Deep Learning for State of Health Estimation of Lithium-Ion Batteries in Electric Vehicles: A Systematic Review. Energies, 18(6), 1463. https://doi.org/10.3390/en18061463

[34] Bockrath, S., Lorentz, V., & Pruckner, M. (2023). State of health estimation of lithium-ion batteries with a temporal convolutional neural network using partial load profiles. Applied Energy, 329, 120307. https://doi.org/10.1016/j.apenergy.2022.120307

[35] Xu, Z., Chen, Z., Yang, L., & Zhang, S. (2024). State of health estimation for lithium-ion batteries based on incremental capacity analysis and Transformer modeling. Applied Soft Computing, 165, 112072. https://doi.org/10.1016/j.asoc.2024.112072

[36] Xiao, H., He, X., Liang, J., & Li, C. (2022). A lithium battery life-prediction method based on mode decomposition and machine learning. Energy Storage Science and Technology, 11(12), 3999-4009. https://doi.org/10.19799/j.cnki.2095-4239.2022.0341 (in Chinese)

[37] Zhang, C., Luo, L., Yang, Z., Du, B., Zhou, Z., Wu, J., & Chen, L. (2024). Flexible method for estimating the state of health of lithium-ion batteries using partial charging segments. Energy, 295, 131009. https://doi.org/10.1016/j.energy.2024.131009

[38] Yang, N., Song, Z., Hofmann, H., & Sun, J. (2022). Robust state-of-health estimation of lithium-ion batteries using a convolutional neural network and random forest. Journal of Energy Storage, 48, 103857. https://doi.org/10.1016/j.est.2021.103857

[39] Lin, M., Wu, J., Meng, J., et al. (2023). State-of-health estimation with an attentional long short-term memory network for lithium-ion batteries. Energy, 268, 126706. https://doi.org/10.1016/j.energy.2023.126706

© 2026 The Author(s). Authors retain copyright in their articles.

[40] Tang, A., Jiang, Y., Yu, Q., & Zhang, Z. (2023). A hybrid neural network model with attention mechanism for state-of-health estimation of lithium-ion batteries. Journal of Energy Storage, 68, 107734. https://doi.org/10.1016/j.est.2023.107734

[41] Guirguis, J., & Ahmed, R. (2024). Transformer-based deep learning models for state of charge and state of health estimation of Li-ion batteries: A survey study. Energies, 17(14), 3502. https://doi.org/10.3390/en17143502

[42] Zhao, H., Chen, Z., Shu, X., Shen, J., Lei, Z., & Zhang, Y. (2023). State of health estimation for lithium-ion batteries based on hybrid attention and deep learning. Reliability Engineering & System Safety, 232, 109066. https://doi.org/10.1016/j.ress.2022.109066

[43] Gao, J., Yang, D., Wang, S., Li, Z., Wang, L., & Wang, K. (2023). State of health estimation of lithium-ion batteries based on Mixers-bidirectional temporal convolutional neural network. Journal of Energy Storage, 73, 109248. https://doi.org/10.1016/j.est.2023.109248

[44] Wei, Y., & Wu, D. (2023). Prediction of state of health and remaining useful life of lithium-ion batteries using a graph convolutional network with dual attention mechanisms. Reliability Engineering & System Safety, 230, 108947. https://doi.org/10.1016/j.ress.2022.108947

[45] Mazzi, Y., Ben Sassi, H., & Errahimi, F. (2024). Lithium-ion battery state of health estimation using a hybrid model based on a convolutional neural network and bidirectional gated recurrent unit. Engineering Applications of Artificial Intelligence, 127, 107199. https://doi.org/10.1016/j.engappai.2023.107199

[46] Li, Z., Zhang, X., & Gao, W. (2024). State of health estimation of lithium-ion battery during fast charging process based on BiLSTM-Transformer. Energy, 311, 133418. https://doi.org/10.1016/j.energy.2024.133418

[47] He, T., & Gong, Z. (2024). State of health estimation for lithium-ion batteries using a hybrid neural network model with Multi-scale Convolutional Attention Mechanism. Journal of Power Sources, 609, 234680. https://doi.org/10.1016/j.jpowsour.2024.234680

[48] Yang, L., He, M., Ren, Y., Gao, B., & Qi, H. (2025). Physics-informed neural network for co-estimation of state of health, remaining useful life, and short-term degradation path in Lithium-ion batteries. Applied Energy, 398, 126427. https://doi.org/10.1016/j.apenergy.2025.126427

[49] Tian, A., He, L., Ding, T., Dong, K., Wang, Y., & Jiang, J. (2025). A generic physics-informed neural network framework for lithium-ion batteries state of health estimation. Energy, 332, 137215. https://doi.org/10.1016/j.energy.2025.137215

[50] Luo, Y., Ju, S., Li, P., & Zhang, H. (2025). A method for estimating lithium-ion battery state of health based on physics-informed hybrid neural network. Electrochimica Acta, 525, 146110. https://doi.org/10.1016/j.electacta.2025.146110

[51] Ye, J., Xie, Q., Lin, M., & Wu, J. (2024). A method for estimating the state of health of lithium-ion batteries based on physics-informed neural network. Energy, 294, 130828. https://doi.org/10.1016/j.energy.2024.130828

[52] Hofmann, T., Hamar, J., Rogge, M., Zoerr, C., Erhard, S., & Philipp Schmidt, J. (2023). Physics-Informed Neural Networks for State of Health Estimation in Lithium-Ion Batteries. Journal of The Electrochemical Society, 170(9), 090524. https://doi.org/10.1149/1945-7111/acf0ef

[53]Jarraya, I., Ben Atitallah, S., Alahmed, F., Abdelkader, M., Driss, M., Abdelhadi, F., & Koubaa, A. (2025). SOH-KLSTM: A hybrid Kolmogorov-Arnold Network and LSTM model for enhanced Lithium-ion battery Health Monitoring. Journal of Energy Storage, 122, 116541. https://doi.org/10.1016/j.est.2025.116541

[54] Lu, J., Xiong, R., Tian, J., Wang, C., & Sun, F. (2023). Deep learning to estimate lithium-ion battery state of health without additional degradation experiments. Nature Communications, 14(1), 2760. https://doi.org/10.1038/s41467-023-38458-w

© 2026 The Author(s). Authors retain copyright in their articles.

[55] Wang, R., Duan, B., Zhao, G., Shen, J., & Liu, S. (2026). Meta-learning driven small-sample lithium-ion battery state of health estimation based on Kolmogorov–Arnold Network and Long Short-Term Memory Network. Journal of Energy Storage, 163, 122207. https://doi.org/10.1016/j.est.2026.122207

[56] Zhang, W., Feng, H., Ren, M., Liu, K., & Yan, G. (2026). LLM-MSIformer: A method for lithium-ion battery state of health estimation based on a large language model and multi-time scale interval transformer. Energy, 342, 139417. https://doi.org/10.1016/j.energy.2025.139417

[57] Ding, H., & Yan, S. (2026). Lithium battery state-of-health prediction method combining a Transformer with diffusion. Electrical Measurement & Instrumentation, 63(5), 140-151. https://doi.org/10.19753/j.issn1001-1390.2026.05.015 (in Chinese)

[58] Guo, Y., Fan, X., Shi, L., et al. (2026). Large-language-model-driven method for lithium-ion battery SOH prediction. Journal of Harbin Institute of Technology, 58(8), 227-239. (in Chinese)

[59] Liang, C., Tao, S., Huang, X., Wang, Y., Xia, B., & Zhang, X. (2025). Stochastic state of health estimation for lithium-ion batteries with automated feature fusion using quantum convolutional neural network. Journal of Energy Chemistry, 106, 205-219. https://doi.org/10.1016/j.jechem.2025.02.030

[60] Lin, M., Xiao, P., Meng, J., Wang, W., & Wu, J. (2026). State of Health Estimation for Battery Packs Based on Ensemble Bayesian Neural Networks Integrating Uncertainty Quantification and Calibration. IEEE Transactions on Transportation Electrification, 12(4), 6188-6199. https://doi.org/10.1109/tte.2026.3683737

[61] Wang, J. (2025). Research progress and development trends in battery remaining useful life prediction methods. Drive System Technique, 39(4), 52-58. (in Chinese)

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).

© 2026 The Author(s). Authors retain copyright in their articles.