

Active-learning-driven error control for data-driven state of charge estimation across the lithium battery lifecycle

Jinwei Xue^{a,b}, Xuzhi Du^{c,*}, Lei Zhao^{a,b}, Zhigang Yang^{a,b,d,*}, Chao Xia^e, Yuan Ma^f, Muhammad Jahidul Hoque^c, Wuchen Fu^c, Xiao Yan^{g,h}, Nenad Miljkovic^{c,i,j,k,l,m,*}

^a School of Automotive Studies, Tongji University, Shanghai, 201804, China

^b Shanghai Automotive Wind Tunnel Center, Tongji University, 201804, China

^c Mechanical Science and Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

^d Beijing Aeronautical Science & Technology Research Institute, Beijing, 102211, China

^e Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Gothenburg 412 96, Sweden

^f Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China

^g Key Laboratory of Low-grade Energy Utilization Technologies and Systems, Chongqing University, Ministry of Education, Chongqing, 400030, China

^h Institute of Engineering Thermophysics, School of Energy and Power Engineering, Chongqing University, Chongqing, 400030, China

ⁱ Institute for Sustainability, Energy, and Environment, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

^j Department of Electrical and Computer Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

^k Materials Research Laboratory, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

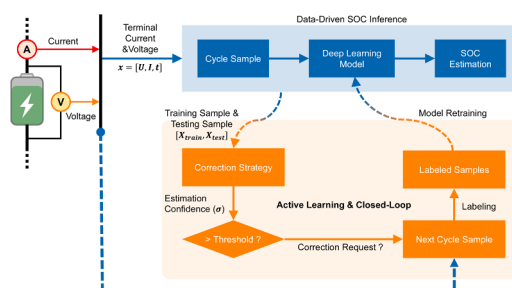
^l Air Conditioning and Refrigeration Center, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

^m International Institute for Carbon Neutral Energy Research (WPI-I2CNER), Kyushu University, 744 Moto-oka, Nishi-ku, Fukuoka 819-0395, Japan

HIGHLIGHTS

- Closed-loop SOC estimation framework with uncertainty-aware correction.
- Active learning minimizes retraining while improving SOC prediction accuracy.
- Model Disagreement index correlates strongly (0.91) with true SOC error.
- Achieves <1.5 % average SOC error with only 4 retrainings under WLTC.
- Framework enables robust SOC tracking across battery aging and fast charging.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:

Lithium-ion battery
State of charge (SOC)
Battery aging
data-driven SOC estimation
Deep learning
Active learning
Closed-loop correction

ABSTRACT

Accurate estimation of lithium-ion battery state of charge (SOC) is crucial for the safe and efficient operation of electric vehicles (EVs). However, both data-driven and model-driven SOC estimation methods face significant challenges under battery aging, which alters internal resistance and electrochemical properties, especially across complex aging trajectories. Most existing deep learning and model-based approaches operate in an open-loop manner, lacking mechanisms for uncertainty quantification, accuracy prediction, or adaptive correction—leading to uncontrolled estimation errors during aging. To address this, we propose an innovative closed-loop SOC estimation framework that integrates active learning with uncertainty-aware correction into deep learning networks, enabling real-time feedback on SOC prediction confidence levels without the need for additional

* Corresponding authors.

E-mail addresses: xuzhidu@illinois.edu (X. Du), zhigang.yang@sawtc.com (Z. Yang), nmiljkov@illinois.edu (N. Miljkovic).

<https://doi.org/10.1016/j.egyai.2025.100549>

Available online 2 July 2025

2666-5468/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

sensors or reference data. Specifically, we quantify the performance degradation of mainstream data-driven methods, including long short-term memory (LSTM) networks and Gaussian process regression (GPR), under complex aging paths. We demonstrate that our model-disagreement-based active learning correction strategy maintains robustness throughout the battery lifecycle. Experimental results show that with only four active retraining sessions over the full aging process, our method reduces average SOC estimation error to below 1.5 %, and maximum cycle-based average error to below 2 %. This work establishes a path toward uncertainty-informed, lifecycle-resilient, and data-efficient SOC estimation, marking a significant advancement in battery management systems for real-world EV applications.

1. Introduction

Lithium-ion batteries (LIBs) are used in electric vehicles (EVs) due to their high voltage, high energy density, low self-discharge, and long lifecycle [1–3]. The efficient operation of EVs depends directly on the battery management system (BMS) [4–7], where the state of charge (SOC) is the critical characteristic parameter. SOC denotes the remaining charge of the battery, defined mathematically by Eq. (1) [1].

$$\text{SOC} = \frac{Q_{\text{available}}}{Q_{\text{rated}}}, \quad (1)$$

where $Q_{\text{available}}$ is the available amount of charge, and Q_{rated} is the full charge capacity.

Accurately estimating the SOC for LIBs is crucial. The rated capacity (Q_{rated}) provided by the battery manufacturer does not represent the true capacity of the battery [8]. Moreover, Q_{rated} varies throughout the battery lifecycle due to factors such as usage time, environmental temperature fluctuations, and complex chemical reactions [9]. Additionally, the availability of sensors capable of directly measuring electrochemical phenomena (such as current, potential, and conductance) inside the battery is limited [10]. Furthermore, manufacturing defects and physical damage during assembly also significantly impact SOC estimation [11]. Due to these uncontrollable factors, high-precision SOC estimation remains a challenging problem.

Mainstream SOC estimation methods can be broadly categorized into two categories: model-driven and data-driven approaches [1]. Model-based SOC estimation methods, also known as the white-box models [12], are constructed based on the principles governing battery charging and discharging. These methods leverage a deep understanding of battery system principles, employing core equations that accurately model these processes [13]. For example, Wang et al. [14–16] proposed two Kalman filter-based architectures to improve the SOC estimation ability and accuracy of LIBs, as well as a limited memory recursive least squares algorithm to improve the accuracy of online parameter identification. The results show that the proposed methods can achieve accurate SOC estimation under different temperatures, operating conditions, and initial SOC values. Therefore, model-based SOC estimation methods typically achieve higher accuracy. However, developing an accurate battery model necessitates expensive and time-consuming experimentation, along with substantial domain expertise and theoretical insight into battery systems [1]. Despite these efforts, limitations in prior knowledge of system dynamics inevitably lead to model imperfections. In recent years, with the continuous improvement in graphics processing unit (GPU) compute power, data-driven methods have garnered increasing attention [17]. Data-driven methods, also known as black-box models [12], rely on empirical observations and require minimal prior knowledge [18]. Data-driven methods heavily leverage available data during analysis, making them suitable for developing SOC estimation models in scenarios where detailed internal battery characteristics and chemical reactions are not known. Compared to model-based methods, data-driven approaches require less time and expertise to model complex nonlinear systems. For example, in the case of inaccurate initial SOC estimation, LSTM networks converge to the true SOC faster than unscented Kalman filters (UKF), achieving root mean square error (RMSE) and mean

absolute error (MAE) levels below 2 % and 1 %, respectively [19]. Furthermore, LSTM models can accurately estimate SOC by monitoring measurable values such as battery current, voltage, and temperature, without relying on detailed information about internal electrochemical reactions and related model parameters [20].

Recently, researchers have made significant advances in SOC estimation through data-driven approaches using neural network models. Tong et al. [21] proposed a neural network model for SOC estimation under three operating modes: idling, charging, and discharging. They utilized the US06 drive cycle for model training and pulse testing for model validation, achieving an average SOC error of 3.8 %. Chaoui et al. [22] evaluated a recurrent neural network (RNN) SOC estimation model under dynamic charge and discharge current profiles and temperature variations. Compared to a multilayer perceptron neural network approach, this model exhibited lower root-mean-square error and higher computational speed. Cui et al. [23] developed an intelligent SOC estimation model based on a wavelet neural network. Using the New European Driving Cycle (NEDC) conditions, this model achieved a mean absolute error (MAE) of 0.59 % and a maximum SOC error of 3.13 %. Yang [18] proposed a recursive neural network with gated recurrent units, which estimates battery SOC from measured current, voltage, and temperature signals. This method provides accurate estimates for data from two mainstream LIB types under dynamic operating profiles and is sensitive to temperature changes. Yang [24] also combined LSTM with UKF, achieving a root-mean-square error of <1.1 % and an average error of <1 %, even for temperatures not included in the training data set. These studies have focused on the generalization performance of models under various operating profiles and temperatures. The training and testing data were collected from adjacent charging and discharging cycles, without systematically considering the impact of battery aging. However, research by Chaoui et al. [22] and Kang et al. [25] has shown that the accuracy of SOC estimation models trained on fresh battery data gradually declines as the battery ages. Since SOC estimation models for practical applications are primarily trained on historical data, their accuracy inevitably decreases due to aging. Therefore, to maintain the accuracy of SOC estimation models throughout the entire battery lifecycle, it is necessary to track the performance degradation of these models. This would enable the timely update or retraining of neural network models, ensuring sustained accuracy of SOC estimation across the lifecycle of the LIB.

To address the challenge of SOC estimation due to battery aging, Mao et al. [26] proposed a joint estimation method for SOC and state of health (SOH), improving SOC accuracy by updating the battery capacity based on estimated SOH. Wu et al. [27] developed a robust LSTM model through feature selection and LSTM structure optimization, achieving a maximum absolute SOC error of 2.721 % when tested on real-world driving data under different temperatures and aging stages. Tao et al. [28] trained an RNN model using data from multiple aging stages, demonstrating its ability to closely approximate the true SOC even on unseen aging stages. Rezaei et al. [29] introduced a UKF method to model the uncertainty of batteries with unknown statistical characteristics, employing fuzzy control to correct SOC estimation as uncertainty increases, thereby mitigating the impact of model inaccuracies on state estimation.

Although promising, these studies face limitations such as challenges

in acquiring data across multiple aging stages and addressing the unknown complex aging pathways induced by variable EV operating profiles. Online corrections applied to model-based methods may lead to error accumulation and insufficient reliability. Addressing these concerns, Kang [25] proposed a straightforward correction approach involving regular battery capacity testing every 20 cycles, with adjustments to the capacity estimation model if errors exceed 3 % in two consecutive tests. While this periodic detection and correction of the capacity estimation model could theoretically be extended to SOC models, frequent testing in practical EV applications is challenging, as it may disrupt normal use due to a lack of efficient mechanisms for informed decision-making.

In summary, data-driven SOC estimation methods—particularly those based on recurrent neural networks such as long short-term memory (LSTM) and probabilistic models like Gaussian process regression (GPR)—have demonstrated notable accuracy and adaptability. However, existing approaches often operate in an open-loop fashion, lacking mechanisms to quantify uncertainty or dynamically adjust to evolving battery aging conditions [30]. As a result, their performance may deteriorate significantly over time, especially under complex aging paths and highly variable EV operating profiles such as the Worldwide harmonized Light vehicles Test Cycle (WLTC).

To overcome these limitations, we propose an innovative closed-loop SOC estimation framework that integrates active learning with uncertainty-aware correction into deep learning networks, enabling real-time feedback on SOC prediction confidence levels without the need for additional sensors or reference data. Specifically, we explore the use of various models—including Model Disagreement, greedy sampling, Gaussian Process Regression (GPR), and Bayesian Neural Networks (BNNs)—to develop active learning strategies for assessing and managing estimation uncertainty. This architecture offers a dynamic, data-driven mechanism to guide efficient model updates and maintain robust SOC estimation accuracy throughout the battery lifecycle. Our framework is validated using the complex WLTC profile, and demonstrates strong generalizability and reliability for SOC estimation in near real-world EV scenarios and nonlinear battery systems.

2. Correction strategies for data-driven SOC estimation models

2.1. SOC estimation based on LSTM-RNN

LSTM [20] is a variant of Recurrent Neural Networks (RNNs) [31] specifically designed for processing sequential data. Fig. 1a illustrates the concept of the hidden state in RNNs. The inputs at each time step from the previous time step in an RNN are collectively referred to as the "hidden state". This hidden state can be understood as the RNN's memory of information from the previous steps in the sequence. The state is updated and propagated as the time steps proceed. Fig. 1b depicts the architecture of an LSTM cell. The LSTM introduces three gates - input, forget, and output gates - to regulate the flow and retention of information, thereby mitigating issues like "vanishing gradient" and "exploding gradient". These gates allow the LSTM to selectively remember, forget, or update information over time steps, enabling it to capture long-term dependencies in sequential data more effectively. This architecture has proven highly effective in various sequence modeling tasks such as language modeling [32], machine translation [33], and speech recognition [34], establishing LSTM as a widely adopted neural network framework in deep learning.

The input gate, forget gate, and output gate can be represented as follows:

$$\mathbf{i}_k = \eta(\mathbf{W}_{\Psi i} \Psi_k + \mathbf{W}_{hi} \mathbf{h}_{k-1} + \mathbf{b}_i), \quad (2)$$

$$\mathbf{f}_k = \eta(\mathbf{W}_{\Psi f} \Psi_k + \mathbf{W}_{hf} \mathbf{h}_{k-1} + \mathbf{b}_f), \quad (3)$$

$$\mathbf{c}_k = \mathbf{f}_k \mathbf{c}_{k-1} + \mathbf{i}_k \tanh(\mathbf{W}_{\Psi c} \Psi_k + \mathbf{W}_{hc} \mathbf{h}_{k-1} + \mathbf{b}_c), \quad (4)$$

$$\mathbf{o}_k = \eta(\mathbf{W}_{\Psi o} \Psi_k + \mathbf{W}_{ho} \mathbf{h}_{k-1} + \mathbf{b}_o), \quad (5)$$

$$\mathbf{h}_k = \mathbf{o}_k \tanh(\mathbf{c}_k). \quad (6)$$

Here, the initial hidden state $\mathbf{h}_0$ is set to a zero matrix, and η represents the sigmoid function. The functions i , f , o , and c denote the input gate, forget gate, output gate and memory cell, respectively. They are referred to as gates because they are sigmoid functions that can be zero, thereby having the ability to suppress the flow of information to the next computational node. Each gate has its own set of network weights, denoted by $\mathbf{W}$. The subscript of $\mathbf{W}$ indicates the transformation between

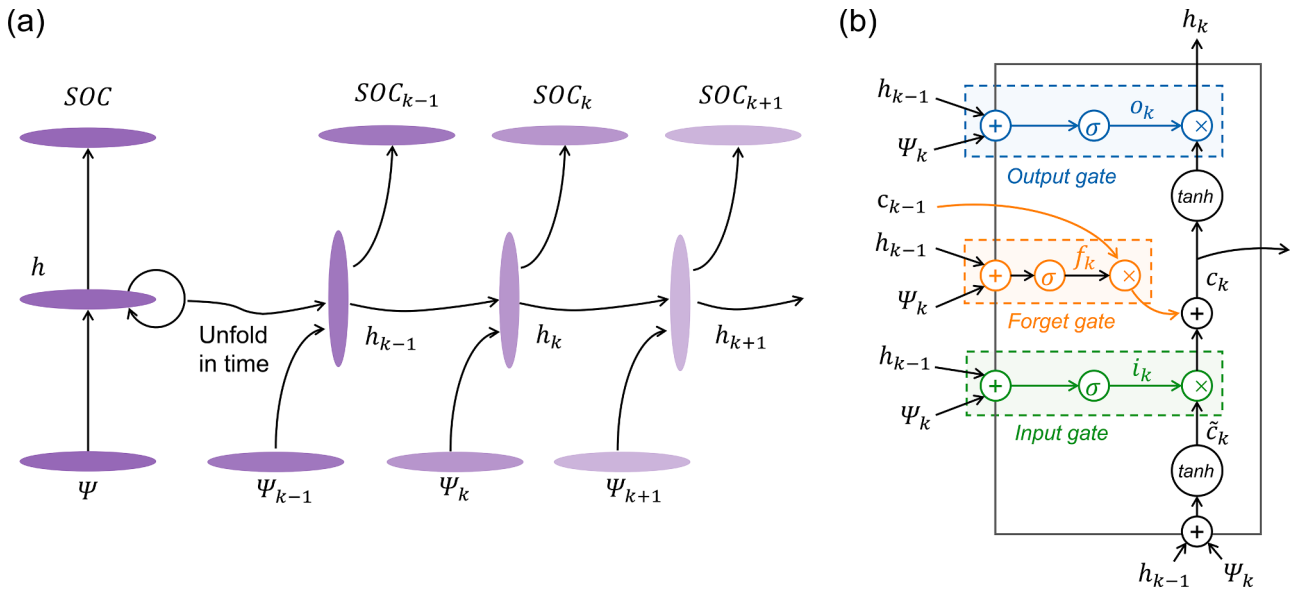


Fig. 1. (a) Architecture of an RNN (left) and architecture of an RNN unfolded in time (right). The input data is given by Ψ , and $\mathbf{h}_{k-1}$ denotes the hidden layer of the LSTM-RNN at time step $k - 1$. The output of the LSTM-RNN is the estimated SOC at each time step. (b) LSTM Cell Architecture, where $\mathbf{i}$, $\mathbf{f}$, $\mathbf{o}$, and $\mathbf{c}$ denote the input gate, forget gate, output gate, and memory cell, respectively.

components. For example, W_{ψ_0} is the input-output gate matrix, and W_{hi} is the hidden-input gate matrix. Each gate also includes a bias term b , enhancing the flexibility of the network in adapting to the data. The final fully connected layer performs a linear transformation on the hidden state tensor h_k to obtain the SOC estimate at time step k , as follows:

$$SOC_{lstm,k} = V_{out} h_k + b_y, \quad (7)$$

where V_{out} and b_y are the weight matrix and biases of the fully connected layer, respectively. For LSTM-based SOC estimation, the input data includes the measured charge voltage and current data over m time steps during the charging process. The collected data is processed by an RNN model, and the output is the estimated SOC at the final sampling point, which can be represented mathematically as:

$$\widehat{SOC}(k) = f_{RNN}(U(k-m+1, k-m+2, \dots, k), I(k-m+1, k-m+2, \dots, k)), \quad m \leq k \leq N, \quad (8)$$

where f_{RNN} represents the RNN model used to estimate SOC, and N is the final time step of the charging process.

2.2. Modeling SOC estimation across the battery lifecycle

During the data preparation stage, ensuring accurate SOC estimation across all SOC levels requires comprehensive coverage of charging information, from the lowest initial SOC to the deepest charging depth. Each cycle, defined as a consecutive charging and discharging process, constitutes a part of the battery lifespan until its end-of-life. Without loss of generality, let $X_n = [x_{n1}, x_{n2}, \dots, x_{nm}]$ denote data samples within the n^{th} cycle, where x represents voltage and current sequences measured over a time window, serving as input to the SOC estimation neural network. Each cycle sample represents a specific aging stage and enters the sample pool sequentially, akin to iterative sample extraction. Additionally, real-time labeling of cycle samples during online SOC estimation is impractical. Therefore, labeling is performed offline for the next cycle sample after correction decisions are made. Given the

approximately consistent degree of aging between adjacent cycles during the normal aging process, this approach is acceptable within this context.

Assuming we have labeled a cycle sample with true SOC values for training the SOC model, we proceed to iteratively extract samples from subsequent cycles as the battery undergoes charging and discharging cycles. While the trained SOC model initially exhibits short-term accuracy, continued battery aging renders predictions inaccurate for subsequently extracted cycle samples. To preempt SOC model failure, appropriate correction strategies are essential to assess whether model adjustments are needed for each cycle sample. Fig. 2 illustrates the lifelong SOC estimation process under correction strategy control, which will be discussed in the subsequent sections.

2.3. Preliminary correction strategies

This section introduces several simple and intuitive correction strategy methods. While these methods have some shortcomings, they are highly feasible and interpretable, making them suitable as reference baselines for active learning correction strategies.

Discharge capacity threshold correction strategy. This strategy involves tracking the total discharge capacity Q since the latest correction. If Q exceeds the threshold σ_Q , a correction is required. Then, Q is set to zero, and the next iteration begins. The discharge capacity is calculated based on the Coulomb counting method [35]:

$$Q = \int_{t_0}^t I(t) dt, \quad (9)$$

where t represents time, t_0 denotes the starting time, and $I(t)$ signifies the current function of t . Notably, the directions of current during the charging and discharging processes are different, which can lead to Coulomb counting cancellation. Therefore, only the discharging process

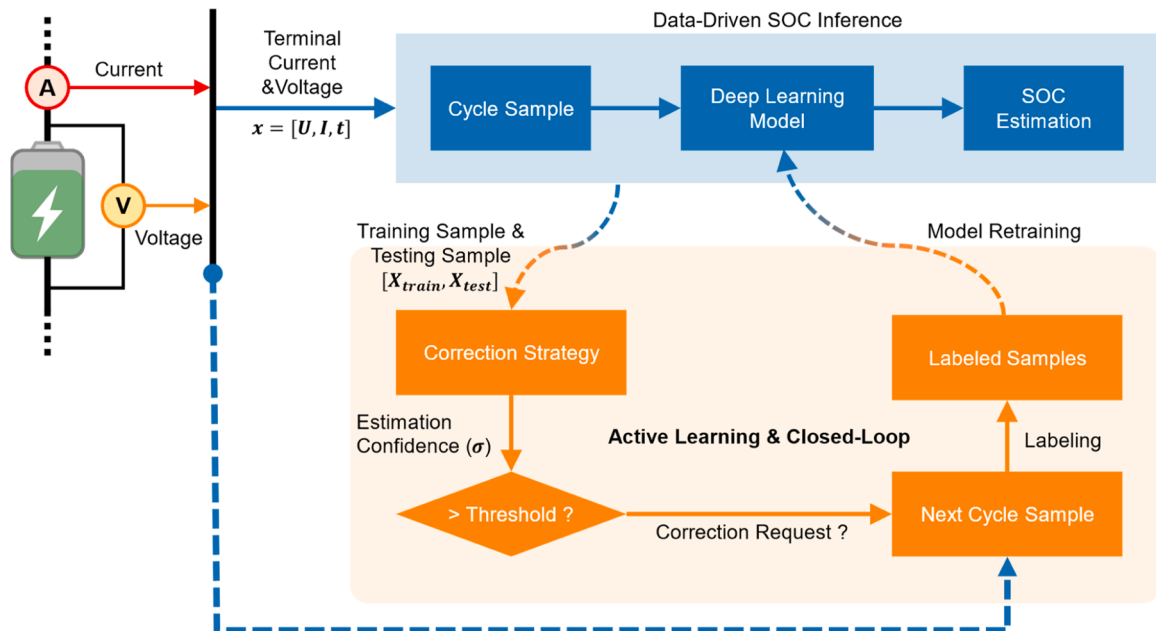


Fig. 2. Data-driven SOC estimation across the battery lifecycle under the guidance of an active learning and closed-loop correction strategy.

is considered. Additionally, in practical applications, this formula needs to be discretized to match the sampling frequency of the sensor.

Cycle count threshold correction strategy. This strategy statistically tracks the total number of cycles the battery has gone through since the latest correction. If the number of cycles exceeds the threshold σ_c , a correction is requested. This method is feasible in experimental settings where the depth of each charge and discharge cycle is relatively consistent. However, in real-world battery usage scenarios, the depth of each charge and discharge can vary significantly, leading to instability in the cycle count statistics and threshold-based correction.

SOH Loss Threshold Correction Strategy. Accurate SOH acquisition is crucial for this strategy and can be achieved through experimental measurement or estimation methods. Experimental measurement of SOH requires specialized discharge testing, which disrupts the normal battery use. Therefore, online estimation methods are recommended, such as training an SOH estimation model like the SOC estimation model, to enable continuous SOH acquisition. Let the SOH of the most recent correction cycle be denoted as SOH_0 . The ΔSOH for any cycle T can be defined as:

$$\Delta SOH_T = SOH_T - SOH_0. \quad (10)$$

If ΔSOH exceeds the threshold σ_h , a correction is requested. This strategy allows for flexible control over the number of corrections throughout the battery lifecycle and offers good interpretability. However, it requires a complete lifecycle dataset of SOH and related input features, along with a well-trained SOH estimation model.

2.4. Active learning-driven correction strategies

The preliminary correction strategies discussed earlier heavily rely on statistical accuracy and are vulnerable to cumulative statistical errors, potentially reducing effectiveness in real-world scenarios. Therefore, focusing on generalization capability, we leverage active learning to address the challenge of maintaining SOC estimation performance. Active learning, a specialized machine learning approach aimed at saving labeling costs and time [36], functions as a targeted search strategy widely applied across diverse research domains. For instance, active learning has been instrumental in identifying the most informative experiments in expansive biological networks to minimize experiment count [37], and in reducing parameter combination evaluations for simulation models [38]. In our context, active learning can be utilized to select the most informative samples to update or retrain the SOC estimation model. This framework decides whether to query the label of the newest sample or to predict it based on previous samples when processing upcoming samples. By strategically labeling and training samples during critical battery life stages, the model accuracy and robustness can be significantly enhanced. In the realm of active learning, maintaining SOC estimation performance resembles a stream-based scenario [36], where unlabeled samples are sequentially and iteratively sampled, rather than being accumulated into a pool all at once (pool-based scenario). Using active learning as a framework, one of the key challenges is modeling the uncertainty associated with estimation. A range of parametric and nonparametric methods have been explored to model uncertainty in power demand [39–41]. For instance, Abbasi [42] achieved a reliable and improved interpretation of power transformer frequency response traces by identifying key characteristics within these traces. However, these uncertainty modeling techniques cannot be directly applied to state of charge (SOC) accuracy, given the unpredictable aging processes of batteries and the time-series nature of their data streams. Herein, we introduce four approaches - Input Discrepancy, GPR Variance, Model Disagreement, and BNN Variance - to quantify the uncertainty inherent in SOC estimation.

2.4.1. Input discrepancy

The Input Discrepancy correction strategy is an active learning approach independent of the SOC model, characterized by low compu-

tational cost. The Input Discrepancy metric inherits the core concept of Greedy sampling on the inputs (GSx) [43,44] which prioritizes the selection of samples farthest from previously labeled samples in each iteration. Under the assumption of generality, let us assume there are k labeled samples. For each of the remaining $N - k$ unlabeled samples, GSx first computes their distances to the k labeled samples:

$$d_{nm}^x = \| \mathbf{x}_n - \mathbf{x}_m \|, \quad m = 1, \dots, k; \quad n = k + 1, \dots, N. \quad (11)$$

The minimum distance d_n^x of sample $\mathbf{x}_n$ to the k labeled samples is then determined as:

$$d_n^x = \min_m d_{nm}^x, \quad n = k + 1, \dots, N. \quad (12)$$

Finally, the sample with the maximum d_n^x is selected and requested for labeling.

GSx is not directly applicable to battery aging correction as it is designed for pool-based scenarios, while the aging correction involves a stream-based scenario. Therefore, we propose the Input Discrepancy metric based on GSx, which quantifies the correction importance of each cycle sample and applies it to stream-based active learning scenarios. Given the irreversible nature of aging, new samples hold greater value than older ones. Thus, during each correction, we replace old samples with newly labeled ones and directly retrain the SOC model with these updated samples. Furthermore, since each cycle sample consists of multiple sub-samples, the distance calculation between cycle samples should involve all sub-samples. Assuming the a^{th} labeled cycle sample X_a with N_a sub-samples and the b^{th} unlabeled cycle sample X_b with N_b sub-samples, we define the distance between each unlabeled sub-sample in X_b and each labeled sub-sample in X_a as follows:

$$d_{nm}^x = \| \mathbf{x}_n - \mathbf{x}_m \|, \quad m = 1, \dots, N_a; \quad n = 1, \dots, N_b. \quad (13)$$

The Input Discrepancy of X_b is calculated as:

$$D_{ba}^x = \frac{1}{N_b} \sum_{i=1}^{N_b} \min_m d_{im}^x, \quad m = 1, 2, \dots, N_a. \quad (14)$$

High Input Discrepancy indicates that the SOC model encounters unfamiliar input features, underscoring the substantial value of labeling such input samples.

2.4.2. GPR variance

GPR is a nonparametric model that leverages a Gaussian Process (GP) prior for regression analysis on data [45]. It offers advantages such as nonparametric modeling, probabilistic prediction, computational efficiency and robustness [46,47], making it widely employed in SOC and SOH estimation [48].

Given a training set $\mathcal{D} = ((\mathbf{x}_i, y_i), i = 1, 2, \dots, n)$ comprising n input-output pairs $(\mathbf{x}_i, y_i)$, where $\mathbf{x}_i$ represents measured voltage and charge difference sequences, and y_i denotes the corresponding SOC at the final sampling time step, we aim to compute the predictive distribution of the unknown observation y^* at input $\mathbf{x}^*$. Here, we define $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]^T$ and $\mathbf{Y} = [y_1, \dots, y_n]^T$. Assuming $y_i = f(\mathbf{x}_i) + \epsilon_i$, where $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ is an independent and identically distributed Gaussian noise, and modeling the output $\mathbf{f} = (f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_N))$ as a Gaussian random field $\mathbf{f} \sim \mathcal{N}(0, \mathbf{K})$, with $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$ representing the covariance kernel, the joint distribution of the training set $\mathcal{D}$ and the predicted test output $(\mathbf{x}^*, y^*)$ is:

$$\begin{bmatrix} \mathbf{Y} \\ \mathbf{y}^* \end{bmatrix} = \mathcal{N} \left(0, \begin{bmatrix} \mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma^2 \mathbf{I} & \mathbf{K}(\mathbf{X}, \mathbf{x}^*) \\ \mathbf{K}(\mathbf{x}^*, \mathbf{X}) & \mathbf{K}(\mathbf{x}^*, \mathbf{x}^*) \end{bmatrix} \right). \quad (15)$$

Given the training set, the mean prediction for the input $\mathbf{x}^*$ is:

$$\bar{\mathbf{y}}^* = \mathbf{K}(\mathbf{x}^*, \mathbf{X}) (\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma^2 \mathbf{I})^{-1} \mathbf{Y}. \quad (16)$$

The prediction variance is:

$$\Delta^2 = \mathbf{K}(\mathbf{x}^*, \mathbf{x}^*) - \mathbf{K}(\mathbf{x}^*, \mathbf{X}) (\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma^2 \mathbf{I})^{-1} \mathbf{K}(\mathbf{X}, \mathbf{x}^*). \quad (17)$$

The standard deviation of the prediction determines the width of the confidence interval, serving as a measure of uncertainty and a key variable for correction decisions. Averaging the standard deviations of the estimated values for all sub-samples within a cycle sample provides the GPR uncertainty for the entire cycle sample.

2.4.3. Disagreement-based methods

The disagreement-based query strategy focuses on selecting data points where there is disagreement among multiple models or experts [49]. In the Model Disagreement strategy, the learner is represented by a hypothesis space $\mathcal{H}$, which contains multiple hypotheses h , where each hypothesis $h \in \mathcal{H}$ represents an SOC estimation model that can map the input $\mathbf{x}$ to the output y . For simplicity, let us assume that the hypothesis space $\mathcal{H}$ contains only two hypotheses, though this setup can be generalized to more complex settings. In a regression problem, the output y is a continuous variable within the range $[0, 1]$. The error of a hypothesis h is given by

$$\epsilon(h) = |h(\mathbf{X}) - \mathbf{Y}|. \quad (18)$$

Define the pseudo-distance between two hypotheses as

$$d(h, h') = |\epsilon(h) - \epsilon(h')|, \quad (19)$$

The Model Disagreement ρ on the input $\mathbf{x}$ is defined as follows:

$$\rho_{h,h'}(\mathbf{x}) = |h(\mathbf{x}) - h'(\mathbf{x})|. \quad (20)$$

Here, the Model Disagreement ρ represents the magnitude of the difference in the predictions made by the different hypotheses for the same input. Specifically, this work uses LSTM and GPR as two contrasting models to produce Model Disagreement index. Model Disagreement directly contributes to the uncertainty in the predictions of the learner, as the learner must consider the outputs of all the hypotheses. When Model Disagreement is large, it becomes challenging to provide a final prediction that accommodates all hypotheses.

2.4.4. BNN variance

BNNs are stochastic artificial neural networks that utilize Bayesian inference [50]. Unlike traditional neural networks, which rely on point estimation methods, BNNs address the issue of trained models potentially generalizing unpredictably and overconfidently to data points outside the training distribution [51]. Such convincing yet incorrect results can be critical in real-world applications. To address this issue, BNNs simulate multiple possible computational pathways by assigning random probability distributions to the weights, thereby better quantifying the uncertainty inherent in the inference process. Figure S1 (see Supplementary Information) illustrates the distinction between Bayesian networks and conventional neural networks in weight calculation, highlighting that a more concentrated result distribution indicates lower uncertainty. This computational process can be summarized as follows:

$$\begin{aligned} \theta &\sim p(\theta), \\ y &= \Phi_\theta(\mathbf{x}) + \epsilon, \end{aligned} \quad (21)$$

where θ represents the network parameters, $p(\theta)$ is the probability distribution of θ , and ϵ denotes random noise, indicating that the function Φ provides an approximate estimation. Training a BNN using a training set D is equivalent to computing the posterior distribution $p(\theta|D)$ of the parameters θ . Assuming independence between the network parameters θ and the input, the posterior distribution of θ can be expressed as:

$$p(\theta|D) = \frac{p(D_y|D_x, \theta)p(\theta)}{\int_{\theta} p(D_y|D_x, \theta)p(\theta)d\theta} \propto p(D_y|D_x, \theta)p(\theta). \quad (22)$$

The Bayesian posterior of complex models such as artificial neural networks is a high-dimensional, highly non-convex probability distribution [52], making direct sampling from the posterior usually intractable. Therefore, variational inference [53] usually serves as a common

alternative. Variational inference is not an exact method [50]. Instead of directly sampling from the posterior, variational inference posits a variational distribution $q_\phi(\theta)$ parameterized by ϕ . The parameters ϕ are optimized through learning to align the variational distribution $q_\phi(\theta)$ as close as possible with the true posterior $p(\theta|D)$. The Kullback-Leibler (KL) divergence (D_{KL}) [54], rooted in Shannon's information theory for measuring the disparity between probability distributions [55], is employed to quantify its proximity:

$$D_{KL}(q_\phi \| p) = \int q_\phi(\theta) \log \left(\frac{q_\phi(\theta)}{p(\theta|D)} \right) d\theta. \quad (23)$$

To avoid the direct computation of $p(\theta|D)$, the Evidence Lower Bound (ELBO) is further derived as the loss function:

$$ELBO = \int q_\phi(\theta) \log \left(\frac{P(\theta, D)}{q_\phi(\theta)} \right) d\theta = \log(P(D)) - D_{KL}(q_\phi \| p). \quad (24)$$

Since $\log(P(D))$ only depends on the prior, minimizing $D_{KL}(q_\phi \| p)$ is equivalent to maximizing the ELBO. Approximating the ELBO using Monte Carlo simulations yields:

$$ELBO \approx \frac{1}{n} \sum_{i=1}^n \log q(\theta^{(i)}) - \log P(\theta^{(i)}) - \log P(D|\theta^{(i)}), \quad (25)$$

where n is the number of Monte Carlo samples, and $P(\theta^{(i)})$ represents the prior probability. We adopt the scale mixture prior distribution proposed by Charles [56]:

$$P(\theta) = \prod_j \pi \mathcal{N}(\theta_j | 0, \sigma_1^2) + (1 - \pi) \mathcal{N}(\theta_j | 0, \sigma_2^2), \quad (26)$$

where θ_j denotes the j^{th} parameter of the network, π is the scaling factor, and $\mathcal{N}$ represents the Gaussian probability density function. The standard deviation of the first Gaussian density (σ_1) is larger than that of the second (σ_2), specifically $\sigma_1 > \sigma_2$, and σ_2 should be much smaller than 1.

Variational inference provides a powerful mathematical framework for Bayesian inference, but its main challenge in deep learning lies in the stochastic nature that impedes gradient backpropagation through internal network nodes [57]. Bayes by Backprop resolves this challenge using a reparameterization technique. It introduces a random variable $\epsilon \sim q(\epsilon)$ as a source of non-variational noise. The network parameters θ are then obtained through a deterministic transformation $t(\epsilon, \phi)$, such that $\theta = t(\epsilon, \phi)$ adheres to the variational distribution $q_\phi(\theta)$. This approach ensures that ϵ , once sampled, remains constant relative to other variables, enabling deterministic computations and standard backpropagation.

When using BNNs for inference, we first sample the posterior distribution $p(\theta|D)$ of the network parameters to obtain $\Theta = \{\theta_i | i \in [0, N]\}$, where $\theta_i \sim p(\theta|D)$. Then, we obtain the predictive results by model averaging:

$$\hat{y} = \frac{1}{|\Theta|} \sum_{\theta_i \in \Theta} \Phi_{\theta_i}(\mathbf{x}). \quad (27)$$

Meanwhile, the standard deviation of these predictions is employed to quantify the uncertainty associated with the estimated results.

2.4.5. Threshold design and debouncing processing

Uncertainty is a unitless quantity derived from variance-based calculations and other metrics (as explained in the Eqs. 14, 17, 20, and 27). The uncertainty represents a relative measure rather than an absolute physical value. The uncertainty measures obtained above are compared against predefined thresholds to assess if corrective adjustments are necessary for the SOC estimation model under the current battery aging conditions. In real-world applications, these thresholds can be flexibly defined based on practical constraints, such as acceptable error levels

and retraining budget. Specifically, these thresholds can be empirically set based on data from batteries of the same model. Our findings indicate that applying consistent thresholds to batteries with diverse aging paths of the same model results in similar maximum SOC estimation errors after correction, supported by detailed experimental data in the Supplementary Information (see Figure S12a and Figure S13a). In addition, thresholds can be dynamically adjusted during battery use. For example, a model correction may be triggered immediately if the driver perceives discrepancies between the estimated remaining range and actual conditions, or when the vehicle mileage reaches a specific threshold. This momentary uncertainty measure then serves as the subsequent threshold.

However, relying solely on simple threshold-based judgments may lack stability in practical application, as outliers in uncertainty could easily surpass the threshold, leading to erroneous decisions. To mitigate this, we propose collecting multiple samples from adjacent cycles and implementing corrections when their uncertainty measures consecutively exceed the set threshold three times. This approach ensures that the correction strategy is only activated when the SOC estimation model uncertainty state is robust, thereby enhancing system reliability and resistance to interference.

3. Battery aging experiments

To conduct battery cycling aging tests, a commercial 18,650 battery was selected, with some key characteristic parameters listed in Table 1. The experimental profiles are divided into two main categories: simple operating profiles and complex operating profiles. The complex operating profiles are designed based on the World Light Vehicle Test Cycle (WLTC) [58]. The WLTC specifies vehicle speed variations within a test cycle, as shown in Figure S2a. This cycle is highly complex, encompassing low, medium, high, and extra-high-speed segments, which comprehensively cover a range of different speeds. Furthermore, the frequent acceleration and deceleration events in the low to medium-speed range closely resemble urban driving conditions.

The resulting current-time curve for an individual battery cell, based on the WLTC, is shown in Figure S2b, which was detailed in our previous work [59]. Notably, during vehicle deceleration, the regenerative braking system is considered, meaning the individual battery cell operation alternates between discharge and charge states. This design choice aligns with the actual usage of EVs and increases the complexity of the operating conditions, enhancing the distinction between the complex and constant-current conditions.

To facilitate a comparative analysis, three test conditions were designed: 1C constant-current, 3C-WLTC, and 6C-WLTC (hereafter referred to as 1C, 3C, and 6C, respectively). The charging process is identical for all three conditions, utilizing a 1C constant-current constant-voltage charging scheme. The differences lie in the discharge processes. The 1C condition follows a 1C constant-current discharge, while the 3C and 6C conditions are scaled versions of the current profile shown in Figure S2b, with maximum discharge rates of 3C (9 A) and 6C (18 A), respectively. All experimental conditions were designed in accordance with the battery specification sheet (see Supplementary Information-1) to ensure safe operation and to avoid cell damage. The specific settings for each condition are summarized in Table 2.

Table 1

Key parameters of the present 18,650 Li-ion battery.

Parameter	Value
Chemical component of cathode	LiNiCoAlO ₂ (NCA)
Chemical component of anode	Graphite
Nominal capacity	3000 mAh
Range of working voltage	2.5 ~ 4.2 V
Nominal voltage	3.7 V
Charging current at 1C	3 A

Table 2

Battery charge and discharge rates and corresponding current values for different operating profiles.

Test profile	Charge/discharge rate	Current magnitude
1C constant current	1C charge/discharge	$I = 3 \text{ A}$
6C-WLTC	1C charge, max 6C discharge	$I_{\max} = 18 \text{ A}$, $I_{\text{ave}} = 3.22 \text{ A}$
3C-WLTC	1C charge, max 3C discharge	$I_{\max} = 9 \text{ A}$, $I_{\text{ave}} = 1.61 \text{ A}$

It is worth noting that the current results should be interpreted as cycling-dominant worst-case validation under the current testing conditions, while slower calendar-driven drift is expected to be even better accommodated by the uncertainty-guided re-calibration mechanism.

4. Results and discussion

In practical EV scenarios, a significant absolute error can render the SOC estimation model inadequate for real-world requirements. This discrepancy becomes particularly problematic during critical battery states, such as near depletion, where a large error could lead to inaccurate mileage estimates, impacting both driving safety and user satisfaction. Therefore, minimizing the error throughout each correction period is essential. Consequently, in the results section, we implement SOC correction based on the cycle-based average error.

This section presents a systematic investigation of correction strategies designed to enhance the accuracy and robustness of battery SOC estimation models over their lifespan. We first explore several intuitive correction approaches based on battery degradation indicators, followed by a comparative analysis of uncertainty-driven strategies using different uncertainty quantification methods. Emphasis is placed on how these strategies interact with single-cycle and full-lifecycle models, and how uncertainty indicators can be leveraged to trigger timely corrections. The goal is to develop a correction framework that balances estimation accuracy with practical applicability under near-real-world usage conditions. The following subsections detail the deployment and performance of these strategies.

4.1. Correction based on preliminary correction strategies

Fig. 3a illustrates the distribution of overall cycle errors and correction points resulting from the preliminary correction strategies proposed in this paper. These strategies include adjustments based on discharge capacity (Q-count), cycle number (cycle-count), and SOH Loss thresholds. The test condition used is 3C, with the single-cycle SOC model trained on instantaneous cycle samples collected. Different correction strategies are represented by distinct colors for their respective curves and correction points. Thresholds designated for the three strategies ensure that the initial correction occurs at the 75th cycle, enhancing the clarity of comparative results. Comparing the strategies, the SOH Loss-based correction strategy yields fewer correction points compared to the fixed discharge capacity strategy. Both strategies effectively suppress maximum SOC estimation errors over the battery lifecycle. Excessive corrections not only increase costs but also impact the normal use of LIBs, making the correction point distribution of the SOH Loss strategy more reasonable. Additionally, compared to the fixed cycle number strategy, the fixed discharge capacity strategy shows progressively delayed subsequent corrections. This delay reflects the reduced maximum capacity of the battery and decreased discharge capacity per cycle as it ages.

Fig. 3b compares the performance of the single-cycle model and the lifecycle model, with the latter trained on full lifecycle data under 6C conditions. The lifecycle model demonstrates an overall lower error level throughout the lifecycle of the battery. In contrast, while the single-cycle model exhibits higher overall errors compared to the

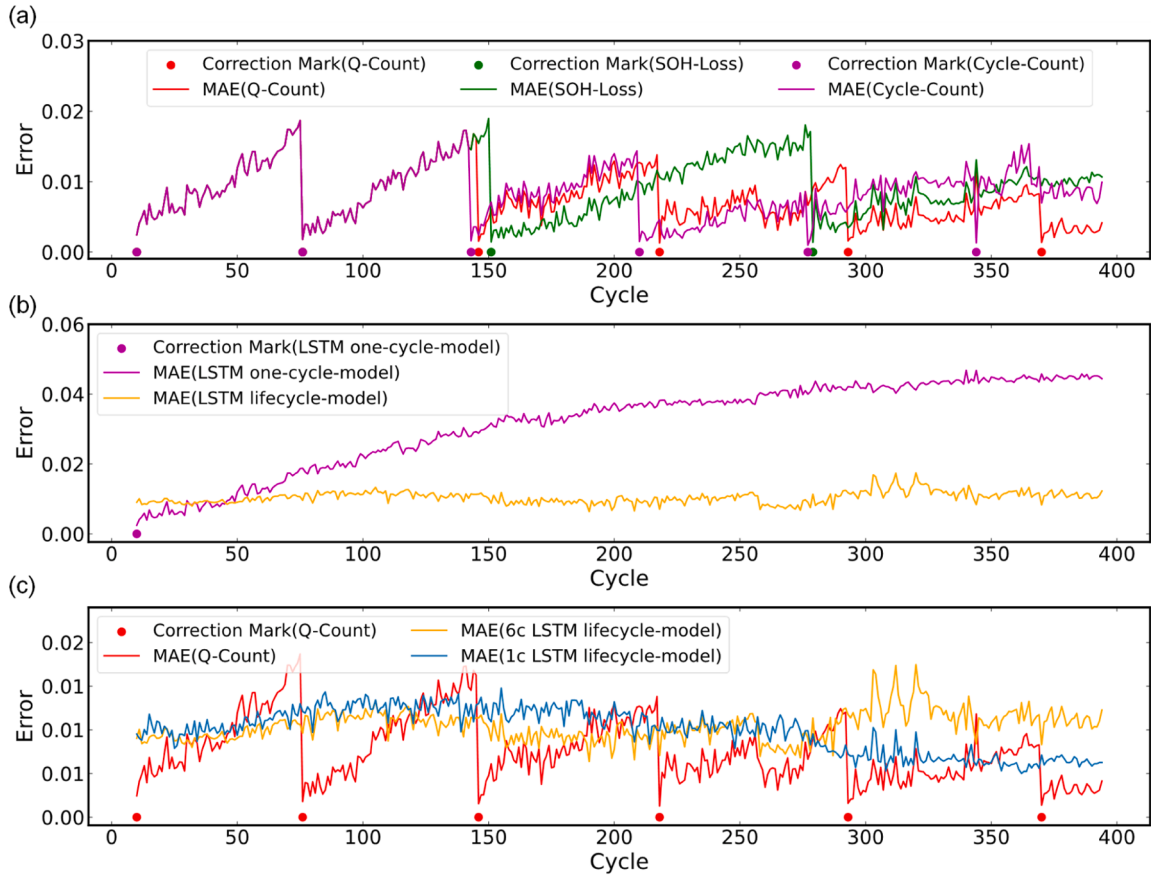


Fig. 3. Test results for the 3C scenario dataset. (a) Lifecycle error curve and correction point distribution based on the preliminary correction strategy. (b) Performance comparison between the lifecycle model and the single-cycle model without any correction. (c) Performance comparison between the lifecycle models and the single-cycle model with corrections advised by Q-count strategy.

lifecycle model, it achieves lower error levels in the few adjacent cycles following a correction. This "adjacent advantage" can be more clearly observed from Fig. 3c: the error for Q-count correction strategy is lower most of the time compared to both the latest lifecycle models trained on the 1C and 6C lifecycle datasets [3]. It can be inferred that with a proper correction strategy and a few corrections, a single-cycle model can outperform a lifecycle model.

It is worth noting that the lifecycle model requires a broad range of training data, and when its performance deviates, effective correction becomes challenging. Conversely, the single-cycle model offers greater flexibility, requiring only one correction to quickly reduce the SOC error to an ideal level. By combining the single-cycle model with a reasonable correction strategy, the entire battery lifecycle can be divided into multiple stages, with a specific SOC estimation model applied in each stage. Consequently, this approach can be expected to maintain good SOC estimation accuracy throughout the lifecycle.

4.2. Uncertainty curves across a full lifecycle without correction

Fig. 4 illustrates the four uncertainty indicators proposed in this paper and their variations with battery aging without correction. Due to the differing scales of the variables, dual y-axes are used to display their changing trends, with the error coordinates on the left and the uncertainty indicator coordinates on the right. The uncertainty indicators are represented by blue curves in each subplot, while the error curves are shown in red. Each blue correction point represents a sample labeling and model training instance. The four indicators are Input Discrepancy, GPR Confidence Intervals, BNN Variance, and Model Disagreement. We observe that the SOC estimation error increases with the cycle count,

and the four indicators exhibit similar growth trends to the SOC estimation error, implying that these indicators can effectively reflect error changes when the error is unknown. The Input Discrepancy indicator and GPR Variance show relatively stable upward trends, followed by Model Disagreement, while the BNN Variance is the most unstable, displaying large oscillations. The stability of these uncertainty indicators is crucial for correction decision-making, as excessive oscillations can obscure the error information they contain. Therefore, filtering the BNN Variance may be necessary before use.

4.3. Input discrepancy correction

Fig. 5 presents the error and uncertainty curves across the entire lifecycle under the Input Discrepancy correction strategy. The infinite threshold ($\sigma = \infty$) denotes the scenario in which no correction is applied throughout the battery life cycle. The other thresholds, for better demonstrating the correction effect over the entire lifecycle, are determined by analyzing the variance trend of the uncertainty indicators across cycles under the no-correction baseline (see Fig. 4). This approach allows us to select a meaningful threshold corresponding to the point where the first correction would occur, thereby enabling a consistent framework for evaluating performance improvements.

Multiple correction points segment the lifecycle into distinct correction stages. The blue curve (Input Discrepancy) and the red curve (error) exhibit similar trends. Each correction results in a significant reduction in both error and uncertainty, while slightly altering the trends of these variables. This change leads to a poor alignment of the two curves, suggesting a deviation in their correlation. As the Input Discrepancy (uncertainty) threshold is lowered, the first correction is

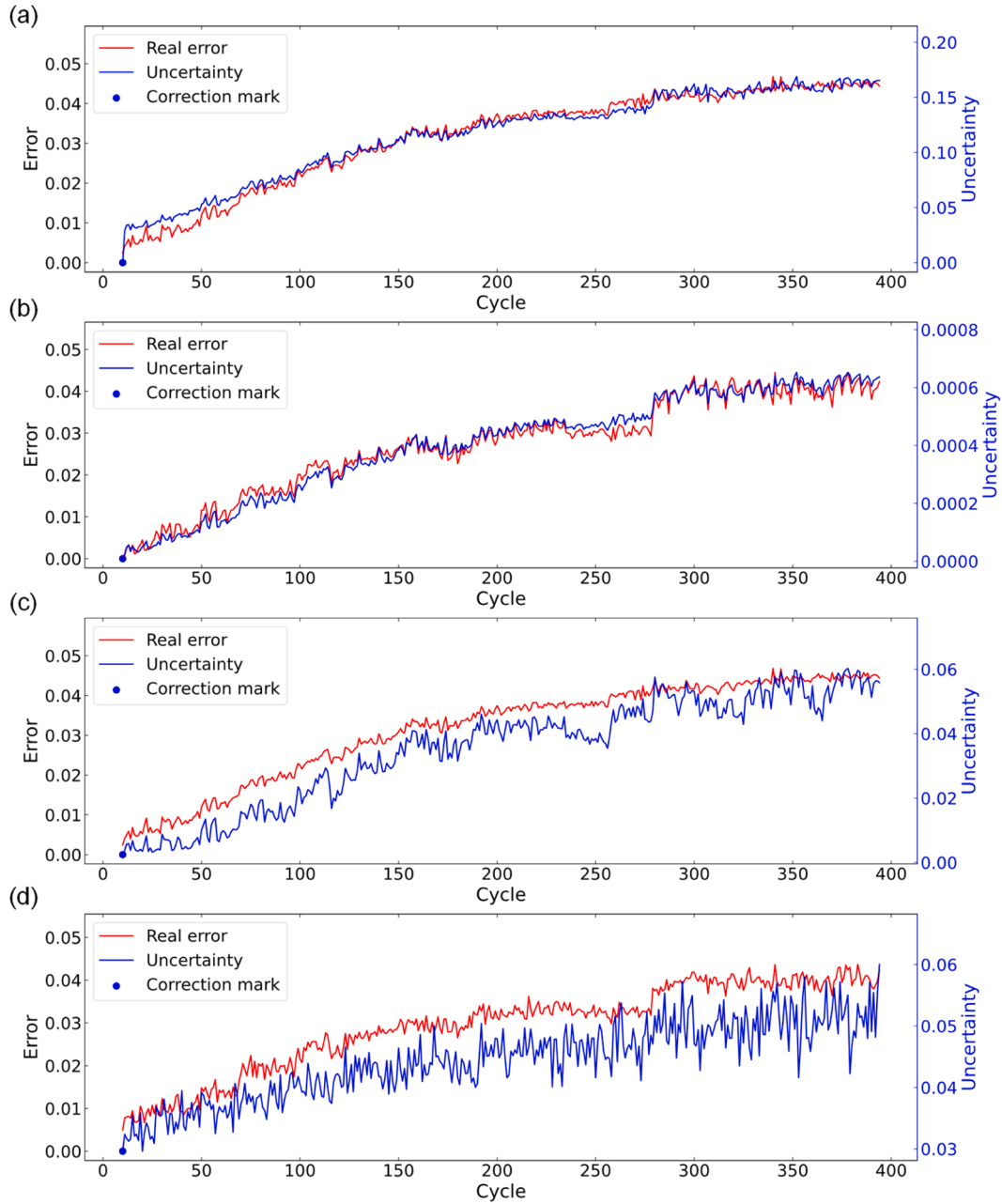


Fig. 4. Variation curves of uncertainty indicators with battery aging for (a) Input Discrepancy; (b) GPR Confidence; (c) Model Dispersion; (d) BNN Variance. Red error curves use the left axis, while blue uncertainty curves use the right axis.

triggered earlier, and the maximum error in the initial correction stage is accordingly reduced. Notably, the uncertainty for the first cycle sample after each correction consistently reaches 0.03, rather than approaching a value closer to 0. This may be due to voltage and current signal measurement errors, as two sets of sensors cannot produce completely identical outputs on the same signal snippet. This strategy effectively controls SOC maximum cycle-based average error to a close level of 0.02, especially for the first and second corrections, demonstrating its efficacy.

Fig. 6 illustrates the relationship between average error and average uncertainty per cycle under the Input Discrepancy correction strategy with varying correction thresholds. The differing data point colors distinguish the correction stages. The data points within the same correction stage align closely around a straight line, with similar regression line slopes across stages, especially for correction #1 and #2, as shown in Fig. 6a and b. This stability in the linear relationship

between error and uncertainty throughout the lifecycle suggests minimal impact from model corrections. In this scenario, establishing a consistent threshold for each correction stage can effectively control the maximum error within each stage, ensuring they remain similar.

Fig. 7 displays the distribution of life-long errors and uncertainties under the Input Discrepancy correction strategy with different correction thresholds. We observe that both error and uncertainty increase with a higher threshold throughout the lifecycle. The uncertainty remains free of outliers and shows a concentrated distribution, indicating its stability and utility for downstream calculations and comparisons.

4.4. GPR confidence interval-based correction

Figure S3 shows the lifecycle error and uncertainty curves under the GPR confidence correction. As can be seen in Figure S3a, in the first correction stage, the two curve segments exhibit significant differences,

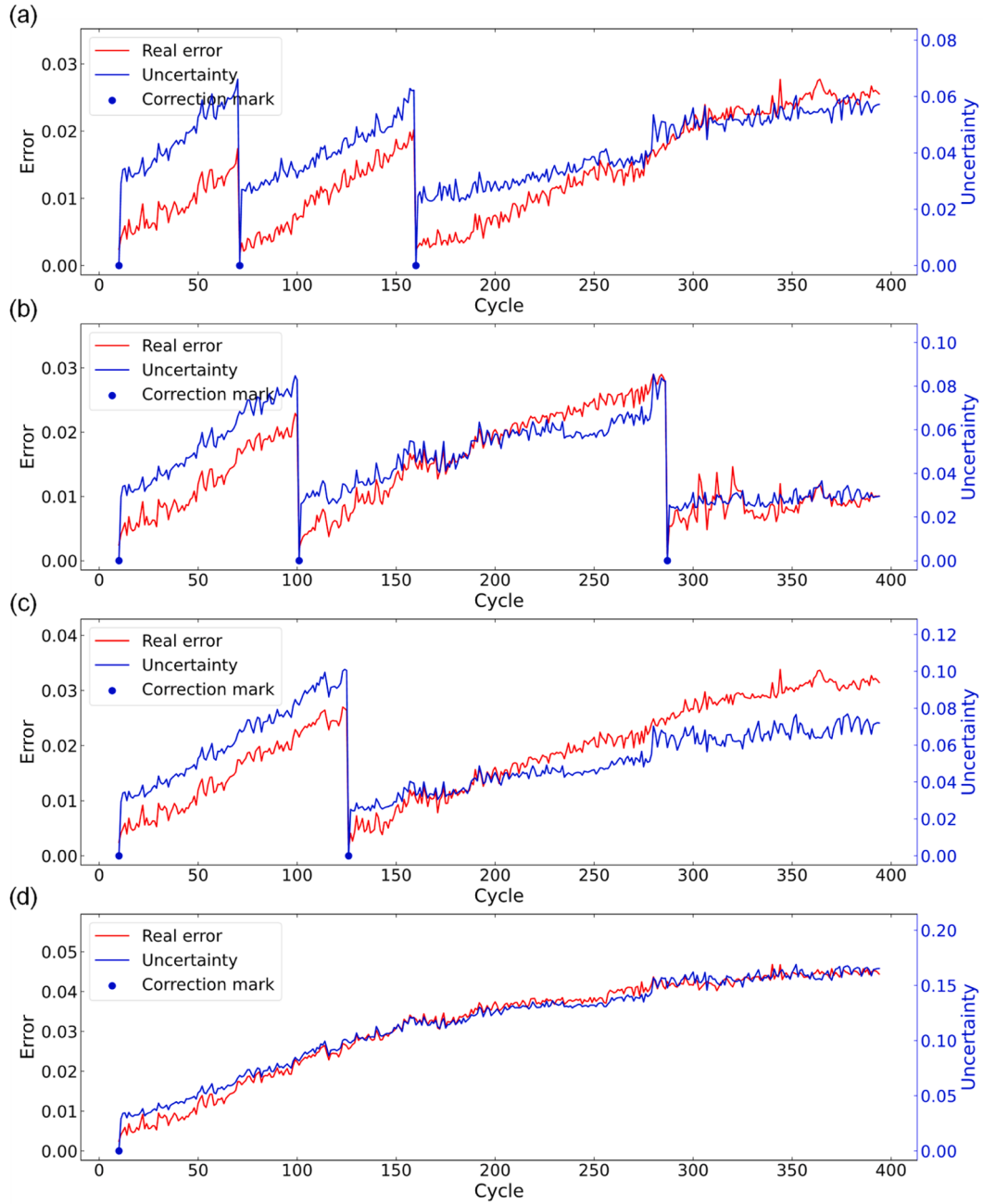


Fig. 5. Error and uncertainty curves throughout the lifecycle under the Input Discrepancy correction strategy (dual y-axis) for thresholds of (a) $\sigma = 0.06$; (b) $\sigma = 0.08$; (c) $\sigma = 0.1$; and (d) $\sigma = \infty$. Red error curves use the left axis, while blue uncertainty curves use the right axis.

whereas the subsequent correction stages show much closer alignment. Specifically, the maximum cycle-based average error in the first stage reaches 0.014, compared to approximately 0.005 in the later stages. This indicates a notable shift in the trend of GPR confidence and error post the initial correction. Consequently, the threshold correction based on GPR confidence fails to effectively control the maximum error within each correction stage and the timing of corrections throughout the lifecycle. For instance, the first stage correction occurs too late, allowing the SOC estimation error to grow too large, potentially failing to meet usage requirements. Additionally, as the battery nears the end of its lifecycle, GPR confidence shows large fluctuations, triggering frequent correction requests. These unnecessary corrections only increase the SOC model maintenance cost without significantly improving performance, contradicting the original intent of the correction strategy. Therefore, GPR confidence may not be suitable for reflecting the SOC

model performance.

Figure S4 further illustrates the significant shift in the linear relationship between GPR confidence and error. The regression line of the scatter points in the first correction stage deviates notably from other stages, and except for this stage, the scatter points in other correction stages are relatively concentrated. This matches the observations in Figure S3, implying that the rapid decline in maximum capacity early in the battery life (as shown in Figure S5, Supplementary Information) may negatively impact the calculation of GPR confidence.

Despite the imperfect linearity between GPR confidence and error, reducing the threshold results in decreased overall GPR confidence and error throughout the lifecycle, as shown in Figure S6. The outliers in the box plot in Figure S6a are mainly due to the significant fluctuations in GPR confidence towards the end of the battery lifecycle. The error outliers in Figure S6b mainly arise from the first correction stage, where

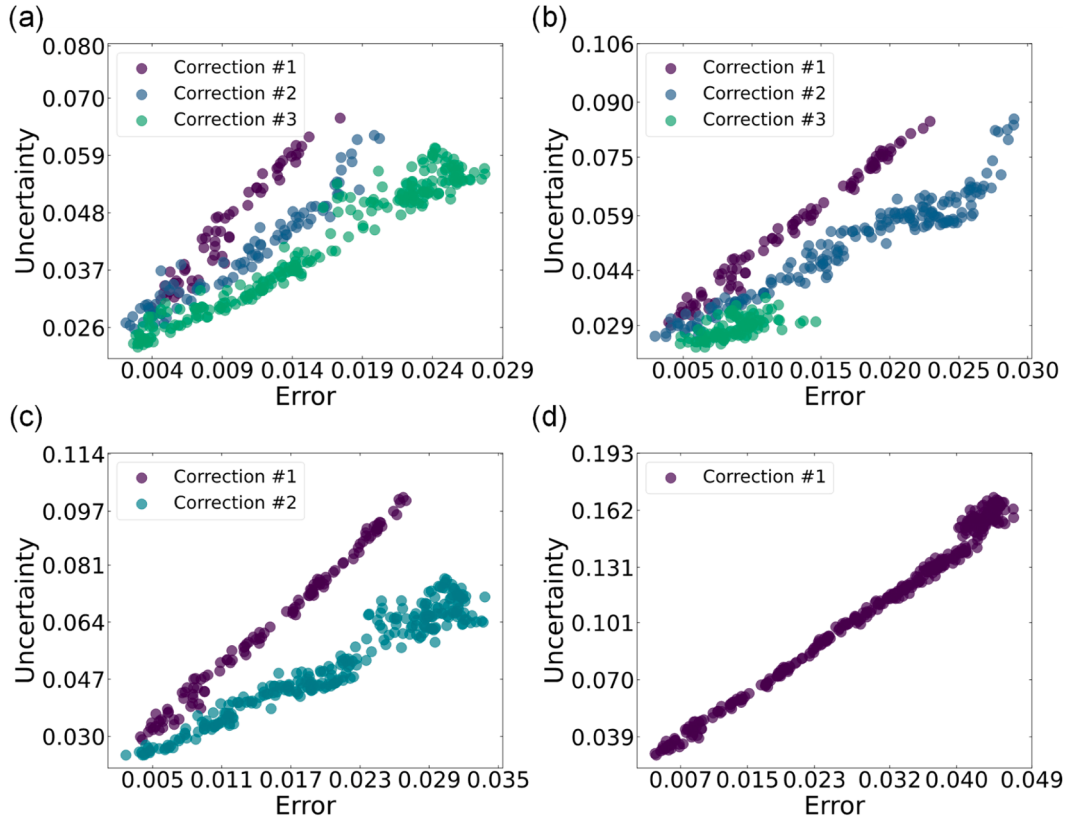


Fig. 6. Scatter plots of the correlation between Input Discrepancy and errors after correction using a threshold of (a) $\sigma = 0.06$; (b) $\sigma = 0.08$; (c) $\sigma = 0.1$; and (d) $\sigma = \infty$.

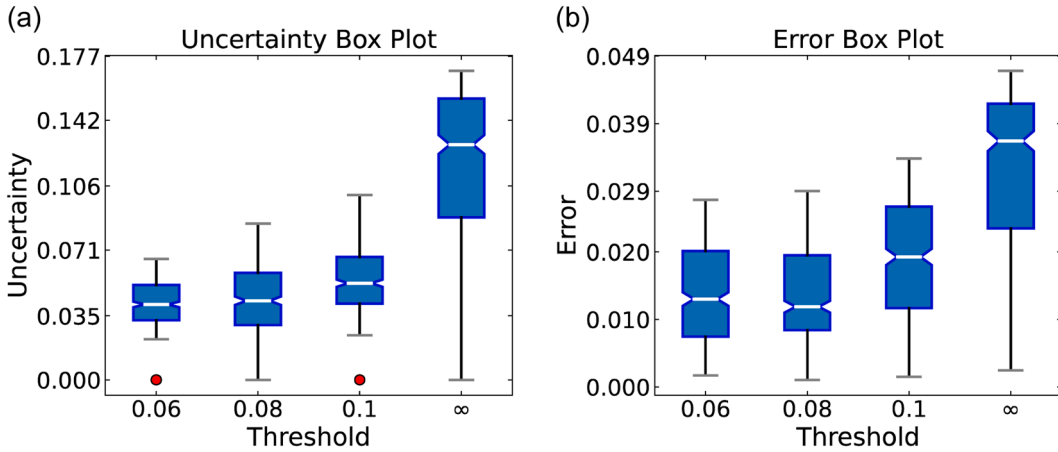


Fig. 7. Distribution box plots of the (a) input discrepancy throughout the lifecycle; and (b) SOC estimation errors throughout the lifecycle.

the error had already reached a relatively high level. However, after the first correction, the correction strategy becomes more aggressive, maintaining lower error levels. As a result, the large error samples from the first correction stage contribute to the outliers above the upper bound of the box plot.

4.5. Model disagreement-based correction

Figure S7 shows the life-long error and uncertainty curves under the Model Disagreement correction strategy. The curves across each three threshold correction results show similar trends, with error and uncertainty well-fitted. As the threshold increases, the number of corrections also rises, which effectively reduces the maximum error. In Figure S7a-c,

the error at each correction point within the same lifecycle is nearly identical, indicating a stable positive correlation between uncertainty and error. This stability allows the strategy to effectively control the maximum error. For example, guided by the Model Disagreement-based correction strategy with an uncertainty threshold of 0.015, the LSTM network requires only four retraining sessions over the entire battery lifecycle to keep the average SOC estimation error below 1.5 % and the maximum cycle-based average error below 2 %. Notably, Model Disagreement remains relatively stable throughout the aging cycles, with very few outliers. As seen in Figure S7c, there is an abnormal fluctuation in Model Disagreement between 150 and 200 cycles, which exceeds the preset threshold multiple times. However, thanks to the high fault tolerance of the proposed threshold debouncing method for

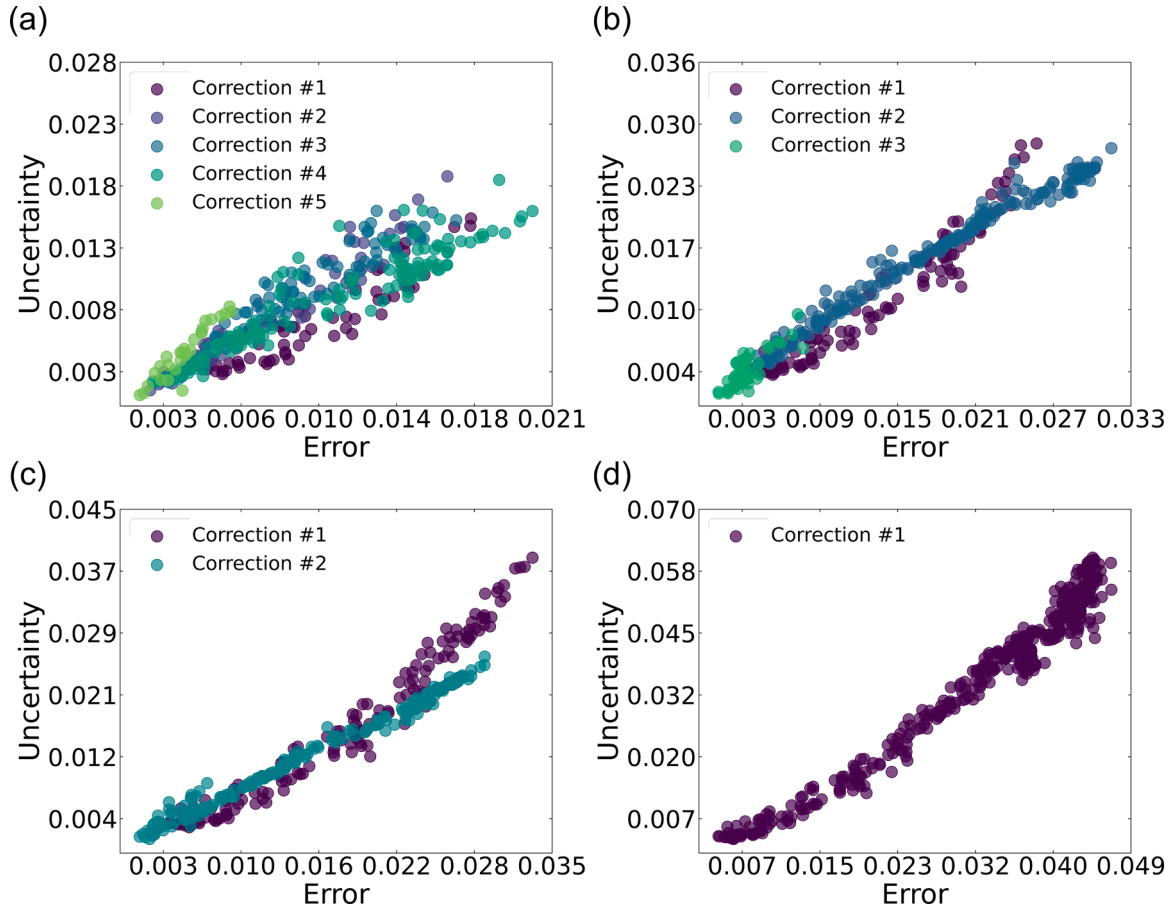


Fig. 8. Scatter plots of the correlation between Model Disagreement and errors after each correction using a threshold of (a) $\sigma = 0.015$; (b) $\sigma = 0.025$; (c) $\sigma = 0.035$; and (d) Threshold $\sigma = \infty$.

sporadic abnormal values, the correction decision remains unaffected by these anomalies.

Fig. 8 presents similar results from the perspective of point distribution. In Fig. 8a-c, the scatter points in each correction stage are tightly clustered along the trend line, with trend lines between correction stages nearly overlapping. This indicates a very stable linear correlation between Model Disagreement and error, suggesting that corrections have not significantly impacted this correlation.

The consistent and stable linear relationship between Model Disagreement and error across correction stages demonstrates that this strategy can effectively control the maximum error by leveraging the reliable correlation. This desirable property allows the correction approach to maintain the SOC model performance throughout the battery lifecycle.

Figure S8 shows that the median distribution of Model Disagreement and errors clearly increases with higher thresholds, indicating robust control over errors by the Model Disagreement correction strategy. By appropriately selecting the Model Disagreement threshold, this approach can provide suitable feedback when the SOC neural network error increases, thereby effectively managing the desired maximum SOC error.

4.6. BNN variance-based correction

Figure S9 presents the error and uncertainty curves throughout the lifecycle under BNN Variance correction. In Figures S9a and b, the uncertainty and error curves are well-aligned and exhibit a similar growth trend. In Figure S9a, the last two corrections occur one after another closely after the third-to-last, despite the corresponding error levels

being relatively low. This is primarily due to significant fluctuations in uncertainty. When peaks of uncertainty occur consecutively, the proposed threshold debouncing method becomes less effective. Despite the instability in BNN Variance, the metric always fluctuates around the true error curve (Figure S9a-c), indicating a certain ability to reflect error. In cases where uncertainty thresholds are slightly raised (Figure S9b and c), the maximum cycle-based average errors across all correction stages remain relatively consistent, reaching 0.025 and 0.03, respectively, depending on the thresholds.

Similarly, Figure S10 shows that the scatter distribution of BNN Variance and error correlation has a high degree of variability, making it difficult to discern a clear regression line. However, despite the scatter distribution's discreteness, there is a relatively large overlap between different correction stages, indicating that the scatter distribution of the correlations between stages has not significantly shifted due to the corrections. This is a desirable property, as it suggests that the correlations are minimally affected by the corrections, which is advantageous for the universality of the scale across different correction stages.

Figure S11 illustrates the distribution of SOC estimation error and BNN Variance over the lifecycle. As indicated in Figure S11a, changes in the threshold do not significantly affect the distribution of uncertainty. However, Figure S11b reveals that the result for the threshold of 0.0375 shows a notable difference compared to 0.04 and 0.043. Referencing Figure S11c, the reduction in error is primarily due to multiple unnecessary corrections. These results suggest that the practical application effect of BNN Variance correction is unsatisfactory. However, this behavior is not due to convergence problems; the BNN generally maintains a low average estimation error (<2 %) across most cycles, suggesting stable prediction performance. The oscillations in BNN

Variance primarily stem from inherent stochasticity in the variational inference process and the limited number of Monte Carlo samples (100 in this study), which was chosen as a trade-off between computational cost and practical feasibility for real-time applications.

4.7. Linear correlation and supplementary validation

Table 3 documents the correlation coefficients between uncertainties and errors under various correction strategies, reflecting the degree of association between these variables. The correlation coefficient metrics analyzed include Pearson correlation coefficient [60], Spearman correlation coefficient [61], and Kendall correlation coefficient [62]. As depicted in Table 3, the Model Disagreement correction with a threshold of 0.035 exhibits the highest values across all three types of correlation coefficients, indicating a robust correlation between uncertainty and error throughout the lifecycle. The Pearson coefficient assesses the strength of linear relationships between continuous variables. In this context, the Model Disagreement correction strategy produces Pearson coefficients exceeding 0.9, indicating a strong positive linear correlation. The Input Discrepancy strategy demonstrates slightly lower Pearson coefficients, yet still around a respectable level of 0.8. The Spearman correlation coefficient evaluates rank correlations, capturing monotonic relationships beyond linear patterns. Here, the Model Disagreement and Input Discrepancy correction strategies perform best, while the BNN Variance strategy shows weaker performance, particularly with a smaller threshold of 0.0375, possibly due to significant fluctuations in BNN Variance. The GPR Confidence exhibits notable fluctuations in correlation with errors post-initial correction, resulting in poorer correlation coefficient metrics. Together, the Model Disagreement strategy exhibits superior error characterization capabilities, maintaining a stable linear relationship minimally impacted by corrections, thereby enhancing the applicability of this correction approach.

To further verify the generalization performance of the method, we adopted the public dataset provided by Severson et al. [63], where commercial LFP/graphite cells (A123 Systems, model AP18650M1A, 1.1 Ah nominal capacity) were cycled in a temperature-controlled environmental chamber (30 °C) under varied fast-charging conditions but identical discharging conditions (4 C to 2.0 V, where 1 C is 1.1 A). This dataset differs from ours in chemistry (LFP vs. NCA), charging protocol (custom multistage fast charge vs. standard CC/CV), and lower sampling rate. For verification, we applied the Model Disagreement correction strategy that performed best on our own data. As can be observed from the error/uncertainty curve shown in Figure S15a, with one initial training and only 3 subsequent corrections, the lifespan SOC estimation error can be controlled lower than 4 %. Figures S14b-d likewise demonstrates a strong correlation between the Model Disagreement uncertainty and the error. The strategy efficiently and

uniformly limits the peak error at each stage, mirroring the results obtained on our in house dataset and suggesting that the method can be generalized to near-real world usage conditions.

Building on the solid results and discussions, our active learning-driven error control strategy demonstrates significant potential for on-line data-driven SOC estimation throughout the lifecycle of LIBs in EVs. EVs, with fewer mechanical components and a simpler drivetrain than internal combustion engine vehicles, have extended maintenance intervals and typically lower maintenance costs. According to the American Automobile Association (AAA), EVs may undergo maintenance annually for regular services such as wheel alignment and brake services [64]. However, due to the complexity of actual operating conditions, the accuracy of data-driven SOC estimation is highly likely to deteriorate prematurely (as shown in Fig. 3b). Our proposed correction strategy enables the automatic identification of this issue for the first time, allowing for smart and timely remedies within closed-loop error detection sessions. Specifically, when deterioration in SOC estimation performance is detected, the maintenance schedule can be flexibly adjusted, either brought forward or delayed. During the subsequent maintenance, a deep charging process is performed, during which common professional sensors can be used to collect the training samples (such as voltage and current curves) required for the data-driven SOC model training. This SOC model correction process only requires the duration of a normal charging session, without significantly increasing the maintenance time or cost. Moreover, our results indicate that the Model Disagreement correction strategy enables robust and high-precision SOC estimation with only four standard corrections across the battery's WLTC lifecycle. Given that the design lifetime of LIB packs in EVs ranges from 6 to 15 years [65], our suggested correction frequency (four times per lifetime) or annual correction would be sufficient for maintaining data-driven SOC estimation performance in real-world EV scenarios. In summary, the innovative active learning and closed-loop correction strategy proposed in this research advances data-driven SOC estimation in EVs in terms of feasibility, reliability, and flexibility.

While our framework demonstrates high flexibility and robustness, we recognize that its performance in overcharging and dynamic temperature scenarios and the impact of cell balancing has yet to be evaluated, primarily due to the lack of corresponding data. We plan to address these in future work by expanding the dataset to include overcharging conditions, thus broadening the applicability of active-deep-learning SOC estimation. We also plan to extend the framework to a module-level or pack-level system, where balancing effects and inter-cell dependencies will be explicitly modeled and addressed in conjunction with SOC estimation and uncertainty correction. Additionally, while we have leveraged data from WLTC profiles to simulate complex driving conditions, there remains a gap between the test data and real-world profiles due to factors such as random driving behaviors and unpredictable charge-discharge cycles. Future studies will further explore the scalability and reliability of our framework under these realistic conditions. Despite these limitations, we believe our framework provides valuable insights into active-deep-learning SOC estimation, offering a flexible and robust solution that sets the foundation for advanced applications in complex EV operating profiles.

5. Conclusions

We systematically evaluate the error trends of mainstream RNN-based SOC estimation models across the entire lifecycle of lithium-ion batteries, using the complex World Light-duty Test Cycle (WLTC) profiles that reflect real-world electric vehicle operation. By analyzing both lifecycle-level and single-cycle training strategies, we reveal that data-driven SOC estimation models experience uncontrolled degradation in accuracy as battery aging progresses. To address this critical limitation, we propose an innovative closed-loop SOC estimation framework powered by active learning, which enables real-time confidence assessment of model predictions without requiring additional reference sensors or

Table 3
Correlation coefficients between correction strategy errors and uncertainties.

Strategy	Threshold	Pearson	Spearman	Kendall
Input discrepancy	0.06	0.84	0.87	0.68
Input discrepancy	0.08	0.87	0.86	0.66
Input discrepancy	0.10	0.77	0.80	0.62
Input discrepancy	∞	0.99	0.99	0.93
GPR variance	0.00010	0.45	0.61	0.44
GPR variance	0.00020	0.47	0.50	0.36
GPR variance	0.00027	0.82	0.84	0.65
GPR variance	∞	0.99	0.99	0.91
Model disagreement	0.015	0.91	0.91	0.74
Model disagreement	0.025	0.98	0.98	0.88
Model disagreement	0.035	0.96	0.98	0.89
Model disagreement	∞	0.98	0.97	0.87
BNN variance	0.0375	0.11	0.06	0.04
BNN variance	0.0400	0.53	0.53	0.37
BNN variance	0.0430	0.63	0.62	0.43
BNN variance	∞	0.89	0.88	0.70

experiments. Our proposed architecture features a Model Disagreement-based active-learning correction strategy that continuously monitors prediction uncertainty and triggers retraining only when necessary. Experimental validation under WLTC conditions and a fast-charging protocol demonstrates that the proposed strategy effectively maintains estimation robustness throughout the battery lifecycle. Notably, the Model Disagreement index achieves a high Pearson correlation coefficient of 0.91 with the actual SOC estimation error, confirming its strong capacity to reveal true prediction uncertainty. For the WLTC dataset, leveraging the proposed mechanism, the LSTM model requires only four retraining sessions across the entire aging cycle to keep the average SOC error below 1.5 % and the maximum cycle-based average error below 2 %. For the fast-charging dataset, the model achieves robust performance with one initial training and only 3 subsequent corrections, keeping the lifetime SOC estimation error below 4 %.

These results underscore the feasibility and effectiveness of our closed-loop, uncertainty-aware architecture in extending the reliability and adaptability of data-driven SOC estimation for EVs under complex and evolving operating conditions. Future work will investigate the generalizability of the proposed framework across different neural network backbones, hyperparameter settings, battery models, environmental temperatures, and real-world EV datasets with calendar aging and cell balancing, to enhance its robustness and applicability. These efforts will further support the validation of the generalization capability of our proposed active-learning-driven error control architecture for data-driven SOC estimation.

CRedit authorship contribution statement

Junwei Xue: Writing – original draft, Validation, Methodology, Formal analysis, Conceptualization, Writing – review & editing, Visualization, Software, Investigation, Data curation. **Xuzhi Du:** Writing – original draft, Validation, Software, Project administration, Investigation, Formal analysis, Conceptualization, Writing – review & editing, Visualization, Supervision, Resources, Methodology, Funding acquisition, Data curation. **Lei Zhao:** Software, Data curation, Methodology. **Zhigang Yang:** Writing – review & editing, Supervision, Resources, Methodology, Formal analysis, Writing – original draft, Software, Project administration, Funding acquisition, Conceptualization. **Chao Xia:** Writing – review & editing, Methodology, Resources, Investigation. **Yuan Ma:** Resources, Writing – review & editing, Methodology. **Muhammad Jahidul Hoque:** Writing – review & editing, Investigation, Methodology. **Wuchen Fu:** Software, Investigation, Writing – review & editing, Methodology, Formal analysis. **Xiao Yan:** Writing – review & editing, Methodology, Formal analysis, Software, Investigation. **Nenad Miljkovic:** Writing – original draft, Resources, Investigation, Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

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.

Acknowledgements

J.X., L.Z., and Z.Y. gratefully acknowledge funding support by the National Key R&D Program of China (Grant No. 2022YFE0208000), the Shanghai Key Laboratory of Aerodynamics and Thermal Environment Simulation for Ground Vehicles (Grant No. 23DZ2229029), and the Shanghai Automotive Wind Tunnel Technical Service Platform (Grant No. 19DZ2290400) and the Fundamental Research Funds for the Central Universities. N.M. gratefully acknowledges the support of the International Institute for Carbon Neutral Energy Research, sponsored by the Japanese Ministry of Education, Culture, Sports, Science and Technology.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.egyai.2025.100549](https://doi.org/10.1016/j.egyai.2025.100549).

Data availability

Data will be made available on request.

References

- [1] How DNT, Hannan MA, Lipu MSH, Ker PJ. State of charge estimation for lithium-ion batteries using model-based and data-driven methods: a review. *Ieee Access* 2019;7:136116–36. <https://doi.org/10.1109/access.2019.2942213>.
- [2] Xu J, Cai X, Cai S, Shao Y, Hu C, Lu S, Ding S. High-energy lithium-ion batteries: recent progress and a promising future in applications. *Energy Environ Mater* 2023;6(5). <https://doi.org/10.1002/eem2.12450>.
- [3] Du X, Zhao L, Yang Z, Jin Z. A closed-loop control on temperature difference of a lithium-ion battery by pulse heating in cold climates. *J Energy Storage* 2023;57. <https://doi.org/10.1016/j.est.2022.106311>.
- [4] Habib AKMA, Hasan MK, Issa GF, Singh D, Islam S, Ghazal TM. Lithium-ion battery management system for electric vehicles: constraints, challenges, and recommendations. *Batteries-Basel* 2023;9(3). <https://doi.org/10.3390/batteries9030152>.
- [5] Singh S, Olyaei M, Jiang K, Gurumukhi Y, Goodson K, Asheghi M, Miljkovic N, Amer Soc Mechanical E. Virtual testbed for economical and reliability analysis of battery thermal management control strategies. In: *Proceedings Of Asme 2023 International Technical Conference And Exhibition On Packaging And Integration Of Electronic And Photonic Microsystems, Interpack2023*; 2023.
- [6] Liu X, Kabirzadeh P, Wu H, Fu WC, Qiu HY, Miljkovic N, Li YM, Wang PF. Machine learning enhanced control co-design optimization of an immersion cooled battery thermal management system. *J Appl Phys* 2024;136(2). <https://doi.org/10.1063/5.0201438>.
- [7] Singh S, Jennings M, Katragadda S, Che JDY, Miljkovic N. System design and analysis methods for optimal electric vehicle thermal management. *Appl Therm Eng* 2023;232. <https://doi.org/10.1016/j.applthermaleng.2023.120990>.
- [8] Yong JY, Ramachandaramurthy VK, Tan KM, Mithulanathan N. A review on the state-of-the-art technologies of electric vehicle, its impacts and prospects. *Renew Sustain Energy Rev* 2015;49:365–85. <https://doi.org/10.1016/j.rser.2015.04.130>.
- [9] Hannan MA, Lipu MSH, Hussain A, Mohamed A. A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: challenges and recommendations. *Renew Sustain Energy Rev* 2017;78:834–54. <https://doi.org/10.1016/j.rser.2017.05.001>.
- [10] Grieshaber D, MacKenzie R, Voerles J, Reimhult E. Electrochemical biosensors – sensor principles and architectures. *Sensors* 2008;8(3):1400–58. <https://doi.org/10.3390/s8031400>.
- [11] Ilott AJ, Mohammadi M, Schauerman CM, Ganter MJ, Jerschow A. Rechargeable lithium-ion cell state of charge and defect detection by in-situ inside-out magnetic resonance imaging. *Nat Commun* 2018;9. <https://doi.org/10.1038/s41467-018-04192-x>.
- [12] Shi Z, O'Brien W. Development and implementation of automated fault detection and diagnostics for building systems: a review. *Autom Constr* 2019;104:215–29. <https://doi.org/10.1016/j.autcon.2019.04.002>.
- [13] Yin S, Ding SX, Xie X, Luo H. A review on basic data-driven approaches for industrial process monitoring. *Ieee Trans Ind Electron* 2014;61(11):6418–28. <https://doi.org/10.1109/tie.2014.2301773>.
- [14] Wang S, Zhang S, Wen S, Fernandez C. An accurate state-of-charge estimation of lithium-ion batteries based on improved particle swarm optimization-adaptive square root cubature kalman filter. *J Power Sources* 2024;624. <https://doi.org/10.1016/j.jpowsour.2024.235594>.
- [15] Wang S, Dang Q, Gao Z, Li B, Fernandez C, Blaabjerg F. An innovative square root - untraced Kalman filtering strategy with full-parameter online identification for state of power evaluation of lithium-ion batteries. *J Energy Storage* 2024;104. <https://doi.org/10.1016/j.est.2024.114555>.
- [16] Wang SL, Wang C, Takyi-Aninakwa P, Jin SY, Fernandez C, Huang Q. An improved parameter identification and radial basis correction-differential support vector machine strategies for state-of-charge estimation of urban-transportation-electric-vehicle lithium-ion batteries. *J Energy Storage* 2024;80. <https://doi.org/10.1016/j.est.2023.110222>.
- [17] Xia C, Wang MJ, Fan YJ, Yang ZG, Du XZ. A hierarchical autoencoder and temporal convolutional neural network reduced-order model for the turbulent wake of a three-dimensional bluff body. *Phys Fluids* 2023;35(2). <https://doi.org/10.1063/5.0137285>.
- [18] Yang F, Li W, Li C, Miao Q. State-of-charge estimation of lithium-ion batteries based on gated recurrent neural network. *Energy* 2019;175:66–75. <https://doi.org/10.1016/j.energy.2019.03.059>.
- [19] Yang F, Song X, Xu F, Tsui K-L. State-of-charge estimation of lithium-ion batteries via long short-term memory network. *Ieee Access* 2019;7:53792–9. <https://doi.org/10.1109/access.2019.2912803>.
- [20] Chemali E, Kollmeyer PJ, Preindl M, Ahmed R, Emadi A. Long short-term memory networks for accurate State-of-charge estimation of Li-ion batteries. *Ieee Trans Ind Electron* 2018;65(8):6730–9. <https://doi.org/10.1109/tie.2017.2787586>.

- [21] Tong S, Lacap JH, Park JW. Battery state of charge estimation using a load-classifying neural network. *J Energy Storage* 2016;7:236–43. <https://doi.org/10.1016/j.est.2016.07.002>.
- [22] Chaoui H, Ibe-Ekeocha CC. State of charge and state of health estimation for lithium batteries using recurrent neural networks. *IEEE Trans Veh Technol* 2017;66(10):8773–83. <https://doi.org/10.1109/TVT.2017.2715333>.
- [23] Cui D, Xia B, Zhang R, Sun Z, Wang W, Sun W, Lai Y, Wang M. A novel intelligent method for the State of charge estimation of lithium-ion batteries using a discrete wavelet transform-based wavelet neural network. *Energies* 2018;11(4). <https://doi.org/10.3390/en11040995>.
- [24] Yang F, Zhang S, Li W, Miao Q. State-of-charge estimation of lithium-ion batteries using LSTM and UKF. *Energy* 2020;201. <https://doi.org/10.1016/j.energy.2020.117664>.
- [25] Kang L, Zhao X, Ma J. A new neural network model for the state-of-charge estimation in the battery degradation process. *Appl Energy* 2014;121:20–7. <https://doi.org/10.1016/j.apenergy.2014.01.066>.
- [26] Mao L, Yang C, Zhao J, Qu K, Yu X. Joint State-of-charge and State-of-health estimation for lithium-ion batteries based on improved lebesgue sampling and division of aging stage. *Energy Technol* 2023;11(10). <https://doi.org/10.1002/ente.202300567>.
- [27] Wu X, Li M, Du J, Hu F. SOC prediction method based on battery pack aging and consistency deviation of thermoelectric characteristics. *Energy Rep* 2022;8: 2262–72. <https://doi.org/10.1016/j.egyr.2022.01.056>.
- [28] Tao S, Jiang B, Wei X, Dai H. A systematic and comparative study of distinct recurrent neural networks for lithium-ion battery State-of-charge estimation in electric vehicles. *Energies* 2023;16(4). <https://doi.org/10.3390/en16042008>.
- [29] Rezaei O, Rahdan A, Sardari S, Dahmardeh M, Wang Z. Papers A fuzzy robust two-stage unscented Kalman filter method for uncertainty and state of charge estimation of lithium-ion batteries. *J Energy Storage* 2023;68. <https://doi.org/10.1016/j.est.2023.107883>.
- [30] Ul Hassan M, Saha S, Haque ME, Islam S, Mahmud A, Mendis N. A comprehensive review of battery state of charge estimation techniques. *Sustain Energy Technol Assess* 2022;54. <https://doi.org/10.1016/j.seta.2022.102801>.
- [31] Chung J, Gülçehre Ç, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *ArXiv abs/1412.3555* 2014.
- [32] Peters ME, Neumann M, Zettlemoyer L, Yih W-T, Assoc Computat L. Dissecting contextual word embeddings: architecture and representation. In: *Conference on Empirical Methods in Natural Language Processing (EMNLP)*; 2018. p. 1499–509.
- [33] Sutskever I, Vinyals O, Le QV. Sequence to Sequence learning with Neural networks. In: *28th Conference on Neural Information Processing Systems (NIPS)*; 2014.
- [34] Graves A, Schmidhuber J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw* 2005;18(5–6):602–10. <https://doi.org/10.1016/j.neunet.2005.06.042>.
- [35] Ng KS, Moo C-S, Chen Y-P, Hsieh Y-C. Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries. *Appl Energy* 2009;86(9):1506–11. <https://doi.org/10.1016/j.apenergy.2008.11.021>.
- [36] Tharwat A, Schenck W. A survey on active learning: state-of-the-art, practical challenges and research directions. *Mathematics* 2023;11(4). <https://doi.org/10.3390/math11040820>.
- [37] Sverchkov Y, Craven M. A review of active learning approaches to experimental design for uncovering biological networks. *PLoS Comput. Biol* 2017;13(6). <https://doi.org/10.1371/journal.pcbi.1005466>.
- [38] Cevik M, Ergun MA, Stout NK, Trentham-Dietz A, Craven M, Alagoz O. Using active learning for speeding up calibration in simulation models. *Med Decis Making* 2016; 36(5):581–93. <https://doi.org/10.1177/0272989x15611359>.
- [39] Akbari K, Rahmani E, Abbasi A, Askari M-R. Optimal placement of distributed generation in radial networks considering reliability and cost indices. *J Intell Fuzzy Syst* 2016;30:1077–86. <https://doi.org/10.3233/IFS-151883>.
- [40] Abbasi AR. Comparison parametric and non-parametric methods in probabilistic load flow studies for power distribution networks. *Electr Eng* 2022;104(6): 3943–54. <https://doi.org/10.1007/s00202-022-01590-9>.
- [41] Ansari J, Homayounzade M, Abbasi AR. Load frequency control in power systems by a robust backstepping sliding mode controller design. *Energy Rep* 2023;10: 1287–98. <https://doi.org/10.1016/j.egyr.2023.08.008>.
- [42] Abbasi AR, Gandhi CP. A novel hyperbolic fuzzy entropy measure for discrimination and taxonomy of transformer winding faults. *IEEE Trans Instrum Meas* 2022;71:1–8. <https://doi.org/10.1109/TIM.2022.3212522>.
- [43] Wu D, Lin C-T, Huang J. Active learning for regression using greedy sampling. *Inf Sci* 2019;474:90–105. <https://doi.org/10.1016/j.ins.2018.09.060>.
- [44] Yu H, Kim S. Passive sampling for regression. In: *2010 IEEE International Conference on Data Mining*; 2010. p. 1151–6.
- [45] Rasmussen CE. Gaussian processes in machine learning. In: Bousquet O, VonLuxburg U, Ratsch G, editors. *Advanced Lectures On Machine Learning*; 2004. p. 63–71. https://doi.org/10.1007/978-3-540-28650-9_4.
- [46] Sahinoglu GO, Pajovic M, Sahinoglu Z, Wang Y, Orlik PV, Wada T. Battery state-of-charge estimation based on regular/recurrent gaussian process regression. *Ieee Trans Ind Electron* 2018;65(5):4311–21. <https://doi.org/10.1109/tie.2017.2764869>.
- [47] Li Y, Liu K, Foley AM, Zulke A, Berecibar M, Nanini-Maury E, Van Mierlo J, Hostler LV. Data-driven health estimation and lifetime prediction of lithium-ion batteries: a review. *Renew Sustain Energy Rev* 2019;113. <https://doi.org/10.1016/j.rser.2019.109254>.
- [48] Zhang Q, Zhang Y, Wang J, Stimming U, Lee AA. Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning. *Nat Commun* 2020;11(1). <https://doi.org/10.1038/s41467-020-15235-7>.
- [49] Cacciarelli D, Kulahci M. Active learning for data streams: a survey. *Mach Learn* 2024;113(1):185–239. <https://doi.org/10.1007/s10994-023-06454-2>.
- [50] Jospin LV, Laga H, Boussaid F, Buntine W, Bannamoun M. Hands-on Bayesian neural networks—a tutorial for deep learning users. *IEEE Comput Intell Mag* 2022; 17(2):29–48. <https://doi.org/10.1109/mci.2022.3155327>.
- [51] Guo C, Pleiss G, Sun Y, Weinberger KQ. On calibration of modern neural networks. In: *34th International Conference on Machine Learning*; 2017.
- [52] Izmailov P, Vikram S, Hoffman MD, Wilson AG. What are bayesian neural network posteriors really like?. In: *International Conference on Machine Learning (ICML)*. *Electr Network*; 2021.
- [53] Blei DM, Kucukelbir A, McAuliffe JD. Variational inference: a review for statisticians. *J Am Stat Assoc* 2017;112(518):859–77. <https://doi.org/10.1080/01621459.2017.1285773>.
- [54] Kullback S, Leibler RA. On information and sufficiency. *Ann Math Stat* 1951;22(1): 79–86.
- [55] Shannon CE. A mathematical theory of communication. *Bell Syst Techn J* 1948;27(3):379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
- [56] B. Charles, C. Julien, K. Koray, W. Daan, 2015, Weight uncertainty in neural network. In *International conference on machine learning*, pp. 1613–1622. PMLR.
- [57] Buntine WL. Operations for learning with graphical models. *J Artif Intell Res* 1994; 2:159–225. <https://doi.org/10.1613/jair.62>.
- [58] Tutuianu M, Bonnel P, Cuffio B, Haniu T, Ichikawa N, Marotta A, Pavlovic J, Steven H. Development of the world-wide harmonized light duty Test cycle (WLTC) and a possible pathway for its introduction in the European legislation. *Transp Res Part D-Transp Environ* 2015;40:61–75. <https://doi.org/10.1016/j.trd.2015.07.011>.
- [59] Zhao L, Du X, Yang Z, Xia C, Xue J, Hoque MJ, Fu W, Yan X, Miljkovic N. Online state of health estimation with deep learning frameworks based on short and random battery charging data segments. *J Electrochem Soc* 2023;170(9):090537. <https://doi.org/10.1149/1945-7111/acf8ff>.
- [60] Pearson VII K. Mathematical contributions to the theory of evolution.—III. Regression, heredity, and panmixia. *Philos Trans Royal Soc London. Series A, Contain Papers Math Phys Character* 1896;(187):253–318.
- [61] Janse RJ, Hoekstra T, Jager KJ, Zoccali C, Tripepi G, Dekker FW, van Diepen M. Conducting correlation analysis: important limitations and pitfalls. *Clin Kidney J* 2021;14(11):2332–7. <https://doi.org/10.1093/cjk/sfab085>.
- [62] Bashir MF, Ma B, Bilal BK, Bashir MA, Tan D, Bashir M. Correlation between climate indicators and COVID-19 pandemic in New York, USA. *Sci Total Environ* 2020;728:138835. <https://doi.org/10.1016/j.scitotenv.2020.138835>.
- [63] Sevenson KA, Attia PM, Jin N, Perkins N, Jiang B, Yang Z, Chen MH, Aykol M, Herring PK, Fraggadakis D, Bazan MZ, Harris SJ, Chueh WC, Braatz RD. Data-driven prediction of battery cycle life before capacity degradation. *Nat Energy* 2019;4(5):383–91. <https://doi.org/10.1038/s41560-019-0356-8>.
- [64] Milchtein CM. How to maintain an electric vehicle. <https://mwg.aaa.com/via/car/how-maintain-electric-vehicle>; 2021.
- [65] Patel AN, Lander L, Ahuja J, Bulman J, Lum JKH, Pople JOD, Hales A, Patel Y, Edge JS. Lithium-ion battery second life: pathways, challenges and outlook. *Front Chem* 2024;12. <https://doi.org/10.3389/fchem.2024.1358417>.