ELECTROCHEMISTRY DEGRADATION REPRESENTATION FOR CAUSAL SOH PREDICTION IN BATTERIES

Anonymous authors
Paper under double-blind review

ABSTRACT

Battery state-of-health (SOH) forecasting is essential for reliable battery management, lifetime assessment, and risk-aware operation in electric vehicles and energy storage systems. However, accurate SOH prediction remains challenging because battery aging is governed by latent electrochemical degradation processes that are only indirectly reflected by heterogeneous cycle-level measurements, while historical SOH information may be incomplete or unreliable during deployment. To address these challenges, we propose **EDR-BAT**, an electrochemistry degradation representation framework for sequential causal SOH prediction in batteries. Guided by physically motivated degradation priors, EDR-BAT first transforms raw cycling measurements into multi-modal electrochemical embedding-based health features and uses series-level identity as a shared prior to query recent electrochemical history. This yields cycle-specific degradation representations that capture both shared aging regularities across similar batteries and cycle-specific variations within individual cells. To reduce drift under online prediction, EDR-BAT further introduces an SOH memory mechanism with memory-anchored training, allowing the model to operate with measured, incomplete, or autoregressively predicted SOH histories and reducing training-inference mismatch. Experiments on five battery aging datasets show that EDR-BAT consistently improves SOH prediction accuracy over representative baselines, reducing MAE/RMSE by **12.9%/11.4%** on average compared with the best-performing deep learning baseline on each dataset.

1 INTRODUCTION

Battery state-of-health (SOH) forecasting is a critical component of battery management systems for electric vehicles, grid-connected energy storage, and other safety-critical storage applications. Accurate SOH prediction enables degradation monitoring, lifetime prediction, maintenance scheduling, and risk-aware operation (Liu et al., 2025). However, battery aging is governed by complex electrochemical processes, such as capacity fade, impedance growth, and nonlinear degradation accumulation. These internal states are not directly observable during operation, but are indirectly reflected by cycle-level measurements including voltage, current, temperature, charge–discharge time, and capacity-related responses (Li et al., 2021c). Reliable SOH forecasting therefore requires inferring latent degradation states from heterogeneous electrochemical observations while using available SOH trajectories as contextual evidence.

Existing SOH prediction methods can be broadly categorized into model-based and data-driven approaches. Model-based methods incorporate electrochemical or equivalent-circuit knowledge, but often require strong assumptions, careful parameter identification, and system-specific calibration (Chen et al., 2023). Data-driven methods learn SOH mappings from cycling measurements, and recent deep models such as recurrent networks (Eddahech et al., 2012), temporal convolutional networks (Bai & Wang, 2023), and Transformers (Gu et al., 2023) further improve nonlinear temporal modeling. Nevertheless, two key challenges remain insufficiently addressed. First, electrochemical observations are often treated as generic time-series signals rather than as evidence of an underlying degradation state (Bockrath et al., 2023). As a result, the learned representations may capture statistical correlations between cycling measurements, but do not explicitly distinguish aging regularities shared across related batteries from cell-specific degradation evolution. Second, SOH history is typically used under a fully supervised setting, while practical deployment often requires sequen-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

tial prediction with unavailable or model-generated SOH values (Qian et al., 2023). This creates a training-inference mismatch, where early prediction errors may be propagated through the historical SOH context and cause drift over long-horizon inference (Bengio et al., 2015).

Motivated by these limitations, we formulate SOH forecasting as a sequential causal SOH prediction problem driven by electrochemical evidence and evolving SOH history. Battery degradation is neither purely universal nor purely cell-specific: batteries from the same series share common aging regularities, while individual cells follow different trajectories due to manufacturing variability, initial conditions, and usage-dependent stress (Baumhöfer et al., 2014). This suggests that SOH forecasting should infer degradation representations that capture series-level commonality while adapting to cell-specific electrochemical evolution (Zhang et al., 2025). Meanwhile, SOH history should be modeled as an evolving memory rather than a fixed ground-truth window, since deployment may involve incomplete, measured, or autoregressively predicted SOH contexts (Cai et al., 2022).

To this end, we propose **EDR-BAT**, an electrochemistry degradation representation framework for causal SOH prediction in batteries, that integrates electrochemical evidence and the evolving SOH history by learning series-aware degradation representations for sequential prediction. First, multi-modal electrochemical embeddings convert raw cycle-level measurements into structured features that summarize temporal, differential, integral, and statistical health indicators, providing degradation-aware evidence beyond generic time-series inputs. Second, a neural degradation representation module learns a cycle-specific degradation state by using a battery-series learnable prior to query recent electrochemical embedding history, thereby capturing shared degradation regularities across similar batteries while adapting to cell-specific aging behavior. A low-rank bilinear degradation interaction couples the learned degradation state with current electrochemical evidence and recent SOH context, enabling expressive yet parameter-efficient modeling of degradation-dependent prediction. Third, history-aware causal SOH prediction incorporates SOH history as trajectory-level context through memory-anchored training and auto-regressive prediction. This design reduces over-dependence on clean ground-truth SOH windows during training and enables the model to construct usable SOH history during inference, thereby mitigating the training-inference mismatch in battery forecasting.

In summary, our main contributions are as follows:

- We formulate battery SOH forecasting as a deployment-aware sequential causal SOH prediction problem, where cycle-level heterogeneous electrochemical observations provide degradation evidence and SOH history serves as an evolving trajectory context.
- We propose an electrochemistry degradation representation framework: multi-modal electrochemical embeddings capture aging responses, while a series-level learnable prior queries recent embeddings to model shared aging regularities and build cycle-specific degradation evolution.
- We introduce a history-aware training and inference paradigm that treats SOH history as an evolving trajectory memory, mitigating mismatch and drift in long-term causal SOH prediction.

2 Related Work

Data-driven SOH prediction. Data-driven SOH prediction learns mappings from battery operating measurements to health states. Classical methods extract health indicators from voltage, current, charge–discharge time, or temperature profiles and perform supervised regression using support vector regression (Yang et al., 2018), Gaussian process regression (Richardson et al., 2018), random forests (Ng et al., 2020), and extreme learning machines (Pan et al., 2018). Deep models such as CNNs (Bai & Wang, 2023), LSTMs (Li et al., 2021b), and GRUs (Chen et al., 2022) further learn degradation-related representations automatically. Recent studies also show that richer degradation features improve health modeling, including early-cycle voltage variations (Severson et al., 2019) and machine-learning pipelines for SOH prediction (Roman et al., 2021). Despite these advances, most methods still perform direct feature-to-SOH regression, with limited explicit modeling of historical degradation context.

Online and cross-cell SOH prediction. Online SOH prediction estimates battery health during in-service operation from sequential observations. In practical electric vehicles or energy storage systems, complete charge–discharge tests are often unavailable, and measurements are partial, ir-

regular, condition-dependent, and sparsely labeled. Prior studies exploit partial charging voltage profiles (Meng et al., 2019), voltage relaxation curves (Zhu et al., 2022), or multi-source fleet features (Liu et al., 2025), showing the feasibility of SOH prediction under practical constraints. However, these methods mainly focus on constructing reliable features from partial observations or updating estimators with newly observed segments, while full-lifecycle cycle-by-cycle prediction using only current and past information remains less explored. Cross-cell SOH generalization further requires models trained on source cells to work on unseen target cells under manufacturing variations, operating-condition shifts, and heterogeneous degradation paths (Zhou et al., 2022; Zhang et al., 2023). Existing studies reduce target-cell dependence through transfer learning, domain generalization, or standardized degradation benchmarks, but causal online prediction without target-cell fine-tuning or calibration remains challenging.

Electrochemistry-informed SOH prediction. Electrochemistry learning has also been explored for battery degradation modeling and SOH estimation. Existing studies incorporate degradation priors, state constraints, or electrochemistry-informed losses into neural models. For example, PINN4SOH combines neural networks with empirical degradation and state-space constraints for stable SOH estimation (Wang et al., 2024). These methods improve physical consistency and interpretability over purely data-driven regression. However, most electrochemistry-informed SOH models emphasize constrained estimation, degradation prognosis, or physically plausible trajectory modeling, while explicit organization of current-cycle electrochemical evidence, evolving historical cycling context, and autoregressive SOH memory for causal online SOH prediction remains less explored.

History-aware degradation Modeling. Deep sequential models capture temporal dependencies in battery degradation using historical windows, cycling records, or charging-segment sequences. LSTM-based models (Li et al., 2020) and attentional LSTMs (Lin et al., 2023) show the usefulness of historical observations for SOH prediction. However, most methods encode current and historical information in a unified sequence model, leaving their functional roles in the current SOH estimate implicit. More structured methods model degradation at the trajectory level, such as one-shot degradation trajectory prediction (Li et al., 2021a) and FESD with Transformer-based long-range modeling and structured state-space diffusion (Wang et al., 2025). Yet they mainly focus on lifetime prediction, future trajectory forecasting, or sequence reconstruction. In contrast, our work targets causal online SOH prediction by explicitly separating current-cycle features, historical cycling context, and historical SOH memory.

3 METHODOLOGY

Problem Statement. Let $\mathcal{B} = \{B_1, B_2, \ldots, B_M\}$ denote a set of batteries, where B_i indexes the i-th battery instance. For battery B_i, we observe a sequence of operational and electrochemical measurements over charge–discharge cycles, denoted as

$$\mathbf{X}_i^{1:n} = \{\mathbf{x}_i(1), \mathbf{x}_i(2), \ldots, \mathbf{x}_i(n)\}, \tag{1}$$

where $\mathbf{x}_i(n)$ denotes the available measurements at cycle n, such as voltage, current, temperature, time, and capacity-related indicators. Let c_i denote the battery-series index of B_i. Given the series index c_i and the historical observations $\mathbf{X}_i^{1:n}$ up to cycle n, the goal of battery SOH prediction is to predict the current SOH value $s_i(n)$:

$$\hat{s}_i(n) = f\big(c_i, \mathbf{X}_i^{1:n}\big), \tag{2}$$

The ground-truth SOH is commonly defined as $s_i(n) = Q_i(n)/Q_i^{\mathrm{rated}}$, where $Q_i(n)$ denotes the available capacity of battery B_i at cycle n, and Q_i^{rated} is the rated capacity.

To solve this problem, we propose an electrochemistry degradation representation framework for battery SOH prediction, as shown in Figure 1. The degradation prior in EDR-BAT is implemented through architecture-level constraints: multi-modal electrochemical embeddings encode observable aging responses, a battery-series learnable prior provides shared series-level information across batteries from the same series, and cross-attention to recent electrochemical history refines this prior into a cycle-specific degradation representation. Accordingly, the framework transforms raw cycle-level measurements into multi-modal electrochemical embeddings, learns a cycle-specific degradation representation by querying recent electrochemical history, and predicts SOH through low-rank bilinear electrochemical and SOH-window branches. We further design a history-aware causal SOH

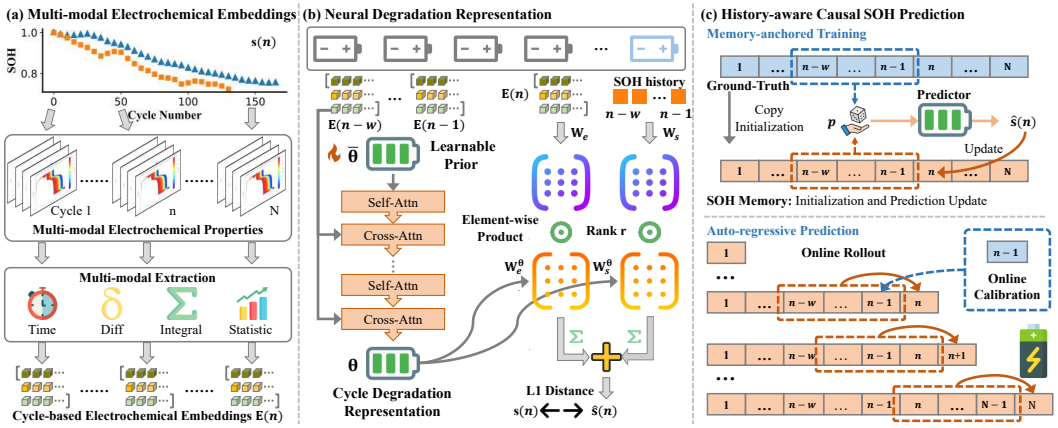


Figure 1: **EDR-BAT: An Electrochemistry Degradation Representation for battery SOH prediction**. (a) Multi-modal raw cycle-level measurements are transformed into cycle-based electrochemical embeddings. (b) A battery-series learnable prior queries recent electrochemical history to learn the cycle-specific degradation representation, while SOH is predicted through low-rank bilinear electrochemical and SOH history branches. (c) History-aware training with memory-anchored SOH sampling reduces training-inference mismatch. During inference, the SOH window is updated autoregressively using previous predictions.

prediction strategy, where memory-anchored training reduces the training-inference mismatch, and auto-regressive inference progressively updates the SOH window with optional online calibration.

3.1 Multi-modal Electrochemical Embeddings

Battery SOH is reflected through multiple electrochemical observations rather than a single measurement. This section describes how raw cycle-level measurements are transformed into cycle-based electrochemical embeddings, which provide physically meaningful observable proxies for the latent battery degradation state.

The historical observation $\mathbf{X}_i^{1:n}$ consists of a sequence of cycle-level measurements, where each $\mathbf{x}_i(n)$ contains multiple electrochemical signals, such as voltage, current, charge/capacity-related indicators, and temperature. Since these signals are heterogeneous and describe battery degradation from different physical perspectives, directly using the raw measurements may obscure the SOH-related electrochemical patterns. Therefore, we first convert each cycle observation $\mathbf{x}_i(n)$ into a compact cycle-based electrochemical embedding $\mathbf{E}_i(n)$. For battery B_i at cycle n, we decompose the cycle-level measurement $\mathbf{x}_i(n)$ into multiple electrochemical signal components (Pastor-Fernández et al., 2017):

$$\mathbf{x}_i(n) = [\mathbf{v}_i(n), \mathbf{i}_i(n), \mathbf{q}_i(n), \mathbf{t}_i(n)], \tag{3}$$

where $\mathbf{v}_i(n)$, $\mathbf{i}_i(n)$, $\mathbf{q}_i(n)$, and $\mathbf{t}_i(n)$ denote the voltage, current, charge/capacity-related indicators, and temperature within cycle n, respectively. Capacity-related indicators refer to curve-derived capacity features, rather than the ground-truth full-cycle capacity used to compute SOH labels. Based on these electrochemical signals, we construct four types of health-indicator representations:

$$\mathbf{h}_i(n) = \left[\mathbf{h}_i^{\text{time}}(n), \mathbf{h}_i^{\text{diff}}(n), \mathbf{h}_i^{\text{int}}(n), \mathbf{h}_i^{\text{stat}}(n)\right]. \tag{4}$$

Here, $\mathbf{h}_i^{\text{time}}(n)$ preserves the original curve-level temporal patterns of the electrochemical signals, $\mathbf{h}_i^{\text{diff}}(n)$ captures differential variations such as slopes or local curve changes, $\mathbf{h}_i^{\text{int}}(n)$ describes accumulated effects through integral quantities, and $\mathbf{h}_i^{\text{stat}}(n)$ summarizes statistical properties of the cycle-level signals. These complementary health indicators characterize SOH-related degradation signatures from curve shape, local variation, accumulated response, and global statistics.

The concatenated health indicator $\mathbf{h}_i(n)$ is projected into a latent feature space, yielding the electrochemical embedding sequence:

$$\mathbf{E}_i^{1:n} = [\mathbf{E}_i(1), \mathbf{E}_i(2), \ldots, \mathbf{E}_i(n)], \qquad \mathbf{E}_i(n) = \phi_{\text{emb}}(\mathbf{h}_i(n)). \tag{5}$$

The intrinsic degradation state is an inherent but not directly observable property of the battery, while the measured electrochemical signals are its external manifestations. Therefore, the cycle-based embedding $\mathbf{E}_i(n)$ provides a compact and physically meaningful expression of SOH-related electrochemical signatures at cycle n, serving as an observable proxy for the latent battery state. The resulting embedding sequence $\mathbf{E}_i^{1:n}$ then forms the basis for subsequent degradation modeling.

3.2 Neural Degradation Representation

The embedding sequence $\mathbf{E}_i^{1:n}$ describes observable electrochemical signatures, but the underlying degradation state is latent and evolves along each battery's historical trajectory. Since batteries from the same series share common electrochemical characteristics, we introduce a series-level learnable prior and refine it by querying the recent electrochemical history of battery B_i. The resulting cycle-specific degradation representation is then coupled with the current electrochemical embedding via a low-rank bilinear interaction to predict the SOH at cycle n.

Cycle Degradation Representation. For each battery dataset or battery series c, we maintain a shared learnable prior $\boldsymbol{\theta}_c \in \mathbb{R}^{d_K \times d_\theta}$, which is shared by all batteries from the same series. For battery B_i, the corresponding series-level learnable prior token $\boldsymbol{\theta}_{c_i}$ is used as the initial query. At cycle n, we collect a recent historical window of cycle-based electrochemical embeddings:

$$\mathbf{E}_i^{n-w:n-1} = [\mathbf{E}_i(n-w), \ldots, \mathbf{E}_i(n-1)], \tag{6}$$

where w is the history window length. This window records the recent electrochemical evolution of battery B_i before cycle n. We use $\boldsymbol{\theta}_{c_i}$ as the initial learnable prior query, and use the historical electrochemical window $\mathbf{E}_i^{n-w:n-1}$ as the key-value sequence. Instead of using a single attention block, we stack multiple self-attention and cross-attention layers to progressively refine the shared learnable prior into a battery- and cycle-specific degradation representation.

For the ℓ-th layer, the representation is first updated by self-attention:

$$\tilde{\boldsymbol{\theta}}_i^{n,\ell} = \mathrm{SelfAttn}^{(\ell)}\left(\boldsymbol{\theta}_i^{n,\ell-1}\right), \quad \boldsymbol{\theta}_i^{n,0} = \boldsymbol{\theta}_{c_i}. \tag{7}$$

and then queries the historical electrochemical embeddings through cross-attention:

$$\boldsymbol{\theta}_i^{n,\ell} = \mathrm{CrossAttn}^{(\ell)}\left(Q = \tilde{\boldsymbol{\theta}}_i^{n,\ell}, K = \mathbf{E}_i^{n-w:n-1}, V = \mathbf{E}_i^{n-w:n-1}\right). \tag{8}$$

After L stacked layers, we obtain $\boldsymbol{\theta}_i^n = \boldsymbol{\theta}_i^{n,L}$, where $\boldsymbol{\theta}_i^n$ denotes the learned degradation representation of battery B_i at cycle n. This representation starts from the shared battery-series identity $\boldsymbol{\theta}_{c_i}$ and is adapted to each battery by querying the recent electrochemical history before cycle n.

Low-Rank Bilinear Degradation Interaction. The representation $\boldsymbol{\theta}_i^n$ encodes the identity-conditioned historical degradation state of battery B_i at cycle n. To predict the current SOH, we couple it with the current electrochemical embedding $\mathbf{E}_i(n)$ through a bilinear interaction (Kim et al., 2018):

$$\tilde{s}_i^{\mathrm{eco}}(n) = (\boldsymbol{\theta}_i^n)^\top \mathbf{W} \mathbf{E}_i(n), \tag{9}$$

where $\mathbf{W}$ parameterizes the interaction between the historical degradation state and the current electrochemical observation. To reduce parameter complexity and improve data efficiency (Kolda & Bader, 2009), we impose a low-rank structure on $\mathbf{W}$:

$$\mathbf{W} = \mathbf{W}_\theta^\top \mathbf{W}_e, \quad \mathbf{W}_\theta \in \mathbb{R}^{r \times d_\theta}, \quad \mathbf{W}_e \in \mathbb{R}^{r \times d_e}, \quad r \ll \min(d_\theta, d_e), \tag{10}$$

where d_θ and d_e are the dimensions of $\boldsymbol{\theta}_i^n$ and $\mathbf{E}_i(n)$. The prediction can then be written as

$$\tilde{s}_i^{\mathrm{eco}}(n) = (\boldsymbol{\theta}_i^n)^\top \mathbf{W}_\theta^\top \mathbf{W}_e \mathbf{E}_i(n) = \sum_{j=1}^{r} \left(\mathbf{W}_\theta \boldsymbol{\theta}_i^n\right)_j \left(\mathbf{W}_e \mathbf{E}_i(n)\right)_j. \tag{11}$$

We capture the coupling between the degradation state and current electrochemical observation by projecting both representations into a shared rank-r space and aggregating their element-wise product to obtain $\tilde{s}_i^{\mathrm{eco}}(n)$, requiring fewer parameters than a full bilinear layer.

3.3 HISTORY-AWARE CAUSAL SOH PREDICTION

The current SOH is reflected by electrochemical observations and constrained by recent SOH trajectory. Therefore, beyond the electrochemical prediction branch $\tilde{s}_i^{\text{eco}}(n)$, we also introduce a SOH-window correction term for the final prediction, where the latter captures the correlation between current degradation and recent SOH evolution.

Memory-Anchored Training. To incorporate recent SOH history, we define a historical SOH window before cycle n as

$$\mathbf{s}_i^{\text{win}}(n) = [\bar{s}_i(n-w), \bar{s}_i(n-w+1), \ldots, \bar{s}_i(n-1)]^\top, \tag{12}$$

where $\bar{s}_i(n)$ denotes the SOH value at cycle n. Given the cycle-specific degradation representation $\boldsymbol{\theta}_i^n$, we model its interaction with the SOH window through a low-rank bilinear decomposition:

$$\tilde{s}_i^{\text{win}}(n) = \sum_{j=1}^{r_s} (\mathbf{W}_\theta^s \boldsymbol{\theta}_i^n)_j \left(\mathbf{W}_{\text{win}}^s \mathbf{s}_i^{\text{win}}(n)\right)_j . \tag{13}$$

Here, r_s is the rank of the SOH-window interaction space. This branch explicitly models the multiplicative coupling between the latent degradation representation and the recent SOH trajectory, while avoiding a full bilinear matrix. The final SOH prediction is obtained by combining the electrochemical branch and the SOH-window branch:

$$\hat{s}_i(n) = \tilde{s}_i^{\text{eco}}(n) + \tilde{s}_i^{\text{win}}(n). \tag{14}$$

During training, directly using ground-truth SOH values in the historical window may cause a distribution mismatch because auto-regressive inference relies on predicted SOH values. To reduce this mismatch, we maintain a prediction-based SOH memory $s_i^{\text{mem}}(n)$ for each battery, initialized by ground-truth SOH values. For each historical position $k < n$, the window input is sampled from either the ground-truth SOH $s_i(k)$ or the detached prediction memory $s_i^{\text{mem}}(k)$:

$$\bar{s}_i(k) = (1-\alpha_{i,k}) s_i(k) + \alpha_{i,k} s_i^{\text{mem}}(k), \quad \alpha_{i,k} \sim \text{Bernoulli}(p), \tag{15}$$

where $p \in [0,1]$ controls the probability of using prediction memory instead of ground-truth SOH history during training. The case $p = 0$ corresponds to fully supervised history, while $p = 1$ matches the auto-regressive inference setting by relying entirely on prediction memory. After predicting $\hat{s}_i(n)$, the memory is updated as $s_i^{\text{mem}}(n) \leftarrow \hat{s}_i(n)$. The model is trained by minimizing the L1 distance between the predicted and ground-truth SOH:

$$\mathcal{L}_{\text{soh}} = \frac{1}{|\mathcal{D}|} \sum_{(i,n)\in\mathcal{D}} \left|\hat{s}_i(n) - s_i(n)\right| . \tag{16}$$

Auto-regressive Prediction. During inference, the historical SOH window is empty at the beginning, so the prediction mainly relies on the electrochemical branch, while missing SOH entries are initialized with a mask token. As predictions are generated, the SOH window is gradually populated and used for subsequent history-aware predictions. Specifically, when predicting the SOH at cycle n, the SOH window only contains historical entries before cycle n. For each historical position $k < n$, the SOH value used in the window is defined as

$$\bar{s}_i(k) = \begin{cases} s_i^{\text{obs}}(k), & \text{if an online measured SOH is available at cycle } k, \\ \hat{s}_i(k), & \text{if a previous prediction is available at cycle } k, \\ s_{\text{mask}}, & \text{otherwise,} \end{cases} \qquad k < n. \tag{17}$$

After predicting $\hat{s}_i(n)$ from the current electrochemical embedding and the available historical SOH window, the SOH window for the next cycle is updated as

$$\mathbf{s}_i^{\text{win}}(n+1) = \text{Update}\left(\mathbf{s}_i^{\text{win}}(n), \hat{s}_i(n)\right). \tag{18}$$

This auto-regressive update allows the model to start without historical SOH records and progressively build its SOH memory during online prediction. When sparse online measured SOH values are available, they replace the corresponding predicted entries in the memory window, so future predictions use calibrated historical values according to Eq. (17).

Table 1: Controlled comparison for full-cycle SOH prediction. MAE/RMSE are percentage points; subscripts denote standard deviations over five seeds. Best and second-best results are bolded and underlined, respectively. SVR is deterministic.

	CALCE		MIT		NASA		Oxford		In-house	
Model	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
SVR	0.55	0.68	1.65	1.80	1.80	2.10	0.47	0.55	4.60	4.96
MLP	$0.80_{0.45}$	$1.01_{0.52}$	$0.72_{0.08}$	$0.81_{0.07}$	$1.28_{0.12}$	$1.46_{0.13}$	$0.75_{0.48}$	$0.82_{0.52}$	$1.06_{0.06}$	$1.19_{0.03}$
CNN	$0.76_{0.10}$	$0.97_{0.12}$	$\underline{0.52}_{0.09}$	$\underline{0.61}_{0.08}$	$1.38_{0.29}$	$1.66_{0.31}$	$0.53_{0.14}$	$0.64_{0.17}$	$6.32_{0.76}$	$6.53_{0.80}$
LSTM	$0.45_{0.06}$	$0.59_{0.07}$	$0.65_{0.03}$	$0.71_{0.04}$	$\underline{0.89}_{0.15}$	$\underline{1.04}_{0.15}$	$0.45_{0.21}$	$0.50_{0.23}$	$\underline{0.92}_{0.09}$	$\underline{1.02}_{0.10}$
Transformer	$\underline{0.42}_{0.08}$	$\underline{0.54}_{0.07}$	$0.67_{0.08}$	$0.71_{0.08}$	$0.97_{0.26}$	$1.19_{0.26}$	$\underline{0.21}_{0.06}$	$\underline{0.25}_{0.07}$	$1.57_{0.53}$	$1.68_{0.53}$
LSTM-W	$0.60_{0.13}$	$1.07_{0.10}$	$0.67_{0.07}$	$0.93_{0.06}$	$1.23_{0.26}$	$1.49_{0.24}$	$0.35_{0.16}$	$0.45_{0.14}$	$2.12_{0.36}$	$2.91_{0.49}$
Trans-W	$0.44_{0.05}$	$0.61_{0.06}$	$0.62_{0.08}$	$0.69_{0.10}$	$0.91_{0.20}$	$1.19_{0.22}$	$0.49_{0.20}$	$0.56_{0.20}$	$1.21_{0.42}$	$1.62_{0.29}$
PINN4SOH	$0.60_{0.05}$	$0.87_{0.07}$	$0.59_{0.20}$	$0.68_{0.20}$	$0.90_{0.15}$	$1.19_{0.14}$	$0.30_{0.05}$	$0.43_{0.05}$	$2.60_{0.43}$	$2.82_{0.46}$
EDR-BAT	$\mathbf{0.40}_{0.03}$	$\mathbf{0.53}_{0.04}$	$\mathbf{0.47}_{0.03}$	$\mathbf{0.50}_{0.03}$	$\mathbf{0.67}_{0.13}$	$\mathbf{0.84}_{0.14}$	$\mathbf{0.18}_{0.02}$	$\mathbf{0.22}_{0.02}$	$\mathbf{0.82}_{0.07}$	$\mathbf{0.96}_{0.08}$

4 EXPERIMENT

4.1 IMPLEMENTATION DETAILS

Baselines and Metrics. We compare EDR-BAT with representative baselines under the same protocol, including SVR (Yang et al., 2018), MLP, CNN (Lee et al., 2023), LSTM (Li et al., 2020), Transformer (Gu et al., 2023), PINN4SOH (Wang et al., 2024), and their fixed-length sliding-window variants LSTM-W and Trans-W. We use MAE and RMSE as evaluation metrics, and report the mean and standard deviation over five random seeds.

Datasets. We evaluate EDR-BAT on five battery degradation datasets: CALCE (Xing et al., 2013), NASA (Saha & Goebel, 2007), Oxford (Howey & Birkl, 2017), MIT (Severson et al., 2019), and an in-house dataset. These datasets cover different battery chemistries, nominal capacities, and cycling protocols. For each dataset, one full-life cell is randomly selected as the target cell for testing, and the remaining cells are used for training. Complete leave-one-cell-out results are reported in Appendix C. Feature columns are standardized using statistics computed from the training cells and applied to the target cell. All evaluations are performed on the full-cycle SOH trajectory without using future test-cell information. Dataset and preprocessing details are provided in Appendix A.

Setup. All models are implemented in PyTorch and trained under the same optimization protocol. During testing, the batch size is fixed to 1 for cycle-by-cycle online inference. For EDR-BAT, core architectural settings are compact and largely shared across datasets, while dataset-specific settings are mainly used for electrochemical and SOH memory lengths. Complete hyperparameters, training details, and computational costs are reported in Appendix B.

4.2 MAIN RESULTS

Quantitative Evaluation. We compare EDR-BAT with current-cycle and sliding-window baselines under the same training/testing split, input features, and evaluation metrics. These baselines represent two common paradigms: estimating SOH from current-cycle electrochemical responses and modeling local degradation trends from fixed-length historical observations. EDR-BAT instead performs causal online SOH prediction by jointly organizing current-cycle evidence, historical cycling context, and historical SOH memory.

As shown in Table 1, EDR-BAT achieves the lowest MAE and RMSE on all five target cells under the unified protocol. Compared with CNN, LSTM, and Transformer, the improvements show that generic feature extraction or sequence encoding alone is limited for unseen-cell SOH prediction. Compared with PINN4SOH, EDR-BAT further highlights the benefit of explicitly modeling the interaction between current-cycle electrochemical responses and accumulated degradation history. The results also show consistent gains over LSTM-W and Trans-W, which directly use sliding-window inputs. Although sliding windows provide local temporal context, they do not explicitly distinguish the roles of current-cycle features, historical cycling behavior, and previous SOH states. EDR-BAT instead decouples current-cycle electrochemical features, historical cycling context, and

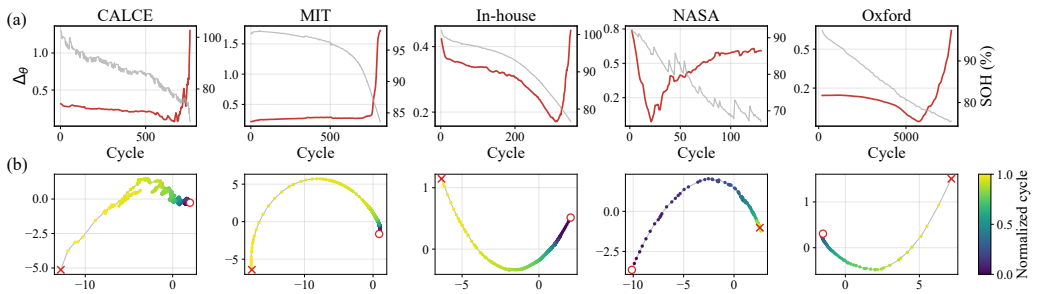


Figure 2: Visualization of the cycle degradation representation θ during online SOH prediction. (a) Relative displacement Δ_θ from the battery-series learnable prior, with true SOH shown on the right axis. (b) PCA trajectories of θ over the test life, where color denotes the normalized cycle.

historical SOH memory, and models their interactions at each prediction step. This design enables adaptive history-conditioned degradation representations and leads to more reliable full-cycle SOH prediction for unseen target cells. Complete leave-one-cell-out results are reported in Appendix C. Averaged over all held-out cells, EDR-BAT achieves consistently lower MAE and RMSE, further supporting cross-cell generalization.

Representation Dynamics Analysis. To examine how EDR-BAT uses historical degradation information, we visualize the learned cycle degradation representation θ during online inference. As shown in Figure 1, the battery-series learnable prior $\bar{\theta}$ provides a global latent representation, while θ is constructed by querying historical electrochemical memory through self-attention and cross-attention. Figure 2 shows the evolution of θ on five target cells. The nonzero and cycle-dependent displacement Δ_θ in Figure 2(a) indicates that historical electrochemical memory actively reshapes the degradation representation. The PCA trajectories in Figure 2(b) further show ordered aging-dependent paths as normalized cycle increases, suggesting that θ captures progressive degradation characteristics as historical evidence accumulates. This dynamic representation allows online SOH prediction to rely on accumulated degradation evidence rather than instantaneous electrochemical responses alone.

4.3 Ablation Study

To analyze the contribution of each component, we report ablation results in Table 2. Here, $p = 0$ uses only ground-truth SOH to construct historical SOH memory during training; w/o E-Q and w/o S-Q remove the electrochemical-history and SOH-memory query interactions, respectively; w/o Hist. removes both historical branches. The full model combines current-cycle electrochemical features, historical cycling, historical SOH memory, and an interaction-based regression head.

Across the five target cells, the full model achieves the lowest average MAE and RMSE, while removing both historical branches leads to the largest average degradation. This result shows that historical degradation evidence is important for full-cycle online SOH prediction. Without historical inputs, the model degenerates toward a single-step estimator and becomes less effective at tracking long-term degradation evolution.

The two interaction terms also contribute to the overall performance on average. Removing either the electrochemical-history query interaction or the SOH-memory query interaction increases the average error, with more pronounced degradation on NASA and the in-house dataset. This indicates that the proposed interaction design helps the model select degradation-relevant information from both historical electrochemical responses and previous SOH states. On Oxford, several variants obtain comparable errors, which is consistent with its relatively smooth degradation trajectories and stronger similarity among source cells.

The hybrid historical-SOH training strategy further improves online robustness. When $p = 0$, the model uses only ground-truth SOH values to construct historical SOH memory during training. This creates a mismatch with test-time online inference, where the memory is built from previous predictions. Allowing predicted SOH values to enter the memory during training better aligns the training and deployment conditions, reducing sensitivity to accumulated prediction errors.

Table 2: Ablation results of EDR-BAT. Best and second-best results are marked in bold and underlined, MAE/RMSE are reported in percentage points, with small subscripts denoting standard deviations over five random seeds.

	CALCE		MIT		NASA		Oxford		In-house	
Variant	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Full	**0.43**$_{0.03}$	**0.64**$_{0.02}$	**0.47**$_{0.03}$	**0.50**$_{0.03}$	**0.67**$_{0.13}$	**0.84**$_{0.14}$	0.21$_{0.04}$	0.26$_{0.05}$	**0.82**$_{0.07}$	**0.96**$_{0.08}$
$p = 0$	0.44$_{0.01}$	0.69$_{0.04}$	0.56$_{0.13}$	0.66$_{0.27}$	0.92$_{0.26}$	1.09$_{0.25}$	0.21$_{0.04}$	0.27$_{0.04}$	1.34$_{0.21}$	1.52$_{0.20}$
w/o E-Q	0.45$_{0.02}$	0.69$_{0.04}$	0.50$_{0.05}$	0.55$_{0.04}$	1.09$_{0.24}$	1.23$_{0.25}$	**0.18**$_{0.03}$	**0.23**$_{0.05}$	1.66$_{0.22}$	1.86$_{0.25}$
w/o S-Q	0.45$_{0.03}$	0.67$_{0.04}$	0.58$_{0.16}$	0.72$_{0.20}$	0.95$_{0.28}$	1.10$_{0.27}$	0.21$_{0.03}$	0.26$_{0.04}$	1.47$_{0.21}$	1.67$_{0.20}$
w/o Hist.	0.51$_{0.03}$	0.80$_{0.09}$	0.51$_{0.05}$	0.57$_{0.06}$	1.16$_{0.20}$	1.34$_{0.17}$	0.19$_{0.05}$	0.26$_{0.05}$	1.89$_{0.26}$	2.08$_{0.26}$

4.4 Analysis

Learnable Prior Representation.

We further visualize the learnable prior to examine whether it encodes battery-series-level characteristics. In EDR-BAT, $\bar{\theta}$ is a learnable latent prior shared by cells from the same battery series. It initializes the query representation used to extract cycle-specific degradation states from historical electrochemical observations. We project the learnable priors from the five datasets into a two-dimensional PCA space. As shown in Figure 3, the latent queries form dataset-dependent structures with partial overlap and visible separation. The overlap indicates shared degradation-related patterns across battery series, whereas the separation is consistent with series-level differences in chemistry, capacity, and cycling protocol. This organization indicates that $\bar{\theta}$ is not a randomly scattered latent basis. Instead, it provides structured battery-series prior information for constructing history-conditioned degradation representations during online SOH prediction.

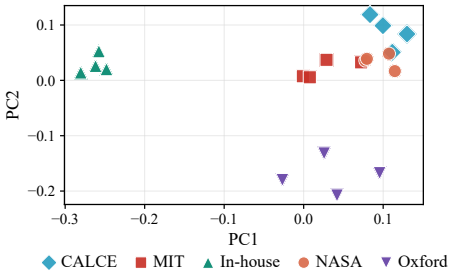


Figure 3: PCA visualization of the learned battery-series learnable prior $\bar{\theta}$. Each point denotes a latent query, and colors indicate datasets.

5 Conclusion

This paper proposes EDR-BAT, an electrochemistry degradation representation framework for causal online battery SOH prediction. Instead of directly regressing SOH from input features, EDR-BAT learns a structured degradation representation that connects current electrochemical observations with accumulated historical evidence. By introducing a battery-series learnable prior and history-conditioned degradation modeling, the framework captures both shared battery characteristics and cycle-specific aging evolution. EDR-BAT further integrates current-cycle features and historical SOH memory for online prediction, with memory-anchored training to reduce the training–inference mismatch. Experiments on five battery degradation datasets demonstrate strong performance under heterogeneous degradation patterns and zero-shot target-cell evaluation. Ablation and visualization results further support the importance of structured historical modeling and cycle-specific degradation representation. The results suggest that sequential battery health prediction benefits from explicitly representing what the battery is currently exhibiting, how its degradation state has evolved, and what SOH trajectory information is available during deployment, rather than encoding all historical information through a single generic sequence representation.

AI Use Statement

The authors used generative AI tools to assist with language polishing, manuscript organization, LaTeX troubleshooting, and formatting of tables and figures. Generative AI tools were also used to help debug auxiliary scripts for experiments and visualizations. All experimental results, analyses,

claims, figures, tables, and manuscript text were reviewed and verified by the authors. The authors take full responsibility for the final content of this submission.

ETHICS STATEMENT

This work may have positive societal impact by improving the reliability of battery SOH prediction, which is important for safer and more efficient battery management in electric vehicles, energy storage systems, and other electrified infrastructures. More accurate online SOH estimation can help extend battery lifetime, reduce unnecessary battery replacement, improve maintenance scheduling, and support the broader deployment of renewable energy and low-carbon transportation. Degradation-aware prediction may also help operators identify abnormal aging trends earlier, reducing potential safety risks caused by unexpected capacity loss or battery failure.

Potential negative impacts may arise if the prediction model is deployed without sufficient validation under real-world operating conditions. Incorrect SOH estimates could lead to inappropriate maintenance decisions, premature replacement, or delayed intervention for degraded batteries. The model may also inherit biases from the training data if certain battery chemistries, usage patterns, or environmental conditions are underrepresented. Therefore, EDR-BAT should be used as a decision-support tool together with domain expertise, safety constraints, and periodic calibration, rather than as the sole basis for high-stakes battery management decisions.

REFERENCES

Tianyou Bai and Huan Wang. Convolutional transformer-based multiview information perception framework for lithium-ion battery state-of-health estimation. *IEEE Transactions on Instrumentation and Measurement*, 72:1–12, 2023.

Thorsten Baumhöfer, Manuel Brühl, Susanne Rothgang, and Dirk Uwe Sauer. Production caused variation in capacity aging trend and correlation to initial cell performance. *Journal of Power Sources*, 247:332–338, 2014.

Samy Bengio, Oriol Vinyals, Navdeep Jaitly, and Noam Shazeer. Scheduled sampling for sequence prediction with recurrent neural networks. *Advances in neural information processing systems*, 28, 2015.

Steffen Bockrath, Vincent Lorentz, and Marco Pruckner. State of health estimation of lithium-ion batteries with a temporal convolutional neural network using partial load profiles. *Applied Energy*, 329:120307, 2023.

Li Cai, Jingdong Lin, and Xiaoyong Liao. A data-driven method for state of health prediction of lithium-ion batteries in a unified framework. *Journal of Energy Storage*, 51:104371, 2022.

Liping Chen, Xinyuan Bao, António M Lopes, Changcheng Xu, Xiaobo Wu, Huifang Kong, Suoliang Ge, and Jie Huang. State of health estimation of lithium-ion batteries based on equivalent circuit model and data-driven method. *Journal of Energy Storage*, 73:109195, 2023.

Zheng Chen, Hongqian Zhao, Yuanjian Zhang, Shiquan Shen, Jiangwei Shen, and Yonggang Liu. State of health estimation for lithium-ion batteries based on temperature prediction and gated recurrent unit neural network. *Journal of Power Sources*, 521:230892, 2022.

Akram Eddahech, Olivier Briat, Nicolas Bertrand, Jean-Yves Delétage, and Jean-Michel Vinassa. Behavior and state-of-health monitoring of li-ion batteries using impedance spectroscopy and recurrent neural networks. *International Journal of Electrical Power & Energy Systems*, 42(1): 487–494, 2012.

Xinyu Gu, Khay Wai See, Penghua Li, Kangheng Shan, Yunpeng Wang, Liang Zhao, Kai Chin Lim, and Neng Zhang. A novel state-of-health estimation for the lithium-ion battery using a convolutional neural network and transformer model. *Energy*, 262:125501, 2023.

David A. Howey and Christoph R. Birkl. Oxford Battery Degradation Dataset 1. University of Oxford, 2017. URL `https://ora.ox.ac.uk/objects/uuid:03ba4b01-cfed-46d3-9b1a-7d4a7bdf6fac`. Available at Oxford Research Archive. Accessed: 2026-04-29.

Jin-Hwa Kim, Jaehyun Jun, and Byoung-Tak Zhang. Bilinear attention networks. *Advances in neural information processing systems*, 31, 2018.

Tamara G Kolda and Brett W Bader. Tensor decompositions and applications. *SIAM review*, 51(3): 455–500, 2009.

Gyumin Lee, Daeil Kwon, and Changyong Lee. A convolutional neural network model for soh estimation of li-ion batteries with physical interpretability. *Mechanical Systems and Signal Processing*, 188:110004, 2023.

Penghua Li, Zijian Zhang, Qingyu Xiong, Baocang Ding, Jie Hou, Dechao Luo, Yujun Rong, and Shuaiyong Li. State-of-health estimation and remaining useful life prediction for the lithium-ion battery based on a variant long short term memory neural network. *Journal of power sources*, 459:228069, 2020.

Weihan Li, Neil Sengupta, Philipp Dechent, David Howey, Anuradha Annaswamy, and Dirk Uwe Sauer. One-shot battery degradation trajectory prediction with deep learning. *Journal of Power Sources*, 506:230024, 2021a.

Weihan Li, Neil Sengupta, Philipp Dechent, David Howey, Anuradha Annaswamy, and Dirk Uwe Sauer. Online capacity estimation of lithium-ion batteries with deep long short-term memory networks. *Journal of power sources*, 482:228863, 2021b.

Weihan Li, Jiawei Zhang, Florian Ringbeck, Dominik Jöst, Lei Zhang, Zhongbao Wei, and Dirk Uwe Sauer. Physics-informed neural networks for electrode-level state estimation in lithium-ion batteries. *Journal of Power Sources*, 506:230034, 2021c.

Mingqiang Lin, Jian Wu, Jinhao Meng, Wei Wang, and Ji Wu. State of health estimation with attentional long short-term memory network for lithium-ion batteries. *Energy*, 268:126706, 2023.

Hongao Liu, Chang Li, Xiaosong Hu, Jinwen Li, Kai Zhang, Yang Xie, Ranglei Wu, and Ziyou Song. Multi-modal framework for battery state of health evaluation using open-source electric vehicle data. *Nature Communications*, 16(1):1137, 2025.

Jinhao Meng, Lei Cai, Daniel-Ioan Stroe, Guangzhao Luo, Xin Sui, and Remus Teodorescu. Lithium-ion battery state-of-health estimation in electric vehicle using optimized partial charging voltage profiles. *Energy*, 185:1054–1062, 2019.

Man-Fai Ng, Jin Zhao, Qingyu Yan, Gareth J Conduit, and Zhi Wei Seh. Predicting the state of charge and health of batteries using data-driven machine learning. *Nature Machine Intelligence*, 2(3):161–170, 2020.

Haihong Pan, Zhiqiang Lü, Huimin Wang, Haiyan Wei, and Lin Chen. Novel battery state-of-health online estimation method using multiple health indicators and an extreme learning machine. *Energy*, 160:466–477, 2018.

Carlos Pastor-Fernández, Kotub Uddin, Gael H Chouchelamane, W Dhammika Widanage, and James Marco. A comparison between electrochemical impedance spectroscopy and incremental capacity-differential voltage as li-ion diagnostic techniques to identify and quantify the effects of degradation modes within battery management systems. *Journal of Power Sources*, 360:301–318, 2017.

Cheng Qian, Binghui Xu, Quan Xia, Yi Ren, Bo Sun, and Zili Wang. Soh prediction for lithium-ion batteries by using historical state and future load information with an am-seq2seq model. *Applied Energy*, 336:120793, 2023.

Robert R Richardson, Christoph R Birkl, Michael A Osborne, and David A Howey. Gaussian process regression for in situ capacity estimation of lithium-ion batteries. *IEEE Transactions on Industrial Informatics*, 15(1):127–138, 2018.

Darius Roman, Saurabh Saxena, Valentin Robu, Michael Pecht, and David Flynn. Machine learning pipeline for battery state-of-health estimation. *Nature Machine Intelligence*, 3(5):447–456, 2021.

Bhaskar Saha and Kai Goebel. Battery Data Set. NASA Ames Prognostics Data Repository, NASA Ames Research Center, Moffett Field, CA, 2007. URL https://www.nasa.gov/intelligent-systems-division/discovery-and-systems-health/pcoe/pcoe-data-set-repository/. Available at NASA Prognostics Center of Excellence Data Set Repository. Accessed: 2026-04-29.

Kristen A Severson, Peter M Attia, Norman Jin, Nicholas Perkins, Benben Jiang, Zi Yang, Michael H Chen, Muratahan Aykol, Patrick K Herring, Dimitrios Fraggedakis, et al. Data-driven prediction of battery cycle life before capacity degradation. *Nature energy*, 4(5):383–391, 2019.

Fujin Wang, Zhi Zhai, Zhibin Zhao, Yi Di, and Xuefeng Chen. Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis. *Nature Communications*, 15 (1):4332, 2024. doi: 10.1038/s41467-024-48779-z.

Qiqi Wang, Min Xie, Rui Wang, and Huadong Mo. Fesd: Feature-enhanced structured-state-space diffusion model for battery soh prediction and imputation. *IEEE Transactions on Instrumentation and Measurement*, 2025.

Yinjiao Xing, Eden WM Ma, Kwok-Leung Tsui, and Michael Pecht. An ensemble model for predicting the remaining useful performance of lithium-ion batteries. *Microelectronics reliability*, 53 (6):811–820, 2013.

Duo Yang, Yujie Wang, Rui Pan, Ruiyang Chen, and Zonghai Chen. State-of-health estimation for the lithium-ion battery based on support vector regression. *Applied Energy*, 227:273–283, 2018.

Han Zhang, Xiaofan Gui, Shun Zheng, Ziheng Lu, Yuqi Li, and Jiang Bian. Batteryml: An open-source platform for machine learning on battery degradation. *arXiv preprint arXiv:2310.14714*, 2023.

Jiarui Zhang, Lei Mao, Zhongyong Liu, Kun Yu, and Zhiyong Hu. A bayesian transfer learning framework for assessing health status of lithium-ion batteries considering individual battery operating states. *Applied Energy*, 382:125260, 2025.

Kaiyang Zhou, Ziwei Liu, Yu Qiao, Tao Xiang, and Chen Change Loy. Domain generalization: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 45(4):4396–4415, 2022.

Jiangong Zhu, Yixiu Wang, Yuan Huang, R Bhushan Gopaluni, Yankai Cao, Michael Heere, Martin J Mühlbauer, Liuda Mereacre, Haifeng Dai, Xinhua Liu, et al. Data-driven capacity estimation of commercial lithium-ion batteries from voltage relaxation. *Nature communications*, 13(1):2261, 2022.

A Datasets

We evaluate the proposed method on four public battery degradation datasets and one in-house dataset, including CALCE, NASA, Oxford, MIT, and the in-house dataset. These datasets cover different cell formats, rated capacities, cycling protocols, and degradation patterns. Figure 4 shows the SOH trajectories used in our experiments.

CALCE. The CALCE dataset is provided by the Center for Advanced Life Cycle Engineering at the University of Maryland. We use the CS2 series cells, including CS2_35, CS2_36, CS2_37, and CS2_38. These cells are small prismatic lithium-ion batteries with a rated capacity of 1.1 Ah and a $LiCoO_2$ cathode. They are charged using a constant-current/constant-voltage protocol, where the current is first held at 0.5C until the voltage reaches 4.2 V, followed by a constant-voltage stage until the current drops below 0.05 A. The cells are then discharged under a constant-current protocol.

NASA. The NASA dataset is provided by the NASA Ames Prognostics Center of Excellence and is widely used for battery SOH prediction and remaining useful life prediction. The dataset contains commercial 18650 lithium-ion cells with a rated capacity of 2 Ah. The cells are charged using a constant-current/constant-voltage protocol, where a 1.5 A current is applied until the voltage reaches 4.2 V, followed by constant-voltage charging until the current drops to 20 mA. The discharge stage is conducted under constant current, and different cells may use different cutoff voltages. We use B0005, B0006, B0007, and B0018 in our experiments.

Oxford. The Oxford Battery Degradation Dataset contains long-term aging data from eight Kokam SLPB533459H4 lithium-ion pouch cells with a rated capacity of 740 mAh. These cells use a graphite anode and an LCO/NCO-based cathode. The dataset is collected under a driving-profile-inspired aging protocol and provides capacity degradation information over long-term cycling. It is commonly used for SOH prediction and degradation diagnosis. In this dataset, one recorded sample corresponds to approximately 100 aging cycles.

MIT. The MIT dataset is from the fast-charging battery degradation dataset constructed by Severson et al. It contains 124 commercial 18650 LFP/graphite cells with a rated capacity of 1.1 Ah. The cells are cycled at 30 °C under different fast-charging protocols. We use three cells from the first batch, corresponding to channels 14, 16, and 24, to evaluate cross-cell online SOH prediction under different fast-charging conditions.

In-house dataset. The in-house dataset is collected from small-capacity lithium-ion cells. It contains three LNMC-based cells with a rated capacity of approximately 0.33 Ah. The cells are cycled under constant-current charging and discharging conditions, and voltage, current, and capacity information are recorded during cycling. This dataset is used to further evaluate the applicability of the proposed method under a self-collected experimental setting and a different capacity scale.

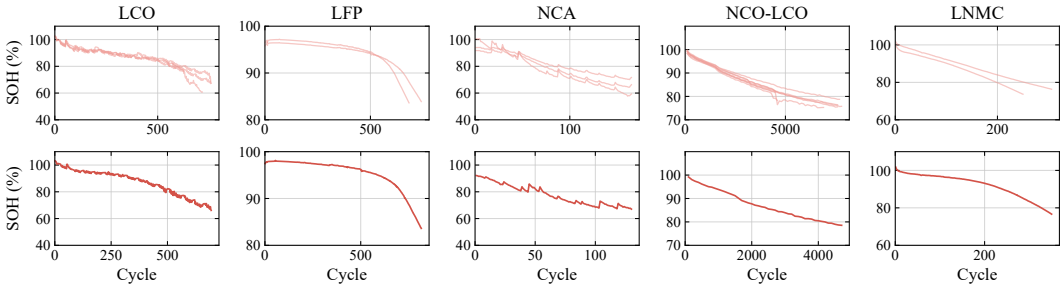


Figure 4: SOH trajectories of the five datasets. The top row shows training cells, and the bottom row shows the target cells used in the main comparison.

B Training Details

All neural models are implemented in PyTorch. We use AdamW with a weight decay of 10^{-4}, cosine annealing learning-rate scheduling, gradient clipping with a threshold of 1.0, and a dropout rate of 0.1. During testing, the batch size is fixed to 1 for cycle-by-cycle online prediction. Computational costs are measured on a workstation with a single NVIDIA GeForce RTX 5060 Ti GPU. The hyperparameters used for EDR-BAT on each dataset are reported in Table 3.

Table 3: Hyperparameters, training details, and computational costs used by EDR-BAT on each dataset.

Parameter	CALCE	MIT	NASA	Oxford	In-house
Number of electrochemical features	10	10	10	10	10
Latent queries	4	4	4	4	4
Attention layers	2	2	2	1	1
Latent dimension	32	32	32	32	32
Attention heads	2	2	2	2	2
Head rank	8	8	4	4	4
Historical electrochemical memory length	8	16	4	4	8
Historical SOH memory length	4	4	2	2	8
Historical SOH embedding dimension	8	8	8	8	8
Feature projection dimension	32	32	32	32	32
SOH roll-out probability $p_{\max}$	1.0	1.0	1.0	1.0	1.0
Learning rate	1e-3	1e-3	1e-3	1e-3	1e-3
Train batch size	32	32	16	4	16
Test batch size	1	1	1	1	1
Training epochs	100	200	200	100	200
Training steps	6900	9200	6600	11300	7200
Total training time (s)	229.86	312.46	258.59	272.19	212.92
Total inference time (s)	7.25	7.69	1.70	0.59	3.23

C Complete Leave-One-Cell-Out Evaluation

To further evaluate cross-cell generalization, we conduct complete leave-one-cell-out experiments on all cells in each dataset. For each run, one cell is held out as the target cell, and all remaining cells from the same dataset are used for training. The same dataset-level hyperparameters reported in Table 3 are used throughout the evaluation.

We compare EDR-BAT with the same baselines used in the main experiments. MAE and RMSE are reported in percentage points. Results are reported as the mean and standard deviation over five random seeds, except for SVR, which is deterministic in our implementation. The complete results are shown in Tables 4–8.

Table 4: Complete leave-one-cell-out results on CALCE.

Cell	Metric	SVR	MLP	CNN	LSTM	Transformer	LSTM-W	Trans-W	PINN4SOH	EDR-BAT
CS2_35	MAE	1.15	$0.87_{0.21}$	$0.73_{0.05}$	$0.56_{0.10}$	$0.52_{0.13}$	$0.60_{0.06}$	$0.48_{0.12}$	$0.75_{0.09}$	$0.65_{0.04}$
	RMSE	3.41	$1.22_{0.38}$	$1.22_{0.22}$	$0.71_{0.13}$	$0.73_{0.14}$	$1.09_{0.07}$	$0.67_{0.16}$	$1.90_{0.59}$	$0.81_{0.12}$
CS2_36	MAE	0.55	$0.80_{0.45}$	$0.76_{0.10}$	$0.45_{0.06}$	$0.42_{0.08}$	$0.60_{0.13}$	$0.44_{0.05}$	$0.60_{0.05}$	$0.40_{0.03}$
	RMSE	0.68	$1.01_{0.52}$	$0.97_{0.12}$	$0.59_{0.07}$	$0.54_{0.07}$	$1.07_{0.10}$	$0.61_{0.06}$	$0.87_{0.07}$	$0.53_{0.04}$
CS2_37	MAE	0.39	$0.61_{0.06}$	$0.60_{0.04}$	$0.45_{0.08}$	$0.41_{0.04}$	$0.53_{0.08}$	$0.38_{0.02}$	$0.40_{0.03}$	$0.33_{0.02}$
	RMSE	0.47	$0.81_{0.08}$	$0.75_{0.04}$	$0.56_{0.09}$	$0.51_{0.04}$	$0.96_{0.07}$	$0.50_{0.03}$	$0.56_{0.06}$	$0.44_{0.02}$
CS2_38	MAE	0.77	$0.69_{0.11}$	$0.65_{0.05}$	$0.55_{0.06}$	$0.51_{0.05}$	$0.62_{0.14}$	$0.55_{0.02}$	$0.75_{0.05}$	$0.43_{0.03}$
	RMSE	1.29	$0.92_{0.06}$	$0.85_{0.05}$	$0.74_{0.02}$	$0.68_{0.10}$	$1.06_{0.09}$	$0.75_{0.06}$	$1.12_{0.06}$	$0.64_{0.02}$

Table 5: Complete leave-one-cell-out results on MIT.

Cell	Metric	SVR	MLP	CNN	LSTM	Transformer	LSTM-W	Trans-W	PINN4SOH	EDR-BAT
channel14	MAE	1.19	$0.50_{0.19}$	$1.62_{0.18}$	$0.24_{0.04}$	$0.60_{0.23}$	$0.23_{0.03}$	$0.48_{0.32}$	$0.40_{0.12}$	$0.32_{0.10}$
	RMSE	1.26	$0.56_{0.16}$	$1.68_{0.19}$	$0.29_{0.04}$	$0.65_{0.23}$	$0.63_{0.02}$	$0.59_{0.29}$	$0.54_{0.14}$	$0.36_{0.10}$
channel16	MAE	0.28	$0.35_{0.10}$	$0.39_{0.16}$	$0.23_{0.10}$	$0.34_{0.11}$	$0.37_{0.19}$	$0.32_{0.13}$	$0.55_{0.15}$	$0.23_{0.04}$
	RMSE	0.37	$0.43_{0.10}$	$0.45_{0.15}$	$0.30_{0.15}$	$0.36_{0.10}$	$0.76_{0.12}$	$0.38_{0.11}$	$0.61_{0.13}$	$0.26_{0.03}$
channel24	MAE	1.65	$0.72_{0.08}$	$0.52_{0.09}$	$0.65_{0.03}$	$0.67_{0.08}$	$0.67_{0.07}$	$0.62_{0.08}$	$0.59_{0.20}$	$0.47_{0.03}$
	RMSE	1.80	$0.81_{0.07}$	$0.61_{0.08}$	$0.71_{0.04}$	$0.71_{0.08}$	$0.93_{0.06}$	$0.69_{0.10}$	$0.68_{0.20}$	$0.50_{0.03}$

Table 6: Complete leave-one-cell-out results on NASA.

Cell	Metric	SVR	MLP	CNN	LSTM	Transformer	LSTM-W	Trans-W	PINN4SOH	EDR-BAT
B0005	MAE	1.44	$0.91_{0.31}$	$1.33_{0.25}$	$0.54_{0.09}$	$0.90_{0.39}$	$0.67_{0.15}$	$1.06_{0.37}$	$1.09_{0.27}$	$0.63_{0.13}$
	RMSE	2.44	$1.21_{0.39}$	$1.75_{0.53}$	$0.71_{0.09}$	$1.11_{0.39}$	$0.92_{0.15}$	$1.30_{0.40}$	$1.37_{0.37}$	$0.87_{0.12}$
B0006	MAE	3.47	$2.35_{0.20}$	$2.58_{0.33}$	$1.91_{0.12}$	$1.25_{0.31}$	$1.74_{0.14}$	$1.10_{0.50}$	$1.80_{0.67}$	$2.01_{0.15}$
	RMSE	4.78	$3.02_{0.28}$	$2.92_{0.36}$	$2.27_{0.17}$	$1.61_{0.30}$	$2.24_{0.16}$	$1.54_{0.49}$	$2.38_{0.84}$	$2.43_{0.15}$
B0007	MAE	1.84	$1.55_{0.24}$	$2.51_{0.23}$	$1.18_{0.13}$	$1.08_{0.30}$	$1.30_{0.30}$	$1.17_{0.29}$	$1.02_{0.08}$	$1.18_{0.26}$
	RMSE	2.16	$1.66_{0.22}$	$2.89_{0.22}$	$1.28_{0.13}$	$1.19_{0.30}$	$1.51_{0.26}$	$1.30_{0.30}$	$1.20_{0.12}$	$1.24_{0.24}$
B0018	MAE	1.80	$1.28_{0.12}$	$1.38_{0.29}$	$0.89_{0.15}$	$0.97_{0.26}$	$1.23_{0.26}$	$0.91_{0.20}$	$0.90_{0.15}$	$0.67_{0.13}$
	RMSE	2.10	$1.46_{0.13}$	$1.66_{0.31}$	$1.04_{0.15}$	$1.19_{0.26}$	$1.49_{0.24}$	$1.19_{0.22}$	$1.19_{0.14}$	$0.84_{0.14}$

Table 7: Complete leave-one-cell-out results on Oxford.

Cell	Metric	SVR	MLP	CNN	LSTM	Transformer	LSTM-W	Trans-W	PINN4SOH	EDR-BAT
Cell1	MAE	0.34	$0.57_{0.11}$	$0.52_{0.13}$	$0.21_{0.05}$	$0.41_{0.40}$	$0.32_{0.15}$	$0.35_{0.12}$	$0.29_{0.14}$	$0.17_{0.04}$
	RMSE	0.42	$0.73_{0.16}$	$0.62_{0.15}$	$0.24_{0.05}$	$0.49_{0.46}$	$0.42_{0.17}$	$0.42_{0.10}$	$0.37_{0.15}$	$0.22_{0.05}$
Cell2	MAE	0.37	$0.74_{0.14}$	$0.69_{0.20}$	$0.41_{0.23}$	$0.36_{0.12}$	$0.65_{0.27}$	$0.39_{0.15}$	$0.35_{0.14}$	$0.33_{0.04}$
	RMSE	0.54	$0.98_{0.20}$	$0.92_{0.24}$	$0.55_{0.24}$	$0.50_{0.16}$	$0.87_{0.32}$	$0.51_{0.13}$	$0.49_{0.15}$	$0.48_{0.06}$
Cell3	MAE	0.36	$0.51_{0.21}$	$0.34_{0.10}$	$0.27_{0.14}$	$0.37_{0.16}$	$0.57_{0.53}$	$0.40_{0.13}$	$0.31_{0.07}$	$0.22_{0.05}$
	RMSE	0.41	$0.64_{0.25}$	$0.42_{0.13}$	$0.32_{0.15}$	$0.43_{0.15}$	$0.66_{0.57}$	$0.47_{0.13}$	$0.39_{0.08}$	$0.27_{0.04}$
Cell4	MAE	0.47	$0.75_{0.48}$	$0.53_{0.14}$	$0.45_{0.21}$	$0.21_{0.06}$	$0.35_{0.16}$	$0.49_{0.20}$	$0.30_{0.05}$	$0.18_{0.02}$
	RMSE	0.55	$0.82_{0.52}$	$0.64_{0.17}$	$0.50_{0.23}$	$0.25_{0.07}$	$0.45_{0.14}$	$0.56_{0.20}$	$0.43_{0.05}$	$0.22_{0.02}$
Cell5	MAE	0.47	$0.74_{0.16}$	$0.76_{0.42}$	$0.41_{0.32}$	$0.22_{0.05}$	$0.34_{0.14}$	$0.43_{0.09}$	$0.22_{0.05}$	$0.23_{0.09}$
	RMSE	0.59	$0.85_{0.20}$	$0.86_{0.43}$	$0.49_{0.38}$	$0.29_{0.06}$	$0.47_{0.14}$	$0.52_{0.05}$	$0.26_{0.05}$	$0.27_{0.10}$
Cell6	MAE	0.55	$0.48_{0.11}$	$0.62_{0.16}$	$0.28_{0.16}$	$0.27_{0.11}$	$0.64_{0.31}$	$0.27_{0.04}$	$0.28_{0.08}$	$0.21_{0.03}$
	RMSE	0.62	$0.61_{0.15}$	$0.72_{0.17}$	$0.39_{0.17}$	$0.38_{0.06}$	$0.78_{0.31}$	$0.40_{0.07}$	$0.38_{0.10}$	$0.34_{0.04}$
Cell7	MAE	0.47	$0.53_{0.27}$	$0.40_{0.11}$	$0.54_{0.33}$	$0.32_{0.24}$	$0.51_{0.40}$	$0.30_{0.11}$	$0.20_{0.06}$	$0.23_{0.09}$
	RMSE	0.57	$0.65_{0.35}$	$0.46_{0.12}$	$0.63_{0.41}$	$0.37_{0.27}$	$0.65_{0.44}$	$0.37_{0.12}$	$0.26_{0.06}$	$0.25_{0.08}$
Cell8	MAE	0.35	$0.46_{0.06}$	$0.71_{0.35}$	$0.36_{0.10}$	$0.23_{0.12}$	$0.45_{0.02}$	$0.41_{0.20}$	$0.28_{0.06}$	$0.21_{0.04}$
	RMSE	0.41	$0.64_{0.11}$	$0.80_{0.33}$	$0.44_{0.11}$	$0.27_{0.13}$	$0.59_{0.04}$	$0.51_{0.20}$	$0.39_{0.09}$	$0.26_{0.05}$

Table 8: Complete leave-one-cell-out results on the in-house dataset.

Cell	Metric	SVR	MLP	CNN	LSTM	Transformer	LSTM-W	Trans-W	PINN4SOH	EDR-BAT
In-house-1	MAE	0.90	$0.90_{0.15}$	$1.03_{0.31}$	$0.50_{0.06}$	$0.43_{0.15}$	$0.68_{0.19}$	$0.52_{0.11}$	$0.87_{0.31}$	$0.58_{0.06}$
	RMSE	1.57	$1.15_{0.20}$	$1.23_{0.36}$	$0.58_{0.07}$	$0.54_{0.18}$	$0.92_{0.21}$	$0.64_{0.13}$	$1.02_{0.29}$	$0.76_{0.02}$
In-house-2	MAE	2.34	$1.02_{0.36}$	$1.78_{0.14}$	$0.93_{0.34}$	$0.70_{0.11}$	$2.53_{0.39}$	$0.61_{0.04}$	$0.97_{0.21}$	$0.85_{0.14}$
	RMSE	2.89	$1.12_{0.36}$	$1.93_{0.13}$	$0.95_{0.34}$	$0.85_{0.17}$	$3.36_{0.54}$	$0.80_{0.06}$	$1.20_{0.24}$	$0.94_{0.14}$
In-house-3	MAE	4.60	$1.06_{0.06}$	$6.32_{0.76}$	$0.92_{0.09}$	$1.57_{0.53}$	$2.12_{0.36}$	$1.21_{0.42}$	$2.60_{0.43}$	$0.82_{0.07}$
	RMSE	4.96	$1.19_{0.03}$	$6.53_{0.80}$	$1.02_{0.10}$	$1.68_{0.53}$	$2.91_{0.49}$	$1.62_{0.29}$	$2.82_{0.46}$	$0.96_{0.08}$

D Additional Prediction Visualizations

We provide additional visualizations on the five target cells used in the main comparison. Figure 5 shows the true-versus-predicted SOH scatter plots of EDR-BAT and the cycle-wise absolute error distributions of different methods. These visualizations provide a more detailed view of the prediction behavior beyond the aggregate MAE and RMSE values reported in Table 1.

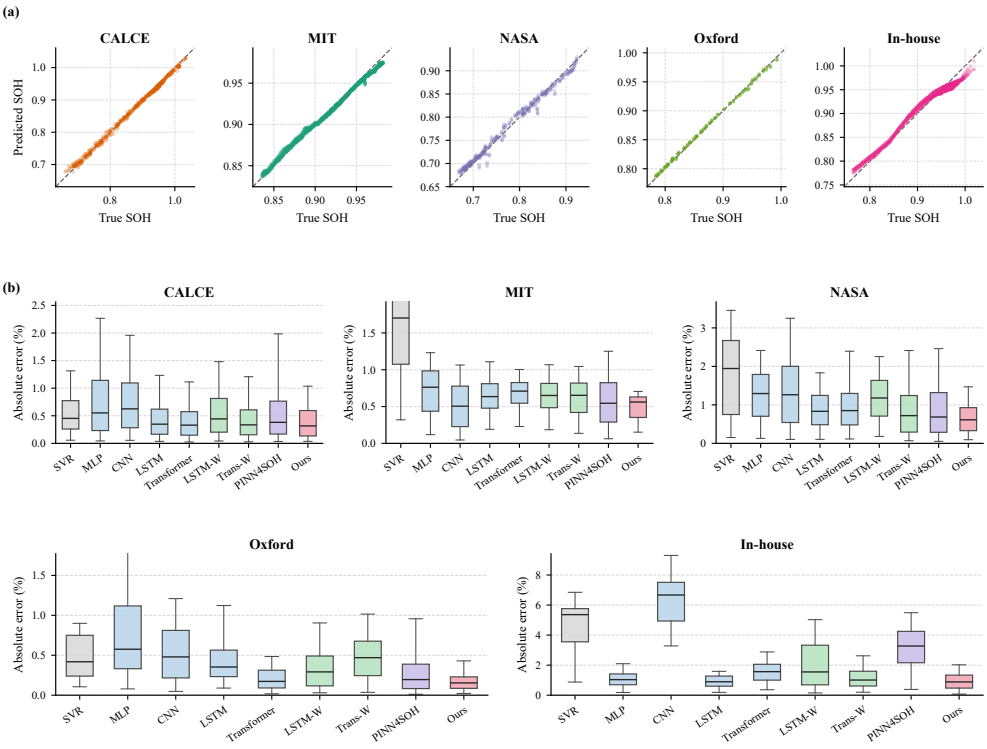


Figure 5: Additional prediction visualizations on the five target cells. (a) True-versus-predicted SOH scatter plots of EDR-BAT, where the dashed diagonal line indicates ideal prediction. (b) Cycle-wise absolute error distributions across different methods, with errors reported in percentage points.

E Sparse Online Bias Calibration

We additionally evaluate a simple sparse online bias calibration setting, where reference SOH observations at the 1st, 5th, 10th, and 50th test cycles are used to adjust subsequent predictions. Anchor cycles are excluded from metric computation. Since this setting uses additional online SOH observations, it is reported only as a practical extension.

Table 9: Sparse online bias calibration results excluding anchor cycles. MAE and RMSE are reported in percentage points, and MAE improvement denotes the relative reduction after calibration.

	No Calib.		Bias Calib.		
Target cell	MAE	RMSE	MAE	RMSE	MAE Improvement
CS2_38	0.439	0.662	0.438	0.655	0.23%
channel24	0.497	0.525	0.472	0.501	5.03%
B0018	0.683	0.861	0.641	0.836	6.15%
Cell8	0.265	0.314	0.243	0.288	8.30%
In-house-3	0.807	0.938	0.699	0.784	13.38%

Figure 6 shows an example on the NASA dataset, where sparse bias calibration reduces part of the prediction offset.

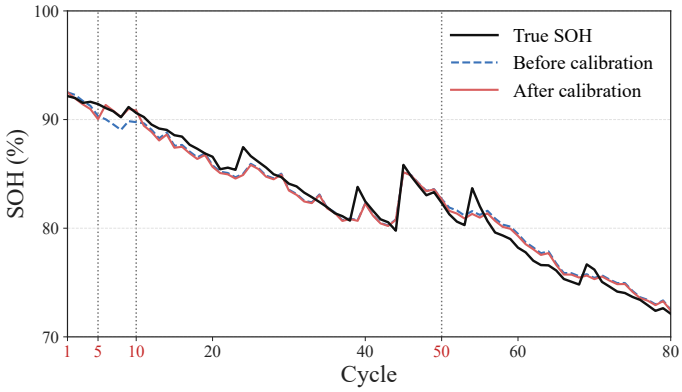


Figure 6: Sparse online bias calibration curve on the NASA dataset. Anchor cycles are set to 1, 5, 10, and 50.

F Sensitivity Analysis of Historical Length

Historical length controls how much past degradation evidence is available for online SOH prediction. We evaluate two memory settings: the maximum historical electrochemical memory length and the historical SOH memory length. As shown in Tables 10–11, the best memory length varies across datasets, reflecting differences in degradation duration, sampling density, and trajectory smoothness. In most cases, moderate historical lengths provide competitive performance, while overly short memories may miss useful degradation context and overly long memories may introduce redundant or less relevant observations.

These results show that EDR-BAT is not highly sensitive to a single fixed memory length, but benefits from using a history range that matches the degradation characteristics of each dataset. Accordingly, we use compact historical lengths that match the degradation scale and sampling characteristics of each dataset.

Table 10: Sensitivity to historical cycle memory length. MAE is reported in percentage points, the best result for each dataset is in bold.

Dataset	1	2	4	8	16
CALCE	0.44	**0.42**	0.44	0.43	0.46
MIT	0.52	0.49	0.51	0.48	**0.47**
NASA	0.92	0.73	0.67	0.66	**0.63**
Oxford	0.18	0.20	0.21	0.20	**0.16**
In-house	0.84	0.85	0.90	**0.82**	0.94

Table 11: Sensitivity to historical SOH memory length. MAE is reported in percentage points, the best result for each dataset is in bold.

Dataset	1	2	4	8	16
CALCE	0.45	0.44	**0.43**	0.46	0.46
MIT	0.45	**0.42**	0.47	0.50	0.44
NASA	0.70	**0.67**	0.72	0.74	0.72
Oxford	0.20	0.21	0.22	**0.16**	0.20
In-house	**0.81**	0.85	0.89	0.82	0.84

Figure 7 provides a more intuitive visualization of these trends.

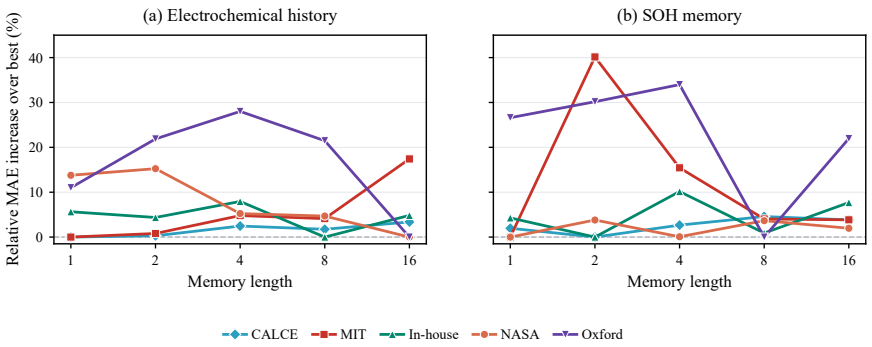


Figure 7: Relative MAE increase over the best memory-length setting. (a) Sensitivity to historical electrochemical memory length. (b) Sensitivity to historical SOH memory length.

G Discussion

Further clarification. EDR-BAT is designed for sequential cross-cell SOH prediction under a strict causal setting: the model only uses current and historical electrochemical observations, without future target-cell information or target-cell fine-tuning. Its main contribution is not a new attention operator, but a structured degradation representation that explicitly separates current electrochemical evidence, historical degradation context, and evolving SOH memory.

Unlike generic sequence models that directly encode a fixed observation window, EDR-BAT constructs a cycle-specific degradation representation by adapting a shared electrochemical prior to recent degradation history, and then explicitly models its interactions with both current-cycle electrochemical evidence and evolving SOH memory. Evidence supports the main claims. Cross-cell generalization is evaluated with complete leave-one-cell-out experiments; robustness to autoregressive SOH history is examined through memory-anchored training ablations; and representation dynamics are analyzed across the battery lifetime.

Future Work. Although EDR-BAT shows promising performance for online SOH prediction, several directions remain open. First, the current experiments are based on representative public and in-house datasets, while real-world battery deployment involves more complex field conditions, including varying temperatures, charging protocols, load profiles, and usage behaviors. Future work will validate EDR-BAT on larger-scale field data, especially fleet-level battery records, to further examine its robustness under diverse operating environments.

Second, EDR-BAT learns degradation-aware representations from multi-modal electrochemical evidence, but the learned representation is still mainly optimized for prediction accuracy. A deeper connection between the representation and physical aging mechanisms remains to be explored. Future studies may examine how the learned degradation representation correlates with interpretable electrochemical factors, such as impedance growth, loss of lithium inventory, capacity fade modes, and nonlinear aging stages.

Third, this work focuses on label-free online prediction for unseen target cells. In practical battery management systems, sparse reference capacity measurements or calibration points may occasionally be available. Future work can extend the current memory-based framework toward adaptive online calibration and cell-specific personalization, allowing the model to better reduce accumulated prediction bias during long-term deployment.