Physics-Constrained Deep Learning for Interpretable Anomaly Detection in Large Battery Packs with Limited Monitoring Granularity

Antoni Plonczak[1], Matthias Wüest [1] and Lilach Goren Huber[1]

[1] *Zurich University of Applied Sciences, Technikumstrasse 9, Winterthur, 8400 Switzerland*
antoni.plonczak@zhaw.ch
matthias.wueest@zhaw.ch
lilach.gorenhuber@zhaw.ch

ABSTRACT

Condition monitoring of large-scale lithium-ion battery packs is often constrained by limited measurement granularity, as operational systems typically rely on aggregated pack-level signals rather than detailed cell-level data. While numerous methods for anomaly detection and health estimation have been developed for individual battery cells, pack-level modeling remains relatively underexplored. Moreover, the impact of reduced monitoring granularity on battery health assessment is still insufficiently understood.

In this work, we present PILSNet, a Physically Interpretable Latent State Network that enables the inference of physically interpretable aging indicators from current-voltage time series. The model combines a convolutional neural network with an equivalent circuit model (ECM), which acts as a constraint to enforce latent variables corresponding to internal resistance and capacity. We apply this hybrid framework to detect abnormal aging behavior in battery packs, achieving both high anomaly detection performance and interpretability of the underlying degradation processes.

Using a physics-based battery simulation framework, we conduct a systematic study of the effect of monitoring data granularity by comparing anomaly detection performance at the cell-level and under aggregated pack-level measurements. The results show that reduced monitoring granularity leads to a significant decrease in anomaly detectability, particularly under realistic scenarios with varying operating conditions and incomplete degradation trajectories. The proposed hybrid model mitigates this performance loss and consistently outperforms purely data-driven and feature engineering-based baselines, especially under constrained data conditions.

In addition to improved detection performance, the inferred latent variables provide direct insight into degradation mechanisms, enabling a pathway from anomaly detection toward fault diagnostics. Beyond battery systems, this work highlights the broader importance of systematically analyzing the relationship between monitoring design and achievable performance in complex systems.

1. INTRODUCTION

Condition monitoring of complex machines and infrastructure in real-world operational systems often requires determining an appropriate monitoring granularity (Zha et al., 2025). In large systems composed of many interconnected components, installing a dense network of sensors that captures the behavior of each individual component can be costly in terms of hardware, data storage, and computational resources. Conversely, reducing the number of sensors or relying on aggregated measurements lowers monitoring resolution and may compromise the quality of condition assessment and the operational decisions derived from it. In particular, the performance of data-driven algorithms such as anomaly detection can be strongly affected when only coarse or aggregated data are available. Despite its practical relevance, the fundamental tradeoff between monitoring granularity and algorithm performance has received limited attention in the scientific literature, with only a few recent studies addressing related questions in areas such as air pollution modeling (Zhalehdoost & Taleai, 2025), electrical load forecasting (Li et al., 2025), wireless sensor networks (Bosman, Iacca, Tejada, Wörtche, & Liotta, 2017), and solar power plant monitoring (Seigneur, 2022; Lüscher, Zgraggen, Guo, Notaristefano, & Huber, 2024).

Battery energy storage systems provide a particularly relevant example of this tradeoff. In large battery installations, measurements are rarely available at the level of detail that would be most informative for condition assessment. A battery pack typically consists of many electrochemically coupled cells

monitored by a battery management system that records voltage, current, and temperature signals. Although individual cell measurements can provide valuable insight into degradation processes, storing and transmitting full cell-level data over long operational lifetimes is often impractical. As a result, many operational systems rely on aggregated pack-level measurements, such as statistical summaries of the cell voltages, rather than detailed signals from each individual cell (Liu, Li, & Yu, 2025).

Reliable monitoring of lithium-ion battery packs is essential in applications ranging from electric mobility and stationary energy storage to industrial power systems. Early detection of abnormal degradation is critical for maintaining safety, extending operational lifetime, and avoiding costly failures. A wide range of anomaly detection (AD) and state-of-health estimation methods have therefore been proposed for battery systems (Guo et al., 2024; Machlev, 2024; Biggio, Bendinelli, Kulkarni, & Fink, 2023; Bajarunas, Baptista, Goebel, & Chao, 2024; Navidi, Bajarunas, Chao, & Hu, 2025). However, the vast majority of these approaches have been developed and evaluated using data from individual battery cells rather than large battery packs.

Extending battery health monitoring methods from individual cells to large-scale battery packs remains an important challenge in current battery research. Most existing pack-level studies continue to assume that individual cell measurements are available (Zhao et al., 2023; Che et al., 2022; Lambert et al., 2026). Only a limited number of approaches operate directly on aggregated pack-level statistics (S. Yang et al., 2024; Qi et al., 2024; Wu, Fang, Dong, & Lin, 2023). Consequently, direct comparisons of anomaly detectability under different monitoring granularities remain scarce. A systematic investigation of the relationship between monitoring granularity and anomaly detection performance has only recently been conducted in our previous work (Wüest, Plonczak, & Goren Huber, 2026).

The present work extends this line of research by introducing a physics-informed deep learning model for anomaly detection and applying it to both cell-level and aggregated pack-level data. This allows us to further investigate how monitoring granularity affects fault detectability in battery packs. In contrast to our previous purely data-driven approach, the proposed model not only detects anomalies but also enables their interpretation in terms of the underlying changes in physical battery properties. Such interpretability provides a first step toward fault diagnostics and enables meaningful analysis of abnormal degradation behavior even when monitoring data are limited to aggregated pack-level measurements.

In the proposed approach, physical interpretability is achieved by incorporating a simple physics-based battery model, an Equivalent Circuit Model (ECM), into the training of a neural network reconstruction model. The ECM acts as a constraint that enforces physically meaningful latent variables corresponding to battery properties such as internal resistance and capacity. While interpretable battery aging indicators have recently been extracted using data-driven approaches for end-of-discharge (Biggio et al., 2023) or battery lifetime (H. Zhang et al., 2025) prediction , these studies focused on individual cells. Similarly, previous work combining ECM-based models with deep learning has primarily addressed cell-level problems (L. Zhang, Xia, & Zhang, 2024; Kuzhiyil, Damoulas, & Dhammika Widanage, 2024; H. Wang & Zhao, 2026). Other interpretable fault detection methods in the battery literature also remain largely restricted to cell-level monitoring (X. Zhang, Liu, Lin, Zhang, & Wang, 2023; Arbaoui, Samet, Ayadi, Mesbahi, & Boné, 2024; Guo et al., 2025; T.-T. Wang, Liu, Peng, & Liu, 2025; Xu et al., 2026).

The hybrid approach proposed in this work combines the strengths of data-driven and physics-based modeling. It uses operational battery signals that are readily available in monitoring systems as inputs to a deep learning model that infers aging-related latent states, while simple physical relationships are used to enable the interpretability of these latent variables. Because the physical model introduces structural constraints, the approach can operate with significantly lower data requirements than purely data-driven models. This makes the method particularly suitable for practical monitoring scenarios where only limited and aggregated measurements are available. As shown here, the same model achieves high AD performance both on individual cells and on aggregated pack-level variables, such as the average, minimum, and maximum of the cell voltages.

The contributions of this work are threefold:

- We introduce a physics-informed deep learning model for detecting abnormal aging behavior in lithium-ion battery packs.
- We present a systematic study of the impact of monitoring data granularity on anomaly detection performance, comparing cell-level measurements with aggregated pack-level data in large-scale battery packs.
- We show that the proposed approach enables diagnostic interpretation of detected anomalies through physically consistent latent aging indicators.

2. DATA DESCRIPTION

To analyze the effect of data aggregation on the detectability of anomalies, we generate a synthetic battery pack dataset using a physics-based simulation framework. The simulations are conducted with `liionpack` (Tranter et al., 2022), which builds on the electrochemical battery modeling library `PyBaMM` (Sulzer, Marquis, Timms, Robinson, & Chapman, 2021) to represent the behavior of individual lithium-ion cells. In total, 20 battery packs are simulated, each consisting of

20 Nickel Manganese Cobalt (NMC) cells (Chen et al., 2020) connected in series and operated over 100 charge-discharge cycles. For each cycle, time-resolved current and voltage measurements are sampled every 30 seconds.

Cell aging is modeled using a standard solid-electrolyte interphase (SEI) growth mechanism (X.-G. Yang, Leng, Zhang, Ge, & Wang, 2017). To capture realistic variability arising from manufacturing tolerances and operating conditions, cells within each pack are assigned heterogeneous initial parameters, including internal resistance, capacity, state of charge, and degradation rate. These parameters are sampled from empirically motivated distributions reported in the literature (Barreras, Raj, Howey, & Schaltz, 2017; Baumhöfer, Brühl, Rothgang, & Sauer, 2014). Abnormal aging is introduced by assigning a single cell in a subset of packs an increased degradation rate, thereby creating a controlled anomaly scenario.

The use of simulated data is essential due to the limited availability of publicly accessible, labeled operational data from large-scale battery packs. The simulation environment enables a systematic investigation of different data aggregation levels under diverse operating regimes, gradually approaching real-world operational conditions.

2.1. Data granularity levels

To reflect different levels of observability in practical battery monitoring systems, two data granularities are considered:

- *Cell-level data*, where the voltage of each individual cell $V_{c,k,t}$ is available for cell c at cycle k and time step t.
- *Pack-level data*, where only aggregated statistics of the cell voltages are available. Specifically, the mean, minimum, and maximum cell voltage are computed at each time step. Since the cells are connected in series, all N_c cells share the same current. The aggregated signals are defined as

$$\begin{aligned} V_{\text{avg},k,t} &= \frac{1}{N_\text{c}} \sum_c V_{c,k,t}, \\ V_{\text{min},k,t} &= \min_c V_{c,k,t}, \\ V_{\text{max},k,t} &= \max_c V_{c,k,t}. \end{aligned} \tag{1}$$

2.2. Data complexity scenarios

We study the effect of data granularity under two data scenarios, representing different data complexity levels:

- *Simplified data.* Battery packs are fully discharged in each cycle. In addition, we assume the availability of the cycle count information as well as complete signals from all past discharge cycles.
- *Towards realistic data.* Inspired by real operational battery data, we study a more realistic scenario in which the depth of discharge (DoD) is not fixed on 100%, and batteries are discharged to a randomly selected SoC level in each cycle (randomly varying DoD). In addition, in this scenario we do not assume having access to the cycle count, nor to a full coverage of the discharge history. This scenario approximates operating conditions in real-world applications, such as electric vehicles and energy storage, where the energy demand per cycle varies, and data from the entire asset lifetime may not be available.

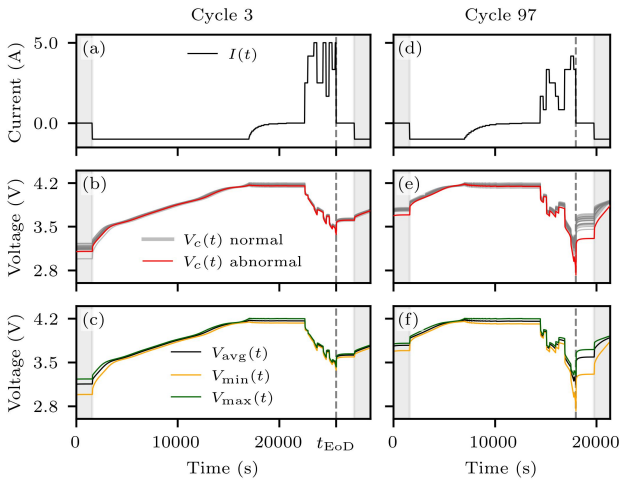

Figure 1. Simulated data examples. Current (top), cell voltages (middle) and aggregated pack voltage statistics (bottom) at early (left) and late (right) cycle counts (Wüest et al., 2026).

Figure 1 shows an example of the simulated current together with the corresponding cell-level and aggregated pack-level voltage signals for an abnormal pack at cycles $k = 3$ and $k = 97$, highlighting the progression of aging effects over time. The dataset is divided into a training set consisting of 10 normal packs and a test set comprising 5 normal and 5 abnormal packs.

3. METHOD

In many battery systems, anomalies do not manifest as entirely new operating regimes, but rather as deviations in the aging behavior of otherwise normal systems. Such anomalies typically correspond to accelerated degradation processes that follow patterns similar to normal aging, but occur earlier or progress more rapidly. This characteristic makes AD based on classical out-of-distribution assumptions challenging. In particular, reconstruction-based approaches such as autoencoders trained on normal data may fail to detect such cases, since abnormal samples can still be reconstructed with low residual error when their behavior remains close to the normal operating manifold.

An alternative strategy is to explicitly infer aging-related information from operational data and detect anomalies by identifying deviations from expected aging trajectories. This idea is closely related to state-of-health (SoH) or End-of-Discharge

(EoD) estimation, where latent indicators of battery degradation are inferred from voltage and current measurements (Biggio et al., 2023). Monitoring the temporal evolution of such indicators over time enables the detection of abnormal aging patterns, even when instantaneous system behavior remains similar to that of normal batteries.

In this work, we adopt this perspective and propose a Physically Interpretable Latent State Network (PILSNet), a physics-informed deep learning approach for extracting latent aging indicators directly from current–voltage time series. A convolutional neural network (CNN) processes the operational signals and produces latent variables representing cycle-specific battery properties. During training, a physical battery model is incorporated to constrain these latent variables to correspond to physically meaningful quantities, internal resistance and effective capacity. Importantly, these constraints are imposed in a fully unsupervised manner: the model requires no ground-truth battery parameters but relies solely on measured current and voltage signals together with a voltage reconstruction objective.

Once trained, the model enables inference of latent battery properties for each cycle via a simple feed-forward evaluation. This allows tracking of aging trajectories for internal resistance and capacity over time. Anomalies are then detected as deviations from expected trajectories of these latent indicators. Expected trajectories are implicitly defined by the distribution of latent variables observed in normal training data. In addition to enabling AD, this representation provides a physically interpretable characterization of the underlying degradation behavior.

The remainder of this section is structured as follows. First, we introduce the physics-informed feature extraction network. Next, we describe the equivalent circuit model used to enforce physical interpretability of the latent variables. Finally, we present the AD procedure based on the inferred latent aging trajectories.

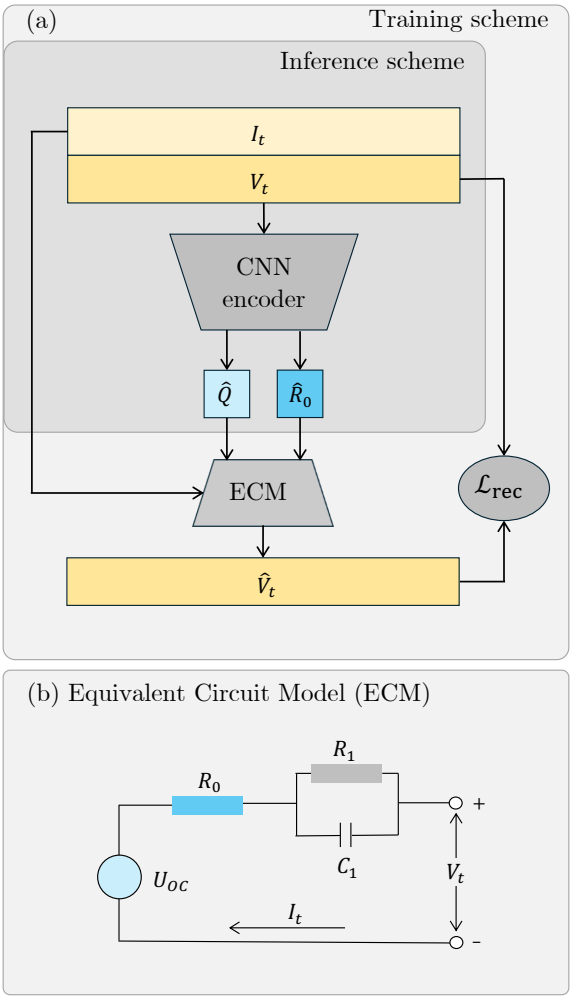

Figure 2. Schematic model overview. (a) Physically interpretable latent state inference network (PILSNet) (b) Equivalent Circuit Model (ECM) used to enforce physically meaningful latent states.

3.1. Physics-Informed CNN for Latent State Inference

In the proposed approach, AD is achieved by inferring battery aging behavior and comparing it to that expected from normal systems. In this way, accelerated degradation can be identified. Moreover, the approach enables separate estimation of key battery parameters, providing insight into the underlying causes of anomalous aging.

As illustrated in Figure 2(a), latent aging indicators are inferred using a CNN that takes current $\{I_t\}_{t=1}^{T}$ and voltage $\{V_t\}_{t=1}^{T}$ signals as input and reconstructs the voltage signal $\{\hat{V}_t\}_{t=1}^{T}$. The inputs are processed by a CNN encoder that extracts temporal features and maps them to a two-dimensional latent space consisting of $\hat{Q}$ and $\hat{R}_0$. These latent variables are then passed through an equivalent circuit model (ECM), which reconstructs the voltage signal given the input current. By imposing an accurate voltage reconstruction, the ECM constrains the inferred latent variables $\hat{Q}$ and $\hat{R}_0$ to correspond to battery capacity and internal resistance, respectively.

The model is trained by minimizing a weighted mean squared reconstruction error of the voltage increments. The increments of the measured and reconstructed voltages, V_t and $\hat{V}_t$, are defined as $\Delta V_t = V_t - V_{t-1}$ and $\Delta \hat{V}_t = \hat{V}_t - \hat{V}_{t-1}$, respectively. The reconstruction loss is given by

$$\mathcal{L}_{\text{rec}} = \sum_i \left((1-\lambda)\,\alpha^{(i)} \sum_t a_t^{(i)}\,\ell_t^{(i)} + \lambda\,\beta^{(i)} \sum_t b_t^{(i)}\,\ell_t^{(i)} \right)$$

where

$$\ell_t^{(i)} = \left(\frac{\Delta V_t^{(i)} - \Delta \hat{V}_t^{(i)}}{\Delta V_t^{(i)}} \right)^2 . \tag{2}$$

The masks $a_t^{(i)}$ and $b_t^{(i)}$ are used to restrict the loss to discharging phases (i.e., $I_t > 0$) and to allow for different weights for time steps with large current variations and smoother regions:

$$\begin{aligned} a_t^{(i)} &= \begin{cases} 1, & \text{if } \Delta I_t > \delta \text{ and } I_t > 0, \\ 0, & \text{otherwise} \end{cases} \\ b_t^{(i)} &= \begin{cases} 1, & \text{if } \Delta I_t \leq \delta \text{ and } I_t > 0, \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{3}$$

Time steps with large current changes emphasize the estimation of the internal resistance R_0, while smoother regions contribute more strongly to the estimation of the capacity Q.

The weights $\alpha^{(i)}$ and $\beta^{(i)}$ normalize the contribution of each sample in the batch and prevent signals with abrupt voltage variations from dominating the loss:

$$\begin{aligned} \alpha^{(i)} &= \left[\sum_t a_t^{(i)} \left| \Delta V_t^{(i)} \right|^{-2} \right]^{-1}, \\ \beta^{(i)} &= \left[\sum_t b_t^{(i)} \left| \Delta V_t^{(i)} \right|^{-2} \right]^{-1}. \end{aligned} \tag{4}$$

The hyperparameters $\delta = 0.1A$ and $\lambda = 0.05$ are selected based on validation performance to achieve accurate voltage reconstruction.

CNN encoder architecture. The encoder consists of three one-dimensional dilated convolutional layers with 32, 16, and 8 channels, respectively. Each layer includes batch normalization, a residual connection, and a ReLU activation function. Exponentially increasing dilation rates (2, 4, and 8) enable efficient extraction of temporal features across multiple time scales. The extracted features are max-pooled, flattened, and passed to a fully connected network with one hidden layer. A sigmoid activation function ensures that the outputs are bounded between 0 and 1, after which they are rescaled to the physically meaningful ranges of capacity and internal resistance. These rescaled latent variables $\hat{Q}$ and $\hat{R}_0$ are then used for ECM-based voltage reconstruction.

With a fixed input length of 200 time steps, the model contains approximately 5,500 trainable parameters, making it suitable for training on limited datasets. To accommodate discharge cycles of varying duration, input sequences shorter than 200 time steps are padded with their zero current and final voltage values. Since training relies solely on reconstruction error, the method is fully unsupervised and does not require labeled data. This is particularly advantageous in real-world settings, where labeled fault data are scarce and datasets may be contaminated with undetected anomalies.

3.2. ECM-Based Physical Constraints for Latent State Interpretability

To constrain the inferred latent variables $\hat{Q}$ and $\hat{R}_0$, we employ a discrete-time first-order Thevenin ECM, illustrated in Figure 2(b):

$$\begin{aligned} V_t &= U_{\mathrm{oc}}(z_t) - R_0 I_t - V_{1,t}, \\ z_t &= z_0 - \frac{\Delta t}{Q} \sum_{j=1}^{t} I_j, \\ V_{1,t} &= V_{1,t-1} + \Delta t \left(-\frac{1}{R_1 C_1} V_{1,t-1} + \frac{1}{C_1} I_t \right). \end{aligned} \tag{5}$$

The voltage V_t is reconstructed as the sum of the open-circuit voltage (OCV) U_{oc}, the ohmic voltage drop $-R_0 I_t$, and the negative polarization voltage $-V_{1,t}$. The OCV is a known function of the state of charge (SOC), z_t, and is commonly assumed to remain unchanged during the battery lifetime. The OCV-SOC curve $U_{\mathrm{oc}}(z_t)$ is provided for all standard batteries by their manufacturer or derived in a designated testing protocol. The SOC at any given time z_t is calculated from the initial SOC z_0 using the accumulated charge $\Delta t \sum_{j=1}^{t} I_j$ divided by the capacity Q. Since we use discharging cycles of fully charged batteries, we can set $z_0 = 1$. The polarization voltage $V_{1,t}$ is modeled by a discrete version of the first-order RC branch dynamics. For simplicity we approximate the capacity Q and the internal resistance R_0 to be constant during each cycle, and vary only across cycles. In addition we fix the values of the polarization parameters $R_1 = 0.02\ \Omega$ and $C_1 = 5000\ \mathrm{F}$ throughout the entire dataset, with their values approximated based on typical range of values from experimental data for NMC Li-ion cells (Thingvad, Calearo, Thingvad, Viskinde, & Marinelli, 2020). While these assumptions introduce minor approximation errors to the inference, they significantly reduce computational complexity.

In the following sections we elaborate on the application of the hybrid PILSNet framework to either cell or pack signals in order to infer interpretable aging indicators that are subsequently used for the detection of abnormal aging.

3.3. Model Adaptation for Cell or Pack Data

The proposed PILSNet model is formulated in a generic manner and can be applied either to individual cell data or to aggregated pack-level signals. This formulation enables a direct comparison of AD performance across different monitoring granularities using a unified model.

Cell-level PILSNet. When current and voltage signals are available for each individual cell, the model takes as input the current $I_{k,t}$ and voltage $V_{c,k,t}$ of cell c at cycle k and time step t. In the series configuration considered here, the current is identical for all cells. The voltage reconstruction is performed under the constraint of a single-cell ECM. As a result, the inferred latent variables $\hat{Q}_{c,k}$ and $\hat{R}_{0,c,k}$ are specific to each cell and cycle.

Pack-level PILSNet. In large-scale battery systems, recording voltage signals for each individual cell is often impractical. Instead, monitoring systems typically store aggregated statistics, as introduced in Eqn. (1). The PILSNet model can be directly applied to these pack-level signals by using $V_{\text{avg},k,t}$, $V_{\text{min},k,t}$, and $V_{\text{max},k,t}$ as inputs.

The latent variables inferred from the average voltage, $\hat{Q}_k^{\text{avg}}$ and $\hat{R}_{0,k}^{\text{avg}}$, can be interpreted as estimates of the mean pack capacity and internal resistance. In contrast, the interpretation of variables inferred from the minimum and maximum voltage signals is less direct. Nevertheless, $\hat{Q}_k^{\text{min}}$, $\hat{R}_{0,k}^{\text{min}}$, $\hat{Q}_k^{\text{max}}$, and $\hat{R}_{0,k}^{\text{max}}$ can be understood as indicators of limiting behavior within the pack. Intuitively, a cell exhibiting increased internal resistance or reduced capacity will influence the aggregated voltage statistics, particularly the minimum voltage during discharge or the maximum voltage during charging. Focusing on discharge cycles, we primarily rely on $\hat{Q}_k^{\text{min}}$ and $\hat{R}_{0,k}^{\text{min}}$ inferred from $V_{\text{min},k,t}$, as these are most sensitive to degraded cells.

It is important to note that the superscript "min" refers to inference based on the minimum voltage signal and does not necessarily correspond to the cell with the minimum capacity or resistance. Since the minimum voltage may originate from different cells at different time steps, these indicators should be interpreted as proxies for the most limiting behavior within the pack. Nevertheless, the minimum voltage during discharge is expected to correlate with cells exhibiting low capacity and/or high internal resistance.

3.4. Anomaly Detection from Latent Aging Trajectories

The inferred latent variables $\hat{Q}$ and $\hat{R}_0$ are used to calculate aging-related indicators. In particular, we focus on the calculation of indicators that capture imbalanced aging among the cells within a pack. These indicators are then used to define cycle-dependent anomaly scores, allowing packs with unusually strong degradation heterogeneity to be identified. The procedure can be applied using either cell-level or pack-level data, depending on the available monitoring signals.

Cell-level indicators. When cell-level voltage signals are available, capacity and internal resistance can be inferred for each cell c and cycle k. Aging indicators are then computed by comparing the inferred cell-specific parameters $\hat{Q}_{c,k}$ and $\hat{R}_{0,c,k}$ to their mean values across all N_c cells in the same pack. The sign of each indicator is chosen such that larger values indicate more abnormal behavior, consistent with the physical expectation that anomalous cells exhibit lower capacity and higher internal resistance:

$$\begin{aligned} f_{c,k}^{(Q)} &= \frac{1}{N_{\text{c}}} \sum_{c'} \hat{Q}_{c',k} - \hat{Q}_{c,k}, \\ f_{c,k}^{(R_0)} &= \hat{R}_{0,c,k} - \frac{1}{N_{\text{c}}} \sum_{c'} \hat{R}_{0,c',k}. \end{aligned} \tag{6}$$

Pack-level indicators. As discussed in the previous section, we focus on aging inference from discharge cycles, for which accelerated degradation is expected to be reflected most strongly in the cell contributing to the minimum pack voltage. We note that an analogous procedure could be applied to the maximum voltage in order to study other degradation patterns. To extract aging-related indicators from pack-level latent variables, we define imbalance metrics using only quantities inferred from aggregated pack-level signals:

$$\begin{aligned} f_{\text{min},k}^{(Q)} &= \hat{Q}_k^{\text{avg}} - \hat{Q}_k^{\text{min}}, \\ f_{\text{min},k}^{(R_0)} &= \hat{R}_{0,k}^{\text{min}} - \hat{R}_{0,k}^{\text{avg}}. \end{aligned} \tag{7}$$

As in the cell-level case, the sign convention is chosen such that larger indicator values correspond to more abnormal behavior, reflecting the physical expectation that anomalous degradation is associated with reduced capacity and increased internal resistance.

Anomaly score derivation. To evaluate AD performance, we derive anomaly scores $s_{p,k}$ per pack p and cycle k from the inferred indicators (6) or (7), depending on the data granularity. We use a unified indicator notation $f_{x,k}^{(y)}$, where $y \in \{Q, R_0\}$ and x denotes the cell index c for cell-level data, and $\min$ for pack-level data. The anomaly score is obtained via the following steps:

1. Smoothing. Moving average over the last $n = 5$ cycles:

$$s_{x,k}^{(y,1)} = \frac{1}{n} \sum_{j=k-n}^{k} f_{x,j}^{(y)}. \tag{8}$$

2. Normalization. For cell-level indicators, we standardize across cells within a pack:

$$s_{x,k}^{(y,2)} = \frac{s_{x,k}^{(y,1)} - \mu_{x,k}^{(y)}}{\sigma_{x,k}^{(y)}}. \tag{9}$$

For pack-level indicators, we normalize the indicators

derived from $V_{\min,k}$ by the ones derived from $V_{\max,k}$:

$$s^{(y,2)}_{\min,k} = \frac{s^{(y,1)}_{\min,k}}{s^{(y,1)}_{\max,k}}. \tag{10}$$

3. Baseline correction. We rescale using the mean $\tilde{\mu}$ and standard deviation $\tilde{\sigma}$ calculated from the training data:

$$s^{(y,3)}_{x,k} = \frac{s^{(y,2)}_{x,k} - \tilde{\mu}^{(y)}_{x,k}}{\tilde{\sigma}^{(y)}_{x,k}}. \tag{11}$$

   In the simplified data scenario, we assume that cycle information is available, thus computing $\tilde{\mu}^{(y)}_{x,k}$ and $\tilde{\sigma}^{(y)}_{x,k}$ per cycle; In the more realistic data scenario we drop this assumption and compute them over all training samples.

4. Pack aggregation. For cell-level models, pack-level scores are obtained via an additional max over the cells in the pack:

$$s^{(y)}_{p,k} = \max_{c \in p} s^{(y,3)}_{c,k}. \tag{12}$$

5. Feature combination and rectification. For each pack and cycle, we take the maximum of the resistance- and capacity-based scores. Based on the physical understanding that anomalies arise either from increased resistance or decreased capacity, we enforce a one-sided anomaly measure by clipping negative values to zero:

$$s_{p,k} = \max\big(0,\ s^{(Q)}_{p,k},\ s^{(R_0)}_{p,k}\big). \tag{13}$$

In the following, the anomaly scores $s_{p,k}$ are used for AD while the physical indicators of Eqns. (6) and (7) provide interpretable insights regarding the contribution of the different physical properties to the aging trajectory.

3.5. Baseline Models

For each data granularity level, we evaluate the AD performance of the proposed hybrid model by comparing it to two baseline models:

1. Physics-Inspired Feature Engineering Model (FEM).
2. Data-driven Deep Latent State Evolution model (DLSE).

The two models are described in length in a separate publication (Wüest et al., 2026). For the sake of completeness we repeat the central ideas here.

Physics-inspired Feature Engineering Model (FEM). To detect abnormal battery aging, we compute a set of physics-inspired features for each cycle and transform them into anomaly scores, where higher scores correspond to a greater likelihood of abnormal behavior. The features are extracted from either cell-level or pack-level data and are evaluated based on their ability to discriminate between abnormal and normal packs. In our previous work we evaluated the AD performance of a large set of features. Here we use only the best performing features as baselines for FEM-based AD performance.

The best performing feature on cell-level is the charge transfer (CT) feature, estimating the difference in charge transferred between the cell voltage and the pack-average voltage over a characteristic voltage interval $[V_1, V_2]$ identified from the incremental capacity analysis (ICA) curve. The CT feature is thus defined as:

$$f^{(\mathrm{CT})}_{c,k} = \Big(\sum_{t \in \mathcal{T}^{\mathrm{avg}}_k} I_k(t) - \sum_{t \in \mathcal{T}^{c}_k} I_k(t) \Big) \Delta t, \tag{14}$$

where $\mathcal{T}^{c}_k = \{t : V_{c,k}(t) \in [V_1, V_2]\}$ and $\mathcal{T}^{\mathrm{avg}}_k$ is defined analogously. Cells with reduced capacity are expected to exhibit larger CT values. ICA-based features are well established in battery diagnostics (Weng, Feng, Sun, & Peng, 2016).

The best performing feature on pack-level is found to be the voltage imbalance (VI) feature, defined as the difference between the average and minimal pack voltage at the end of discharge (EoD):

$$f^{(\mathrm{VI})}_{\min,k} = V_{\mathrm{avg},k}(t_{\mathrm{EoD}}) - V_{\min,k}(t_{\mathrm{EoD}}). \tag{15}$$

The anomaly scores are derived from the physics-inspired features analogously to the above derivation of the PILSNet anomaly scores.

Data-driven Deep Latent State Evolution model (DLSE) As a data-driven baseline, we employ a deep latent-state model inspired by recent approaches for extracting aging-related information from battery time-series data (Bajarunas et al., 2024; Biggio et al., 2023; Bajarunas et al., 2025). The model processes current–voltage time series of a discharge cycle and learns a compact latent representation that separates fast electrical dynamics from slowly evolving battery health. Specifically, a temporal convolutional encoder maps the past current and voltage signals to latent variables, which are then propagated forward in time conditioned on the future current input to predict the corresponding voltage trajectory. The model is trained on cycles from normal packs, using a voltage prediction loss combined with a regularization that enforces slow temporal variation of the health-related latent variables. After training, the slow latent state is projected to a scalar health indicator using principal component analysis (PCA). Deviations of this inferred health indicator from the population mean are used as inputs for anomaly scoring. Two variants of this model are considered: a cell-level model, trained on individual cell voltage signals, and a pack-level model, which operates on aggregated pack voltage statistics (average, minimum, and maximum). In both cases, features are derived from differences between inferred health indicators and subsequently converted to anomaly scores following the lines of

Section 3.4.

The full model description and implementation details are provided in our previous work (Wüest et al., 2026).

4. RESULTS

The proposed PILSNet model is trained on data from the 10 normal training packs, each containing 20 cells cycled to 100 charging-discharging cycles. We evaluate the model in two aspects: (i) AD performance (ii) interpretability of the AD outcomes. In both cases we explicitly analyze the effect of monitoring data granularity by comparing model performance under cell-level and aggregated pack-level data.

4.1. Fault Detectability with Different Data Granularities

The AD performance of the proposed model is evaluated by assigning a normal or abnormal label to each cycle of every test pack, resulting in a test set of 1000 samples. Based on the anomaly scores (13) assigned by the model to each cycle independently, we compute the Receiver Operating Characteristic (ROC) curve. The ROC curve plots the true positive rate (TPR) against false positive rate (FPR) across all decision thresholds, where,

$$\begin{aligned} \text{TPR} &= \frac{\text{TP}}{\text{TP}+\text{FN}}, \\ \text{FPR} &= \frac{\text{FP}}{\text{FP}+\text{TN}} \end{aligned} \tag{16}$$

Here, TP, TN, FP, and FN denote the numbers of true positives, true negatives, false positives, and false negatives, respectively. The area under the ROC curve (AUROC) is used as the threshold-independent metric for evaluating the AD performance of the proposed model and comparing it with the baseline models across the different data granularities.

Table 1 summarizes the AUROC results for all models and

Table 1. AUROC comparison of different models across data granularities. Best results are in bold and second-best are underlined.

Simplified data			
Granularity	Feature Extraction Baseline	Data-driven Baseline	PILSNet (proposed)
Cell-level	**1.00**	0.97	<u>0.99</u>
Pack-level	**0.97**	0.92	**0.97**
Towards realistic data			
Granularity	Feature Extraction Baseline	Data-driven Baseline	PILSNet (proposed)
Cell-level	**0.98**	0.89	**0.98**
Pack-level	<u>0.87</u>	0.77	**0.92**

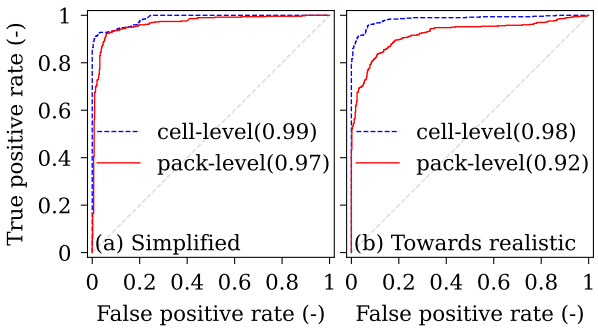

Figure 3. Anomaly detection performance of the PILSNet model. ROC curves with cell- (dashed blue) and pack-level (solid red) data for the (a) simplified and (b) towards realistic data scenarios with corresponding AUROC scores in brackets.

data granularities. The upper half of the table corresponds to a simplified data scenario with fixed operating conditions (i.e full DoD) and full access to the degradation trajectory. The lower half presents results for a more realistic scenario with varying DoD and limited data availability. Below we analyze the AD results under the two scenarios.

Simplified data scenario. A consistent observation across all models is the presence of a *detectability gap* between cell-level and pack-level data. Reducing monitoring granularity leads to a decrease in AD performance, with AUROC drops ranging from 0.02 for PILSNet to 0.05 for the DLSE baseline.

Comparing the models, PILSNet outperforms the purely data-driven DLSE approach for both data granularities. At the cell-level, PILSNet achieves an AUROC of 0.99 compared to 0.97 for DLSE. This performance gap becomes more pronounced under pack-level monitoring, where PILSNet achieves 0.97 compared to 0.92 for DLSE.

In comparison to the physics-inspired feature engineering baseline (FEM), PILSNet achieves comparable performance. At the cell-level, FEM slightly outperforms PILSNet (1.00 vs. 0.99 AUROC), while both methods reach similar performance under pack-level monitoring (0.97 AUROC). This indicates that the proposed hybrid approach maintains high detection accuracy without the need for manual feature engineering.

Towards realistic data scenario. The lower half of Table 1 shows the results under more realistic conditions, where the operating conditions vary and data availability is reduced. In this setting, two additional challenges arise: (i) a learning challenge due to strong variation in DoD across cycles, and (ii) the absence of explicit cycle count information, which prevents a per-cycle comparison with normal training reference trajectories. This scenario reflects practical monitoring conditions in many real-world battery systems, but is often overlooked in

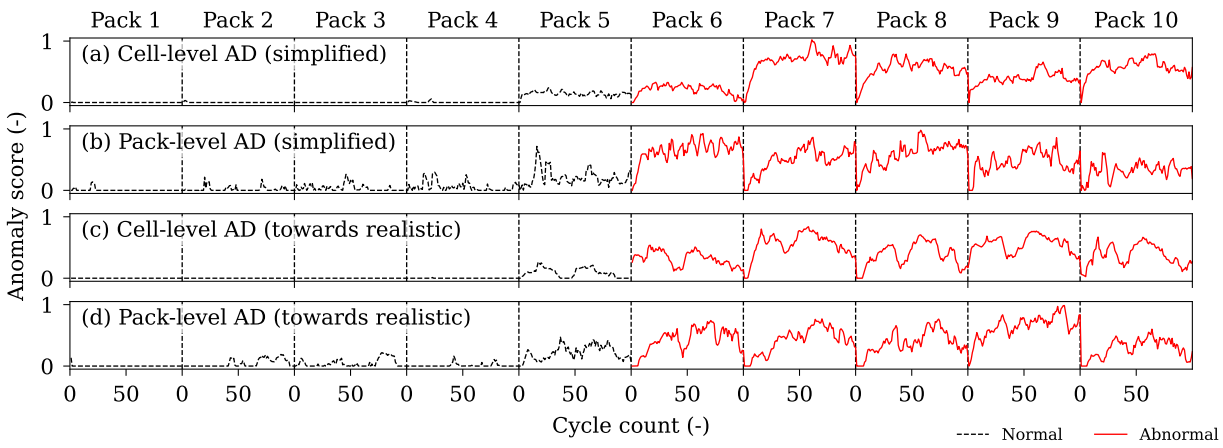

Figure 4. Anomaly scores derived for different data variants. Score evolution vs. cycle count is plotted for 5 normal (left) and 5 abnormal (right) battery packs with 100 cycles per pack, using the proposed PILSNet model with (a) the simplified cell-level data, (b) simplified pack-level data, (c) towards realistic cell-level data and (d) towards realistic pack-level data. Scores of abnormal packs are marked in red.

present research when relying on access to the cycle count.

As expected, the increased difficulty of this setting is reflected in consistently lower AUROC values across all models and data granularities, compared with the simplified scenario. Furthermore, the effect of data aggregation becomes more pronounced, with detectability gaps between cell and pack models ranging from 0.06 for PILSNet to 0.12 for the baseline methods.

A comparison of the models reveals a clear dependence on data granularity. When cell-level data are available, PILSNet and FEM achieve similarly high performance (AUROC of 0.98), both significantly outperforming the DLSE model (0.89). However, under pack-level monitoring, PILSNet clearly outperforms both baselines, achieving an AUROC of 0.92 compared to 0.87 for FEM and 0.77 for DLSE. This highlights the advantage of the proposed hybrid approach, which effectively combines data-driven and physics-based information, particularly under constrained data conditions.

Overall, these results demonstrate that the proposed method reduces the detectability drop due to limited monitoring granularity, suggesting strong potential for practical deployment in real-world battery systems.

The ROC curves of the PILSNet model for both data granularities are shown in Fig. 3 for (a) the simplified and (b) the realistic data scenarios. In both cases, cell-level performance (dashed blue) is compared with pack-level performance (solid red), with corresponding AUROC values indicated.

A more detailed comparison of AD behavior is provided in Fig. 4, which shows anomaly scores for each pack and cycle based on the inferred latent states $\hat{Q}_k$ and $\hat{R}_{0,k}$.

Row (a) shows results obtained using cell-level data. Packs 1-5 (left) contain only normal cells, while packs 6-10 (right) each contain one abnormal cell with accelerated degradation. A clear separation between normal and abnormal behavior is observed, with consistently higher anomaly scores for abnormal cells across all cycles.

Row (b) presents the corresponding results for pack-level data. The anomaly scores increase during the early cycles and stabilize at elevated levels for abnormal packs. A large fraction of normal cycles yields an anomaly score of zero, which is a consequence of the one-sided anomaly definition in (13).

Rows (c) and (d) show the corresponding results under the more realistic data scenario. While anomalies remain detectable, their separation from normal behavior is somewhat reduced compared to the simplified scenario, consistent with the observed decrease in AUROC.

In addition to its strong AD performance, the proposed PILSNet approach offers practical advantages over manual feature engineering. In particular, it avoids the need for extensive design and tuning of handcrafted features, which can become unreliable in the presence of realistic variability and noise (Wüest et al., 2026). Instead, the model provides a unified framework that can be applied consistently across different data types, including both cell-level and aggregated pack-level signals. This flexibility is particularly beneficial in operational settings with heterogeneous and limited data availability.

4.2. Physical Interpretability of AD Outcomes

While the previous section demonstrated that PILSNet achieves strong AD performance across different data granularities, a key advantage of the proposed approach lies in

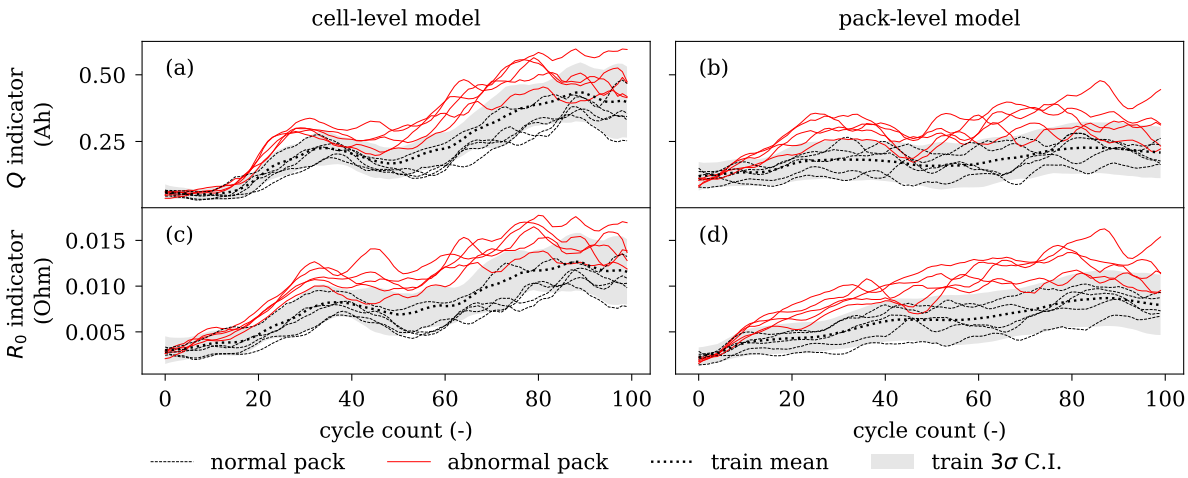

Figure 5. Interpretable aging indicators at cell- and pack-level. The cycle dependent evolution of the inferred capacity Q (upper panels) and internal resistance R_0 (lower panels) indicators for the 10 test battery packs are derived with the cell-level model ((a) and (c)) and the pack-level model ((b) and (d)). The mean and confidence intervals (CI) of the training packs are shown as a dotted line and shaded grey respectively, and abnormal test packs are plotted in red.

the interpretability of its outputs. In contrast to the black-box DLSE model, PILSNet provides direct access to physically meaningful latent variables, namely the inferred capacity $\hat{Q}$ and internal resistance $\hat{R}_0$. These quantities enable a more detailed analysis of the detected anomalies and can provide insight into their underlying causes.

Beyond the anomaly scores themselves, the method allows for tracking the temporal evolution of the latent aging indicators. Figure 5 illustrates the evolution of indicators derived from $\hat{Q}_k$ (panels (a) and (b)) and $\hat{R}_{0,k}$ (panels (c) and (d)) as a function of the cycle count k. The left panels (a) and (c) correspond to the cell-level model (Eqn. (6)), while the right panels (b) and (d) show the corresponding results for pack-level data (Eqn. (7)). In all cases, normal packs are shown in black and abnormal packs in red. The dotted black line indicates the mean value of the latent variable over the training data, while the shaded grey region represents the corresponding 3σ confidence interval. For both cell-level and pack-level models, a single indicator per pack is shown; in the cell-level case, this corresponds to the maximum value across all cells.

The inferred dynamics of the aging indicators depend on the monitoring granularity. This is partly due to differences in inference accuracy between cell-level and pack-level models, and partly due to differences in the physical interpretation of the inferred variables. For cell-level data, aging indicators are estimated for each individual cell, and pack-level indicators are obtained by aggregating these values (e.g., via the maximum across cells). In contrast, for aggregated pack-level data, the indicators inferred from statistics such as the minimum voltage do not directly correspond to the minimum capacity or maximum resistance of individual cells, but rather serve as proxies for bounds on these quantities.

Despite this limitation, the availability of physically interpretable aging indicators enables a detailed analysis of their temporal behavior throughout the battery lifetime. This provides additional insight into degradation processes and supports the interpretation of detected anomalies. In particular, separating the contributions of capacity fade and resistance increase enables distinguishing between different degradation mechanisms, which is essential for actionable diagnostics in practical battery systems. As such, the proposed approach offers a pathway from AD toward fault diagnostics, even under limited monitoring granularity. A deeper study of this capability, comparing different types of anomalies and degradation modes, is the topic of our future research.

5. CONCLUSIONS

Health monitoring of lithium-ion battery packs has received increasing attention in both research and industrial applications. While numerous methods for fault detection and state-of-health estimation have been developed for individual battery cells, only a limited number of approaches have been validated on aggregated measurements from large-scale battery packs, despite the strong practical relevance of such data in real-world systems.

In this work, we proposed a physics-informed hybrid deep learning model capable of inferring physically interpretable aging indicators from both cell-level data and low-granularity pack-level measurements, where only aggregated voltage statistics are available.

The main findings and contributions of this study can be sum-

marized as follows:

- We demonstrate the importance of systematically analyzing the effect of monitoring data granularity on the observability of battery health indicators. Such analyses remain largely absent in current battery research, despite their critical role in designing effective monitoring strategies. Our results show that reduced data granularity leads to a consistent decrease in anomaly detection performance when moving from cell-level to pack-level monitoring.
- We further show that the effect of monitoring granularity becomes more pronounced under more realistic conditions. By explicitly studying scenarios with varying operating conditions and incomplete degradation trajectories, we highlight the importance of accounting for reality gaps when evaluating anomaly detection methods based on simulated data.
- The proposed physics-informed hybrid model mitigates the performance loss associated with reduced data granularity. It consistently outperforms both purely data-driven and feature engineering-based approaches across different monitoring scenarios.
- The advantage of the hybrid approach is particularly pronounced under realistic conditions, where monitoring granularity is low, data availability is limited, and operating conditions vary significantly.
- By providing physically interpretable latent variables corresponding to capacity and internal resistance, the proposed approach enables insight into the underlying degradation mechanisms. This establishes a pathway from anomaly detection toward fault diagnostics, extending beyond purely detection-based methods. This path will be pursued in future research by including multiple aging mechanisms in the simulated data.

The aim of the present work is to highlight important trends and dependencies. Studying model performance under varying data granularities, while gradually closing the sim-to-real gap marks the beginning of this effort. Future work will continue by further enhancing the simulation environment and extending the model's applicability. This includes incorporating additional physical effects in the data, such as temperature dependence, arbitrary initial state-of-charge levels as well as realistic noise sources. In addition extending the framework to infer a broader set of ECM parameters and allowing their time dependence within the cycle may improve the inference performance. Finally, the ultimate step to close the sim-to-real gap is validating the approach on real-world operational battery data.

Beyond the battery domain, this work highlights the broader importance of systematically studying the relationship between monitoring data granularity and algorithm performance. Such tradeoffs are relevant across a wide range of applications, including power systems, transportation networks, industrial production systems, communication networks, and environmental monitoring. Despite their practical importance, these effects remain insufficiently explored in many fields, and further research is needed to better understand how monitoring design influences achievable performance in complex systems and in particular in data-driven condition monitoring systems.

ACKNOWLEDGEMENTS

We would like to thank Manuel Arias Chao, Kristupas Bajarunas and Jokin Alcibar for the insightful discussions.

REFERENCES

Arbaoui, S., Samet, A., Ayadi, A., Mesbahi, T., & Boné, R. (2024). Data-driven strategy for state of health prediction and anomaly detection in lithium-ion batteries. *Energy and AI*, *17*, 100413.

Bajarunas, K., Baptista, M. L., Goebel, K., & Chao, M. A. (2024). Health index estimation through integration of general knowledge with unsupervised learning. *Reliability Engineering & System Safety*, *251*, 110352.

Bajarunas, K., Navidi, S., Dimitrios, Z., Hu, C., Goebel, K., & Chao, M. A. (2025). Degradation-dynamics-informed learning for future-aware battery health prognostics. *Available at SSRN 5928714*.

Barreras, J. V., Raj, T., Howey, D. A., & Schaltz, E. (2017). Results of screening over 200 pristine lithium-ion cells. In *2017 ieee vehicle power and propulsion conference (vppc)* (pp. 1–6).

Baumhöfer, T., Brühl, M., Rothgang, S., & Sauer, D. U. (2014). Production caused variation in capacity aging trend and correlation to initial cell performance. *Journal of Power Sources*, *247*, 332–338.

Biggio, L., Bendinelli, T., Kulkarni, C., & Fink, O. (2023). Ageing-aware battery discharge prediction with deep learning. *Applied Energy*, *346*, 121229.

Bosman, H. H., Iacca, G., Tejada, A., Wörtche, H. J., & Liotta, A. (2017). Spatial anomaly detection in sensor networks using neighborhood information. *Information Fusion*, *33*, 41–56.

Che, Y., Deng, Z., Li, P., Tang, X., Khosravinia, K., Lin, X., & Hu, X. (2022). State of health prognostics for series battery packs: A universal deep learning method. *Energy*, *238*, 121857.

Chen, C.-H., Planella, F. B., O'regan, K., Gastol, D., Widanage, W. D., & Kendrick, E. (2020). Development of experimental techniques for parameterization of multi-scale lithium-ion battery models. *Journal of The Electrochemical Society*, *167*(8), 080534.

Guo, F., Couto, L. D., Mulder, G., Trad, K., Hu, G., Capron, O., & Haghverdi, K. (2024). A systematic review of electrochemical model-based lithium-ion battery state estimation in battery management systems. *Journal of*

Energy Storage, *101*, 113850.

Guo, F., Xu, K., Zhang, Z., Zhou, H., Chen, G., Hu, J., ... Mo, H. (2025). Battery soh prediction under different conditions via mblstm and itransformer with anomaly detection and explainability. *IEEE Open Journal of the Computer Society*.

Kuzhiyil, J. A., Damoulas, T., & Dhammika Widanage, W. (2024). Neural equivalent circuit models: Universal differential equations for battery modelling. *Applied Energy*, *371*, 123692.

Lambert, P., Drummond, R., Ross, J. P., Tredenick, E. C., Howey, D. A., & Duncan, S. R. (2026). Detecting faulty lithium-ion cells in large-scale parallel battery packs using current distributions. *Communications Engineering*, *5*(1), 17.

Li, H., Heleno, M., Zhang, W., Sun, K., Garcia, L. R., & Hong, T. (2025). A cross-dimensional analysis of data-driven short-term load forecasting methods with large-scale smart meter data. *Energy and Buildings*, 115909.

Liu, S., Li, K., & Yu, J. (2025). Battery pack condition monitoring and characteristic state estimation: Challenges, techniques, and future prospectives. *Journal of Energy Storage*, *105*, 114446.

Lüscher, M. F., Zgraggen, J., Guo, Y., Notaristefano, A., & Huber, L. G. (2024). Data scarcity in fault detection for solar tracking systems: the power of physics-informed artificial intelligence. In *Phm society european conference* (Vol. 8, pp. 8–8).

Machlev, R. (2024). Ev battery fault diagnostics and prognostics using deep learning: Review, challenges & opportunities. *Journal of Energy Storage*, *83*, 110614.

Navidi, S., Bajarunas, K., Chao, M. A., & Hu, C. (2025). Forecasting battery capacity for second-life applications using physics-informed recurrent neural networks. *eTransportation*, *25*, 100432.

Qi, Q., Liu, W., Deng, Z., Li, J., Song, Z., & Hu, X. (2024). Battery pack capacity estimation for electric vehicles based on enhanced machine learning and field data. *Journal of Energy Chemistry*, *92*, 605–618.

Seigneur, H. (2022). *Lcoe reduction through proactively optimized monitoring of pv systems (final technical report)* (Tech. Rep.). Univ. of Central Florida, Orlando, FL (United States).

Sulzer, V., Marquis, S. G., Timms, R., Robinson, M., & Chapman, S. J. (2021). Python battery mathematical modelling (pybamm). *Journal of Open Research Software*, *9*(1).

Thingvad, M., Calearo, L., Thingvad, A., Viskinde, R., & Marinelli, M. (2020). Characterization of nmc lithium-ion battery degradation for improved online state estimation. In *2020 55th international universities power engineering conference (upec)* (p. 1-6). doi: 10.1109/UPEC49904.2020.9209879

Tranter, T., Timms, R., Sulzer, V., Planella, F., Wiggins, G., Karra, S., ... others (2022). liionpack: A python package for simulating packs of batteries with pybamm. *Journal of Open Source Software*, *7*(70).

Wang, H., & Zhao, H. (2026). Differentiable physics information neural network with equivalent circuit model constraints for lithium-ion battery parameter identification. *Journal of Energy Storage*, *155*, 121685.

Wang, T.-T., Liu, K.-Y., Peng, H.-J., & Liu, X. (2025). Interpretable machine learning for battery prognosis: Retrospect and prospect. *Advanced Energy Materials*, *15*(48), e03067.

Weng, C., Feng, X., Sun, J., & Peng, H. (2016). State-of-health monitoring of lithium-ion battery modules and packs via incremental capacity peak tracking. *Applied Energy*, *180*, 360–368.

Wu, J., Fang, L., Dong, G., & Lin, M. (2023). State of health estimation for lithium-ion batteries in real-world electric vehicles. *Science China Technological Sciences*, *66*(1), 47–56.

Wüest, M., Plonczak, A., & Goren Huber, L. (2026). Detecting anomalies under data scarcity: Data granularity tradeoffs in large battery packs. In *2026 ieee swiss conference on data science and ai (sds)* (p. 91-98). doi: 10.1109/SDS70563.2026.00020

Xu, W., Mao, R., Han, P., Yuan, N., Li, Y., Guo, Y., & Zhang, H. (2026). Graph neural network modeling of lithium-ion battery capacity with physics-guided features. *Journal of Energy Storage*, *155*, 121627.

Yang, S., Zhang, C., Chen, H., Wang, J., Chen, D., Zhang, L., & Zhang, W. (2024). A hierarchical enhanced data-driven battery pack capacity estimation framework for real-world operating conditions with fewer labeled data. *Journal of Energy Chemistry*, *91*, 417–432.

Yang, X.-G., Leng, Y., Zhang, G., Ge, S., & Wang, C.-Y. (2017). Modeling of lithium plating induced aging of lithium-ion batteries: Transition from linear to nonlinear aging. *Journal of Power Sources*, *360*, 28–40.

Zha, D., Bhat, Z. P., Lai, K.-H., Yang, F., Jiang, Z., Zhong, S., & Hu, X. (2025). Data-centric artificial intelligence: A survey. *ACM Computing Surveys*, *57*(5), 1–42.

Zhalehdoost, A., & Taleai, M. (2025). Unravelling the importance of spatial and temporal resolutions in modeling urban air pollution using a machine learning approach. *Scientific Reports*, *15*(1), 27708.

Zhang, H., Li, Y., Zheng, S., Lu, Z., Gui, X., Xu, W., & Bian, J. (2025). Battery lifetime prediction across diverse ageing conditions with inter-cell deep learning. *Nature Machine Intelligence*, *7*(2), 270–277.

Zhang, L., Xia, B., & Zhang, F. (2024). Adaptive fault detection for lithium-ion battery combining physical model-based observer and bilstmnn learning approach. *Journal of Energy Storage*, *91*, 112067.

Zhang, X., Liu, P., Lin, N., Zhang, Z., & Wang, Z. (2023). A novel battery abnormality detection method using inter-

pretable autoencoder. *Applied Energy*, *330*, 120312.

Zhao, D., Zhou, Z., Zhang, P., Zhang, Y., Feng, Z., Yang, Y., & Cao, Y. (2023). Health condition assessment of satellite li-ion battery pack considering battery inconsistency and pack performance indicators. *Journal of energy storage*, *60*, 106604.