Aging States Estimation and Monitoring Strategies of Li-Ion Batteries Using Incremental Capacity Analysis and Gaussian Process Regression

Moritz Landwehr, Patrick Hoher and Johannes Reuter

HTWG University of Applied Sciences, Konstanz, Germany
moritz.landwehr@htwg-konstanz.de
phoher@htwg-konstanz.de
jreuter@htwg-konstanz.de

ABSTRACT

Existing approaches for battery health forecasting often rely on extensive cycling histories and continuously monitored cells. In contrast, many real-world scenarios provide only sparse information, e.g. a single diagnostic cycle. In our study, we investigate state of health (SoH)- and remaining useful life (RUL) estimation of previously unseen lithium-ion cells, relying on cycling data from begin of life (BOL) to end of life (EOL) of multiple similar cells by using the publicly available Oxford battery aging dataset. The estimator applies incremental capacity analysis (ICA)-based feature extraction in combination with data-efficient regression methods. Particular emphasis is placed on a multi-model Gaussian process regression ensemble approach (GPR_n), which also provides uncertainty quantification. Due to a rather cell invariant behaviour, the mapping of ICA features to SoH estimation is highly precise and points out a normalized mean absolute error (NMAE) of 1.3%. The more cell variant mapping to RUL estimation is challenging, reflecting in a NMAE of 5.3%. Using the estimation results, a RUL monitoring strategy is derived. The objective is to safely operate a battery cell from BOL to EOL by only taking sparse diagnostic measurements. On average, only four diagnostic measurements are required during a cell's lifetime of 3300 to 5000 cycles.

1. INTRODUCTION

1.1. Motivation

Lithium-ion batteries are a key technology for electric vehicles, stationary storage, and portable electronics and their aging directly impacts system reliability, safety, and cost. Capacity fade and resistance increase reduce usable energy and power, while uncontrolled degradation can lead to unplanned downtime and safety concerns. Accurate estimation of SoH and RUL is therefore essential for predictive maintenance, lifetime extension, and economic decision-making, especially in fleet and second-life applications. In many practical scenarios, however, only a few diagnostic cycles are available instead of continuous, long-term monitoring of individual cells. Methods that rely on dense cycling histories are thus difficult to apply, which creates a need for data-efficient approaches that extract information-rich features from sparse measurements and provide reliable SoH- and RUL estimates with quantified uncertainty.

1.2. Battery Modelling and Estimation Methods

In (Demirci, Taskin, Schaltz, & Acar Demirci, 2024) and (Schmitt, 2022) an overview about different methods of SoH estimation is shown, whereas RUL estimation methods are grouped and summarized in (Elmahallawy, Elfouly, Alouani, & Massoud, 2022). Both, SoH- and RUL estimation, deal with long term dynamic behaviour of battery cells, which is in contrast to state of charge (SoC) estimation. The methods for SoH- and RUL estimation can be grouped in model-based and experimental methods. Regarding model-based methods, battery aging can be modelled using a broad spectrum of approaches with different complexity, accuracy and computational intensity (Hamar, 2024). Most widespread modelling approaches are electro-chemical models (Santhanagopalan, Guo, Ramadass, & White, 2006), electric equivalent circuit models (Guo & Shen, 2021), electrochemical impedance models (Jossen, 2006), open circuit potential models, and data-driven models. For SoH estimation there are also several experimental methods, dividing in direct measurement methods, like impedance measurement and coulomb counting, and indirect analysis methods, like ICA and differential voltage analysis (DVA) (Demirci et al., 2024). Electro-chemical models typically require extensive parametrization, detailed knowledge of internal cell chemistry, and significant computational

effort, which limits their scalability and applicability across different cell types and operating conditions. For electric equivalent circuit models, the aging behaviour of the circuit elements has to be mathematically described and optimized using a large, high resolution dataset. Data-driven, specifically machine learning (ML)-approaches, have therefore gained increasing attention (Dar & Singh, 2025).

1.3. Machine Learning in Battery State Estimation

(Al-Hashimi et al., 2025) review, that different neural networks topologies, like fully connected feedforward neural networks (FFNNs), long short-term neural networks (LSTMs), support vector regressions (SVRs), Gaussian process regressions (GPRs), and deep learning architectures have been successfully applied to SoH- and RUL estimation. Mainly voltage, current, temperature, and capacity over time from charging and discharging cycles are used as regressor inputs. They are comparably easy to measure. Intuitively, these features can be directly fed into one of the previously mentioned regressors. However it is a challenge to map the sparse information density of raw measurement data to meaningful features, describing the battery aging process. In the literature there are some examples given. In (Jha, Dorkar, Biswas, & Emadi, 2024) a deep learning transformer network directly estimates the RUL with time based features as inputs. Also (Nguyen Van & Quang, 2023) processes time based input data with a LSTM for SoH estimation. There are also regressors, like convolutional neural networks (CNNs), including inherently feature extraction (Zhou, Vuylsteke, Anderson, & Sun, 2025). All this approaches are based on observing batteries from BOL to a certain age as training data source, to afterwards predict the development of the aging until EOL. In contrast, we deal with the task of a SoH- and RUL estimation of previously unseen cells using cycling data from BOL to EOL of multiple similar cells. In this case, due to cell-to-cell variance the direct processing of sparse information dense time based features is critical. Here ICA has emerged as a powerful non-destructive diagnostic technique to extract high information dense aging-related features.

1.4. Battery State Estimation Using ICA

As (She et al., 2023) describe, the ICA method transforms voltage plateaus and inflection points on the voltage-capacity curves to identifiable peaks and valleys on the corresponding incremental capacity (IC) curves. The characteristic shape of the IC curves directly corresponds to degradation mechanisms, like loss of lithium inventory (LLI), loss of active material (LAM) and ohmic resistance increase (ORI). LLI is generally the primary source, leading to direct capacity fade (Ansean et al., 2019). So ICA curves retain a certain degree of physical interpretability. Thanks to these compelling properties, during the last ten years SoH- and RUL estimation based on ICA features have become popular. (Wang, Cui, Li, Cui, & Yuan, 2023) maps the extracted ICA features simply by a low order polynomial fit to the SoH estimation. They are selecting one or more features depending on charging or discharging cycle and the initial SoC. A further advanced method is suggested in (Li, Yuan, & Wang, 2020), where the mapping from ICA features to SoH is done with a trained SVR. In both papers, training data is also taken from similar cells, while the test cell is regarded as previously unseen. Also RUL estimation is already done using ICA features, (Xia, Wang, & Chen, 2023) apply a bidirectional gated recurrent unit (BiGRU) for RUL estimation. (Pang et al., 2021) even use GPR as regression method, both papers train single cell aging models with cycles from BOL to a certain threshold and test for cycles from the threshold to EOL.

1.5. Proposed Solution

In a typical application, a battery returns from an arbitrary unknown usage. After completing one (partial) charging cycle, SoH and RUL should be estimated under assessable uncertainties. Based on estimation and uncertainty, a decision on the further use can be taken. A proper method for uncertainty estimation in ML regressors are Gaussian processes (GPs) (Rasmussen & Williams, 2008). There are publications like (Li, Yuan, Li, & Wang, 2020), also using ICA and GPR for SoH estimation. However, they perform single cell estimation using the first 55% of the cycles as training data to predict further development. In our study, we focus on a scenario where the regression maps one charging cycle of a previously unseen battery cell to the SoH- and RUL estimation, by knowing cycling data from BOL to EOL of multiple similar cells. The main contributions of this paper are

- Derivation of a GPR_n regressor, combining single training cell GPs,
- Comparison of GPR_n performance to other standard ML regressors, regarding SoH- and RUL estimation,
- Derivation of a RUL monitoring strategy based on GPR_n uncertainty estimation,
- Hyperparameter tuning and quantitative evaluation of the RUL monitoring strategy.

The proposed RUL monitoring strategy manages controlled aging until EOL by only taking measurements on a few charging cycles during battery lifetime.

1.6. Organization of the Paper

The paper is organized as follows: Section 2 explains the ICA process, coming from raw data to ICA features. Further, it describes the theory of our GPR_n approach, serving as SoH- and RUL estimator and explains the RUL monitoring strategy. In section 3 the GPR_n based SoH- and RUL estimators are quantitively compared to other ML estimators. Also the result of the RUL monitoring strategy is presented. Key conclusions and an outlook are drawn in section 4.

2. METHODS

2.1. Oxford Battery Dataset

In our study, the Oxford Battery Degradation Dataset 1 is used (Birkl, 2017). It comprises aging experiments conducted on eight Kokam SLPB533459H4 lithium-polymer pouch cells with a nominal capacity of 740 mAh and rated voltage of 3.6 V. Each cell employs a lithium cobalt oxide positive electrode and a graphite negative electrode. The tests were performed in a temperature chamber maintained at 40 °C, using a constant current (CC) and constant voltage (CV) charging protocol with upper and lower cut-off voltages of 4.2 V and 2.6 V, respectively. The charging rate in CC is 1C. Discharging is done with defined driving profiles. Voltage, charge, and cell temperature were recorded during diagnostic cycles conducted every 100 cycles at a sampling rate of 1 Hz. The practical EOL of batteries in mobile applications often is defined at 80% SoH. This is reached by the eight cells in between 3300 and 5300 cycles (Birkl, 2016).

2.2. Incremental Capacity Analysis

With ICA, battery aging mechanisms are analyzed using their peaks and valleys. The IC is defined as shown in equation 1, where Q represents the charge and V the voltage across the battery terminals.

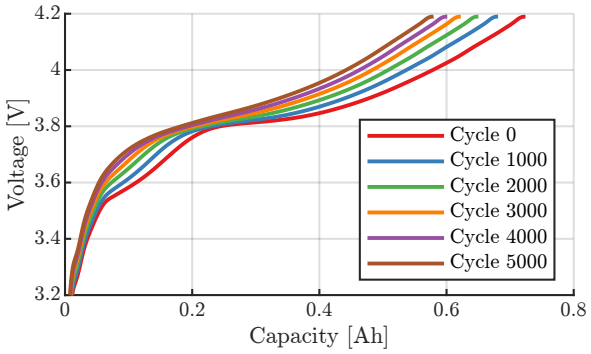
$$\mathrm{IC} = \frac{\mathrm{d}Q}{\mathrm{d}V} = \frac{I \cdot \mathrm{d}t}{\mathrm{d}V} = I \cdot \frac{\mathrm{d}t}{\mathrm{d}V} \tag{1}$$

So IC in words means, how much charge is stored into the battery per voltage change. The ICA method converts the voltage plateaus in the voltage curve into clearly identifiable peaks on the IC curve (Ansean et al., 2019). To illustrate the advantage of transforming the measurement data into an IC curve, figure 1 compares the standard charging curve with the IC curve of a single cell's life. The comparison clearly shows better interpretability of cell aging in the IC curve. Even if only partial charging segments are available (e.g. $3.5\ \mathrm{V} < V_{\mathrm{meas}} < 3.9\ \mathrm{V}$), relevant information for aging prognosis is still included. For the standard charging curve, the most visible indicator is the capacity at end of charge, which in fact is a direct representative of the SoH, due to the definition (Yao et al., 2021)

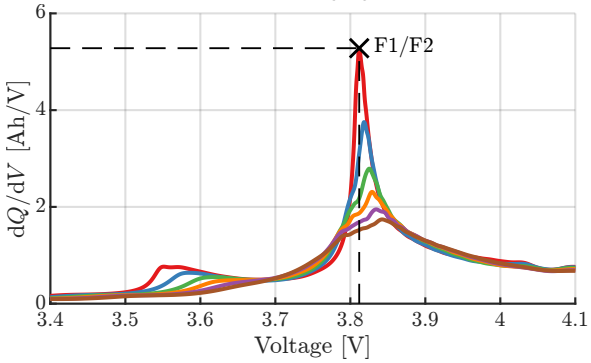
$$\mathrm{SoH} = \frac{Q_m}{Q_r} \cdot 100\%, \tag{2}$$

where Q_m is the current maximum available capacity of the battery, measured under rated conditions, and Q_r is the rated capacity. So the standard charging curve only provides highly interpretable information about cell aging when the battery is fully charged from $\mathrm{SoC} = 0$ to $\mathrm{SoC} = Q_m$.

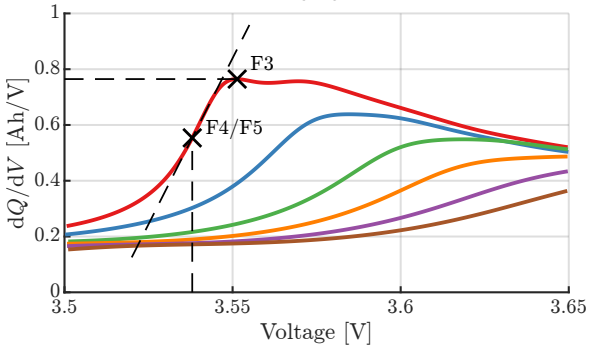
Normally, only the CC segment of the charging cycle is considered, which implicitly defines I in equation (1). As shown in (Wang et al., 2023) the charging rate C influences the distinctiveness of characteristic features of the IC curve significantly. High currents heavily influences cell reactions and distort the IC curve characteristic (Stroe & Schaltz, 2020). When the current I is fixed by the CC charging protocol, only the time derivative of the voltage $\frac{\mathrm{d}V}{\mathrm{d}t}$ has to be calculated from raw measurement data. There is always an inevitable error and noise in battery operation and battery voltage measurement. Also, for a proper sampling frequency, the voltage change in between two time steps is $\Delta V = V_k - V_{k-1} \approx 0$, making the direct numeric derivative $\frac{\mathrm{d}V}{\mathrm{d}t}$ of the measurement data extremely noisy. Filtering methods are needed to smooth IC curves (Chen et al., 2024). (Beatty, Strickland, & Ferreira, 2024) reviews several filter methods, including moving average, gaussian filter, Kalman filter, butterworth low-pass



(a) Standard charging curve



(b) IC charging curve



(c) IC charging curve zoom

Figure 1. Comparison of charging curves

filter and Savitzky–Golay filter. In our case a 4th degree butterworth low-pass with a cut frequency of $f_c = 0.01$ Hz is applied to filter the voltage measurement time series by using MATLAB's *filtfilt()* zero-phase digital filtering. The resulting IC curve is finally smoothed using a moving average filter.

Features describing the aging of the battery cell have to be defined. For smaller datasets like the given one, feature definition is done manually by data observation and analysis. Looking at figure 1 b), the most obvious feature F1 is the maximum peak of the IC curve $\mathrm{IC}_{\mathrm{peak}}$, as it significantly decreases with further aging. Feature F2 is the corresponding voltage $V(\mathrm{IC}_{\mathrm{peak}})$. Zooming into the interval $3.5\ \mathrm{V} < V_{\mathrm{bat}} < 3.65\ \mathrm{V}$ at figure 1 c), three more features are derived: F3 is the maximum occurring IC, named $\mathrm{IC}_{\mathrm{max}}$ for this given interval. F4 is the maximum positive slope of the IC curve in this interval, $\frac{\mathrm{dIC}}{\mathrm{d}V}_{\mathrm{max}}$. F5 is is the corresponding voltage $V(\frac{\mathrm{dIC}}{dV}_{\mathrm{max}})$, where F4 occurs.

The chosen features are evaluated using the Spearman rank correlation coefficient with the single feature as input quantity and the true SoH and RUL as output quantity. RUL is defined as the number of remaining charge–discharge cycles, until the battery's SoH reaches the EOL threshold of 80 %, i.e.,

$$\mathrm{RUL} = N_{\mathrm{EOL}} - N_{\mathrm{current}}, \quad \text{with} \quad \mathrm{SoH}_{\mathrm{EOL}} = 0.8. \tag{3}$$

The Spearman rank correlation coefficient is a non-parametric measure of a monotonic association between two variables ("Spearman Rank Correlation Coefficient", 2008). It can be expressed as

$$\rho_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}, \tag{4}$$

where $d_i = R[X_i] - R[Y_i]$ is the difference in ranks (sorted positions) of the i-th data pair and n is the number of observations. Unlike the Pearson coefficient, Spearman's ρ_s does not assume linearity or normality and is therefore robust for assessing monotonic but non-linear relationships. Shown in table 1, the correlation of a single feature to SoH is always greater than the correlation to RUL. This is due to higher cell-to-cell variances in mapping the features to RUL than to SoH. Figure 2 shows the mapping of the dominant feature $\mathrm{IC}_{\mathrm{peak}}$ to the true SoH and RUL respectively.

Table 1. Spearman rank correlation coefficients of IC features

	F1	**F2**	**F3**	**F4**	**F5**
SoH	0.9974	-0.9817	0.9873	0.9830	-0.9908
RUL	0.9702	-0.9488	0.9814	0.9811	-0.9715

2.3. Gaussian Process Regression

GPR is used in this work to probabilistically model the non-linear relation between ICA-based features and the aging quantities SoH and RUL. A Gaussian Process defines a distribution over functions

$$f(\mathbf{x}) \sim \mathcal{GP}\big(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')\big), \tag{5}$$

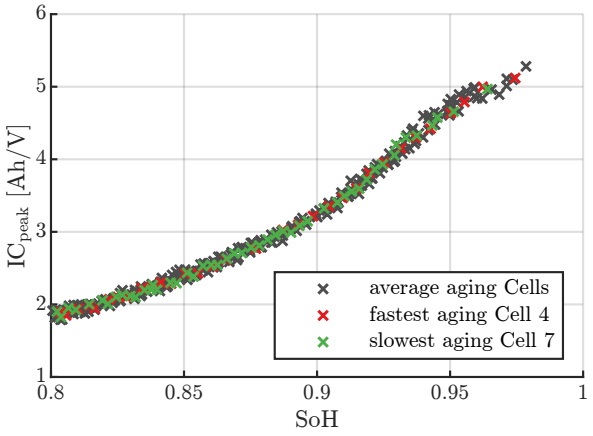
which is fully specified by the mean function $m(\mathbf{x})$ and the covariance function $k(\mathbf{x}, \mathbf{x}')$ (Rasmussen & Williams, 2008).

For the ICA feature vectors $\mathbf{x} \in \mathbb{R}^d$, a radial basis function (RBF), squared exponential kernel for d-dimensional feature is employed.

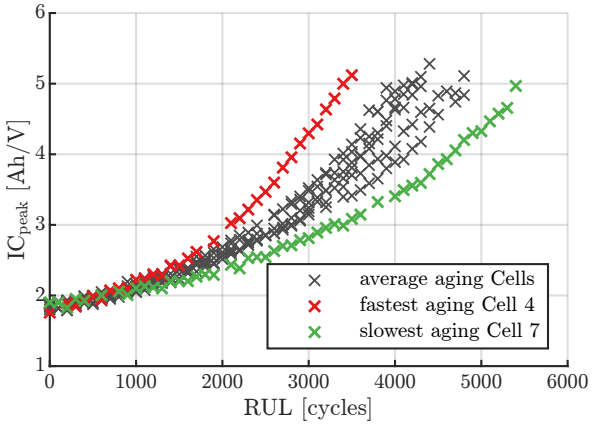
$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{1}{2}\sum_{j=1}^{d}\frac{\left(\mathbf{x}_j - \mathbf{x}'_j\right)^2}{\ell_j^2}\right). \tag{6}$$

RBF kernels are suitable for smooth underlying functions, such as those encountered when mapping ICA features to SoH and RUL. σ_f^2 denotes the signal variance and ℓ_j are the characteristic length scales of the individual ICA features (Li, Yuan, Li, & Wang, 2020). By setting individual length scales for every feature, automatic relevance determination (ARD) defines which features dominate the regression performance.

The observed target quantities of all n training samples, i.e.



(a) SoH cell variance



(b) RUL cell variance

Figure 2. Cell variance of feature F1 regarding SoH and RUL

SoH or RUL, are denoted as y_i. They are modelled as noisy function evaluations from the hypothetical underlying latent function $f(\cdot)$

$$y_i = f(\mathbf{x}_i) + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma_n^2), \quad i = 1, ..., n. \tag{7}$$

The prior distribution can be denoted as

$$\mathbf{y} \sim \mathcal{N}\big(m(\mathbf{x}),\, \mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}\big), \tag{8}$$

with the n-dimensional symmetric positive definite kernel matrix $\mathbf{K}$ elements $\mathbf{K}_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$. For a new ICA feature vector $\mathbf{x}_*$, the joint prior distribution of $\mathbf{y}$ and $f_* = f(\mathbf{x}_*)$ is given by

$$\begin{bmatrix} \mathbf{y} \\ f_* \end{bmatrix} \sim \mathcal{N}\left(m(\mathbf{x}), \begin{bmatrix} \mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I} & \mathbf{k}_* \\ \mathbf{k}_*^\top & k_{**} \end{bmatrix} \right), \tag{9}$$

with $\mathbf{k}_* = [\, k(\mathbf{x}_1, \mathbf{x}_*), \ldots, k(\mathbf{x}_n, \mathbf{x}_*)\,]^\top$ and $k_{**} = k(\mathbf{x}_*, \mathbf{x}_*)$. From that, the posterior distribution of the test output data $\mathbf{f}^*$, knowing training data $\mathbf{x}$, $\mathbf{y}$ and test input features $\mathbf{x}^*$, is

$$p(\mathbf{f}^* \mid \mathbf{x}, \mathbf{y}, \mathbf{x}^*) = \mathcal{N}\big(\bar{\mathbf{f}}^*, cov(\mathbf{f}^*)\big), \tag{10}$$

where the prediction mean $\bar{\mathbf{f}}^*$ and the prediction covariance $cov(\mathbf{f}^*)$ are given by

$$\bar{\mathbf{f}}^* = \mathbf{K}(\mathbf{x}, \mathbf{x}^*)^{\mathrm{T}} \left[\mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}\right]^{-1} \mathbf{y} \tag{11}$$

and

$$\begin{aligned} cov(\mathbf{f}^*) &= \mathbf{K}(\mathbf{x}^*, \mathbf{x}^*) \\ &\quad - \mathbf{K}(\mathbf{x}, \mathbf{x}^*)^{\mathrm{T}} \left[\mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}\right]^{-1} \mathbf{K}(\mathbf{x}, \mathbf{x}^*). \end{aligned} \tag{12}$$

The hyperparameter set $\boldsymbol{\theta} = \big\{\sigma_f^2,\, \ell_1, \ldots, \ell_d,\, \sigma_n^2\big\}$ is obtained by maximizing the logarithmic marginal likelihood (LML)

$$\begin{aligned} \log p(\mathbf{y} \mid \mathbf{x}, \boldsymbol{\theta}) = &-\frac{1}{2}\, \mathbf{y}^\top \big(\mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}\big)^{-1} \mathbf{y} \\ &- \frac{1}{2}\, \log\big|\mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}\big| - \frac{n}{2} \log(2\pi), \end{aligned} \tag{13}$$

using gradient-based optimization with Pythons *Adam()* optimization environment.

2.4. GPR Ensemble Approach $\mathrm{GPR_n}$

A single GPR model trained on pooled data from all cells cannot capture the pronounced cell-to-cell variability observed in the Oxford dataset, in particular for the mapping from ICA features to RUL. To address this, an ensemble of cell-specific GPRs, denoted as $\mathrm{GPR_n}$, has been applied in this work. Ensemble learning methods based on GPR can be found in literature for different applications (Bayati, Nguyen, & Cheriet, 2020), (Xue, Zhang, Zhang, & Li, 2022). In our apllication, for each of the T training cells, a separate GP expert is fitted, using only the ICA feature vectors and target labels (SoH or RUL) of that particular cell. Assume, that for a new feature vector $\mathbf{x}^*$, each of the T GP experts provides a univariate Gaussian predictive distribution

$$p_i(y \mid \mathbf{x}^*) = \mathcal{N}\big(m_i, s_i^2\big), \qquad i = 1, \ldots, T, \tag{14}$$

with mean m_i and variance s_i^2. The $\mathrm{GPR_n}$ ensemble combines these experts as a finite mixture of Gaussians with non-negative weights π_i that sum to one, $\sum_{i=1}^{T} \pi_i = 1$,

$$p(y \mid \mathbf{x}^*) = \sum_{i=1}^{T} \pi_i\, \mathcal{N}\big(m_i, s_i^2\big). \tag{15}$$

If no further information is given, all training cells are weighted equally, resulting in $\pi_i = \frac{1}{T}$. The predicted mean of the mixture is the expectation of y under $p(y \mid \mathbf{x}^*)$,

$$\mu = \mathbb{E}[y] = \sum_{i=1}^{T} \pi_i\, \mathbb{E}_i[y] = \sum_{i=1}^{T} \pi_i m_i. \tag{16}$$

The predicted variance is obtained from variance definition equation

$$\mathrm{Var}(y) = \mathbb{E}[y^2] - \big(\mathbb{E}[y]\big)^2. \tag{17}$$

After some simple manipulations the variance results in

$$\begin{aligned} \mathrm{Var}(y) &= \mathbb{E}[y^2] - \mu^2 \\ &= \sum_{i=1}^{T} \pi_i \big(s_i^2 + m_i^2\big) - \left(\sum_{i=1}^{T} \pi_i m_i\right)^2 \\ &= \underbrace{\sum_{i=1}^{T} \pi_i s_i^2}_{\text{epistemic uncertainty}} + \underbrace{\left(\sum_{i=1}^{T} \pi_i m_i^2 - \left(\sum_{i=1}^{T} \pi_i m_i\right)^2\right)}_{\text{aleatoric uncertainty}}. \end{aligned} \tag{18}$$

The first term corresponds to the average predictive variance of the individual experts and captures the epistemic uncertainty of the data. It is caused by a lack of knowledge, because information about the true SoH and RUL are only available at finite samples. The second term is the variance of the single GP expert means and measures how strongly the experts disagree. This is interpreted as aleatoric uncertainty arising from inherent random cell-to-cell variance.

To prevent the single GPs from overfitting, the number of training epochs n of the GPR is restricted. If the number of epochs is too large, single GPs overfit, resulting in a higher mean absolute error (MAE) when evaluating the $\mathrm{GPR_n}$ approach with training data. Therefore n is set, where the MAE over n is at its minimum.

2.5. RUL Monitoring Strategy

In real-world applications, including second-life applications, continuous cell monitoring is not always guaranteed, respec-

tively expensive. Our proposed RUL monitoring strategy enables informed operational decisions based on sparse measurements, typically three to five diagnostic cycles over the entire battery lifetime. The procedure leverages the GPR_n based variance including RUL estimation and the SVR based SoH estimation to balance safe operation with maximal utilization. The monitoring workflow is illustrated in figure 3

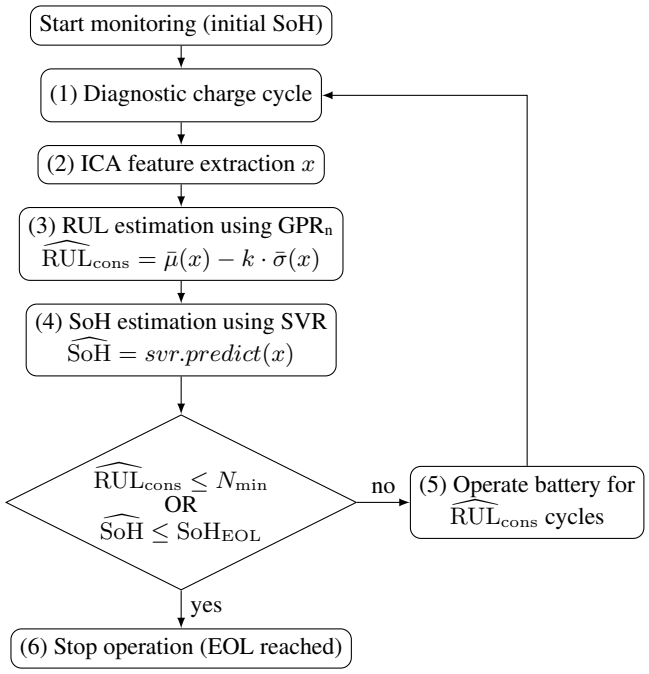


Figure 3. Program flow chart of the iterative RUL monitoring strategy using ICA features and the GPR_n ensemble.

and operates iteratively: (1) perform a single diagnostic charging cycle at arbitrary points in the battery's lifetime, (2) extract ICA features and feed them into the trained GPR_n ensemble, (3) compute a conservative RUL estimate by subtracting a margin proportional to the uncertainty, (4) estimate the current SoH using a SVR model, (5) operate exactly $\widehat{\text{RUL}}_{\text{cons}}$ cycles before the next measurement, (6) repeat until the battery approaches EOL. EOL is defined by $\widehat{\text{RUL}}_{\text{cons}} \leq N_{\min}$, or if the estimated SoH is already smaller than SoH_{EOL}. With $\text{N}_{\min}$, it is possible to define application-specific values for the RUL, at which a further diagnostic cycle is worthwhile. In this case, $N_{\min} = 40$ was selected. This adaptive scheme allows operators to tolerate uncertainty quantification and trigger additional diagnostics only, when uncertainty becomes prohibitively large.

The following key performance indicators (KPIs) assess the effectiveness of the monitoring strategy across all cells:

- **Utilization Grade**
  $U = \frac{N_{\text{cycles,strategy}}}{N_{\text{cycles,possible}}}$, measures the percentage of available cycles actually used before triggering EOL. Higher values indicate effective exploitation of remaining capacity.
- **Number of Monitoring Steps**
  M, counts how many diagnostic cycles were required to reach EOL. Lower M reduces measurement burden.
- **Overcycling Rate**
  P_{over}, fraction of cells that fall below the $\text{SoH} = 0.80$ threshold. Ideally $P_{\text{over}} = 0$ to avoid cell damage.
- **End-of-Life Cycle Deviation**
  $\Delta N_{\text{EOL}} = N_{\text{true,EOL}} - N_{\text{strategy,EOL}}$, absolute error in cycle count at EOL. Small deviations demonstrate predictive accuracy.
- **End-of-Life SoH Deviation**
  $\Delta\text{SoH}_{\text{EOL}} = \text{SoH}_{\text{strategy,EOL}} - \text{SoH}_{\text{true,EOL}}$, residual SoH error at strategy termination. Ideally positive and close to zero.

The interplay between these KPIs can be tuned by setting hyperparameter k and reveals the trade-off between aggressive utilization and safe operation. A highly conservative strategy (large k) yields high safety but low utilization, conversely, an aggressive strategy risks overcycling. Therefore, optimal tuning is essential for practical deployment.

3. RESULTS AND DISCUSSION

3.1. SoH and RUL Regression Setup

Different regression approaches are benchmarked for SoH and RUL prediction using the same ICA feature set and a common cross-cell evaluation protocol. For each experiment, six cells are used for training and two cells for testing, and all $\binom{8}{2}$ train-test combinations are evaluated. Performance is reported as MAE averaged over all train-test combinations and absolute maximum error occurring.

The following commonly used regression models are benchmarked against the shown GPR approach from literature and our novel GPR_n approach:

- **Poly1D**
  - One-dimensional polynomial regression using only the dominant feature IC_{peak} as input. Hyperparameter: Polynomial degree = 3.
- **PolyMulti**
  - Multivariate polynomial regression over the full ICA feature vector. Hyperparameter: Polynomial degree = 3, feature interaction terms degree = 2.
- **FFNN**
  - Hyperparameter: Two hidden layers with 64 neurons each; ReLU activations, trained with *Adam()*, 1000 epochs, learning rate 10^{-3}.
- **SVR**
  - SVR with RBF kernel. Regularisation parameter C, tube width ε and feature-wise kernel length scales tuned by Bayesian Optimization and manually fine tuning based on T-fold cross-validation loss.

- **$\mathrm{GPR_{LOCO}}$**
  - Similar model and hyperparameter as GPR, hyperparameter optimization by minimizing leaving one cell out (LOCO) validation loss.

3.2. SoH Regression Results

For SoH estimation, the dataset extends ICA features from $\mathrm{SoH} = 100\%$ to $\mathrm{SoH} = 80\%$ for all eight cells. The SoH results in table 2 show that all investigated regression models achieve comparatively small average errors, but differ substantially in robustness and worst-case behaviour. The Poly-

Table 2. Performance metrics of SoH estimation

	MAE train in %	**MAE test** in %	**Max Error test** in %
Poly1D	0.723	0.764	3.02
PolyMulti	0.159	0.341	4.03
FFNN	0.220	0.305	1.66
SVR	0.184	0.261	1.01
GPR	0.214	0.271	1.20
$\mathrm{GPR_{LOCO}}$	0.202	0.317	5.64
$\mathrm{GPR_n}$	0.412	0.463	1.93

nomial fitting approaches provide a reasonable reference MAE, but still exhibit relatively large maximum deviations, making them mainly a simple baseline. The commonly used ML methods FFNN, SVR and GPR outperform polynomial fitting in lower MAE and lower maximum error. SVR is evaluated as best performing method. Obviously, $\mathrm{GPR_n}$ performs worse than GPR. This can be explained by only minor cell-to-cell variance when ICA features are mapped to SoH, as shown in figure 2. The single pooled GPR therefore uses approximately six times more data points to describe the SoH–feature relationship, reducing epistemic uncertainties. Figure 4 il-

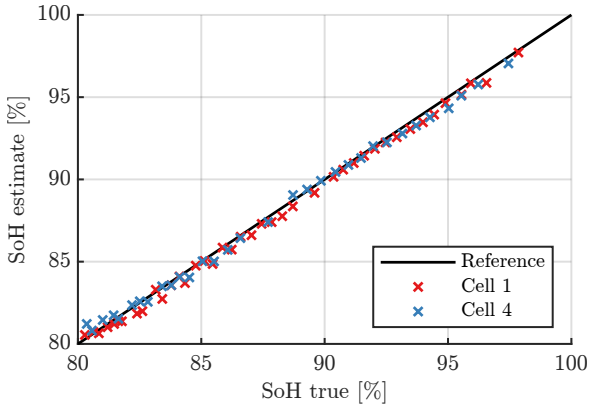


Figure 4. SoH regression example using SVR

lustrates the SoH regression using the best performing SVR approach for two test cells (Cell 1 and Cell 4) of the Oxford dataset. The regression results are highly satisfactory for both cells.

3.3. RUL Regression Results

For RUL estimation, the dataset extends ICA features from $N_{\mathrm{cycles}} = 0$ to $N_{\mathrm{cycles}} = N_{\mathrm{EOL}} + 400$ for all eight cells. In contrast to the SoH estimation, the RUL results in table

Table 3. Performance metrics of RUL estimation

	MAE train in cycles	**MAE test** in cycles	**Max Error test** in cycles
Poly1D	335	375	1468
PolyMulti	111	299	2770
FFNN	226	247	1729
SVR	236	271	1390
GPR	36	323	2748
$\mathrm{GPR_{LOCO}}$	133	249	1734
$\mathrm{GPR_n}$	201	232	1277

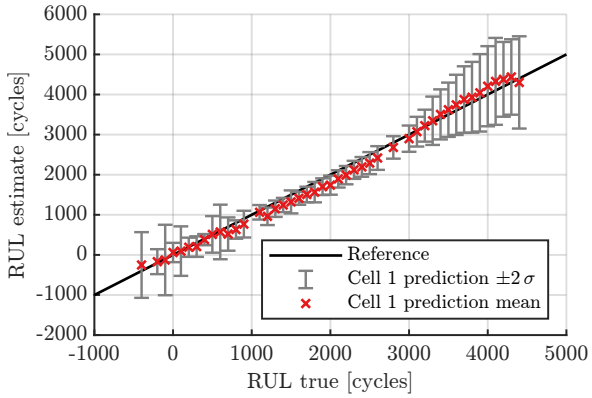
3 highlight that direct remaining-life prediction is strongly affected by cell-to-cell variability. Cell-to-cell variability is averaged best by $\mathrm{GPR_n}$, also FFNN and SVR perform satisfying. Here, the pooled GPR completely overfits, which can be seen at the extreme low MAE train compared to the MAE test. Already $\mathrm{GPR_{LOCO}}$ reduces the overfitting by the cross validation mechanism, but still there are single outliers causing a high maximum error in the test dataset. The $\mathrm{GPR_n}$ ensemble, which explicitly models the cell-to-cell variability, gets rid of overfitting. This leads to noticeably lower mean and maximum RUL errors across all train-test splits. Additionally, the ensemble structure of $\mathrm{GPR_n}$ provides the uncertainty components, which are later exploited by the RUL monitoring strategy. Figure 5 shows the RUL regression using the best performing $\mathrm{GPR_n}$ approach for the same test cells as in the SoH regression evaluation shown in figure 4. The regression results are highly cell depended. For Cell 1, the RUL estimation is performing much better than for Cell 4. Looking at figure 2 b), one becomes aware, that Cell 1 has approximately average aging behaviour, whereas Cell 4 is aging fastest. Due to the averaging effect of $\mathrm{GPR_n}$, the RUL estimation performance improves, as more the test cell behaves like the average cell in the dataset.

3.4. Comparsion of SoH and RUL Regression

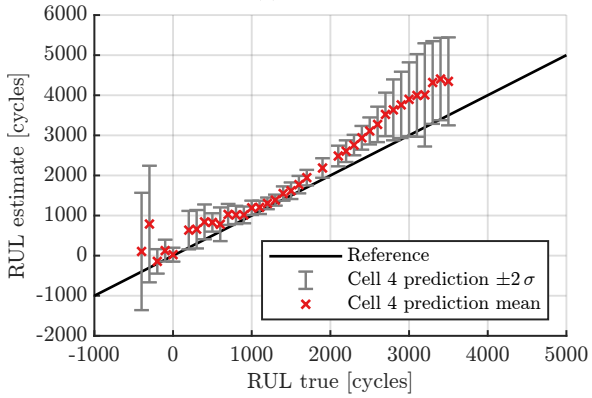
To compare SoH estimation with RUL estimation a new measure has to be introduced, the normalized mean absolute error (NMAE). It is defined by the mean absolute error divided by the range of the true values (Goldberg, Roeder, Gupta, & Perkins, 2001). The resulting metric becomes unitless, like for SoH estimation

$$\mathrm{NMAE_{SoH}} = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|\widehat{\mathrm{SoH}}_i - \mathrm{SoH}_i\right|}{\mathrm{SoH_{max}} - \mathrm{SoH_{min}}}. \tag{19}$$

Using NMAE allows a direct numerical comparison of SoH and RUL regression performance, even though SoH is ex-



(a) Cell 1



(b) Cell 4

Figure 5. RUL regression example using GPR_n

pressed in % and RUL in absolute cycle counts. The cell-to-cell variances makes mapping of ICA features to RUL more challenging, resulting in a NMAE of 5.33%, which is around four times higher than for SoH (1.31%).

3.5. RUL Monitoring Strategy

Based on the results of RUL estimation with GPR_n regression, the derived monitoring strategy is evaluated. It is assumed that aging data are available for all cells in the dataset, except for the single test cell. The monitoring strategy basically inherits two tunable parameters. First, the number of epochs n of the single GPs training. It is defined finding a minimum in the training data MAE over all train-test combinations. Figure 6 shows, that the minimum is located at $n \approx 20$.

The second tunable parameter is k, trading-off between aggressive utilization and safe operation. All in section 2.5 defined KPIs are effected from the selection of k, illustrated in figure 7. In these five plots, mean results of the KPIs over all train-test combinations are presented. Using low k, the utilization grade is set aggressive, resulting in $U > 1$, meaning in more than half of the examples the battery gets over-cycled ($P_\text{over} > 0.5$). Therefore ΔN_EOL also points out

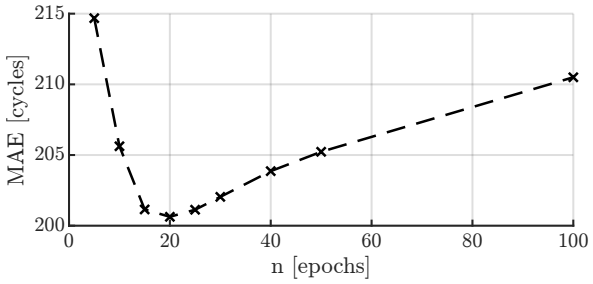


Figure 6. Tuning number of epochs based on MAE trraining loss

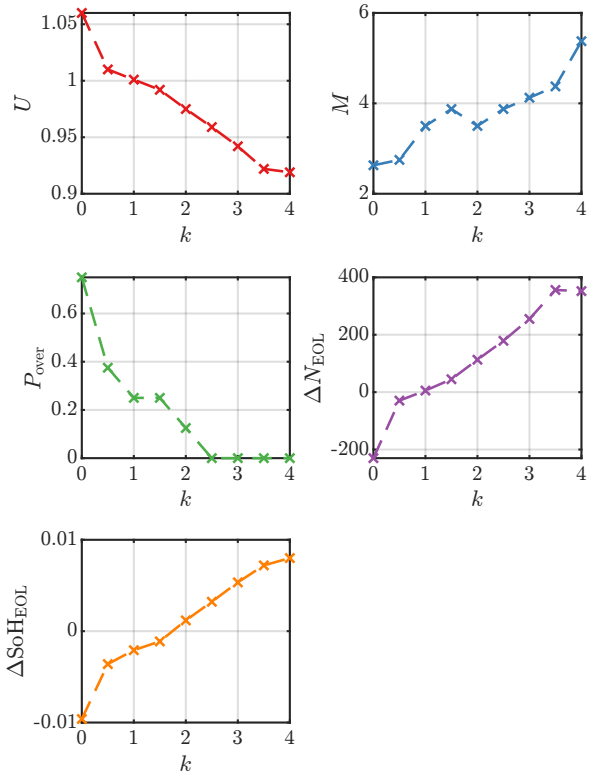


Figure 7. Tuning hyperparameter k based on KPIs

negative values. However, the number of monitoring steps $M \approx 3$ is rather small. With increasing k, the overcycling rate P_over consequently reduces due to focus on safer operation. M therefore rises. A smooth trade-off is found for $k = 2$. Here ΔN_EOL has a small, positive value, the overcycling rate $P_\text{over} = \frac{1}{8}$ is already sufficiently small and $M = 3.5$ still compromises a low diagnostic effort. Looking at figure 7, the overcycling in terms of SoH is limited for all chosen k to $\Delta\text{SoH}_\text{EOL} \leq \pm 1\%$. Even the maximum overcycling of a single train-test combination for $k = 0$ is $\text{SoH}_\text{EOL} = -3.6\%$, meaning no battery is aged to values $\text{SoH} \leq 76.4\%$. Therefore, all selections of k can be described as safe operation modes, which leaves the choice up to the user. For further presented examples $k = 2$ is used. Observing the RUL monitoring strategy from a cell behaving in an average manner in

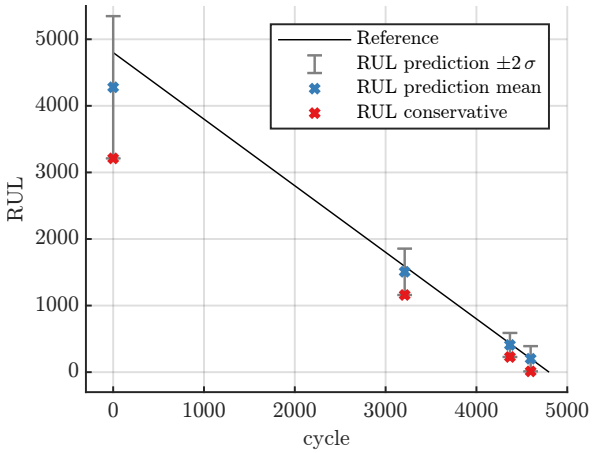


Figure 8. RUL monitoring strategy Cell 3

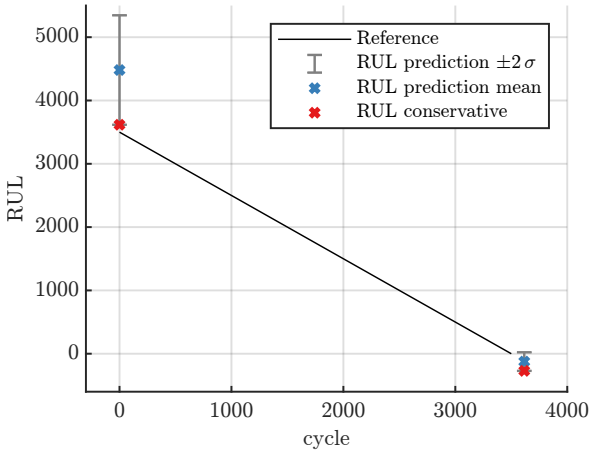


Figure 9. RUL monitoring strategy Cell 4

figure 8, one notes, that the RUL predictions nearly match the true value. Through the conservative selection of $k = 2$, four monitoring steps are required. The operation cycles between two diagnostic charge cycles decreases monotonically with cell age. There is no overcycling, when EOL of the monitoring strategy is reached. With $\Delta N_{\mathrm{EOL}} = 203$, already 95.8% of the possible cycles were executed. If one looks at the fastest aging Cell 4 in figure 9, the initial RUL prediction noticeably deviate from the true values. Nevertheless, due to the conservative strategy, there is only a small overcycling of $\Delta N_{\mathrm{EOL}} = -116$ cycles. Taking a closer look at the slowest aging Cell 7 in figure 10, consequently the initial RUL prediction underestimates the battery life. From the second diagnostic cycle onwards, the RUL is estimated properly. Here $\Delta N_{\mathrm{EOL}} = 17$ makes almost perfect use of the actual battery life, 99.7% of the possible cycles were executed.

For all eight cells the KPIs of the RUL monitoring strategy are summarized in table 4. To achieve, on average, the same utilization rate U and safety against overcycling in terms of a comparable $\Delta\mathrm{SoH}_{\mathrm{EOL}}$ using a fixed-interval inspection strategy, the interval would have to be set to 200, resulting in $M \approx 20$ monitoring steps. This clearly demonstrates the diagnostic benefits of the monitoring strategy described.

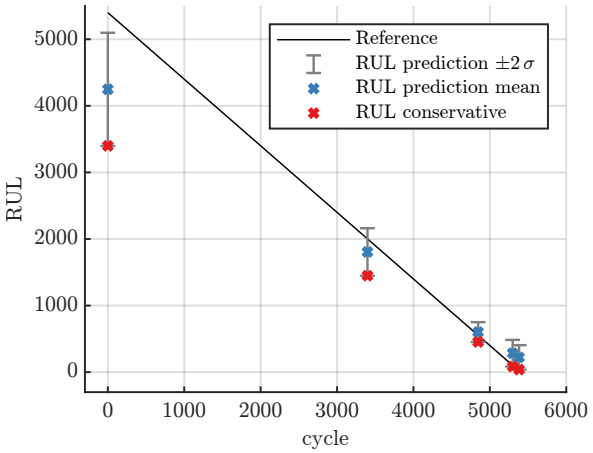


Figure 10. RUL monitoring strategy Cell 7

Table 4. RUL monitoring strategy results for $k = 2$

Cell	U	M	P_{over}	ΔN_{EOL}	$\Delta\mathrm{SoH}_{\mathrm{EOL}}$
1	0.9195	3	0	354	0.0104
2	0.9717	4	0	119	0.0025
3	0.9577	4	0	203	0.0020
4	1.0331	2	1	-116	-0.0074
5	0.9370	3	0	271	0.0061
6	0.9875	3	0	50	0.0008
7	0.9969	5	0	17	0.0003
8	0.9975	4	0	12	-0.0052

4. CONCLUSION AND FUTURE WORK

The presented work demonstrates that ICA-based features, combined with data-efficient regression methods, enable accurate and uncertainty-aware estimation of both SoH and RUL for previously unseen lithium-ion cells using only single diagnostic charge cycles and aging data from a small set of similar reference cells. For SoH estimation, the mapping from ICA features to SoH is found to be almost cell invariant, which leads to low NMAE values around 1.3% and allows even comparatively simple regression models such as SVR and pooled GPR to achieve highly satisfying accuracy with small maximum errors. In contrast, RUL estimation exhibits pronounced cell-to-cell variance, resulting in a higher NMAE of about 5.3% and clearly revealing the need for models that explicitly account for heterogeneous degradation trajectories. Here, the $\mathrm{GPR}_{\mathrm{n}}$ ensemble successfully averages over individual cell models, reduces overfitting compared to pooled GPR, and provides a meaningful decomposition into epistemic and aleatoric uncertainty, with the latter dominating due to intrinsic cell-to-cell variability.

Building on these uncertainty-aware RUL estimates, the proposed monitoring strategy shows that a battery can be safely operated from BOL to EOL by performing only a few (typically three to five) diagnostic cycles, while exploiting more than 90% of the available cycle budget and keeping SoH deviations at strategy EOL well below 1% for most cells. By tuning the strategy parameter k, operators can trade off utilization grade, number of monitoring steps, and overcycling risk, which makes the approach directly applicable to scenarios with sparse measurements, such as second-life deployment after unknown first-life usage.

Future work could focus on the ICA feature extraction. Studies on feature ablation could help determine the dominance and significance of certain features that correlate with specific physical properties of the cell. Automatic feature extraction is of great interest as well. Further hidden information within the ICA curve could be identified. Both approaches could then be compared in terms of the accuracy of the resulting state estimation or in terms of noise sensitivity in the input data disturbances.

Next investigations could also extend the current RUL monitoring strategy, where predictions only rely on a single-cycle measurement. By incorporating information from multiple already taken diagnostic cycles of the same test cell instead of ignoring them, for example by storing past SoH estimates, degradation trends can be constructed. If there is a bigger dataset available, the degradation trend of the currently monitored cell can be compared to degradation trends of all training cells. Using this information one can adapt the training cell weights π_i in Eq. (15) within the GPR_n ensemble. This concept could reduce the dominance of aleatoric uncertainty in the RUL estimation. Success depends heavily on the size of the data set. Eight cells are far from being sufficient to describe cell variant aging trends in a statistically reliable manner.

Generally, also the progression of the ICA features on aging cells under varying operating conditions should be investigated. Incorporating different charging rates, discharging profiles and ambient temperatures generalizes the presented method for more practical applications. Larger datasets and greater variability in operating conditions are the key factors enabling this academic research to be deployed in real-world scenarios.

REFERENCES

Al-Hashimi, Z., Khamis, T., Al Kouzbary, M., Arifin, N., Mokayed, H., & Abu Osman, N. A. (2025, March). A decade of machine learning in lithium-ion battery state estimation: a systematic review. *Ionics*, *31*(3), 2351–2377.

Ansean, D., Garcia, V. M., Gonzalez, M., Blanco-Viejo, C., Viera, J. C., Pulido, Y. F., & Sanchez, L. (2019, May). Lithium-Ion Battery Degradation Indicators Via Incremental Capacity Analysis. *IEEE Transactions on Industry Applications*, *55*(3), 2992–3002.

Bayati, A., Nguyen, K.-K., & Cheriet, M. (2020). Gaussian process regression ensemble model for network traffic prediction. *IEEE Access*, *8*, 176540-176554.

Beatty, M., Strickland, D., & Ferreira, P. (2024, August). A Review of Methods of Generating Incremental Capacity–Differential Voltage Curves for Battery Health Determination. *Energies*, *17*(17), 4309.

Birkl, C. (2016). *Diagnosis and Prognosis of Degradation in Lithium-Ion Batteries* (Unpublished doctoral dissertation). University of Oxford, Oxford.

Birkl, C. (2017). *Oxford Battery Degradation Dataset 1*. University of Oxford.

Chen, H., Yue, W., Bin, G., Jiang, Q., Shao, W., & She, C. (2024, November). Filter methods comparation for incremental capacity analysis in lithium-ion batteries health prediction. *Journal of Energy Storage*, *101*, 113878.

Dar, T. H., & Singh, S. (2025, April). A comprehensive review, perspectives and future directions of battery characterization and parameter estimation. *Journal of Applied Electrochemistry*, *55*(4), 837–863.

Demirci, O., Taskin, S., Schaltz, E., & Acar Demirci, B. (2024, August). Review of battery state estimation methods for electric vehicles-Part II: SOH estimation. *Journal of Energy Storage*, *96*, 112703.

Elmahallawy, M., Elfouly, T., Alouani, A., & Massoud, A. M. (2022). A Comprehensive Review of Lithium-Ion Batteries Modeling, and State of Health and Remaining Useful Lifetime Prediction. *IEEE Access*, *10*, 119040–119070.

Goldberg, K., Roeder, T., Gupta, D., & Perkins, C. (2001, July). Eigentaste: A Constant Time Collaborative Filtering Algorithm. *Information Retrieval*, *4*(2), 133–151.

Guo, R., & Shen, W. (2021, December). A Review of Equivalent Circuit Model Based Online State of Power Estimation for Lithium-Ion Batteries in Electric Vehicles. *Vehicles*, *4*(1), 1–29.

Hamar, J. C. (2024). *Data-driven battery state estimation* (Doctoral dissertation, Technische Universität München TUM School of Engineering and Design, München). PhD thesis, Technical University of Munich.

Jha, A., Dorkar, O., Biswas, A., & Emadi, A. (2024, June). iTransformer Network Based Approach for Accurate Remaining Useful Life Prediction in Lithium-Ion Batteries. In *2024 IEEE Transportation Electrification Conference and Expo (ITEC)* (pp. 1–8). Chicago, IL, USA: IEEE.

Jossen, A. (2006, March). Fundamentals of battery dynamics. *Journal of Power Sources*, *154*(2), 530–538.

Li, X., Yuan, C., Li, X., & Wang, Z. (2020, January). State of health estimation for Li-Ion battery using incremental capacity analysis and Gaussian process regression. *Energy*, *190*, 116467.

Li, X., Yuan, C., & Wang, Z. (2020, July). State of health estimation for Li-ion battery via partial incremental capacity analysis based on support vector regression. *Energy*, *203*, 117852.

Nguyen Van, C., & Quang, D. T. (2023, June). Estimation of SoH and internal resistances of Lithium ion battery based on LSTM network. *International Journal of Electrochemical Science*, *18*(6), 100166.

Pang, X., Liu, X., Jia, J., Wen, J., Shi, Y., Zeng, J., & Zhao, Z. (2021, December). A lithium-ion battery remaining useful life prediction method based on the incremental capacity analysis and Gaussian process regression. *Microelectronics Reliability*, *127*, 114405.

Rasmussen, C. E., & Williams, C. K. I. (2008). *Gaussian processes for machine learning* (3. print ed.). Cambridge, Mass.: MIT Press.

Santhanagopalan, S., Guo, Q., Ramadass, P., & White, R. E. (2006, June). Review of models for predicting the cycling performance of lithium ion batteries. *Journal of Power Sources*, *156*(2), 620–628.

Schmitt, J. (2022). *State of health estimation for lithium-ion batteries* (Doctoral dissertation, TUM School of Engineering and Design der Technischen Universität München, München). PhD thesis, Technical University of Munich.

She, C., Zhang, L., Wang, Z., Sun, F., Liu, P., & Song, C. (2023, February). Battery State-of-Health Estimation Based on Incremental Capacity Analysis Method: Synthesizing From Cell-Level Test to Real-World Application. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, *11*(1), 214–223.

Spearman rank correlation coefficient. (2008). In *The concise encyclopedia of statistics* (pp. 502–505). New York, NY: Springer New York.

Stroe, D.-I., & Schaltz, E. (2020, January). Lithium-Ion Battery State-of-Health Estimation Using the Incremental Capacity Analysis Technique. *IEEE Transactions on Industry Applications*, *56*(1), 678–685.

Wang, G., Cui, N., Li, C., Cui, Z., & Yuan, H. (2023, December). A state-of-health estimation method based on incremental capacity analysis for Li-ion battery considering charging/discharging rate. *Journal of Energy Storage*, *73*, 109010.

Xia, F., Wang, K., & Chen, J. (2023, August). State of health and remaining useful life prediction of lithium-ion batteries based on a disturbance-free incremental capacity and differential voltage analysis method. *Journal of Energy Storage*, *64*, 107161.

Xue, Z., Zhang, Y., Zhang, L., & Li, H. (2022). Ensemble learning embedded with gaussian process regression for soil moisture estimation: A case study of the continental u.s. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-17.

Yao, L., Xu, S., Tang, A., Zhou, F., Hou, J., Xiao, Y., & Fu, Z. (2021, August). A Review of Lithium-Ion Battery State of Health Estimation and Prediction Methods. *World Electric Vehicle Journal*, *12*(3), 113.

Zhou, Q., Vuylsteke, G., Anderson, R. D., & Sun, J. (2025, June). *Battery State of Health Estimation and Incremental Capacity Analysis under General Charging Profiles Using Neural Networks.* (arXiv:2502.19586 [eess])