Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Nested cell-level voltage-window selection for partial-charge state-of-health estimation of lithium-ion batteries

Huy Hoang Le
Powermore Ltd.
Kim-Anh Nguyen
nkanh@dut.udn.vn
University of Da Nang

Research Article

Keywords: lithium-ion battery, state of health, incremental capacity, partial charging, ridge regression, robustness

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11580964/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Nested cell-level voltage-window selection for partial-charge state-of-health estimation of lithium-ion batteries

Anonymous authors

Abstract

Estimating the state of health (SOH) of lithium-ion batteries from a short part of the charging curve is attractive for battery management systems, because vehicles and devices are rarely charged from empty to full. Published errors are often very low, but the voltage window and the data split are usually chosen by the authors, and the sensitivity to measurement conditions is seldom reported. We study a simple pipeline: five features from a voltage window of the constant-current charging phase feed a ridge regression, and the window is selected using only the training cells (nested leave-one-cell-out). On three public datasets of LCO/NMC-type cells (lithium cobalt oxide or nickel–manganese–cobalt oxide cathodes; NASA, 4 cells; CALCE, 6 cells; Oxford, 8 cells), the mean root-mean-square error (RMSE) of SOH is 3.01, 2.28 and 0.72 %, against 5.9, 2.8 and 2.3 % for a linear model that knows only the cycle index. Ridge regression is statistically indistinguishable from random forest, gradient boosting and support-vector regression (paired Wilcoxon $p \approx 0.42$, 18 cells) and is much cheaper. The best window differs between datasets; a window shift of 0.02 V increases the error by a factor of 1.6 to 3.3 on NASA and Oxford; 10 mV of voltage noise doubles the NASA error unless the voltage is filtered; and models do not transfer between datasets. On 119 lithium iron phosphate (LFP) cells (Supplementary Material), ridge does not beat the cycle-index baseline, and tree ensembles are only moderately better.

Keywords: lithium-ion battery, state of health, incremental capacity, partial charging, ridge regression, robustness

Introduction

State of health (SOH), the ratio of the present to the initial capacity, is central to the safe and economic use of lithium-ion batteries; here, SOH is expressed in percent of the initial capacity. Measuring capacity directly requires a full discharge, which is impractical in service, so many studies estimate SOH from signals that the battery management system (BMS) already records; reviews are given in [1–4]. The charging curve is a convenient source, because charging is regular and routinely logged. Machine-learning pipelines that extract features from segments of the charge voltage and current have been demonstrated on large sets of cells [5], and voltage-indexed capacity features have been used to predict cycle life [6].

In practice, charging rarely starts at the same state of charge (SOC) or ends at full charge. Estimators have therefore been built on a *partial* charging curve, that is, a voltage window or a short segment of the constant-current (CC) phase. Existing approaches use incremental-capacity (IC) features of partial CC data [7–12], optimised voltage ranges [13,14], voltage reconstruction [15], or neural networks on short or random segments [16–23]. Reported errors are small: for example a root-mean-square error (RMSE) of 1.6 % for a 50 mV segment [21], a mean error below 5 % for 10 mV segments (for cells cycled like the training cells) [20], and errors as low as 0.5 % for a multiple linear regression on randomly selected partial charges [24].

These results, however, depend on how the estimators are evaluated. Reviews stress that the evaluation design, the robustness under distribution shift and the uncertainty must be reported [25,26]. A benchmark on 32 NASA cells adopts a leakage-averse, cross-battery protocol [27], and a comparison of charging-phase indicators reports a 119 % performance gap between 5-fold and leave-one-battery-out validation [28]. Leakage and selection bias are known pitfalls of machine learning [29–31]. In partial-charge studies the window is often chosen by the authors, for example from the

position of IC peaks, and the sensitivity of the estimate to this choice and to small voltage offsets is rarely quantified.

This paper does not propose a new network. It examines a simple pipeline under one strict rule: the held-out cell is never used to fit the model or to select the window. We ask four research questions (RQ1–RQ4):

1. How should the voltage window be chosen, and how much does the choice matter across cell types?
2. Is a lightweight linear model sufficient compared with nonlinear regressors under a cell-level validation?
3. How robust is the estimate to a shift of the window and to voltage noise, and can this be mitigated?
4. Does a model fitted on one dataset transfer to another, or do labelled cells of other types help?

The main text covers three datasets of LCO/NMC-type cells (lithium cobalt oxide or nickel–manganese–cobalt oxide cathodes) and the Supplementary Material covers one dataset of lithium iron phosphate (LFP) cells. Every estimator is compared with age-only baselines, that is, models that know only the cycle index. Voltage-range optimisation has precedent [13,14]. Our contribution is its evaluation under a nested, cell-level protocol across cell types, with cell-level statistics, a sensitivity analysis for window shift and voltage noise, and an account of the limits of cross-dataset transfer.

Section 2 reviews related work, Section 3 describes the data and methods, Section 4 gives the results, Section 5 discusses them and Section 6 concludes.

Related work

IC analysis and partial-charge features. IC and differential-voltage analysis are established tools for tracking ageing from the charging curve [32–34] and have been combined with Gaussian process regression [35]. For partial charging, support-vector regression on filtered partial IC curves reached errors below 2 % [7]; IC features and voltage entropy from partial CC data with different initial voltages were fused by a neural network [8]; IC analysis at C/3, 1C and 2C (1C is the current that charges the nominal capacity in one hour) and an integrated-voltage method worked on partial data from large-format cells [9]; and a linear relation between IC features and SOH, calibrated on full-range data, was applied to partial data [10]. Filtering-based and model-based IC methods have been compared [36]. Other approaches use voltage reconstruction [15], an incremental-energy curve [37], feature fusion with Gaussian process regression on NASA and Oxford cells [38] and physics-informed models [39].

Choice of the voltage window. Single and multiple voltage ranges were optimised by grid search and a genetic algorithm on three NMC cells [13], and a charging voltage section was optimised for capacity estimation [14]. An experimental study of partial IC analysis on three NMC/Si–graphite cells found that the window width has a significant influence on the results and that partial and full charges give different results [11]. A 0.1 V window was selected dynamically from IC peaks in [12].

Short and random segments. Many neural networks have been applied to short or random segments of the charging curve. They include convolutional neural networks (CNN) [16,17, 21,22], recurrent [18,19, 23] and transformer-based [40] models, and a method validated on 75 cells under varied temperatures and discharge rates [20]. Transfer learning was studied for segments at varied starting positions [41], and a hybrid CNN–attention model for a constrained window under fast charging [42]. Simpler alternatives are a single-feature linear estimator [43] and a relative-SOC preprocessing [44].

Evaluation and gap. Reviews discuss evaluation methods, uncertainty, calibration and robustness under distribution shift [26], and the evidence needed for trustworthy and transferable estimators [25].

A benchmark on 32 NASA cells under blocked GroupKFold splits (cross-validation in which all records of a cell stay in the same fold) found that sequence models were the most accurate, that gradient-boosted ensembles were competitive and much cheaper, and that about one third of the cells had elevated errors linked to noise or regime shifts [27]. A study of indicators from the CC and the constant-voltage (CV) charging phases found the combined set best under leave-one-battery-out validation [28]. The partial-charge literature is thus crowded, and voltage-range optimisation has precedent, so we claim no new feature or architecture. What is less documented is how the window can be selected without test cells, how it varies between cell types, how sensitive the estimator is to a window shift and to noise, and whether simple linear models suffice under a cell-level protocol. These are the points addressed here.

Data and methods

Framework

Figure 1 summarises the framework, with NASA cell B0005 as a worked example. Each charge is reduced to its CC phase (a), from which the incremental capacity is computed (b). A voltage window $[V_1, V_2]$, with start V_1, end V_2 and width $w = V_2 - V_1$, is cut out and five features are extracted from it. The window is chosen automatically. For each candidate position and width, a ridge model is scored by leaving out each training cell in turn (c). The best-scoring window is kept, provided that at least 90 % of the training records cover it (d). A ridge model is then fitted on all training cells at this window and applied once to the held-out cell (e). The estimate is compared with reference models, tested at the level of cells and stressed with window shifts, voltage noise and dataset changes.

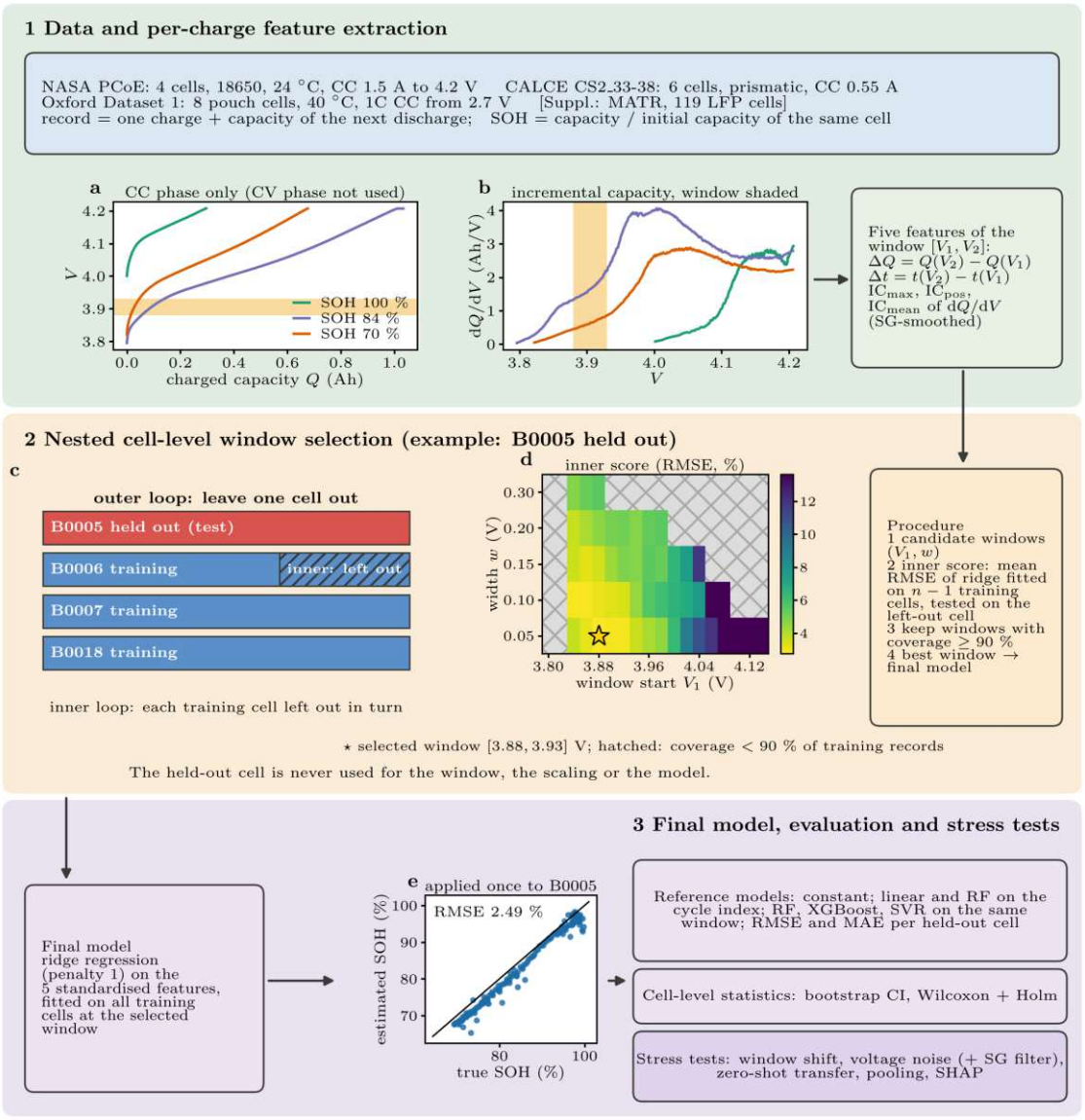


Fig. 1 Framework of the study, with NASA cell B0005 held out as example: (1) feature extraction from a voltage window of the CC phase; (2) nested cell-level window selection on the training cells; (3) final model, evaluation and stress tests. SG: Savitzky–Golay; CI: confidence interval; RF: random forest; SVR: support-vector regression.

Datasets

Three public datasets form the main study (Table 1): the battery data set of the NASA Prognostics Center of Excellence (PCoE), the CS2 cells of the Center for Advanced Life Cycle Engineering (CALCE) of the University of Maryland, and the Oxford Battery Degradation Dataset 1. A fourth, the fast-charging dataset of [6] (hosted at data.matr.io; we call it MATR), requires a different protocol and is reported in the Supplementary Material. The datasets differ in chemistry, charging protocol and sampling on purpose: the aim is to test whether the same *procedure*, not the same window, works across them.

Each charge is paired with the capacity of the following discharge (NASA: reported capacity; CALCE: integral of the discharge current); for Oxford, the capacity comes from the 1C discharge of each characterisation. SOH is the capacity divided by the initial capacity of the same cell. For NASA, charges without a usable CC phase were dropped. For CALCE, we kept cycles with a valid capacity, a CC phase reaching at least 4.15 V and a start voltage below 4.0 V. The analysis is limited to SOH ≥ 70 %, because the cells continue to about 27 %, a range of little practical interest. The Oxford cells are

characterised only every 100 ageing cycles and thus provide few records; they are also the most homogeneous dataset (one temperature, one protocol, similar trajectories) and should be read as the easiest case. CALCE files contain no temperature channel, so temperature is not used as a feature for any dataset.

Table 1 Datasets used. SOH is the capacity divided by the initial capacity of the same cell.

Dataset	Cells	Cell type, conditions	Charging used	Records per cell	Lowest SOH reached
NASA PCoE [45]	B0005, B0006, B0007, B0018	Cylindrical 18650, 2 Ah nominal, 24 C, 2 A discharge	CC 1.5 A to 4.2 V, then CV	132–168	56.7–74.1 %
CALCE CS2 [46,47]	CS2_33–38	Prismatic, about 1.1 Ah (temperature not recorded in the files used)	CC 0.55 A (0.5C) to 4.2 V, then CV	599–778 (SOH $\geq$ 70 %)	70 % (limit of the analysis; cells continue to ≈27 %)
Oxford Dataset 1 [48,49]	Cells 1–8	Kokam pouch, 740 mAh, 40 C	1C CC from 2.7 to 4.2 V, every 100 ageing cycles	46–78	62–80 %

Feature extraction

The CC phase is identified by a current band (NASA 1.4–1.6 A; CALCE 0.50–0.60 A; Oxford: the single CC phase), and the voltage is made monotone by a running maximum. The charged capacity $Q(V)$ at voltage V is the integral of the charging current I over time t (Oxford: the recorded charge). For a window $[V_1, V_2]$ covered by the CC phase, five features are computed:

- $\Delta Q = Q(V_2) - Q(V_1)$, the capacity charged inside the window;
- Δt, the time spent inside the window;
- $\mathrm{IC}_{\max}$, $\mathrm{IC}_{\mathrm{pos}}$ and $\mathrm{IC}_{\mathrm{mean}}$: the maximum, the position (relative to V_1) and the mean of the incremental capacity $\mathrm{d}Q/\mathrm{d}V$ inside the window, computed on a 41-point uniform voltage grid and smoothed with a Savitzky–Golay filter [50] (Fig. 2).

A charge *covers* a window if its CC phase spans the whole window; charges that do not cover a window are not used for that window, and the *coverage* of a window is the fraction of records that cover it. In this paper, a *record* is one charge paired with the capacity of the next discharge.

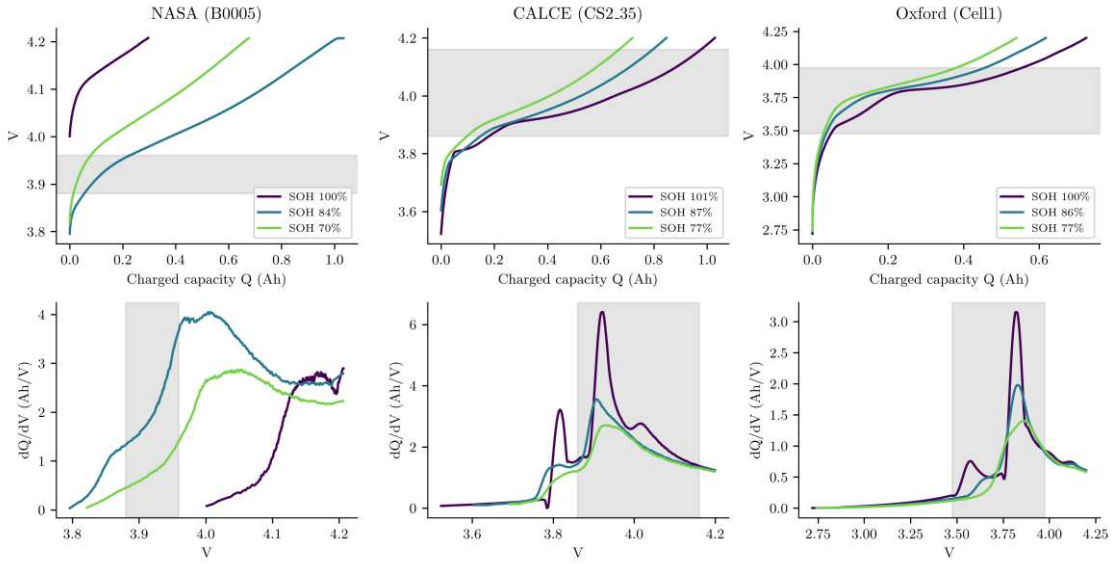


Fig. 2 Charge curves (top) and incremental capacity (bottom) at three SOH levels; the grey band is the median selected window.

Nested window selection by cell

The window is defined by its start V_1 and width w, taken from a candidate grid that depends on the dataset (NASA: V_1 from 3.80 to 4.19 V in steps of 0.02 V, widths 0.05–0.30 V; CALCE: V_1 from 3.60 V, same steps and widths; Oxford: V_1 from 3.0 to 4.2 V in steps of 0.05 V, widths 0.05–0.50 V; $V_2 \leq 4.19$–4.20 V). Choosing a hyper-parameter inside the training data and evaluating on held-out data is the standard remedy for selection bias [30,31]; we apply it at the level of cells. The outer evaluation is leave-one-cell-out. For each held-out cell:

1. The remaining cells form the training set.
2. For each candidate window, an inner leave-one-cell-out loop on the training cells fits a ridge regression on all but one training cell and predicts the remaining one; the mean RMSE over the training cells is the score of the window.
3. Only windows covered by at least 90 % of the training records are admissible (97 % for MATR, Supplementary Material), so that the estimator can run on most charges.
4. The window with the lowest score is used to fit the final model on all training cells, which is then applied once to the held-out cell.

The held-out cell is not used for the window, the scaling or the model, and cells, not cycles, are the unit of splitting. Some design choices were nevertheless fixed after exploratory analyses that included all cells: the grids, the 90 % coverage threshold, the SOH ≥ 70 % cut-off, the choice of CS2_33–38, the ridge penalty and the use of ridge regression as the default model. The reported errors are therefore not strictly out-of-sample with respect to these choices.

Models and metrics

The default model is a ridge regression [51] (penalty 1.0) on standardised features. For comparison, we trained a random forest (RF; 300 trees, minimum leaf size 2), extreme gradient boosting (XGBoost [52]; 300 trees, depth 3, learning rate 0.05, subsample 0.8) and a support-vector regression (SVR) with a radial-basis-function (RBF) kernel (regularisation parameter $C = 30$, tube width $\varepsilon = 0.5$), using scikit-learn [53] and the XGBoost library. In the main comparison all models use the window selected with ridge regression; a check in which each model selects its own window is reported in Section 4.4. Stochastic models were run with three seeds and averaged per cell. The hyper-parameters were fixed in advance and not tuned, which favours no model in particular but does not give the nonlinear models their best accuracy. For a held-out cell with N records, the error of record i is $e_i = \widehat{\mathrm{SOH}}_i - \mathrm{SOH}_i$, where $\widehat{\mathrm{SOH}}_i$ is the estimate. We report the root-mean-square error $\mathrm{RMSE} = \sqrt{N^{-1}\sum_i e_i^2}$ and the mean absolute error $\mathrm{MAE} = N^{-1}\sum_i |e_i|$, both in percentage points of SOH (shown as %).

Statistical analysis

The unit of analysis is the cell. 95 % confidence intervals (CI) of the mean RMSE are obtained by bootstrapping cells [54] (10 000 resamples). Methods are compared with the paired Wilcoxon signed-rank test [55] on per-cell RMSE, within each dataset and over all 18 cells; pairs with zero difference are dropped by the test. The p-values are corrected with the Holm method [56] over a family of six comparisons within each scope (the sixth, 0.05 V against 0.30 V fixed width, is not shown in Table 6). With 4–8 cells per dataset the attainable p-values are limited (for n cells, the smallest exact two-sided p is 0.125 for $n = 4$, 0.031 for $n = 6$ and 0.008 for $n = 8$), and bootstrap intervals based on four cells are themselves unreliable. We therefore use the intervals descriptively.

Robustness, transfer and interpretation tests

Window shift. The window chosen from the training cells is moved by ± 0.02 and ± 0.04 V when features are extracted from the test cell, imitating a voltage offset or a misaligned segmentation.

Voltage noise. Gaussian noise ($\sigma = 2$, 5 and 10 mV, two realisations per level) is added to the voltage of the processed curves; current noise is not added. Models are either trained on clean data, trained on data with the same noise, or applied after a Savitzky–Golay filter (15 samples, order 2) has been used on the voltage of both training and test curves.

Transfer and pooling. A model with the window chosen on a source dataset is applied to the other datasets (zero-shot). In a second test, labelled cells of the other datasets are added to the training set together with k cells of the target dataset.

Interpretation. SHapley Additive exPlanations (SHAP) values [57,58] are computed for the ridge model (linear SHAP) and for XGBoost (tree SHAP) fitted on all cells of a dataset at the window selected from all cells.

Results

Accuracy with nested window selection

Table 2 gives the outer-fold results and Fig. 3 the predictions for the held-out cells. The mean RMSE of SOH is 3.01 % (NASA), 2.28 % (CALCE) and 0.72 % (Oxford). A held-out charge whose CC phase does not cover the selected window cannot be estimated and is skipped; this concerned 23 of 5151 records (coverage of at least 97.7 % in every fold).

The errors are dominated by cell-specific offsets rather than by random scatter. The predictions for a held-out cell often run parallel to the identity line (for example, B0007 and B0018 are over-estimated and CS2_34 is under-estimated over most of its life). Cell-to-cell differences in the relation between window features and capacity are therefore the main source of error. NASA cell B0006, which degrades to 56.7 % SOH, lies outside the range of the other training cells and is over-estimated at its lowest SOH (about 61 % predicted for a true value near 57 %).

Table 2 Leave-one-cell-out results of the ridge model (SOH, %); the bracket is the 95 % cell-bootstrap interval of the mean RMSE.

Dataset	Held-out cell	Selected window (V)	RMSE	MAE
NASA	B0005	3.88–3.93	2.49	2.11
	B0006	3.88–3.98	2.84	2.44
	B0007	3.88–3.98	2.28	1.95
	B0018	3.84–3.94	4.45	3.58
	Mean		**3.01** [2.39, 3.96]	**2.52**
CALCE	CS2_33	3.86–4.16	2.48	1.46
	CS2_34	3.76–4.06	3.81	3.19
	CS2_35	3.86–4.16	1.80	1.22
	CS2_36	3.86–4.16	1.71	1.03
	CS2_37	3.86–4.16	1.94	1.19
	CS2_38	3.86–4.16	1.94	0.93
	Mean		**2.28** [1.83, 2.94]	**1.51**
Oxford	Cell 1	3.45–3.95	0.75	0.69
	Cell 2	3.45–3.95	0.61	0.49
	Cell 3	3.40–3.90	0.39	0.31
	Cell 4	3.50–4.00	0.54	0.52
	Cell 5	3.30–3.80	1.49	0.54
	Cell 6	3.50–4.00	0.33	0.20
	Cell 7	3.50–4.00	0.93	0.91
	Cell 8	3.50–4.00	0.75	0.74
	Mean		**0.72** [0.51, 0.98]	**0.55**

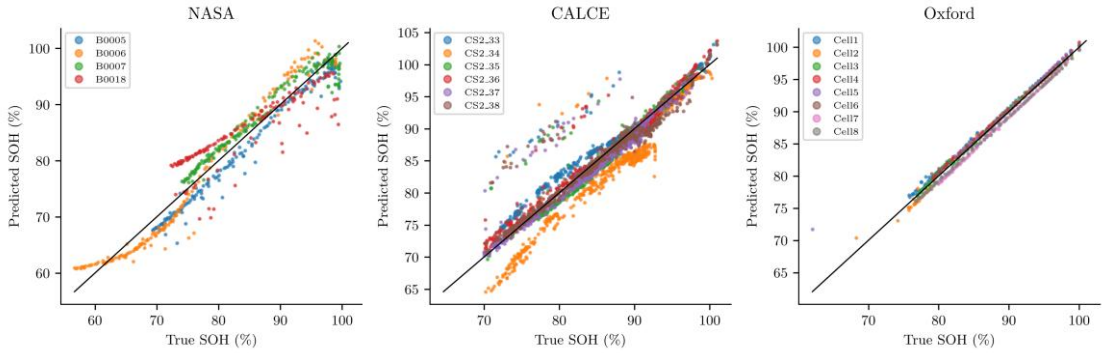


Fig. 3 Predicted against true SOH on the held-out cells (ridge, nested window selection).

Comparison with age-only baselines

A useful estimator must beat models that know only the cycle index. Table 3 reports a constant predictor and linear and random-forest models of the cycle index (the ordinal number of the record for NASA and CALCE, the ageing cycle for Oxford and MATR). The window features reduce the error of the linear age model by about 49 % (NASA), 18 % (CALCE) and 69 % (Oxford). The gain is thus large for NASA and Oxford but modest for CALCE, and on the LFP cells ridge is slightly worse than the cycle index (Supplementary Material). Cycle counts are not always available in service and ignore the usage history, so this baseline is a reference rather than a deployable competitor, but accuracy claims should be read against it.

Table 3 Mean RMSE (SOH, %) of age-only baselines and of the window-based ridge model (MATR: eight groups of cells).

Dataset	Constant	Linear in cycle	Random forest on cycle	Window ridge
NASA	11.17	5.87	5.84	**3.01**
CALCE	6.97	2.78	2.79	**2.28**
Oxford	6.73	2.29	1.93	**0.72**
MATR (LFP)	2.84	**2.00**	–	2.15

Choice of window (RQ1)

The best window width depends on the dataset. Table 4 and Fig. 4 give the RMSE when the width is fixed and only the position is selected, and when both are selected. The best fixed width is about 0.15 V for NASA and 0.30 V for CALCE; for Oxford, the nested procedure chose the 0.50 V width, which is not in the fixed-width grid, in every fold. A single narrow window is therefore not universally adequate. On CALCE, for example, the 0.05 V window has almost twice the error of the selected window (4.27 vs. 2.28 %). On NASA, the nested procedure (3.01 %) is worse than a fixed 0.15 V width (2.57 %), presumably because the inner loop with only three training cells is noisy; the procedure never uses the test cell but cannot guarantee the best window. Finally, the advantage of the nested choice over a fixed 0.05 V window (0.75 points over the 18 cells; 95 % CI of the difference $[-1.67, -0.19]$; raw $p = 0.010$) does not survive Holm correction ($p = 0.059$), and on NASA a fixed narrow window is not worse (Table 6).

Table 4 Mean RMSE (SOH, %) for fixed widths and for the full nested procedure.

Dataset	0.05 V	0.10 V	0.15 V	0.20 V	0.30 V	Nested
NASA	3.16	2.87	**2.57**	2.63	3.59	3.01
CALCE	4.27	2.77	2.58	2.71	**2.28**	**2.28**
Oxford	0.85	0.98	0.93	0.98	0.90	**0.72**

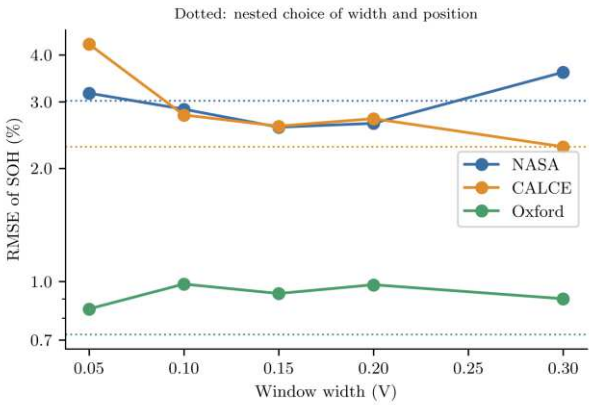


Fig. 4 Mean RMSE against window width (ridge); dotted: nested choice of width and position.

Linear versus nonlinear models (RQ2)

On these three datasets, ridge regression is as accurate as the nonlinear regressors. It is best on NASA and Oxford and within 0.16 points of the best model on CALCE (Table 5). Over the 18 cells, the mean paired RMSE differences of ridge against random forest, XGBoost and SVR are -0.24, -0.30 and -0.31 points; all three 95 % intervals include zero, and the Wilcoxon p-values are 0.42–0.44 (Table 6). We therefore conclude that ridge regression is *as accurate as*, not more accurate than, the nonlinear models, while being much cheaper. This holds for the present protocol: with fixed hyper-parameters and only 3–7 training cells, the nonlinear models may be under-tuned.

Table 5 Mean RMSE (SOH, %) of different regressors; the nonlinear models were not tuned.

Dataset	Ridge	Random forest	XGBoost	SVR
NASA	**3.01**	3.76	3.75	4.17
CALCE	2.28	2.35	**2.12**	2.27
Oxford	**0.72**	0.83	1.15	0.85

Table 6 Paired comparisons of per-cell RMSE (negative: first method better), with Wilcoxon and Holm-corrected p-values.

Comparison	Scope (n cells)	Mean difference [95 % CI]	Cells won	p	p (Holm)
Ridge − RF	All (18)	−0.24 [−0.71, 0.04]	10	0.417	1.00
Ridge − XGBoost	All (18)	−0.30 [−0.97, 0.20]	7	0.417	1.00
Ridge − SVR	All (18)	−0.31 [−0.95, 0.07]	7	0.442	1.00
Nested − fixed 0.05 V	CALCE (6)	−1.99 [−4.34, −0.68]	6	0.031	0.156
Nested − fixed 0.05 V	All (18)	−0.75 [−1.67, −0.19]	12	0.010	0.059
Nested − fixed 0.30 V	Oxford (8)	−0.18 [−0.24, −0.11]	8	0.008	0.047
Nested − fixed 0.30 V	All (18)	−0.21 [−0.51, 0.06]	11	0.028	0.140

Because the windows in Table 5 were selected with ridge regression, they could favour ridge. We therefore repeated the comparison with each model selecting its own window (lighter versions for speed: random forest with 100 trees, XGBoost with 150 trees and learning rate 0.1). The mean RMSE was then 3.35 (random forest) and 3.20 (XGBoost) on NASA, 3.79 and 3.72 on CALCE, and 0.70 and 0.77 on Oxford, against 3.01, 2.28 and 0.72 for ridge. The nonlinear models thus do not clearly beat ridge (the only gain is 0.02 points, for the random forest on Oxford). On CALCE they are clearly worse than with the ridge windows (2.35 and 2.12 in Table 5), which shows that selecting windows with high-variance models from few cells is unstable. The LFP cells behave differently (Section 4.8).

Robustness to window shift and voltage noise (RQ3)

The estimate is sensitive to a shift of the window. A shift of +0.02 V raises the RMSE from 3.01 to 5.25 % (NASA) and from 0.72 to 2.39 % (Oxford), whereas it leaves the 0.30 V-wide CALCE window unchanged (2.28 %; 2.69 % for −0.02 V). A shift of 0.04 V gives 9.99, 4.66 and 4.85 % on NASA, Oxford and CALCE (Fig. 5, left). The voltage reading and the segmentation must therefore be controlled in practice (sensor accuracy, rest-voltage handling, polarisation after rest).

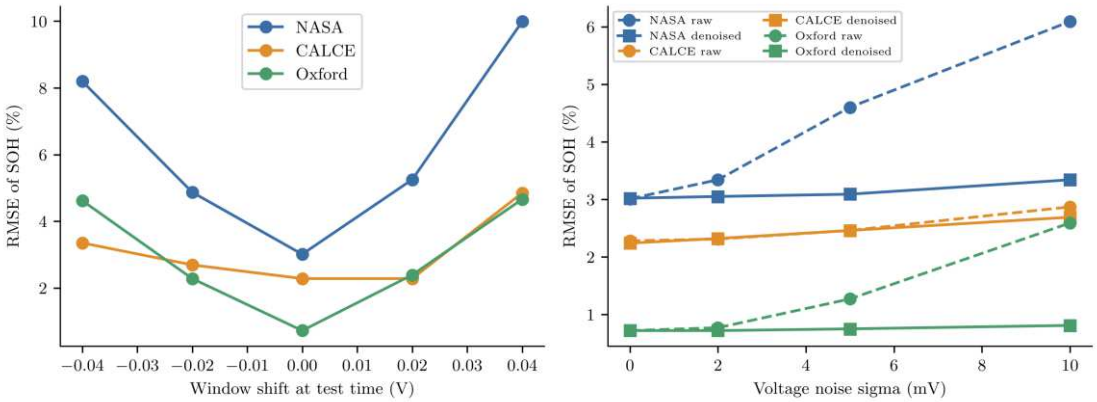


Fig. 5 Left: RMSE under a shift of the selected window. Right: RMSE under voltage noise, without (dashed) and with (solid) Savitzky–Golay filtering.

Voltage noise has a similar effect on models trained on clean data (Fig. 5, right; Table 7): at 10 mV the RMSE reaches 6.09 % on NASA and 2.59 % on Oxford, but only 2.87 % on CALCE, presumably because its wide window averages the noise. Training with the same noise level restores most of the accuracy on NASA (4.05 %) and Oxford (0.98 %) but not on CALCE (3.06 %). Filtering the voltage, identically for training and test curves, is better still on all three datasets (3.34, 0.81 and 2.69 % at

10 mV) and costs almost nothing without noise (3.02, 0.72 and 2.24 %). The filter must be applied identically: in an exploratory run on the NASA data, filtering only the test curves gave 8.5 % even without added noise, because the filtered features differ in distribution from those seen in training.

Table 7 Mean RMSE (SOH, %) under voltage noise for clean-trained, noise-matched and denoised models.

Dataset	Training	0 mV	2 mV	5 mV	10 mV
NASA	clean	3.01	3.34	4.60	6.09
	noise-matched	3.01	3.08	3.40	4.05
	denoised	3.02	3.05	3.09	**3.34**
CALCE	clean	2.28	2.31	2.46	2.87
	noise-matched	2.28	2.41	2.64	3.06
	denoised	2.24	2.32	2.46	**2.69**
Oxford	clean	0.72	0.77	1.27	2.59
	noise-matched	0.72	0.74	0.80	0.98
	denoised	0.72	0.72	0.75	**0.81**

Interpretation

The IC features carry most of the information. For the ridge model on NASA and CALCE, and for XGBoost on CALCE, the IC features (peak and mean) carry most of the attribution; XGBoost on NASA relies mostly on ΔQ, and on Oxford the capacity and time features share the attribution with the IC mean (Fig. 6). A feature-group ablation with the same nested protocol (Table 8) shows that the three IC features alone are as good as all five on every dataset, whereas ΔQ alone is 0.05–0.5 points worse. Because $\Delta Q = I\,\Delta t$ during CC charging (with constant current I), ΔQ, Δt and $\mathrm{IC_{mean}}$ are strongly collinear, and the most important single feature changes between folds (for example NASA: $\mathrm{IC_{mean}}$, ΔQ, $\mathrm{IC_{max}}$, ΔQ). The attributions should therefore be read per feature group, not per feature.

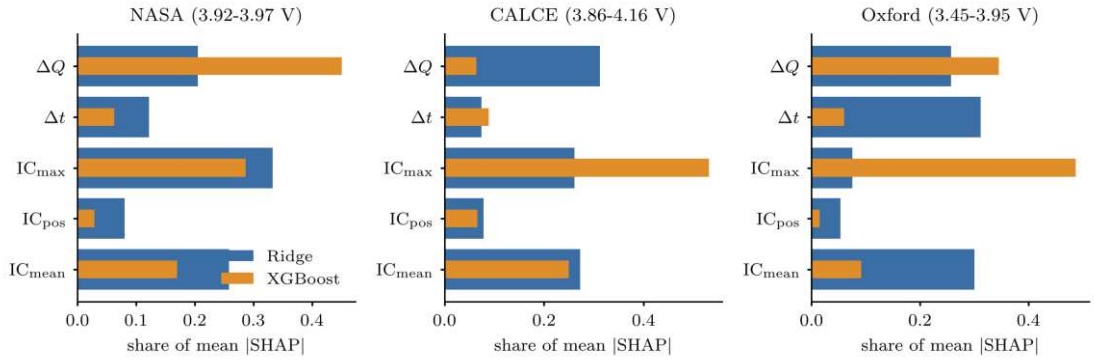


Fig. 6 Share of the mean absolute SHAP value per feature for ridge (blue) and XGBoost (orange).

Table 8 Feature-group ablation: mean RMSE (SOH, %) of ridge regression with nested window selection.

Features	NASA	CALCE	Oxford
ΔQ	3.20	2.73	0.76
ΔQ, Δt	3.26	2.74	0.76
ΔQ, $\mathrm{IC_{max}}$	3.10	2.35	0.70
$\mathrm{IC_{max}}$, $\mathrm{IC_{pos}}$, $\mathrm{IC_{mean}}$	**2.99**	**2.22**	**0.71**
All five	3.01	2.28	0.72

Transfer between datasets (RQ4)

Models do not transfer between datasets. We applied the window and the model fitted on one dataset to another, without any labelled cell of the target dataset (zero-shot). The RMSE ranged from 6.9 % (NASA→CALCE; 7.5 % without the recalibration described next) to 62.9 % (Oxford→CALCE). A simple recalibration subtracts a constant so that the first three predictions equal 100 %, the known SOH of a new cell. It reduced several errors (for example Oxford→CALCE from 62.9 to 23.8 %), but none became acceptable. Normalising features by each cell's first-cycle value did not help either. Adding labelled cells of the other datasets to the training set did not help either: with two target cells the RMSE was 3.54 % (CALCE) and 2.24 % (Oxford) when the other datasets were added, against 2.42 % and 0.98 % with target cells only; for NASA the effect was small (3.12 vs. 3.33 %). The estimator should therefore be fitted per cell type. Generalisation across datasets has been attempted with a voltage-drop-based feature and a small neural network blind-tested on public datasets [59]; we did not test such an approach.

LFP fast-charging cells (summary of the Supplementary Material)

On the 119 LFP cells of the MATR dataset [6], analysed with a modified protocol (Supplementary Material), the conclusions differ. A window on the final 1 A charging stage gives an RMSE of 2.15 % for ridge, against 2.00 % for the cycle-index model, and 1.65–1.68 % for tree ensembles. A 0.3 V window on the constant-current discharge gives 1.56 % for ridge and about 1.0 % for trees (Supplementary Table S1). The sensitivities to shift and noise are similar, with narrow charge windows failing under noise unless the voltage is filtered (Supplementary Table S2). The Supplementary Material also describes a coverage pitfall that made a first analysis look better than it was.

Computational cost

On a single processor core (scikit-learn, one sample at a time), feature extraction took about 0.8 ms per charge. The ridge model took about 0.2 ms and has six parameters (five coefficients and an intercept); the random forest took about 14 ms and has about 10^5 tree nodes. These timings come from a cloud container and an earlier six-feature version of the pipeline. They indicate relative cost, not the performance on a BMS microcontroller, which we did not measure.

Discussion

Main findings. (1) On three LCO/NMC-type datasets, a ridge regression on five window features, with the window selected from training cells only, estimates SOH with an RMSE of 0.7–3 %. The gain over a model of the cycle index is large for NASA and Oxford but only 18 % for CALCE. (2) Linear and nonlinear models are statistically indistinguishable, so the cheaper model is preferable for an embedded estimator. This differs from the benchmark of [27], in which sequence models were the most accurate and gradient-boosted trees were mid-tier on 32 NASA cells. That study used raw sequences, many more training cells and a blocked 5-fold GroupKFold protocol, whereas our estimators see five window features and 3–7 training cells; nonlinear models may therefore gain an advantage that our data cannot show. (3) There is no universal window: the best width and position depend on the cell type, and the benefit of nested selection over a fixed window is clear only for some datasets and does not remain significant after correction for multiple comparisons. (4) The estimator is sensitive to a shift of the window and to voltage noise unless the voltage is filtered, and it does not transfer between datasets; pooling labelled cells of other types hurts. (5) On the 119 LFP cells of MATR the conclusions do not carry over: the flat LFP plateau and the heterogeneous fast-charging histories leave little information in the final 1 A stage, ridge does not beat the cycle index, and tree ensembles are only moderately better (Supplementary Material).

Reading the results. Very low errors should be read together with their validation design. Our Oxford result (0.72 %) is the most favourable but also the least demanding: the cells were aged with a repeated

drive-cycle profile and characterised every 100 cycles, and a related caution is that models trained on repeated profiles may exploit the repetition [24].

Practical implications. (i) Select the window with a procedure that cannot see the test cells, and report the selected windows. (ii) Report the sensitivity to a window shift and to voltage noise: a 0.02 V shift multiplied the error by 1.6–3.3 on the NASA and Oxford windows. (iii) Fit the estimator per cell type, and filter the voltage identically in training and deployment. (iv) Report confidence intervals at the level of cells; with 4–8 cells per dataset, many differences between methods cannot be established.

Limitations. (1) There are few cells per dataset (4, 6 and 8), and the inner selection for NASA uses only three training cells. (2) Several design choices were made after exploratory analyses on all cells (Section 3.4), so the errors are optimistic to an unknown degree. (3) Only the CC phase is used; the CV phase and rest voltages are not exploited, and the windows can be compared only indirectly with the voltage-range optimisation of [13,14], which used other cells and objectives. (4) CALCE was restricted to CS2_33–38 and SOH ≥ 70 %; the other CS2 cells were examined but not used, because their files mix characterisation and partial cycles, so the capacity of a cycle is not the cell capacity. (5) The shift and noise tests used processed curves, two noise realisations and voltage noise only; real sensor errors may be correlated and include offsets. (6) The nonlinear models were not tuned, and the window-selection check used lighter versions of them. (7) Cross-temperature behaviour was checked only in a preliminary run (zero-shot RMSE 14–24 % on NASA cells aged at 4 C, with an earlier six-feature version), which suggests that temperature-dependent features or per-temperature models are needed. (8) The cost figures come from a cloud container and an earlier version of the pipeline, and there is no hardware-in-the-loop test or uncertainty quantification beyond bootstrap intervals. (9) The MATR analysis uses a different evaluation design (eight groups of cells, every tenth cycle, 119 of the original 124 cells) and is not pooled with the other datasets.

Future work. Future work should (i) include the CV phase and rest voltages; (ii) use the early part of the charge or the discharge curve for LFP cells, where the final 1 A stage carries little information; (iii) add uncertainty quantification, for example conformal prediction; (iv) test few-shot calibration with a single labelled cell; and (v) measure the cost on an embedded target.

Conclusion

We studied partial-charge SOH estimation with five window features and a ridge regression, selecting the window from training cells only. On three public LCO/NMC-type datasets, the RMSE of SOH was 3.01 %, 2.28 % and 0.72 % under leave-one-cell-out validation. Ridge regression was statistically indistinguishable from random forest, gradient boosting and support-vector regression, at a much lower cost. The gain over an age-only baseline was large for NASA and Oxford, modest for CALCE and absent for ridge on 119 LFP cells. The best window differs between datasets, the estimate is sensitive to a window shift and to voltage noise unless the voltage is filtered, and models do not transfer across cell types. These findings support window selection that never sees the evaluation cell, fitting per cell type, and explicit reporting of robustness.

References

[1] Berecibar, M., Gandiaga, I., Villarreal, I., Omar, N., Van Mierlo, J., Bossche, P.: Critical review of state of health estimation methods of Li-ion batteries for real applications. *Renewable and Sustainable Energy Reviews* **56**, 572–587 (2016) https://doi.org/10.1016/j.rser.2015.11.042

[2] Xiong, R., Li, L., Tian, J.: Towards a smarter battery management system: A critical review on battery state of health monitoring methods. *Journal of Power Sources* **405**, 18–29 (2018) https://doi.org/10.1016/j.jpowsour.2018.10.019

[3] Li, M., Xu, W., Zhang, S., Liu, L., Hussain, A., Hu, E., Zhang, J., Mao, Z., Chen, Z.: State of Health Estimation and Battery Management: A Review of Health Indicators, Models and Machine Learning. *Materials* **18**(1), 145 (2025) https://doi.org/10.3390/ma18010145

[4] Rauf, H., Khalid, M., Arshad, N.: Machine learning in state of health and remaining useful life estimation: Theoretical and technological development in battery degradation modelling. *Renewable and Sustainable Energy Reviews* **156**, 111903 (2022) https://doi.org/10.1016/j.rser.2021.111903

[5] Roman, D., Saxena, S., Robu, V., Pecht, M., Flynn, D.: Machine learning pipeline for battery state-of-health estimation. *Nature Machine Intelligence* **3**(5), 447–456 (2021) https://doi.org/10.1038/s42256-021-00312-3

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

[7] Li, X., Yuan, C., Wang, Z.: State of health estimation for Li-ion battery via partial incremental capacity analysis based on support vector regression. *Energy* **203**, 117852 (2020) https://doi.org/10.1016/j.energy.2020.117852

[8] Wei, Z., Ruan, H., Li, Y., Li, J., Zhang, C., He, H.: Multistage State of Health Estimation of Lithium-Ion Battery With High Tolerance to Heavily Partial Charging. *IEEE Transactions on Power Electronics* **37**(6), 7432–7442 (2022) https://doi.org/10.1109/tpel.2022.3144504

[9] Jenu, S., Hentunen, A., Haavisto, J., Pihlatie, M.: State of health estimation of cycle aged large format lithium-ion cells based on partial charging. *Journal of Energy Storage* **46**, 103855 (2022) https://doi.org/10.1016/j.est.2021.103855

[10] He, J., Meng, S., Li, X., Yan, F.: Partial Charging-Based Health Feature Extraction and State of Health Estimation of Lithium-Ion Batteries. *IEEE Journal of Emerging and Selected Topics in Power Electronics* **11**(1), 166–174 (2023) https://doi.org/10.1109/jestpe.2022.3143831

[11] Singh, A., Barai, A.: Parametric investigation of partial incremental capacity analysis for Lithium-ion cells. *Journal of Energy Storage* **113**, 115668 (2025) https://doi.org/10.1016/j.est.2025.115668

[12] Luan, W., Cai, H., Wang, X., Zhao, B.: State of Health for Lithium-Ion Batteries Based on Explainable Feature Fragments via Graph Attention Network and Bi-Directional Gated Recurrent Unit. *Sensors* **25**(19), 5953 (2025) https://doi.org/10.3390/s25195953

[13] Meng, J., Cai, L., Stroe, D.-I., Luo, G., Sui, X., Teodorescu, R.: Lithium-ion battery state-of-health estimation in electric vehicle using optimized partial charging voltage profiles. *Energy* **185**, 1054–1062 (2019) https://doi.org/10.1016/j.energy.2019.07.127

[14] Tian, Y., Dong, Q., Tian, J., Li, X., Li, G., Mehran, K.: Capacity estimation of lithium-ion batteries based on optimized charging voltage section and virtual sample generation. *Applied Energy* **332**, 120516 (2023) https://doi.org/10.1016/j.apenergy.2022.120516

[15] Yang, S., Zhang, C., Jiang, J., Zhang, W., Gao, Y., Zhang, L.: A voltage reconstruction model based on partial charging curve for state-of-health estimation of lithium-ion batteries. *Journal of Energy Storage* **35**, 102271 (2021) https://doi.org/10.1016/j.est.2021.102271

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

[17] Liu, Y., Fan, G., Zhou, B., Chen, S., Sun, Z., Wang, Y., Zhang, X.: Rapid and flexible battery capacity estimation using random short-time charging segments based on residual convolutional networks. *Applied Energy* **351**, 121925 (2023) https://doi.org/10.1016/j.apenergy.2023.121925

[18] Zhao, L., Du, X., Yang, Z., Xia, C., Xue, J., Hoque, M.J., Fu, W., Yan, X., Miljkovic, N.: Online State of Health Estimation with Deep Learning Frameworks Based on Short and Random Battery Charging Data Segments. *Journal of The Electrochemical Society* **170**(9), 090537 (2023) https://doi.org/10.1149/1945-7111/acf8ff

[19] Tian, A., Chen, Z., Pan, Z., Yang, C., Wang, Y., Dong, K., Gao, Y., Jiang, J.: Li-Ion Battery State of Health Estimation Based on Short Random Charging Segment and Improved Long Short-Term Memory. *IET Signal Processing* **2023**, 8839034 (2023) https://doi.org/10.1049/2023/8839034

[20] Deng, Z., Hu, X., Xie, Y., Xu, L., Li, P., Lin, X., Bian, X.: Battery health evaluation using a short random segment of constant current charging. *iScience* **25**(5), 104260 (2022) https://doi.org/10.1016/j.isci.2022.104260

[21] Zheng, K., Meng, J., Yang, Z., Zhou, F., Yang, K., Song, Z.: Refined lithium-ion battery state of health estimation with charging segment adjustment. *Applied Energy* **375**, 124077 (2024) https://doi.org/10.1016/j.apenergy.2024.124077

[22] Ke, L., Lin, M., Meng, J., Wu, J., Wang, W., Wang, F.: A sparse SOH estimation framework based on clustered voltage segments and multi-head CNN under arbitrary charging condition. *Energy* **344**, 140057 (2026) https://doi.org/10.1016/j.energy.2026.140057

[23] Zhang, L., Zhang, J., Gao, T., Lyu, L., Wang, L., Shi, W., Jiang, L., Cai, G.: Improved LSTM based state of health estimation using random segments of the charging curves for lithium-ion batteries. *Journal of Energy Storage* **74**, 109370 (2023) https://doi.org/10.1016/j.est.2023.109370

[24] Vilsen, S.B., Stroe, D.-I.: On the use of randomly selected partial charges to predict battery state-of-health. *Batteries* **10**(6), 193 (2024) https://doi.org/10.3390/batteries10060193

[25] Rizanov, S., Stoynova, A., Mihov, G.: Toward Trustworthy and Transferable SOH/RUL Estimation for Lithium-Ion Batteries: A Critical Review and Multi-Fidelity Validation Framework from Laboratory Cells to Real-World Packs. *Batteries* **12**(7), 255 (2026) https://doi.org/10.3390/batteries12070255

[26] Sohail, M., Tanveer, M., Kim, H.S.: Trustworthy Machine Learning and Mathematical Modelling for Lithium-Ion Battery State-of-Health Estimation. *Mathematics* **14**(11), 1879 (2026) https://doi.org/10.3390/math14111879

[27] Madani, S.S., Hébert, M., Boulon, L., Lupien-Bédard, A., Allard, F.: Comparative Analysis of ML and DL Models for Data-Driven SOH Estimation of LIBs Under Diverse Temperature and Load Conditions. *Batteries* **11**(11), 393 (2025) https://doi.org/10.3390/batteries11110393

[28] Anonymous: Reference withheld for double-blind review (2026)

[29] Kapoor, S., Narayanan, A.: Leakage and the reproducibility crisis in machine-learning-based science. *Patterns* **4**(9), 100804 (2023) https://doi.org/10.1016/j.patter.2023.100804

[30] Varma, S., Simon, R.: Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* **7**(1), 91 (2006) https://doi.org/10.1186/1471-2105-7-91

[31] Cawley, G.C., Talbot, N.L.C.: On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research* **11**, 2079–2107 (2010)

[32] Dubarry, M., Svoboda, V., Hwu, R., Yann Liaw, B.: Incremental Capacity Analysis and Close-to-Equilibrium OCV Measurements to Quantify Capacity Fade in Commercial Rechargeable Lithium

Batteries. *Electrochemical and Solid-State Letters* **9**(10), 454 (2006) https://doi.org/10.1149/1.2221767

[33] Weng, C., Cui, Y., Sun, J., Peng, H.: On-board state of health monitoring of lithium-ion batteries using incremental capacity analysis with support vector regression. *Journal of Power Sources* **235**, 36–44 (2013) https://doi.org/10.1016/j.jpowsour.2013.02.012

[34] Zheng, L., Zhu, J., Lu, D.D.-C., Wang, G., He, T.: Incremental capacity analysis and differential voltage analysis based state of charge and capacity estimation for lithium-ion batteries. *Energy* **150**, 759–769 (2018) https://doi.org/10.1016/j.energy.2018.03.023

[35] Li, X., Yuan, C., Li, X., Wang, Z.: State of health estimation for Li-Ion battery using incremental capacity analysis and Gaussian process regression. *Energy* **190**, 116467 (2020) https://doi.org/10.1016/j.energy.2019.116467

[36] He, J., Bian, X., Liu, L., Wei, Z., Yan, F.: Comparative study of curve determination methods for incremental capacity analysis and state of health estimation of lithium-ion battery. *Journal of Energy Storage* **29**, 101400 (2020) https://doi.org/10.1016/j.est.2020.101400

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

[38] Fan, L., Wang, P., Cheng, Z.: A remaining capacity estimation approach of lithium-ion batteries based on partial charging curve and health feature fusion. *Journal of Energy Storage* **43**, 103115 (2021) https://doi.org/10.1016/j.est.2021.103115

[39] Kohtz, S., Xu, Y., Zheng, Z., Wang, P.: Physics-informed machine learning model for battery state of health prognostics using partial charging segments. *Mechanical Systems and Signal Processing* **172**, 109002 (2022) https://doi.org/10.1016/j.ymssp.2022.109002

[40] Shu, X., Yang, H., Liu, X., Feng, R., Shen, J., Hu, X.: State of Health Estimation for Lithium-Ion Batteries Based on Voltage Segment and Transformer. *Journal of Energy Storage* **108**, 115200 (2025) https://doi.org/10.1016/j.est.2024.115200

[41] Li, Y., Gao, Y., Wang, X., Wang, C., Wang, Z., Wu, C., Wang, J.: Transfer learning-based SOH estimation from partial charging data with polarization-aware modeling. *Scientific Reports* **16**(1), 23076 (2026) https://doi.org/10.1038/s41598-026-48906-4

[42] Zhao, J., Li, D., Li, Y., Shi, D., Nan, J., Burke, A.F.: Battery state of health estimation under fast charging via deep transfer learning. *iScience* **28**(5), 112235 (2025) https://doi.org/10.1016/j.isci.2025.112235

[43] Shi, M., Xu, J., Lin, C., Mei, X.: A fast state-of-health estimation method using single linear feature for lithium-ion batteries. *Energy* **256**, 124652 (2022) https://doi.org/10.1016/j.energy.2022.124652

[44] Jo, S., Jung, S., Roh, T.: Battery State-of-Health Estimation Using Machine Learning and Preprocessing with Relative State-of-Charge. *Energies* **14**(21), 7206 (2021) https://doi.org/10.3390/en14217206

[45] Saha, B., Goebel, K.: Battery data set. NASA Ames Prognostics Data Repository. NASA Ames Research Center, Moffett Field, CA (2007)

[46] CALCE Battery Research Group: Battery data. Center for Advanced Life Cycle Engineering, University of Maryland. Accessed 2026. https://calce.umd.edu/battery-data

[47] He, W., Williard, N., Osterman, M., Pecht, M.: Prognostics of lithium-ion batteries based on Dempster–Shafer theory and the Bayesian Monte Carlo method. *Journal of Power Sources* **196**(23), 10314–10321 (2011) https://doi.org/10.1016/j.jpowsour.2011.08.040

[48] Birkl, C.R.: Oxford Battery Degradation Dataset 1. University of Oxford, Oxford University Research Archive. Dataset (2017). https://doi.org/10.5287/bodleian:KO2kdmYGg

[49] Birkl, C.R.: Diagnosis and prognosis of degradation in lithium-ion batteries. PhD thesis, University of Oxford (2017)

[50] Savitzky, A., Golay, M.J.E.: Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry* **36**(8), 1627–1639 (1964) https://doi.org/10.1021/ac60214a047

[51] Hoerl, A.E., Kennard, R.W.: Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **12**(1), 55–67 (1970) https://doi.org/10.1080/00401706.1970.10488634

[52] Chen, T., Guestrin, C.: XGBoost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794 (2016). https://doi.org/10.1145/2939672.2939785

[53] Pedregosa, F., et al.: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011)

[54] Efron, B.: Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics* **7**(1), 1–26 (1979) https://doi.org/10.1214/aos/1176344552

[55] Wilcoxon, F.: Individual Comparisons by Ranking Methods. *Biometrics Bulletin* **1**(6), 80 (1945) https://doi.org/10.2307/3001968

[56] Holm, S.: A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* **6**(2), 65–70 (1979)

[57] Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. In: *Advances in Neural Information Processing Systems 30*, pp. 4765–4774 (2017)

[58] Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I.: From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence* **2**(1), 56–67 (2020) https://doi.org/10.1038/s42256-019-0138-9

[59] Sahoo, S., Hariharan, K., Agarwal, S., et al.: Transfer learning based generalized framework for state of health estimation of Li-ion cells. *Scientific Reports* **12**(1), 13173 (2022) https://doi.org/10.1038/s41598-022-16692-4

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- codeandresults.zip
- 03SupplementaryMaterialAppendixandTables.docx