



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Capacity and Internal Resistance of lithium-ion batteries: full degradation curve prediction from Voltage response at constant Current at discharge

Citation for published version:

Ibraheem, R, Strange, C & Dos Reis, G 2023, 'Capacity and Internal Resistance of lithium-ion batteries: full degradation curve prediction from Voltage response at constant Current at discharge', *Journal of Power Sources*, vol. 556, 232477. <https://doi.org/10.1016/j.jpowsour.2022.232477>

Digital Object Identifier (DOI):

[10.1016/j.jpowsour.2022.232477](https://doi.org/10.1016/j.jpowsour.2022.232477)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Journal of Power Sources

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

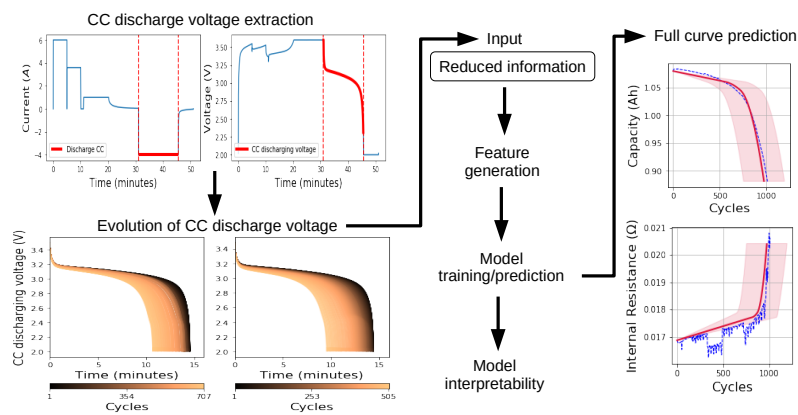
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Graphical Abstract




Capacity and Internal Resistance of lithium-ion batteries: full degradation curve prediction from Voltage response at constant Current at discharge

👤 Rasheed Ibraheem, 👤 Calum Strange, 👤 Gonçalo dos Reis



Highlights

Capacity and Internal Resistance of lithium-ion batteries: full degradation curve prediction from Voltage response at constant Current at discharge

 Rasheed Ibraheem,  Calum Strange,  Gonçalo dos Reis

- Early-life prediction of Capacity and Internal Resistance degradation curves
- Only the voltage response at CC during the discharge phase is used as input
- Prediction of full degradation trajectory
- Provide interpretability and feature analysis
- Impact analysis of Voltage data subsampling in the machine learning models

Capacity and Internal Resistance of lithium-ion batteries: full degradation curve prediction from Voltage response at constant Current at discharge

 Rasheed Ibraheem^a,  Calum Strange^a,  Gonçalo dos Reis^{a,b,*}

^a*School of Mathematics, University of Edinburgh, The Kings buildings, Edinburgh, EH9 3JF, UK*

^b*Centro de Matematica e Aplicacoes (CMA), Faculdade de Ciencias e Tecnologia, Campus da Caparica, Caparica, 2829-516, Portugal*

Abstract

The use of minimal information from battery cycling data for various battery life prognostics is in high demand with many current solutions requiring full in-cycle data recording across 50-100 cycles. In this research, we propose a data-driven, feature-based machine learning model that predicts the entire capacity fade and internal resistance curves using only the voltage response from constant current discharge (fully ignoring the charge phase) over the first 50 cycles of battery use data. This approach is applicable where the discharging component is controlled and consistent, but sufficiently general to be applicable to settings with controlled charging but noisy discharge as is the case of electric vehicles.

We provide a detailed analysis of the impact of the generated features on the model. We also investigate the impact of sub-sampling the voltage curve on the model performance where it was discovered that taking voltage measurements at every 1 minute is enough for model input without loss of quality. Example performance includes Capacity's and Internal Resistance's end of life being predicted with a mean absolute error (MAE) of 71 cycles and $1.5 \times 10^{-5} \Omega$ respectively.

Keywords: Capacity degradation, Internal resistance degradation, Prediction of full degradation curve, Voltage response under constant current at discharge, Lithium-ion cells, Machine learning, Remaining useful life

*Corresponding author

1. Introduction

Lithium-ion battery modelling is a fast growing research field. This can be linked to the fact that lithium-ion batteries have desirable properties such as affordability, high longevity and high energy densities [1, 2, 3]. In addition, they are deployed to various applications ranging from small devices including smartphones and laptops to a more complicated and fast growing products such as electric vehicles. However, as batteries age their ability to store energy (capacity) fades by the influence of different mechanisms: usage, storage, environment, chemistry and combinations thereof. For many cell chemistries and use cases the degradation throughout time is nonlinear [4, 5]. This calls for the development of tools able to capture the degradation pattern necessary for effective battery management system, battery longevity classification, and quality control. Linked to capacity fade is the internal resistance (IR) rise curve which quantifies the amount of opposition to the flow of current in and out of a battery [6]. A considerable volume of work has been done to understand [5, 7], detect [8, 6] and predict [8, 6, 9, 10] key quantities relating to the evolution of cell capacity and IR.

Such models draw on rich datasets [11] from cell-cycling testing containing current, voltage and temperature response time series across many cycles of charge and discharge, plus reference performance tests to determine capacity and IR. Of high relevance to this manuscript are data-centric feature-based modelling approaches whose main advantage is explainability. This approach has been used extensively where statistical, in-field and physics-informed features are engineered, selected based on their relevance, and fed into a machine learning predictive algorithm [12, 8, 13, 14, 15]. Of these, we give strong emphasis to Paulson et al. [15] where using the first 100 cycles of cycling data (spanning through NMC111, NMC532, NMC622, NMC811, HE5050 and 5Vspinel lithium-ion cells) they combine in-field knowledge with statistics to generate a total of 396 features from the cycling data, after which they applied a sequential feature selection technique to pick the best features for modelling purpose. After considering several data-centric models for prediction purposes they predict the cycle life with mean absolute error (MAE) of 78 cycles and include an exhaustive analysis of the importance of each of the features in the prediction. In alternative, other models taking a feature free approach have been proposed utilising deep learning approaches [9, 16, 17]. These show promising predictive performance from limited data but offer limited explainability.

The lab data on which the above approaches were developed, contains consistently controlled charge and discharge profiles. For many real world

applications the charge or discharge component may be less predictable. For example, the charging component of electric vehicles (EVs) use is completely controlled by the charging point and battery management system (BMS) [18]. On the other hand, the discharge component depends on the chosen route, level of traffic, driver habits and is non-monotonic (regenerative braking). For EVs it is thus easier to design prognostics based on the charging component. As for the situation where the discharging component is consistent, a good example is the storage for renewable energy. In this case, the charging component relies on the availability of the natural sources such as the intensity of the sun in solar panels and wind speed in wind turbines); however, the quantity of derived energy is controlled and monitored [19]. It is thus of interest that the dataset used in this research contains multiple differing charge profiles, but predictions can be made from the common discharge component.

The aim of this paper is to develop a model to predict the full capacity and IR trajectory predictive model (including end of life (EOL)) taking into account limitations seen in real-life battery usage and requiring a limited number of input cycles (and data). The datasets chosen for this research, [12, 20, 6], have a large number of different charging profiles across the considered cells but a consistent discharge profile for all cells. For this reason we focus on prediction models using only the constant current (CC) discharge Voltage response part of the cycling data. We aim to develop a quick, cheap, computationally efficient, and highly practical model.

To the best of our knowledge, the idea of using CC discharging Voltage for capacity/IR early life curve prediction is new and has not been addressed in the research domain except for the very recent [17] (carried out independently) and that we discuss in more detail in Section 4.5 below. Work to estimate state-of-health (SOH) under these constraints has been conducted, mainly making use of various deep learning techniques to estimate in-cycle capacity and impedance spectra [21, 22, 23]. Inspired in [8] we investigate how many cycles of input data are needed for effective predictions and also the effect of frequency of data-recording (i.e., subsampling the voltage response) in the predictor model. In fact, up to [8], feature-based data-driven models took 100 or more cycles worth of input data for prognostics and in [8] it was shown that 50-cycles was enough. Then in [9] it was shown that one-single cycle of input data was enough to achieve an accuracy rate better than that of [8], the model used in [9] is a convolutional deep neural networks that is less explainable. Inspired by [15], we develop a model based on explainable feature mechanisms and provide an analysis of the impact of generated features on model accuracy. This will give an insight

into what features are important for curve prediction and thus speaking to a parsimonious model.

We structure this paper as follows: Section 2 gives information about the data used in this study and Section 3 provides detail of the methods adopted for feature generation and model building. Section 4 presents the results and their rationalization, Section 5 provides concluding remarks and potential future study.

2. Data Description

The data used in this study are those presented in [12], [20] and [24]. Although cells in each of these papers differ slightly in terms of cycling process, they have almost identical cycling conditions and experimental recorded data are presented in the same format. Thus, their combination is suitable. They contain data for Lithium Ferrophosphate (LFP)/graphite A123 APR18650M1A cells each with nominal capacity of $1.1Ah$ and nominal voltage of $3.3V$. As described in the three works, all cells are cycled at a constant temperature of $30^{\circ}C$ and, *critically for our work: the same discharging protocol was used throughout, but with a large variety of different fast-charging policies*. About 80 different charge policies were tested and they are explicitly described in Table 9 of the supplementary material to [12] – see additionally [11] for a description of the datasets as whole.

The generated data is presented in eight batches, namely batches 1 to batch 8: as a convention, each cell is named as $bMcN$ where M and N denote the batch number and cell number respectively. It is important to note that cells from batches 4 to 7 are not cycled to EOL, thus, we excluded them from this research; cells from batch 8 are cycled pass their EOL, for consistency with batches 1 to 3, we only use data up to the EOL. In addition, we excluded 11 outlier cells that last for more than 1200 cycles (inline with other research [14, 25]) which leaves us with 158 cells.

Several measurements are included in the data but we only extract the CC discharging voltage, discharging time, per cycle measure of capacity and IR. As batch 8 does not contain IR measurements we use the corresponding data from [24] as substitute. We present the evolution of the CC discharging voltage for some selected cells in Figure 1. The maximum and minimum voltage across the cells can be seen to range between approximately $3.4V$ and $2.0V$. In addition, the curve diminishes as each of the cells ages and degrades. This is a key observation, which later forms the basis for the techniques used for feature generation.

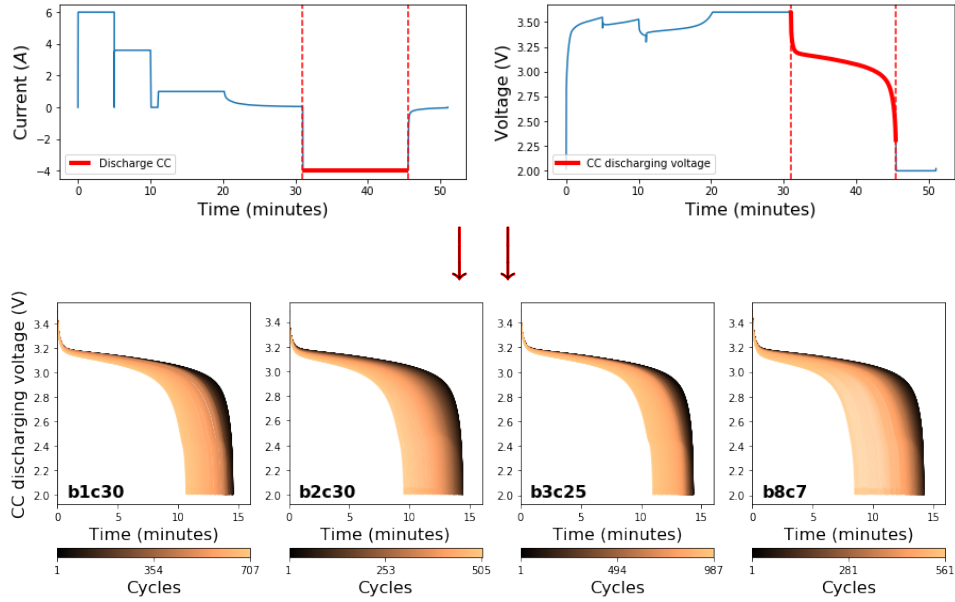


Figure 1: Top: Current-voltage profile for sample cell at sample cycle. Bottom: Evolution of CC discharging voltage response curve for selected cells across full cycle life (note: the profiles we use are close but different from CC-CV ones).

3. Modelling and Methodology

In this section, we present the modelling approaches for the early prediction of capacity fade and IR rise curves. First we discuss the in-field informed feature generation approach used in this research, then give details about the choice of machine learning algorithm, the number of cycles needed for prediction and the technique adopted in predicting the entire curves.

3.1. Feature generation

Our main feature of consideration is the CC discharge voltage and its evolution. To construct features based on this we used statistical (mean, variance, kurtosis, skewness, min and max) and gradient based approaches (area under curve and slope). The calculation of gradient and area under the curve is not obvious and requires a careful processing of the CC discharge voltage and its corresponding time. Values for voltage are not measured at regular time intervals, thus we fitted a one-dimensional linear interpolation to the values and the associated time (using Numpy in *interp1d* function [26]). We then used the resulting interpolated function to obtain voltage

values at equal time steps (every 4 seconds) while taking the initial and final times into consideration. Following this approach, area under the curve and the rate of change of the voltage curve were computed using the *trapz* and *gradient* functions of the Numpy library. Partly inspired by [15], for each cell and cycle (i) we construct the following set of features:

$$F_i^{(\text{cell})} := \begin{cases} \min\{V_k, k = 1, 2, \dots, l_i\}, & \max\{V_k, k = 1, 2, \dots, l_i\}, \\ \text{mean}\{V_k, k = 1, 2, \dots, l_i\}, & \text{var}\{V_k, k = 1, 2, \dots, l_i\}, \\ \text{kurt}\{V_k, k = 1, 2, \dots, l_i\}, & \text{skew}\{V_k, k = 1, 2, \dots, l_i\} \\ \min \left\{ \left(\frac{dV}{dt} \right)_k, k = 1, 2, \dots, l_i \right\}, \\ \max \left\{ \left(\frac{dV}{dt} \right)_k, k = 1, 2, \dots, l_i \right\}, \\ \frac{h}{2} (V_1 + 2(V_2 + \dots + V_{l_i-1}) + V_{l_i}), \\ \left(\frac{dV}{dt} \right)_{k=1}, \quad \left(\frac{dV}{dt} \right)_{k=l_i}; \end{cases} \quad (1)$$

where V is the interpolated CC discharge voltage for cycle i of the selected cell and l_i denotes its length; *var*, *kurt* and *skew* denote the usual variance, kurtosis and skewness operations respectively; h is the time step (in minutes).

To include information about how features from in-cycle data compare with others, we calculated multi-cycle features that find the difference between values of features at two different cycle numbers (gradient information). As gradient features are more sensitive to noise, we consider median values from windows of width i_{10} and multi-cycle features are calculated according to

$$\begin{cases} f_0 := \text{median}\{F_i^{(\text{cell})}, i = 1, 2, \dots, i_{10}\}, \\ f_{n/2} := \text{median}\{F_i^{(\text{cell})}, i = n/2 - i_{10}, n/2 - i_{10} + 1, \dots, n/2 + i_{10}\}, \\ f_n := \text{median}\{F_i^{(\text{cell})}, i = n - i_{10}, n - i_{10} + 1, \dots, n\}, \\ f_{n-0} := f_n - f_0, \\ f_{diff} := f_n - 2f_{n/2} - f_0; \end{cases} \quad (2)$$

where n is the number of cycles selected for feature generation (we choose $n = 50$) and i_{10} is the cycle number corresponding to 10% of n . In total we generated 54 features using this approach, and their description is provided in Table 1.

Table 1: A comprehensive list of generated features: <multi> signifies any of the identifiers defined in Equation (2), V is the CC discharge voltage measured at time t .

Feature	Description
min-ccv-<multi>	minimum CC discharge voltage
max-ccv-<multi>	maximum CC discharge voltage
mean-ccv-<multi>	mean of CC discharge voltage
var-ccv-<multi>	variance CC discharge voltage
skew-ccv-<multi>	skewness of CC discharge voltage
kurt-ccv-<multi>	kurtosis CC discharge voltage
area-ccv-<multi>	area under the CC discharge voltage curve
grad-ccv-start-<multi>	dV/dt at start time
grad-ccv-end-<multi>	dV/dt at final time
grad-ccv-min-<multi>	minimum value of dV/dt
grad-ccv-max-<multi>	maximum value of dV/dt

3.2. Machine learning model

Various machine learning models were explored in this research: Decision Trees [27], Random Forest [28], Support Vector Regression [29], Extremely Randomized Trees [30], and Extreme Gradient Boosting (XGBoost) [31]. However, out of all these algorithms, XGBoost was able to reduce overall model variance and over-fitting, capture non-linear relationships in the data, and had the best performance (both in sample and out of sample) in terms of the metrics considered. For the purpose of brevity and the fact that this study is not meant for the comparison of model performances, only the results of the XGBoost model are presented and discussed.

Main model is XGBoost. It is a non-linear ensemble model which builds regression trees (called weak learners) sequentially and the variance of the previous tree is corrected in the new tree to finally build a strong learner [31]. The sci-kit learn [32] implementation was used.

Targets to predict. We used the approach discussed in [9, Section 3] to obtain a list of targets for our machine learning model. In short, each capacity curve is summarised by certain cogent points: knee-onset (k-o) and knee-point (k-p) [8]; the same is applied to IR curves: initial point, elbow-onset (e-o) and elbow-point (e-p) [6]. Knees, Elbows and their onsets are identified via the Bacon-Watts model [8, 6]. In addition, capacity at k-o (Qatk-o) and k-p (Qatk-p); IR at e-o (IRate-o), e-p (IRate-p) and EOL (IRatEOL). Capacity at EOL is known (taken as 80%) and not included in the list of targets to be predicted. See Table 2 for a summary.

Two block models, ‘cycle-at’ and ‘value-at’, were trained (see summary

Table 2). Firstly, we trained an XGBoost model to predict knees, elbows, and EOL jointly, which we call ‘cycle-at’ model. This combination is natural as the values of these targets are larger in comparison to those of capacity/IR; all are measured in cycles and are correlated. Secondly, a model to jointly predict capacity and IR at knees, elbows, and EOL values was trained on the log-transformed values of these targets and we call it the ‘value-at’ model. The log transformation was considered due to the small values of capacity (ranging between 1.1 to 0.88Ah) and IR (ranging between approximately 0.015 to 0.03Ω).

Table 2: Description and nomenclature for the machine learning models.

Model class type	Prediction targets
cycle-at	k-o, k-p, e-o, e-p, EOL
value-at	Qatk-o, Qatk-p, IRate-o, IRate-p, IRatEOL

Model performance is measured using the mean absolute error (MAE) and root mean squared error (RMSE), see metrics in Equation (3). We present the values of these metrics for each of the models both for training and test dataset.

Training and tuning. Choosing hyper-parameters for each of the machine learning models involved two techniques: (1) trial and error approach, (2) grid search method. The former was done to search for feasible parameters values while the latter considers a range of values around those obtained in the first stage. The second stage was performed using the GridSearchCV object¹ from the scikit-learn Python library. An exhaustive search was not performed for all the parameters in each of the considered models, however, we selected those that most effect the model performance during the course of trial and error approach. The algorithm makes use of squared-error loss function. It was found that n -estimator, max-depth, reg-alpha, learning-rate, and min-samples-split were the most important to tune for model improvement. The description of these hyperparameters can be found in the XGBoost documentation².

¹GridSearchCV is an exhaustive search mechanism which takes a range of values of parameters (parameter space), a method of searching, cross-validation scheme, and a scoring function. It returns the parameter setting which best improves the metric used.

²<https://xgboost.readthedocs.io/en/stable/parameter.html>

3.3. Choosing cycle number threshold – How much longitudinal data?

To see the effects of using different cycle thresholds on the prediction of each of the targets, we split errors based on whether they are associated to capacity or IR. Maximum threshold of 100 cycles was considered and we generated features by extracting the first n cycles, $10 \leq n \leq 100$, where n is an integer. For each n , we built two different XGBoost models as discussed in Section 3.2 and evaluated them through cross-validation on the training set using the metrics of Equation (3). Figure 2 presents the cross-validation errors on training data together with their 90% confidence intervals. By inspecting this figure, we note that model performance is not sensitive to the number of cycles selected. Thus, inline with [8], we selected 50 cycles as the input size for our model.

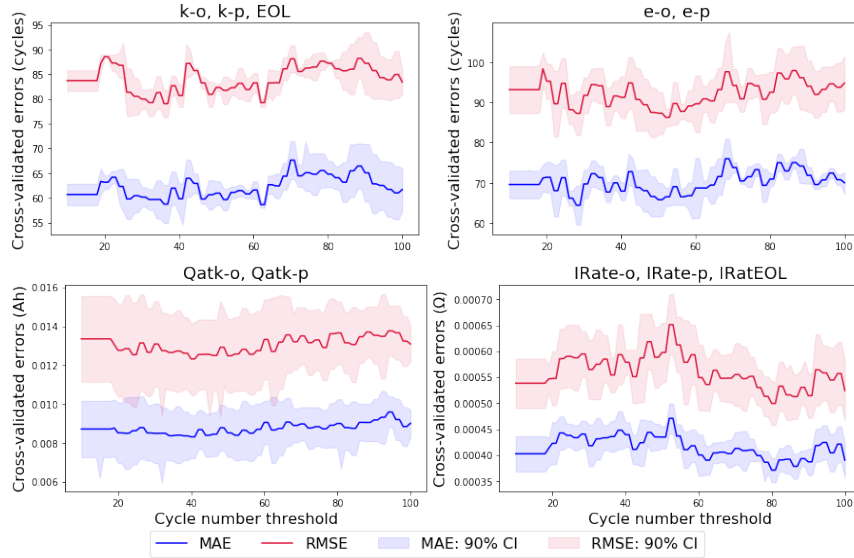


Figure 2: Cross-validated errors on training data showing the impact of using different cycle number threshold for feature generation and modelling.

3.4. Prediction of the entire curves

We recover the Capacity and IR degradation curves by employing a methodology similar to that of [9]. Concretely, after obtaining the knees, elbows and EOL together with the corresponding capacity and IR from the predictions of the model (Table 2), we fitted a modified quadratic spline to these points to get the entire capacity fade and IR curves – see Equations (4)-(5) in Methods section below.

4. Results and Discussion

In this section, we present the results of our findings together with the rationalizations behind them. It covers the presentation of model performance results, and an account of feature importance analysis. Recall Section 3.2 for nomenclature.

4.1. Model performance

The performance of the cycle-at model on training and test data is given in Table 3. To account for a comprehensive performance, we closely look at the errors in predicting each of the targets by providing performance metrics and their corresponding confidence intervals. With respect to training metrics, the model is able to capture the underlying pattern with a best MAE and RMSE (in the case of k-o and k-p prediction) of 10 and 13 cycles respectively. As for the performance on the test set, similar pattern is also observed: it gives the lowest MAE and RMSE of 57 and 78 cycles respectively for k-o prediction, whereas highest values are seen for e-o (with MAE and RMSE of 81 and 103 cycles, respectively). These results show that the model predicts knees more accurately than elbows. This is most likely linked to the nature of the process of obtaining knees and elbows: k-o and k-p are extracted from the capacity fade curve with less noise in comparison to the IR rise curve from where e-o and e-p are extracted; the noisiness in IR measurements of [12] has been thoroughly discussed in [24]. In addition, the fact that we obtain the smallest test errors on k-o prediction can be associated to the closeness of this point to the cycle from which we are predicting from (which is 50 cycles).

The summary of the performance of value-at model is shown in Table 4. In terms of capacity at knees prediction, we note that the model achieves approximately the same MAE and RMSE on the training set for the prediction of both k-o and k-p. However, different out-of-sample performance is observed. As for the prediction of IR at elbows and EOL, the model yields approximately similar MAE and RMSE on the training set for all the IRs. On the other hand, the best performance on the hold-out data is attained for the prediction of e-o and e-p with MAE and RMSE of approximately $4.13 \times 10^{-4}\Omega$ and $5.77 \times 10^{-4}\Omega$, respectively. The similarities in errors on the training set in both capacity and IR estimations can be associated to the positions of knees/elbows on the capacity fade/IR curve: k-o and k-p as well as e-o and e-p are points with few cycles away from each other. Thus, their corresponding capacity/IR values will not be differed significantly.

The parity plots of the predicted versus observed values are presented in Figure 3. In both train and test data, most of the points are remarkably close to the parity line. The embedded histogram also shows that most residuals/errors are distributed around zero which indicates a reliable performance of the models.

Table 3: Performance metrics and their corresponding 95% confidence intervals of cycle-at model in predicting k-o, k-p, e-o, e-p and EOL.

	MAE (cycles)		RMSE (cycles)	
	Train	Test	Train	Test
k-o	10 ± 1.5	57 ± 15.4	13 ± 2.0	78 ± 23.4
k-p	10 ± 1.5	69 ± 20.2	13 ± 1.9	99 ± 34.9
e-o	12 ± 1.8	81 ± 18.5	15 ± 2.2	103 ± 20.7
e-p	12 ± 1.9	68 ± 16.7	16 ± 2.4	89 ± 20.5
EOL	11 ± 1.8	71 ± 21.0	15 ± 2.3	101 ± 36.5

Table 4: Performance metrics together with 95% confidence intervals of value-at model in predicting Qatk-o, Qatk-p, IRate-o, IRate-p and IRatEOL.

	MAE		RMSE	
	Train ($\times 10^{-4}$)	Test ($\times 10^{-3}$)	Train ($\times 10^{-4}$)	Test ($\times 10^{-3}$)
Qatk-o	7.79 ± 1.30	5.373 ± 1.20	10.50 ± 2.10	6.778 ± 1.60
Qatk-p	7.79 ± 1.30	4.590 ± 1.10	10.52 ± 2.20	6.002 ± 1.20
IRate-o	0.15 ± 0.02	0.413 ± 0.12	0.20 ± 0.035	0.577 ± 0.17
IRate-p	0.14 ± 0.025	0.414 ± 0.11	0.19 ± 0.035	0.570 ± 0.17
IRatEOL	0.15 ± 0.025	0.440 ± 0.13	0.20 ± 0.030	0.624 ± 0.21

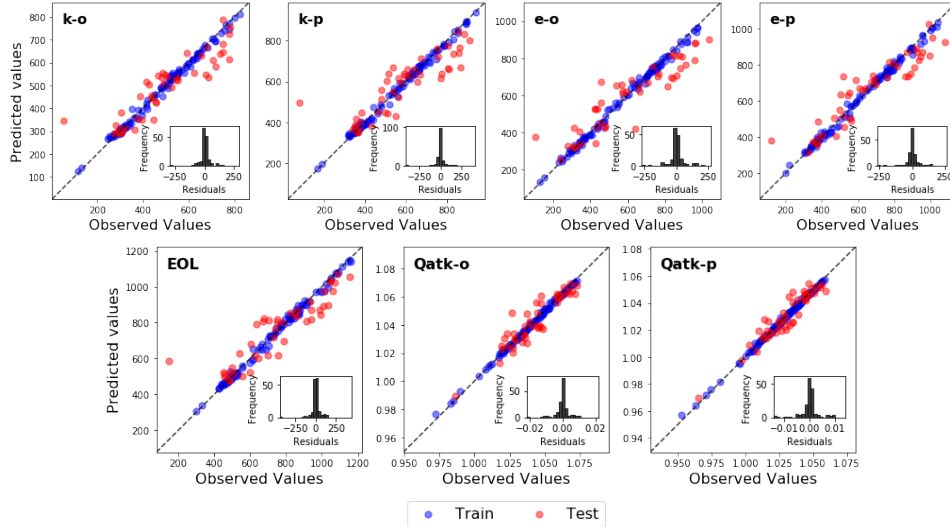
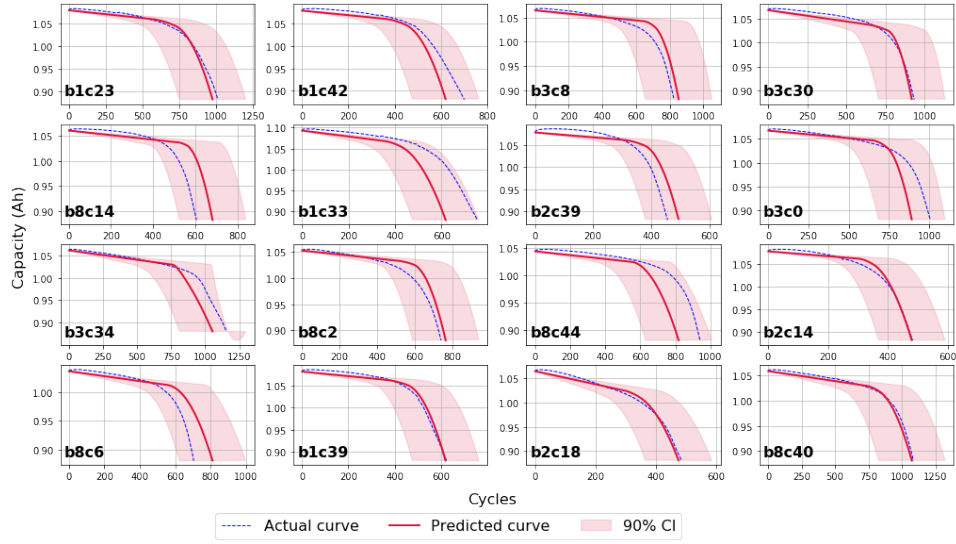


Figure 3: Parity plots of the predicted values from the cycle-at and value-at models versus observed values. Histograms of corresponding residuals embedded in each plot.

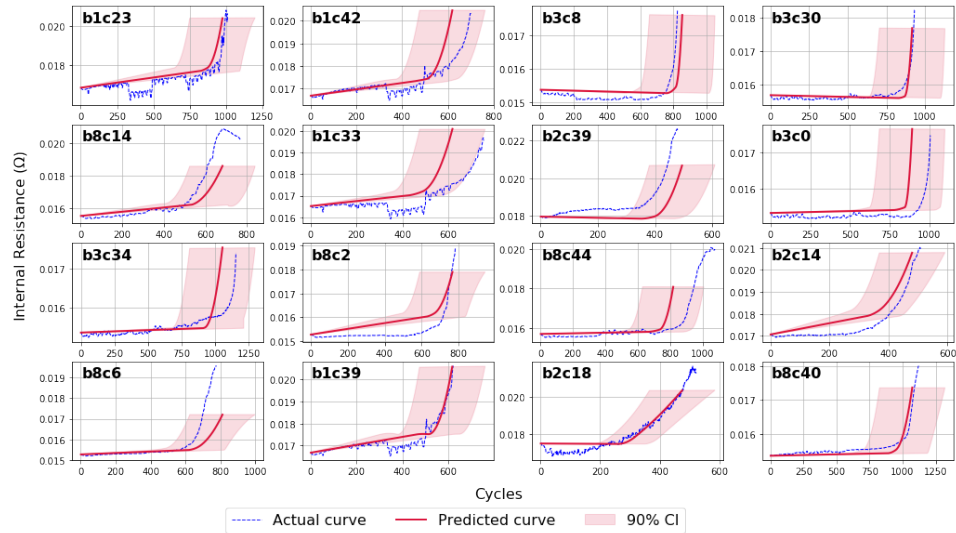
4.2. Entire curve prediction

After obtaining the knees/elbows and their corresponding capacity and IR values, we used the modified quadratic spline described in Section 3.4 to fit these values. The results of this procedure, applied to random cells in test set, together with 90% confidence interval are shown in Figure 4. For each of the curves, the associated confidence interval was obtained by calculating prediction intervals for the test set. These intervals were calculated by training and predicting with multiple independent copies of models: each of these models gave different predictions due to stochastic nature of the XGBoost algorithm. Subsequent to this, we apply the spline on the predicted intervals for k-o/e-o, k-p/e-p and EOL. It is observed that the method performs better in predicting capacity fade curves than IR curves. The reason for this is not far-fetched. The IR values in the utilised dataset are very noisy and thus any model using these values is likely to struggle in capturing the underlying pattern (see, e.g., b1c33 and b2c39 in Fig 4b); However, the predicted IR curves do qualitatively capture the observed trends. We also note an underestimation of IRatEOL for the IR curves of cells in batch 8. In addition to the measured IR data, we make use of the predicted IR values in [24] for cells in batch 8. This will introduce an inherent uncertainty/variance in their prediction into our model and additionally illustrates differences from predicting from real data (that of [12]) and syn-

thetic data (that of [6]) – we emphasize that without such synthetic data no quantification of IR prediction would be possible for the [20] dataset.



(a)



(b)

Figure 4: Predicted (a) capacity fade and (b) IR curves for random selected test cells with confidence interval.

4.3. Feature relevance

We now report on the feature importance through correlation and model results as means to discuss the most relevant features out of the generated predictor models.

We, firstly, obtain the Pearson correlation (ρ) between the features and targets; results are given in Figure 5. We aim at looking at features which have a linear relationship with target values. Most features are found to have high correlation with targets except in few cases such as min-ccv-f0, min-ccv-f25 and max-ccv-fdiff. Looking in more detail at the correlations, features corresponding to the mean of the CC discharge voltage are seen to have a strong positive correlation with knee-onset/point, elbow-onset/point and EOL (with $\rho > 0.5$), while strong negative correlations are observed for IRs at e-o, e-p and EOL (with $\rho < -0.5$). The correlations observed here reflect practical scenarios. It is expected that cells which decay slower (and thus have a later knee-onset/point, elbow-onset/point and EOL) would have a higher average voltage during discharge and thus a positive correlation. On the other hand, cells that decay quickly will have a faster increase in IR, and hence, a negative correlation with the average voltage. This discovery is also seen in the correlations of all features belonging to kurtosis; few extracted from area under the curve (area-ccv-f50-0 and area-ccv-fdiff); and some corresponding to gradients (grad-ccv-start-f0/f25/f50, grad-ccv-min-diff and grad-ccv-max-diff). However, the reverse is the case for all features measuring the variance and skewness of the curve, grad-ccv-end-f25, and grad-ccv-min-f25.

Secondly, we look at the feature importance in each of the proposed XGBoost models: cycle-at and value-at models. Detailed information about how feature importance is computed by the XGBoost algorithm can be found in [31]. We provide the plot of this scaled importance (with least and highest importance as 0 and 1 respectively) in Figure 6. With respect to the cycle-at model, feature measuring the difference between the variance of the CC discharge voltage at the initial cycle and cycle 50 (var-ccv-f50-0) is found to be used most frequently in tree construction with highest importance value across all the predicted targets. This feature does not only capture the dynamics of the voltage curve at the measured points, but also compare their variability at the two extreme points of prediction (at first and 50th cycle). This rationalization is responsible for this remarkable performance. Looking at the literature that makes use of this idea of variance to measure variability, a variance model using only the logarithm of the variance of

$\Delta Q_{100-10}(V)$ ³ was built in [12] for EOL prediction. Said model achieves an MAPE of about 15% and 11% on their primary and secondary test sets, respectively.

As for the value-at model, the features' relative importance is spread across a larger number of features. Features with low relative importance might capture different information or pattern in the data, which does not have a direct impact on the targets in consideration. Looking at the prediction of Qatk-o and Qatk-p, the area under the CC discharge voltage curve is found to have the highest feature importance. This can be linked to the relationship between this feature and capacity. The time integral of discharge voltage is proportional to the energy delivered by the battery since the current is kept constant over the discharge process. This energy is in turn influenced by the capacity of the battery: the energy produced by a battery is controlled by the amount of electricity generated as a result of electrochemical reactions in the battery. On the other hand, the associated with the rate of change of the CC discharge voltage with time is discovered to be the most relevant in predicting IRs at the elbows and EOL when compared to other features in the model. This is most likely attached to the inverse relationship between this predictor and IR. As IR increases, the battery temperature rises and voltage drops thus a negative rate. The reverse is the case when the IR falls with time.

³ $\Delta Q_{100-10}(V) = Q_{100}(V) - Q_{10}(V)$, where $Q_i(V)$ is the value of the interpolated discharge capacity (Q)-voltage (V) curve corresponding to cycle i .

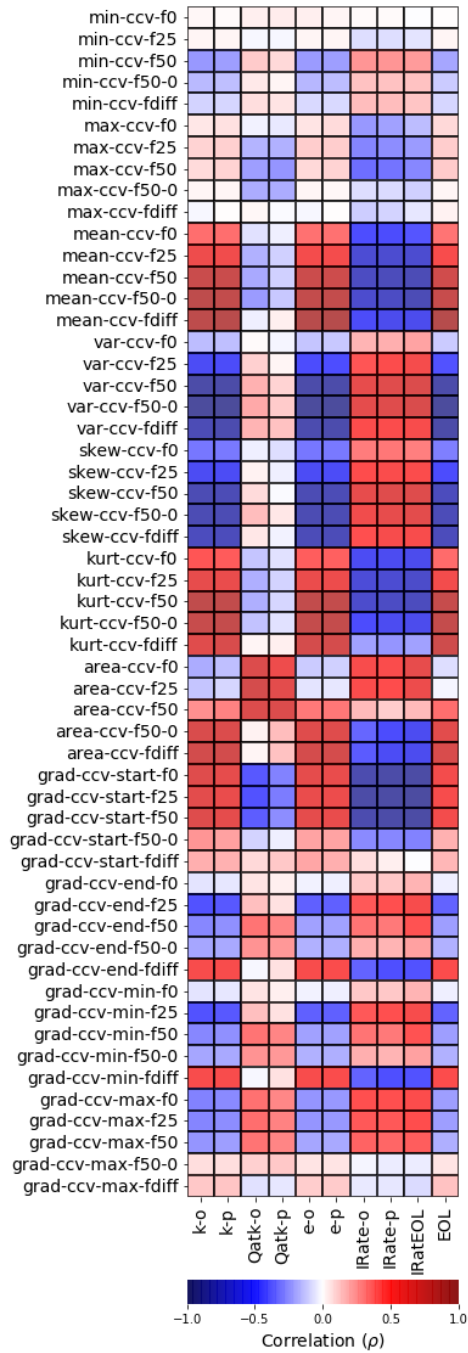


Figure 5: Pearson correlation between all generated features and targets.

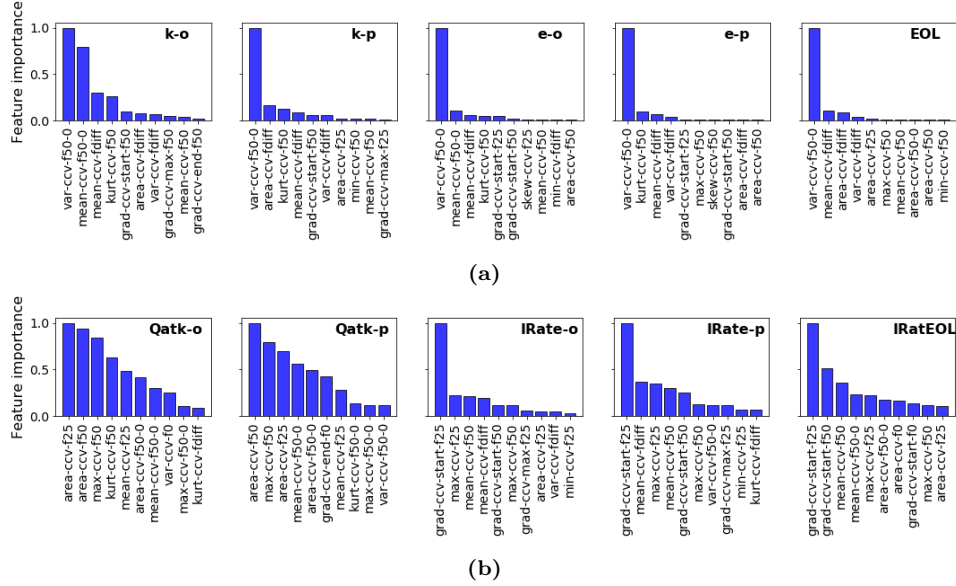


Figure 6: Feature importance bar charts of the first 10 most important features (as metrised by the XGBoost algorithm [31]) in the case of (a) cycle-at model and (b) value-at model.

4.4. Sub-sampling of CC discharging voltage

For the dataset at hand and illustrated in Figure 1, the frequency of data recording is approximately 4 seconds and all models developed so far made full use of the data. In this section we investigate how much sub-sampling in the CC discharge voltage response curve can be done without a significant loss of quality in the models. We focus on a single model predicting the capacity EOL cycle number of the cells and evaluate it through cross-validation. Figure 7 shows, as expected, errors to be increasing as the time steps rise (i.e., sub-sampling in wider intervals). The errors are more pronounced when a time step of more than 1 minute is used for sub-sampling. A direct implication of this strategy is that CC discharge voltage does not need to be measured in a small-time interval to obtain a good enough model, and we note that it is sufficient to take voltage response measurement at the end of every 1 minute. This reduces cost of data acquisition, storage and computational time during the modelling process.

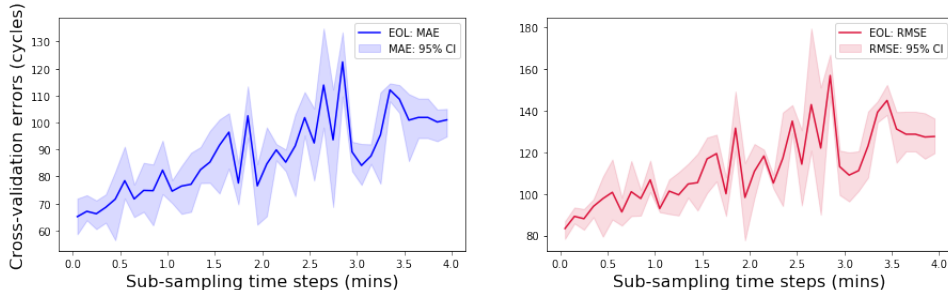


Figure 7: Cross-validation results using different time steps for CC discharge voltage sub-sampling. The 95% confidence intervals on the mean errors were calculated via bootstrapping. It is observed that the model yielded an MAE and RMSE of 65 and 83 cycles respectively at the minimum time step (4 seconds) whereas errors are 87 and 109 cycles respectively when the CC discharge voltage curves are sub-sampled at every 1 minute. This indicates that the model does not lose significant quality at 1 minute time step.

4.5. Comparison with the literature

In this section, we compare our approach to other methods found in the literature making use of the same datasets [12] and feature-based technique to predict EOL. To bring all the approaches to a common ground as much as possible, we only use batches 1, 2, and 3 (i.e., [12]) for this comparison as they were used in the referenced work. We present the summary of this comparative analysis in Table 5.

Severson et al. [12] worked on data-driven prediction of EOL for capacity degradation. The authors generated features and built three linear regression models which they called variance, discharge and full model. A key note about their study is that they made use of the first 100 cycles of the data. The variance model, which used logarithm of the variance of $\Delta Q_{100-10}(V)$ for EOL prediction, achieved a MAPE of about 15% and 11% on the primary and secondary test sets respectively. The discharge model, in addition to the previous features, used predictors obtained from voltage and current readings during the first 100 cycles, and attains an MAPE of 10.1% and 8.6% on the primary and secondary test data respectively. Their full model, using a combination of features from the previous two and features derived from temperature and IR, achieved 7.5% and 10.7% MAPE on the primary and secondary test data respectively.

In a paper closely related to that of [12], Saxena et al. [17] used the first 100 cycles of the discharge voltage-capacity curves as inputs to a CNN model to predict the entire capacity fade curve via an exponential curve parameterisation (no IR is predicted). On testing the accuracy of their model, a mean absolute percentage error (MAPE) of approximately 22% was recorded

in predicting the EOL of the cells involved. Although it should be emphasized that their model was not trained to predict the EOL directly.

A broad learning-extreme learning machine was proposed by Ma et al. [33] to predict the capacity and EOL using the first 100 cycles of data for feature extraction. They recorded an MAPE of 9% on the test data.

Another approach similar to [12] is the work of Shen et al. [34]: they made use of the Relevance Vector Machine to generate synthetic data for cells characterized by longer cycle-lives. Subsequent to this, this synthetic data together with the actual data were used to build a feature-based convolutional neural network (CNN) using the gradient of the interpolated capacity-voltage curve $Q(V)$ corresponding to the first 100 cycles to predict EOL. An average MAPE of 11.7% was reported in their paper.

In the same vein, the EOL of the considered cells was predicted in [35] using the Gradient Boosting Regression Trees algorithm trained on features extracted from different sources namely: entire voltage, capacity, and temperature. The model evaluation yielded a mean average percentage error of 7%.

Contrary to all the aforementioned strategies which used the entire curves of a specific battery chemistry, our approach makes use of 50 cycles of only CC voltage response corresponding to the discharge part of the curve. Generated features were fed into an extreme gradient boosting regression algorithm which resulted in a mean absolute percentage error of 12% on the test data.

Table 5: Comparison of the performance of our model in predicting EOL. For comparison purpose, we restrict model building to only cells in batches 1, 2, and 3. Papers are ordered as an increasing function of number of cycles used and whether they use full or part of the measured data; data used are measured voltage (V), discharge capacity (Q), current (I), temperature (T), IR (IR), discharge capacity-voltage curve ($Q(V)$), state-of-health (SOH) and charge time (ct). ‘Full at discharge’ means that the corresponding curves are obtained using the whole measurements when energy is being derived from the cells; whereas ‘CC at discharge’ depicts the data is derived only from the CC component of the discharge curve. Metrics in parenthesis correspond to the secondary test scores from the paper.

Papers	Data used	Cycles used	Nature of curve	MAPE (%)
This work	V	50	CC at discharge	12.0
[17]	$Q(V)$	100	Full at discharge	~ 22.0
Variance [12]	$Q(V)$	100	Full at discharge	15.0 (11.0)
Discharge [12]	$Q(V), V, I$	100	Full at discharge	10.1 (8.6)
Full [12]	$Q(V), V, I, T, IR$	100	Full at discharge	7.5 (10.7)
[33]	SOH, $Q(V)$, IR , ct	100	Full at discharge	9.0
[34]	$Q(V)$	100	Full at discharge	11.7
[35]	SOH, $Q(V)$, V, T	100	Full at discharge	7.0

5. Conclusion

The prediction of degradation capacity fade and IR rise curves from the CC voltage response during the discharge phase has been addressed. The model developed provides a stark benefit over previous approaches that require information of full charge/discharge (or both) cycles. Many real-world applications do not allow for consistent charge/discharge cycles and thus being able to predict from just one of the (charge or discharge) phases (or portions of it) in isolation provides great value.

In the context of a reduction in data dimension, our method uses only CC discharge voltage corresponding over the first 50 cycles of data with competitive accuracy. We found that minutewise recording of the voltage response suffices for prediction, contrasting with the literature reviewed in this study that requires all the available data from charging/discharging (or both) and with high-frequency data recording. These model characteristics are especially valuable for practical deployment where only limited data can be accessed. Our findings also show how much (or how little) information is needed for cheap, informative and accurate early-life prediction.

On the access to various remaining useful life quantifiers, our strategy leverages the accurate prediction of knees, elbows and EOL. This indicates the efficiency and versatility of our approach as one obtains a full trajectory

curve prediction and various battery life prognostics. Regarding feature importance analysis, we strike a balance between model complexity and interpretability as we can analyse the contribution of each of the features generated to the overall model performance. This is extremely useful from the industrial application point of view as we find that a minimal number of features can capture the main behaviours of the degradation profile.

In terms of future study, our approach could be extended to consider CC charging situations, as it typical for electric vehicles.

Methods

Machine Learning Performance

The metrics considered in this study are *mean absolute error* (MAE), *mean absolute percentage error* (MAPE) and *root mean squared error* (RMSE). They are defined as follows:

$$\left\{ \begin{array}{l} \text{MAE}(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \\ \text{MAPE}(y, \hat{y}) = \frac{100\%}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \\ \text{RMSE}(y, \hat{y}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \end{array} \right. \quad (3)$$

where y_i, \hat{y}_i are the actual and predicted values for sample i respectively, and n is the number of samples.

Modified Quadratic Spline

The idea behind the construction of this spline is to have three different polynomials defined in three different domains which have the knees, elbows and EOL as their end-points. The first of these polynomials is a straight line which joins the start of the curve to the knee/elbow onset. While the second (a quadratic curve) connects the knee/elbow onset to the knee/elbow point, the last polynomial (also quadratic) completes the curve estimation by joining the knee/elbow point to the EOL. Concretely, this spline is defined in terms of polynomials P_0 , P_1 , and P_3 as:

$$\mathcal{S}_2 := \begin{cases} P_0(n) = a_0 + b_0n + c_0n^2, & n_0 \leq n < n_1; \\ P_1(n) = a_1 + b_1n + c_1n^2, & n_1 \leq n < n_2; \\ P_2(n) = a_2 + b_2n + c_2n^2, & n_2 \leq n \leq n_3; \end{cases} \quad (4)$$

constrained by

$$\left\{ \begin{array}{l} P_0(n_i) = f_i, \quad i = 0, 1; \\ P_1(n_i) = f_i, \quad i = 1, 2; \\ P_2(n_i) = f_i, \quad i = 2, 3; \\ P'_0(n_1) - P'_1(n_1) = 0; \\ P'_1(n_2) - P'_2(n_2) = 0; \\ c_0 = 0. \end{array} \right. \quad (5)$$

Here, n depicts cycles and a_i, b_i, c_i are all real numbers. The points $n_i, i = 0, \dots, 3$ correspond to the initial cycle number, k-o/e-o, k-p/e-p and EOL respectively. Similarly $f_i, i = 0, \dots, 3$ are the initial capacity/IR at these points: Qatk-o/IRatk-o, Qatk-p/IRate-p and QatEOL/IRatEOL respectively. The first three constraints of Equation (5) ensure that the spline coincide with the actual values f_i while derivative constraints ensure the continuity of the spline at the interior points. The last constraint fixes the first polynomial P_0 to be a straight line and is also needed to obtain values of all the coefficients. To obtain the entire spline, we make the assumption that the initial cycle and the corresponding capacity/IR values are known. In addition, \mathcal{S}_2 is only defined in the interval $[n_0, n_3]$. This is not rigid as the EOL (which corresponds to n_3) can be redefined other than cycle number at 80% of initial capacity. The linear equations generated for each of the cells are solved computationally using the NumPy solver [26].

References

- [1] B. Dunn, H. Kamath, J.-M. Tarascon, Electrical energy storage for the grid: a battery of choices, *Science* (2011) 928–935.
- [2] B. Nykvist, M. Nilsson, Rapidly falling costs of battery packs for electric vehicles, *Nat. Clim. Change* (5) (2015) 329–332.
- [3] R. Schmich, R. Wagner, G. Hörpel, T. Placke, M. Winter, Performance and cost of materials for lithium-based rechargeable automotive batteries, *Nat. Energy* (3) (2018) 267–278.
- [4] C. R. Birkl, M. R. Roberts, E. McTurk, P. G. Bruce, D. A. Howey, Degradation diagnostics for lithium ion cells, *Journal of Power Sources* 341 (2017) 373–386.
- [5] P. M. Attia, A. Bills, F. B. Planella, P. Dechent, G. dos Reis, M. Dubarry, P. Gasper, R. Gilchrist, S. Greenbank, D. Howey, O. Liu,

- E. Khoo, Y. Preger, A. Soni, S. Sripad, A. G. Stefanopoulou, V. Sulzer, Review—“Knees” in lithium-ion battery aging trajectories, *Journal of The Electrochemical Society* 169 (6) (2022) 060517.
- [6] C. Strange, S. Li, R. Gilchrist, G. Dos Reis, Elbows of internal resistance rise curves in li-ion cells, *Energies* 14 (4) (2021) 1206.
- [7] W. Diao, J. Kim, M. Azarian, M. Pecht, [Degradation modes and mechanisms analysis of lithium-ion batteries with knee points](#), *Electrochimica Acta* 431 (2022) 141143. doi:<https://doi.org/10.1016/j.electacta.2022.141143>.
URL <https://www.sciencedirect.com/science/article/pii/S0013468622013007>
- [8] P. Fermín-Cueto, E. McTurk, M. Allerhand, E. Medina-Lopez, M. F. Anjos, J. Sylvester, G. dos Reis, Identification and machine learning prediction of knee-point and knee-onset in capacity degradation curves of lithium-ion cells, *Energy and AI* 1 (2020) 100006.
- [9] C. Strange, G. dos Reis, Prediction of future capacity and internal resistance of li-ion cells from one cycle of input data, *Energy and AI* 5 (2021) 100097.
- [10] S. Sohn, H.-E. Byun, J. H. Lee, [Cnn-based online diagnosis of Knee-point in Li-ion battery capacity fade curve](#), *IFAC-PapersOnLine* 55 (7) (2022) 181–185, 13th IFAC Symposium on Dynamics and Control of Process Systems, including Biosystems DYCOPS 2022. doi:<https://doi.org/10.1016/j.ifacol.2022.07.441>.
URL <https://www.sciencedirect.com/science/article/pii/S2405896322008424>
- [11] G. Dos Reis, C. Strange, M. Yadav, S. Li, Lithium-ion battery data and where to find it, *Energy and AI* 5 (2021) 100081.
- [12] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggedakis, M. Z. Bazant, S. J. Harris, W. C. Chueh, R. D. Braatz, Data-driven prediction of battery cycle life before capacity degradation, *Nature Energy* 4 (5) (2019) 383–391.
- [13] D. Roman, S. Saxena, V. Robu, M. Pecht, D. Flynn, Machine learning pipeline for battery state-of-health estimation, *Nature Machine Intelligence* 3 (5) (2021) 447–456.

- [14] S. Greenbank, D. A. Howey, Automated feature selection for data-driven models of rapid battery capacity fade and end of life, arXiv e-prints (2021) arXiv–2101.
- [15] N. H. Paulson, J. Kubal, L. Ward, S. Saxena, W. Lu, S. J. Babinec, Feature engineering for machine learning enabled early prediction of battery lifetime, *Journal of Power Sources* 527 (2022) 231127.
- [16] W. Li, N. Sengupta, P. Dechent, D. Howey, A. Annaswamy, D. U. Sauer, [One-shot battery degradation trajectory prediction with deep learning](#), *Journal of Power Sources* 506 (2021) 230024. doi:<https://doi.org/10.1016/j.jpowsour.2021.230024>. URL <https://www.sciencedirect.com/science/article/pii/S0378775321005528>
- [17] S. Saxena, L. Ward, J. Kubal, W. Lu, S. Babinec, N. Paulson, [A convolutional neural network model for battery capacity fade curve prediction using early life data](#), *Journal of Power Sources* 542 (2022) 231736. doi:<https://doi.org/10.1016/j.jpowsour.2022.231736>. URL <https://www.sciencedirect.com/science/article/pii/S0378775322007303>
- [18] P. Li, Z. Zhang, R. Grosu, Z. Deng, J. Hou, Y. Rong, R. Wu, An end-to-end neural network framework for state-of-health estimation and remaining useful life prediction of electric vehicle lithium batteries, *Renewable and Sustainable Energy Reviews* 156 (2022) 111843.
- [19] A. Aitio, D. A. Howey, Predicting battery end of life from solar off-grid system field data using machine learning, *Joule* 5 (12) (2021) 3204–3220.
- [20] P. M. Attia, A. Grover, N. Jin, K. A. Severson, T. M. Markov, Y.-H. Liao, M. H. Chen, B. Cheong, N. Perkins, Z. Yang, P. K. Herring, M. Aykol, S. J. Harris, R. D. Braatz, S. Ermon, W. C. Chueh, Closed-loop optimization of fast-charging protocols for batteries with machine learning, *Nature* 578 (2020) 397–402.
- [21] M. Park, M. Seo, Y. Song, S. W. Kim, Capacity estimation of li-ion batteries using constant current charging voltage with multilayer perceptron, *IEEE Access* 8 (2020) 180762–180772.
- [22] Y. Duan, J. Tian, J. Lu, C. Wang, W. Shen, R. Xiong, Deep neural network battery impedance spectra prediction by only using constant-current curve, *Energy Storage Materials* 41 (2021) 24–31.

- [23] N. Yang, T. Yu, Q. Luo, K. Wang, Fast and accurate health assessment of lithium-ion batteries based on typical voltage segments., *Front. Energy Res.* 10 (2022) 925947.
- [24] C. Strange, S. Li, R. Gilchrist, G. dos Reis, Elbows of internal resistance rise curves in li-ion cells, *Energies* 14 (2021) 1206.
- [25] P. Dechent, S. Greenbank, F. Hildenbrand, S. Jbabdi, D. U. Sauer, D. A. Howey, Estimation of li-ion degradation test sample sizes required to understand cell-to-cell variability, *Batteries & Supercaps* 4 (12) (2021) 1821–1829.
- [26] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, T. E. Oliphant, Array programming with NumPy, *Nature* 585 (2020) 357–362.
- [27] J. R. Quinlan, Induction of decision trees, *Machine Learning* 1 (2006) 81–106.
- [28] L. Breiman, Random forests, *Machine Learning* 45 (2004) 5–32.
- [29] A. J. Smola, B. Schölkopf, A tutorial on support vector regression, *Statistics and Computing* 14 (2004) 199–222.
- [30] P. Geurts, D. Ernst, L. Wehenkel, Extremely randomized trees, *Machine Learning* 63 (2006) 3–42.
- [31] T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, in: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*, ACM, New York, NY, USA, 2016, pp. 785–794.
- [32] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research* 12 (2011) 2825–2830.
- [33] Y. Ma, L. Wu, Y. Guan, Z. Peng, The capacity estimation and cycle life prediction of lithium-ion batteries using a new broad extreme learning machine approach, *Journal of Power Sources* 476 (2020) 228581.

- [34] S. Shen, V. Nemani, J. Liu, C. Hu, Z. Wang, A hybrid machine learning model for battery cycle life prediction with early cycle data, in: 2020 IEEE Transportation Electrification Conference & Expo (ITEC), 2020, pp. 181–184.
- [35] F. Yang, D. Wang, F. Xu, Z. Huang, K.-L. Tsui, Lifespan prediction of lithium-ion batteries based on various extracted features and gradient boosting regression tree model, *Journal of Power Sources* 476 (2020) 228654.

Funding

This project was funded by an industry-academia collaborative grant *EPSRC EP/R511687/1* awarded by *EPSRC & University of Edinburgh* program *Impact Acceleration Account (IAA)*.

R. Ibraheem is a PhD student in EPSRC’s MAC-MIGS Centre for Doctoral Training. MAC-MIGS is supported by the UK’s Engineering and Physical Science Research Council (grant number EP/S023291/1).

G. dos Reis acknowledges partial support from the *Fundação para a Ciência e a Tecnologia* (Portuguese Foundation for Science and Technology) through the project UIDB/00297/2020 and UIDP/00297/2020 (Center for Mathematics and Applications, CMA/FCT/UNL). G. dos Reis acknowledges support the Faraday Institution [grant number FIRG049]

Author contributions

All authors provided domain expertise, edited and reviewed the manuscript.

R.I.: methodology, software, visualisation; data curation, visualisation; writing, review and editing.

C.S.: conceptualisation, visualisation; writing, review and editing.

G.d.R.: conceptualisation, supervision, funding acquisition; writing, review and editing.

Competing interest declaration

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Additional information

This work has no supplementary information file.