Prediction of future capacity and internal resistance of Li-ion cells from one cycle of input data

Energy and AI 5 (2021) 100097

Contents lists available at ScienceDirect

Energy and AI

journal homepage: www.elsevier.com/locate/egyai

Prediction of future capacity and internal resistance of Li-ion cells from one cycle of input data

Calum Strange [a], Gonçalo dos Reis [a,b,*]

[a] School of Mathematics, University of Edinburgh, The King's Buildings, Edinburgh, EH9 3FD, UK
[b] Centro de Matemática e Aplicações (CMA), FCT, UNL, Quinta da Torre, 2829-516 Caparica, Portugal

ARTICLE INFO

Keywords:
Capacity degradation
Internal resistance degradation
Prediction of full degradation curve
Knee and elbow-points
Lithium-ion cells
Machine learning
Remaining useful life

ABSTRACT

There is a large demand for models able to predict the future capacity retention and internal resistance (IR) of Lithium-ion battery cells with as little testing as possible. We provide a data-centric model accurately predicting a cell's entire capacity and IR trajectory from one single cycle of input data. This represents a significant reduction in the amount of input data needed over previous works. Our approach characterises the capacity and IR curve through a small number of key points, which, once predicted and interpolated, describe the full curve. With this approach the remaining useful life is predicted with an 8.6% mean absolute percentage error when the input-cycle is within the first 100 cycles.

1. Introduction

Sales of electric vehicles and energy storage systems are undergoing a marked growth as battery costs continue to fall and governments around the world introduce increasingly strict emissions regulations.

Of importance to all applications is a cell's state-of-health (SOH). In many applications the key metric for cell health is capacity retention. In this regard, SOH is often interpreted as the current capacity of a cell as a percentage of its rated capacity. As the capacity degrades over time so does the cell's usefulness eventually reaching a point at which the cell is no longer deemed useful for its current application. This point, called the end-of-life (EOL), is often a predefined capacity level. Another key health indicator is the internal resistance (IR) of the cell: as the cell degrades its IR increases, impairing the cell's ability to provide and receive charge. Capacity degradation and IR rise of a Li-ion cell are often not linear throughout its lifetime [1,2]. Cell capacity typically starts to degrade in a linear manner until reaching a critical point, called the 'knee' (referred to henceforth as the knee-point), at which the rate of capacity degradation increases considerably [3–5]. In [1] the additional variable 'knee-onset' is introduced (along with an alternative identification mechanism) to provide a useful indication of the start of this rate of increased degradation. Building on this idea, [2] introduced the variables 'elbow-onset' and 'elbow-point' describing the same phenomena as the knee-onset and -point but for the IR rise curve. Accurate identification and prediction of the occurrence of knee-onset and -point can provide essential guidance for scheduling of replacements and cell maintenance to prolong service life. However, knee-points (and knee-onset) may appear before or after the EOL is reached and their occurrences are also cell chemistry dependent [6]. The same holds for elbow-onsets and points. Other degradation metrics, such as the remaining useful life (RUL) or the whole capacity trajectory, thus have to be considered collaboratively for a comprehensive view.

Much research has been dedicated to the modelling of Li-ion cells and, in particular, lifetime prediction such as EOL and RUL. Broadly, there are two approaches to this problem either *model-based* or *data-centric*. The model-based approach encapsulates empirical models, Equivalent circuit models and physics-based models. It includes electrochemical type models where the cell's internal physical degradation mechanisms are simulated (see [7] for a review), and parametric/semi-parametric type models where empirical models are fitted to realised capacity fade curves and combined with advanced filtering techniques to predict future degradation [8,9]. The data-centric approach consists of machine learning and statistical models trained on in-cycle and cycle-to-cycle measurement data such as voltage, current, capacity, temperature and internal resistance. Feature based approaches allow for expert input on essential features [10–13] but may also take a purely data-driven feature selection approach. Feature free approaches use deep learning techniques such as Convolutional neural networks (CNN) to process 'raw' cycle data. The data-centric approach typically requires larger data sets for training than model-based approaches, nonetheless, this approach is showing great potential [14–16]. Physics-informed models need to be calibrated to the cell's data (a non-trivial

* Corresponding author at: School of Mathematics, University of Edinburgh, The King's Buildings, Edinburgh, EH9 3FD, UK.
E-mail address: G.dosReis@ed.ac.uk (G. dos Reis).

https://doi.org/10.1016/j.egyai.2021.100097
Received 14 May 2021; Received in revised form 18 June 2021; Accepted 19 June 2021
Available online 26 June 2021

problem) and if too simplified lose the inherent conservation laws. The computational requirements of semi-parametric type modelling are much lower than those of electrochemical-models or the data-centric approach. For both data-centric and semi-parametric modelling, the available data for training/calibration is a methodological limitation in itself, be it on the quality of fitting, extrapolation or simply the method one can use.

Data-driven research has mainly focused on the prediction of EOL or RUL [1,10] in contrast the literature on the prediction of the complete capacity trajectory is sparse. We point notably to [17] who make use of a simple feed forward neural network to enhance the *slope and bias correction* model migration technique. At its base, this approach uses a parametric model for the capacity fade curve then a neural network is trained to migrate this fitted model from the first 30% (50–150 cycles depending on cell) of data into a prediction at a given future cycle. With this approach, they are able to accurately describe the full capacity fade curve of their 4 test cells.

Our study aims to determine the smallest number of cycles needed to accurately predict the whole capacity/IR trajectory of a cell. We found that the information contained in any one cycle of charge/discharge data was sufficient. This speaks directly to cell manufacturers who need to grade batteries and buyers of cells performing quality control.

This aim stems from several gaps in the current literature. Firstly, we address a gap that electrochemical/physics-based modelling is yet to close. Concretely, the prediction of the lifespan of a cell from a single cycle of input data. Take for instance the well-known Doyle–Fuller–Newman (DFN) model for lithium-ion batteries and its variants (see [18] for a review). Parameterising a DFN model from cycling data is impossible (much less from one single cycle) without many material assumptions: one would need stoichiometries of the two electrodes which cannot be obtained from cycle data [19]; one would need to dissemble the cell to carry specific measurements which may take around 3 months [20,21]. Without disassembling the cell, one needs to rely on current–voltage response, for which case many of the parameters are not well identifiable — work on sensitivity analysis and optimal excitation for parameter identification can be found in [22]. Secondly, to the best of our knowledge, the machine learning prediction models developed so far require gradient information for prediction. This implies longitudinal data spanning a large number of cycles, e.g., 50 to 100 cycles to predict EOL are needed [1,10] usually involving a feature generation step. In terms of the amount of input data, the current best art for quantitative early prediction of RUL is using only 4 cycles of data achieving a 10.6% *mean absolute percentage error* (MAPE) [23], and this result marks a non-trivial improvement over earlier work. However, reducing this number further would represent a further reduction in testing times and costs. Thirdly, the majority of the literature deals solely with the prediction of EOL or RUL. RUL is a key indicator of cell health and the EOL of cell quality but neither is a complete picture: both fail to capture non-linear dynamics in the capacity fade trajectory. And lastly, the vast majority of work focuses on capacity retention and ignores questions on IR degradation, both are important SOH indicators and neither is a complete picture on their own. In fact, the 80%-capacity level for EOL is an industry postulation while the knee/elbow-point (and knee/elbow-onset) [1,2] reflects better traceable physics/eletrochemical causal changes. There is thus space for new approaches to predict the full capacity and IR trajectory, and to do so from a reduced number of measurements.

The main contribution of this work addresses the above four limitations from the data-driven modelling point of view. The model proposed uses a CNN which jointly predicts, from a single cycle of data, the full capacity fade trajectory and the full IR rise trajectory.

The rest of this study is organised as follows. Section 2 describes this study's datasets and Section 3 contains a full description of the proposed modelling approaches and insights leading to it. An account of the model's performance is given in Section 4 including a comparison with existing art: methods and approaches used, presented results, used features, mode of feature selection and the number of cycles used for prediction. Section 5 concludes this work.

2. Data description

We work with the datasets of [10] and [24] found at https://data.matr.io/1; detailed descriptions provided there. The data consists of high-throughput cycling data for 8 batches of commercial lithium iron phosphate (LFP)/graphite cells cycled under fast-charging conditions: [10] provides data for 3 batches of approximately 48 cells each (referred to as batches 1 to 3); [24] provides data for 5 batches, of between 45 and 48 cells each (referred to as batches 4 to 8). *Cell code Notation*: across the 8 batches of cells, $bXcY$ refers to cell Y of batch X.

All cells in batches 1, 2, 3 and 8 are cycled close to or past their EOL, defined as 80% of initial capacity, in a temperature controlled environment ($T = 30°$) with a variety of charge/discharge profiles. It is important to note that for each individual cell, its charge/discharge profile was kept constant from cycle to cycle. Batches 4–7 were only run for 100–120 cycles and thus do not reach an EOL. The dataset contains in-cycle measurements of temperature, current, charge and discharge capacity, as well as per-cycle measurements of capacity, internal resistance and charge time. Batch 8 does not include IR data, for these cells we use the predicted IR data of [2] available at https://doi.org/10.7488/ds/2957.

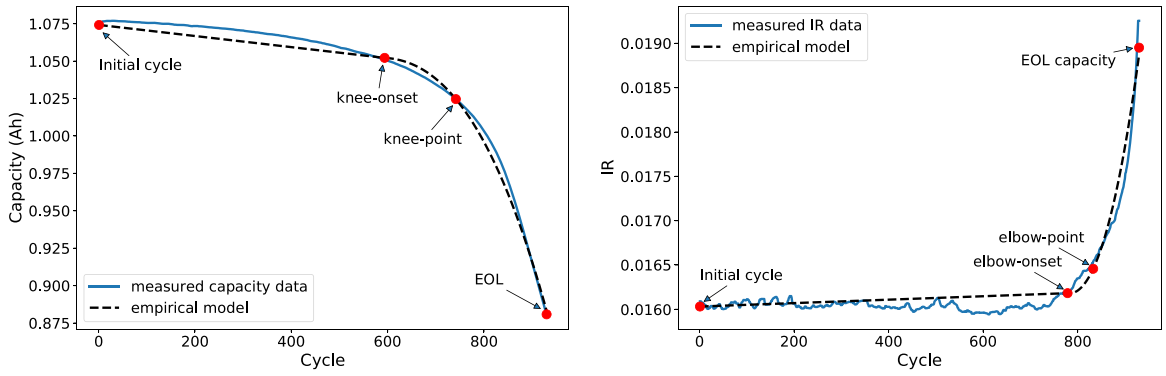
The datasets of [10] and [24] both consist of the same type of LFP/graphite cells cycled in a consistent experimental setting with data logged in the same format. Combination of the two is thus natural. For comparison with previous works, we present results both with and without the inclusion of the data from [24]. A fuller description of these datasets can be found by consulting the relevant papers or in the recent review paper [25, Section 2.1.3].

3. Predicting future capacity and internal resistance

A question present in all battery applications is: what does the future degradation of a particular cell look like, when will a cell no longer be suitable for its current application and at what speed will this degradation occur? When solely considering cycle-ageing, ideally one would know the future capacity and internal resistance of a cell any number of cycles into the future, up to (and perhaps beyond) the EOL.

Following on from previous work [1,2], we describe the capacity degradation by use of the knee-onset, knee-point and EOL, and the IR rise curve by the elbow-onset and elbow-point. Fermin et al. [1] proposed the use of the Bacon–Watts and double Bacon–Watts model to identify the cycle at which the knee-point and knee-onset occur, respectively. Strange et al. [2] proposed an additional smoothing process prior to deploying the Bacon–Watts models — this process involves fitting an empirical line-plus-exponential model to an isotonic regression of the data. The linear relationships between these points are also explored in the cited papers. Here, we use the second approach (with smoothing). Additionally, as we are interested in describing the full curves, we must select the capacity and IR values at the knees and elbows. Since the recorded data (in particular the IR) is noisy, we take these capacity/IR values from the smoothed (line-plus-exponential) curves and not the raw data.

We now propose a simple empirical model with which we can describe the full capacity and IR curves. In addition to the knees (for the capacity) and elbows (for the IR) we need the first and last points of both curves. As described previously, data was recorded until the cells reached 80% of their nominal capacity (~ 0.88Ah). So, we describe the capacity curve by four points: the current cycle (measured capacity), knee-onset (empirical capacity), knee-point (empirical capacity) and EOL (0.88Ah). And, we describe the IR curve by the current cycle (measured IR), elbow-onset (empirical IR), elbow-point (empirical IR) and capacity-EOL (empirical IR). Our proposed approach (assuming the current cycle is sampled ahead of the knee/elbow-onset and -point) is as follows: (1) fit a cubic spline between the four points; (2) take the cubic spline as the approximation between the last three points; (3)



(a) capacity **(b)** internal resistance

Fig. 1. Empirical model fitted to the capacity and IR curves of cell b3c12.

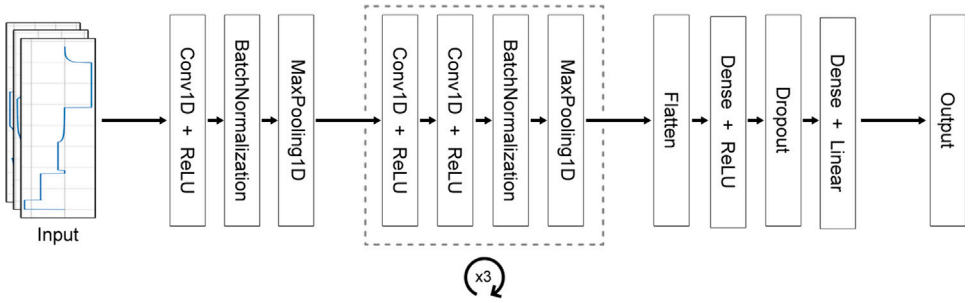


Fig. 2. Representation of the CNN architecture. The '×3' notation denotes three repeated blocks with the displayed configuration.

approximate between the first two points by a straight line. Applied to the measured capacity fade curves of batches 1, 2, 3 and 8 (up to EOL), and taking cycles 1 to 100 as the 'current cycle', this model obtains a root mean square error (RMSE) of 0.0039 and a coefficient of determination (R^2) value of 0.9931, and applied to the measured IR curves it obtains a RMSE of 0.00015 and an R^2 value of 0.9838 showing a strong agreement with the measured values. Up to the onset points the degradation is linear, and thus the straight line approximation (step 3) performs well. Comparing after the knee-onset the model obtains a RMSE of 0.0046 and an R^2 value of 0.9902, and restricted to after the elbow-onset the model obtains a RMSE of 0.00017 and an R^2 value of 0.9832. So, our simple interpolation accurately describes the true curves. An example of the approximated capacity and IR curves is given in Fig. 1.

Then, in order to predict the entire capacity fade and IR rise curves, it is enough to have a measurement (or prediction) of the current capacity/IR value and predictions of the number of cycles until (and the remaining capacity/IR values at) the knee-onset, knee-point, EOL, elbow-onset and elbow-point. To illustrate these quantities, we predict the '*time to knee-onset*' (ttk-o), '*time to knee-point*' (ttk-p), RUL, '*time to elbow-onset*' (tte-o) and '*time to elbow-point*' (tte-p). In addition, we predict the remaining capacities at knee-onset (Q@k-o) and knee-point (Q@k-p), and the IR values at elbow-onset (IR@e-o), elbow-point (IR@e-p) and the IR at (capacity) EOL (IR@EOL). The retained capacity at EOL is known and thus does not need prediction. The empirical model described above can then be fitted through the predicted points giving a prediction of the entire capacity degradation and IR rise curves. It should be pointed out that the characteristic points we select are stylistic in nature and that not all cells display knees or elbows. However, the authors believe that the basic idea of identifying and predicting key points before fitting an empirical model should be applicable to a wide range of ageing modes. Different ageing modes may require a different stylised model (this is left to future research).

We restrict most of our discussion to 'early' prediction, here defined as the first 100 cycles of data (*initial* setting). This is a more difficult setting than prediction at later points (*full* setting) and allows for a direct comparison with previous works in the literature. Indeed, as expected, our model performs better as the cycle from which predictions are made approaches the actual cycle of the predicted quantity. For illustrative purposes, our model's performance predicting the RUL versus the distance from the EOL is presented in Fig. 5.

3.1. Modelling approach

Ideally, the testing required to make a prediction should be limited in time, making prediction fast and convenient. Thus, we restrict to the prediction from a single cycle of data. This removes the need for past knowledge of a cell, a problem faced in so-called 'second-life' applications, where the historical cycling profiles of most cells are unknown. Of course, (if present) it would not be difficult to incorporate this knowledge into our approach. In addition, this restriction massively multiplies the available data for training. For each cell (when restricting to the first n cycles of data) we have n cycles of data and n corresponding distinct values for each item we are predicting. Predicting from one cycle of data, we thus have n examples from each train/test cell. This multiplication of the training set allows us to use deep learning techniques.

We will now describe our proposed model. We take a data-driven and feature-free approach. We propose a model consisting of a convolutional 'feature extraction' block followed by two densely connected layers, displayed in Fig. 2 and described in Table 1. As output, this model can be trained to predict values jointly (*joint prediction*) or separately. As input our model takes a single cycle of voltage, current and SOC data (obtained by the coulomb counting method, from one full cycle). Our model was implemented in Python using TensorFlow via the Keras API [26]. All layer names given in Fig. 2 refer to the

Table 1
Proposed architecture of the CNN model for the prediction of ttk-o, Q@k-o, ttk-p, Q@k-p, RUL, tte-o, IR@e-o, tte-p, IR@e-p and IR@EOL. The model can be trained to predict multiple points at once (joint prediction) or separately. Hyper-parameters are given in the format: filters, kernel size, activation for conv1d layers; pool size for max_pooling; dropout for dropout; nodes, activation for dense layers.

Layer name	Input size	Hyper-parameters	Output size
conv1d_1	926 × 3	24, 6, *ReLU*	921 × 24
batch_normalisation_1	921 × 24	–	921 × 24
max_pooling_1	921 × 24	2	460 × 24
conv1d_2	460 × 24	32, 3, *ReLU*	458 × 32
conv1d_3	458 × 32	32, 3, *ReLU*	456 × 32
batch_normalisation_2	456 × 32	–	456 × 32
max_pooling_2	456 × 32	2	228 × 32
conv1d_4	228 × 32	32, 3, *ReLU*	226 × 64
conv1d_5	226 × 64	32, 3, *ReLU*	224 × 64
batch_normalisation_3	224 × 64	–	224 × 64
max_pooling_3	224 × 64	2	112 × 64
conv1d_6	112 × 64	32, 3, *ReLU*	110 × 64
conv1d_7	110 × 64	32, 3, *ReLU*	108 × 64
batch_normalisation_4	108 × 64	–	108 × 64
max_pooling_4	108 × 64	2	54 × 64
flatten_1	54 × 64	–	3456
dense_1	3456	64, *ReLU*	64
dropout_1	64	0.4	64
dense_2	64	*n_outputs*, *linear*	*n_outputs*

corresponding Keras layers. To assist training, batch normalisation was used before each MaxPooling1D layer. The model was trained using the *Adam* optimiser for 100 epochs with a batch size of 512. And the *mean absolute error* was set as the loss function. A *learning rate scheduler* (described in Eq. (1)) was used during training with a 'decay rate' of 0.9 and a 'decay step' of 5 epochs.

Since the in-cycle data was not recorded at consistent time intervals or for the same number of time-steps (after cleaning) the data was interpolated and *nan*-extended to a consistent length and time-step. The interpolation was performed with the *SciPy* [27] function *interp1d* and interpolated to one measurement every four seconds. The data was then allocated at random (by cell) into an 80–20 train-test split. The training and testing sets were then restricted to the first 100 cycles of data before being standard scaled (when the model was trained in the full setting, the restriction step was dropped). Hyper parameters were optimised prior to testing using randomly selected validation subsets of cells from the training set.

When using the model to predict future capacity (Q@k-o and Q@k-p) and IR (IR@e-o, IR@e-p and IR@EOL) the model was trained to predict the *loss in capacity* and *rise in IR*, respectively. Given a measurement (or prediction) of current capacity and current IR these can easily be converted into predictions of future capacity and IR. The loss in capacity from the start to EOL is of the order of 0.1. Similarly, the loss in capacity to knee-onset and knee-point is quantitatively small. The IR rises are even smaller, of order 0.001. Small target values can mean small values in a model's loss function which can negatively impact training. Thus to improve performance, our model was trained to predict 10000×(loss in capacity) and 2000000×(rise in IR). These multiplications were accounted for when converting to a prediction of future capacity and IR. The multiplicative constants selected here are largely arbitrary (chosen to roughly match the range of RUL values) and their exact specification did not significantly impact model performance.

3.2. Prediction intervals via the forward-dropout method

Here we briefly describe the approach taken to provide prediction intervals. A simple approach is to train and predict with multiple independent copies of a model calculating prediction intervals from the independent predictions. Here by 'independent' we mean models trained separately: due to the stochastic nature of the training each trained model will provide different predictions. This approach is often referred to as an 'ensemble' approach, and this is the approach taken to produce the performance metrics displayed in this paper. However, there are notable issues which may make such an approach unattractive or unfeasible. Firstly, there is the computational cost and time associated with training a model repeatedly and independently. And secondly, there is the cost of storing multiple models in memory — which poses a particular barrier in storage limited applications such as integrated chipsets.

Another approach, superior to the ensemble approach in both aspects described (although not necessarily in terms of accuracy), is to deploy *dropout* during the forward pass of a network (forward dropout). That is, predicting with a trained model multiple times each time with a random dropout (of a pre-specified rate) applied and calculating prediction intervals from these predictions. This approach can be viewed as a Bayesian approximation of a Gaussian process [28]. The rate at which dropout is applied during prediction is optimised such that the distribution of 'residuals' from dropout prediction to the median dropout prediction matches the distribution of residuals from the model without dropout's prediction to the true value. This optimisation can be performed prior to deployment (on a validation set) or 'on the fly' after deployment as predictions are made and then compared with realised results. This is our preferred approach when using the model to predict the full capacity fade and IR rise curves. When this approach was applied to our model the dropout layer present in Table 1 and Fig. 2 was used to apply dropout in training with our selected training dropout and then in prediction with a separately optimised prediction dropout rate.

4. Model performance

4.1. Performance metrics

We now present our model's performance metrics when predicting each quantity in isolation. The figures for the prediction of ttk-o, ttk-p, RUL, tte-o and tte-p are presented in Table 2. For the capacity related predictions, the cycle error (MAE and RMSE) is lower for the points which are temporarily closer to the cycle from which the prediction is made: the knee-onset and knee-point. However, in percentage terms, the model performs better predicting the RUL than the ttk-o or ttk-p. This is explained by the larger target value and thus smaller percentage error for a given cycle error. The larger percentage error for the knee-onset prediction is explained by the smaller target value. The prediction of the IR related quantities is quantitatively worse than the capacity related ones; this is explained by the (much) noisier IR measurements which in turn effect the elbow identification [2, page 8]. A more granular view of the errors can be found in Fig. 4b (for the RUL) and Fig. 3.

The results of comparing the predicted Q@k-o, Q@k-p, IR@e-o, IR@e-p and IR@EOL with the values from the empirical fitted curve are presented in Table 3 where it is shown that our model can accurately predict these values from a single cycle of data.

4.2. Full curve prediction

We now inspect the performance of the model to predict the full capacity and IR curves. For this we trained three models to predict (jointly) related quantities. These three models were a 'time to' model (predicting ttk-o, ttk-p, RUL, tte-o and tte-p), a 'capacity' model (Q@k-o and Q@k-p) and an 'IR' model (IR@e-o, IR@e-p and IR@EOL). In this way we recover the performance metrics of individual prediction and avoid issues such as the knee-point being predicted before the knee-onset. For each of these models forward dropout rates for each of their outputs were optimised by training and testing on subsets of the training data. For the 'time to' model the selected dropout rates were 0.45, 0.325, 0.35, 0.35 and 0.30 for the ttk-o, ttk-p, RUL, tte-o

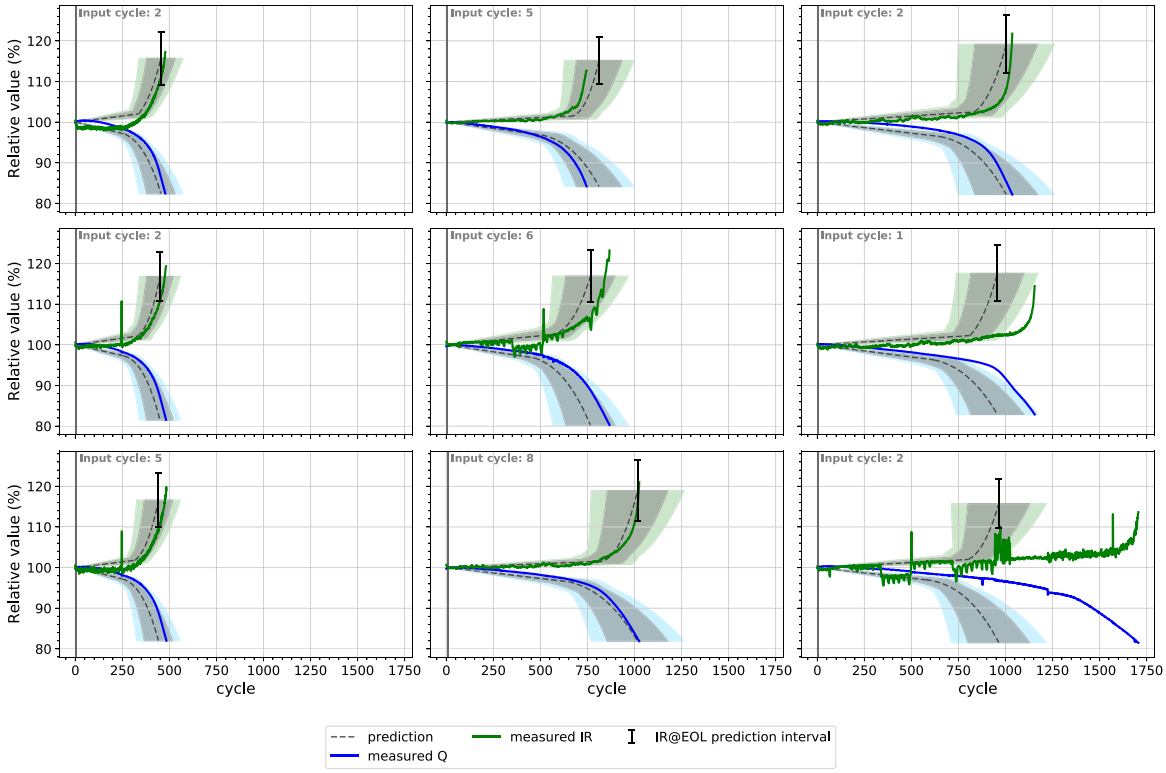


Fig. 3. Plots of model predictions for nine randomly chosen test cells at randomly selected input cycles. Model predictions are produced from data of a single cycle, given the measured capacity/IR at that cycle. The prediction intervals (95% and 80%) are calculated under a normality assumption from 100 predictions with forward dropout applied. Moving from left to right, and then down a row repeating, the 'test cell' displayed is b2c30, b8c20, b3c9, b2c14, b1c7, b3c34, b2c45, b3c26 and b1c4. Here, we present the IR and capacity values relative to the values measured at the input cycle.

Table 2
Performance of proposed model to predict Capacity's: ttk-o, ttk-p and RUL; and IR's: tte-o and tte-p.

	RMSE (cycles)		MAE (cycles)		MAPE (%)	
	Train	Test	Train	Test	Train	Test
ttk-o	38 ± 5.0	84 ± 12.0	21 ± 2.7	55 ± 6.5	4.9 ± 0.72	12.6 ± 1.44
ttk-p	41 ± 4.8	83 ± 14.4	26 ± 3.4	55 ± 6.1	4.2 ± 0.48	9.7 ± 0.94
RUL	50 ± 4.7	100 ± 19.3	32 ± 3.1	66 ± 7.8	3.8 ± 0.33	8.6 ± 0.95
tte-o	51 ± 3.7	112 ± 23.8	30 ± 2.6	71 ± 10.2	4.6 ± 0.37	11.9 ± 1.87
tte-p	50 ± 4.5	105 ± 28.3	31 ± 3.1	68 ± 11.7	4.2 ± 0.40	10.1 ± 1.54

Table 3
Performance of model to predict future capacity and IR, when current capacity/IR is known.

	RMSE		MAE		MAPE (%)	
	Train	Test	Train	Test	Train	Test
Q@k-o	0.0018 ± 2.4e−4	0.0082 ± 6.7e−4	0.0013 ± 1.9e−4	0.0041 ± 2.6e−4	0.13 ± 0.02	0.40 ± 0.03
Q@k-p	0.0022 ± 2.8e−4	0.0075 ± 4.6e−4	0.0017 ± 2.5e−4	0.0040 ± 2.1e−4	0.17 ± 0.02	0.41 ± 0.02
IR@e-o	5.4e−5 ± 6.2e−6	0.00019 ± 2.1e−5	3.3e−5 ± 3.8e−6	0.00014 ± 1.6e−5	0.20 ± 0.02	0.84 ± 0.10
IR@e-p	6.7e−5 ± 6.7e−6	0.00021 ± 2.5e−5	4.4e−5 ± 5.0e−6	0.00015 ± 2.1e−5	0.26 ± 0.03	0.85 ± 0.12
IR@EOL	0.00020 ± 4.9e−5	0.00041 ± 5.5e−5	0.00014 ± 2.0e−5	0.00032 ± 4.0e−5	0.72 ± 0.10	1.71 ± 0.21

and tte-p, respectively; for the 'capacity' model 0.3 and 0.15, Q@k-e and Q@k-p; and, for the 'IR' model 0.75, 0.7 and 0.5, IR@e-o, IR@e-p and IR@EOL. Final models were then trained on the full training set and multiple predictions made with the selected dropout rates. Example plots of the predictions produced by this model are presented in Fig. 3, where we present plots for 9 randomly chosen cells from the 35 test cells at random cycles from between cycle 1 and 10. The full curve prediction intervals presented in this plot were calculated by fitting our empirical model (Fig. 1(b)) to the prediction intervals calculated from the dropout predictions of knee/elbow -onsets and points, and the EOL. We do not address the impact of measurement noise on our input data. A simple approach to address this issue would be to allow a normal distribution around the measured capacity/IR with variance calibrated to the training data, or to take capacity values from several cycles as input.

It is clear from Fig. 3 and Fig. 4b that, on average, our model performs worse on cells with longer cycle lives and Fig. 4c shows the reason why: a significantly lower amount of training data is available for cells with an EOL above 1200 cycles.

A related problem to the prediction of RUL is the classification of cells by expected lifetime. For example, manufacturers may wish to select only the best performing cells to place in a battery pack. In the context of the Severson dataset we point to [10] who report a 4.9% test error classifying cells by 'long-lived' (EOL > 550 cycles) and 'short lived' (EOL < 550 cycles) from 5 cycles of data, and [1]

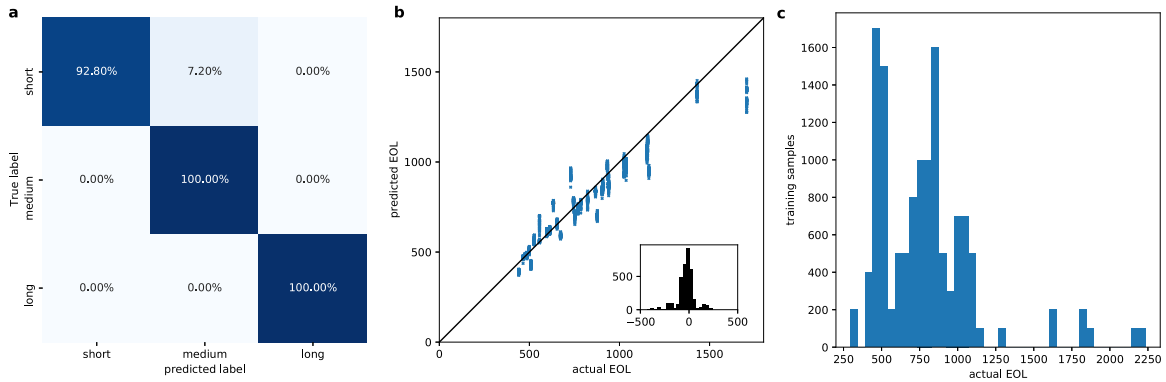


Fig. 4. **a.** Confusion matrix for prediction of EOL classes 'short life' (EOL < 550), 'medium life' (550 < EOL < 1200) and 'long life' (EOL > 1200), based on quantitative prediction for test set. **b.** Actual EOL vs predicted EOL with distribution of errors, showing a slight bias to under-predict EOL especially for longer lived cells. **c.** Distribution of training data for EOL showing a significantly lower amount of training data for long lived cells.

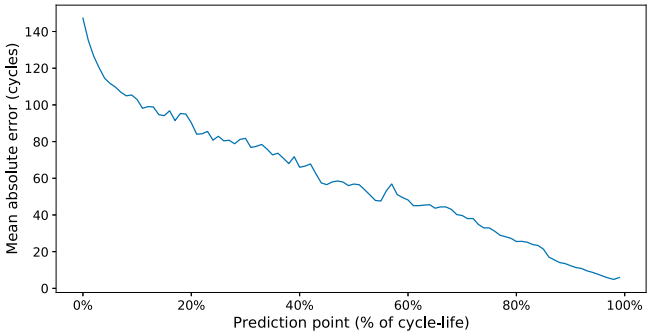


Fig. 5. Plot of model's performance to predict RUL in the full setting, showing a mostly linear improvement in performance as the prediction point approaches the EOL.

who achieve an accuracy of 88% classifying the batteries into 'short' (knee-point < 500 cycles), 'medium' (knee-point between 500–1100 cycles) and 'long-range' (knee-point > 1100 cycles) from 5 cycles of data. For comparison with these results, Fig. 4a displays a confusion matrix obtained from converting our model's cycle predictions into three classes 'short' (EOL < 550 cycles), 'medium' (550-1200 cycles) and 'long' lived (> 1200 cycles). We see that our model achieves a comparable level of accuracy while performing the classification from a single cycle of data. This shows that, while not performing well on long-lived cells (in the regression problem), the prediction is competitive for the classification problem. The barrier of 550 matches that used in [10], and the barrier of 1200 is in line with that used in [1] (as the EOL occurs somewhat after the knee-point).

In Fig. 5 we note that the model performance improves in a largely linear manner as the cycle from which the prediction was made approaches the EOL.

4.3. Comparison with prior art

We report a comparative study of our work to existing art utilising the Severson dataset. A review of works presenting models for prediction of EOL and RUL utilising different datasets can be found in the introduction (and references therein). Model performance may vary due to different employed datasets. In this regards, we can only meaningfully compare our results to these in terms of methodology. However, we choose to provide a quantitative comparison against literature drawing on the same dataset analysed in this work. In particular, we discuss [10,23,29–31] to review methods and approaches used, results presented, features selected, mode of feature selection and the number of cycles required for prediction. A summary of this comparison can be found in Table 4. We separate EOL and RUL prediction as these problems are quantitatively distinct.

Prediction of RUL (summarised in Table 4 (a))

Hong et al. [23] propose a Dilated CNN to predict the RUL. Their approach is feature free. As input their model takes 4 cycles of voltage, current and temperature data. They introduce two settings: the 'initial' setting where they restrict to data from the first 100 cycles, and the 'full' setting where they restrict to data before EOL. They obtain a reported MAPE of 10.6% and 19.7% in the initial and full settings, respectively.

In order to compare our model's performance with that of [23] in Table 5 we present our model's performance in the initial and full setting. Where we obtain an MAPE of 9.6% and 12.8%, respectively. Outperforming the model of [23] in both settings using fewer cycles of data. We emphasise the methodological difference here to [23]: their use of 4 cycles is explicitly to capture inter-cycle cross-data correlations and temporal patterns; our results need only 1 cycle which does not contain any type of gradient information.

Prediction of EOL (summarised in Table 4 (b))

Severson et al. [10] predict the EOL using a Regularised Linear Model trained on features extracted from the first 100 cycles of data. They emphasise the importance of including voltage data in their regression models, in particular capacity as a function of voltage (Q(V)); they propose three candidate models to predict EOL all utilising features extracted from the gradients of voltage discharge curves. The best performing of these models (utilising the most features) obtains a reported MAPE of 9.1%

Ma et al. [29] propose a new 'Broad Learning-Extreme Learning Machine'. This model is tested to predict the capacity and EOL on three data sets. For the Severson dataset they present results only for the prediction of EOL. Like Severson et al., their model takes as input features extracted from the first 100 cycles of data. With this model they obtain a reported MAPE of 9%.

Shen et al. [30] predict the EOL using a Relevance Vector Machine (RVM) to enhance the dataset by generating 'artificial cells' with long cycle-lives. The enhanced dataset is used then to train a CNN. As input the CNN takes the same gradients of Q(V) as Severson et al., thus we consider it a 'feature based' CNN. Evaluating their model on a primary and secondary test set, they report an average MAPE of 11.7%.

In contrast to the approaches described above we take a feature extraction free approach, utilising a convolutional neural network to learn the 'optimal' features. As input our model takes a single cycle of voltage and current data, thus our model sees no gradient information. The performance of our model to predict the EOL and RUL (when restricted to batches 1, 2 and 3) is presented in Table 5.

Prediction of the IR rise curve

To the authors' knowledge the only other work predicting the IR rise curve for the Severson dataset is [2], where a RVM was used to predict the elbow-onset and -point from the first 50 cycles of data. For

Table 4
Comparison of results from works using the Severson dataset. For comparison purposes, reported results exclude data from batch 8. Ordered by number of cycles used and reported MAPE. Inputs listed are in-cycle measurements of voltage (V), current (I), temperature (T), capacity as a function of voltage (Q(V)); and cycle-to-cycle measurements of capacity (SOH), internal resistance (IR) and time to charge (ttc).

Paper	Inputs	Cycles used	MAPE	What is predicted
This work	V, I	**1**	**8.8** %	EOL
			9.6 %	RUL
[23]	V, I, T	4	10.6%	RUL
[29]	SOH, Q(V), IR, ttc	100	9.0%	EOL
[10]	SOH, Q(V), IR, ttc, T	100	9.1%	
[30]	Q(V)	100	11.7%	
[31]	SOH, Q(V), V, T	250	7.0%	

Table 5
Performance of proposed model to predict EOL and RUL when restricted to batches 1, 2 and 3. We report performance both in the initial setting (input cycle < 100 cycles) and the full setting.

	RMSE (cycles)		MAE (cycles)		MAPE (%)	
	Train	Test	Train	Test	Train	Test
EOL	55.0 ± 5.8	110 ± 24.4	33 ± 3.4	73 ± 12.4	3.5 ± 0.35	8.8 ± 1.43
RUL (initial)	55.0 ± 5.8	110 ± 24.4	33 ± 3.4	73 ± 12.4	3.7 ± 0.43	9.6 ± 1.47
RUL (full)	38 ± 2.7	99 ± 34.8	23 ± 2.5	59 ± 12.6	5.3 ± 0.47	12.8 ± 1.26

the prediction of elbow-onset they achieved a MAPE of 14.0% and a MAE of 91.3, and for the elbow-point a MAPE of 11.5% and a MAE of 83.4. We have improved on this previous work in terms of accuracy and number of input cycles.

Prediction of the entire capacity fade curve

To the authors' knowledge the only other work predicting the entire capacity fade curve for the Severson dataset can be found in Herring et al. [32]. Presenting a python library for the prognosis and cycle life prediction of Li-ion cells. As an example of their libraries performance, they predict the evolution of cell capacity for the Severson dataset. They train a multi-task linear model to predict the number of cycles until a cell reaches a range of SOH levels. This model takes as input the features in [10] covering 100 cycles of data. No performance metrics were provided.

5. Conclusions

The prediction of future capacity loss and IR rise is a problem of great importance. Current capacity prediction algorithms demand input data across many tens of full charge–discharge cycles to work and IR rise prediction has received little attention in the literature. In our framework, the remaining useful life and the entire capacity/IR trajectory are accurately predicted from a single input data cycle. This reduction entails a significant increase in prognostics procedures' affordability through reduced testing times, and stands to benefit academics and industry.

Differentiating from existing methods, we use key quantities as a dimension reduced description of the capacity fade and IR rise curve, which combined with an empirical model describe the full curves. Regarding model selection and simplification, we effectively demonstrate that gradient information is not required for the prediction of future capacity degradation. To the best of our knowledge this is in stark contrast to all previous work in this domain, which explicitly or implicitly require gradient information for prediction. Lastly, our model shows competitive performance compared with prior art, demonstrating the power of deep learning unlocked by considering each data cycle individually.

In terms of future work, the methodology we present could be deployed to electrochemical impedance spectroscopy (EIS) data which, a priori, is easier to gather.

Methods

Learning rate scheduler. Starting from the default Keras learning rate, the learning rate scheduler updates the learning rate every 'decay step' number of epochs as described in Eq. (1)

$$\text{new learning rate} = \text{previous learning rate} \times \text{decay rate}. \tag{1}$$

Machine learning performance scores. The mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean square error (RMSE) are defined as follows: where $\boldsymbol{y}$ the vector of true values and $\hat{\boldsymbol{y}}$ is the vector of predicted values

$$\text{MAE}(\boldsymbol{y}, \hat{\boldsymbol{y}}) = \frac{1}{n_{\text{samples}}} \sum_{i=1}^{n_{\text{samples}}} |\hat{y}_i - y_i|, \tag{2}$$

$$\text{MAPE}(\boldsymbol{y}, \hat{\boldsymbol{y}}) = \frac{100\%}{n_{\text{samples}}} \sum_{i=1}^{n_{\text{samples}}} \frac{|\hat{y}_i - y_i|}{y_i}, \tag{3}$$

$$\text{RMSE}(\boldsymbol{y}, \hat{\boldsymbol{y}}) = \sqrt{\frac{1}{n_{\text{samples}}} \sum_{i=1}^{n_{\text{samples}}} (\hat{y}_i - y_i)^2}. \tag{4}$$

CRediT authorship contribution statement

Calum Strange: Conceptualization, Methodology, Software, Provided domain expertise, Writing - review & editing. **Gonçalo dos Reis:** Conceptualization, Provided domain expertise, Writing - review & editing, Supervised the work, Funding acquisition.

Declaration of competing interest

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to https://doi.org/10.1016/j.egyai.2021.100097. The University Court of the University of Edinburgh, has filed a patent related to this work: UK Patent Application No. 2020829.4, dated 31 December 2020.

Acknowledgements

C.S. & G.d.R thank R. Gilchrist and S. Li (both University of Edinburgh) for the helpful suggestions.

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)*.

G. dos Reis acknowledges support from the *Fundação para a Ciência e a Tecnologia* (Portuguese Foundation for Science and Technology, Portugal) through the project UIDB/00297/2020 (Centro de Matemática e Aplicações CMA/FCT/UNL).

References

[1] Fermín P, McTurk E, Allerhand M, Medina-Lopez E, Anjos MF, Sylvester J, et al. Identification and machine learning prediction of knee-point and knee-onset in capacity degradation curves of lithium-ion cells. Energy AI 2020;100006.
[2] Strange C, Li S, Gilchrist R, dos Reis G. Elbows of internal resistance rise curves in li-ion cells. Energies 2021;14(4):1206.
[3] Diao W, Saxena S, Han B, Pecht M. Algorithm to determine the knee point on capacity fade curves of lithium-ion cells. Energies 2019;12:2910.
[4] Neubauer J, Pesaran A. The ability of battery second use strategies to impact plug-in electric vehicle prices and serve utility energy storage applications. Lancet 2011;196:10351–8.
[5] Ecker M, Nieto N, Käbitz S, Schmalstieg J, Blanke H, Warnecke A, et al. Calendar and cycle life study of Li(NiMnCo)O2-based 18650 lithium-ion batteries. J Power Sources 2014;248:839–51.
[6] Cook R, Swan L, Plucknett K. Failure mode analysis of lithium ion batteries operated for low earth orbit CubeSat applications. J Energy Storage 2020;31:101561.
[7] Waag W, Fleischer C, Sauer DU. Critical review of the methods for monitoring of lithium-ion batteries in electric and hybrid vehicles. J Power Sources 2014;258:321–39.
[8] Chang Y, Fang H, Zhang Y. A new hybrid method for the prediction of the remaining useful life of a lithium-ion battery. Appl Energy 2017;206:1564–78.
[9] Richardson RR, Osborne MA, Howey DA. Gaussian process regression for forecasting battery state of health. J Power Sources 2017;357:209–19.
[10] Severson K, Attia P, Jin N, Perkins N, Jiang B, Yang Z, et al. Data-driven prediction of battery cycle life before capacity degradation. Nat Energy 2019;4:1–9.
[11] Al-Dulaimi A, Zabihi S, Asif A, Mohammadi A. Hybrid deep neural network model for remaining useful life estimation. In: ICASSP 2019-2019 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE; 2019, p. 3872–6.
[12] Ren L, Zhao L, Hong S, Zhao S, Wang H, Zhang L. Remaining useful life prediction for lithium-ion battery: A deep learning approach. IEEE Access 2018;6:50587–98.
[13] Richardson RR, Osborne MA, Howey DA. Battery health prediction under generalized conditions using a Gaussian process transition model. J Energy Storage 2019;23:320–8.
[14] Wu B, Widanage WD, Yang S, Liu X. Battery digital twins: Perspectives on the fusion of models, data and artificial intelligence for smart battery management systems. Energy AI 2020;100016.
[15] Chen C, Zuo Y, Ye W, Li X, Deng Z, Ong SP. A critical review of machine learning of energy materials. Adv Energy Mater 2020;10(8):1903242.
[16] Ng M-F, Zhao J, Yan Q, Conduit GJ, Seh ZW. Predicting the state of charge and health of batteries using data-driven machine learning. Nat Mach Intell 2020;1–10.
[17] Tang X, Liu K, Wang X, Gao F, Macro J, Widanage WD. Model migration neural network for predicting battery aging trajectories. IEEE Trans Transp Electrif 2020.
[18] Brosa Planella F, Sheikh M, Widanage WD. Systematic derivation and validation of a reduced thermal-electrochemical model for lithium-ion batteries using asymptotic methods. Electrochim Acta 2021;388:138524. http://dx.doi.org/10.1016/j.electacta.2021.138524, https://www.sciencedirect.com/science/article/pii/S0013468621008148.
[19] Chen C-H, Planella FB, O'regan K, Gastol D, Widanage WD, Kendrick E. Development of experimental techniques for parameterization of multi-scale lithium-ion battery models. J Electrochem Soc 2020;167(8):080534.
[20] Ecker M, Tran TKD, Dechent P, Käbitz S, Warnecke A, Sauer DU. Parameterisation of a physico-chemical model of a lithium-ion battery part i: Determination of parameters. J Electrochem Soc 2015;162(9):A1836–48.
[21] Ecker M, Käbitz S, Laresgoiti I, Sauer DU. Parameterization of a physico-chemical model of a lithium-ion battery: II. Model validation. J Electrochem Soc 2015;162(9):A1849.
[22] Park S, Kato D, Gima Z, Klein R, Moura S. Optimal experimental design for parameterization of an electrochemical lithium-ion battery model. J Electrochem Soc 2018;165(7):A1309.
[23] Hong J, Lee D, Jeong E-R, Yi Y. Towards the swift prediction of the remaining useful life of lithium-ion batteries with end-to-end deep learning. Appl Energy 2020;278:115646.
[24] Attia PM, Grover A, Jin N, Severson KA, Markov TM, Liao Y-H, et al. Closed-loop optimization of fast-charging protocols for batteries with machine learning. Nature 2020;578(7795):397–402.
[25] dos Reis G, Strange C, Yadav M, Li S. Lithium-ion battery data and where to find it. Energy AI 2021;100081. http://dx.doi.org/10.1016/j.egyai.2021.100081, https://www.sciencedirect.com/science/article/pii/S2666546821000355.
[26] Chollet F, et al. Keras: The python deep learning library. Ascl 2018;ascl–1806.
[27] Jones E, Oliphant T, Peterson P, et al. Scipy: Open source scientific tools for Python. 2001, http://www.scipy.org/.
[28] Gal Y, Ghahramani Z. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In: International conference on machine learning, 2016. p. 1050–9.
[29] Ma Y, Wu L, Guan Y, Peng Z. The capacity estimation and cycle life prediction of lithium-ion batteries using a new broad extreme learning machine approach. J Power Sources 2020;476:228581.
[30] Shen S, Nemani V, Liu J, Hu C, Wang Z. A hybrid machine learning model for battery cycle life prediction with early cycle data. In: 2020 IEEE transportation electrification conference & expo (ITEC). IEEE; 2020, p. 181–4.
[31] Yang F, Wang D, Xu F, Huang Z, Tsui K-L. Lifespan prediction of lithium-ion batteries based on various extracted features and gradient boosting regression tree model. J Power Sources 2020;476:228654.
[32] Herring P, Gopal CB, Aykol M, Montoya JH, Anapolsky A, Attia PM, et al. BEEP: A python library for battery evaluation and early prediction. SoftwareX 2020;11:100506.