

---

# HISTORY-AGNOSTIC BATTERY DEGRADATION INFERENCE

---

Mehrad Ansari<sup>1,2</sup>, Steven B. Torrisi<sup>1</sup>, Amalie Trewartha<sup>1</sup>, and Shijing Sun<sup>1,\*</sup>

<sup>1</sup> Energy & Materials Division, Toyota Research Institute, Los Altos, CA 94022, USA.

<sup>2</sup> Department of Chemical Engineering, University of Rochester, Rochester, NY, 14627, USA.

## ABSTRACT

Lithium-ion batteries (LIBs) have attracted widespread attention as an efficient energy storage device on electric vehicles (EV) to achieve emission-free mobility. However, the performance of LIBs deteriorates with time and usage, and the state of health of used batteries are difficult to quantify and to date are poorly understood. Having accurate estimations of a battery’s remaining life across different life stages would benefit maintenance, safety, and serve as a means of qualifying used batteries for second-life applications. Since the full history of a battery may not always be available in downstream applications, in this study, we demonstrate a deep learning framework that enables dynamic degradation trajectory prediction, while requiring only the most recent battery usage information. Specifically, our model takes a rolling window of current and voltage time-series inputs, and predicts the near-term and long-term capacity fade via a recurrent neural network. We exhaustively benchmark our model against a naive extrapolating model by evaluating the error on reconstructing the discharge capacity profile under different settings. We show that our model’s performance in accurately inferring the battery’s degradation profile is *agnostic* with respect to cell cycling history and its current state of health. This approach can provide a promising path towards evaluating battery health in running vehicles, enhance edge-computing battery diagnostics, and determine the state of health for used batteries with unknown cycling histories.

## 1 Introduction

Lithium-ion batteries (LIBs) are an enabling technology for clean energy innovation and zero-emissions transportation, which will play a critical role in reducing pollution and mitigating climate change. However, a key limitation is that the capacity of LIBs – the amount of charge a battery can reversibly hold in one discharge or usage cycle – deteriorates nonlinearly with time and usage and so LIBs must eventually be replaced. This reality makes awareness of a battery’s state of health (SOH) an important part of practical large-scale deployment of batteries in consumer, engineering, and industrial settings. Tools to monitor battery health and forecast future battery performance will be important for critical applications like grid-level energy storage systems [1], electric vehicles (EV) [2–4], aerospace [5], and solar power generation [6]. These applications may involve deployment in contexts with low computational resources, or where the full history of the battery is not available, introducing new practical design constraints for prognostic tools. Therefore, a computationally efficient model which can make useful predictions about battery degradation while using limited information about the cell in question would be highly desirable. In this work, we present a model architecture and label design built to satisfy these criteria that can predict the imminent degradation of the battery with minimal prior knowledge of the cell required.

The task of forecasting battery degradation with limited historical data is challenging by the fact that the LIB degradation process is highly path dependent, with various nonlinear and intersecting failure modes across time-scales, environmental conditions, use-cases, cell designs, and battery chemistries [7, 8]. This complexity makes it difficult to decompose aging mechanisms into a coherent physico-chemical model that can accurately predict and explain degradation. Approximate prediction of battery state of health (SOH) without detailed knowledge of the internal mechanisms can take many forms,

---

\*shijing.sun@tri.global

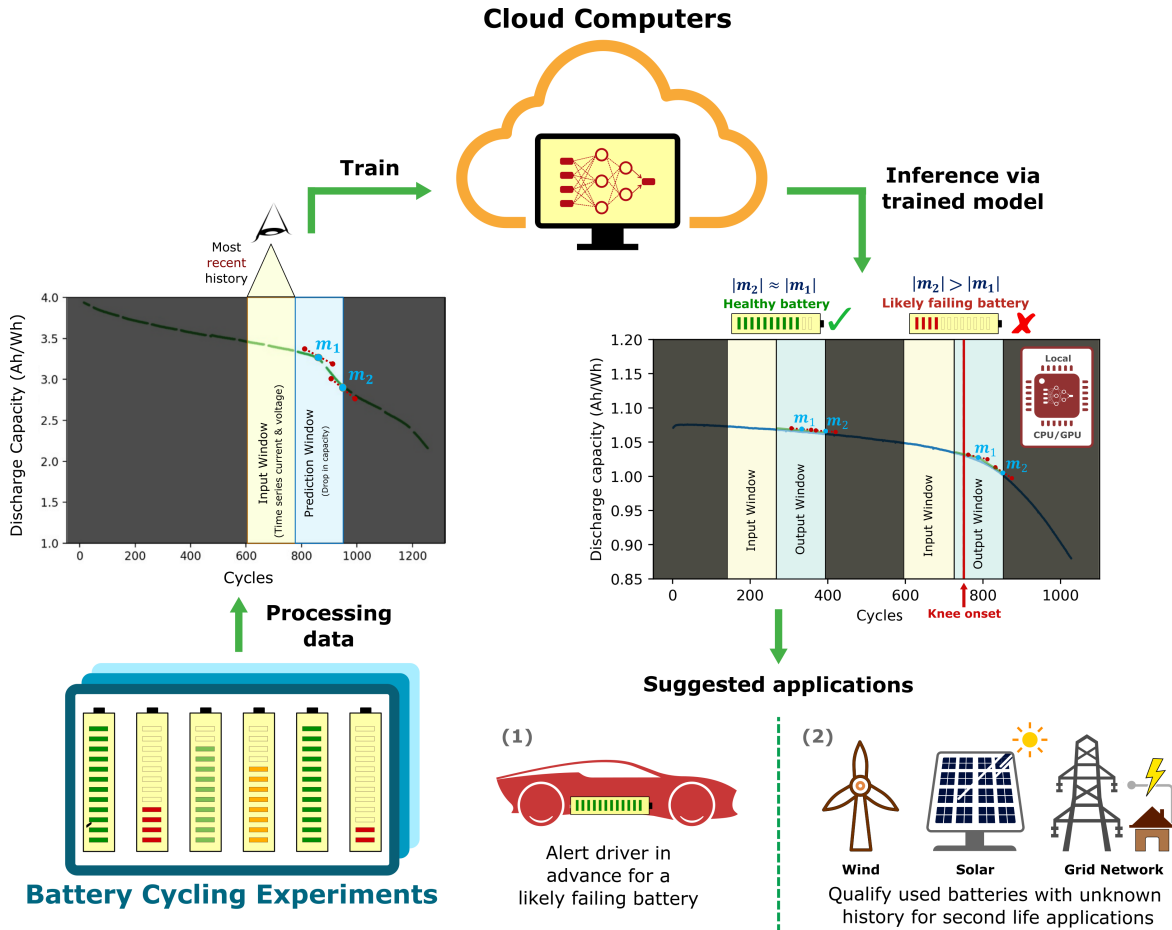


Figure 1: Overview of the proposed framework. Our recurrent neural network model inputs time-series data for current and voltage coming from a single, movable window of recent cycles, predicting near-term ( $m_1$ ) and long-term ( $m_2$ ) as slopes of the capacity fade curve in the consecutive prediction window. Approximately equal values of slopes (with small magnitudes) accounts for a healthy battery at early cycles, whereas diverging slopes show a likely failing battery at the vicinity of the knee onset. The resource-expensive training is done on cloud computers, however, the cheap-to-evaluate inference model can effectively run on low commodity CPU/GPUs, allowing for an on-board vehicle inference opportunity for cells in operation. Additionally, this approach can be insightful for qualifying used LIBs with unknown cycling histories for second life applications.

ranging from the diagnostic –estimating current capacity and/or power fade– to the prognostic — predicting the number of cycles remaining before discharge capacity drops below a particular threshold [9–11]. Accurate estimations of battery SOH during deployment not only aids maintenance, safety and resource optimization, but also serves as a means of qualifying batteries for different field usages, second-life applications, and new supply chain options. Moreover, early and frequent inferences on life cycle statistics can facilitate timely LIB research by allowing researchers to accelerate performance testing in response to cell design or process parameters, validating a particular configuration without having to go through exhaustive time-consuming and resource-intensive cycling measurements [12, 13].

The extensive research on LIB capacity degradation modeling can be categorized into mathematical physics-based models, and data-driven models. Comprehensive overviews of physics-based modeling of LIB degradation can be found in Refs. [14–16], but we will briefly summarize the attributes as well as strengths and weaknesses of some common models. The first category relies on empirical formulae with constraints set by physical principles, which can sometimes yield advantages of immediate interpretability after fitting to data. Particular advantages and disadvantages vary across common models like equivalent circuit models (ECM) [17–23], electrochemical models (EM) [24–29], or empirical models [30–32]. Equivalent circuit models can approximate the dynamic and static characteristics of a battery, and their establishment is dependent on the impedance data [33, 34]. Despite being computationally inexpensive, their application is limited by their inaccuracy under a wide range of C-rates [35]. Electrochemical models combine

the conventional pseudo two dimensional (P2D) model [36], single particle (SP) [37] or extended SP [38] model, with specific degradation mechanisms. They describe the transport phenomena using a system of partial differential equations, discretized over space and time, making the solution memory and computationally intensive, especially when phenomena of interest occur over several length scales. While electrochemical models can be useful in predicting battery life cycle, they tend to lack knowledge of the degradation mechanisms, and other challenges exist: the model parameterization process, balancing the trade-off between model complexity and prediction accuracy, and having to be re-solved with constantly changing operating conditions (i.e. temperature and power) limit their widespread use in online applications with cheap embedded hardware. Finally, capacity-based exponential empirical models analyze correlations from a large amount of data, and use different regression models to fit the declining capacity trend and forecast the degradation state. Despite being easy to establish and their demonstrated wide applicability, empirical degradation models are known to have high generalization error, unless their parameters are updated for different case studies.

On the other hand, machine learning and data-driven models like the one in this manuscript can be considered as part of the next generation of empirical models, where modern advances in hardware and software allow us to fit models with many parameters to approximate complex and nonlinear trends. Data-driven prognostics attempt to extract hidden correlations from a large amount of data, and generate a predictive model without relying on a mathematical physics-based model with a fixed analytical form. With advances in computing power, open-source tools, and available data, such models have received considerable attention in battery degradation studies over the past decade [39–45]. However, developing accurate prognostic models are challenging due to two reasons; 1) There exist only relatively small datasets available that span a limited range of lifetimes [46]. 2) More importantly, these datasets are skewed towards the early linear capacity-fade region, where there is negligible capacity degradation. Severson et al. [47], offered a significant demonstration with a cycling dataset of lithium-iron-phosphate (LFP) commercial batteries from A123 Systems. Using a simple regularized regression with hand-built features, they achieved high accuracy predictions, with RMSE of  $\sim 100$  cycles. Using different feature engineering techniques, subsequent papers have demonstrated the prediction of knee-point rollover on the same dataset. Fermín-Cueto et al. [48] developed a machine learning model that achieves early prediction of the knee-point, using only information from the first 50 cycles of the cells' life. Strange and Dos Reis [49] also used the Severson dataset and implemented a feature-free approach, utilizing a convolutional neural network to predict the entire capacity degradation curve with current and voltage data obtained from only from a preliminary cycle (though with limited accuracy).

Our work relies on a simple heuristic about the total observed degradation of discharge capacity: that the degradation of a fresh battery typically occurs in three regimes. First, a slow, linear decay in discharge capacity at the early life-cycle, second, accelerated nonlinear decay (a.k.a *knee point roll-over*) [50], and finally, fast linear decay. Thus, the knee point can also be identified using predictions about coarse changes in the curvature of the discharge capacity. The main contributions in this work address two limitations from the data-driven modeling and practical battery management points of view. 1) Our approach does not require having access to the full cycling history of the cell. This can be of importance in real-world applications, where the prior knowledge on the battery's life cycle is primarily limited to the most recent cycling data. More importantly, this advantage can provide a promising path towards a robust estimation of end-of-life (EOL) of used batteries with limited available cycling history, as well as suggesting second-life applications, given cells' inferred health status. 2) We demonstrate a feature-free deep learning model using an architecture that combines recurrent neural networks (Bi-directional Long Short-Term Memory), and convolutional neural networks. Our model uses windows of raw current and voltage time-series data as input and yields outputs of two scalars describing near- and long-term capacity loss. This can be critical in edge-computing battery diagnostics, where non-engineered features directly coming from the battery management system can be used for vehicle on-board inference. This can simply make the trained inference models computationally more tractable, and promote EOL estimations for operating cells in a running vehicle. It is also important to note that obtaining and engineering the most informative set of non-correlated features and is an important but challenging goal in real-world battery diagnostics applications that is the focus of ongoing research [7].

This manuscript is organized as follows. In Section 2, we start by describing the battery cycling dataset, and deep learning model architecture, as well as the choices for the model hyperparameters. This is followed in Section 3 by showing results on reconstructing the discharge capacity profile and evaluating the model performance in a comparative setting with a naive extrapolating model under different scenarios. Finally, we conclude the paper with a discussion of the implications of our findings, suggested model applications, and outlooks.

## 2 Materials and Methods

### 2.1 Dataset

Time-series data for current and voltage are obtained from aging tests of 139 commercial lithium-iron-phosphate (LFP)/graphite cells [47]. Variations in battery degradation are presented by plotting cells’ discharge capacity fade profiles in Figure S1 in the Supporting Information. An exclusion criteria is used to disregard short-lasting cells, and cycles where the discharge capacity falls below 80% of its nominal value (1.10 Ah). 100 cells are used for training, 25 for validation and hyperparameter tuning, and the remaining 14 cells are used for assessing model’s generalization error. For more details on the experimental data, readers are encouraged to refer to [47].

### 2.2 Model Inputs and Outputs

We first introduce our inputs and outputs to the model before describing the model itself. We consider time-series data for current and voltage coming from a single, movable window of cycles as the input to our model. The number of cycles in the input window is set at training time, which we vary in this study across  $\{10, 25, 50, 75, 100\}$  cycles. In other words, our model takes in as input “snippets” of voltage and current information from a set of sequential battery cycles, which we call the “input window”.

The outputs are, for successive “prediction windows” of cycles, the slope from a linear fit to the “degradation profile” (i.e. the discharge capacity observed over cycles) for two successive output windows of variable size set at training time, which in this study we vary across models from  $\{10, 25, 50, 75, \dots, 200\}$ . We refer to the predicted linearized dropoff interchangeably via variables called  $m$  or *slopes*, borrowed from the familiar form of  $y = mx + b$  from elementary linear fit. Note that rather than making predictions of the actual values of discharge capacities in the following output window of cycles, our model predicts the linearized magnitude of drop in the cell’s discharge capacity; in other words, for the linearized near-future of the battery, our model predicts a slope  $m$  but not the offset  $b$ . The dropoffs in capacity that we use as our output labels are measured by two slopes of two linear fits to the discharge capacity curve between the end of the last cycle in the input window and the mid-point of the prediction window, denoted as slope  $m_1$ , and from the mid-point to the final cycle of the following prediction window, denoted as slope  $m_2$ . Having multiple slopes enable us to better account for the changes in rate of degradation within the prediction window. See Figure 4 for a visual example on reconstructing the discharge capacity profile. In practice,  $m$  will always be a negative value as the battery falls off in capacity over time. The magnitude of  $m$  (how negative it is), thus, corresponds to a greater falloff in capacity. The motivation for this choice is that rather than explicitly predicting the falloff with full accuracy, we instead reduce the task to determining the onset of nonlinearity– which can be approximated by the difference between the two slopes. We note that the disagreement between the two slopes is proportional to the finite-difference approximation to the second derivative, wherein if  $x_0$ ,  $x_1$ , and  $x_2$  are at the beginning, midpoint, and end of the output prediction window respectively,  $f$  is the function describing the discharge capacity profile, and  $\Delta x$  is a step-size in the x-axis (here a particular number of cycles),

$$m_2 - m_1 \propto \frac{m_2 - m_1}{\Delta x} = \frac{f(x_2) - f(x_1)}{(\Delta x)^2} - \frac{f(x_1) - f(x_0)}{(\Delta x)^2} \approx f''(x_0)$$

$$\implies m_2 - m_1 \propto f''(x_0),$$

and thus, because the knee point is characterized by a period of nonlinear decay (where  $f'' < 0$ ), this quantity can be used as a means of detecting the onset of the knee point.

In summary, across the output window  $m_1$  and  $m_2$  respectively account for the short-term and long-term drop in the discharge capacity. In a healthy pre-knee cell, where discharge capacity drops linearly as a function of cycles, the absolute values for  $m_1$  and  $m_2$  are approximately equal ( $|m_2| \approx |m_1|$ ) and individually small in magnitude. On the other hand, in a degrading cell, where the discharge capacity drops nonlinearly in the vicinity of the knee point, the absolute value of  $m_2$  is expected to be greater than  $m_1$  ( $|m_2| > |m_1|$ ), and after the knee, even if falloff is linear ( $|m_2| \approx |m_1|$ )  $m_1$  and  $m_2$  will themselves be larger in magnitude (see Table 1 for a summary). This approach allows us to infer battery degradation using cell’s most recent cycling data with no prior knowledge of the degradation mechanisms, nor having access to the cell’s full cycle history, which is typically unavailable in practice.

### 2.3 Model Architecture

The model architecture is presented in Figure 2. The time-series current and voltage data for fixed number of  $N$  cycles is normalized and defined as the input. The temporal correlation between past and future data within each window of cycles are learned using a convolutional neural network (CNN). We apply the CNN to each cycle within the input

Table 1: Comparison between  $m_1$  and  $m_2$  at different stages of battery degradation.

Region	Capacity decay type	$m_1, m_2$ behavior
Pre-knee	Linear (slow decay)	$m_1 \approx m_2$ (small magnitude)
Knee	Nonlinear (rapid decay)	$ m_2  >  m_1 $ (increasing magnitude)
Post-knee	Linear (rapid decay)	$m_1 \approx m_2$ (large magnitude)

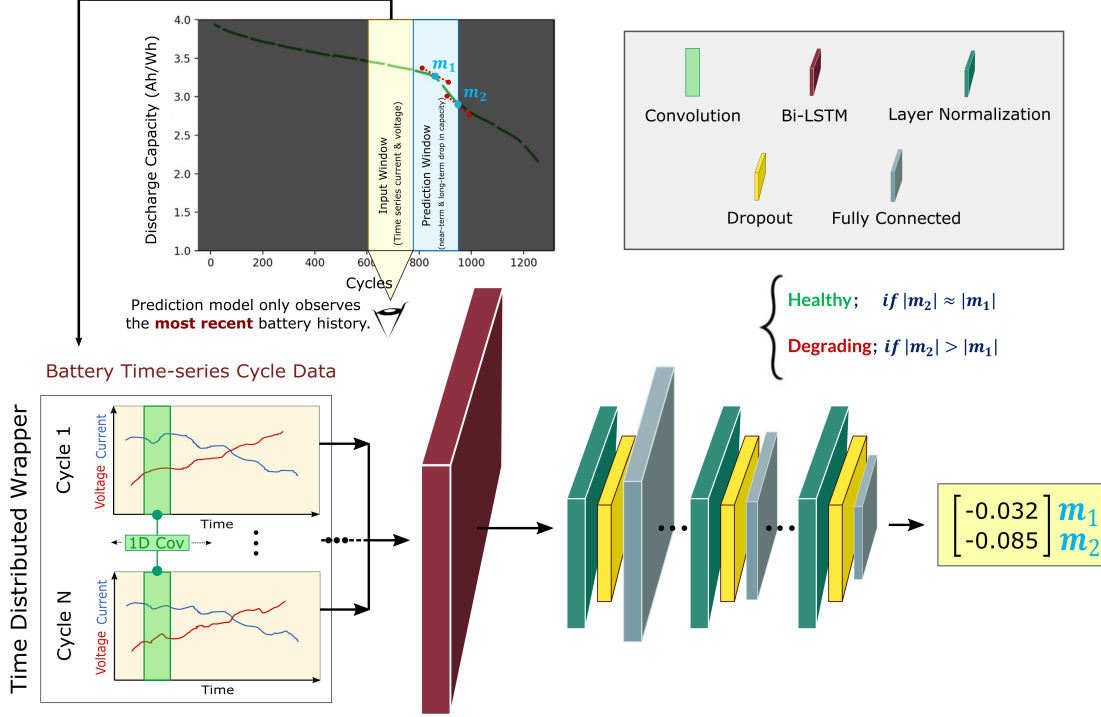


Figure 2: RNN model architecture. Model only observes times-series voltage and current information in a rolling input window of cycles, and makes prediction on the short-term ( $m_1$ ) and long-term drop ( $m_2$ ) in the discharge capacity. For a degraded battery, it is expected that  $|m_2| > |m_1|$ , otherwise the two values should be approximately equal before hitting the knee point roll-over.

window via a time distributed wrapper. This wrapper allows to apply any given layer (in our case a CNN) to every temporal slice of the input. To learn distant-based information, as well as the similarities and dissimilarities between different cycles within every input window, the output from the wrapper is fed to a double stacked Bi-directional Long Short-Term Memory (Bi-LSTM) neural network. The output from the Bi-LSTM is then normalized and fed to a dropout layer with a rate of 10%, followed by a dense neural network with a tanh activation function. This is repeated three times, until the final layer’s output is downsized to a vector of size 2, predicting the regressed values for the two slopes ( $m_1$  and  $m_2$ ), respectively.

The training is done via the Adam optimizer [51] with mean absolute error (MAE) as the loss function. The hyperparameters for the model are chosen based on the random search that resulted the lowest MAE in inferring the two slopes for the validation cells. We use an adaptive learning rate schedule, where the initial learning rate of  $10^{-3}$  is reduced by a factor of 0.9 and patience of 5 epochs when the monitored validation loss is not improving, until the learning rate reaches  $10^{-5}$ . To avoid overfitting, we add early stopping with patience of 10 that restores the best model weights from the epoch with the lowest MAE on the validation cells during training.

### 3 Results and Discussion

In this section, we run a systematic analysis with variations of sliding input and prediction window sizes, and evaluate the performance of our RNN model in predicting  $m_1$  and  $m_2$  on unobserved test cells. This is followed by using the

two predicted slopes to reconstruct the discharge capacity profile in the prediction window (snippet) via two linear fits with  $m_1$  and  $m_2$ , as demonstrated in Section 2.2. In essence, we use the predictions of  $m_1$  and  $m_2$  for forward propagation of the discharge capacity curve within the prediction window. We consider a rolling window of cycles with widths varying from  $\{10, 25, 50, 75, 100\}$  as input, and make predictions for  $m_1$  and  $m_2$  in their corresponding following prediction window with widths  $\{10, 25, 75, 100, 125, 150, 175, 200\}$ . As an example, if we choose an input window size of 50, and prediction window size of 50, the model is trained by slicing the discharge capacity profile into non-overlapping 50-cycle input, 50-cycle prediction windows for all the 100 training cells, where  $m_1$  and  $m_2$  are at the end of windows of the first and second 25 cycles each. Note that the input window only contains time-series information on current and voltage, and *not* the discharge capacities of previous cycles.

### 3.1 Predictive Performance on the Slopes

We show the parity plots for input window size of 50 and prediction window size of 50 in Figure 3, comparing RNN model predictions with the experimentally observed ground-truth values for all the test cells. In this Figure, panels 3a and b show results for  $m_1$  and  $m_2$ , respectively. Additionally, histogram plots for the error in predictions (i.e.  $m - \hat{m}$ , difference between ground-truth and predicted value) are represented in the right corner of panels 3a and b. Our RNN model provides generally accurate predictions on the slopes, and except for a small number of outliers, the error for majority of predictions fall into the  $\sim 0$  error bin for both  $m_1$  and  $m_2$ . Kernel density estimation plots for  $m_1$  and  $m_2$  predictions can be found in Figure S2 of the Supporting Information.

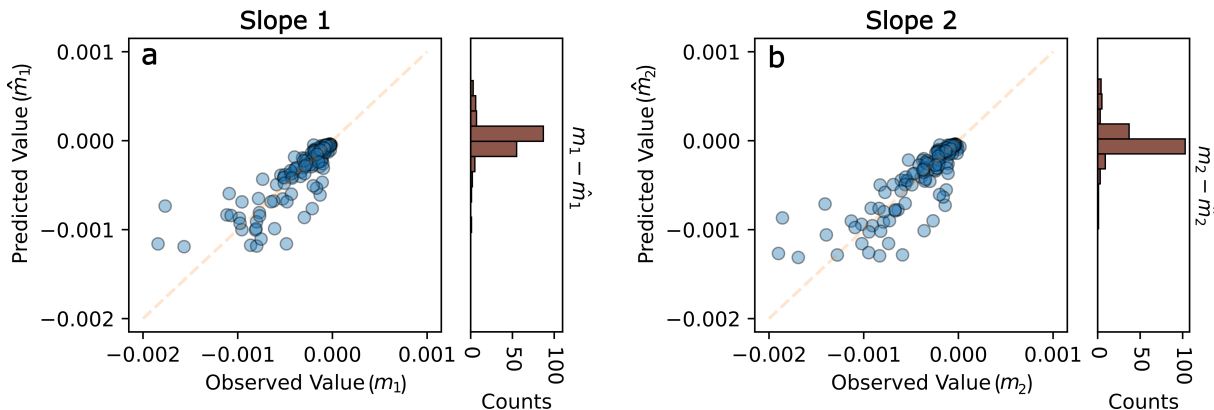


Figure 3: Parity plots for evaluating generalization error for slopes  $m_1$  and  $m_2$  on test cells, with histograms of the error presented on the right.  $m$  is the observed ground-truth, and  $\hat{m}$  denotes the predicted value for the slope.

### 3.2 Continuous Discharge Capacity Reconstruction

Next, we use the predictions for  $m_1$  and  $m_2$  slopes to reconstruct the discharge capacity profile within the prediction snippet as a function of cycles. This is simply done by constructing two lines using the predicted slope, and setting the offset to align the lines with the value of the discharge capacity at the beginning of the output window. In other words, the first line is defined based on the final known ground-truth discharge capacity from the input window as the intercept, and the slope of  $m_1$ , followed by having the second line intersect the first line at the cycle in the middle of the prediction window with the slope of  $m_2$ . Figure 4 shows the discharge capacity profile reconstructed on a single test cell for 6 snippets at different state of health of a single test cell, with input window size of 100, and prediction window size of 100. This input and prediction window sizes are selected for better visualization purposes. Our model is able to capture the capacity fade before, and even beyond the knee point roll-over, and successfully reconstruct the discharge capacity profile. By predicting  $m_1$  and  $m_2$  in a rolling fashion, we are able to continuously reconstruct the discharge capacity loss throughout a battery’s lifetime using the most recent cycling history, regardless of what state of health the battery is in. This also allows us to predict the drop in the discharge capacity even on cells currently under operation in a vehicle. A real-world application would be an additional vehicle warning built in the battery management system, where the driver is alerted if the battery’s discharge capacity is likely to fall below a certain threshold (i.e. 80%) within a small arbitrary number of charge-discharge cycles in the near future.

#### 3.2.1 Model Benchmarking for Varying Sliding Window Sizes

Next, we evaluate the performance of our RNN model by systematically changing the input and prediction window sizes. To benchmark our approach, we run a similar analysis using a *naive* model that extrapolates the discharge capacity

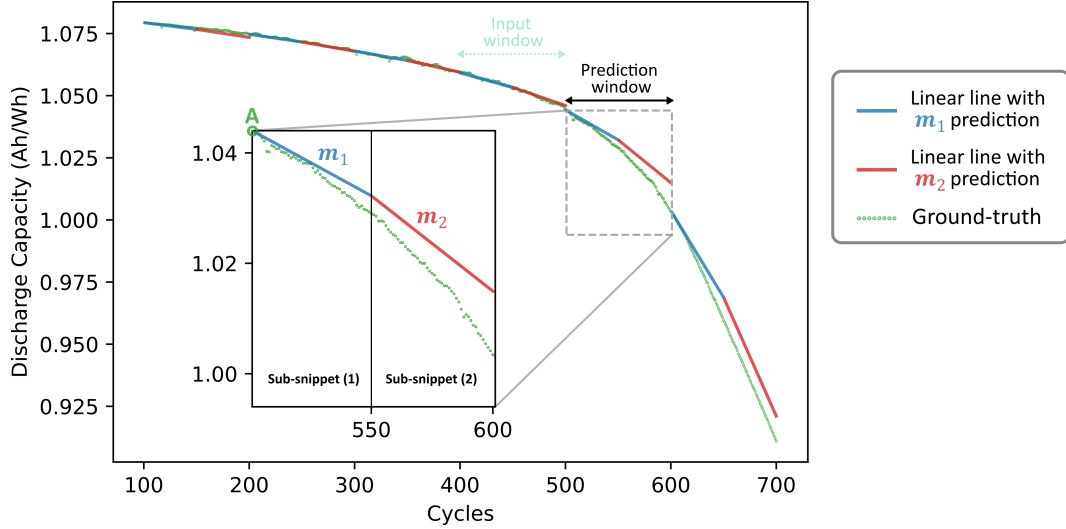


Figure 4: Discharge capacity reconstruction for a test cell, with input window size of 100, and prediction window size of 100 cycles. Green shows the ground-truth, and blue and red show the discharge capacity profile reconstructed using the predicted  $m_1$  and  $m_2$  slopes, respectively. From the input window, we only obtain data on current, voltage and not the discharge capacity, except for point A. Point A denotes the last known observed (ground-truth) discharge capacity coming from the final cycle of the input window, from which two linear lines are reconstructed using the predicted  $m_1$  and  $m_2$  slopes, as shown in sub-snippets (1) and (2), respectively.

profile in a simple linear fashion. This model represents the simplest practical way to extrapolate the discharge capacity fade. From this point of the manuscript onwards, we refer to the linearly extrapolating model as the *naive* model. A key difference between our RNN model and the naive model is the input. Rather than inputting voltage and current time-series data from the input window as the RNN model does, the naive model exclusively uses the entire discharge capacity profile within the input window, and makes a prediction on the discharge capacity profile in the following prediction window via linear extrapolation. In this setting, we define a second metric to evaluate the performance of our model, and assess its generalization error in reconstructing the discharge capacity profile. This error is quantified via the root-mean squared error (RMSE) between our model’s reconstructed discharge capacity profile, and the ground-truth profile from cycling experiments of test cells for each prediction window (snippet).

The discharge capacity reconstruction RMSE is represented as a contour plot for variations of the input and prediction window sizes in Figure 5, where panels 5a and b show the error by the RNN and the naive extrapolating model, respectively. For easier comparison, panel c visualizes the difference between the first two panels. Given different input and prediction window sizes, the RMSE reported at each point of the grid is averaged over all resulting snippets for all test cells. All RMSE values in between the grid are calculated through linear interpolation. The discharge capacity reconstruction RMSE evaluates how the predicted discharge capacity profile deviates from the ground-truth profile within the prediction windows.

It is important to note that we neglect the data from the first 2% of discharge capacity drop (compared to the nominal capacity), when defining the prediction window size, which would cause the overall performance of the naive model (and likely the RNN model) to be overestimated. The rationale behind this is to avoid including the earliest part of the battery’s first linear capacity drop region, where  $|m_2 - m_1| \approx 0$  (this is equivalent to the very first few snippets in Figure 4), and focus more on the prediction windows at and beyond the knee point roll-over, where the expected magnitude of error is significantly larger. For the sake of comparison, a similar analysis was done by removing this minimum cut-off and having no restrictions on where the prediction window is throughout the entire battery cycling history (see Figure S3 in the Supporting Information). Regardless of the input and prediction window size, our RNN model outperforms the naive extrapolating model, by having a much lower RMSE in reconstructing the discharge capacity profile (Figure 5a). In fact, the maximum observed RMSE from the RNN model ( $\approx 0.0171$ , marked as point A in Figure 5) is less than half of the maximum RMSE using the naive model. More importantly, as observed in Figure 5c, with the increase of the input window size while setting the prediction window size fixed, the RMSE from the RNN model decreases, whereas we see the opposite trend in the naive model (Figure 5b). The reason for this trend can be attributed to the fact that a larger input window for a linear fit increases the influence of the earlier values of the curve (where the curve is virtually always flatter), especially in the vicinity of knee-point (where the nonlinearity is at a global

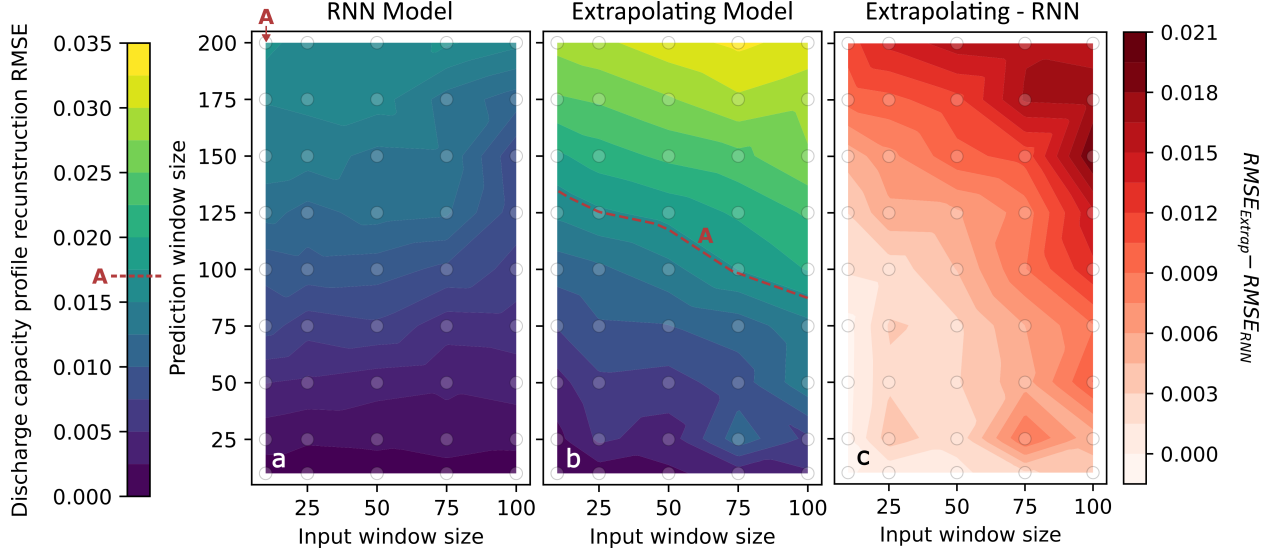


Figure 5: Discharge capacity profile reconstruction RMSE with minimum discharge capacity cut-off of 2%. a) Our RNN model, b) Naive linearly extrapolating model. c) Difference between panels a and b. RNN outperforms the naive model, especially at larger prediction window sizes. Surprisingly, the growth in RMSE with the increase of prediction window size using our RNN is much smaller than the naive model. Point A shows the largest RMSE ( $\approx 0.0171$ ) from the RNN model (highlighted by red dashed line in panel b).

maximum). It is important to note that our RNN model can learn this nonlinear behavior with the increase of the input window size, especially at larger prediction sizes (100-200), where the RMSE is lower. And finally, with the increase of the prediction window size, the RMSE increase from our RNN model is much slower than the naive model. This is a critical strength of our approach, as in real-world battery diagnostic applications, we care about predicting the capacity fade in the *largest* possible number of cycles in the future, while having access to the *least* amount of information on cell’s cycling history.

### 3.3 History-agnostic Inference at Varying State of Health

Finally, we benchmark our model for degradation prediction at different states of health. This is done by evaluating the capacity reconstruction RMSE for each prediction window, as a function of the absolute difference between the predicted slopes (i.e.  $|m_2 - m_1|$ ), as represented in Figure 6. The newly defined  $|m_2 - m_1|$  prediction determines the level of nonlinearity of the capacity fade within the prediction window, and is mostly equal to zero. This is expected since most of the cell’s cycling history occurs during a period of linear capacity fade, which is either slow linear decay before the knee, or rapid linear decay beyond the knee point. Every column in Figure 6, has RMSE results for small, medium, and large prediction window sizes (i.e. 10, 100, 200, respectively). Similarly, rows are sorted based on small, medium, and large input window sizes (i.e. 10, 50, 100, respectively). Each point is color coded based on the normalized value of the maximum  $|m_2 - m_1|$  observed within the history of cycles in each test cell normalized by the global maximum  $|m_2 - m_1|$  for a particular cell ( $M_{observed}/M_{max}$ ). This allows us to have a better assessment of the model performance at the regions specifically close to the knee point roll-over, where the  $M_{observed}/M_{max}$  reaches 1.0. This region is represented by points marked with red color. The blue points ( $M_{observed}/M_{max} \approx 0$ ) correspond to the linear capacity fade at the early cycles of the cells. Colors in-between account for transition region between linear and nonlinear domains of the discharge capacity profile. Considering the first column in Figure 6, the RMSE for small prediction windows are by an order of magnitude smaller than medium and large ones, regardless of the choice for input window size. Surprisingly, for prediction window size of 10, the magnitude of this error is *agnostic* to the current state of health of the tested battery, as the majority of points with red and blue colors are colocated at the very low RMSE. Note that we use a simple method of double linear fits via  $m_1$  and  $m_2$  to reconstruct the capacity fade profile. Consequently, for larger prediction window sizes (i.e. 100 and 200), and where the knee point roll-over falls within the prediction window (red points), it is expected that RMSE grows. It is naturally expected that deploying more linear fits within the prediction window can help reduce RMSE at larger prediction windows, especially at the vicinity of the knee point roll-over. It is important to note that the top right panel in Figure 6 exactly corresponds to the point A in Figure 5a, where we have pushed our RNN model to its lowest-performing regime. On the other hand, at the linear-nonlinear



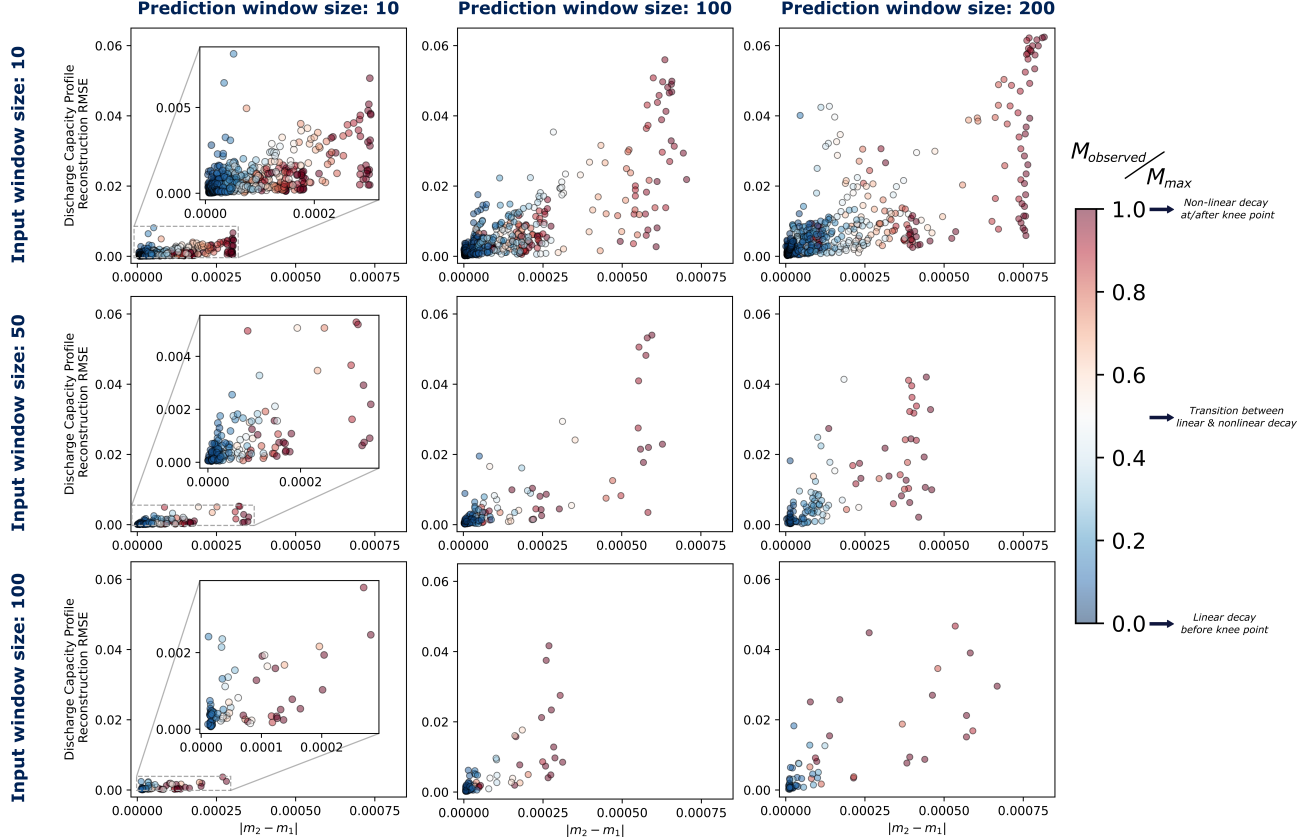


Figure 6: History-agnostic inference at varying state of health. Columns (prediction window size) and rows (input window size) are sorted by size. Each point represents an input/output window pair “snippet” where y-axis values represent reconstruction RMSE (see Figure 4 for depiction of reconstruction using linear fits) and x-axis values represent  $|m_2 - m_1|$  and therefore indicate the linearity in the output window (see Table 1 for a summary of what different values mean). Each point represents reconstruction RMSE for the prediction window at  $|m_2 - m_1|$ .  $M_{observed}$  denotes the maximally observed  $|m_2 - m_1|$  within the history of cycles for a particular cell as a way to differentiate points before and after the knee point which are approximately linear i.e. those with small  $|m_2 - m_1|$ . Color represents the ratio between the  $M_{observed}$  and the global maximum  $|m_2 - m_1|$  value  $M_{max}$  for a particular cell. The color bar’s  $M_{observed}/M_{max}$  value starts at 0 in the first linear regime, saturates to 1 at the knee point, then all successive points remain at 1. Thus, blue points account for the region with small capacity fade at the early cycles. White points represent the transition between linear and nonlinear decay regions. Red points represent both snippets at the knee point and afterwards in the second linear, more rapid, decay region.

transition region, marked by the white points, where  $M_{observed}/M_{max}$  monotonically approaches 1.0 but  $|m_2 - m_1|$  is still small, we see low RMSE compared to the points around the knee where  $|m_2 - m_1|$  is at its largest. This shows that the expected error from the RNN model is low before and after the knee. Even, at the nonlinear knee region, this error can be substantially reduced by an educated choice of the hyperparameters, such as the prediction window size. While RMSE rises at the knee point (especially at the larger prediction window sizes), the discharge capacity can still be reconstructed with reasonable confidence, as demonstrated in Figure 4.

### 3.4 Applications and Outlook

This model can be effectively used to enhance degradation and failure predictions for improved maintenance, and reliable operation of battery systems. Specifically, our approach can provide insights for battery’s end-of-life assessments in settings where the entire cycling history is either expensive to evaluate via low commodity resources, or is unknown; 1) The former setting can be promising in the design of low-power edge computing battery diagnostics, built in the battery management systems of vehicles in operation. With the continuous improvement of hardware (i.e Moore’s law [52]) and some minor compromises on model size and complexity, our approach can be used to predict the magnitude of cell capacity fade using the most recent charge-discharge cycling history, and alert the driver for a scheduled maintenance

in case of a likely degrading battery. The use of *most recent* cycling history allows for resource-efficient on-board vehicle inference. It is also important to note that the computational resource-intensive model training step is completed on high-performance or cloud computers, whereas the inference model, which is typically cheap to evaluate, can be loaded into the battery management system. 2) In the latter setting, our approach can be applied to real-world used battery second-life application assessments, where typically a cell's entire cycling history is unknown. By only using a limited recent cycling history, the used battery can be qualified for different small-scale applications in stationary energy storage systems that typically require less-frequent battery cycling (i.e. 100 to 300 cycles per year). This can help battery manufacturers and recyclers develop more established plans for used batteries, and thus, prevent the current mass disposal strategy of LIBs that poses many environmental and financial concerns.

## 4 Conclusions

Data-driven modeling is a promising route for battery diagnostics and early end-of-life estimations by bypassing the requirement for time-consuming and resource-intensive battery life cycle testing. In this work, we develop a deep learning model that is capable of dynamically predicting short and long-term discharge capacity fade in commercial graphite/LiFePO<sub>4</sub> batteries[47] cycled under fast-charging conditions. Our model takes as input time-series voltage and current data from an arbitrary window of cycles, and reconstructs the discharge capacity fade for the following window of cycles, *without* the need for access to the full cycling history of the battery. This approach can complement battery management systems by allowing for an on-board resource-efficient inference of the end-of-life for operating cells, identify abnormal cell behaviors, and even warn for the time that the battery requires replacement. More importantly, this approach can provide a promising path towards estimating end-of-life of used batteries, where there are typically limited cycling history available, and suggest second-life applications given the cells' state of health.

## Conflict of Interest Statement

The authors are *in the process of filing* a patent on the work described in this paper.

## Acknowledgements

S.S. acknowledges Xiao Cui, Patrick Asinger from the Data-Driven Design of Li-ion Batteries (D3BATT) program and Jens Hummelshøj from the E&M Division at Toyota Research Institute for fruitful discussions.

## Data Availability

The datasets used in this study are available at <https://data.matr.io/1>. BEEP structured data is used to pre-process the data in this study.

## Code Availability

All source code used for data processing and producing the results in this study are publicly available at <https://github.com/TRI-AMDD/bye-cycle>.

## References

- [1] Tianmei Chen, Yi Jin, Hanyu Lv, Antao Yang, Meiyi Liu, Bing Chen, Ying Xie, and Qiang Chen. Applications of lithium-ion batteries in grid-scale energy storage systems. *Transactions of Tianjin University*, 26(3):208–217, 2020.
- [2] Bruce Dunn, Hareesh Kamath, and Jean-Marie Tarascon. Electrical energy storage for the grid: a battery of choices. *Science*, 334(6058):928–935, 2011.
- [3] Jiuchun Jiang and Caiping Zhang. *Fundamentals and applications of lithium-ion batteries in electric drive vehicles*. John Wiley & Sons, 2015.
- [4] Richard Schmuch, Ralf Wagner, Gerhard Hörpel, Tobias Placke, and Martin Winter. Performance and cost of materials for lithium-based rechargeable automotive batteries. *Nature Energy*, 3(4):267–278, 2018.

- [5] Linda J Bolay, Tobias Schmitt, Omar S Mendoza-Hernandez, Yoshitsugu Sone, Arnulf Latz, and Birger Horstmann. Degradation of lithium-ion batteries in aerospace. In *2019 European Space Power Conference (ESPC)*, pages 1–3. IEEE, 2019.
- [6] Boucar Diouf and Ramchandra Pode. Potential of lithium-ion batteries in renewable energy. *Renewable Energy*, 76:375–380, 2015.
- [7] Noah H Paulson, Joseph Kubal, Logan Ward, Saurabh Saxena, Wenquan Lu, and Susan J Babinec. Feature engineering for machine learning enabled early prediction of battery lifetime. *Journal of Power Sources*, 527: 231127, 2022.
- [8] Aijun Yin, Zhibin Tan, and Jian Tan. Life prediction of battery using a neural gaussian process with early discharge characteristics. *Sensors*, 21(4):1087, 2021.
- [9] Datong Liu, Hong Wang, Yu Peng, Wei Xie, and Haitao Liao. Satellite lithium-ion battery remaining cycle life prediction with novel indirect health indicator extraction. *Energies*, 6(8):3654–3668, 2013.
- [10] Shengjin Tang, Chuanqiang Yu, Xue Wang, Xiaosong Guo, and Xiaosheng Si. Remaining useful life prediction of lithium-ion batteries based on the wiener process with measurement error. *energies*, 7(2):520–547, 2014.
- [11] Ghassan Zubi, Rodolfo Dufo-López, Monica Carvalho, and Guzay Pasaoglu. The lithium-ion battery: State of the art and future perspectives. *Renewable and Sustainable Energy Reviews*, 89:292–308, 2018.
- [12] Mohammad Arif Ishtiaque Shuvo, Gerardo Rodriguez, Md Tariqul Islam, Hasanul Karim, Navaneet Ramabadran, Juan C Noveron, and Yirong Lin. Microwave exfoliated graphene oxide/tio<sub>2</sub> nanowire hybrid for high performance lithium ion battery. *Journal of Applied Physics*, 118(12):125102, 2015.
- [13] Masahiro Kawasaki, Viratchara Laokawee, Thapanee Sarakonsri, Takashi Hashizume, and Makoto Shiojiri. Structural investigation of sisn/(reduced graphene oxide) nanocomposite powder for li-ion battery anode applications. *Journal of Applied Physics*, 120(20):204302, 2016.
- [14] Chien-Fan Chen, Pallab Barai, and Partha P Mukherjee. An overview of degradation phenomena modeling in lithium-ion battery electrodes. *Current Opinion in Chemical Engineering*, 13:82–90, 2016.
- [15] Yang Li, Dulmini Ralahamilage, Mahinda Vilathgamuwa, Yateendra Mishra, Troy Farrell, San Shing Choi, and Changfu Zou. Model order reduction techniques for physics-based lithium-ion battery management: A survey. *IEEE Industrial Electronics Magazine*, 2021.
- [16] Jacqueline S Edge, Simon O’Kane, Ryan Prosser, Niall D Kirkaldy, Anisha N Patel, Alastair Hales, Abir Ghosh, Weilong Ai, Jingyi Chen, Jiang Yang, et al. Lithium ion battery degradation: what you need to know. *Physical Chemistry Chemical Physics*, 23(14):8200–8221, 2021.
- [17] Gregory L Plett. Sigma-point kalman filtering for battery management systems of lipb-based hev battery packs: Part 2: Simultaneous state and parameter estimation. *Journal of power sources*, 161(2):1369–1384, 2006.
- [18] Hongwen He, Rui Xiong, and Hongqiang Guo. Online estimation of model parameters and state-of-charge of lifepo<sub>4</sub> batteries in electric vehicles. *Applied Energy*, 89(1):413–420, 2012.
- [19] Juergen Remmlinger, Michael Buchholz, Thomas Soczka-Guth, and Klaus Dietmayer. On-board state-of-health monitoring of lithium-ion batteries using linear parameter-varying models. *Journal of Power Sources*, 239: 689–695, 2013.
- [20] Christian Fleischer, Wladislaw Waag, Hans-Martin Heyn, and Dirk Uwe Sauer. On-line adaptive battery impedance parameter and state estimation considering physical principles in reduced order equivalent circuit battery models: Part 1. requirements, critical review of methods and modeling. *Journal of Power Sources*, 260:276–291, 2014.
- [21] Weihai Li, Monika Rentemeister, Julia Badedo, Dominik Jöst, Dominik Schulte, and Dirk Uwe Sauer. Digital twin for battery systems: Cloud battery management system with online state-of-charge and state-of-health estimation. *Journal of energy storage*, 30:101557, 2020.
- [22] Lei Zhang, Wentao Fan, Zhenpo Wang, Weihai Li, and Dirk Uwe Sauer. Battery heating for lithium-ion batteries based on multi-stage alternative currents. *Journal of Energy Storage*, 32:101885, 2020.
- [23] Hujun Peng, Jianxiang Li, Kai Deng, Andreas Thul, Weihai Li, Lars Lowenstein, Dirk Uwe Sauer, and Kay Hameyer. An efficient optimum energy management strategy using parallel dynamic programming for a hybrid train powered by fuel-cells and batteries. In *2019 IEEE vehicle power and propulsion conference (VPPC)*, pages 1–7. IEEE, 2019.
- [24] Nima Lotfi, Jie Li, Robert G Landers, and Jonghyun Park. Li-ion battery state of health estimation based on an improved single particle model. In *2017 American Control Conference (ACC)*, pages 86–91. IEEE, 2017.

- [25] Weihan Li, Decheng Cao, Dominik Jöst, Florian Ringbeck, Matthias Kuipers, Fabian Frie, and Dirk Uwe Sauer. Parameter sensitivity analysis of electrochemical model-based battery management systems for lithium-ion batteries. *Applied Energy*, 269:115104, 2020.
- [26] Florian Ringbeck, Marvin Garbade, and Dirk Uwe Sauer. Uncertainty-aware state estimation for electrochemical model-based fast charging control of lithium-ion batteries. *Journal of Power Sources*, 470:228221, 2020.
- [27] Weihan Li, Yue Fan, Florian Ringbeck, Dominik Jöst, Xuebing Han, Minggao Ouyang, and Dirk Uwe Sauer. Electrochemical model-based state estimation for lithium-ion batteries with adaptive unscented kalman filter. *Journal of Power Sources*, 476:228534, 2020.
- [28] Jorn M Reniers, Grietus Mulder, and David A Howey. Unlocking extra value from grid batteries using advanced models. *Journal of Power Sources*, 487:229355, 2021.
- [29] Joel C Forman, Scott J Moura, Jeffrey L Stein, and Hosam K Fathy. Genetic identification and fisher identifiability analysis of the doyle–fuller–newman model from experimental cycling of a lifepo4 cell. *Journal of Power Sources*, 210:263–275, 2012.
- [30] Johannes Schmalstieg, Stefan Käbitz, Madeleine Ecker, and Dirk Uwe Sauer. A holistic aging model for li (nimnco) o2 based 18650 lithium-ion batteries. *Journal of Power Sources*, 257:325–334, 2014.
- [31] Ira Bloom, BW Cole, JJ Sohn, Scott A Jones, Edward G Polzin, Vincent S Battaglia, Gary L Henriksen, Chester Motloch, Richardson Richardson, T Unkelhaeuser, et al. An accelerated calendar and cycle life study of li-ion cells. *Journal of power sources*, 101(2):238–247, 2001.
- [32] Dirk Uwe Sauer and Heinz Wenzl. Comparison of different approaches for lifetime prediction of electrochemical systems—using lead-acid batteries as example. *Journal of Power sources*, 176(2):534–546, 2008.
- [33] Jürgen Remmlinger, Michael Buchholz, Markus Meiler, Peter Bernreuter, and Klaus Dietmayer. State-of-health monitoring of lithium-ion batteries in electric vehicles by on-board internal resistance estimation. *Journal of power sources*, 196(12):5357–5363, 2011.
- [34] Akram Eddahech, Olivier Briat, Eric Woirgard, and Jean-Michel Vinassa. Remaining useful life prediction of lithium batteries in calendar ageing for automotive applications. *Microelectronics Reliability*, 52(9-10):2438–2442, 2012.
- [35] Maricela Best Mckay, Brian Wetton, and R Bhushan Gopaluni. Learning physics based models of lithium-ion batteries. *IFAC-PapersOnLine*, 54(3):97–102, 2021.
- [36] Thomas F Fuller, Marc Doyle, and John Newman. Simulation and optimization of the dual lithium ion insertion cell. *Journal of the electrochemical society*, 141(1):1, 1994.
- [37] Bala S Haran, Branko N Popov, and Ralph E White. Determination of the hydrogen diffusion coefficient in metal hydrides by impedance spectroscopy. *Journal of Power Sources*, 75(1):56–63, 1998.
- [38] Weilin Luo, Chao Lyu, Lixin Wang, and Liqiang Zhang. A new extension of physics-based single particle model for higher charge–discharge rates. *Journal of Power Sources*, 241:295–310, 2013.
- [39] Selina SY Ng, Yinjiao Xing, and Kwok L Tsui. A naive bayes model for robust remaining useful life prediction of lithium-ion battery. *Applied Energy*, 118:114–123, 2014.
- [40] Ji Wu, Chenbin Zhang, and Zonghai Chen. An online method for lithium-ion battery remaining useful life estimation using importance sampling and neural networks. *Applied energy*, 173:134–140, 2016.
- [41] Lei Ren, Li Zhao, Sheng Hong, Shiqiang Zhao, Hao Wang, and Lin Zhang. Remaining useful life prediction for lithium-ion battery: A deep learning approach. *Ieee Access*, 6:50587–50598, 2018.
- [42] Yongzhi Zhang, Rui Xiong, Hongwen He, and Michael G Pecht. Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries. *IEEE Transactions on Vehicular Technology*, 67(7):5695–5705, 2018.
- [43] Sheng Shen, Mohammadkazem Sadoughi, Xiangyi Chen, Mingyi Hong, and Chao Hu. A deep learning method for online capacity estimation of lithium-ion batteries. *Journal of Energy Storage*, 25:100817, 2019.
- [44] Xiaoyu Li, Lei Zhang, Zhenpo Wang, and Peng Dong. Remaining useful life prediction for lithium-ion batteries based on a hybrid model combining the long short-term memory and elman neural networks. *Journal of Energy Storage*, 21:510–518, 2019.
- [45] Weihan Li, Neil Sengupta, Philipp Dechent, David Howey, Anuradha Annaswamy, and Dirk Uwe Sauer. One-shot battery degradation trajectory prediction with deep learning. *Journal of Power Sources*, 506:230024, 2021.
- [46] B Saha and K Goebel. Battery data set, NASA Ames prognostics data repository. NASA Ames, Moffett Field, CA, USA, 2007.

- [47] Kristen A Severson, Peter M Attia, Norman Jin, Nicholas Perkins, Benben Jiang, Zi Yang, Michael H Chen, Muratahan Aykol, Patrick K Herring, Dimitrios Fraggedakis, et al. Data-driven prediction of battery cycle life before capacity degradation. *Nature Energy*, 4(5):383–391, 2019.
- [48] Paula Fermín-Cueto, Euan McTurk, Michael Allerhand, Encarni Medina-Lopez, Miguel F Anjos, Joel Sylvester, and Gonçalo dos Reis. Identification and machine learning prediction of knee-point and knee-onset in capacity degradation curves of lithium-ion cells. *Energy and AI*, 1:100006, 2020.
- [49] Calum Strange and Goncalo Dos Reis. Prediction of future capacity and internal resistance of li-ion cells from one cycle of input data. *Energy and AI*, 5:100097, 2021.
- [50] Stephen J Harris, David J Harris, and Chen Li. Failure statistics for commercial lithium ion batteries: A study of 24 pouch cells. *Journal of Power Sources*, 342:589–597, 2017.
- [51] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [52] Chris A Mack. Fifty years of moore’s law. *IEEE Transactions on semiconductor manufacturing*, 24(2):202–207, 2011.