

DETC2021-71407

Physics-Informed Machine Learning for Degradation Diagnostics of Lithium-Ion Batteries

Adam Thelen*

Department of Mechanical
Engineering,
Iowa State University
Ames, IA, USA

Yu Hui Lui*

Department of Mechanical
Engineering,
Iowa State University
Ames, IA, USA

Sheng Shen

Department of Mechanical
Engineering,
Iowa State University
Ames, IA, USA

Simon Laflamme

Department of Civil, Construction,
and Environmental Engineering,
Iowa State University
Ames, IA, USA

Shan Hu

Department of Mechanical
Engineering,
Iowa State University
Ames, IA, USA

Chao Hu†

Department of Mechanical
Engineering,
Iowa State University
Ames, IA, USA

ABSTRACT

State of health (SOH) estimation of lithium-ion batteries has typically been focused on estimating present cell capacity relative to initial cell capacity. While many successes have been achieved in this area, it is generally more advantageous to not only estimate cell capacity, but also the underlying degradation modes which cause capacity fade because these modes give further insight into maximizing cell usage. There have been some successes in estimating cell degradation modes, however, these methods either require long-term degradation data, are demonstrated solely on artificially constructed cells, or exhibit high error in estimating late-life degradation. To address these shortfalls and alleviate the need for long-term cycling data, we propose a method for estimating the capacity of a battery cell and diagnosing its primary degradation mechanisms using limited early-life degradation data. The proposed method uses simulation data from a physics-based half-cell model and early-life degradation data from 16 cells cycled under two temperatures and C rates to train a machine learning model. Results obtained from a four-fold cross validation study indicate that the proposed physics-informed machine learning method trained with only 60 early life data (five data from each of the 12 training cells) and 30 high-degradation simulated data can decrease estimation error by up to a total of 9.77 root mean square error % when compared to models which were trained only on the early-life experimental data.

Keywords: Lithium-ion battery, State of health, Degradation diagnostics, Physics-informed machine learning

1. INTRODUCTION

Lithium-ion (Li-ion) batteries are an attractive mobile energy storage device due to their high energy density, long cycle life, and continuously falling cost [1]–[3]. Despite the advantages, Li-ion battery cells degrade over time due to detrimental and irreversible internal electrochemical reactions during operation. To ensure safe and reliable operation within a system, it is important to monitor the state of health (SOH) of the Li-ion battery cell. The SOH of a battery cell is often defined as the ratio of its present cell capacity/resistance to its initial capacity/resistance. Accurate estimation of SOH is a crucial factor in implementing timely maintenance and replacement of battery cells/modules/packs.

Over the past few years, many health diagnostic methods have been developed based on readily available measurements (i.e., voltage, current, and temperature) from the cells. Generally, SOH estimation methods can be classified into two categories: (1) model-based methods [4]–[8] and (2) data-driven methods [9]–[14].

Model-based capacity estimation methods involve the use of empirical [4],[8]/equivalent circuit [8]/electrochemical [5],[6] models to fit the experimental data. Typically, the battery model is combined with advanced filtering techniques, such as Kalman filter and particle filter, to first estimate the model parameters through tracking the voltage (V) vs. capacity (Q) curve (VQ curve), then use the model parameters to estimate the SOH. While these model-based methods have shown successes in SOH estimation, constructing an accurate battery

* The authors contributed equally to this work.

† Corresponding author.

Emails: chaohu@iastate.edu; huchaostu@gmail.com.

model is a challenging task, given it often requires knowledge on the battery working principles or experimental data under carefully designed and well-controlled conditions.

Recently, data-driven methods have begun to emerge as an appealing alternative due to advances in data generation and increasing public availability of battery degradation data. A growing number of studies in the literature have applied machine learning techniques on SOH estimation by learning the complex dependency between the SOH of a cell and the characteristic features extracted from battery degradation measurements. In particular, popular machine learning methods such as linear regression [12], support vector machine (SVM) [12], relevance vector machine (RVM) [13], Gaussian process [14], and neural network (NN) [10], [11] have been successfully applied to the capacity estimation of Li-ion batteries.

While there have been many successes in estimating cell capacity using model-based and data-driven methods, these methods fall short in providing insight on the underlying degradation mechanisms which cause capacity fade. Three commonly reported degradation mechanisms in a Li-ion battery cell are loss of lithium inventory (LLI) and loss of active materials in the positive and negative electrodes (LAM_{PE} and LAM_{NE}, respectively) [15]–[18]. Limiting SOH estimation to available cell capacity does not fully reflect battery health. Diagnosing the underlying degradation mechanisms can provide better insight into battery health, which can facilitate prediction of future battery health [19], [20] and optimize the usage of battery cells [16]. Several degradation diagnostics methods have been proposed in the past. Han et al. proposed using membership functions to quantify the areas under the peak locations of the differential capacity (dQ/dV) curve and correlating these capacities to LLI and LAM_{NE} [16]. Birkel et al. proposed a diagnostic algorithm for estimating the degradation mechanisms (LLI, LAM_{PE} and LAM_{NE}) of a battery cell [15]. The proposed algorithm was experimentally validated through reconstructing the pseudo-OCV curve of a coin-cell with artificial degradation mechanisms. Dubarry et al. proposed an online SOH estimation method by comparing features (extracted from differential voltage (dV/dQ) and dQ/dV curves) from an online cell to features in a simulated offline path degradation database [17]. Tian et al. trained a convolutional neural network (CNN) to learn the relationship between a partial 1C charge curve and the electrode aging parameters (obtained from a pseudo-OCV curve) [18]. These electrode aging parameters can be used to quantify the degradation parameters of the battery cell.

Despite the promising results, these works either required long-term degradation data, were demonstrated solely on artificially constructed cells, or exhibited high errors when estimating degradation mechanisms at a late aging stage. To alleviate the time and cost in collecting long-term aging data, it is of great significance to develop methods that can enable accurate degradation diagnostics at a late aging stage using only early aging data. Such methods are beneficial for those who have limited budget and time for conducting aging tests.

In this work, we propose a methodology for constructing a data-efficient estimator that can estimate the capacity of a battery cell and diagnose its primary degradation mechanisms. The methodology incorporates physical knowledge of battery cell degradation into machine learning models by combining simulation data from a physics-based model with experimental data from the early cycle life only. The use of early cycle life data is to simulate a common scenario where experimental time is limited, resources are scarce, and long-term cycling is not an option. Different combinations of simulation data and experimental data are considered to investigate their effects on the long-term health estimation accuracy of the machine learning models. The estimation results suggest that accurate SOH estimation models only require training on a small amount of early-life, light-degradation experimental data and high-degradation simulated data. The role of the high-degradation simulated data is to loosely expand the physically meaningful degradation parameter space to a space that is larger than the space observed in the light-degradation data, enabling physics-informed machine learning. By expanding the observed space of the training data, the models are able to interpolate between low and high degradation data points, and more accurately estimate late-life capacity and degradation parameters.

The rest of the paper is arranged in the following manner. Section 2 presents the three machine learning models used in this study. Section 3 describes the methodologies of estimating degradation parameters using half-cell models, different configurations of the training dataset, and evaluation criteria of the model performance. Section 4 presents the experimental setup for aging data acquisition. Section 5 shows and discusses the results in estimating the health parameters (i.e., degradation parameters and cell capacity). Section 6 summarizes the key ideas of this study and offers concluding remarks.

2. MATHEMATICAL PRELIMINARIES

Consider the SOH estimation problem is to learn the relationship between the inputs and outputs of N distinct training samples $\{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^N$, where $\mathbf{x}_i \in \mathbb{R}^L$ and $\mathbf{y}_i \in \mathbb{R}^T$. Here L is the number of features, which is the number of dQ/dV readings per training/test sample in this study, and T is the number of tasks, which is equal to four in this study. The t th output can be represented by an output function $y_t = f_t(\mathbf{x})$, where $t = 1, \dots, T$.

2.1. Multi-output Gaussian process

Multi-output Gaussian process (MOGP) extends the Gaussian process (GP) to jointly model a vector of outputs [21]. Let us first consider a single-output GP defined by its zero mean trend function $m(\mathbf{x})$ and covariance function $k(\mathbf{x}, \mathbf{x}')$. We are interested in the target output $f_t(\mathbf{x})$ which can be expressed as

$$f_t(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \quad (1)$$

For simplicity, we only consider the squared exponential covariance function in this study. This kernel function takes the following form

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}')^\top P^{-1}(\mathbf{x} - \mathbf{x}')\right) \quad (2)$$

where the input variance is σ_f^2 and the length scales are encoded in $P \in \mathbb{R}^{L \times L}$.

Now we consider an MOGP with T outputs. We aim to model the isotopic training set which is now $\mathbf{x}_{t,i} = \mathbf{x}_i$ and $X_1 = \dots = X_T = \bar{X}$. Just like the single-output GP, the T outputs denoted as $\mathbf{f} = [f_1, \dots, f_T]^T$ also follow a GP model

$$\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k_M(\mathbf{x}, \mathbf{x}')) \quad (3)$$

where the new multi-output covariance function $k_M(\mathbf{x}, \mathbf{x}') \in \mathbb{R}^{T \times T}$ is defined to be

$$k_M(\mathbf{x}, \mathbf{x}') = \begin{bmatrix} k_{11}(\mathbf{x}, \mathbf{x}') & \dots & k_{1T}(\mathbf{x}, \mathbf{x}') \\ \vdots & \ddots & \vdots \\ k_{T1}(\mathbf{x}, \mathbf{x}') & \dots & k_{TT}(\mathbf{x}, \mathbf{x}') \end{bmatrix}. \quad (4)$$

Each element $k_{tt'}(\mathbf{x}, \mathbf{x}')$ represents the covariance between outputs, or in other words, the similarity between the tasks. Just like a single-output GP, we define the MOGP relationship to be

$$\mathbf{y}_t(\mathbf{x}) = f_t(\mathbf{x}) + \epsilon_t \quad (5)$$

where the term ϵ_t is iid Gaussian noise for each of the T outputs such that $\epsilon_t \sim \mathcal{N}(0, \sigma_t^2)$. The corresponding multi-output likelihood function is written as

$$p(\mathbf{y}|\mathbf{f}, \mathbf{x}, \Sigma_s) = \mathcal{N}(\mathbf{f}(\mathbf{x}), \Sigma_s) \quad (6)$$

where Σ_s is a diagonal matrix which contains the noise information for each output, i.e., $\Sigma_s = \text{diag}(\sigma_1^2, \dots, \sigma_T^2)$. Given the original multi-output training dataset, the posterior distribution at a new point $\mathbf{x}_*$ can be written

$$\mathbf{f}(\mathbf{x}_*)|\mathbf{x}, \mathbf{y}, \mathbf{x}_* \sim \mathcal{N}(\hat{\mathbf{f}}(\mathbf{x}_*), \Sigma_*) \quad (7)$$

The corresponding prediction mean and covariance are, respectively, given as

$$\hat{\mathbf{f}}(\mathbf{x}_*) = K_{M*}^\top [K_M(\bar{X}, \bar{X}) + \Sigma_M]^{-1} \mathbf{y} \quad (8)$$

$$\Sigma_* = K_M(\mathbf{x}_*, \mathbf{x}_*) - K_{M*}^\top [K_M(\bar{X}, \bar{X}) + \Sigma_M]^{-1} K_{M*} \quad (9)$$

where $K_{M*} = K_M(\bar{X}, \mathbf{x}_*)$ and has blocks $K_{tt'}(\bar{X}, \mathbf{x}_*) = [k_{tt'}(\mathbf{x}_i, \mathbf{x}_*)]$ for $t, t' = 1, \dots, T$ and $i = 1, \dots, n$.

2.2 Multitask Lasso and elastic net

Lasso and elastic net are two regression methods with built-in variable selection and regularization. Let x_{ij} denote the j th reading (or discretized measurement) in the i th dQ/dV charge curve where $i = 1:N$ and $j = 1:L$. In a multitask model with T tasks, $\mathbf{y}_{ti}$ with $t = 1, \dots, T$ represents an instance of a single task output. With the prior knowledge that the T tasks are related to one another, we define a multitask regression problem to map from a dQ/dV curve to the four tasks. The multitask regression model consists of a single design matrix X and multiple task-specific weights β_t :

$$\mathbf{y}_t = X\beta_t + \epsilon_t \quad (10)$$

where ϵ_t is normalized random noise. To improve the model performance and reduce overfitting, the $\mathbf{x}_j$'s are first standardized to have zero mean and unit variance. To find the optimal weights β_t for the T tasks, multitask Lasso solves the following:

$$\min_{\beta} J_1(\beta) = \frac{1}{2} \sum_{t=1}^T \|\mathbf{y}_t - X\beta_t\|_2^2 + \alpha \sum_{j=1}^L \|\beta_{tj}\|_2 \quad (11)$$

where the term $\sum_{j=1}^L \|\beta_{tj}\|_2$ is a mixed ℓ_1/ℓ_2 -norm which encourages sparsity across related tasks and α is a hyperparameter which controls the extent to which the mixed ℓ_1/ℓ_2 -norm is enforced in the loss function. After iterative

testing, it was found that 0.1 is an acceptable value for α . By including another quadratic penalty $\sum_{j=1}^L \|\beta_{tj}\|_2^2$ in the loss function along with an additional hyperparameter γ for adjusting the ratio of the two loss penalties, we arrive at the objective function for a multi-task elastic net model with the following formulation:

$$\min_{\beta} J_2(\beta) = \frac{1}{2} \sum_{t=1}^T \|\mathbf{y}_t - X\beta_t\|_2^2 + \alpha\gamma \sum_{j=1}^L \|\beta_{tj}\|_2 + \alpha(1-\gamma) \sum_{j=1}^L \|\beta_{tj}\|_2^2 \quad (12)$$

Through iterative testing, it was determined that optimal values for α and γ are 0.05 and 0.05, respectively.

2.3 Extreme learning machine

Extreme learning machine (ELM) is often used as a single hidden layer feedforward network (SLFN) with random input weights and biases that do not change throughout the training process. Mathematically, the SLFN with N distinct training samples $\{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^N$, where $\mathbf{x}_i \in \mathbb{R}^L$ and $\mathbf{y}_i \in \mathbb{R}^T$, can be formulated as

$$\mathbf{o}_i = \sum_{j=1}^D \beta_j g(\mathbf{w}_j^\top \cdot \mathbf{x}_i + b_j) = \mathbf{y}_i + \epsilon_i, \quad i = 1:N \quad (13)$$

where $\mathbf{o}_i$ is the output of the network, $\mathbf{w}_j = [w_{j1}, w_{j2}, \dots, w_{jL}]^\top$ is the hidden layer weight vector connecting the input neurons and the j th hidden neuron, b_j is the bias for j th hidden neuron, D is the total number of hidden neurons, $g(\cdot)$ is the activation function that is used to introduce nonlinearity to the SLFN, $\beta_j = [\beta_{1j}, \beta_{2j}, \dots, \beta_{Tj}]^\top$ is the output bias vector that connects the j th hidden neuron to output neurons, and ϵ is noise that includes both random noise and noise that depends on variables other than the inputs [22].

The above N equations can be written in the following matrix form for compact representation,

$$\mathbf{H}\beta = \mathbf{O} \quad (14)$$

$$\mathbf{H} = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \dots & g(\mathbf{w}_D \cdot \mathbf{x}_1 + b_D) \\ \vdots & \ddots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \dots & g(\mathbf{w}_D \cdot \mathbf{x}_N + b_D) \end{bmatrix}_{N \times D} \quad (15)$$

$$\beta = \begin{bmatrix} \beta_1^\top \\ \vdots \\ \beta_D^\top \end{bmatrix}_{D \times T} \quad \text{and} \quad \mathbf{O} = \begin{bmatrix} \mathbf{o}_1^\top \\ \vdots \\ \mathbf{o}_N^\top \end{bmatrix}_{N \times T} \quad (16)$$

where $\mathbf{H}$ is called the hidden layer output matrix.

During the training process, ELM uses the Moore-Penrose inverse (pseudoinverse) to find a solution to the linear system shown in equation (4). For many machine learning problems, the number of training instances (i.e., N) is much larger than the number of hidden neurons (i.e., D). Therefore, the linear system is an overdetermined system, and a unique solution is obtained by minimizing the ℓ_2 norm of the vector of training errors:

$$\hat{\beta} = \mathbf{H}^\dagger \mathbf{Y} \quad (17)$$

$$\mathbf{H}^\dagger = (\mathbf{H}^\top \mathbf{H})^{-1} \mathbf{H}^\top \quad (18)$$

where $\mathbf{H}^\dagger$ is the pseudoinverse of matrix $\mathbf{H}$, and $\hat{\beta}$ is the solution of the overdetermined system of linear equations $\mathbf{H}\beta = \mathbf{Y}$. Compared to the model training with back propagation which requires the iterative updating of input

weights and biases, the model training with ELM is much faster as it requires only a single process, pseudoinverse, to find the optimum solution.

3. METHODOLOGY

3.1 Half-cell model

The half-cell model is a non-destructive degradation analysis method that estimates the three commonly reported degradation mechanisms (i.e., LAM_{PE}, LAM_{NE}, and LLI) in a Li-ion battery cell by reconstructing the measured full-cell V (dV/dQ) vs. Q curve with positive and negative half-cell V (dV/dQ) vs. Q curves [23]–[25]. The VQ curve analysis ensures the upper and lower cutoff voltages of the simulated curve match those of the experimental curve. In contrast, the dV/dQ curve unveils the phase transition information of the battery electrode materials as identifiable peaks during charge/discharge and amplifies small changes in the voltage curve, thus facilitating the identification and quantification of degradation mechanisms of the battery cell. We use both VQ curve analysis and dV/dQ curve analysis to robustly identify and quantify the degradation mechanisms. The equations for VQ curve analysis and dV/dQ curve analysis are shown in equation (19) and (20), respectively:

$$V_c(Q)|_{Q=Q_c} \approx V_p(q_p)|_{q_p=\frac{Q_c-\delta_p}{m_p}} - V_n(q_n)|_{q_n=\frac{Q_c-\delta_n}{m_n}} \quad (19)$$

$$\left. \frac{dV(Q)}{dQ} \right|_{Q=Q_c} = \left. \frac{1}{m_p} \frac{dV_p(q_p)}{dq_p} \right|_{q_p=\frac{Q_c-\delta_p}{m_p}} - \left. \frac{1}{m_n} \frac{dV_n(q_n)}{dq_n} \right|_{q_n=\frac{Q_c-\delta_n}{m_n}} \quad (20)$$

where $V_c(Q_c)$ is the full-cell VQ curve with Q_c denoting the cell capacity, $V(q)$ is the half-cell curve, q is the specific capacity (mAh/g), m is the active mass (g), and δ is the half-cell curve slippage (mAh). The subscripts p and n correspond to the positive and negative electrodes (PE and NE), respectively. Slippage δ_p / δ_n quantifies the horizontal distance the left endpoint of the positive/negative half-cell curve with respect to $Q_c = 0$ mAh [20], [26]. When the value of δ increases (decreases), the half-cell curve shifts to the left (right). The active masses in the PE and NE control the capacity of the corresponding half-cell curves, which shrink when active masses decrease.

In this study, we identify three degradation parameters from the half-cell model to quantify the three degradation mechanisms in a Li-ion battery cell. These degradation parameters are m_p and m_n , used to quantify LAM_{PE} and LAM_{NE}, respectively, and the lithium inventory indicator (LII), which is defined as $LII = Q_p - (\delta_p - \delta_n)$, used to quantify LLI [19].

3.2 Input and output variables

This study aims to estimate the three degradation parameters and capacity of a cell based on its dQ/dV curve by using machine learning models. The input features of these models are the dQ/dV readings, calculated by differentiating the

capacity over the voltage. A sampling interval, ΔV , is defined to calculate the dQ/dV value at the specified voltage, $[V_{LC}, V_{LC} + \Delta V, V_{LC} + 2\Delta V, \dots, V_{LC} + (L-1)\Delta V]$ within the lower and upper cutoff voltages (V_{LC} and V_{UC} , respectively), where $(L-1) = \left\lfloor \frac{V_{UC}-V_{LC}}{\Delta V} \right\rfloor$. dQ/dV analysis possesses several benefits compared to VQ analysis and dV/dQ analysis. Compared to VQ analysis, dQ/dV analysis transforms the phase transitions of active electrode materials to identifiable peaks in the dQ/dV curve which are sensitive to small changes in the materials. These small changes can be learned by machine learning models to improve the diagnostic accuracy. Unlike the dV/dQ and VQ curves, the dQ/dV curve is dependent on a fixed, directly measurable voltage range (e.g., 3.4 V – 4.075 V considered in this study) instead of a capacity range, which could vary over the course of cell aging and be unmeasurable due to an unknown initial capacity in a partial charge cycle.

The input variables for one sample are split into L equally spaced voltage steps and the corresponding discretized dQ/dV values are considered as the input variables to be fed into a machine learning model. The larger the value L , the higher resolution the dQ/dV curve possesses.

Two datasets are considered in this study: 1) simulation dataset generated from half-cell model and 2) experimental dataset collected from a high precision charger (see Table 1). The simulation data are generated by sampling the three degradation parameters in a predefined design space using Latin hypercube sampling. This design space is selected such that it sufficiently encompasses possible extremes in the experimental data. The experimental data consists of 16 implantable-grade Li-ion battery cells that are aged in a long-term cycling test for up to 2.5 years. These cells are divided into four groups each running a unique test configuration (see the test matrix in Table 3).

Table 1 Summary of the two datasets used in this study.

Dataset	Simulation data	Experimental data
Number of cells	NA	16
Data size	10,000	372 (up to 2.5 years of aging test)
Source	Half-cell model	Implantable-grade Li-ion cells (see Table 3)

3.3 Training dataset configurations

In this subsection, we identify a number of training dataset configurations that serve the objective of constructing an accurate health (degradation mechanisms and capacity) estimator based on early-life degradation data by using a machine learning model. Each configuration consists of experimental data, simulation data, or both. The use of simulation data is to overcome the difficulty of extrapolation in data-driven models. Simulation data from a physics-based model can indirectly incorporate degradation physics into the data-driven models, thus facilitating the estimation of late-life health parameters. We label each configuration as EXPN_{EXP}

$\text{Sim}N_{\text{Sim}}$, where N_{EXP} indicates that the first N_{EXP} data points (early-life data) from each experimental cell are used in the training dataset and N_{Sim} is the number of randomly selected simulation data for inclusion in the training dataset.

In the first few configurations, we explore models trained using simulation data only. These configurations explore the accuracy of the half-cell model in its ability to mirror the true degradation observed in the experimental cells. Next, we investigate models trained using only early-life experimental data. These tests will help expose the models' extrapolation limits in estimating late-life degradation. As a baseline for training dataset configurations consisting of both experimental data and simulation data, we fix the early-life experimental data to be the first five data points from each cell ($N_{\text{EXP}} = 5$). For the experimental data used in this study, five data points are equivalent to approximately three months of lab testing. For the next set of training dataset configurations, we augment the baseline dataset with various amounts of simulation data selected either randomly from the simulation dataset or limited to high degradation simulation data (labelled as HiDeg (short for High Degradation)). For each training dataset consisting of both experimental and simulation data, we investigate the benefit of feature dimension reduction by using principal component analysis (PCA). PCA is used to reduce the original dimension of dQ/dV features from 100 to 20.

3.4 Error metric

To evaluate the performance of the machine learning models, we use a four-fold cross validation (CV). In particular, the complete experimental dataset consisting of 16 cells was divided into four mutually exclusive folds with the test data in each fold consisting of one battery cell (see Table 2) from each group in the test matrix (Table 3). The overall test errors $\varepsilon_{\text{RMSE}\%}^{\text{ALL}}$ of the machine learning models were estimated by taking the average of the individual test errors across the four training/test runs:

$$\varepsilon_{\text{RMSE}\%}^{\text{ALL}} = \sqrt{\frac{1}{\sum_{k=1}^4 N_k} \sum_{k=1}^4 \sum_{i=1}^{N_k} \left[\left(\frac{\hat{y}_{ti} - y_{ti}}{y_{ti}} \right) \times 100\% \right]^2} \quad (21)$$

where the subscript t denotes the t th health parameter, $\hat{y}_{ti}$ and y_{ti} denote the predicted value and true value for the t th degradation parameter at the i th test point, respectively. Using a normalized error metric is important when the outputs have different magnitudes and comparison among them is desired.

Table 2 CV folds indicating the online cells (test data).

Discharge rate	C/24		C/3	
Temperature	37 °C	55 °C	37 °C	55 °C
Fold 1	G1C1	G2C1	G3C1	G4C1
Fold 2	G1C2	G2C2	G3C2	G4C2
Fold 3	G1C3	G2C3	G3C3	G4C3
Fold 4	G1C4	G2C4	G3C4	G4C4

4. EXPERIMENTAL SETUP

We conducted a cycle aging test on 16 fresh implantable-grade Li-ion battery cells. The cells were cycled at two ambient temperatures and two discharge rates, a total of four test configurations (see Table 3). A temperature of 37 °C was chosen to simulate the normal working temperature of the implantable-grade battery cells and 55 °C was chosen to accelerate the capacity fade while maintaining the stability of the battery materials [27]. For each temperature, two charge/discharge conditions are considered, i.e., (C/3, C/24) and (C/3, C/3). Charging was carried out via a constant-current (CC), constant voltage (CV) step, where the cells were charged at C/3 to an upper cutoff voltage of 4.075 V, at which point, charging continued at a constant voltage of 4.075 V until either the charge current was C/50 or the charge time reached 30 min. Following the CC-CV charging step, the cells were CC discharged at C/3 or C/24 until the voltage reached a lower cutoff of 3.4 V. Compared to the nominal operating temperature and discharge rate of an implantable cell, 55 °C and C/3 conditions highly accelerate the aging of the cells [28].

During cycling, a characterization test is conducted at an ambient temperature of 40 °C every 2 weeks during the first 3 months and every 4 weeks thereafter. The characterization test is used to determine the battery cell capacity and voltage curve at a slow rate. The recorded capacity and voltage curves are used for degradation parameter quantification. The characterization test includes four sequential steps: 1) CC and CV charge to 4.075 V at C/3 with a cutoff current of C/50, 2) CC discharge to 3.4 V at C/50 and rest for 30 min, 3) CC charge to 4.075 V at C/50 and rest for 30 min, and 4) repeated CC discharge to a voltage corresponding to 10% of the state of charge reduction at C/10 followed by a 1-hour rest period until the cell voltage reached 3.4 V. The temperature (40 °C) and charging rate (C/50) are selected to be consistent with the test condition of the half-cells [19]. The capacity evolution of the 16 battery cells is shown in Fig. 1. Cell C1 and C2 from group G1 and G3, and cell C3 and C4 from group G2 and G4 were removed roughly halfway through the test for destructive analysis.

The charge curve at step 3 of the characterization test was used to quantify the degradation parameters with half-cell curve analysis. Both the cycling and characterization tests were conducted using a high precision charger from NOVONIX with a voltage range of 0–5 V and a maximum current output of 2 A for high-quality aging data acquisition [29].

Table 3 Test matrix of the experimental data.

Group	Charge rate	Discharge rate	Temperature	Number of cells
G1	C/3	C/24	37 °C	4
G2	C/3	C/24	55 °C	4
G3	C/3	C/3	37 °C	4
G4	C/3	C/3	55 °C	4

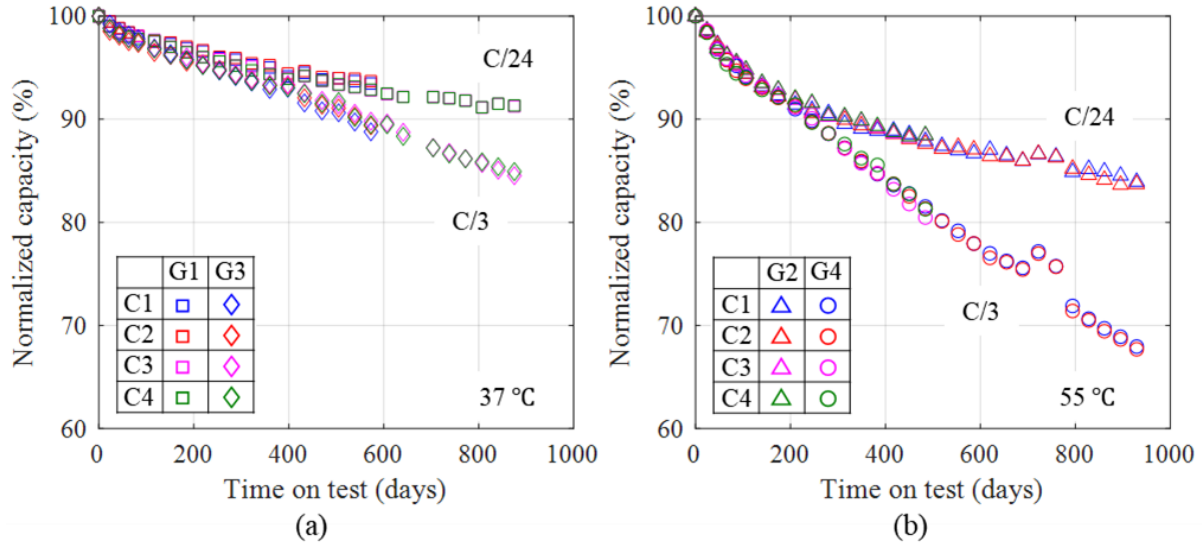


Fig. 1 Capacity fade plots of battery cells tested at (a) 37°C and (b) 55°C. Each cell has anywhere between 17 and 30 data points, depending on whether or not the cell has been removed for destructive analysis [19].

5. RESULTS AND DISCUSSION

5.1 Physics informed machine learning study

Figure 2 shows the health parameters estimation accuracy for models trained on either simulation data only or experimental data only. Despite the simulation dataset encompassing a much wider range of degradation scenarios (as described in subsection 3.2) than the early-life experimental dataset, the models trained only with simulation data performed worse when compared to those trained only with early-life experimental data. This can be attributed to the fact that the half-cell model, being a simplified, largely imperfect physics-based model, cannot accurately reconstruct the voltage curves of the experimental cells especially when the cells have aged heavily. We have discussed such observation in our previous paper [19]. Due to the disagreement between the simulation and the experimental data, the health parameters estimation errors do not possess clear decreasing trends as the amount of training data increases (see Fig. 2(a-d)). In contrast, the estimation errors of the models trained with early-life experimental data show obvious decreasing trends as more experimental data are used for training. However, using 12 or more experimental data points (an approximately 14-month or longer test time) is often not feasible because of the extensive time, money, and manpower required to

perform the tests. In response to this issue, we limit the next set of tests to include only the first five sequential data points from each cell (approximately three months of test time).

Now, with limited experimental data, we aimed to improve model accuracy further by including half-cell model simulation data in the training dataset for each machine learning model. The results are displayed in Fig. 4 and the best results for each model are compared in Table 4. The first five entries in Fig. 4 show the results for the inclusion of random simulation data. We observe that the small quantity of random data points selected from the large simulation dataset generally do not improve model accuracy. In the MOGP and ELM models, we observe that LLI and capacity have improved accuracy, but at the cost of reduced accuracy on m_p and m_n . These results are consistent with the results seen in Fig. 2(a-d), where the models were trained on simulation data only. The simulation data covers a very large degradation space, and when choosing points at random, there is no guarantee the chosen points will reflect the degradation trends observed in the experimental cells. In effect, the random simulation data prove to corrupt the experimental data, making it more difficult for the models to learn the correct mapping from dQ/dV features to late-life degradation parameters.

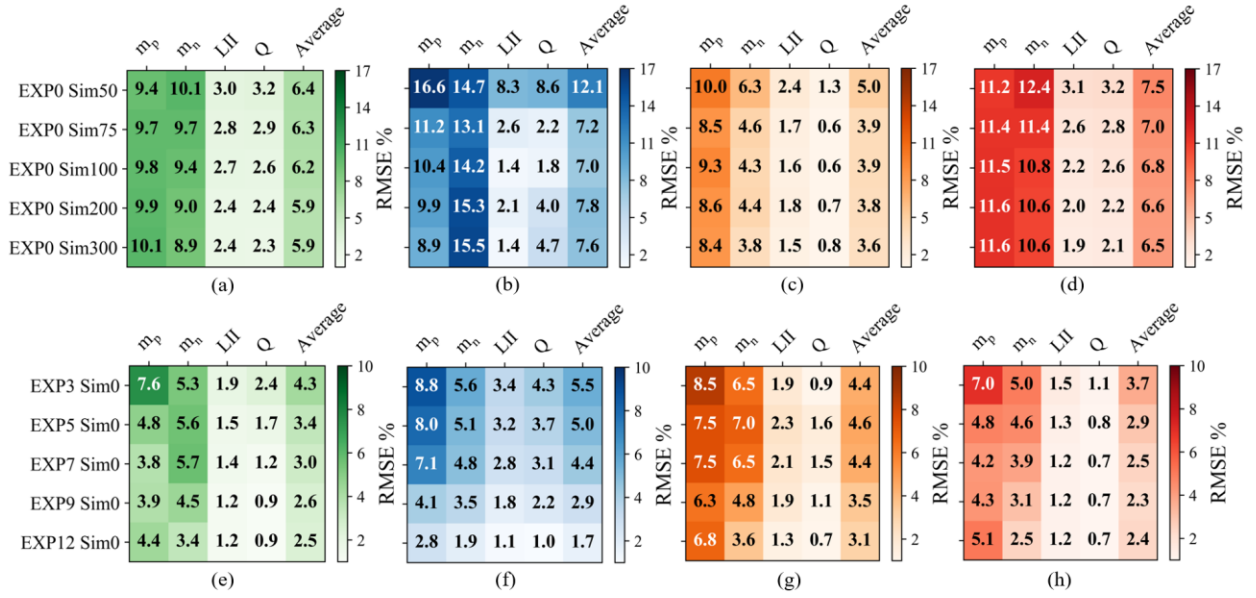


Fig. 2 RMSE % of models when trained exclusively on half-cell model simulation data for (a) Lasso, (b) MOGP, (c) ELM, and (d) elastic net. RMSE % of models when trained exclusively on experimental data for (e) Lasso, (f) MOGP, (g) ELM, and (h) elastic net.

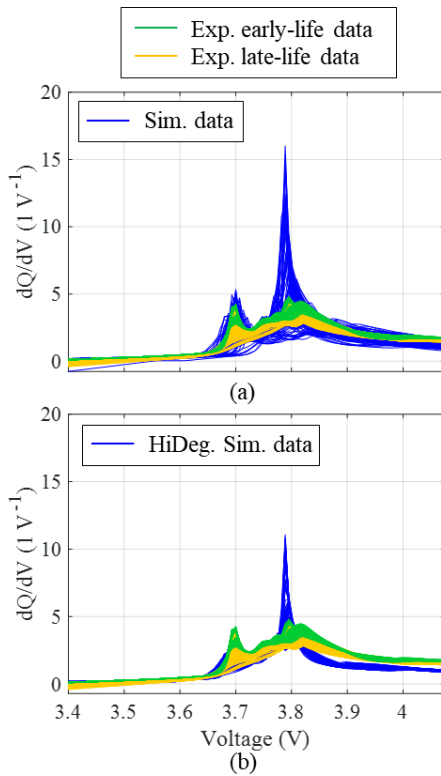


Fig. 3 Comparison of the 30 simulated dQ/dV curves randomly selected from (a) the entire simulation dataset and (b) the high degradation dataset in relation to the early- and late-life experimental data.

Inspired by the way simulation data caused the models to learn an incorrect degradation trend, we investigated whether the models could be coerced into learning a late-life degradation trend that more closely resembled the observed experimental cells. From the large simulation dataset, we filtered to only data points corresponding to the highest 20% degradation in m_p , m_n , and LII . The new subset of simulation data, labelled as HiDeg, were added in small amounts to the early-life experimental data. The results are shown in the next four rows in Fig. 4. It is observed that the HiDeg simulation data improved m_p and m_n estimation accuracy for Lasso, MOGP, and ELM when compared to their respective models without any simulation data. Likewise, MOGP and ELM showed a sizeable increase in LII and capacity estimation accuracy.

The inclusion of HiDeg simulation data into the models' training dataset did not produce consistent and large accuracy improvements across all models. For MOGP and ELM, this method was more effective. In this work, we only considered an MOGP model with a zero-mean trend function. As a result, when estimating late-life cell degradation which is far from the training data, the MOGP model returns to the mean, zero. Including the high degradation simulation data served to extend the training data space, allowing the MOGP model to interpolate between the high degradation simulation data and the early-life experimental data. The same can be said about ELM. Neural networks like ELM are conditioned on the training data distribution and are only intelligent at estimation within the domain of the training data. With the inclusion of high-degradation simulated training data, the ELM is able to bridge the gap between early-life and

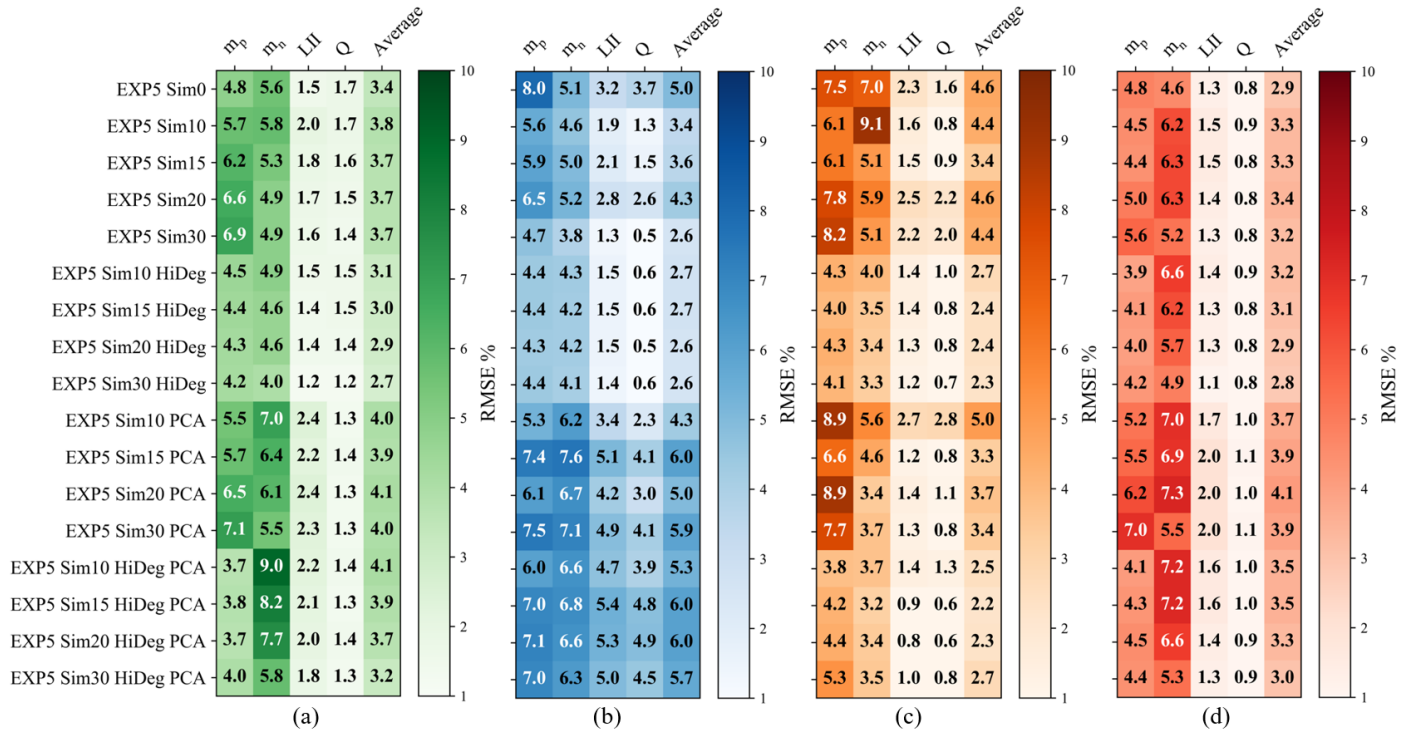


Fig. 4 RMSE % of models trained with different training data combinations for (a) Lasso, (b) MOGP, (c) ELM, and (d) elastic net.

Table 4 Estimation accuracy of best-performing machine learning models.

Training Data	Model	m_p	m_n	LII	Q
EXP5 Sim0	Elastic net	4.77	4.58	1.32	0.76
EXP5 Sim30 HiDeg	Elastic net	4.23	4.86	1.14	0.77
EXP5 Sim0	Lasso	4.82	5.62	1.50	1.69
EXP5 Sim30 HiDeg	Lasso	4.19	4.02	1.17	1.24
EXP5 Sim0	MOGP	8.03	5.09	3.17	3.66
EXP5 Sim30 HiDeg	MOGP	4.35	4.08	1.44	0.57
EXP5 Sim0	ELM	7.53	6.98	2.30	1.63
EXP5 Sim30 HiDeg	ELM	4.06	3.25	1.24	0.74
EXP5 Sim15 HiDeg PCA	ELM	4.20	3.15	0.87	0.62

late-life degradation parameter trends and can more accurately interpolate within the new combined distribution.

As for the Lasso and elastic net models, the method of increasing the training space by adding high degradation simulation data was less effective at improving accuracy. This is likely due the parametric models' innate ability to extrapolate more consistently than the other methods. Using only the first few early-life experimental data points proved to be nearly enough to get acceptable late-life degradation estimates. When adding high-degradation simulation data to the models, the estimation accuracy improved only slightly due to the models' improved understanding of late-life degradation trends.

Last, the results in Fig. 4 indicate PCA was ineffective at increasing model accuracy. Lasso and elastic net already have built-in feature normalization and selection mechanisms, which made PCA less effective. As for MOGP, the model performed worse when the PCA transformation was applied to the data. MOGP likely relied on the un-altered feature values in the dQ/dV curve to learn trends. Reducing the number of input features to 20 principal components likely altered the feature data in a way which made fitting the MOGP model more difficult. ELM was the only model to exhibit similar performance both with and without PCA for feature reduction. However, PCA negligibly increased ELM accuracy.

5.2 Feature reduction study

A key parameter in the generation of half-cell data is the voltage increment between successive dQ/dV samples. To determine whether 100 dQ/dV samples were sufficient for optimal model performance, we varied the input feature size for each model and recorded the average error across the four outputs. Each model was four-fold cross validated 50 times with a mixed dataset, namely EXP5 Sim20 HiDeg. The results are shown in Fig. 5(a) and the dQ/dV feature vectors of different lengths are visualized in Fig. 5(b). Relative to a feature vector length of 100, there was minimal effect on model performance until the number of features was reduced to 25 or less. This indicates that 100 dQ/dV samples were enough to represent the curve's major features (peaks and valleys) and those that are useful for degradation diagnostics. Likewise, using too many dQ/dV samples proved to be counterproductive, and the models had a difficult time determining the most important features.

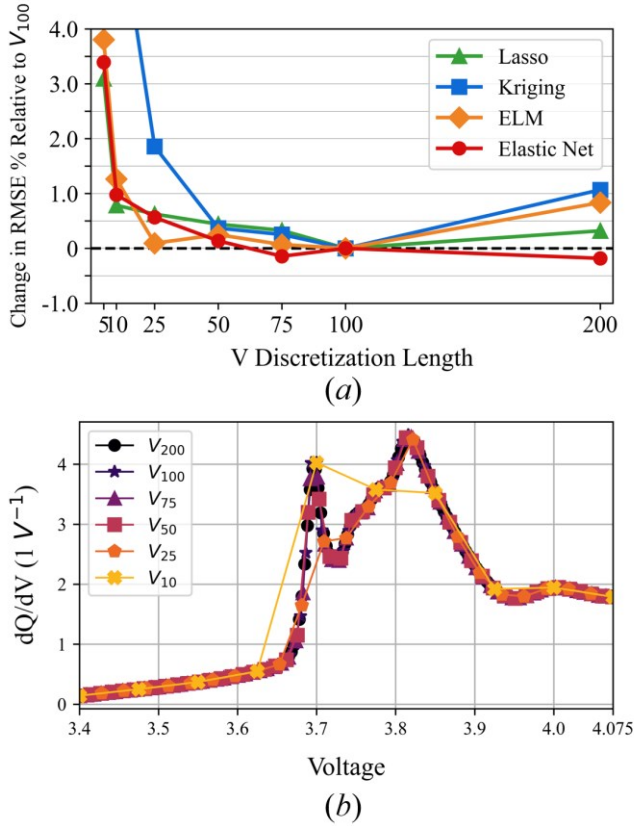


Fig. 5 (a) Effect of dQ/dV discretization length on model accuracy. (b) Visualizations of dQ/dV curves sampled to different lengths.

6. CONCLUSION

This study has demonstrated the possibility of accurately estimating lithium-ion battery capacity and degradation parameters by training a machine learning model using both limited early-life degradation data obtained through cycling tests and high-degradation simulation data from a half-cell model.

The resulting physics-informed machine learning models exhibit improved accuracy compared to models trained strictly on early-life degradation data. The preliminary results in this study suggest that the proposed methodology can significantly reduce the amount of experimental degradation data required for accurate late-life degradation estimation. The best models use only 60 experimental data (five data from each of the 12 training cells) and 30 high-degradation simulation data. In turn, this can reduce the expenses, manpower, and time required to characterize cell degradation in a laboratory setting for the purpose of online degradation diagnostics over a cell's lifetime.

ACKNOWLEDGEMENTS

This research was in part supported by the US National Science Foundation (NSF) Grant Nos. ECCS-1611333 and ECCS-2015710. Any opinions, findings, or conclusions in this paper are those of the authors and do not necessarily reflect the views of the sponsoring agency. The authors would also like to thank Dr. Gaurav Jain and Dr. Hui Ye at Medtronic for providing the implantable-grade Li-ion battery cells used to collect the cycling data for this study.

REFERENCES

- [1] B. Nykvist and M. Nilsson, "Rapidly falling costs of battery packs for electric vehicles," *Nat. Clim. Chang.*, vol. 5, no. 4, pp. 329–332, Mar. 2015, doi: 10.1038/nclimate2564.
- [2] X. Zeng *et al.*, "Commercialization of Lithium Battery Technologies for Electric Vehicles," *Advanced Energy Materials*, vol. 9, no. 27, Wiley-VCH Verlag, Jul. 19, 2019, doi: 10.1002/aenm.201900161.
- [3] B. Dunn, H. Kamath, and J. M. Tarascon, "Electrical energy storage for the grid: A battery of choices," *Science (80-.)*, vol. 334, no. 6058, pp. 928–935, 2011, doi: 10.1126/science.1212741.
- [4] G. L. Plett, "Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs - Part 3. State and parameter estimation," *J. Power Sources*, vol. 134, no. 2, pp. 277–292, 2004, doi: 10.1016/j.jpowsour.2004.02.033.
- [5] S. J. Moura, N. A. Chaturvedi, and M. Krstić, "Adaptive partial differential equation observer for battery state-of-charge/state-of-health estimation via an electrochemical model," *J. Dyn. Syst. Meas. Control. Trans. ASME*, vol. 136, no. 1, pp. 1–11, 2014, doi: 10.1115/1.4024801.
- [6] R. Xiong, L. Li, Z. Li, Q. Yu, and H. Mu, "An electrochemical model based degradation state identification method of Lithium-ion battery for all-climate electric vehicles application," *Appl. Energy*, vol. 219, no. 5, pp. 264–275, 2018, doi: 10.1016/j.apenergy.2018.03.053.
- [7] C. Hu, B. D. Youn, and J. Chung, "A multiscale framework with extended Kalman filter for lithium-ion battery SOC and capacity estimation," *Appl. Energy*, vol. 92, pp. 694–704, 2012, doi: 10.1016/j.apenergy.2011.08.002.

- [8] T. Feng, L. Yang, X. Zhao, H. Zhang, and J. Qiang, "Online identification of lithium-ion battery parameters based on an improved equivalent-circuit model and its implementation on battery state-of-power prediction," *J. Power Sources*, vol. 281, pp. 192–203, 2015, doi: 10.1016/j.jpowsour.2015.01.154.
- [9] A. Braun *et al.*, "The electronic, chemical and electrocatalytic processes and intermediates on iron oxide surfaces during photoelectrochemical water splitting," *Catal. Today*, 2015, doi: 10.1016/j.cattod.2015.07.024.
- [10] S. Shen, M. Sadoughi, M. Li, Z. Wang, and C. Hu, "Deep convolutional neural networks with ensemble learning and transfer learning for capacity estimation of lithium-ion batteries," *Appl. Energy*, vol. 260, no. July 2019, p. 114296, 2020, doi: 10.1016/j.apenergy.2019.114296.
- [11] S. Shen, M. Sadoughi, X. Chen, M. Hong, and C. Hu, "A deep learning method for online capacity estimation of lithium-ion batteries," *J. Energy Storage*, vol. 25, no. March, p. 100817, 2019, doi: 10.1016/j.est.2019.100817.
- [12] M. Bercibar *et al.*, "Online state of health estimation on NMC cells based on predictive analytics," *J. Power Sources*, vol. 320, pp. 239–250, 2016, doi: 10.1016/j.jpowsour.2016.04.109.
- [13] C. Hu, G. Jain, C. Schmidt, C. Strief, and M. Sullivan, "Online estimation of lithium-ion battery capacity using sparse Bayesian learning," *J. Power Sources*, vol. 289, pp. 105–113, 2015, doi: 10.1016/j.jpowsour.2015.04.166.
- [14] D. Yang, X. Zhang, R. Pan, Y. Wang, and Z. Chen, "A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve," *J. Power Sources*, vol. 384, pp. 387–395, Apr. 2018, doi: 10.1016/j.jpowsour.2018.03.015.
- [15] C. R. Birkl, M. R. Roberts, E. McTurk, P. G. Bruce, and D. A. Howey, "Degradation diagnostics for lithium ion cells," *J. Power Sources*, vol. 341, pp. 373–386, 2017, doi: 10.1016/j.jpowsour.2016.12.011.
- [16] X. Han, M. Ouyang, L. Lu, J. Li, Y. Zheng, and Z. Li, "A comparative study of commercial lithium ion battery cycle life in electrical vehicle: Aging mechanism identification," *J. Power Sources*, vol. 251, pp. 38–54, 2014, doi: 10.1016/j.jpowsour.2013.11.029.
- [17] M. Dubarry, M. Bercibar, A. Devie, D. Anseán, N. Omar, and I. Villarreal, "State of health battery estimator enabling degradation diagnosis: Model and algorithm description," *J. Power Sources*, vol. 360, pp. 59–69, 2017, doi: 10.1016/j.jpowsour.2017.05.121.
- [18] J. Tian, R. Xiong, W. Shen, and F. Sun, "Electrode ageing estimation and open circuit voltage reconstruction for lithium ion batteries," *Energy Storage Mater.*, 2021, doi: 10.1016/j.ensm.2021.02.018.
- [19] Y. H. Lui *et al.*, "Physics-based prognostics of implantable-grade lithium-ion battery for remaining useful life prediction," *J. Power Sources*, vol. 485, p. 229327, Feb. 2021, doi: 10.1016/j.jpowsour.2020.229327.
- [20] A. Downey, Y.-H. Lui, C. Hu, S. Laflamme, and S. Hu, "Physics-based prognostics of lithium-ion battery using non-linear least squares with dynamic bounds," *Reliab. Eng. Syst. Saf.*, vol. 182, 2019, doi: 10.1016/j.res.2018.09.018.
- [21] H. Liu, J. Cai, and Y. S. Ong, "Remarks on multi-output Gaussian process regression," *Knowledge-Based Syst.*, vol. 144, pp. 102–121, 2018, doi: 10.1016/j.knosys.2017.12.034.
- [22] G. Bin Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006, doi: 10.1016/j.neucom.2005.12.126.
- [23] H. M. Dahn, a. J. Smith, J. C. Burns, D. a. Stevens, and J. R. Dahn, "User-Friendly Differential Voltage Analysis Freeware for the Analysis of Degradation Mechanisms in Li-Ion Batteries," *J. Electrochem. Soc.*, vol. 159, no. 9, pp. A1405–A1409, 2012, doi: 10.1149/2.013209jes.
- [24] I. Bloom *et al.*, "Differential voltage analyses of high-power, lithium-ion cells 1. Technique and application," *J. Power Sources*, vol. 139, no. 1–2, pp. 295–303, 2005, doi: 10.1016/j.jpowsour.2004.07.021.
- [25] R. Fathi *et al.*, "Ultra High-Precision Studies of Degradation Mechanisms in Aged LiCoO₂/Graphite Li-Ion Cells," *J. Electrochem. Soc.*, vol. 161, no. 10, pp. A1572–A1579, 2014, doi: 10.1149/2.0321410jes.
- [26] K. Honkura, H. Honbo, Y. Koishikawa, and T. Horiba, "State Analysis of Lithium-Ion Batteries Using Discharge Curves," *ECS Trans.*, vol. 13, no. 19, pp. 61–73, 2008, doi: 10.1149/1.3018750.
- [27] L. Lu, X. Han, J. Li, J. Hua, and M. Ouyang, "A review on the key issues for lithium-ion battery management in electric vehicles," *J. Power Sources*, vol. 226, pp. 272–288, 2013, doi: 10.1016/j.jpowsour.2012.10.060.
- [28] C. Hu, H. Ye, G. Jain, and C. Schmidt, "Remaining useful life assessment of lithium-ion batteries in implantable medical devices," *J. Power Sources*, vol. 375, pp. 118–130, Jan. 2018, doi: 10.1016/j.jpowsour.2017.11.056.
- [29] 2020, "Cyclers." <http://www.novonix.ca/cyclers> (accessed Apr. 06, 2020).