Journal of The Electrochemical Society, 2021 **168** 030535

Anode Potential Estimation in Lithium-Ion Batteries Using Data-Driven Models for Online Applications

Jacob C. Hamar,[1,2,z] Simon V. Erhard,[2] Christoph Zoerr,[2] and Andreas Jossen[1]

[1]*Technical University of Munich, Munchen, Baveria 80333, Germany*
[2]*BMW Group, Munchen, Bavaria 80788, Germany*

Three anode estimation methods are presented and evaluated for their accuracy and storage requirements. After generating training data using a Pseudo-2D Physiochemical model, these models are fit and trained to estimate the anode potential during fast charge events. A simplified linear and non-linear model show an estimationerror of ca. 13 mV and the lowest memory demand, however, a novel random forest model reduces the error to 2.6 mV. The empirical methods are suitable for a lithium plating warning detection system during fast charging and are further evaluated for over-fitting and robustness using an out-of-sample dataset.
 [DOI: 10.1149/1945-7111/abe721]

Manuscript submitted December 6, 2020; revised manuscript received January 16, 2021. Published March 22, 2021.

One of the most pressing customer concerns to the adoption of electric vehicles is the expected charging time. Auto manufacturers are developing fast charge protocols to bring the charging time of a 60 kWh battery (equivalent to 320 km) under 30 mins by 2025.[1] A major barrier to reducing the charging time of electric vehicles is the accelerated aging and increased safety risk which are inherent in higher current charging. Aging during fast charge results from multiple factors, including solid electrolyte interface (SEI) growth and the deposition of lithium on the surface of the anode, or lithium plating.[2,3] In order to develop faster charging profiles it is necessary to monitor the state of the battery to prevent the damaging effects from these aging mechanisms.

Of particular concern during fast charge is aging caused by lithium plating which reduces the available lithium for cycling and, in extreme cases, could lead to a short circuit due to the build up of lithium dendrite formations on the anode puncturing the separator.[3–5] Plating is believed to occur when the anode potential, $\Phi_{an,s}$ is reduced to near 0 V against Li/Li^+.[6] The equilibrium anode half-cell potential is decreased during charging as the electrode is lithiated. At the same time, due to diffusion and charge transfer limitations and ohmic loses, the over-potential is also increased, further decreasing the anode potential against Li/Li^+. With a reduced potential against Li/Li^+, there is an increased likelihood that the lithium ions do not intercalate into the anode structure, but instead deposit onto the surface at the electrode-electrolyte interface.[7–9]

It should be mentioned here, that lithium plating is not completely understood.[6] According to recent publications[10] the plating onset condition is likely to be a function of temperature, pressure and concentration. Still there is a thermodynamic dependency between lithium plating and the potential difference between the negative electrode and the electrolyte. Therefore the minimum potential difference—between the solid phase at the anode surface ($\Phi_s(x,\ t)$) and the liquid phase in the electrolyte ($\Phi_l(x,\ t)$) as a function of the electrode thickness (x-dimension) and time—is considered to be the best indicator for the occurrence of lithium plating, where the onset condition is considered to be $\Phi_s(x,\ t) - \Phi_l(x,\ t) \leqslant 0$ V vs Li/Li^+.

In automotive applications, there are no methods for directly measuring if lithium plating is occurring. Capacity loss resulting from plating can only be confirmed during a post-mortem analysis of the cell, where the formation of the lithium deposits can be directly measured. In practice, a reference electrode is inserted between the anode and the separator in the electrolyte in order to measure the potential difference between these two electrodes. The insertion of reference electrodes is expensive and impractical for many commercial applications. for this reason there are currently no sensors available which can reliably detect lithium plating in automobile applications.

One other method would be to build an accurate model of the lithium-ion cell to simulate the anode potential.[11–13] This method however requires considerable effort to parameterize the model and validate the accuracy, and with few exceptions (Sturm et al.[14]), these models are too large to operate on conventional battery management systems. For this reason, manufacturers are looking for alternatives to estimate or predict when plating might occur during a fast charge event using data-driven methods.

Data-driven methods offer an attractive alternative when coupled with sufficiently accurate, large, and representative data sets because they have the potential to accurately estimate the anode potential under a wide variety of conditions. Data-driven methods for modeling battery behavior have recently grown in popularity with contributions focusing on predicting the SOH or state-of-charge of the battery[15–21] It is to the best of the authors knowledge that there currently exists only one publication, by author Xianke Lin[20] which adapts data-driven methods to the prediction of anode potential. In their approach a long short-term (LSTM) algorithm was trained using a P2D model to generate training data from a range of constant current and standard driving profiles. The LSTM is a suitable algorithm because the memory feature in the neural networks is well suited to modeling time-dependent processes in battery dynamics. The goal of this paper is to addresses the assumption that the memory feature of the LSTM is necessary in data-driven approaches and to evaluate a range of suitable regression algorithms specifically in the context of fast-charging to better understand their computational efficiency, accuracy and suitability for online implementation.

In the following sections three methods of varying complexity and computation requirements are presented and evaluated based on their performance estimating the anode potential during fast-charging. The core of this work is detailed below. In this section the data generation procedure, feature selection process, algorithm development, and the validation method will be discussed. The results of the work and their discussion can be found in at the end of the work and, along with the conclusion, will provide a summary and suggestion of the next steps.

Method

In this method, a conventional P2D model is used to generate training and validation data. It is not the goal of this paper to investigate the accuracy of P2D models but rather use a validated model to train data-driven models (See Fig. 1 for an overview of the method). As such, it was assumed that the P2D model used in this paper accurately represents the cell behavior and is suitable for training a data-driven model. More information regarding the P2D model used can be found in Appendix.

[z]E-mail: jacob.hamar@tum.de

P2D model parameterization.—The use of a P2D model offers a reliable estimation of cell behavior including: current, voltage and temperature, as well as other variables which can not be easily measured, such as anode potential. In addition, the P2D model offers a quick and flexible alternative to data generation when compared to laboratory measurements. The P2D model used in this paper is based on the Newman model[11–13] The parameterization set (see Table IV), is derived from a lithium-ion cell with an nickle-manganese-cobalt (NMC) cathode. For a deeper discussion into the equations used in the P2D model see the Appendix or Ennifar et al.[22]

Training data simulation.—The anode potential estimation models were trained and validated on P2D simulation data. This has the significant advantage that data can be collected quickly and cost effectively compared to laboratory measurements. One of the largest drawbacks, however, is if the simulation conditions or model do not adequately capture the cell performance under real-world conditions. To help mitigate any error arising between simulation and real-world conditions it would be advisable to include measurement data along with the simulation data for model training. One difficulty of generating measurement data for model training though is that, in addition to significant time and cost considerations, the anode potential is not easily measured during operation. Due to time and cost restrictions, it was not possible to include validation on measurement data.

The training data generated represents a range of expected operating conditions for fast-charge events which could occur during the normal lifetime of an electric vehicle. The training data matrix is composed of various starting SOC and temperatures which are chosen at intervals broad enough to avoid over-fitting, but also narrow enough to provide accurate and robust predictions. The profiles simulated were based on the multiple-step constant current charging procedure (MSCC). The MSCC is a charging profile designed specifically to avoid conditions inducing lithium plating by reducing heat generation and reducing mechanical stress when lithium diffusion is constrained.[6,23] In order to limit the required amount of training data, only charging (MSCC) conditions were used for training and testing. The simulation-space focused on profiles with high C-rates, as well as temperatures above 15 °C as these are likely profiles to be encountered in automobile applications. By not considering other charging profiles, temperature and currents, the algorithms performance can only be evaluated within the variable ranges provided in Table I.

An overview of the initial conditions for the simulations can be seen in Table I. Each training data point (marked with an 'x' in Table I) consists of five fast charge profiles with varying initial currents, and changes in the step duration of the MSCC profile. These profiles were defined similar to those implemented by charge controllers in the automobile industry. Between the training intervals, out-of-sample validation data was generated in order to test the model performance on data not seen during training. These points are marked with an open circle "o" in in Table I. Interruption data, (marked with an "x" in Table I), were also simulated to further test the robustness of the data-driven models under unlikely operation conditions. These profiles include: linearly increasing current, current disruption, sawtooth current and a rapid current reduction. More detail on these profiles can be found below.

The limits of the training matrix exclude conditions not expected during normal operation, and which would be prohibited by most charging controllers. Initial starting temperatures below 15 °C were excluded due to the severe risk of lithium plating at cooler temperatures. For the specific Li-NMC cell simulated, currents above 2.0 C were considered to be outside of the normal operation range, and were therefore also excluded. Initial currents below 0.9 C and initial SOC above 65% do not qualify as fast charge events for the training of this model. The matrix resulted in 80 fast charge profiles, at four initial temperatures and four initial SOC, with five profiles simulated at each point.

Each simulation has a sampling rate of 1 Hz, resulting in 58033 samples of 7 independent variables and the simulated anode potential (see Table II for a full list of simulation variables).

Feature Selection.—The P2D model outputs several variables which could be used as inputs for a data-driven model (Table II). In this work, the same set of variables were used as independent variables for each of three the data-driven models. This section provides an overview on the variable selection process which includes a correlation analysis and a measure of collinearity.

When building a data-driven regression model, the independent variables must be related to the dependent variable. One of the most widely used indicators of this dependence, or correlation, between two variables, y and x, is the Pearson Correlation Coefficient, r. The Pearson Coefficient measures the degree of which the predictor variable increases (or decreases) as a response to an increase (or decrease) of an independent variable with values ranging from -1 to 1 indicating the strength and direction of the correlation.[24]

$$r = \frac{\Sigma_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{(N-1)\sigma_x \sigma_y} \quad [1]$$

where N is the number of samples, i is the index within the sample set, and $\bar{x}$ and $\bar{y}$ are the average of x and y. The variable σ is the standard deviation,

$$\sigma = \sqrt{\frac{\Sigma_{i=1}^{n}(x - \bar{x})^2}{N-1}}. \quad [2]$$

By plotting the available independent variables from the simulation against the dependent variable, see Fig. 2, and calculating the Pearson Coefficient, a few trends can already be seen. First, there are a number of variables which are highly correlated with the anode potential.

The variables with the highest correlation are: voltage ($r = -0.83$), capacity ($r = -0.64$), OCV ($r = -0.78$), SOC ($r = -0.64$), current ($r = -0.39$), temperature ($r = 0.30$) and time

Table I. Training matrix indicating the initial conditions for the P2D simulations. The P2D simulation generated the data necessary to train and validate the data-driven methods.

		Temperature/°C							
		15	20	25	29	33	38	43	
SOC/%	75	x		x		x		x	
	65		i				o		x—Training
	50	x		x		x		x	o—Validation
	30	i			o			i	i—Interruption
	25	x		x		x		x	
	15		o				i		
	0	x		x		x		x	

Table II. An overview of the output variables from the P2D model which are selected from to use as independent variables for each of the data-driven models.

Variable Name	Variable Description
Cell Voltage	OCV plus over-potentials: ohmic, diffusion resistance and electrode kinetics.
Capacity	Integral of the charging current
OCV	Average equilibrium potential of cathode minus average equilibrium potential of anode.
SOC	Concentration of lithium ions in the cathode related to the maximum concentration.
Temperature	Temperature considered uniform between anode and cathode in the x dimension.
Current	Current at the current collector of the elementary cell.
Time	Simulation time.

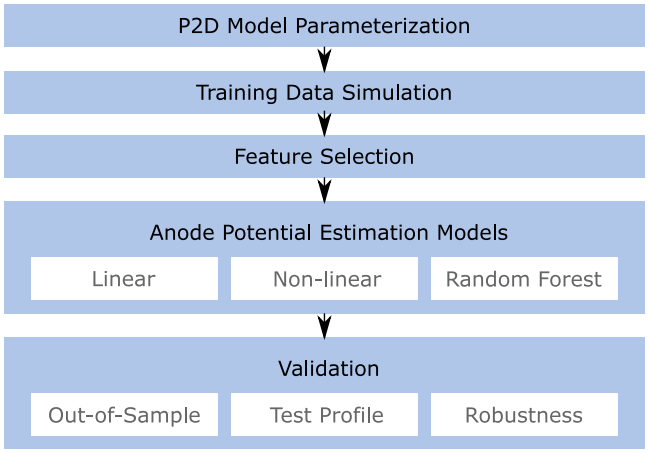

Figure 1. The proposed method includes five steps: (2.1) P2D Model Parameterization- adapting a validated P2D model to a known cell, (2.2) Training Data Simulation- generating data required for training and validating data-driven models, (2.3) Feature Selection- analysis of potential independent variables for regression models, (2.4) Anode Potential Estimation Models- development of multiple data-driven models, and (2.5) Validation- compare models against one another and their suitability for automobile applications

($r = 0.50$). The Pearson Correlation assumes a linear correlation, however, the dependence of anode potential on some of the factors may be better described by a non-linear relationship. The effects of assuming linear dependency is discussed at the end. The statistical hypothesis test, p, is also evaluated with a null hypothesis that the true correlation between the two variables is zero, with p-values less than 0.01 indicating a high probability that the correlation predicted is likely.

Figure 2 also highlights potential issues regarding collinearity between variables. Collinearity means that two or more independent variables are nearly linearly dependent on each other, which leads to the regression coefficients being unstable and sensitive to small random errors.[25] The output variables from the P2D simulation are tested for their correlation with the anode potential, and each other to check for collinearity.

When selecting the independent variables for the regression model the minimum number of factors required to explain the variance in the dependent variable should be selected. When independent variables are highly correlated with themselves there tends to be a large standard errors for the partial regression coefficients of the independent variables, and also a reduction in the statistical significance of both.[25] Removing collinear variables not only improves the stability of the model, it also reduces RAM and CPU required.

Aside from looking at the Pearson Coefficent between the dependent variables, estimating the Variable Inflation Factor (VIF) is the standard method for identifying dependent variables which are collinear within a model.[25] The VIF is calculated from the regression coefficient, r as follows, with larger values indicating a higher instability:

$$VIF = \frac{1}{1 - r_i^2} \quad [3]$$

It was found that when considering all variables in a multiple regression model: voltage ($VIF = 3.5 \cdot 10^2$), capacity ($VIF = 3.5 \cdot 10^6$), SOC ($VIF = -4.4 \cdot 10^2$), and current ($VIF = 3.5 \cdot 10^2$) are collinear, and as such only one should be selected as a factor in the linear regression model. Additionally, the low VIF for temperature ($VIF = 3.2 \cdot 10^0$) and current ($VIF = 3.1 \cdot 10^1$) indicate they describe unique variance in the residual of the dependent variable, and should both be considered.

Taking into account the Pearson Coefficient and the Variable Inflation Factor, three variables were chosen for the use in this study: voltage (U), current (I) and temperature (T). Voltage was chosen over similarly highly correlated factors, such as SOC and capacity because it is readily measurable during vehicle operation.

One omission to the potentially relevant set of predictor variables is the battery state-of-health (SOH). As the battery ages, the development and growth of various aging mechanisms will influence the charging behavior of the battery, for example, more heat will be generated due to higher internal resistance or a sudden roll-over in capacity loss.[26,27] In order to address the influence the change in the electro-chemical behavior has on the anode potential during fast charging, training data over a wider range of SOH could be incorporated, or, as many battery management units also provide an SOH estimation, having the SOH as an input parameter would also help to capture any variance caused by an aging cell.

Anode potential estimation models.—Three data-driven regression models were developed to predict the anode potential from multiple variables. Each method was implemented using R statistical programming software[28] and are described in detail in this section.

Linear model.—As a first approach, a multiple linear regression model was used to predict anode potential. Using three dependent variables: current, voltage and temperature, a simple model was fit on the training data and tested with an out-of-sample data set. Multiple linear regression is a widely applied method to generate fitting functions to data, however, they are most appropriate under a few conditions: the regression model is linear in its parameters, no perfect multicollinearity exists, residuals demonstrate homoscedastitcity and no autocorrelation.[29] Each of these points will be investigated in more detail in the later sections.

The proposed multiple linear regression equation for anode potential estimation is as follows,

$$\hat{\phi}_{an} = \hat{b}_0 + \hat{\beta}_U U + \hat{\beta}_T T + \hat{\beta}_I I \quad [4]$$

x where $\hat{\phi}_{an}$ is the predicted anode potential, U, T and I are independent variables, $\hat{b}_0$ is the y-intercept and $\hat{\beta}_U$, $\hat{\beta}_T$, $\hat{\beta}_I$ are the fitted coefficients for the dependent variables. The linear model in

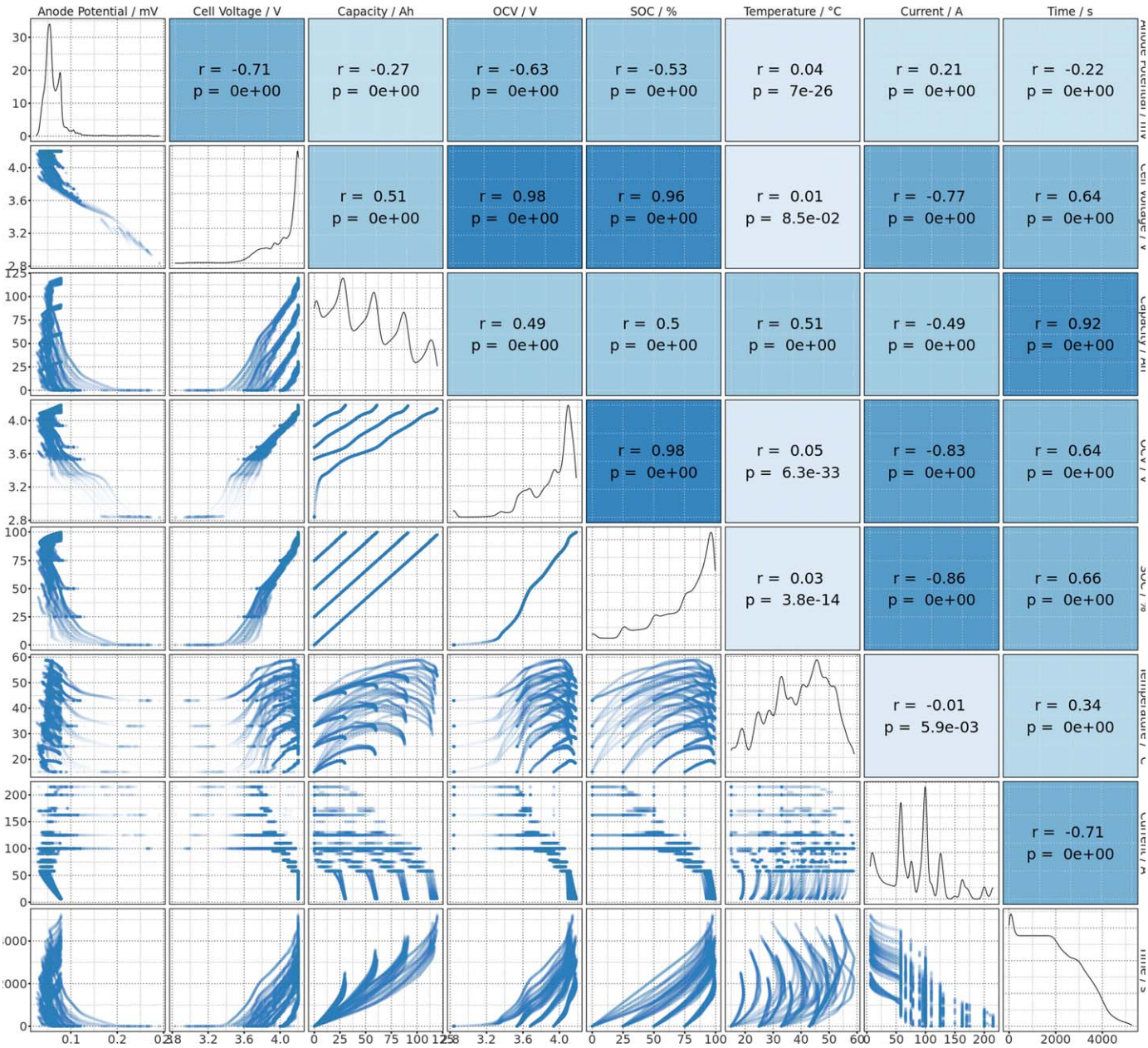

Figure 2. Using all of the result from the P2D model, each variable is plotting against the other to evaluate the correlation and collinearity of the independent variables. The darker the (blue) box, and the higher the r-value, indicate a higher correlation. The p-value is also measured to check the assumption that the variables are somehow dependent on each other.

this work was fit using a least square regression function. Least sum of square models minimize the residual sum of squares function,

$$RSS = \Sigma_{i=1}^{n}(Y_i - \hat{\phi}_{an})^2. \quad [5]$$

The resulting equation has the following regression coefficient fits: voltage ($\beta_U = -170$), temperature ($\beta_T = 0.11$) and current ($\beta_I = -0.44$), with a y-intercept of 796.

Although two parameters, voltage and temperature, show distinct non-linear dependence, the relatively high Pearson Coefficient suggests that there exists sufficient linear dependence to justify this assumption, however, this assumption will be further discussed with the non-linear model.

Non-linear model.—Non-linear models should be used when the independent variables show a non-linearity with respect to the dependent variable or a known coupling between two variables exists. As discussed in the previous section, voltage and temperature do not exhibit purely linear behaviors (see Fig. 2). The following section will introduce a non-linear model for predicting the anode potential and will serve to investigate in more detail the dependence of anode potential on voltage and temperature.

The same independent variables used for the development of the linear model were again used for the non-linear model. This method requires an assumption of the form of the model and is sensitive to the initial parameters. The form of the function was manually determined based on a visual analysis of the independent variable correlation with the dependent variable from Fig. 2.

The suggested model is the superposition of the three non-linear models from the three independent variables:

$$\phi(U) = b_1 + M_1 U + A\sin(UB), \quad [6]$$

$$\phi(T) = b_2 + \frac{C}{T}, \quad [7]$$

$$\phi(I) = b_3 + M_2 I. \quad [8]$$

yielding a combined model,

$$\phi(U, T, I) = b_0 + M_1 U + A\sin(UB) + \frac{C}{T} + M_2 I \quad [9]$$

where b_0 ($b_0 = 797$) is the superposition of all three y-intercepts (b_1, b_2 and b_3), M_1 ($M_1 = 170$) and M_2 ($M_2 = -0.42$) are the slopes of linear equations and A ($A = 2.88$), B ($B = -623$) and C ($C = -115$) are constants, each fitting an independent variable (voltage (U), current (I) and temperature (T)) to the dependent variable ϕ.

The single variable non-linear models in Eq. 9 were chosen to model the dependency between the selected independent variable and the anode potential. The voltage shows an almost linear dependency with a slight sinusoidal perturbation. Temperature shows a reciprocal relationship with the anode potential. Finally, a linear model was chosen to represent the dependence on current as no discernible alternative was seen from the data (see Fig. 2).

While this method attempts to capture the non-linear influence of voltage and temperature, it is still limited by the ability of the chosen functions to mimic the physiochemical behaviors dictating the anode potential. Equations 6–8 are one possible set of candidate functions but there are potentially infinite functions which could yield a more accurate model.

In this approach, the form of the equation must be specified which would lead to a time-consuming process to manually identify the most accurate combination of variables and their exponents. One alternative approach would be to use Principal Component Analysis,[24] which automatically condenses all available variables into a new set, drastically reducing the parameter space for model identification. The focus in this work however, is to introduce the multiple non-linear approach and validate its applicability for online anode potential estimation, and as such a more in depth optimization is left for a future work.

The accuracy of this non-linear model and the comparison against the linear model is continued in below.

Random forest model.—Random forest models are often the go-to algorithm for many machine learning regression problems because of their intuitive implementation, robust accuracy and applicability to a diverse problem-set.[30] The random forest method is chosen in this paper as an example algorithm to highlight the core trends of increasing model accuracy and complexity when switching from linear/non-linear models to machine learning methods. Many common programming languages including R, Python and Matlab have built-in packages to make using a random forest algorithm easily accessible. Other methods, such as neural networks, would also be applicable, however, these methods are often better suited to computer vision tasks and over-complicate simple regression problems. In this section the general theory of random forest models will be explained, as well as, the specific implementation with the P2D simulation data.

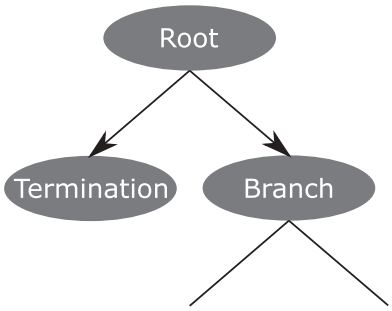

Figure 3. A random fores is comprised of multiple decisions trees which divide random partitions of the training data into branches based on binary splitting conditions which then result in a final value a the termination node.

A random forest is a collection of decision trees which are assigned a random partition of the overall data-set. In this section, the basic principles of a random forest model are discussed; for a more comprehensive review, refer to Murphy et al.[31] A decision tree, shown in Figure 3, stems from the root, splitting into multiple branches and ending at a termination node, or leaf. The entire data-set (or the random partition of the data-set, in the random forest case) starts at the root. The decision tree algorithm then evaluates all possible binary splits in the node which result in the lowest variance of the dependent variable according to

$$\min e(x_i) = \frac{1}{n}\Sigma_i^n (x_i - \bar{x}_i)^2 \quad [10]$$

where e is the residual or variance of the dependent variable x, n is the number of samples in the split and $\bar{x}$ is the average of the variable. The decision tree continues to split into branches with fewer samples in each node until an end criterion is reached-typically a specified number of branches. If there is no split at a node (or leaf), then the there is a value assigned based on the function

$$m(x) = \Sigma_i^l (k_i I(x \epsilon (D_i))) \quad [11]$$

where m is the average, k is a constant, $I(.)$ is a binary operator returning 1 or 0, l is a leaf and D is a randomly selected partition of the original data, which gives a fitted average of all the dependent variable values grouped in the termination node. A forest is then the

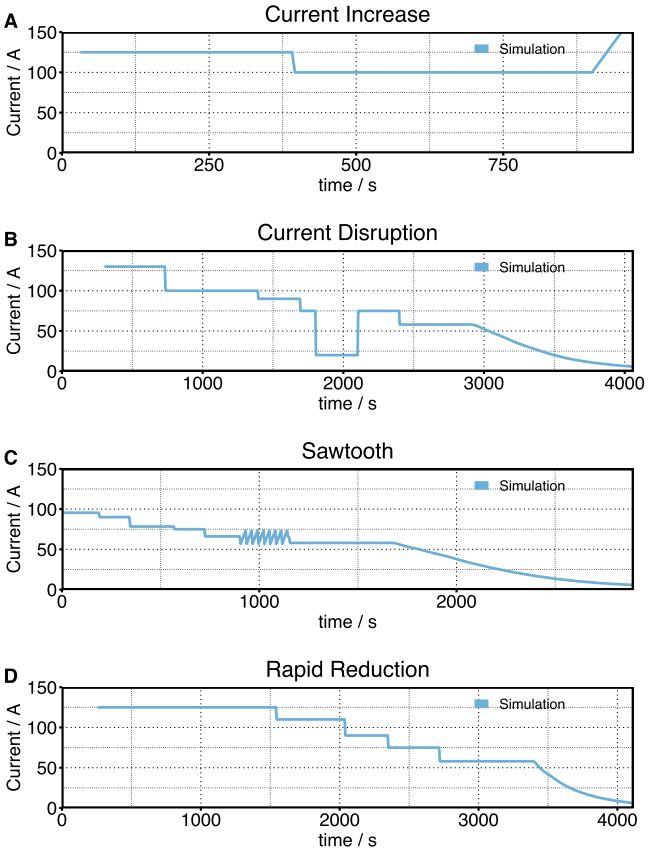

Figure 4. Four current profiles were used to test the robustness of the data-driven models to uncharacteristic charging behaviors. (A) current increase near end of charging, (B) current disruption for 6 minutes during charging, (C) sawtooth wave form and (D) a rapid reduction in steps.

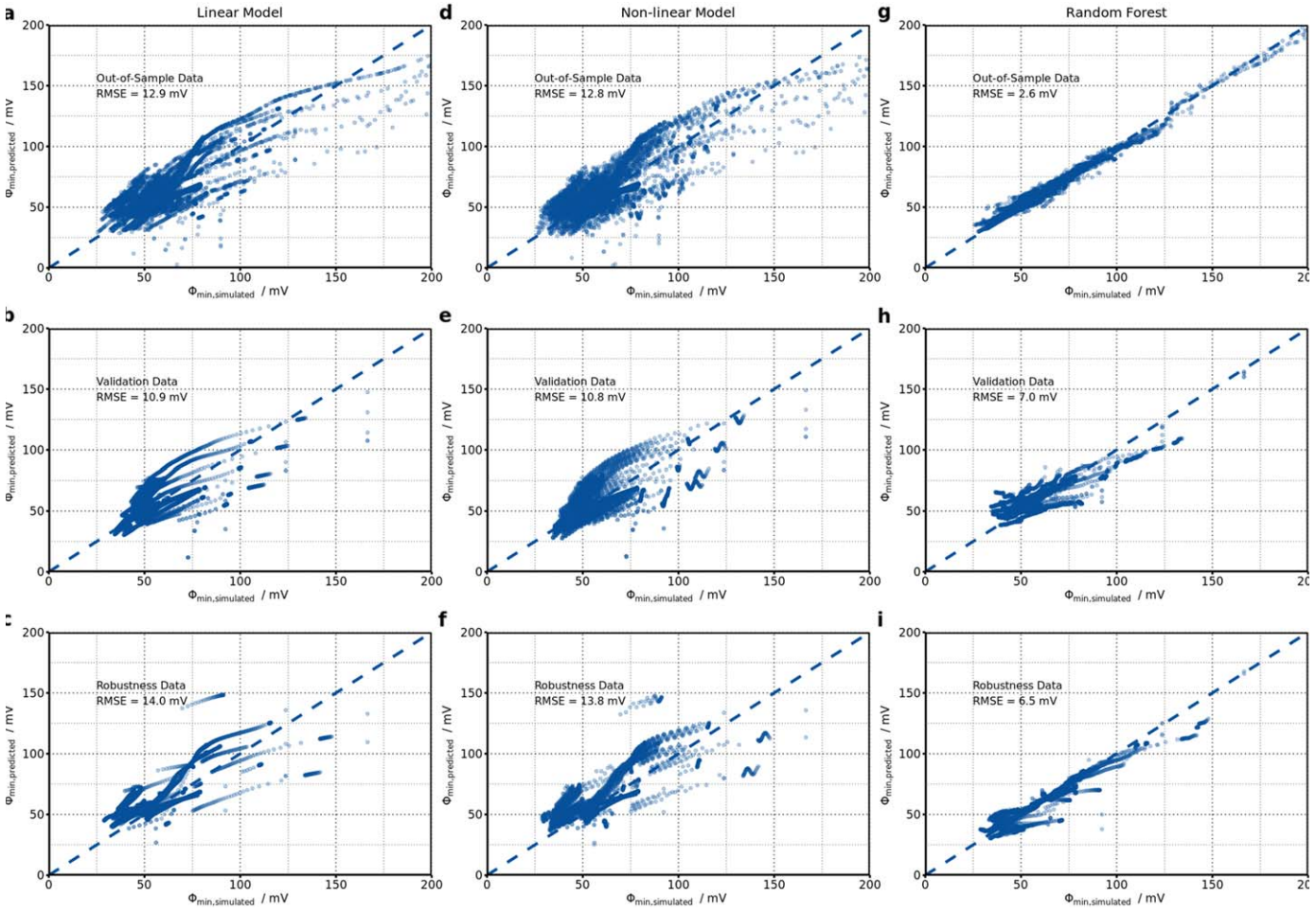

Figure 5. Validation matrix showing the predicted vs simulated anode potential over the three validation data-sets for the three models. (a-c) show the results for the linear model, (d-f) show the results for the non-linear model and (g-i) show the results for the random forest model.

combination of the results from many decision trees, each being initialized with a random partition of the overall dataset. This distributed method helps the algorithm achieve a minimum cost with reduced over fitting to the dataset as well as a reduced computation cost per tree. The number of trees, branches and the cost function are typical hyperparameters which can be optimized for a given problem. For the estimation of the anode potential, an intuitive explanation of how the algorithm works is as follows. First, the training dataset is broken up into multiple random partitions each including measurement samples of voltage, temperature, current and the simulated anode potential. From this first branch, a binary split occurs on one of the input features, for example, all measurements with a temperature below 20 °C would follow the left branch, and all above would follow the right branch. This binary branching will continue following the criterion from Eq. 10 until the designated number of branches is reached. At the end of this tree, there are groups of readings all with similar voltage, temperature and current, and it is from these samples that an average is taken from their simulated anode potential. When a new reading is fed to the random forest model, it will follow the binary split until it reaches the termination node and will pull the estimated anode potential from this leaf. In this respect, a random forest is a highly efficient algorithm at generating a large look-up table.

Validation.—Since no anode potential measurement was available for validation, the ground truth used to evaluate the model accuracy was the anode potential simulation from the P2D model. In order to evaluate the performance of the various models, three groups of out-of-sample data sets were generated along side the training data using the P2D model. In addition to splitting the training data into two groups for training and validating the random forest model, a second validation data set was generated at temperatures and starting SOC which were not included in the training data set for an additional investigation into model performance on unknown data. A third validation set was also generated, again at temperatures and starting SOC outside of the training data set, with charging profiles behaving significantly differently from those used in the training dataset. The initial conditions of these two groups of validation sets can be found in Table I. The profiles are described in more detail below and the results can be seen in the following section.

Out-of-sample.—In this method, all data used for the training of the model was split into two groups. A random selection of 75% (43, 524 samples) was reserved for a new training data-set and the remaining 25% (14, 509 samples) for validation using the base *sample* function in R. Each model was then re-trained with the smaller training data set. This approach helps to identify any autocorrelation in the prediction because the validation data no longer contains any time dependence.

Profiles.—A group of simulations were also run at starting SOC and temperatures which were not included in the training data set (marked as "o" in Table I). These profiles help to test the sparseness of the training matrix to better understand if a finer mesh is needed for accurate prediction.

Robustness.—The last group of profiles are intended to test the response of the prediction algorithms to severe interruptions and

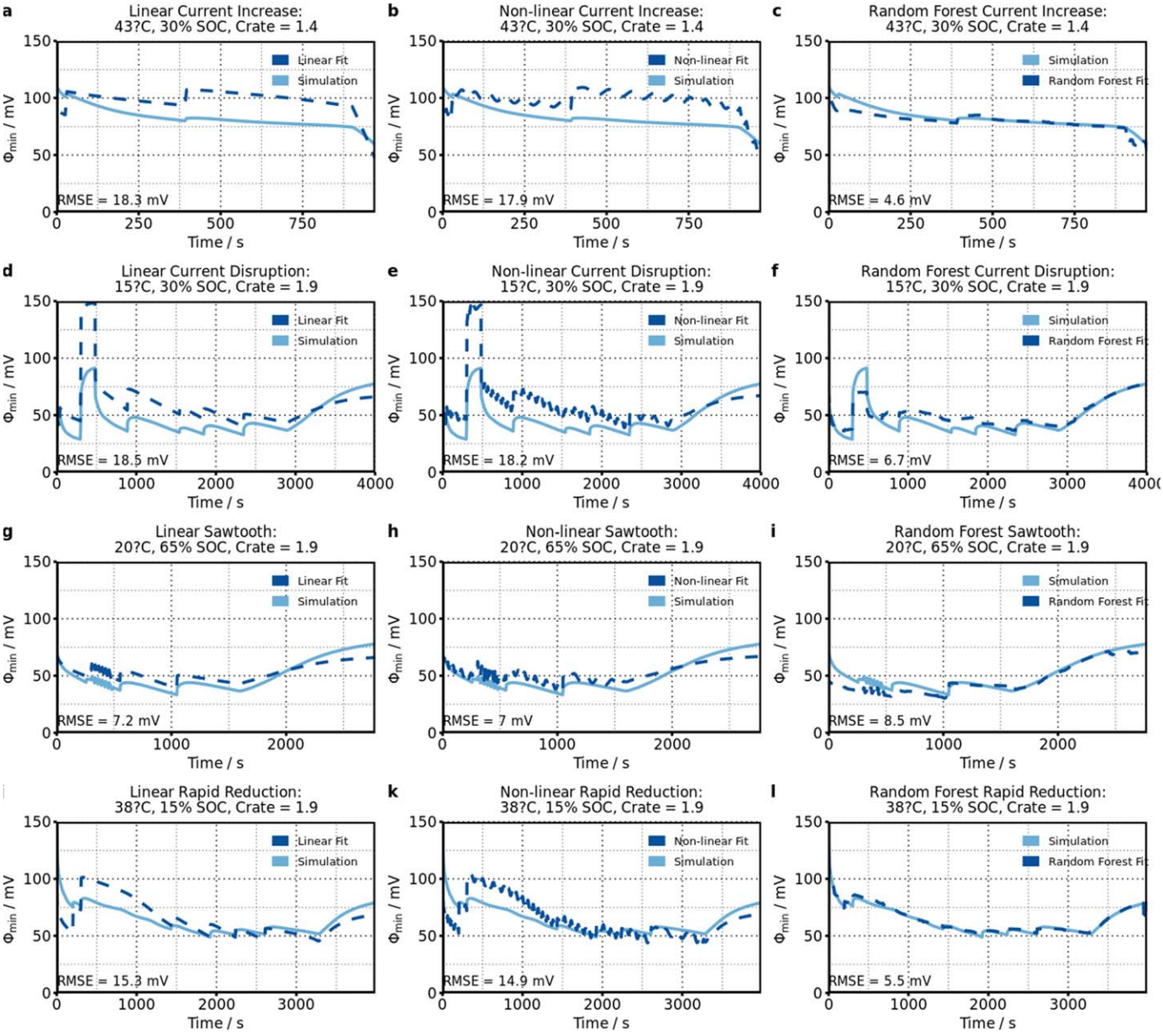

Figure 6. The results of the robustness tests: (a-c) Current Increase, (d-f) Current Disruption (g-i) Sawtooth, (j-l) Rapid Reduction—for each of the tree models: (a,d,g,j) linear model, (b,e,h,k) non-linear model, (c,f,i,l) random forest.

unexpected cell behavior not captured in the training data. These are meant to be worst case scenarios which should be detected and avoided by the charge controller. Four interruption charge profiles were used: current disruption, sawtooth, increasing current, and a rapid reduction. Each of the profiles can be seen in Figure 4. For the current disruption case, current drops to zero. The sawtooth profile has an amplitude of 10 A and a frequency of 1 min^{-1}. The increasing current profile, increases with a slope of 10 $A \cdot min^{-1}$. Finally, the rapid reduction profile switches from between multiple profiles in quick succession. The robustness profiles have various starting SOC and temperatures according to the training matrix in Table I (robustness profiles are marked as "i"). Each disruption lasts for 5 minutes, except for the last case, where the profile simply changes. A total of 16 robustness profiles are simulated.

Results and Discussion

Three models were developed to estimate the anode potential during a fast charge event. Each model was fit using the same data-set generated from the P2D model describe in the Appendix. In order to compare model accuracy, the RSME was calculated for the predicted values against simulation values using three sets of validation data detailed in the previous section. The RSME is calculated by

$$RMSE = \frac{\sqrt{\Sigma_{i=1}^{2}(y_i - \hat{y}_i)^2}}{n}, \quad [12]$$

where y_i is the dependent variable, $\hat{y}_i$ is the prediction, n is the number of samples and i is the measurement index.

The results for each model over each testing method are shown in Fig. 5. The multiple linear regression model has the highest average RMSE over the three validation sets with the highest RMSE (14.0 mV) shown when predicting the robustness profiles. Marginal improvement is seen using the non-linear model with an RMSE value of ca. 0.2 mV improvement compared to the linear model. The lowest RMSE value results when using the random forest model with 500 trees, showing

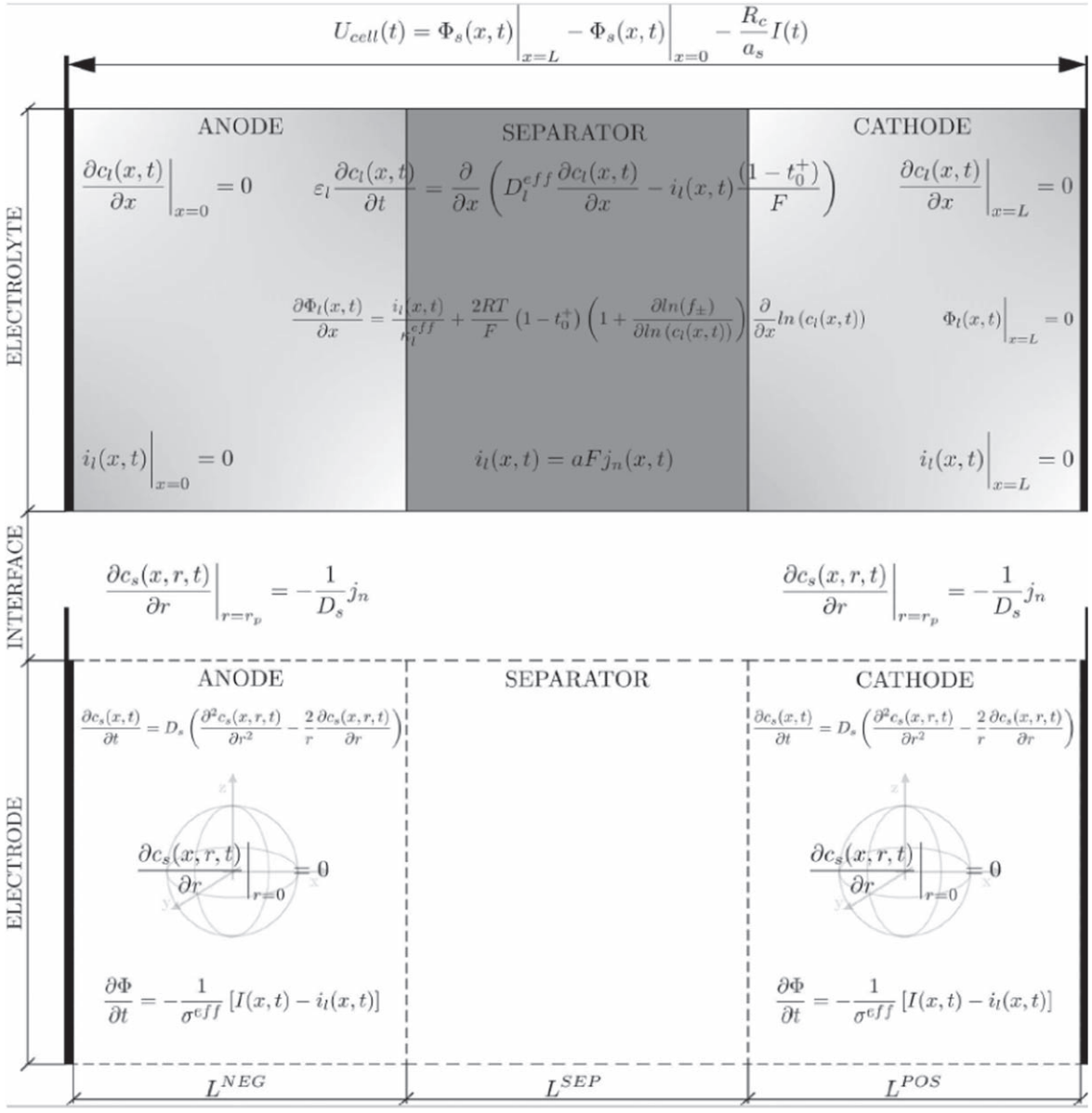

Figure 7. Dimension related equations of the so-called Newman-type P2D model (According to from Ref. 32).

less than 2.6 mV accuracy when tested using the out-of-sample method. The results testing the validation and robustness profiles are similar at 7.0 mV and 6.5 mV, respectively.

The linear and non-linear models show poor prediction accuracy especially at higher anode potentials, which occur at the beginning of a charging event (see Fig. 5). The prediction error was also particularly poor for these two models when considering higher temperatures and lower charge rates as seen in Fig. 6 for the charging profile at 43 ° C and 30% SOC with an initial charge C-rate of 1.4 $h^{-}1$. This is likely due to the fact that these profiles exhibit a relatively high anode potential, where these two models are known to perform poorly, as seen in the RMSE error plots. This is likely due to the fact that the bulk of the training data is in regions with lower anode potentials, where the average anode potential of the training data is 63 mV potentials leading to an under-fitting at higher anode potentials. This discrepancy also results in higher errors at the onset of fast-charging.

For the non-linear and linear fit models, there is higher inaccuracy at the on-set of the charging event (within the first few minutes), with improved accuracy as the voltage stabilizes. The accuracy decreases again during the CV phase. The inaccuracy at anode potentials above ca. 50 mV is most certainly a result of the high cell dynamics within the cell as it responds to a sudden change in condition. This behavior is dominated by short-term ohmic resistance and the ability of the electrode to accept more ions (concentration gradients). As the cell warms and the diffusion reaches equilibrium, the dynamics stabilize.

Since there is little improvement by assuming a non-linear model it could be concluded that the assumption of sufficient linear correlation between the independent variables and the dependent variable is valid. It should also be mentioned that the non-linear model could be potentially improved by choosing a different fitting function, however, capturing the highly irregular behavior of the anode potential with one function would be a daunting task, which is one reason why machine learning methods, such as the proposed random forest model are so powerful, because they are able to model this non-linear behavior with no prior information to this behavior required.

The linear model, and the non-linear model were significantly outperformed by the random forest model. The average error from Fig. 5 does not capture large prediction errors which occur during a few of the simulated fast charge profiles. The maximum error of the random forest model (using 250 trees) on the simulated robustness profiles was

Table III. Training matrix initial conditions for P2D simulation.

Model	RMSE	Training time	Prediction time	RAM
Linear Model	12.9 mV	0.05 s	0.001 s	3 kb
Non-Linear Model	12.9 mV	2.4 s	0.001 s	8 kb
Decision Tree, 1 Tree	22 mV	16 s	0.01 s	1068 kb
Random Forest, 30 Trees	15 mV	41 s	0.08 s	7288 kb
Random Forest, 100 Trees	11 mV	130 s	0.10 s	21416 kb
Random Forest, 250 Trees	5.5 mV	348 s	0.15 s	52034 kb
Random Forest, 500 Trees	2.6 mV	660 s	0.15 s	103382 kb

15 mV, also during the current disruption. As can be seen from the results in Fig. 5, this method yielded highly accurate results, even when considering profiles well outside of the expected use conditions.

The random forest model significantly out-performs either the linear or non-linear models. This increase in accuracy, however, comes with added computation cost. For both the linear and non-linear models less than 10 kb of memory is required compared with 103 Mb for the random forest model with 500 trees. The offline fitting time of the models, as well as the estimated online prediction time are less than 10 ms for the linear and non-linear models. Depending on the desired accuracy and storage capability, the random forest model accuracy improves exponentially whereas the RAM, training time and prediction time increase linearly. A summary of these results can be seen in Table III. As a comparison, an on board P2D model presented by Sturm et al.[14] would require 10 Mb RAM and one minute to compute. The on board model, however, consists of a system of highly non-linear differential equations, each requiring a time consuming order-reduction to be able to run on a micro-processor. An alternative machine learning approach was taken by Lin et al., which results in ca. 3.5 mV error using long short-term memory regression algorithms.[20] Once trained, their algorithm could predict the anode potential, using the same input variables used in this work, with an on-board algorithm requiring 345 Mb.[20] The accuracy of the linear and non-linear models show, however, that a sufficient early warning detection for lithium plating could be implemented with minimal computation costs and modeling effort as it could detect anode potentials nearing a threshold. The low accuracy might render this method ineffective for producing optimized fast-charge protocols which are highly sensitive to the anode potential accuracy.

Conclusions

With a prediction accuracy between 2.6 mV and 12.9 mV the models would be able to alert the battery control module as the anode potential nears 0 mV against Li^+. As plating does not occur instantaneously as the anode potential dips below 0 mV, a warning system with any of the above the proposed anode potential prediction models would be capable of detecting sudden drops or anode potentials within a threshold, and as such it would be sufficient to alert the controller to a risk of plating. As plating is a long-term process, by estimating the amount of time the battery has a low anode potential, the controller can track and estimate the severity of plating over time. The more accurate the estimation is, for example with the random forest estimation, the better the plating accumulation estimation will be. The low computation and storage requirement and estimation time, even from the more complex random forest models, indicate good online applicability. For the non-linear and linear models, the storage requirement of only a few kilobyte means that there should be no difficulty embedding the function of a battery management unite. Although battery management units are used in a wide variety of applications with different capabilities and limitations, the 100 Mb required by the 500 tree random forest would be more challenging to implement in automotive applications as there is a strict limit and competition for available controller space. In order to further improve the linear and non-linear models, the heteroscedasicity of the model should be considered.[29] Heteroscedasicity refers to relationship of the residuals of the independent variables to the dependent variable. An appropriate model will show the same random variation, or noise, across all variables. Typical factors influencing the heteroscedasicity of a model are autocorrelation, missing independent variables and skew in the independent variable distribution.[29]

Further research will include an expanded investigation into the training data used for the machine learning model development, the incorporation of a lag factor to offset minor autocorrelation dependencies and hardware validation of the machine learning algorithm in a real system. Additionally, a study on the applicability of this method for use over the life-time of a battery will also be investigated in including SOH as an independent variable.

Acknowledgments

This work was funded by the BMW Group and was performed in cooperation with the Technical University of Munich.

Appendix: P2D Model

The literature reveals plenty of models to describe the behavior of a Li-ion battery.[33] In this work the generally accepted and extensively discussed Newman-type physiochemical model is used. The so-called pseudo-two-dimensional (P2D) model describes the Li-ion cell on a macroscopic scale (x-dimension) with two porous insertion electrodes, an insulating separator and the liquid electrolyte. The diffusion of Li-ions within the electrode's active material particles is modeled by an additional pseudo-dimension (r-dimension).

The DUALFOIL model, which is a Fortran based version of the P2D model, and is derived from the work of Doyle[11] is taken as a reference for this study. The isothermal model is coupled to a zero-dimensional thermal model to treat the temperature dependency of the reaction kinetics, and the transport parameters accurately. Since the Newman-type model approach is well-known in literature, it will only be given a short introduction to the key equations of the model here. A list of all P2D parameters and thermal model parameters used in this model and their values are given in Table IV and Table V. Thermalparameters assumed for the NMC-graphite cell respectivly. A more precise description is given in the publications of Doyle, Fuller and Newman.[11–13]

In general, the dependent variables of the P2D model are considered to be the Li-ion concentration in the liquid and in the solid phase c_l and c_s, the potential in the electrolyte and in the active material particles Φ_l and Φ_s, the current density in the liquid and in the solid phase i_l and i_s, as well as the ionic flux (pore-wall flux) j_n.

The Li-ion concentration in the liquid electrolyte domain c_l is given by

$$\varepsilon_l \frac{\partial c_l(x,t)}{\partial t} = \frac{\partial}{\partial x}\left(D_l^{eff} \frac{\partial c_l(x,t)}{\partial x} - i_l(x,t)\frac{(1-t_0^+)}{F} \right) \quad [13]$$

with a volume fraction of the active material ε_l, a concentration dependent effective electrolyte diffusivity D_l^{eff}, and a transference number of the cations in the solution t_0^+.

Table IV. P2D model parameterization set for the investigated NMC-graphite cell.

Electrochemical Model Geometry		*Anode*	*Separator*	*Cathode*
Thickness L	m	7.35e-05	1.40e-05	5.55e-05
Particle radius r_P	m	7.00e-06		7.00e-06
Active material volume fraction ε_s	···	0.7164		0.6465
Inert filler volume fraction $\varepsilon_{s,na}$	···	0.0900		0.0840
Porosity ε_l	···	0.2350	0.3900	0.2560
Bruggeman coefficient α_B	···	2.60	2.50	2.10
Thermodynamics				
Stoichiometry at 100 % SOC $\theta_{100\%}$	···	0.8330		0.1890
Stoichiometry at 0 % SOC $\theta_{0\%}$	···	0.0367		0.9300
Maximum lithium concentration $c_{s,max}$	mol m^{-3}	3.11e+04		4.92e+04
Electrode equilibrium potential E_{eq}	V vs Li/Li^+	14		14
Entropic coefficient $\partial E_{eq}/\partial T$	V K^{-1}	14		14
Transport				
Diffusion coefficient in solid D_s	m^2 s^{-1}	1.60e-14		1.00e-13
Diffusion coefficient activation energy E_{a,D_s}	J mol^{-1}	3.00e+04		3.00e+04
Diffusion coefficient in liquid D_l	m^2 s^{-1}		Eq. 26	
Electrical conductivity of the solid matrix σ	S m^{-1}	25.00		0.25
Kinetics				
Reference Reaction rate $k_{a/c}$	m s^{-1}	6.40e-09		1.68e-09
Reaction rate activation energy $E_{a,k}$	J mol^{-1}	6.80e+04		5.00e+04
Transfer coefficient $\alpha_{a/c}$	···	0.50		0.50

Table V. Thermal parameters assumed for the NMC-graphite cell.

Thermal Model		
Density ρ	kg m^{-3}	2354.42
Heat transfer coefficient h	W m^{-2} K^{-1}	6.00
Specific heat capacity C_p	J kg^{-1} K^{-1}	880.00
Surface area A	m^2	7.60e-02
Volume V	m^{-3}	4.20e-04

In contrast the Li-ion concentration in the electrode active material particle domain c_s is calculated by

$$\frac{\partial c_s(x, r, t)}{\partial t} = D_s\left(\frac{\partial^2 c_s(x, r, t)}{\partial r^2} + \frac{2}{r}\frac{\partial c_s(x, r, t)}{\partial r}\right) \quad [14]$$

where r is the radius of the active material particles and D_s is the diffusivity of the solid phase. The x-dimension is coupled to the r-dimension using a boundary condition that applies at the particle's surface. It therefore characterizes the diffusion of Li-ions from the electrolyte into the active material particles. The equation corresponding to the boundary condition is given by

$$\left.\frac{\partial c_s(x, r, t)}{\partial r}\right|_{r=r_p} = -\frac{1}{D_s} j_n(x, t) \quad [15]$$

where j_n is the ionic flux and D_s is the diffusivity at the surface of the idealized active material particles. Considering the diffusivity to be constant reduces the solid-phase diffusion to a linear problem. The equations can therefore be solved using a superposition technique.[34] The DUALFOIL model by default uses the Duhamel Superposition Integral (DSI) method. More detailed information about the DSI is given by Doyle et al.[34]

The Butler-Volmer equation relates the reaction-rate and thus the ionic pore-wall flux j_n to the spatial surface overpotential η. The pore-wall flux j_n is given by

$$j_n(x, t) = \frac{i_0(x, t)}{F}\left[\exp\left(\frac{\alpha_a F}{RT}\eta(x, t)\right) - \exp\left(\frac{\alpha_c F}{RT}\eta(x, t)\right)\right] \quad [16]$$

with the exchange current density i_0 that describes the reversibility of the reaction. It depends on the concentration of reactants and products[13] and on the reaction-temperature. In contrast the transfer coefficients α_a and α_c relate to how an applied potential gradient favors one reaction-direction over the other.

The surface overpotential η expresses the deviation of the electrode potential at the electrode-electrolyte interface from its thermodynamic equilibrium state E_{eq} and is given by

$$\eta(x, t) = \Phi_s(x, t) - \Phi_l(x, t) - E_{eq}(\theta(x, t)) \quad [17]$$

where $\Phi_s - \Phi_l$ is the potential of the electrode compared to an arbitrarily chosen location in the liquid phase. The electrode's equilibrium potential, respectively its open-circuit potential, depends on the stoichiometry θ and therefore on the electrode's state of charge (SOC).

The potential gradient in the electrolyte $\partial\Phi_l/\partial x$ is considered the driving force for the flow of current in the liquid phase. In accordance with the work of Newman[13] it is given by

$$\frac{\partial\Phi_l(x, t)}{\partial x} = -\frac{i_l(x, t)}{\kappa^{eff}} + \frac{2RT}{F}(1 - t_0^+)\left(1 + \frac{\partial \ln(f_\pm)}{\partial \ln(c_l)}\right)\frac{\partial}{\partial x}\ln(c_l) \quad [18]$$

with the effective conductivity κ^{eff} and the activity coefficient $\partial \ln(f_\pm)/\partial \ln(c_l)$ to be considered as functions of the Li-ion concentration in the electrolyte.

The mathematical equation for the potential in the solid phase is directly derived by Ohm's law. In terms of the physiochemical P2D model, Ohm's law relates the current density in the solid phase i_s to the potential gradient in the solid phase. The spatial potential gradient in the electrode domain is therefore given by

$$\frac{\partial \Phi_s(x,\,t)}{\partial x} = -\frac{1}{\sigma^{eff}}[I(t) - i_l(x,\,t)] \quad [19]$$

with the electric conductivity of the porous composite electrode σ^{eff}, and the current density in the solid phase i_s calculated by $I(t) - i_l(x,\,t)$ according to Kirchhoff's law.

The current density in the electrolyte i_l can be derived by Faraday's law. It is therefore related to the pore-wall flux j_n and given by

$$\frac{\partial i_l(x,\,t)}{\partial x} = aFj_n(x,\,t) \quad [20]$$

where a specifies the electrode-electrolyte interfacial area that is calculated by the quotient of the particle's surface and its volume weighted with the volume fraction ε_l.

Since the current density in the electrode domain i_s is related to the potential gradient $\nabla\Phi_s$ by Ohm's law, one can derive the solid phase current density i_s by

$$i_s(x,\,t) = -\sigma^{eff}\frac{\partial \Phi_s(x,\,t)}{\partial x}. \quad [21]$$

Thermal Model

The thermal behavior of the Li-ion cell affects the electrochemical reactions substantially. In consequence, it is essential to take the heat generation and dissipation within the cell into account. Therefore, a simplified energy balance is applied that is derived from the work of Bernardi[35] and given by

$$\rho V C_p\frac{\partial T}{\partial t} = \left(U_0 - V - T\frac{\partial U_0}{\partial T}\right)I + hA(T - T_\infty) \quad [22]$$

where the irreversible heat generation is expressed by the term

$$(U_0 - V)I. \quad [23]$$

Here V is the cell voltage, U_0 is the OCV and I is the applied current. The reversible heat generation that is directly related to the cell's entropy is given by $TI(\partial U_0/\partial T)$. The parameters ρ, V, and C_p characterize the cell's density, its volume, and the specific heat capacity, respectively. The assumed values can be found in Table V.

The heat transfer between the Li-ion cell and the ambient air is given by the term

$$hA(T - T_\infty) \quad [24]$$

where h is the heat-transfer coefficient, A is the intersectional area, and T_∞ is the ambient air temperature.

The zero-dimensional thermal model is coupled to the physiochemical due to a temperature dependency of the diffusion coefficients in the electrode and the electrolyte D_s^{eff}, and D_l^{eff}, the reaction rate constants k_a, and k_c that contribute the exchange current density i_0, the ionic conductivity κ^{eff}, and the activity coefficient $\partial\ln(f_\pm)/\partial\ln(c_l)$.

The temperature dependencies of the solid phase diffusion coefficient D_s^{eff}, and the reaction rate constants k_a, and k_c are modeled by the general Arrhenius equation

$$\Psi = \Psi_{ref}\exp\left(\frac{E_{a,\Psi}}{R}\left(\frac{1}{T_{ref}} - \frac{1}{T}\right)\right). \quad [25]$$

In this context, Ψ specifies the temperature-dependent variable, and Ψ_{ref} is a reference value that was determined at a corresponding reference temperature T_{ref}. $E_{a,\Psi}$ is the activation energy of the process, and T is the temperature given by the zero-dimensional thermal model.

In contrast the temperature dependencies of the liquid phase diffusion coefficient D_l^{eff}, the ionic conductivity κ^{eff}, and the activity coefficient $\partial\ln(f_\pm)/\partial\ln(c_l)$ are given by the following equations:

$$D_l(c_l,\,T) = 10^{-4}\cdot 10^{-4.43 - 0.22\cdot 10^{-3}c_l - \frac{54}{T - 229 - 5\cdot 10^{-3}c_l}} \quad [26]$$

$$\begin{aligned}\kappa(c_l,\,T) = {} & c_l\cdot 10^{-4}(-10.5 + 0.074T - 6.96\cdot 10^{-5}T^2 \\ & + 0.668\cdot 10^{-3}c_l - 0.0178\cdot 10^{-3}c_lT \\ & + 2.8\cdot 10^{-8}c_lT^2 + 0.494\cdot 10^{-6}c_l^2 \\ & - 8.86\cdot 10^{-10}c_lT)^2\end{aligned} \quad [27]$$

$$\begin{aligned}\frac{\partial\ln(f_\pm)}{\partial\ln c_l} = {} & 0.601 - 0.24(c_l\cdot 10^{-3})^{0.5} + 0.982[1 \\ & - 0.0052(T - T_{ref})(c_l\cdot 10^{-3})^{1.5}] \\ & \cdot(1 - t_+)^{-1} - 1.\end{aligned} \quad [28]$$

Note that the parameters in Eqs. 26–28 are adopted from the literature.[36] Since the P2D model parametrization is beyond the scope of this study, it is referred to Valøen et al.[36] for more information about the parameter fitting.

ORCID

Jacob C. Hamar https://orcid.org/0000-0002-0845-3667
Simon V. Erhard https://orcid.org/0000-0002-5029-7477
Christoph Zoerr https://orcid.org/000-0002-6361-2296
Andreas Jossen https://orcid.org/0000-0003-0964-1405

References

1. LLP Deloitte, Deloitte. New market. New entrants. New challenges. Battery Electric Vehicles (2019), https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/manufacturing/deloitte-uk-battery-electric-vehicles.pdf.
2. F. Leng, Z. Wei, C. M. Tan, and R. Yazami, "Heirarchical degradation processes in lithium-ion batteries during aging." *Electrochemica Acta*, **256**, 52 (2017).
3. J. Vetter, P. Novák, M. R. Wagner, C. Veit, K.-C. Möller, J. O. Besendhard, M. Winter, M. Wohlfahr-Mehrens, C. Volger, and A. Hammouche, "Ageing mechanisms in lithium-ion batteries." *Journal of Power Sources*, **147**, 269 (2005).
4. R. Bhattacharyya, B. Key, H. Chen, A. S. Best, A. F. Hollenkamp, and C. Grey, "In situ nmr observation of the formation of metallic lithium microstructures in lithium batteries." *Nat. Mater.*, **9**, 504 (2010).
5. C. Monroe and J. Newman, "Dendrite growth in lithium/polymer systems: A propagation model for liquid electrolytes under galvanostatic conditions." *J. Electrochem. Soc.*, **150**, A1377 (2003).
6. W. Mai, A. M. Colclasure, and K. Smith, "Model-instructed design of novel charging protocols for the extreme fast charging of lithium-ion batteries without lithium plating." *J. Electrochem. Soc.*, **167**, 1 (2020).
7. J. Newman and K. E. Thomas-Alyea, *Electrochemical Systems* (John Wiley and Sons, Inc., Hoboken, New Jersey) 3rd ed. (2004).
8. Z. Li, J. Huang, B. Y. Liaw, V. Metzler, and J. Zhang, "A review of lithium deposition in lithium-ion metal secondary batteries." *Journal of Power Sources*, **254**, 168 (2014).
9. P. Arora, R. E. White, and M. Doyle, *Capacity fade mechanisms and side reactions in lithium-ion batteries*, **145**(10), 3647 (1998).
10. D. R. Baker and M. W. Verbrugge, "Modeling overcharge at graphite electrodes: Plating and dissolution of lithium." *J. Electrochem. Soc.*, **167**, 1 (2020).
11. C. M. Doyle, "Design and Simulation of Lithium Rechargeable Batteries." *Dissertation* (University of California, Berkeley) (1995).
12. T. F. Fuller, "Simulation and optimization of the dual lithium ion insertion cell." *J. Electrochem. Soc.*, **141**, 1 (1994).
13. J. Newman and K. E. Thomas-Alyea, *Electrochemical Systems* (Wiley-Interscience, Hoboken, NJ) (2012), s.l.,3. aufl. edition.
14. J. Sturm, S. Ludwig, J. Zwirner, C. Ramirez-Garcia, B. Heinrich, M. F. Horsche, and A. Jossen, "Suitability of physiochemical models for embedded systems regarding a nickel-rich, silicon-graphite lithium-ion battery." *Journal of Power Sources*, **436**, 1 (2019).
15. D. Liu, X. Yin, Y. Song, W. Liu, and Y. Peng, "An on-line state of health estimation of lithium-ion battery using unscented particle filter." *IEEE Access*, **6**, 40990 (2018).

16. R. Richardson, M. Osborne, and D. Howey, "Gaussian process regression for forcasting battery state of health." *Journal of Power Sources*, **357**, 209 (2017).
17. Y. Zhang, R. Xiong, H. He, and M. G. Pecht, "Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries." *IEEE Transactions on Vehicular Technology*, **67**, 5695 (2018).
18. A. G. Kashkooli, H. Fathiannasab, Z. Mao, and Z. Chen, "Application of artificial intelligence to state-of-charge and state-of-health estimation of calendar-aged lithium-ion pouch cells." *J. Electrochem. Soc.*, **166**, A605 (2019).
19. Kailong. Yi Li, Aoife M. Liu, Alana. Foley, Maitane. Zülke, and Elise. Berecibar, "Nanini-Maury, Joeri van Mierlo, and Harry E. Hoster. Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review." *Renew. Sustain. Energy Rev.*, **113**, 109254 (2019).
20. X. Lin, "Real-time prediction of anode potential in li-ion batteries using long short-term neural networks for lithium plating prevention." *J. Electrochem. Soc.*, **166**, A1893 (2019).
21. K. A. Severson et al., "Data-driven prediction of battery cycle life before capacity degradation." *Nat. Energy*, 1 (2019).
22. H. Ennifar, "Electrochemical State Estimation of a Lithium-Ion Battery: Using a Pseudo-2D Model and an Extended Kalman Filter." *Masterarbeit*, Technische Universität München, München (2017).
23. A. Tomaszewska et al., "Lithium-ion battery fast charging: A review." *eTransportation*, **1**, 100011 (2019).
24. B. Everitt and T. Hothorn, *An Introduction to Applied Multivariate Analysis with R* (Springer, New York, New York, NY) (2011).
25. J. O. Rawlings, S. G. Pantula, and D. A. Dickey, *Applied Regression A Research Tool* (Springer, Berlin) 2nd ed. (1998).
26. A. Raj, M. F. Rodrigues, and D. P. Abraham, "Rate-dependent aging resulting from fast charging li-ion cells." *J. Electrochem. Soc.*, **2020**, 1 (2020).
27. D. Anseán, M. Dubarry, A. Devie, B. Y. Liaw, V. M. Garcia, J. C. Viera, and M. González, "Fast charging technique for high power lifepo4 batteries: A mechanistic analysis of aging." *J. Electrochem. Soc.*, **321**, 201 (2016).
28. R Core Team, R: A language and environment for statistical computing (2019).
29. P. Bruce and A. Bruce, *Practical Statistics for Data Scientists: 50 Essential Concepts* (O'Reilly Media, Inc., Sebastopol, CA) (2017).
30. A. Burkov, *The Hundred Page Machine Learning Book* (Andriy Burkov, Online) (2019), http://themlbook.com/.
31. K. P. Murphy, *Machine Learning: A Probabilistic Perspective* (The MIT Press, Cambridge) (2012).
32. R. Klein, N. A. Chaturvedi, J. Christensen, J. Ahmed, R. Findeisen, and A. Kojic, "State estimation of a reduced electrochemical model of a lithium-ion battery." *IEEE Xplore, 2010 American Control Conference*, 6618 (2010).
33. V. Ramadesigan, P. W. C. Northrop, S. De, S. Santhanagopalan, R. D. Braatz, and V. R. Subramanian, "Modeling and simulation of lithium-ion batteries from a systems engineering perspective." *J. Electrochem. Soc.*, **159**, R31 (2012).
34. C. M. Doyle, T. F. Fuller, and J. Newman, "Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell." *J. Electrochem. Soc.*, **140**, 1526 (1993).
35. D. Bernardi., "A general energy balance for battery systems." *J. Electrochem. Soc.*, **132**, 5 (1985).
36. L. O. Valøen and J. N. Reimers, "Transport properties of lipf6-based li-ion battery electrolytes." *J. Electrochem. Soc.*, **152**, A882 (2005).