

Special  
Collection

# Early Battery Performance Prediction for Mixed Use Charging Profiles Using Hierarchical Machine Learning

M. Ross Kunz,<sup>[a]</sup> Eric J. Dufek,<sup>\*[b]</sup> Zonggen Yi,<sup>[b]</sup> Kevin L. Gering,<sup>[b]</sup> Matthew G. Shirk,<sup>[b]</sup> Kandler Smith,<sup>[c]</sup> Bor-Rong Chen,<sup>[b]</sup> Qiang Wang,<sup>[b]</sup> Paul Gasper,<sup>[c]</sup> Randy L. Bewley,<sup>[b]</sup> and Tanvir R. Tanim<sup>[b]</sup>

A key step limiting how fast batteries can be deployed is the time necessary to provide evaluation and validation of performance. Using data analysis approaches, such as machine learning, the validation process can be accelerated. However, questions on the validity of projecting models trained on limited data or simple cycling profiles, such as constant current cycling, to real-world scenarios with complex loads remains. Here, we present the ability to predict performance with less than 1.2% mean

absolute percent error when trained on cells aged using complex electric vehicle discharge profiles, and either AC Level 2 charge or DC Fast charge profiles, using only the first 45 cycles, namely 5% of the total testing time. While error is low across the projections, this study also highlights that battery lifetime analysis using only cycling data may not extrapolate safely to certain real-world conditions due to the impact of calendar degradation.

## 1. Introduction

The utility of Li-ion batteries continues to evolve. This evolution is driven by trifold advances in life, increased energy content and reduced cost.<sup>[1–3]</sup> While cell-level performance has distinctly improved, methods to understand cell degradation throughout life are still emerging. In some applications such as electric vehicles, there exist standard life and performance evaluation methods including those employed by the United States Advanced Battery Consortium.<sup>[4]</sup> In other areas, such as stationary storage applications, a few procedures for common use cases such as peak shaving and frequency regulation have been implemented.<sup>[5]</sup> Where such methods have been established, life projection typically still requires extensive data sets. For established chemistries and cell designs, it can take 6 to 9 months of aging data to develop sufficient battery life models. For new chemistries and cell designs, obtaining sufficient test data for life estimation can take well over a year and require large experimental designs.<sup>[6,7]</sup> The extended time needed for life evaluations can lengthen the technology development process and subsequently the deployment of

new systems. There is also uncertainty when specific designs are used for new applications or when duty cycles change after deployment of a battery system.

Advanced methods for analysis of battery performance provide the opportunity to significantly shorten the time needed to make predictions about cell performance. Indeed, recent work has shown high fidelity predictions using as few as 100 cycles to achieve life prediction that varied from experimental results by less than 10 percent.<sup>[8]</sup> Aligning advanced data analysis with physical meaning is more difficult, as battery aging is a complex phenomenon that result from diverse physical and chemical mechanisms.<sup>[9]</sup> In part this can be linked to the fact that there is no “universal” physics-based life model.

Life models and prediction techniques can usually be grouped into either data-driven empirical models or physics-based models.<sup>[6,7,10–17]</sup> Physics-based models need to span many different length scales to fully encapsulate specific electrochemical signatures that are linked to battery performance.<sup>[17–19]</sup> Empirical data-driven models are typically not directly tied to physical mechanisms, as such they often do an incomplete job in separating different aging factors,<sup>[9]</sup> including those linked to both cycle and calendar aging.<sup>[20]</sup> This work looks to bridge some of the gaps between life models and advanced statistical methods that can be applied to predict performance at different points in life. In doing so, the work looks to better align physical models with experiments targeting specific features linked to materials and cell performance.<sup>[18,21]</sup> However, battery aging data comes in multiple dimensions, e.g., per cell, per cycle, and per factor, that makes it difficult for application of approaches based on lumped data values. Thus, rather than using summary statistics data that capture singular values for cycles, mean values, or intermittent data collected during reference performance tests (RPTs), here an approaching using physically-targeted, transient time series data is used to train models. These time series include capacity, post-charge volt-

[a] Dr. M. R. Kunz  
Biological & Chemical Processes Department  
Idaho National Laboratory  
Idaho Falls, ID 83415, USA

[b] Dr. E. J. Dufek, Dr. Z. Yi, Dr. K. L. Gering, M. G. Shirk, Dr. B.-R. Chen,  
Dr. Q. Wang, R. L. Bewley, Dr. T. R. Tanim  
Energy Storage & Advanced Transportation Department  
Idaho National Laboratory  
Idaho Falls, ID 83415, USA  
E-mail: eric.dufek@inl.gov

[c] Dr. K. Smith, Dr. P. Gasper  
Energy Conversion and Storage Systems Center  
National Renewable Energy Laboratory  
Golden, CO 80401, USA

**Special Collection** An invited contribution to a Special Collection on Artificial Intelligence in Electrochemical Energy Storage

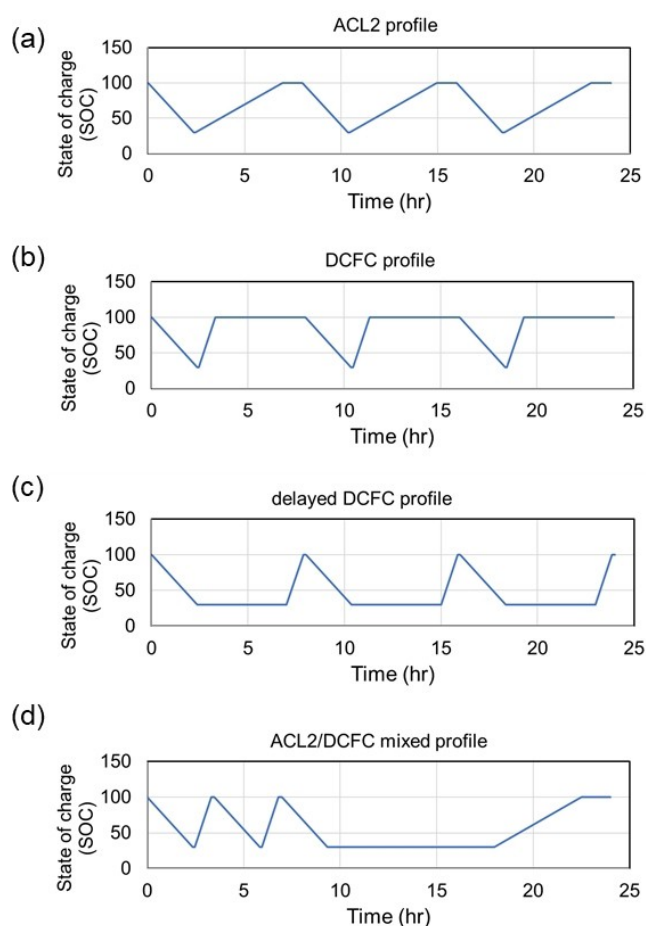
age, and temperature. The proposed methodology in this study may be considered a hierarchical approach. It first estimates the lower dimension representation of each time series per cell that may then be used to predict performance for varied use scenarios per cell. Similar methods have recently been used to shorten the time needed for classification of failure modes using as few as 25 cycles.<sup>[22]</sup>

As more vehicles with fast charge capabilities become available it is increasingly likely that use profiles will incorporate an intermittent mix of both fast and slow charging.<sup>[23–26]</sup> Charging scenarios which have intermittent use vary drastically from most standard testing scenarios and can create complications when predicting life. As described here, using the combination of summary and transient time series data provides the opportunity to predict performance using less than 50 cycles for an assortment of different charging conditions. As new battery use cases emerge this ability to predict from reduced data sets, will be vital to accelerate technology development and for supporting battery controls in advanced energy storage systems.

## 2. Results

**Cell aging protocols and data extraction.** The majority of cells analyzed here were evaluated as described in Tanim *et al.*<sup>[27]</sup> Briefly, 24 graphite/lithium manganese oxide-based pouch cells (33.1 Ah) were obtained from a Nissan Leaf battery pack. These cells were cycled using Maccor 4000 series testers using four different charge profile variants including AC Level 2 (ACL2), DC Fast Charge (DCFC), delayed DCFC, and a combined ACL2 and DCFC charging protocols, as shown in Figure 1. Detailed description about the charging protocols and testing temperatures can be found in the experimental section. All cells were periodically evaluated using an RPT at 30 °C. The RPTs were performed every 32 days (96 cycles). The RPT included hybrid pulse power characterization (HPPC) and C/3 charge-discharge cycle between 4.2 V and 2.5 V. In addition to capturing electrochemical data, the temperature for each cell was acquired throughout the course of the testing. Cells were fixtured using Al plates (~30 kPa) to enhance thermal uniformity during cycling.<sup>[28]</sup>

Selecting appropriate data elements that have utility in physics-guided aging analysis is vital to enhance robustness of predictions. The testing protocol and data acquisition, therefore, needs to enable consistent methods for reproducible data extraction. Other ancillary issues, such as temperature variation during charging, need careful quantification to ensure adequate thermal information is being captured. Lastly, data acquisition according to conditions of convenience which do not require significant secondary processing allow for ready analyses of aging trends since they are not constrained by large lapses in time between data events as seen for RPTs. For the current work, particular interest has been paid to the charging data. To facilitate prediction using the machine learning algorithms, data points for each cell were extracted on a cycle-by-cycle basis. The respective data points extracted included

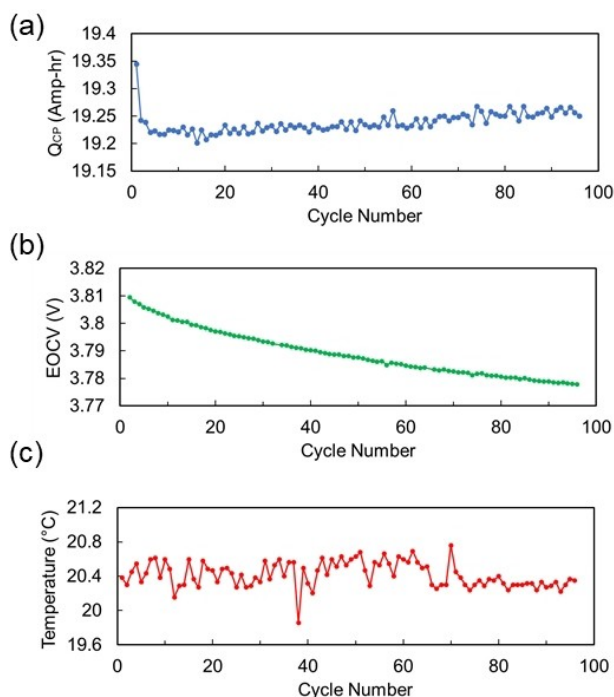


**Figure 1.** Charge protocol variants using (a) ACL2 charging, (b) DCFC charging with the rest at the top of charge (c) delayed DCFC with the rest at a low state of charge (denoted as the delayed DCFC profile), and (d) mixed use of DCFC and ACL2 (denoted as the mixed profile).

the temperature ( $T$ ) at the end of each constant power (CP) component of the charge, the Ah count at the end of each CP charge component ( $Q_{CP}$ ), and the voltage at the end of each post-charge rest period (EOCV). These metrics reflect aspects of cell aging trends, as shown in Figure 2, that can provide indication of the state of the battery electrodes and aging processes without the need for significant data reduction.<sup>[22,29,30]</sup> While not described here, a corresponding set of data can be generated for cell discharge conditions.

**RPT capacity.** To ensure training and test data sets, the results from the various test conditions were split into two groups. Both the ACL2 and DCFC data sets (Figures 3a and 3b, respectively) were used for training purposes, where specific elements from the charge profiles were extracted to highlight key physical features related to cell degradation processes. As shown previously, the dominant aging factors for both DCFC and ACL2 cells are loss of lithium inventory (LLI) and loss of active material at the negative electrode ( $LAM_{LINE}$ ).<sup>[27]</sup>

While the RPT data as provided in Figure 3 is valuable for predictions, the addition of coarse data from each cycle provides enhanced opportunity to capture physically linked



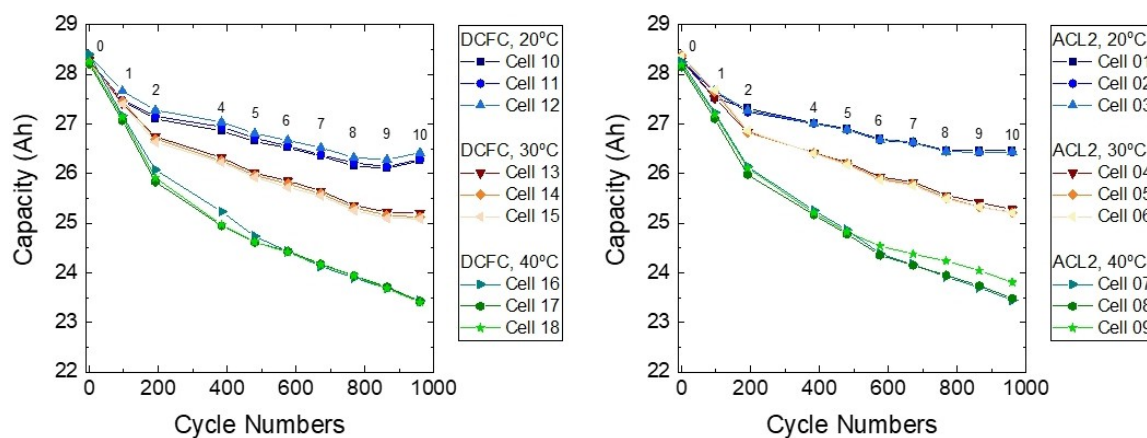
**Figure 2.** An example showing the trend of (a)  $Q_{cp}$ , (b) EOCV, and (c) Temperature changes with cycle number as the aging test proceeds. The data presented here is extracted from the first 96 DST cycles of Cell 03 tested under ACL2 profile at 20°C.

processes. For this reason, data was preprocessed as described above to generate cycle-by-cycle information related to the electrochemical signatures, namely  $Q_{cp}$ , EOCV, and  $T$  (examples in Figure 2). From visual inspection, it is apparent that EOCV and  $Q_{cp}$  have a significant trend with respect to the cycle number where a combination of summary statistics will not suffice in describing the trend. However, the  $T$  value could be described as a random trend with respect to minor environmental chamber fluctuations. For each of the extracted values, the predictor variable is tested for any individual outlier

information within the series as described below and in the statistical methodology section in the Experimental Procedures. If a potential outlier is flagged, then linear interpolation is applied between the nearest data points and the interpolated point replaces the individual point. It should be noted that this technique is not intended for determining if groups of outliers exist and may be improved upon in future work.

The training data consists of a total of 18 different cells that are equally split into groups using either the ACL2 or DCFC charging profile, as showed in Table 4 in the experimental section. The hold-out validation subset is never used in the building of the training model. The validation dataset includes a total of six cells; three delayed DCFC charging profiles and three mixed ACL2/DCFC profiles. The training of the model only leverages a maximum of 96 cycles contained within the first DST cycle set for the prediction of the final RPT capacity. The final predictor matrix,  $X$ , consists of the mean  $T$ , mean EOCV, mean  $Q_{cp}$ , the CV time, and the auto regressive integrated moving average (ARIMA) coefficients for the EOCV and the  $Q_{cp}$  to fully describe the trend with respect to cycle number. This results in the final dimensions of the training predictor matrix to be of 18 cells and 8 variables, i.e.  $X \in \mathbb{R}^{18 \times 8}$ .

**ARIMA regression on the electrochemical signatures.** Before any prediction of the RPT values, all information is preprocessed by performing ARIMA on the EOCV and  $Q_{cp}$  effects. No information is shared between cells, variables or cycles and as such can be performed on individual cells regardless of the information being used in training or in validation of the model. After taking the difference of the series and performing outlier detection, several values of combinations of the number of AR and MA terms are tested with a maximum of three for  $p$  and  $q$ . A consensus based on the ARIMA model Bayesian information criterion (BIC) value between all cells results in selecting a  $p$  and  $q$  to be 1. This allows for direct comparison of the ARIMA coefficients between different variables as well as reduces the potential to overfit the series. To reiterate from above, the AR component is the optimal weight on the previous lag point that describes the series while the MA piece



**Figure 3.** RPT discharge capacity at C/3 for cells cycled using either a DCFC (a) or an ACL2 charge (b) for all charging events. Each of the data points were measured during RPT from RPT 0 to 10, as marked in the graph. RPT0 was measured at the begging of life. Note that RPT 3 data point were not shown due to the use of a different RPT sequence which resulted in a higher C/3 discharge capacity.

is the exponential smoothing of the errors of the self-regression. Table 1 describes the mean of the coefficients for each different charging protocol subgroup. Notably, the ACL2/DCFC combination protocol exhibits a strong exponential effect in the  $Q_{CP}$  compared to the rest of the subgroups. As this subgroup will be part of the validation set, this could potentially make the prediction of the final RPT values difficult. Within the ACL2, DCFC and delayed DCFC groups, the AR coefficients for the  $Q_{CP}$  reveals that the DCFC charge has the largest exponential effect followed by the delayed DCFC and ACL2.

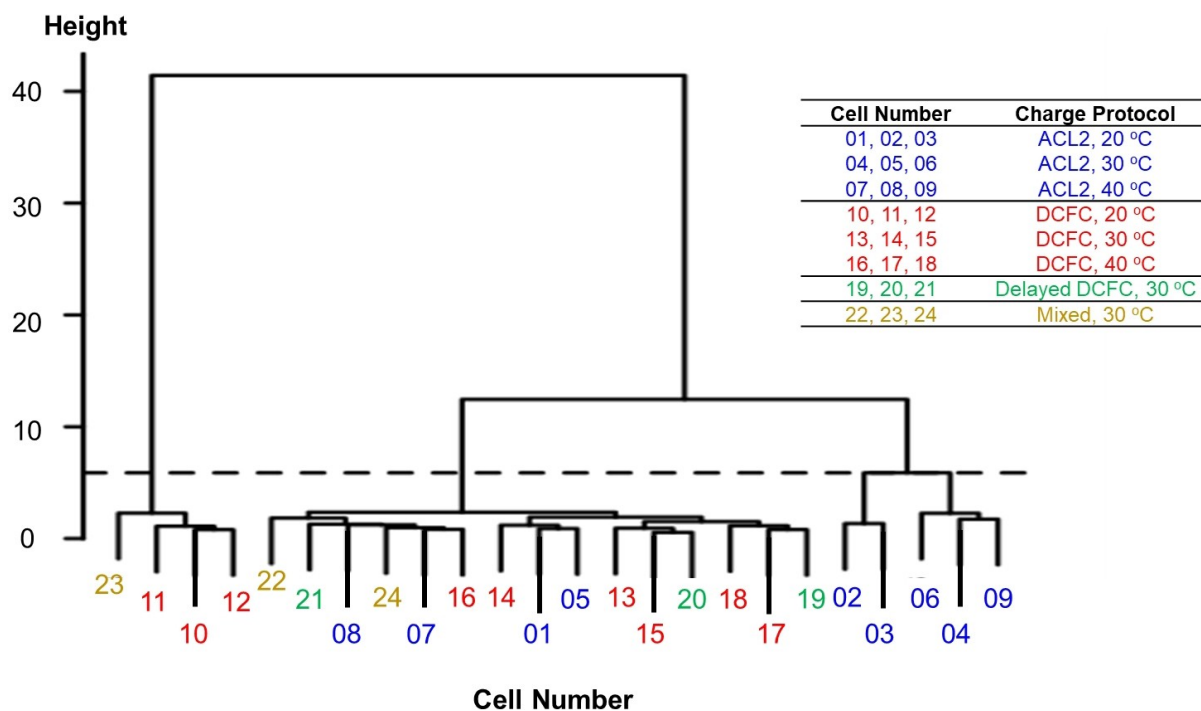
Combining cell data and then performing hierarchical clustering can give some *a priori* knowledge of the different individual cell behavior, as shown in Figure 4. Each branch in Figure 4 is an individual cell (as labeled in Table 1) where the height indicates the distance between clusters. Three different notable outcomes are obtained from the clustering analysis. First, the left-most cluster is composed of cells with a short CV time. As such, the CV time must be considered as a variable within the RPT regression even though it does not relate to the degradation to the cell. Second, all ACL2 cells are in the largest right-hand cluster. While some of the ACL2 cells are clustered

very closely together (Cell 02, 03, 04, 06, 09), the other 4 are dissimilar from the larger cluster and from each other (Cell 05, 01 are similar, and 07, 08 are similar). For the ACL2 cells in these clusters there is no obvious relationship between testing temperature and the clustering, indicating that the lifetime of these cells may be difficult to predict in training. Lastly, all cells that underwent DCFC at  $T > 20^\circ\text{C}$ , including DCFC, delayed DCFC, and mixed DCFC/ACL2, are located in the middle cluster. This indicates that the use of fast charging impacts cell degradation and is a dominant contribution to overall degradation. Thus, even for variations in duty cycle, the application of adaptive Least Absolute Shrinkage and Selection Operator (LASSO) should provide a reasonable prediction. As such, this type of approach may be used in early testing of cells to determine if a new cell is within the training data set sample space. If outside the space, it is feasible to adapt the experiment design to account for the variation in cell design or duty cycle.

**Prediction of RPT capacity.** The prediction of the RPT capacity must first be compared against the set of potential regression results with increasing assumptions on the linear model. Table 2 displays the prediction results of the standard ordinary least squares (OLS) solution, generalized LASSO and the generalized adaptive LASSO. This relative change in the standard OLS solution, compared to the generalized LASSO techniques in the validation dataset, describes the requirement that the capacity must be related to the  $T$ ,  $Q_{CP}$  and EOCV. The correlation is an exponential relationship as described by the physics and modeled here using the ARIMA approach. Furthermore, the additional ability of adaptive LASSO to consistently select the appropriate variables further reduces the

**Table 1.** ARIMA coefficients for the different use cases, showing the AR and MA components.

Charge Group	$Q_{CP,AR}$	$Q_{CP,MA}$	$EOCV_{AR}$	$EOCV_{MA}$
ACL2	−0.32	−0.29	0.67	−0.57
DCFC	−0.45	−0.69	0.82	−0.63
Delayed DCFC	−0.40	−0.75	0.60	−0.71
ACL2/DCFC	−0.99	−1	−0.21	0.08



**Figure 4.** Cluster dendrogram for each of the cell uses evaluated. Each branch is a given cell number, and the height indicates the distance between clusters. Clustering can clearly identify cells with different aging and test considerations.



**Table 2.** Prediction results for the training and validation set of data, in terms of MAPE value, using different analysis methods for the RPT capacity at 864 cycles.

	OLS	Generalized OLS	LASSO	Generalized LASSO	Generalized Adaptive LASSO
Training	0.241 %	0.269 %	0.285 %	0.303 %	0.313 %
Validation	0.664 %	0.618 %	0.527 %	0.414 %	0.408 %

validation error when compared to LASSO. Figure 5 compares the observed and predicted values of the RPT capacity at 864 cycles. The final validation of the delayed DCFC and mixed ACL2/DCFC cells results in a mean absolute percentage error (MAPE) of 0.41 % with a maximum absolute percent error of 1.14 % given in Cell 21, one of the delayed DCFC cells.

### 3. Discussion

**Effect of different number of DST cycles in RPT capacity prediction.** With the appropriate regression model chosen as the generalized adaptive LASSO, the RPT capacity prediction may be compared over different number of DST cycles in training and different RPT cycle counts in validation. This determines the effects of the prediction model when using a limited number of cycles and provides insight into the total amount of experiment time required to make accurate lifetime predictions. Table 3 displays the MAPE according to how many DST cycles were used in the training of the model when predicting the RPT capacity at 864 cycles. The training error is similar, regardless of the number of DST cycles used. However, the predictions using 45 DST cycles and 96 DST cycles resulted in approximately half of the validation error compared to using only 10 DST cycles. As such, a balance of the training and validation error must be considered to determine if overfitting

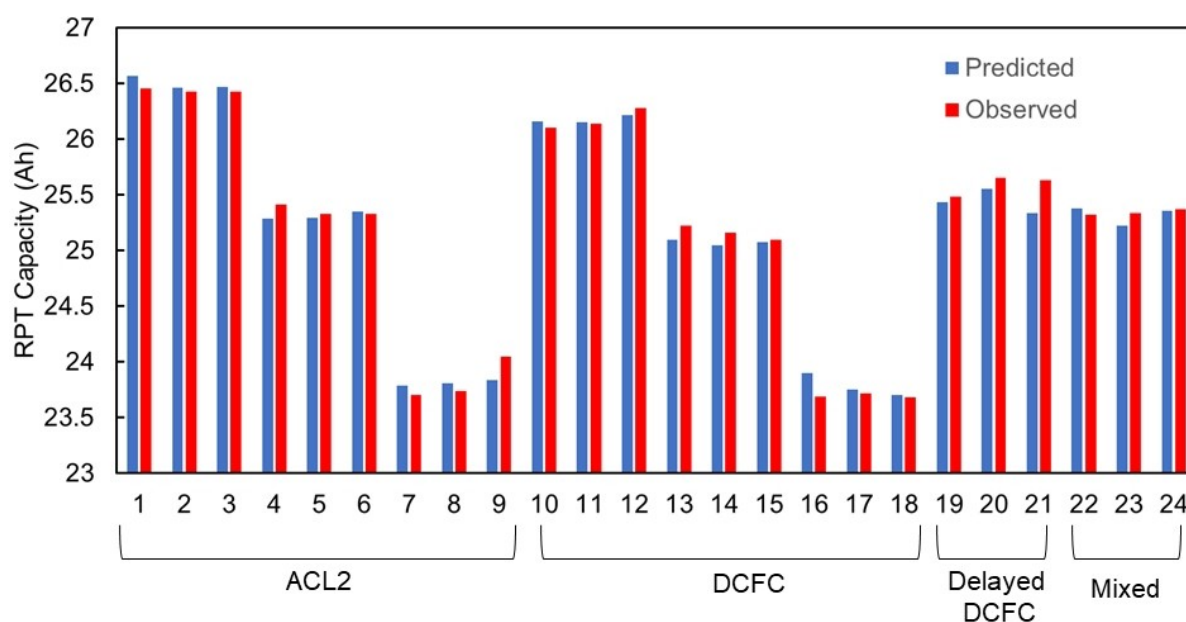
**Table 3.** Training and validation MAPE values for predicting RPT capacity at 864 cycle using only the first 10, 45, and 96 DST cycles.

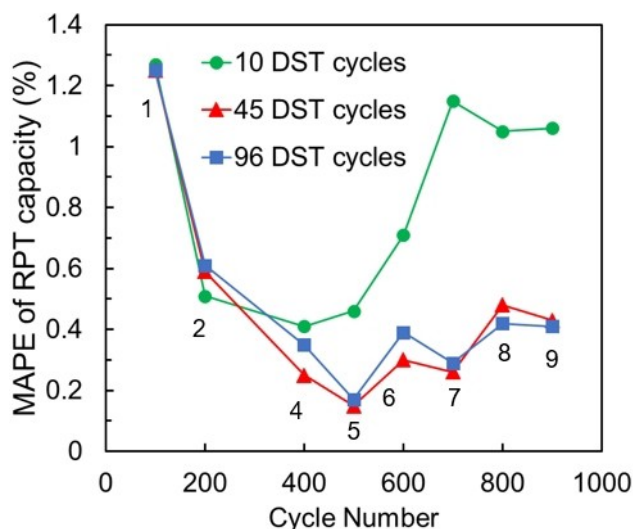
	10 DST cycles	45 DST cycles	96 DST cycles
Training MAPE	0.28 %	0.23 %	0.31 %
Validation MAPE	1.06 %	0.43 %	0.41 %

is occurring. For example, the 10 and 45 DST cycle models may be overfitting the training data resulting in poor performance in validation.

The validation sets used variants of the cycling procedures used in the training sets by either changing the state of charge (SOC) at which the cells rested, or by using a combination of DCFC and ACL2 charging. While the cycling regimes were different, using the more standard cycling profiles in the training sets still enabled good prediction with the average MAPE across both validation sets remaining below 0.5 %. These low MAPE values show the utility of using the adaptive LASSO method to perform predictions on data sets that encapsulate charge procedures not directly involved in data training. This is vital in the long-term prediction of performance as lab-based testing often varies distinctly from actual deployed use cases.

Figure 6 shows the variation of the RPT capacity MAPE for all cells (training and validation) for each of the models using only the first 10, 45, and 96 DST cycles. The predicted capacity value was extrapolated to various cycle counts across 100 to

**Figure 5.** Predicted and actual RPT capacity values at 864 cycles using the first 96 DST cycles for each of the cells in the training and validation data sets. The predicted data are compared to the observed capacity at RPT 9, showing consistency between the two across different cells.



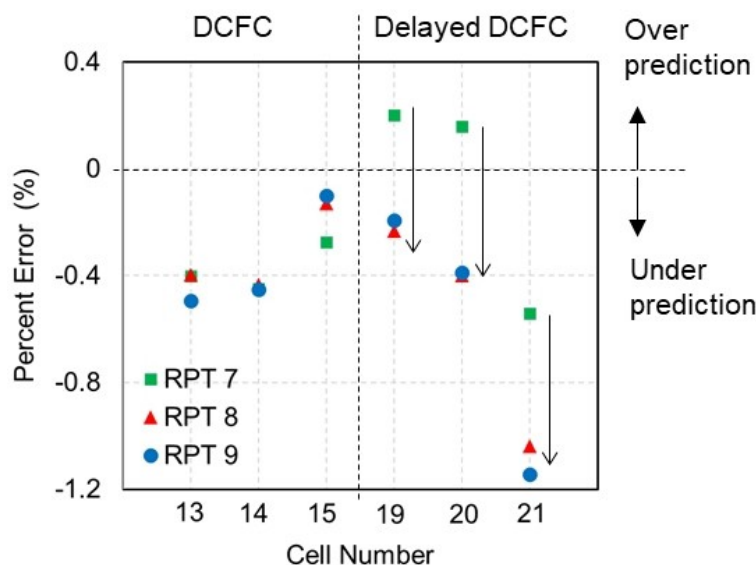
**Figure 6.** The mean absolute percentage error (MAPE) of the predicted RPT capacity for all the cells as a function of the length of DST cycles used for prediction (up to 10, 45, and 60 cycles). The predicted RPT capacity was extrapolated to various cycle counts across 100 to 864 cycles as 9 individual RPTs.

864 cycles. All models perform similarly until cycle 400, at which point the model trained on only the first 10 cycles begins to perform worse. Thus, while short-term degradation behavior can be safely extrapolated from only the first 10 cycles (~3 days) of testing, the long-term degradation behavior cannot be inferred until sometime between the 10<sup>th</sup> and 45<sup>th</sup> cycles. This corroborates some of the findings from Severson et al.,<sup>[8]</sup> which required approximately 60 cycles for model convergence, despite large differences in the cell chemistry, cell format, and testing protocol. This hints at a more general conclusion that battery lifetime may be accurately predicted

from early life (<10% of total lifetime) regardless of cell chemistry or size. When considering the variation in cell behavior and the convolved effects of beginning-of-life and long-term degradation modes, some threshold number of cycles needs to be met before accurate long-term predictions can be made. In practice, the total number of DST cycles may be adjusted for the optimization of the cost of experiment compared to the error of the RPT capacity predictions.

In addition to the experimental resources needed, there is also a need to understand the computational resources required for different analyses. Here, the computation time of building the prediction model with the largest cycle count and over the 9 different RPT responses took a maximum and average computation time of 1.036 and 0.834 seconds, respectively. This was performed on a 2.5 GHz Intel Core i7 without parallelization. This computation time may be decreased through parallelization of the ARIMA component per cell. As such, the proposed methodology does not rely on significant computing resources and large number of samples that may be required in other machine learning techniques such as deep neural networks.

**Prediction error and possible calendar aging effects.** While the variations are minor, differences between prediction and observed variations (Figure 7) suggest that while the overall MAPE is low, the over or under prediction of RPT capacity appears to have a trend with respect to aging of the cell not distinctly related to the active cycling. As an example, when comparing both the DCFC and delayed DCFC cells at 30 °C, the prediction goes from slightly over predicting the RPT capacity fade to underpredicting the fade for the delayed DCFC condition. To further understand the significance of this observation a t-test was performed on the prediction percent error of RPTs 1–5 and RPTs 7–9 for the two conditions. In performing the test, a p-value of 0.002 is obtained indicating



**Figure 7.** Percent error of the predicted capacity from RPT 7, 8, and 9 (at 672, 768, and 864 cycles) for cells 13–15 (DCFC, 30 °C) and 19–21 (delayed DCFC, 30 °C). There is a progressive transition to a more negative percent error between the predicted versus the actual experimental observations in the Delayed DCFC dataset, as indicated by the black arrows.

there is a statistically valid difference between the predictions for the two sets of data. In Figure 7, for the DCFC cells at 30 °C (Cell 13, 14, and 15), there is a relatively random distribution of error when moving from RPT 7 to RPT 9. However, for the delayed DCFC cells at 30 °C (Cell 19, 20, and 21), there is a progressive transition to a more negative percent error between the predicted and actual experimental observations, as indicated by the arrows in Figure 7. This trend toward a larger negative percent error is indicative of a progressively worse underprediction in RPT capacity (overprediction of the capacity fade) as this set of cells age.

A plausible explanation for the distinct trending in the delayed DCFC cells but not in the DCFC cells at 30 °C is that the current framework does not fully capture degradation due to calendar aging. All cells in the training set are held at 100% SOC between cycles, while cells in the validation set are held for long times at 25% SOC, which should lead to greater calendar aging in the training set than in the validation set. During the first 6 RPTs, the impact from this difference in the calendar degradation rate was less prominent. However, during the last 3 months of the cycling experiment between RPT 7–9, the reduced stress from sitting at a lower SOC led to a progressive overprediction of the capacity fade. The modeling framework presented here could likely account for the impact of rest SOC on long-term degradation if the rest SOC was varied within the training set. Creating a more robust prediction will require more direct accounting, especially for LLI, as LLI associated with SEI growth is known to vary distinctly with EOC and T.<sup>[14,31]</sup>

Note that while the trend of increased error during later RPTs for the delayed DCFC is clearly visible, the adaptive LASSO approach still predicts the overall fade within ~1% percent error. This low level of error holds true even under the situation that different cycling regimes and calendar aging conditions between the training and validation sets. For short or intermediate length estimates of life, this level of error would minimally impact prediction purposes. However, as pointed out recently by Gasper *et al.*, even small prediction errors can lead to sizeable variation when projecting to the 10+ year life need for most automotive and stationary applications.<sup>[14]</sup> Thus, future work needs to more effectively capture aspects related to calendar aging variations, and to validate model predictions or

update model parameters with real-world, long-term data whenever possible.

## 4. Conclusion

In this work, we used a hierarchical machine learning approach to predict cell performance to the later stages of cell life for mixed use profiles. A training set of cells that underwent either ACL2 or DCFC cycling was used to predict the discharge capacity of delayed DCFC and mixed DCFC/ACL2 at cycle 864 (RPT 9) to within 1.2% error using only the first 45 cycles (~5% of total test time). These results show the possibility for high fidelity prediction of performance for non-standard cycling conditions using common cycling test sets. The results also show that, while the prediction in cycle life is encouraging, there is a progressively worse underprediction in RPT capacity for the delayed DCFC cells which spent a more significant period of time at low SOC. This is significant as predicting performance for a battery system that may need a 10+ year performance life using models relying solely on cycling cells may be problematic. As approaches such as the one presented here advance, the need to encapsulate combinations of calendar and cycle life fade more fully will be needed. Ultimately, this method will accelerate the transitioning from development to deployment when integrating new battery technologies or deploying battery systems to new use cases.

## Experimental Procedures

**Charging protocol utilized and experimental analysis.** The majority of cells analyzed here were evaluated as described in Tanim *et al.*<sup>[27]</sup> Briefly, the cells were graphite/lithium manganese oxide-based pouch cells (33.1 Ah), obtained from a Nissan Leaf battery pack. These cells were cycled using Maccor 4000 series testers using different charge profile variants including AC Level 2 (ACL2), DC Fast Charge (DCFC), delayed DCFC, and a combined ACL2 and DCFC charging protocols (details in Figure 1 and Table 4). Regardless of charging method each condition was limited to three charging and discharging events per day. All tests were performed in environmental chambers controlled to  $\pm 1$  °C of the set-point for each condition. The respective charging profiles were scaled to the cell level from pack level values of 3.3 kW for ACL2 and 50 kW for DCFC. In the case of ACL2 charging a constant power (CP) of

**Table 4.** Cell numbers, test conditions, and subjected charge protocols. The detailed charge profiles can be found in Figure 2. The cells are split into training and test datasets in the machine learning analysis.

Cell Number	Charge Protocol	T [°C]	Dataset	Total Length of Charge + Rest	Total Length of Discharge + Rest
01, 02, 03	ACL2 <sup>[a]</sup>	20	Training	333 min	147 min (~5 min rest)
04, 05, 06	ACL2	30	Training	333 min	147 min (~5 min rest)
07, 08, 09	ACL2	40	Training	333 min	147 min (~5 min rest)
10, 11, 12	DCFC <sup>[b]</sup>	20	Training	333 min	147 min (~5 min rest)
13, 14, 15	DCFC	30	Training	333 min	147 min (~5 min rest)
16, 17, 18	DCFC	40	Training	333 min	147 min (~5 min rest)
19, 20, 21	Delayed DCFC <sup>[c]</sup>	30	Test	60 min	420 min (~280 min rest)
22, 23, 24	Mixed <sup>[d]</sup>	30	Test	62 min for DCFC 360 min for ACL2	147 min (~5 min rest) for DCFC 662 min (~565 min rest) for ACL2

<sup>[a]</sup> ACL2: AC Level 2 protocol. <sup>[b]</sup> DCFC: DC fast charge protocol. <sup>[c]</sup> delayed DCFC: DCFC charge protocol but with rest period at low SOC. <sup>[d]</sup> Mixed: a mixture of ACL2 and DCFC protocols.

17.2 W was used with a cutoff voltage of 4.11 V followed by a constant voltage (CV) step of 60 min or until the current decayed to a C/20 charging rate. For the DCFC, a CP of 260.4 W was applied until the voltage reached 4.11 V. As with the ACL2 charge a CV step at 4.11 V of 60 min or until a C/20 charge current was met followed the CP charge portion. Discharge was accomplished using a dynamic stress test (DST) discharge with a peak discharge power of 250 W, as described in the USABC Battery Test Manual for Electric Vehicles.<sup>[4]</sup> For this work one DST discharge cycle equated to the net removal of 74.35 Wh. To provide three cycles per day for each condition, tests were constructed so that the total time for each charge plus post charge rest was kept constant. Similarly, the total time for each discharge plus post discharge rest was maintained over the course of the cycling experiment. The total charge + rest and discharge + rest for each protocol are given in Table 4. Keeping a constant time for the combined charge or discharge and rest was used to account for capacity fade as the cells age. Keeping just the rest constant would have progressively lead to more cycles per day.

All cells were periodically evaluated using an RPT at 30 °C. The RPTs were performed every 32 days (96 cycles). The RPT included hybrid pulse power characterization (HPPC) and C/3 charge-discharge cycle between 4.2 V and 2.5 V. In addition to capturing electrochemical data, the temperature for each cell was acquired throughout the course of the testing. Cells were fixtured using Al plates (6.35-mm, ~30 kPa) to enhance thermal uniformity during cycling.<sup>[28]</sup>

## Statistical Methodology

**Dimension reduction via time series analysis.** From the collected set of data, an objective was established to predict the end state RPT capacity solely based on the first 10 to 96 DST cycles. This equates to predictions using the first 1%–10% of the test data. The full evaluation period for these cells was 10 months. Setting the data constraints in this manner allows for early estimates of battery performance for unique use profiles, at a future time. The data constraints also provide a reduction in costs and time associated with validation. As such, the prediction of the RPT capacity may be described as a supervised learning problem with two time series variables, EOCV and  $Q_{cp}$ , as predictors over a series of cycles that are influenced by noise, outliers, and missing information. T is excluded as a time series, as it typically varies randomly around the set experimental T. To account for any outlier and noise issues, the data was passed through a set of statistical preprocessing steps.

To set the prediction of the RPT capacity problem in a mathematical framework, let  $y$  be a vector of RPT capacity values after 864 DST cycles. Let  $X$  be a matrix of a set of predictor variables including the EOCV polarization,  $Q_{cp}$  and T effects for each cell. The addition of the cycle information for these variables creates a tensor structure where the predictors are within three dimensions rather than the matrix structure required for most machine learning algorithms. More concretely, this problem can be stated as:

$$y \in R^{n \times 1}, \quad \chi \in R^{n \times m \times c} \quad (1)$$

where  $y$  is the set of RPT capacity measurements for  $n$  cells and  $\chi$  are the extracted DST values associated with  $m$  variables (EOCV,  $Q_{cp}$ , and T) within  $c$  number of cycles. In contrast, the standard ordinary least squares (OLS) regression model is in the form of:

$$y = X\beta + \varepsilon \quad (2)$$

and is solved as:

$$\min_{\beta} \|y - X\beta\|_2^2 \quad (3)$$

where  $\beta$  are the regression coefficients,  $\varepsilon$  is a vector of independent and identical normally distributed error and  $X$  is a matrix of predictors. While methods currently exist for the application of tensor regression, interpretability with respect to the mechanism may be lost and is outside the scope of this paper.<sup>[32–34]</sup> Here we first discuss dealing with the tensor structure as a dimension reduction problem. The goal is to explain the maximal amount of variance within the DST cycling in a condensed manner, while leveraging the stochastic nature of the data. Then statistical preprocessing techniques are applied to determine if any outliers exist within cycles and between cells. Finally, generalized penalized regression techniques are applied to account for variables that may be unimportant in the capacity prediction and describe the physics-based dependence. The combination of each step provides a framework for the accurate prediction of the RPT capacity at a given number of cycles.

**ARIMA model describe the trends of  $Q_{cp}$ , EOCV, and T.** As noted above, to reduce the dimension of the DST information in a conformable manner, the  $Q_{cp}$  and EOCV with respect to cycle number is treated as a time series of information. As such, the predictors in a single cell deviate with respect to time, i.e. cycle numbers, based on the previous time points and a change in the mean values. For example, Figure 2 describes the variation of the recorded  $Q_{cp}$ , EOCV, and T for each cycle number of a cell (Cell 03) in the first 96 cycles of the cycling test. Visually, the rate of change of the EOCV and  $Q_{cp}$  trends is based on previous time points and could be leveraged to describe the whole series, while the T varies randomly around the mean. These three variables, as time series, can be described using the auto regressive integrated moving average (ARIMA) time series model, which is a mathematical method for measuring the dependence of a series on previous time points.<sup>[35,36]</sup> The ARIMA model is composed of three distinct parts: autoregressive (AR) component explaining the reliance of the current time value on a previous lag, non-stationarity integrated (I) ensuring the mean of the series does not depend on time, and moving average (MA) component describing how much the white noise changes over lag points. More concretely, let  $x_t$  be one of the predictor variables with respect to cycle number and  $w$  designate the noise, then the ARIMA model decomposes the time series into the following components:

- Autoregressive

$$x_t = \phi_{t-1}x_{t-1} + \dots + \phi_{t-p}x_{t-p} + w_t \quad (4)$$

- Integrated

$$x'_t = x_t - x_{t-1} \quad (5)$$

- Moving Average

$$x_t = \theta_{t-1}w_{t-1} + \dots + \theta_{t-q}w_{t-q} \quad (6)$$

where  $\phi$  and  $\theta$  are regression coefficients for each lag respective to the autoregressive and moving average components. Based on the previous example in Figure 2, the  $Q_{cp}$  and the EOCV effects will be heavily affected by all three ARIMA components as the degradation of the battery as the mean of the series depends on time and has an underlying monotonic decreasing effect. The T exhibit polarized random fluctuations with a constant mean between cycles based on minimal fluctuations in the environmental chamber control. The



ARIMA requires stationarity of the series which may be verified through techniques such as the Anderson-Darling (AD) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for seasonal and trend respectively.<sup>[37,38]</sup> For each predictor, the time series can be reduced to a small set of coefficients, i.e.  $p$  and  $q$ , based on the autoregressive and moving average components. The total number of lags used for each  $p$  and  $q$  given the same order of differencing are determined through the use of the Akaike or Bayesian Information Criterion (AIC/BIC) providing an optimal model fit based on the tradeoff of the error and the total number of variables in the model.<sup>[39,40]</sup>

**Outlier detection.** One key assumption to the use of machine learning in general is that there does not exist outliers that will have significant impact on the model. The presence of outliers will result in misleading predictions of the response. The DST cycle information must be pre-processed in way to automatically flag potential outliers within each cycling series and to determine if the cells are consistent within the sample set given. The more robust training set given in the production of the model will dictate the ability to predict future samples. Missing information within the individual cycling of the EOCV and  $Q_{cp}$  is a non-issue as the ARIMA coefficients that describe the series are used as a surrogate in the regression model. However, those coefficients may still be influenced by the outliers within the cycling series. A variety of proven time series outlier detection algorithms exist in methodology and code implementation for the detection of individual and groups of outliers.<sup>[41–43]</sup> In this study, we use the simple yet effective automatic estimation of an outlier through the likelihood ratio test of the studentized residuals of the fit ARIMA model.<sup>[44]</sup> As such, if there exists a data point that is significantly different from the residual distribution, this data point can be flagged as an outlier. Following from the correction of individual cycle information, the set of ARIMA coefficients derived from the DST per cell must then be analyzed for potential outliers for the eventual prediction of the RPT capacity. This information may include individual or groups of outliers, where the use of analysis on the studentized residuals may fail to detect. As such, common unsupervised screening methods, e.g. clustering, may be used to determine which samples are distinctly different from the whole space. Here, we apply hierarchical clustering to measure the similarity between each cell.<sup>[45]</sup> This method specifically uses a minimax distance as the linkage function which provides interpretable clusters as well prototypical cells that best describe a given behavior or charging profile.

**Variable Selection.** With the reduction in dimension and detection of outliers, we treat the coefficients obtained from each  $p$  and  $q$  as a predictor in the analysis of the final RPT value where  $X \in \mathbb{R}^{n \times [4 + 2(p+q)]}$  is now a matrix, i.e. the mean T, mean EOCV, mean  $Q_{cp}$ , charge time and the autoregressive/moving average components for the  $Q_{cp}$  and EOCV. With a small sample size of the number of cells, this still quickly becomes a high dimensional problem with respect to the total number of predictors being greater than the total number of samples. This is a concern as the estimation of the regression coefficients  $\beta$  in Equation 3 relies on the inverse of  $X^T X$  which is singular in high dimensions. Additionally, some of the variation captured within each predictor obtained through the ARIMA model may not have an impact on the final prediction of the RPT. Penalized least squares models extends the nature of Equation 3 to account for multicollinearity between variables and variable selection by introducing norm constraints on the regression coefficients, i.e.,

$$\min_{\beta} \|y - X\beta\|_2^2 + \lambda \|\beta\|_2^2 \quad (7)$$

where  $\|\cdot\|_2^2$  is the penalty on the regression coefficients in the linear model based on the tuning parameter  $\lambda$ . The specific form of Equation 4 refers to ridge regression (RR) where the regression coefficients are penalized by the 2-norm resulting in a reduction in variance of the coefficients and dealing with multicollinearity between the predictors.<sup>[46]</sup> If the 2-norm is replaced by the 1-norm, i.e. the regression coefficients are penalized by the Manhattan distance rather than the Euclidean distance, the result is Least Absolute Shrinkage and Selection Operator (LASSO) regression allowing for exact zeros in the coefficients for selection of the important predictors in the available set.<sup>[47]</sup> With the addition of another tuning parameter and the use of both norms, Equation 4 becomes the elastic-net model.<sup>[48]</sup> The tuning parameter controls the shrinkage of the regression coefficients for either reducing the variance or variable selection. For example, when  $\lambda$  is equal to zero Equation 4 is just the ordinary least squares model in Equation 3 and as the value of  $\lambda$  increases, the coefficients are shrunk towards zero. The optimization of  $\lambda$  is commonly performed using  $k$ -fold cross validation where a set of samples are left out as a validation set and the rest are used to build the regression model.<sup>[49]</sup> This is repeated for each fold over a grid of  $\lambda$  values where the selected  $\lambda$  is based on the minimization of the mean square error in predicting the validation set at each iteration. One drawback of LASSO is the selected predictors may vary based on the random nature of the tuning parameter selection. This may cause slight differences in the coefficients and the selection of predictors especially when highly correlated. Adaptive LASSO is an improvement to the original LASSO through weighting the regression coefficients such that the variable selection is consistent.<sup>[50]</sup> This modifies Equation 4 to

$$\min_{\beta} \|y - X\beta\|_2^2 + \lambda w|\beta|_1 \quad (8)$$

where  $w$  is a vector of weights for each regression coefficient. The values of  $w$  are typically chosen as the transformation of the OLS regression coefficients or the ridge coefficients when the problem is high dimensional, i.e.,

$$\hat{w} = \frac{1}{|\beta_{OLS}|^\gamma}, \quad \text{or} \quad \frac{1}{|\beta_{RR}|^\gamma} \quad (9)$$

where  $\gamma > 0$ . The additional step of calculating the regression coefficients for OLS or RR requires a two-step process but is minimal in computation time for the benefit of providing consistent coefficient estimates.

**Generalized linear models.** Adaptive LASSO specifically highlights the important factors, and those that can be discarded, through an optimized weight associated with the final capacity. However, this model can be enhanced through the inclusion of the capacity physical interpretation in the assumptions of the linear model. Two basic physical assumptions are taken into account with respect to the measured factors: the natural log relationship with the capacity and the polynomial effects or interactions between each measured factor. The former assumption is addressed by including interaction terms within the linear model, i.e., the element-wise multiplication of different factors as a new factor within the predictor matrix. For example, the interaction of T and EOCV may have a specific physical interpretation. In this way, the predictor matrix may be augmented with each polynomial interaction term after the application of the ARIMA model while still holding a linear relationship with respect to the capacity. To handle the assumption of the natural log relationship with the capacity, a transform must take place on the capacity or the measured factors. However, if we

assume the capacity measurements includes normally distributed noise, then direct application of the natural log will also result in the transformation of the noise or error term within the linear model. This conflicts with one of the base assumptions of the linear model where we expect the noise to be independent and identical normally distributed. As such, generalized linear models will be used to extend the basic form of the objective function given in Equation 3.<sup>[51]</sup> The generalized linear model uses a link function  $g(\mu) = X\beta$  to connect the mean of the response,  $\mu$ , to the measured factors. In this specific case, the link function is the natural log changing the original objective function to:

$$\|y - g^{-1}(X\beta)\|_2^2 \quad (10)$$

In this way, assumptions are maintained with respect to the physical relationship of capacity and the linear model for accurate estimates of the regression coefficients.

In summary, Algorithm 1 describes the machine learning framework to provide accurate estimates of the RPT capacity. First, within cycle measurements are tested for any outlier information prior to the application of ARIMA. Following this, hierarchical clustering is used as a determination if a given cell may be outside the model training space. Finally, generalized adaptive LASSO is applied in the prediction of the RPT capacity.

#### Algorithm 1 RPT capacity estimation framework

**Input:**  $x$ : experimental response for  $Q_{cpr}$ , EOCV and  $T$ ,  $y$ : RPT capacity at end of life

1. **for** Each predictor by cycle **do**
2. Treat outliers within time series
3. Application of ARIMA
4. Reform predictors as  $X$
5. Hierarchical clustering on each cell within  $X$
6. Ridge regression on RPT using DST
7. Generalized Adaptive LASSO with RR coefficients

#### Author Contributions

ED, RK – Conceptualization. RK, ZY, KG, ED – Machine learning data analysis. RK, ED, BRC, QW, KG, ZY initial writing. All contributed to discussion of results and in editing the final document. TT, MS, and RB data collection and initial analysis. ED, KS Supervision and funding acquisition

#### Acknowledgements

Research has been supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Vehicle Technologies of the U.S. Department of Energy through the Advanced Battery Development, System Analysis, and Testing program technology managers Samuel Gillard and Simon Thompson. INL is operated by Battelle Energy Alliance under Contract Nos. DE-AC07-05ID14517 for the U.S. Department of Energy. NREL is operated by operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the

United States Government retains a nonexclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. Government purposes.

#### Conflict of Interest

The authors declare no conflict of interest.

**Keywords:** Battery performance prediction • Machine learning • Elastic Net • Calendar aging • Cycle life

- [1] J. Liu, Z. Bao, Y. Cui, E. J. Dufek, J. B. Goodenough, P. Khalifah, Q. Li, B. Y. Liaw, P. Liu, A. Manthiram, Y. S. Meng, V. R. Subramanian, M. F. Toney, V. V. Viswanathan, M. S. Whittingham, J. Xiao, W. Xu, J. Yang, X.-Q. Yang, J.-G. Zhang, *Nat. Energy* **2019**, *4*, 180–186.
- [2] J. E. Harlow, X. Ma, J. Li, E. Logan, Y. Liu, N. Zhang, L. Ma, S. L. Glazier, M. M. E. Cormier, M. Genovese, S. Buteau, A. Cameron, J. E. Stark, J. R. Dahn, *J. Electrochem. Soc.* **2019**, *166*, A3031–A3044.
- [3] B. Nykvist, F. Sprei, M. Nilsson, *Energy Policy* **2019**, *124*, 144–155.
- [4] J. Christophersen, Battery Test Manual for Electric Vehicles, U.S. Advanced Battery Consortium, **2015**, Rev. 3 INL/EXT-15-34184, <https://doi.org/10.2172/1186745>.
- [5] D. R. Conover, A. J. Crawford, J. C. Fuller, S. R. Gourisetti, V. V. Viswanathan, S. Ferreira, D. Schoenwald, D. Rosewater, Protocol for Uniformly Measuring and Expressing the Performance of Energy Storage Systems. **2016** PNNL-22010 Rev. 2 TE1400000, <https://doi.org/10.2172/1249270>.
- [6] K. L. Gering, *Electrochim. Acta* **2017**, *228*, 636–651.
- [7] K. Smith, A. Saxon, M. Keyser, B. Lundstrom, C. Ziwei, A. Roc, *2017 American Control Conference (ACC)*; IEEE, **2017**; pp 4062–4068.
- [8] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggedakis, M. Z. Bazant, S. J. Harris, W. C. Chueh, R. D. Braatz, *Nat. Energy* **2019**, *4*, 383–391.
- [9] C. R. Birkel, M. R. Roberts, E. McTurk, P. G. Bruce, D. A. Howey, *J. Power Sources* **2017**, *341*, 373–386.
- [10] J. M. Reniers, G. Mulder, D. A. Howey, *J. Electrochem. Soc.* **2019**, *166*, A3189–A3200.
- [11] S. Su, H. J. Chen, *IOP Conf. Ser. Earth Environ. Sci.* **2017**, *93*, 012040.
- [12] R. R. Richardson, M. A. Osborne, D. A. Howey, *J. Energy Storage* **2019**, *23*, 320–328.
- [13] X. Hu, L. Xu, X. Lin, M. Pecht, *Joule* **2020**, *4*, 310–346.
- [14] P. Gasper, K. Gering, E. Dufek, K. Smith, *J. Electrochem. Soc.* **2021**, *168*, 020502.
- [15] D. P. Finigan, J. Zhu, X. Feng, M. Keyser, M. Ulmefors, W. Li, M. Z. Bazant, S. J. Cooper, *Joule* **2021**, *5*, 316–329.
- [16] K. Liu, Y. Shang, Q. Ouyang, W. D. Widanage, *IEEE Trans. Ind. Electron.* **2021**, *68*, 3170–3180.
- [17] M. Aykol, C. B. Gopal, A. Anapolsky, P. K. Herring, B. van Vlijmen, M. D. Berliner, M. Z. Bazant, R. D. Braatz, W. C. Chueh, B. D. Storey, *J. Electrochem. Soc.* **2021**, *168*, 030525.
- [18] Y. Liu, B. Guo, X. Zou, Y. Li, S. Shi, *Energy Storage Mater.* **2020**, *31*, 434–450.
- [19] S. Shi, J. Gao, Y. Liu, Y. Zhao, Q. Wu, W. Ju, C. Ouyang, R. Xiao, *Chinese Phys. B* **2016**, *25*, 018212.
- [20] T. Raj, A. A. Wang, C. W. Monroe, D. A. Howey, *Batteries & Supercaps* **2020**, *3*, 1377–1385; *Supercaps* **2020**, *3*, 1377–1385.
- [21] Y. Liu, J. Wu, M. Avdeev, S. Shi, *Adv. Theory Simulations* **2020**, *3*, 1900215.
- [22] B.-R. Chen, M. R. Kunz, T. R. Tanim, E. J. Dufek, *Cell Reports Phys. Sci.* **2021**, *2(3)*, 100352.
- [23] S. Ahmed, I. Bloom, A. N. Jansen, T. R. Tanim, E. J. Dufek, A. Pesaran, A. Burnham, R. B. Carlson, F. Dias, K. Hardy, M. Keyser, C. Kreuzer, A. Markel, A. Meintz, C. Michelbacher, M. Mohanpurkar, P. A. Nelson, D. C. Robertson, D. Scofield, M. Shirk, T. Stephens, R. Vijayagopal, J. Zhang, *J. Power Sources* **2017**, *367*, 250–262.
- [24] A. Burnham, E. J. Dufek, T. Stephens, J. Francfort, C. Michelbacher, R. B. Carlson, J. Zhang, R. Vijayagopal, F. Dias, M. Mohanpurkar, D. Scofield, K. Hardy, M. Shirk, R. Hovsapien, S. Ahmed, I. Bloom, A. N. Jansen, M.

- Keyser, C. Kreuzer, A. Markel, A. Meintz, A. Pesaran, T. R. Tanim, *J. Power Sources* **2017**, 367, 237–249.
- [25] T. R. Tanim, E. J. Dufek, M. C. Evans, C. Dickerson, A. N. Jansen, B. J. Polzin, A. R. Dunlop, S. E. Trask, R. Jackman, I. Bloom, Z. Yang, E. Lee, *J. Electrochem. Soc.* **2019**, 166, A1926–A1938.
- [26] A. M. Colclasure, T. R. Tanim, A. N. Jansen, S. E. Trask, A. R. Dunlop, B. J. Polzin, I. Bloom, D. Robertson, L. Flores, M. C. Evans, E. J. Dufek, K. Smith, *Electrochim. Acta* **2020**, 337, 135854.
- [27] T. R. Tanim, M. G. Shirk, R. L. Bewley, E. J. Dufek, B. Y. Liaw, *J. Power Sources* **2018**, 381, 56–65.
- [28] R. L. Bewley, E. J. Dufek, S. E. Egan, D. K. Jamison, C. Ashton, C. Ho, M. C. Evans, T. L. Bennett, *J. Electrochem. Soc.* **2019**, 166, E263–E265.
- [29] P. P. Paul, E. J. McShane, et al. *Adv. Energy Mater.* **2021**, 11(17), 2100372.
- [30] M. Dubarry, B. Y. Liaw, *J. Power Sources* **2009**, 194, 541–549.
- [31] L. Kolzenberg, A. Latz, B. Horstmann, *ChemSusChem* **2020**, 13, 3901–3910.
- [32] X. Zhang, L. Li, *Technometrics* **2017**, 59, 426–436.
- [33] H. Zhou, L. Li, H. Zhu, *J. Am. Stat. Assoc.* **2013**, 108, 540–552.
- [34] R. Guhaniyogi, S. Qamar, D. B. Dunson, *J. Mach. Learn. Res.* **2017**, 18, 2733–2763.
- [35] R. H. Shumway, D. S. Stoffer, *Time Series Analysis and Its Applications*; Springer, **2016**.
- [36] R. H. Shumway, D. S. Stoffer, *ARIMA Models*, Springer, 75–163, **2017**.
- [37] T. W. Anderson, D. A. Darling, *Ann. Math. Stat.* **1952**, 23, 193–212.
- [38] D. Kwiatkowski, P. C. B. Phillips, P. Schmidt, Y. Shin, *J. Econom.* **1992**, 54 (1–3), 159–178.
- [39] H. Akaike, *Information Theory and an Extension of the Maximum Likelihood Principle*, Springer, 199–213, **1998**.
- [40] G. Schwarz, *Ann. Stat.* **1978**, 6, 461–464.
- [41] Y. She, A. B. Owen, *J. Am. Stat. Assoc.* **2011**, 106, 626–639.
- [42] W. W. Wei, *Multivariate Time Series Analysis and Applications*; John Wiley & Sons, **2018**.
- [43] P. J. Huber, *Robust Statistics*; John Wiley & Sons, **2004**.
- [44] V. Gómez, A. Maravall, *Seasonal Adjustment and Signal Extraction in Economic Time Series*; Banco de España. Servicio de Estudios, **1998**.
- [45] J. Bien, R. Tibshirani, *J. Am. Stat. Assoc.* **2011**, 106, 1075–1084.
- [46] A. E. Hoerl, R. W. Kennard, *Technometrics* **1970**, 12, 55–67.
- [47] R. Tibshirani, *J. R. Stat. Soc. Series. B* **1996**, 58, 267–288.
- [48] H. Zou, T. Hastie, *J. R. Stat. Soc. Series. B* **2005**, 67, 301–320.
- [49] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning*; Springer Series in Statistics; Springer New York: New York, NY, **2009**.
- [50] H. Zou, *J. Am. Stat. Assoc.* **2006**, 101, 1418–1429.
- [51] A. Agresti, *Categorical Data Analysis, 2nd Edition.*, Wiley, **2003**.

Manuscript received: April 1, 2021  
Revised manuscript received: April 30, 2021  
Accepted manuscript online: May 3, 2021  
Version of record online: May 17, 2021