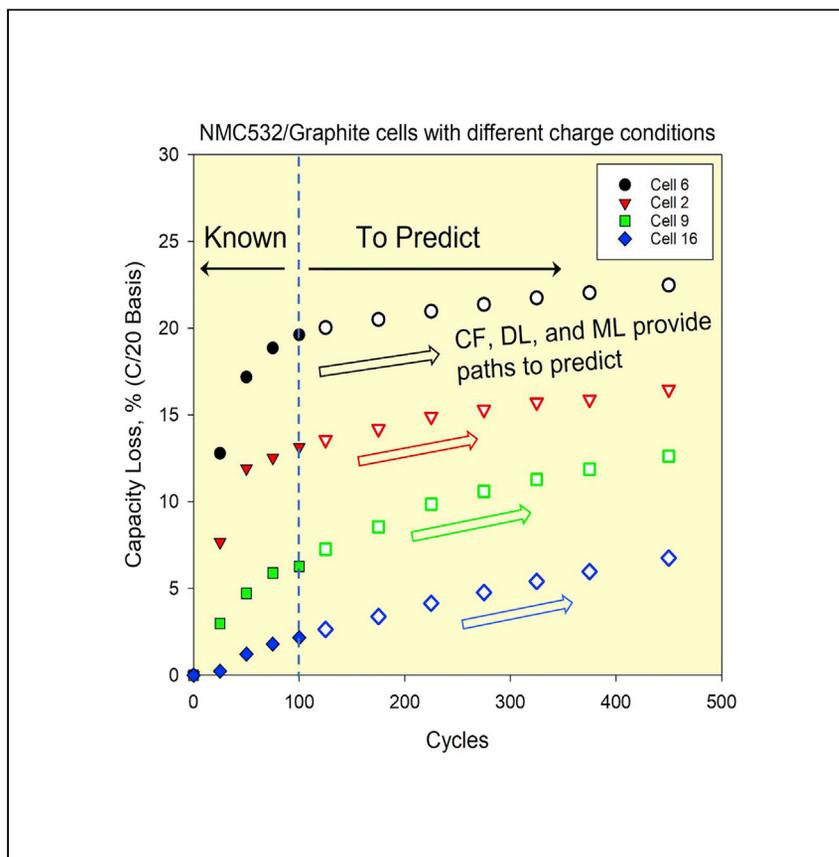CellPress
OPEN ACCESS

Article

Accelerated battery life predictions through synergistic combination of physics-based models and machine learning

Kim et al. report methods to accelerate prediction of battery life on the basis of early-life test data. This allows timely decisions toward managing battery performance loss and related use conditions. This approach provides insights into battery design and operation strategies and may further improve robustness and reliability of batteries.

Sangwook Kim, Zonggen Yi, M. Ross Kunz, Eric J. Dufek, Tanvir R. Tanim, Bor-Rong Chen, Kevin L. Gering

kevin.gering@inl.gov

Highlights

Battery life prediction is accelerated on the basis of using early-life capacity loss data

Deep learning, advanced curve fitting, and machine learning are compared

Methods are demonstrated on NMC/graphite cells tested for fast charge

Small percentage deviations are seen between extended test data and models

Article

Accelerated battery life predictions through synergistic combination of physics-based models and machine learning

Sangwook Kim,[1] Zonggen Yi,[1] M. Ross Kunz,[1] Eric J. Dufek,[1] Tanvir R. Tanim,[1] Bor-Rong Chen,[1] and Kevin L. Gering[1,2,*]

SUMMARY

There are tremendous economic and technical benefits to shortening battery test periods through robust predictive methods. Accurate long-term forecasting of battery life enables proactive planning of battery management (e.g., cell replacements) and preemptive actions to modify operating conditions to improve safety and life. The ever-evolving landscape of battery materials and applications ensure an abiding need for early capture of aging mechanisms. Herein we report on accelerated determination of battery aging mechanisms together with prediction of future capacity loss. Sigmoidal rate expressions (SREs) are used as diagnostic and predictive engines to evaluate aging mechanisms at play. We demonstrate three methods by which SRE parameters are early assessed. Overall results indicate that for cases dominated by loss of lithium inventory we can predict end-of-test capacity loss using less than three weeks of data. In many cases, predictions are within 5%–10% relative error and to within 1%–2% absolute error of observed performance.

INTRODUCTION

Battery energy storage (BES) is undergoing prolific growth into new areas and within existing areas such as vehicles and stationary scenarios. Over the next decade, there is planned a massive transition to electrification of applications that would otherwise be powered by fossil fuels.[6] This move carries with it the need to jointly increase deployment of energy storage technologies and improve battery management with near real-time battery health diagnostics and life predictions. As such, rapid technology validation and intelligent battery management will become an essential element in the ever growing market for electric vehicles and stationary energy storage systems, with the ultimate goal of zero emission in the foreseeable future. A preferred battery management system (BMS) would not only enable battery performance to be continuously monitored but would also allow the prediction of battery aging trends and ultimate lifetime, as well as failures during anticipated modes of operation. A thoughtful understanding and an accurate prediction will extend battery lifetime and increase safety by avoiding detrimental degradation. From the viewpoint of battery research and development (R&D), this management will provide insights into battery design and operation strategies, further improving the robustness and reliability of batteries in an accelerated development cycle. Information gathered from BMSs and coupled with physics-based models can also be leveraged to determine if batteries that have aged past their first application can be repurposed into a second application or undergo materials recycling. Therefore, high-fidelity battery life predictions have

[1]Energy Storage & Electric Transportation, Idaho National Laboratory (INL), Idaho Falls, ID 83415, USA

[2]Lead contact

*Correspondence: kevin.gering@inl.gov

https://doi.org/10.1016/j.xcrp.2022.101023

We seek to diagnose and predict LLI and LAM progression through physics-guided ML.
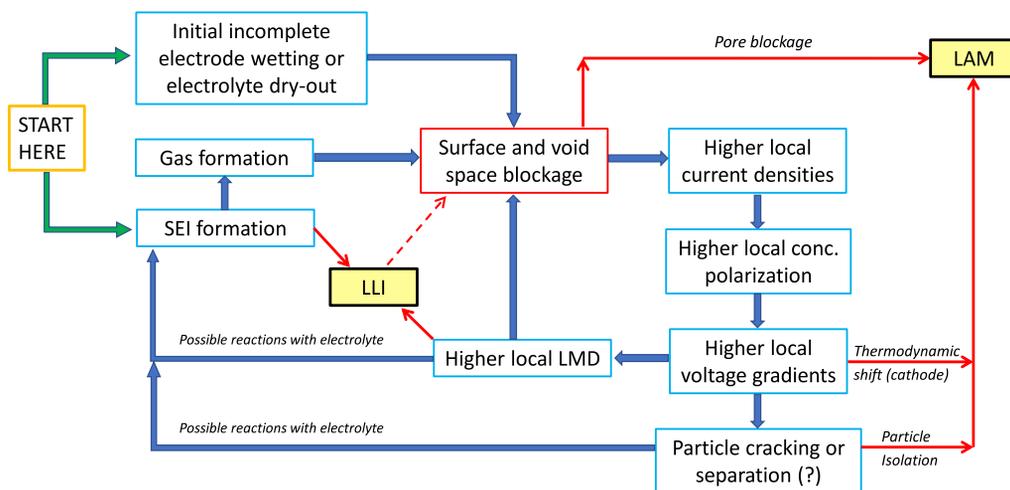

Figure 1. Process flow diagram for lithium-ion battery aging
This diagram shows aging contributions to lithium-ion cells as they relate to LLI and LAM and interconnections thereof. Note that LMD is included as a contributor to LLI.

become a technical and economic necessity for BES to become a thriving component of our energy future.

The primary target for this study is an accelerated determination of battery aging for cells which undergo extreme fast charging (XFC), wherein there can be a combination of foremost mechanisms under three broad "aging modes": loss of lithium inventory (LLI) and loss of active materials (LAM) as segregated between the positive and negative electrodes (LAM_{PE} and LAM_{NE}). The presence of fast charge also can cause lithium metal plating or deposition (LMD), which we generally place under the LLI category, as LMD is largely a surface-driven process that occurs relatively early in battery life.[7–10] We seek not only to diagnose these various mechanisms but also to predict their emergence and growth over extended cycling and time, while also predicting other conventional metrics of aging (e.g., capacity loss). Without a comprehensive understanding of the root causes of degradation, battery lifetime prediction remains a "black box," leaving battery developers and end users with little guidance for cell chemistry improvements and corrective strategies to decrease battery aging.

Figure 1 gives an overview of aging consequences for lithium-ion cells, where connections are drawn between LLI and LAM outcomes. This figure follows a logical progression of consequences under an electrochemical batch reactor scenario, given the materials and typical use conditions for lithium-ion cells. The description given here for Figure 1 is complementary to the LLI and LAM discussion by Gering.[11] It is seen that surface-to-electrolyte interphase (SEI) formation and LMD are primary sources of LLI. Again, these are largely surface-centric processes that occur relatively early in battery aging. Electrode surface and pore blockage can arise from initial incomplete surface wetting by the electrolyte, continued SEI film formation, gas formation, and also the occurrence of LMD. Pore blockage can lead to trapped active material that no longer ionically communicates with the rest of the cell, resulting in one category of LAM. Progression of surface and pore blockage leads to a cascading set of consequences that promote further aging in terms of LLI and LAM, affecting

both electrodes. The cascade sequence is estimated through higher local current densities, larger local concentration gradients with related polarization, higher local voltage gradients, and particle cracking or separation. These contributions cause increased stress on materials due to being driven further from equilibrium (at rest) states. It is noted that cathode particle cracking and formation of LMD may well initiate further surface passivation reactions with the electrolyte. Another outcome from this overall aging landscape is that the negative-to-positive ratio can be shifted away from the original specification, where the direction of shift will reveal which electrode is more profoundly aged. To conclude, Figure 1 gives a concise view of aging processes within lithium-ion cells, with inter-connections noted. When LLI, LAM, and LMD "mechanisms" are mentioned herein, these terms refer to all contributing processes to LLI, LAM, and LMD. Such net contributions to each mechanism are the focus of our work, and not the particular sub-processes (e.g., individual surface reactions for LLI).

Within our consideration of battery aging, we recognize that the following may have a profound influence on aging consequences and the evaluation thereof: condition dependence (path dependence), cell chemistry dependence, and effects from irreversible versus reversible performance losses (polarization). Path dependence will emerge as the operational and environmental conditions of use vary over time. Cell chemistry dependence will reflect the unique aging response to stress factors on the basis of any of several material-specific parameters in the electrodes, electrolyte, or separator. Last, it is important to segregate irreversible and reversible causes of diminished state of health, such as accounting for polarization effects when testing a cell for available capacity.

In previous works, we have showcased frameworks that allow specific aging modes, namely LLI and LAM, to be identified and classified on the basis of the electrochemical data collected during battery life cycles. Multiple electrochemical signatures such as capacity loss, voltage response, and increment capacity (IC) curves can be combined together into a machine learning framework for aging classification.[9,12–15] By combining IC analysis and deep learning together, the percentages of aging in the cathode or anode can be quickly quantified. These tools are a collective complement to foundational work done to establish SRE as a viable means to diagnose and predict battery aging mechanisms.[11]

Herein we deploy SRE as compact, robust mathematical forms that contain three variables each (a, b, M), as discussed in Figure 2, that capture the thermodynamic and kinetic "thumbprint" of the mechanism progression within the context of a batch reactor scenario.[11] As battery cells transfer only current and heat while the materials remain inside, they qualify as batch reactors and SRE can be applied to describe the aging thereof. Per Figure 2, the SRE parameters (a, b, M) are physical parameters representing chemical kinetic and thermodynamic terms (rate constant, order of reaction, and maximum extent of degradation) for a given mechanism. Thus, a given set of (a, b, M) represents a particular mechanism in response to a set of stress factors, and hence will be unique as related to a given cell chemistry. These terms and related mathematics are discussed at length below and in Note S1 and Note S2.

Other works have recently used different ML architectures and related methods for early life prediction.[16–18] Uniqueness of our approach is found in the numerical methods by which we accelerate the identification and prediction of failure modes and performance by determining the (a, b, M) sets and assign them to particular aging mechanisms such as LLI versus LAM. We also derive uniqueness from translation
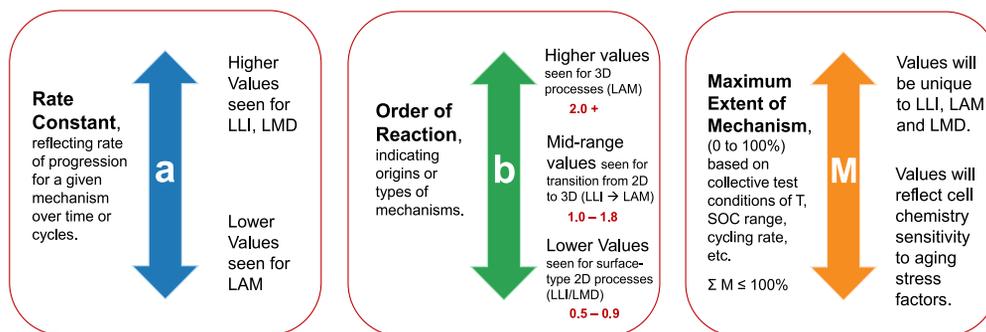
$$\psi\,(i,\mathrm{i}^*,t) = \left\{\left(1 - \theta_t(i,\mathrm{i}^*)\right) \stackrel{!}{=} M_i' + 2\left(M_i - M_i'\right)\left[\frac{1}{2} - \frac{1}{1 + exp((a_i\,t)^{b_i})}\right]\right\}$$

Figure 2. Description of SRE parameters as applied to battery aging analyses
Values of these parameters reveal chemical kinetic and thermodynamic metrics tied to the contributing aging mechanisms. See also Note S1 for further explanation.

between data types (daily cycle-by-cycle versus reference performance tests [RPTs]) to infer SRE (a, b, M) parameters for use within a ML framework. Collectively, these elements enable early forecasting of aging outcomes (type of mechanism and extent of aging) far ahead of actual end-of-life test results. A related benefit is the reduced cost associated with battery life testing, as our accelerated prediction methods can significantly decrease testing time.

On the basis of our SRE foundation, we investigate three separate approaches to accelerated life prediction that require only modest amounts of early data such as capacity loss: (1) a predictive curve-fitting (CF) technique, (2) deep learning (DL) paired with a Monte Carlo method, and (3) machine learning (ML). We consider robustness and the applicability of each as applied to various aging trends as expressed in capacity loss data and gain insights on method accuracy, resilience and practical limitations. Figure 3 summarizes inputs, outputs, and distinct features for these three methods, showing that data requirements generally increase with a move toward ML. We conclude that data-rich applications can be well suited for ML treatments, yet data-sparse scenarios point to using predictive curve-fitting or deep learning methods. In terms of computing demand, the predictive curve-fitting approach requires the least amount of computing time per dataset.

RESULTS

Examples of SRE applied to battery life data

As a demonstration of SRE utility, Figure 4 has a collection of plots that compare capacity loss data (symbols) to SRE regression results (solid or dashed curves) on the basis of full RPT data. Herein we see cases with

- LLI dominance (e.g., P462 cells 9 and 12, Figures 4F and 4E);
- LLI plus a secondary mechanism, most likely LAM (e.g., P462 cells 8 and 11, Figures 4C and 4D); and
- LMD as part of the LLI (P462 cells 4, 5, 6, and 7, Figures 4A and 4B).

From these cases, we see that the values of (a, b, M) vary according to aging trends seen in the test data, here from measurements of capacity at the RPT condition of full C/20 discharges. In cases in which there are two mechanisms, there are two sets of

Increased Need for Experimental Data

	Deep Learning Prediction	Curve Fitting Prediction	Machine Learning Prediction
Method	Simulation + DL	Predictive Curve Fitting (CF)	ML
Input	Early cycle RPT data & SRE simulation data	Early cycle RPT data	Training data includes experimental data with early-to-late life information
Output	Capacity fade & LLI prediction	Capacity fade & LLI prediction	Capacity Fade & LLI & LAM prediction
Distinct Features	A created model can work for all cells in the same group	Each cell needs a specific set of model parameters	All cells in different groups are used together

It is difficult to predict LAM using early cycle data because LAM mostly becomes significant in later cycles.

Figure 3. Comparison of predictive methods used for this work
The relative need for experimental data is highlighted, where DL and CF are well suited for sparse data and cell-wise cases, while the ML method requires more data to facilitate populating the training sets. The nature and breadth of datasets will determine which method should be applied for predictive aging analysis.

SRE parameters denoted by (a, b, M) and (c, d, N). Regression fidelity is high, achieving R^2 values of at least 0.994. The SRE parameters reveal aging mechanisms that are more surface driven (b < 1.0) and those that are tied to 3-dimensional processes (b $\geq$ 2). For the sake of demonstration, the regressed LLI and LAM contributions are plotted for cell 8 in Figure 4C. As expected, LLI dominates as an early aging mechanism, with LAM emerging later in time to eventually dominate the net capacity loss. The progression and relative magnitudes of LLI and LAM are informative about the sensitivity of the cell chemistry to the chosen operating conditions and could be used to advise cell design modifications and corrective actions taken by a BMS.

Summary of accelerated SRE parameter estimations for RPT-based capacity loss

Three different approaches involving DL, CF, and ML were used to predict capacity fade in P462, P492, and P533 cells. Attributes of these cells and the nature of their cycle-life testing are summarized in Note S3. Figure 5 summarizes outcomes for end-of-test (EOT) capacity loss prediction errors, while Figure 6 shows similar type results for the LLI contribution. Note that Figures 5 and 6 account for absolute prediction error between model predictions and experimental data, while Figures S1 and S2 account for relative error. Because of the limited RPT datasets, one additional RPT (at 225 cycles) was used in CF prediction for P492, whereas the other two approaches used RPT cycling data from 0 through 125 cycles. The accuracy of the prediction was measured by the difference between capacity fade percentage from the experiment and the predicted percentage at the EOT at 450 or 425 cycles.

Compared with the prediction performance in P492 and P533, prediction error in P462 is relatively high; average prediction errors using DL, CF, and ML are 2.91%, 2.81%, and 1.72%, respectively. Higher prediction error is caused mainly by reliance on a single mechanism-based SRE, which assumes that a single aging mode (i.e., LLI) dominates capacity fade. On the basis of our previous results, most cells in P462, P492, and P533 showed LLI-dominant aging cases.[14] However, the SRE diagnostic analyses tell us that some cells in P462 and P492 show clear cases of LLI + LAM aging and they require two SRE equations to include both, as shown in Figures 4C and 4D. In DL and CF, only early stage cycling data are used, wherein the effect of LAM is barely observed, thus it is difficult to predict capacity fade accurately because of development of a second mechanism (i.e., LAM) at the later stages of cycling. On
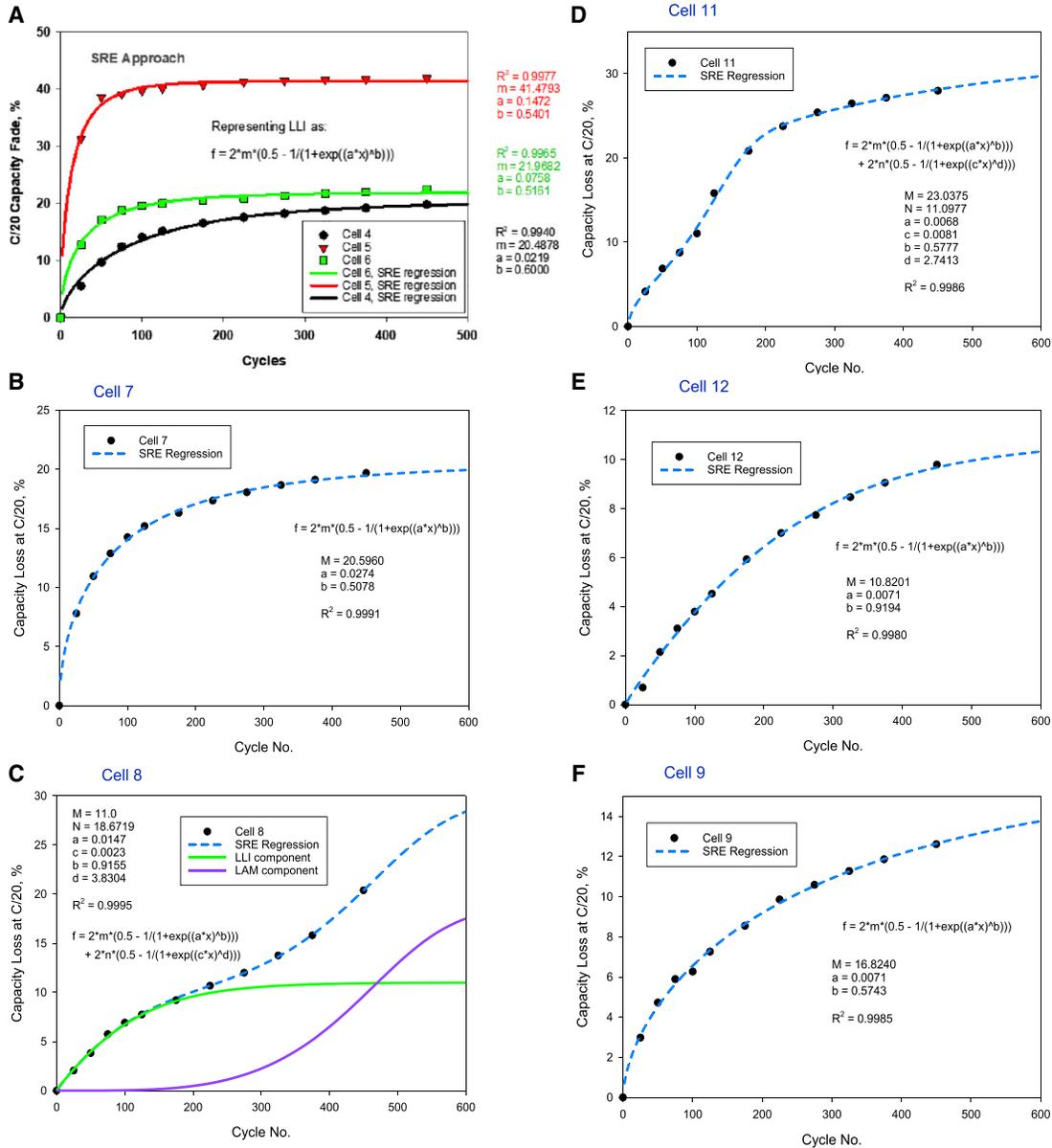
Figure 4. Examples of SRE regression outcomes from P462 cells

(A–F) In most cases, the aging is dominated by a single contribution such as LLI (A, B, E, and F), while other cases show a clear indication of a second mechanism presumed to be LAM (C and D), where both (a, b, M) and (c, d, N) parameters are shown. (C) Constituent contributions from LLI and LAM for the sake of demonstration.

the other hand, the ML approach shows relatively higher prediction accuracy because it is a supervised learning approach and training datasets include the life information at late cycles as the training targets. Moreover, according to the previous studies from our group, clear evidence of LMD was present in more than half of the cells in P462.[15] Thus, complexity and uncertainty from LMD also causes relatively higher prediction error in P462.

It is also observed that there is a strong relationship between change in capacity at 125 cycles and prediction accuracy at EOT (Figure 5). Change in capacity is calculated by the increased capacity fade with respect to the capacity fade at
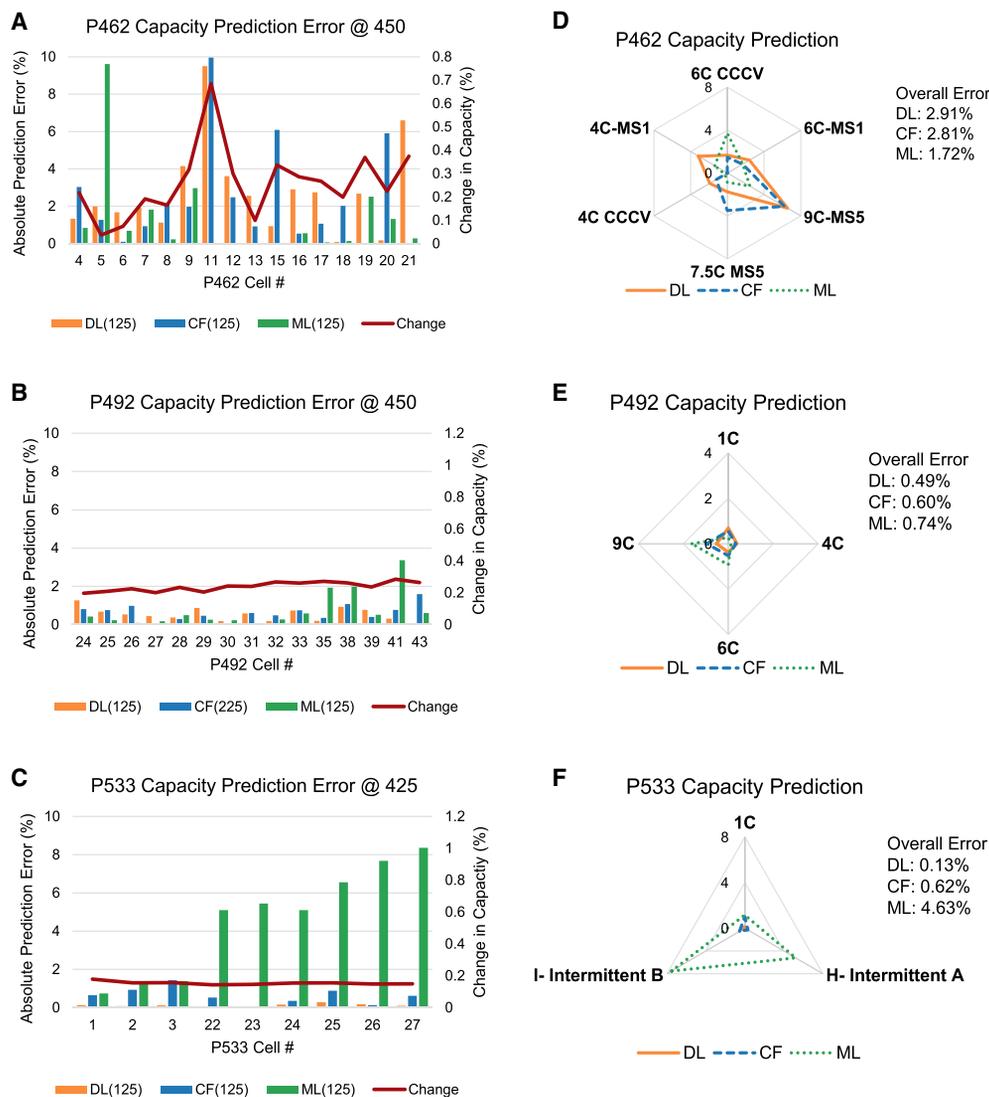
Figure 5. Absolute error of capacity prediction

(A–F) Values are determined at end of test from different methods for (A) P462, (B) P492, and (C) P533. (A–C) Individual cells and (D–F) errors according to different charging protocols. Errors are lowest for the cells with thinner electrodes (P492) while cells with thicker electrodes (e.g., P462) have more cell-to-cell variability, presumably caused in part by inconsistent or inadequate wetting of electrode laminates by electrolyte. ML predictions show increased error for P533 (C and F) because training sets do not have adequate coverage of the intermittent-type charging conditions.

previous RPT: $[(Cap_{RPT}^n - Cap_{RPT}^{n+1})/Cap_{RPT}^n] \times 100$ (%) (Cap$n$th RPT – Cap[n + 1]th RPT)/Cap$n$th RPT × 100 (%). For example, the change in the capacity of cell 11 in P462 is 69% at the sixth RPT, which corresponds to 125 cycles, and the highest prediction error is measured. On the other hand, the averaged change in capacity in P492 and P533 are relatively low, at 24% and 15%. In these cells, DL and CF predict capacity at EOT very accurately as shown in Figures 5 and S1. Despite the low change in capacity in P533, ML exhibits relatively larger error, 4.63%, because the ML model has been trained by the cells in P462 and P492 because of a limited total number of testing cells in P533. From this, performance of ML prediction is highly affected by training datasets, whereas DL and CF prediction accuracy are more affected by change in capacity.
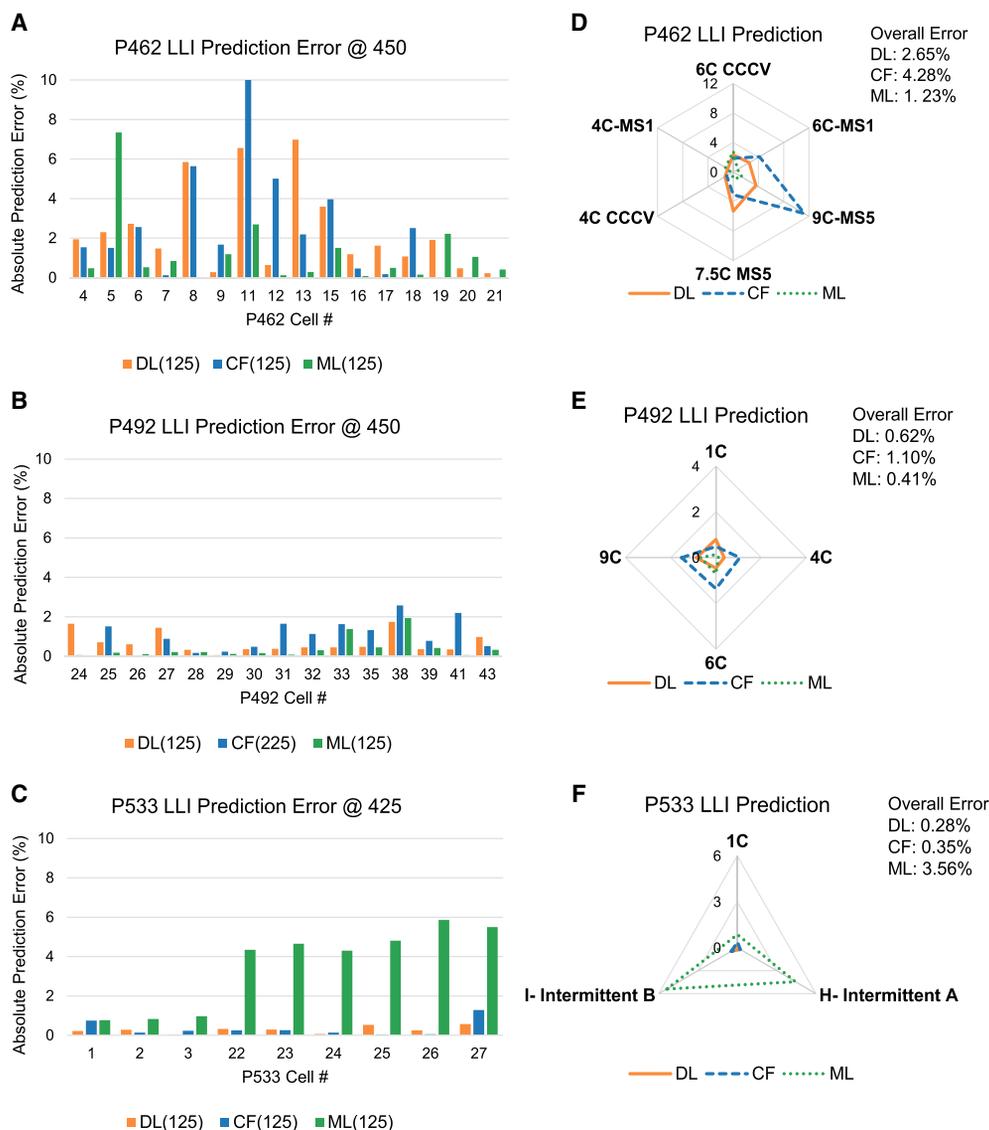
Figure 6. Absolute error of LLI prediction

(A–F) Values are determined at end of test from different methods for (A) P462, (B) P492, and (C) P533. (A–C) Individual cells and (D–F) errors according to different charging protocols. As seen in Figure 5, errors are lowest for the cells with thinner electrodes (P492), while cells with thicker electrodes (e.g., P462) have more cell-to-cell variability, presumably caused in part by inconsistent or inadequate wetting of electrode laminates by electrolyte. ML predictions show increased error for P533 (C and F) because training sets do not have adequate coverage of the intermittent-type charging conditions.

Our three approaches are also used to predict LLI at EOT on the basis of 125 cycles of data, as seen in Figures 6 and S2. In this study, supporting LLI data were obtained from DL incremental capacity (DL-IC) analysis.[14] LLI prediction error is similar with capacity prediction error seen in Figures 5 and S1, as LLI comprises most of the capacity loss in cells of P462, P492, and P533. Therefore, trends of LLI are similar to that of capacity fade during aging cycles. As observed in capacity fade prediction, LLI prediction accuracy was relatively low in the LLI + LAM aging cells.

In this study, capacity fade and LLI at EOT are successfully predicted under different cycling conditions by using only the first 25%–30% of data over time (approximately the first 125–150 cycles) with less than 5% absolute prediction error. Combining
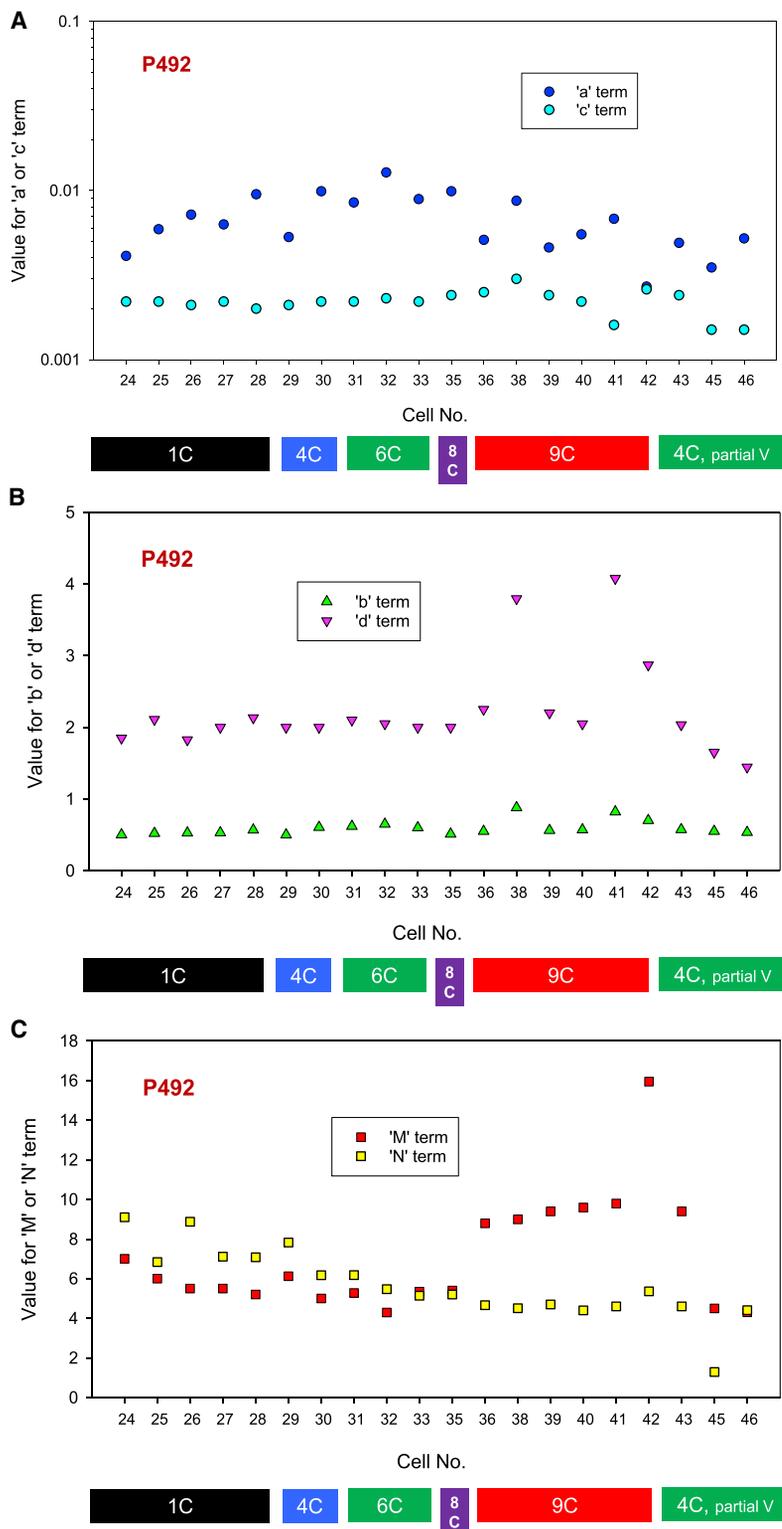
Figure 7. Cell-wise map of SRE parameters, P492

(A–C) This summary of SRE regression parameters covers LLI (a, b, M) and LAM (c, d, N) contributions, showing the relationship between cells and their charge conditions (colored boxes). (A) Rate constant terms, (B) orders of reaction, and (C) maximum extents of degradation under the LLI and LAM routes.

accelerated prediction methods with DL-IC framework will add to the robustness and expandability. Early aging modes that are classified and quantified in DL-IC framework are used to predict the capacity and aging modes at the later stage of cycling data. This will support timely decisions and actions to ensure battery safety and longer life and eventually reduce battery development effort, time, and cost.
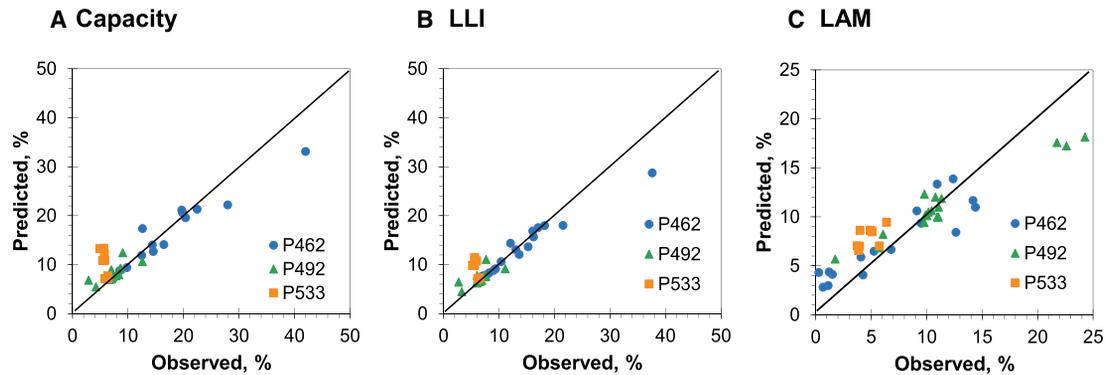
Examples of SRE results for full RPT sets

In addition to accelerated determination of SRE parameters, we recognize the benefit from using multiple SRE to describe full-life aging trends from historical data. In many cases, two SREs are able to accurately render the aging trends for LLI + LAM, while three SREs might be required for cases in which there is an additional early-life mechanism involving slow lithium release from lithium-enriched cathode stoichiometries. In Figure 7, cell-wise aging from P492 was used as a basis to obtain regression outcomes for the (a, b, M) and (c, d, N) parameters for LLI and LAM, respectively. Figure 7A shows results for the rate constants a and c, Figure 7B shows the reaction order terms b and d, and Figure 7C covers the mechanism maximum extent terms M and N. Colored boxes are used to indicate the charge rate used during testing for each subgroup of cells. The parameter plots are generally sensitive to test conditions, providing a good basis for diagnostic aging analysis, as well as predictive (extrapolative) analysis. In Figure 7B it is seen that the b term is consistently below 1.0, indicating that the related aging mechanism is indeed exhibiting kinetics related to surface-driven processes as would produce LLI and LMD, while the d term value is about 2.0 for most P492 cells, with some values higher and lower. Values of d in proximity of 2.0 are indicative of aging processes in 3-dimensional space, whereas values that are appreciably greater than 2.0 may indicate an aspect of damage to active materials in addition to normal aging toward LAM. Another interesting outcome is how parameters respond to the transition from 8 C charging to 9 C, as seen most clearly in Figures 7B and 7C, which suggests the initiation of further LLI possibly because of LMD or electrolyte passivation on surfaces of freshly fractured electrode particles.

ML outcomes

For the data-driven estimation of the RPT responses (i.e., capacity, LLI, and LAM), the P462 and P492 cell data were used as the training/test sets, while P533 was held out as a validation set. This resulted in a total of 29 samples to train and test the random forest model and 3 samples for validation with a different charging protocol. Because of discrepancies of RPT schedules, the end-of-life (EOL) RPT responses differ between each cell type. Table S4 describes the EOL cycle number as well as the total number of cycles at selected RPTs. Three different feature sets for the capacity and the estimated SRE parameters using the cycle-by-cycle data were built on the basis of the total amount of information per RPT (i.e., RPT 1, 3, and 4). The single RPT does not contain any seasonal RPT effects and as such, the features were generated from a simple autoregressive integrated moving average (ARIMA) (1,1,0) model. However, the features generated after 3 and 4 RPTs used the seasonal ARIMA model, as discussed in Experimental procedures, to account for the RPT variation on the cycle-by-cycle series. The total number of RPTs used is restricted to only 4 RPTs, as that is the point in which P533 reaches the designated end-of-life.

Starting with the 3 RPT feature set, the random forest model was built on approximately 125 cycles. This is considered the minimum number of RPTs required for measuring the seasonal ARIMA coefficients as there are only a total of three periods. Figure 8 displays the parity plots for each predicted response in all cell types that correspond to R^2 values of 0.91, 0.95, and 0.94, respectively. Through visual

Figure 8. ML outcomes
Parity plots are shown from the ML method for each predicted response using cycle-by-cycle information that spans three RPTs. The table represents the median absolute prediction errors for each RPT response on the basis of the amount of information per number of RPTs used to generate the cycle-by-cycle features. Note that values for "3 RPT" correspond with the parity plots.

Response	1 RPT	3 RPT	4 RPT
Capacity	3.02%	0.92%	0.88%
LLI	1.21%	0.61%	0.51%
LAM	3.02%	2.16%	2.35%

inspection, it can be noted that there may exist several outliers within the observed values as each response increases in value. These points do depreciate the overall model performance but are still included until further cell investigation can confirm if the points are truly outliers. The maximum absolute error including the validation set for each response was 8.8%, 8.8%, and 6.1% respectively.

The performance of the random forest model was tested over the multiple generated feature sets that includes increasing amount of cycle-by-cycle information. The feature set with only 1 RPT can be considered the control case as there is no seasonality within the cycle-by-cycle data. The table in Figure 8 displays the median absolute prediction errors for each response and feature set combination. In general, as the total number of included cycles on the basis of the RPT increases, the amount of prediction error decreases. This can be attributed to the seasonal ARIMA model to account for the variation in the cycle-by-cycle data caused by each RPT. However, the LAM prediction does not have a strictly decreasing error rate. This could be contributed to the LAM degradation mechanism being more influential later in cycle count, but not necessarily exhibited in capacity curve alone. Different charging features that are more correlated to the LAM can be included to further increase the accuracy in prediction.

DISCUSSION

The three computational methods described herein demonstrate that accurate prediction of future capacity loss and LLI can be achieved through using SREs as predictive engines within different protocols. In particular, accelerated determination of the SRE (a, b, M) parameters was achieved through the DL and CF methods. Accuracy is best attained for aging pathways along constant test conditions (e.g.,

charging protocol, voltage range, temperature, and discharge conditions remain constant within a test group) and that exhibit dependence on one dominant aging mechanism such as LLI. Overall results indicate that for cases dominated by LLI we can predict the EOT capacity loss by using only the first 2–3 weeks of data (out of 8–12 weeks of testing), in many cases to within 5%–10% relative error and to within 1%–2% absolute error. In particular, capacity loss averages for the three packs are summarized per method as follows: P462 (DL, 2.91%; CF, 2.81%; ML, 1.72%), P492 (DL, 0.49%; CF, 0.60%; ML, 0.74%), and P533 (DL, 0.13%; CF, 0.62%; ML, 4.63%).

There appear to be simple reasons why the methods show high accuracy under some conditions but lower accuracy for others. For example, all methods saw low absolute prediction error for P492, a pack that has cells with thinner electrodes and better overall transport rate capabilities. In comparison, P462 has thicker electrodes that have contributed to more cell-to-cell variability, apparently because of issues involving heterogeneity in achieving consistent and thorough electrolyte wetting of the porous regions. Hence, P492 is not as affected by LMD as P462, for example, because of speculated more consistent electrode wetting and decreased over-potential that would otherwise push thermodynamics at the anode toward local LMD. With this in mind, we note that in Figures 5, 6, S1, and S2, there is greater average predictive error in our CF and DL methods for P462 versus P492 and P533. Another observation is that there is a higher incidence of error in P462 for CF and DL methods for MS5 charge cases (see cells 11, 12, 13, and 15 in Figures 5A, 6A, S1A, and S2A). The reason for this is not immediately clear, aside from the fact that these four cells were charged at rates up to 7.5 and 9 C, which would have encouraged more LMD (e.g., cells 11 and 13 had appreciable LMD on the basis of post-test inspection). Last, the ML method provides good predictive accuracy overall with the exception of a few outliers, such as P462 cell 5 and P492 cell 41, but did poorly with P533 cells that underwent intermittent charge profiles (cells 23, 24, 25, 26, and 27). This discrepancy is likely tied to the choice of training versus test sets assigned to the ML, as no training sets were implemented that had intermittent-type charge conditions. In conclusion, the attributes and availability of test data will determine which computational method (or combination thereof) will be better suited for accelerated life prediction. As summarized in Figure 3, DL and predictive CF can manage sparse datasets with SRE parameter analyses done on a cell-by-cell basis. The ML approach relies on more data-rich scenarios but can predict off-matrix conditions with careful selection of training versus test sets.

In addition to expedited predictive capabilities, the SRE evaluations allow diagnostic insights for how the material-driven aging responses varied with cell design and the chosen test conditions. The parameters (a, b, M) (and [c, d, N] where applicable) are a diagnostic thumbprint based on chemical kinetics and thermodynamics, making them suitable to support materials diagnostics and design optimization. For example, it is overall desirable to consider design concepts that will lead to relatively low (a, M) terms, denoting a slow degradation process and a low maximum extent of degradation. Thus, (a, b, M) can be tracked over a design matrix to identify optimal design-to-aging outcomes (this could be done, for example, between P462 and P492 SRE outcomes to look at electrode thickness design attributes).

A key advantage of the methods described here is a broad flexibility to use on a cell-to-cell basis across chemistries and designs including both research and commercial

cells. As clearly shown above predictions were performed using the SRE method for each individual cell and each specific use case. The ability to apply individual mechanism-wise SRE is a distinct advantage when cells are expected to perform under harsh use conditions. The ability to compensate for minor cell-to-cell variation is well suited for design of battery packs and in understanding how other manufacturing and design processes such cell wetting, electrode alignment, and fixturing could ultimately affect performance. The ability to perform broad predictions also provides the opportunity to reduce error when drivers of variability are limited such as in closely matched commercial cells or for mild use cases. Last, we foresee no limitations in applying of our methods to other cell chemistries, because of the generalized batch reactor mathematical basis of the SRE treatments.

In practice, our methods possess high potential to reduce battery test times and related expense, promote rapid validation and market entry, provide diagnostic insights into new cell chemistry performance, and enable timely feedback to battery developers and end users. It should be noted that complicated duty cycles, mixed charging profiles, and varying environmental factors over the timeline may still produce LLI-dominated aging that can be treated through the general SRE approaches given herein. For more complicated aging behavior such as LLI + LMD + LAM, and others, the batch reactor analogy still holds, and other complementary SRE-related mathematical treatments can be practiced if multiple mechanisms are suspected and if test conditions vary along the timeline to produce path-dependent aging. In addition, our techniques are adaptable to cell conductance loss, a parallel term to capacity loss for assessing battery state of health. We recognize the critical importance of such impedance-related terms, as seen in a previous publication that has EIS data along with SEI and charge transfer resistances for P462 and P492 cells.[1] These other factors will be the subject of a follow-on paper.

To conclude, using 2–3 weeks of data with application of SRE-type mathematics for accelerated battery life predictions has been demonstrated within simple constraints, as demonstrated with research-grade Gr/NMC532 pouch cells. This involved methods with and without machine learning elements, depending on whether rich or sparse datasets were used, respectively. The basis of SRE establishes the connection to physics, as the model parameters pertain to chemical kinetics and thermodynamics of an aging mechanism. In most cases, predictions were within 5%–10% relative error and within 1%–2% absolute error for capacity loss and LLI for most of the 40 cells included in our study. These results varied somewhat depending on the chosen method and early-life data trends. Our predictive methods produce best results when early-life aging is dominated by a single mechanistic path such as LLI. Also, our three methodologies apply regardless of the design metrics and charging profile (cell aging done under a single or constant charge condition versus mixed use charging profile), and cell design. The work also shows applicability to extending tools developed using standard cycling to non-uniform, intermittent cases, although care should be taken to have appropriate training sets for ML to enhance prediction capability for anticipated mixed-use cases. In parallel to predictive capabilities, the SRE basis allows the (a, b, M) parameters to be used for diagnostic analyses, with possible connection to improve cell materials design.

Benefits foreseen from this work are numerous because of reduced test times, good accuracy of predictions, adaptability to different data types (e.g., RPT versus CBC), diagnostic insights for cell chemistry response to stress factors, and overall reduced cost for deploying new battery chemistries into the market. The CF and DL methods

are patent pending and are available through license from Idaho National Laboratory (INL).

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Kevin Gering (Kevin.Gering@inl.gov).

Materials availability

This study did not generate new materials.

Data and code availability

- Laboratory data for cell capacity loss reported in this paper will be shared by the lead contact upon reasonable request.
- This paper does not report original code.
- Laboratory procedures and conditions for battery testing are given in the Supplemental Experimental Basis: Battery Testing (Note S3), with key citations given therein.[1–5] Summaries of cell design metrics, cycling protocols and detailed charging profiles are given in Tables S1, S2, and S3, respectively.

Methods to accelerate capture of SRE parameters (a, b, M) with fewer cycles

The overarching objective is to minimize the time (or cycles) required to obtain reliable SRE parameters that support prediction of battery aging over extended time at least 3–4 times longer than the initial test data. Methods are discussed that make use of early-stage RPT data versus use of cycle-by-cycle data performed at higher cycling rates. There are advantages and disadvantages to either approach, as will be discussed below.

Analytical-based extrapolation methods using specialized functions

The general predictive curve-fitting method (see Figure 9) involves five components:

1. early-time RPT data (e.g., capacity loss) of at least three points;
2. applying specialized interpolative/extrapolative mathematical functions that synthesize numerous capacity loss data in between and past RPT data points of item 1;
3. using the outcome of item 2 (at different progression of time or cycle count) to perform SRE regression to obtain SRE parameters (a, b, M) at different points of progression of time or cycle count;
4. plotting the individual plots of a, b, and M gained from item 3 over different points of progression of time or cycle count to ascertain convergence behavior of these terms (we seek to determine convergent values for a, b, and M using minimal RPT data); and
5. using the outcome of item 4 to predict battery aging over extended time or cycle count by implementing the determined set of (a, b, M) within an SRE expression.

We demonstrate this approach by using a single SRE that would denote a single dominating aging mechanism, yet we acknowledge that a generalized application would allow for multiple mechanisms with a corresponding number of SRE expressions. The choice of the special functions in item 2 is for demonstration purposes and is arbitrary, as some forms may provide better extrapolative behavior for particular datasets. These special functions can be customized toward particular SRE

(a) Find (a_n, b_n, M_n) ➡ **(b) Predict (a_{12}, b_{12}, M_{12}) at 450 cycles** ➡ **(c) Prediction at 450 cycles**

A

SRE_n: $F = 2 \cdot M_n \cdot (0.5 - 1./(1 + \exp((a_n \cdot \text{cycle})^{b_n})));$

Data	Cycle	Capacity
1	0	0
2	25	12.78135
3	50	17.179
4	75	18.8563
5	100	19.61334
6	125	20.02941
7	175	20.49761
8	225	20.97087
9	275	21.36115
10	325	21.73976
11	375	22.03629
12	450	22.47789

M_1, a_1, b_1 (3points)
M_2, a_2, b_2 (4points)
M_3, a_3, b_3 (5points)
M4, a4, b4 (6points)

M12, a12, b12
(To be estimated)

B

$y = 1.434\ln(x) + 13.063$

$y = 0.1048x^{-0.101}$

$y = 1.3675x^{-0.103}$

C

$SRE_{n=12}$: $F = 2 \cdot M_{n=12} \cdot (0.5 - 1./(1 + \exp((a_{n=12} \cdot \text{cycle})^{b_{n=12}})))$

- ·×· Experiment
- ○ Cap. (100 cycles)
- △ Cap. (125 cycles)
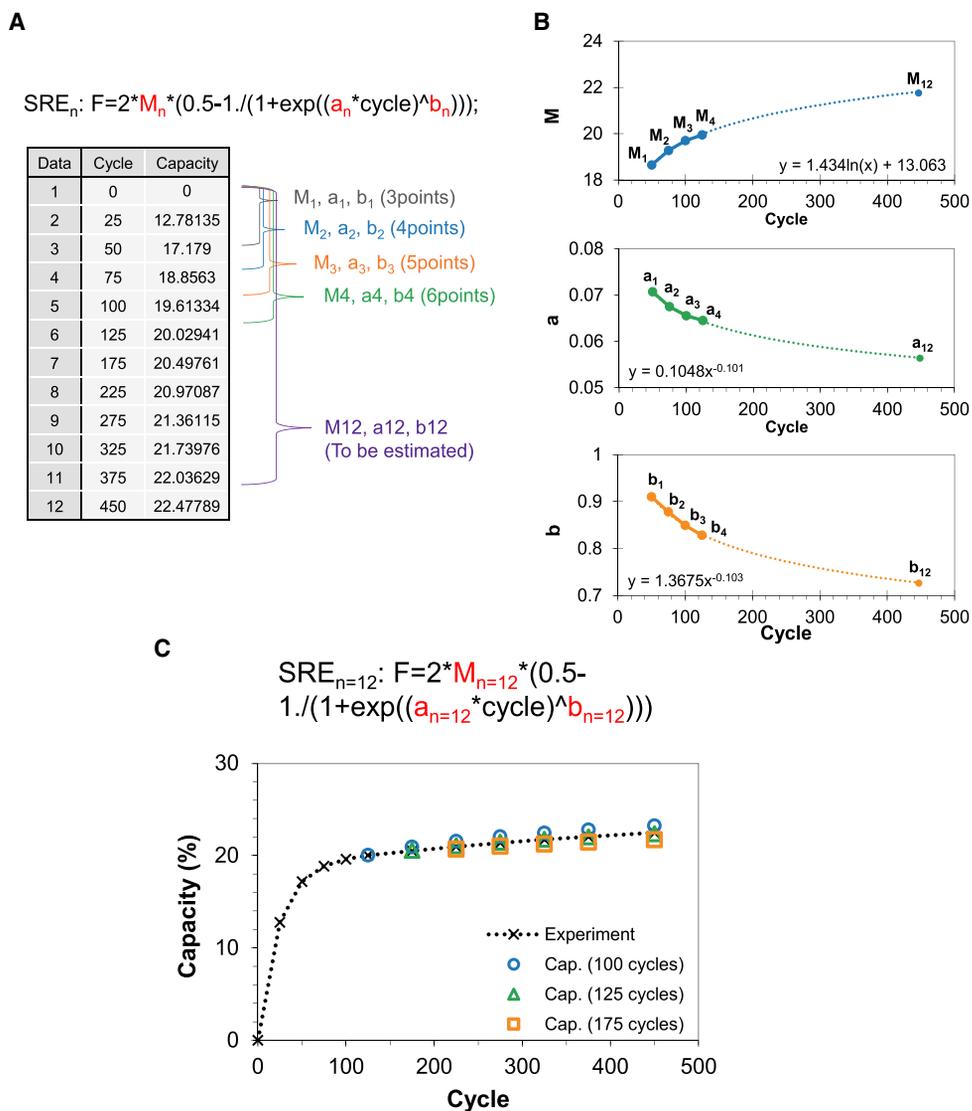- □ Cap. (175 cycles)

Figure 9. Overview of CF method
Shown are procedures and results for accelerated battery life prediction, as demonstrated for P462 cell 6. Note that the parameters (a, b, M) may each require distinct mathematical formulae to enable accurate forecasting on the basis of a small number of initial data points.

parameters. Last, we recognize that datasets outside of RPT conditions could serve for item 1 provided the data are obtained under conditions that do not incur significant polarization effects.

On the basis of this general approach, the following describes a case with particular elements of application. This method incorporates an SRE model with curve fitting to predict capacity loss and LLI at EOT. Three SRE parameters (a, b, M) representing entire aging trends (i.e., from beginning to EOT) are predicted via curve fitting at the early stage of cycling and used to predict capacity and LLI at EOT.

First, SRE parameters $(a, b, M)_n$ are calculated on the basis of $RPT_0 - RPT_i$ ($i = n + 1$) using the curve fitting method. Note that $(a, b, M)_1$ are SRE parameters that best
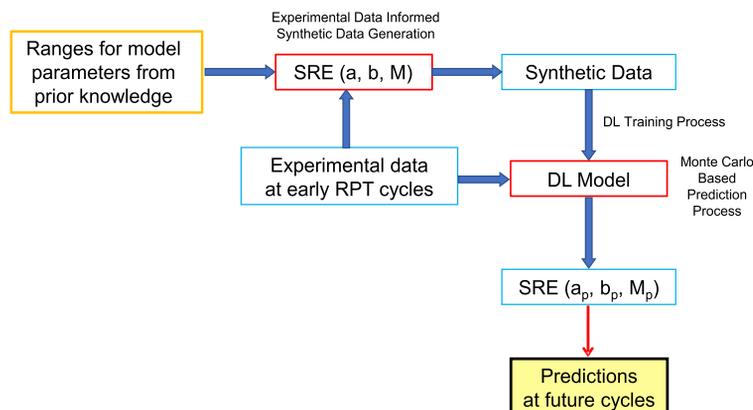
Figure 10. Overview of DL method
This method uses a SRE basis for accelerated battery life prediction, employing a DL training process connected to Monte Carlo inputs than span numerous sets of (a, b, M) values.

represents aging trends between RPT_0 and RPT_2, and $(a, b, M)_2$ is based on the trends between RPT_0 and RPT_3. The parameters are limited according to the general ranges of parameters (a, b, M) of the SRE model. In this study, the non-linear least squares method (Levenberg-Marquardt algorithm) from extensive library SciPy available in Python is used.[19] It requires at least three data points to derive fitting parameters, and fitting performance is improved by including more data points via step 2 above. Thus, capacity and LLI at every cycle number are estimated with the spline interpolation operation[19] using the capacity and LLI at each RPT.

Second, a set of a, b, and M at EOT is predicted on the basis of the trends in each parameter. In most cells, logarithmic and exponential trends are observed in three or four sets of SRE parameters, which require five and six RPT data, including RPT_0, as shown in Figure 9. Trendline equations for each SRE parameter are derived using a Microsoft Excel spreadsheet. After obtaining SRE parameters (a, b, M) at EOT, capacity and LLI are predicted using SRE equation.

Method that couples deep learning with Monte Carlo methods
An accelerated battery life prediction framework is developed by combining the physics-based simulation using SRE forms and deep learning technology.[20] This framework provides an approach which integrates the experimental data, simulations, physics-based model, and deep learning together to predict battery life with early cycle data.

In this framework, the training of DL prediction model relies on synthetic data, which is generated by following an experimental data informed generation process as shown in Figures 10 and S3. More specifically, the synthetic data are generated on the basis of SRE models and include both synthesized capacity fade or LLI data cycle-by-cycle until the end of life (e.g., 600 cycles). Separated sets of synthetic data have been created for different packs (e.g., P462, P492, and P533). They are used to create DL models for predicting capacity fade and LLI, respectively. The DL model requires the inputs of general ranges of parameters in SRE model from prior physical knowledge, and also the information for experimental data at early RPT cycles. According to general ranges of parameters (a, b, M) of SRE model, a grid search method is applied to generate various combinations of (a, b, M), which are fed into SRE model for calculate data at specific cycles. The known experimental data at early cycles is used to ensure the synthetic data aligns with the degradation

behavior at early cycles which has already happened, the simulated data which is not within given thresholds at early cycles is discarded. This type of synthetic data generation process helps create a physically meaningful simulation dataset by injecting a small amount of experimental data, which makes the DL model's training process more efficient. The use of synthetic data on the DL training process alleviates the demand on experiment data a lot, which makes the application of DL methods in accelerated battery life prediction feasible.

The DL model in this proposed framework is a regression model that predicts the parameters (a, b, M) of a SRE model. A one-dimensional (1-D) convolutional neural network (CNN) is constructed to conduct this task, which includes three convolutional layers, two max-pooling layers and three fully connected layers as illustrated in Figure S4. The input data of this DL model has the dimension of 126×1 (this is the dimension when making predictions with the use of RPT data by cycle 125), which is preprocessed (i.e., normalization and interpolation) from given RPT cycles. There are three outputs from this model, which associate with the three parameters in the SRE model. This DL model is trained only using the synthetic data with detailed training and validation process as illustrated in Figure S5. The mean-square error, which is the average of the square of the difference between the original (a, b, M) values and the predicted values, is used to evaluate the accuracy of parameters prediction model. The synthetic data are divided into training and validation datasets (e.g., 80% of synthetic data for training and 20% of them for validation), which aim to reduce the overfitting during the training process. Early stopping technique, which involves updating model checkpoints only when validation accuracy is improved, is also applied to avoid overfitting. Models that have the best validation accuracy on synthetic data are selected, and their model parameters are saved for restoring during the testing and prediction process when using experimental data. The DL model is implemented using Python on the basis of the PyTorch library.[21] The models that make predictions in the Results section are trained using Adam optimizer,[22] learning rate of 0.001, total epochs of 500, and batch size of 128, and so on.

A Monte Carlo-based DL prediction process is developed to predict the capacity fade or LLI at later cycles as illustrated in Figure S6. The inputs of this prediction process are the experimental data of capacity fade or LLI at early RPT cycles and a trained DL model to estimate the parameters in SRE model. By combining the Monte Carlo simulation technique, each given experimental data point serves as the mean value of a uniform distribution or a normal distribution. The variance of constructed input distributions is selected to represent the uncertainties or perturbation on given experimental inputs. By sampling the constructed distributions, several groups of perturbated experimental RPT data are generated. Before feeding them into the trained DL model, the interpolation operation is applied to these groups of RPT data to make sure the inputs have the information at each cycle (e.g., the data array with RPT cycle 0, 25, 75, 125 [a vector with dimension 4 × 1] will be interpolated to a data array with information at each cycle, i.e., a vector with dimension [126 × 1]). The trained DL model from synthetic data will take the experimental input data to predict values of parameters (a, b, M). Each generated input from Monte Carlo simulation will create a prediction on these three parameters. Using these predicted parameters, the corresponding SRE models are established so that the capacity fade information in the future cycles can be inferred from the SRE-based prediction models. For a given future cycle (e.g., cycle 450), multiple predictions on capacity fade or LLI could be derived. From these predictions, a mean value and SD can be calculated, which stand for the mean prediction on capacity fade or LLI and its prediction uncertainties. This shows the proposed approach will not only provide the

mean values on the target prediction, but also can model the uncertainties on predictions. The uncertainties would be necessary and valuable to describe the confidence on predictions for future battery life. Therefore, because of the combination of simulations, physics-based model and DL technologies, this approach reduces the need for experimental data to predict battery life. This accelerates battery life prediction by using very early cycle testing data.

In conclusion, the methods above demonstrate that high-fidelity prediction of EOT capacity loss can be obtained through accelerated determination of the SRE (a, b, M) parameters. This outcome is best captured for aging pathways along constant test conditions that are dependent on one dominant mechanism such as LLI. Other mathematical considerations can be invoked if more than one mechanism is suspected and if test conditions vary along the timeline to produce path-dependent aging. These other cases will be the subject of a future paper.

Machine learning approach

Multi-dimensional testing information can be arranged into a ML problem that allows linkage between test conditions and data types (e.g., daily cycle-by-cycle to monthly RPT data transitions). This enables cycle-wise interpolations between RPT points, thereby producing a much-enriched data field for machine learning objectives. Linkage to SRE formulae then connects this ML framework to physics-defined aging processes.

Data-driven application of machine learning to the prediction of the capacity, LLI, and LAM is highly dependent on the quantity and quality of information obtained within the cells testing period. Additionally, there are complications on how to structure the generated data into a machine learning problem. This stems from the fact that the response that is to be predicted (e.g., LLI, LAM, or aging) is per cell, while the testing information consists of a time series based on the cycle over multiple variables (e.g., temperature, charging conditions, capacity, and voltage responses). Rather than having a matrix of information, commonly used in most ML algorithms, the testing information consists of a 3-dimensional set or tensor: information per cell, information per variable within the cell, and the information per variable with respect to cycle.[12,15] Additionally, the RPT used to measure the state of the battery induces cycle-by-cycle variation within a variable where each RPT may be taken at a different cycle number. As such, this section devoted to connecting the cycle-by-cycle information to the RPT LLI and LAM through data-driven machine learning.

One of the first issues to be considered is converting the structure of the cycle-by-cycle information for each cell into a matrix. In a way, this is a form of dimension reduction to convert the tensor of information into a matrix. The method of choice for this paper is to use the autoregressive moving average coefficients to describe the temporal aspect of a single variable (e.g., the capacity or columbic efficiency per cycle).[12,15] The choice of lags within the ARIMA may be tuned to the problem through an information criterion. However, as noted previously, the RPT has a non-periodic effect on the cycle-by-cycle information. For example, a cell may have an RPT measured at 25, 50, and 75 cycles while another may be measured at 25, 75, and 150 cycles creating disparate cycle behaviors. Early cycle-by-cycle information prior to the first RPT may be used without detriment, but attributes such as the LAM may only be apparent after multiple RPTs. To address this issue, a single cell's cycle-by-cycle data in-between each RPT is interpolated to be of the same size as the smallest RPT period. This creates a periodic RPT effect where the seasonal ARIMA model may be applied to determine the additive or multiplicative effects of each RPT.[23] With the ARIMA coefficients obtained and the seasonal RPT effect

removed, the data can be interpolated back to the original cycle length. Combining the ARIMA coefficients results in a final matrix where each row consists of a single cell and each column is an ARIMA coefficient per variable.

In addition to the ARIMA coefficients, a set of SRE a, b, and M values were estimated using the CBC information within the respective RPT range. This was done to help inform the model of the non-linearities introduced by the logistic nature of the a, b, and cycle number. The values of a, b, and M were optimized through minimizing the root-mean-square error between the cycle-by-cycle capacity and the SRE functional response of the capacity. The cycle-by-cycle capacity does not follow the exact RPT capacity, but it is expected that this information can be used to infer the RPT result.

On the basis of the aspects of the data, the random forest algorithm was chosen to be used in prediction of the capacity, LLI, and LAM.[24] This is mainly because the data are small and exhibit potentially imbalanced cell types. However, the collection of trees that drives the prediction is restrictive in interpretation and will be reported as a prediction value rather than explaining which effects constitute an increase in either LLI or LAM in the degradation mechanism. Both the number of variables sampled at each decision tree split (mtry) and the number of trees used were both optimized using the out-of-bagging error (OOBE).

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xcrp.2022.101023.

AUTHOR CONTRIBUTIONS

Conceptualization, K.L.G., S.K., Z.Y., and M.R.K.; methodology (computation), K.L.G., S.K., Z.Y., and M.R.K.; lab data, T.R.T. and B.-R.C.; investigation, K.L.G., S.K., Z.Y., and M.R.K.; writing – original draft, K.L.G., S.K., Z.Y., E.J.D., B.-R.C., T.R.T., and M.R.K.; writing – review & editing, K.L.G.; funding acquisition, E.J.D.; resources, E.J.D.; supervision, K.L.G., T.R.T., and E.J.D.

DECLARATION OF INTERESTS

K.L.G., Z.Y., and S.K. have a patent pending on the CF and DL methods of this work.

REFERENCES

1. Chinnam, P.R., Colclasure, A.M., Chen, B.-R., Tanim, T.R., Dufek, E.J., Smith, K., Evans, M.C., Dunlop, A.R., Trask, S.E., Polzin, B.J., and Jansen, A.N. (2021). Fast-charging aging considerations: incorporation and alignment of cell design and material degradation pathways. ACS Appl. Energy Mater. *4*, 9133–9143. https://doi.org/10.1021/acsaem.1c01398.

2. Wohlfahrt-Mehrens, M., Vogler, C., and Garche, J. (2004). Aging mechanisms of lithium cathode materials. J. Power Sources *127*, 58–64. https://doi.org/10.1016/j.jpowsour.2003.09.034.

3. Du Pasquier, A., Huang, C.C., and Spitler, T. (2009). Nano $Li_4Ti_5O_{12}$–$LiMn_2O_4$ batteries with high power capability and improved cycle-life. J. Power Sources *186*, 508–514. https://doi.org/10.1016/j.jpowsour.2008.10.018.

4. Han, X., Lu, L., Zheng, Y., Feng, X., Li, Z., Li, J., and Ouyang, M. (2019). A review on the key issues of the lithium ion battery degradation among the whole life cycle. eTransportation *1*, 100005. article 100005. https://doi.org/10.1016/j.etran.2019.100005.

5. Mai, W., Colclasure, A.M., and Smith, K. (2020). Model-instructed design of novel charging protocols for the extreme fast charging of lithium-ion batteries without lithium plating. J. Electrochem. Soc. *167*, 080517. https://doi.org/10.1149/1945-7111/ab8c84.

6. U. S. DOE (2020). Energy Storage Grand Challenge Roadmap. DOE/PA-0022.

7. Katzer, F., Jahn, L., Hahn, M., and Danzer, M.A. (2021). Model-based lithium deposition detection method using differential voltage analysis. J. Power Sources *512*, 230449. https://doi.org/10.1016/j.jpowsour.2021.230449.

8. Waldmann, T., Hogg, B.-I., and Wohlfahrt-Mehrens, M. (2018). Li plating as unwanted side reaction in commercial Li-ion cells – a review. J. Power Sources *384*, 107–124. https://doi.org/10.1016/j.jpowsour.2018.02.063.

9. Tanim, T.R., Paul, P.P., Thampy, V., Cao, C., Steinrück, H.G., Nelson Weker, J., Toney, M.F., Dufek, E.J., Evans, M.C., Jansen, A.N., et al. (2020). Heterogeneous behavior of lithium plating during extreme fast charging. Cell Rep. Phys. Sci. *1*, 100114. https://doi.org/10.1016/j.xcrp.2020.100114.

10. Chinnam, P.R., Tanim, T.R., Dufek, E.J., Dickerson, C.C., and Li, M. (2022). Sensitivity and reliability of key electrochemical markers for detecting lithium plating during extreme fast charging. J. Energy Storage *46*, 103782. https://doi.org/10.1016/j.est.2021.103782.

11. Gering, K.L. (2017). Novel method for evaluation and prediction of capacity loss metrics in Li-ion electrochemical cells. Electrochim. Acta *228*, 636–651. https://doi.org/10.1016/j.electacta.2017.01.052.

12. Kunz, M.R., Dufek, E.J., Yi, Z., Gering, K.L., Shirk, M.G., Smith, K., Chen, B., Wang, Q., Gasper, P., Bewley, R.L., and Tanim, T.R. (2021). Early battery performance prediction for mixed use charging profiles using hierarchal machine learning. Batter. Supercaps *4*, 1186–1196. https://doi.org/10.1002/batt.202100079.

13. Nagpure, S.C., Tanim, T.R., Dufek, E.J., Viswanathan, V.V., Crawford, A.J., Wood, S.M., Xiao, J., Dickerson, C.C., and Liaw, B. (2018). Impacts of lean electrolyte on cycle life for rechargeable Li metal batteries. J. Power Sources *407*, 53–62. https://doi.org/10.1016/j.jpowsour.2018.10.060.

14. Kim, S., Yi, Z., Chen, B.-R., Tanim, T.R., and Dufek, E.J. (2022). Rapid failure mode classification and quantification in batteries: a deep learning modeling framework. Energy Storage Mater. *45*, 1002–1011. https://doi.org/10.1016/j.ensm.2021.07.016.

15. Chen, B.-R., Kunz, M.R., Tanim, T.R., and Dufek, E.J. (2021). A machine learning framework for early detection of lithium plating combining multiple physics-based electrochemical signatures. Cell Rep. Phys. Sci. *2*, 100352. https://doi.org/10.1016/j.xcrp.2021.100352.

16. Severson, K.A., Attia, P.M., Jin, N., Perkins, N., Jiang, B., Yang, Z., Chen, M.H., Aykol, M., Herring, P.K., Fraggedakis, D., et al. (2019). Data-driven prediction of battery cycle life before capacity degradation. Nat. Energy *4*, 383–391. https://doi.org/10.1038/s41560-019-0356-8.

17. Birkl, C.R., Roberts, M.R., McTurk, E., Bruce, P.G., and Howey, D.A. (2017). Degradation diagnostics for lithium ion cells. J. Power Sources *341*, 373–386. https://doi.org/10.1016/j.jpowsour.2016.12.011.

18. Gasper, P., Gering, K., Dufek, E., and Smith, K. (2021). Challenging practices of algebraic battery life models through statistical validation and model identification via machine-learning. J. Electrochem. Soc. *168*, 020502. https://doi.org/10.1149/1945-7111/abdde1.

19. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods *17*, 352. https://doi.org/10.1038/s41592-020-0772-5.

20. LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature *521*, 436–444. https://doi.org/10.1038/nature14539.

21. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al. (2019). Pytorch: an imperative style, high-performance deep learning library. Adv. Neural Inf. Process. Syst. *32*, 1–12. 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada. https://doi.org/10.48550/arXiv.1912.01703.

22. Kingma, D.P., and Adam, J.B. (2017). A method for stochastic optimization. Preprint at arXiv. Published as a conference paper at ICLR 2015. https://doi.org/10.48550/arXiv.1412.6980.

23. Shumway, R.H., and Stoffer, D.S. (2000). Time Series Analysis and its Applications*3* (Springer).

24. Breiman, L. (2001). Random forests. Mach. Learn. *45*, 5–32. https://doi.org/10.1023/A:1010933404324.

Supplemental information

Accelerated battery life predictions

through synergistic combination

of physics-based models and machine learning

Sangwook Kim, Zonggen Yi, M. Ross Kunz, Eric J. Dufek, Tanvir R. Tanim, Bor-Rong Chen, and Kevin L. Gering

<u>**Note S1**</u>: **Supplemental Mathematical Background and Application to Battery Aging**

Three methods that have a basis in SRE are investigated for comparative accelerated life predictions wherein only small amounts of early-life data such as capacity loss are required. A discussion on SRE basis and utility is given for the sake of technical context for battery systems.

   1.1. SRE Basis and Background

For battery systems, SREs are a chemical engineering means to describe reaction progression and consequences in a batch reactor, for single (Eq. 1) and multiple (Eq. 2) aging mechanisms[6]:

$$\psi\,(i, i^{*}, t) = \left\{ \left(1 - \theta_t(i, i^{*})\right) \overset{!}{=}\ M_i' + 2\,(M_i - M_i')\left[\frac{1}{2} - \frac{1}{1 + exp((a_i\,t)^{b_i})}\right]\right\} \qquad (1)$$

$$\sum_{i}^{z} \psi\,(i, i^{*}, t) = \left\{ \sum_{i}^{z}\left(1 - \theta_t(i, i^{*})\right) \overset{!}{=}\ \sum_{i}^{z}\left(M_i' + 2\,(M_i - M_i')\left[\frac{1}{2} - \frac{1}{1 + exp((a_i\,t)^{b_i})}\right]\right)\right\} \qquad (2)$$

where $\psi\,(i, i^{*}, t)$ represents a life metric such as capacity loss that varies from an initial state, usually a new cell at beginning of life (BOL), although provision is made to start the analysis from an interim aging state through inclusion of the M_i' term. Each SRE contains three physical parameters (a,b,M) that collectively describe the progression of a given mechanism along a time regime or other counting-type regime such as cumulative battery cycles (see Fig. 2). The 'a' term represents the rate constant for the mechanism, the 'b' term represents the order of said reaction or mechanism, and the 'M' term is the maximum possible extent (0 to 1 or 0 to 100%) of the mechanism under the conditions of the aging test. When aging data (such as percent capacity loss over time) is regressed using SREs, there are some general guidelines that should be followed to keep (a,b,M) within realistic physical boundaries. First, (a,b,M) should all be positive. Values of 'a' should track with the relative strength of the prevailing stress factors that drive forward a mechanism, e.g., becoming larger at higher temperatures. The 'b' term will reflect the origin or mode of mechanism, such as LLI vs LAM, by virtue of its magnitude (see Fig. 2). Also, similar to the 'a' term, the 'M' term should track with the relative strength of the prevailing stress factors. That is, for battery aging in general we anticipate that M will become larger under battery operating conditions of higher temperatures and larger SOC. We note that in all cases the sum of M contributions from all mechanisms must be within 0 to 100 percent, that is $\Sigma M_i \leq 100\%$. Because (a,b,M) are unique to a given chemistry under particular stress factors, they provide a physics basis for cataloguing aging response and can then be used to elucidate (1) cell chemistry improvements, and (2) changes to cell use conditions that would moderate the rate and extent of aging. SRE-type expressions are becoming increasingly recognized as concise yet powerful predictive tools for describing battery aging behavior.[13] This is an encouraging trend, as it speaks to the move to adopt physics-based methods and to lessen the dependence on empirical mathematics.

In the expressions above 'i' represents the ith aging mechanism, i* represents arbitrary aging conditions and 't' is time. Equivalently, cycle count can be substituted for time provided there is a translation for cycles-per-day in the 'a' term. The summation form of Eq. (2) infers that aging mechanisms are additive. This appears to be a reasonable assumption for LLI and LAM, as they tend to progress along different mechanistic pathways with LLI being surface-driven and LAM being more reliant on 3-dimensional attributes. We recognize, however, that there can be inter-dependence among aging mechanisms, such as where LLI might need to reach a particular threshold before a geometric-type LAM component is initiated. In such cases, data analyses through SRE (a,b,M)$_i$ can reveal if there is a transition region between the occurrence of LLI and a mature LAM mechanism.

Regarding the progression of LLI and LAM, it is generally understood that surface-driven reactions such as electrode passivation occur at a quicker rate than those aging mechanisms that progress along 3-dimensional domains within electrode materials. Thus, we anticipate that battery aging will typically have LLI as a predominant early mechanism, followed by LAM.[6] The extent to which these mechanisms manifest over time is controlled by the combinatorial effects of battery materials design and chemical formulations thereof as well as the prevailing environmental and use conditions. Given such relative emergence between LLI and LAM, an SRE expression for each provides adequate descriptions of the combined aging over time. This summed SRE outcome generally assumes that the LLI and LAM contributing mechanisms are independent, although the approach can be adapted to accommodate inter-dependency between LLI and LAM, such as when LLI is sufficiently advanced to reach a threshold that initiates LAM. Lastly, there are cases where a third capacity change mechanism is at play, but rather than a loss mechanism there is capacity gain from early-life thermodynamic stabilization of the cathode, wherein the cathode slowly excludes "excess" lithium as it arrives at a more thermodynamic-stable crystalline state.[6,14-16] This is often seen for cathodes that are engineered with excess stoichiometric lithium. For this work we restrict SRE application to LLI and LAM components. A special caveat for LLI is that it may contain contributions from LMD during cell charging. Since LMD is a surface driven consequence, its impact on capacity loss may be indistinguishable from routine LLI (that is the result of surface passivation from oxidation-reduction reactions involving solvents on electrode surfaces) without additional data or analysis.

Note S2: Supplemental Data Management and Requirements

Uniqueness of solution is always an underlying question for regression schemes. Under normal aging progression, LLI (a 2-dimensional consequence on electrode surfaces) would dominate at early time, while LAM (a 3-dimensional consequence within electrode substrates) would take longer to manifest. This is general knowledge for lithium-ion systems discussed by Gering.[6] Per the SRE parameters described in Figure 2 we can confirm the order of mechanisms from regression outcomes of the (a,b,M) parameters. Regarding overfitting, it would be difficult to determine a specific minimum amount of data for a deep learning model at the beginning of analysis. Usually, while training a deep learning model we try utilizing as much data as we can to achieve a best performance. After doing this, various techniques, e.g. early stopping, regularization, etc. can be used to avoid overfitting.

On the practical side of data utilization, we are aware that too much data can result in less accurate estimation and sometimes outlier data needs to be excluded for better estimation. However, usually excessive data that "breaks" the modeling framework is because too many features are used as model inputs, not just because the amount of data is used for model training. If some irrelevant features are used for model training and construction, this will compromise the model's performance. In CF and DL prediction, we have utilized early cycling RPT data (with as little as only 3-5 data points), thus we did not have the problem of too much data causing overfitting. So, the usable data for our model's inputs is limited. As shown in Fig. 5, we derived $(a,b,M)_{1-4}$ using 3 – 6 data points and predict $(a,b,M)_{12}$ based on trend captured from $(a,b,M)_{1-4}$. Therefore, a little deviation or scatter in the trend (e.g., capacity loss at 100 cycles in cell 9, shown in Fig. 4f) in capacity data or derived (a,b,M) affects prediction accuracy significantly in CF prediction. This results in relatively higher prediction error in P462 Cells 4, 11, and 15. On the other hand, the DL prediction model is trained using the generated simulated data based on SRE, which helps to minimize effect of data scatter. That is one of main reason why DL prediction outperform CF prediction without excluding data.

As a parallel point, the incorporation of SRE physics-based models provides benefits in terms of guiding ML over cycle counts, thus reducing the number of cells needed to train the ML architecture. In this way our methods are distinct from traditional ML approaches that are data-intensive and we reduce the testing burden.

A complication is arrived at in some early-life battery aging data in terms of mixed-mode LLI comprised of SEI and LMD contributions. A general starting point to deconvolute LMD from "normal" SEI-driven LLI is to consider a baseline testing condition at a charging rate slow enough to avoid creating LMD-forming conditions. That would be used to compare against data obtained at higher charge rates, keeping the number of cycles as the basis. Subtraction of {higher rate} – {slower rate} capacity loss data can give an early estimate for the contribution due to LMD. Choice of these cycling conditions is important and might take some early trial and error for a particular chemistry. For example, when considering cells with thicker electrodes our LMD-free charging rate was assigned at around 4C, while 6C conditions clearly produced LMD in many cells. There are other methods that make use of electronic signatures to detect the occurrence of LMD, such as was published by our group.[5]

Thus, aging estimation accuracy is affected by mechanistic complexity of aging processes within the system, and this gives shape to the edge of understanding for battery chemistries. The presented methods in this study show better estimation accuracy in cells with less or no lithium plating, possibly due in part to CF and DL having utilized single SRE for all cells. For example, we observed significant lithium plating on the anode after 450 cycles from cells 4-15 in P462, and less or no evidence of lithium plating from cells 16-21 in a previous publication,[10] which corresponds to estimation accuracy as shown in Fig. 8a. The progression from LLI to LAM is also an area of heightened complexity, as there may well be interdependence of LAM on LLI. We see the value of exploring the enticing complexity of these topics in a follow-on paper.

Note S3: Supplemental Experimental Basis: Battery Testing

3.1 Background

The purpose of the underlying research was to determine performance and aging of lithium-ion cell chemistries designed for fast-charge applications. As described elsewhere, cells were tested under multiple conditions, many defined by CC-CV charging, wherein there is an initial constant-current (CC) portion followed by a constant-voltage (CV) condition at an upper voltage.[4,9,10] In addition to single-condition CC-CV charge rates, some cells underwent other "mixed use" charge profiles that had either multiple charge steps or other combinations either within the same charging event or periodically over a specified number of cycles. Mixed use charge cases allow us to orient charging profiles more closely to match capabilities of a cell chemistry (to increase charge performance and decrease aging), and to approach more realistically what charge conditions will eventually be seen in vehicle charging facilities.

3.2 Description of cell chemistry and differences therein (P462, P492, P533)

The single layer pouch cells used in this study were built by Cell Analysis, Modeling, and Prototyping (CAMP) Facility at Argonne National Laboratory and had Graphite (1506T) and NMC532 (Toda America) as anode and cathode materials, respectively. The Toda NMC532 has a polycrystalline architecture with no coatings. The cells were built with two different combinations of loading or thickness: low loading (L_{low}) and moderate loading ($L_{moderate}$). Detailed information of the cell materials and architectures can be found in Table S1. The tested cells' names are identified by a test plan number (starting with a "P" following by three digits) and a cell number. Cells with the same P number are fabricated using the same design, cell chemistry, and tested in the same batch.

The choice of cell materials was driven by the U.S. DOE XCEL program, covering development of battery fast-charge chemistries. The overarching goal of this program is to develop lithium-ion cell chemistries that facilitate a 10-minute fast charge at room temperature while using high-energy electrode loadings. While the NMC532 cathode architecture proved to be a reasonable and informative early choice for the XCEL program, the evolving focus has since moved more toward cobalt-reduced cathode materials such as NMC811. In more recent work we confirmed application of our predictive methods to cells with NMC811 and other anode materials.

3.3 Test Conditions and Charge Protocols

Table S2 lists the various conditions for cell formation, life-cycle aging and charging protocols. It is seen that the three test packs underwent unique combinations of charge conditions. Table S3 contains more detailed descriptions for the various charging methods. Cells were tested with a MACCOR series 4000 Automated Test System at 30 ±1°C. During each life cycle, the cells were charged and discharged as designated, with 15-minutge rests between cycling legs.

Cells were cycled with different charging profiles up to 600 times, while discharging at a C/2 rate in between. For the L_{low} cells time limited constant current-constant voltage (CC-CV) charge profiles between 1C to 9C were evaluated (See Table S2). For the $L_{moderate}$ cells, besides CC-CV, several different versions of fast charge profiles, i.e., 2-step, MS5, voltage ramp, and intermittent profiles were investigated (refer to Table S2). A subset of $L_{moderate}$ cells were cycled with different fast charge scheduling described in Table S2. The intermittent profiles are a mixture of several simpler profiles. In intermittent profile A1 in every 5 cycles were charged at the 7.5 C voltage ramp charge profile[17] (see Table S3), and the rest of 4 cycles were charged using 1C CC-CV. In intermittent profile B, the first 5 cycles in each of the cycle life set (continues cycles before RPT) were charged using the voltage ramp 7.5 C profile, while the rest of cycles follow 1C CC-CV profile.

3.4 RPT versus cycle-by-cycle (CBC) bases

Good supporting data is the currency for good analyses. In some cases, selected data conditioning is done to prepare data for one or more mathematical analyses. For example, smoothing is done for data that would otherwise reflect undue CBC variance. Also, polarization effects are subtracted from higher rate cycling data to better isolate true irreversible capacity loss. Ultimately, such data conditioning provides a more consistent platform for data entering SRE and ML architectures, enabling more stable and reliable predictions. In addition to the daily CBC testing, periodic RPT were conducted every 25–50 cycles initially, and every 100–125 cycles in the later stages over the course of a total of 450–600 cycles. RPTs consisted of a C/20 charge-discharge cycle with an 1 h rest time between each the charge and discharge components of the cycle. For cells charged with aggressive conditions, the cell recovers during the slow RPT cycles from the polarization caused by the fast-charging lifecycles and gains a more efficient Li utilization pattern. This recovery process is visible in the cycle-by-cycle trend of capacity, as intermittent spikes up to 5–7 cycles immediately after each of the RPT intervals. This phenomenon is also reported and rationalized in previous works.[4,8,9] An outcome of operating at both RPT and CBC conditions is that the resultant capacity data has unequal bases owing from the CBC data being impacted by polarization and the dynamic recovery introduced by such.

3.5 Data suitable for use in accelerated life predictions with PGML

We endeavor to identify irreversible aging outcomes through our testing and modeling techniques. Data that is obtained at relatively higher cycling rates (here, C/2 versus C/20) will be encumbered with reversible polarization effects that confuse the isolation of true aging progression. Cycles allowed to continue under such conditions may well develop a polarization hysteresis that is cumulative over cycles. For these reasons we generally use slow-rate RPT capacity data (e.g., C/20 basis) as the "truth" for irreversible aging. However, numerical treatment of faster-rate data (via smoothing, compensation for polarization, etc.) was used herein to provide information that is usable within ML as training data. The benefit of using CBC data obtained at more common, faster rates is one of convenience and economy, in terms of not disrupting cycle-life testing (as much) for the sake of time-consuming monthly diagnostics. However, in order to use CBC data within a meaningful quantitative framework, polarization attributes must be managed to prevent introducing reversible polarization effects into data that is meant to capture true irreversible aging.
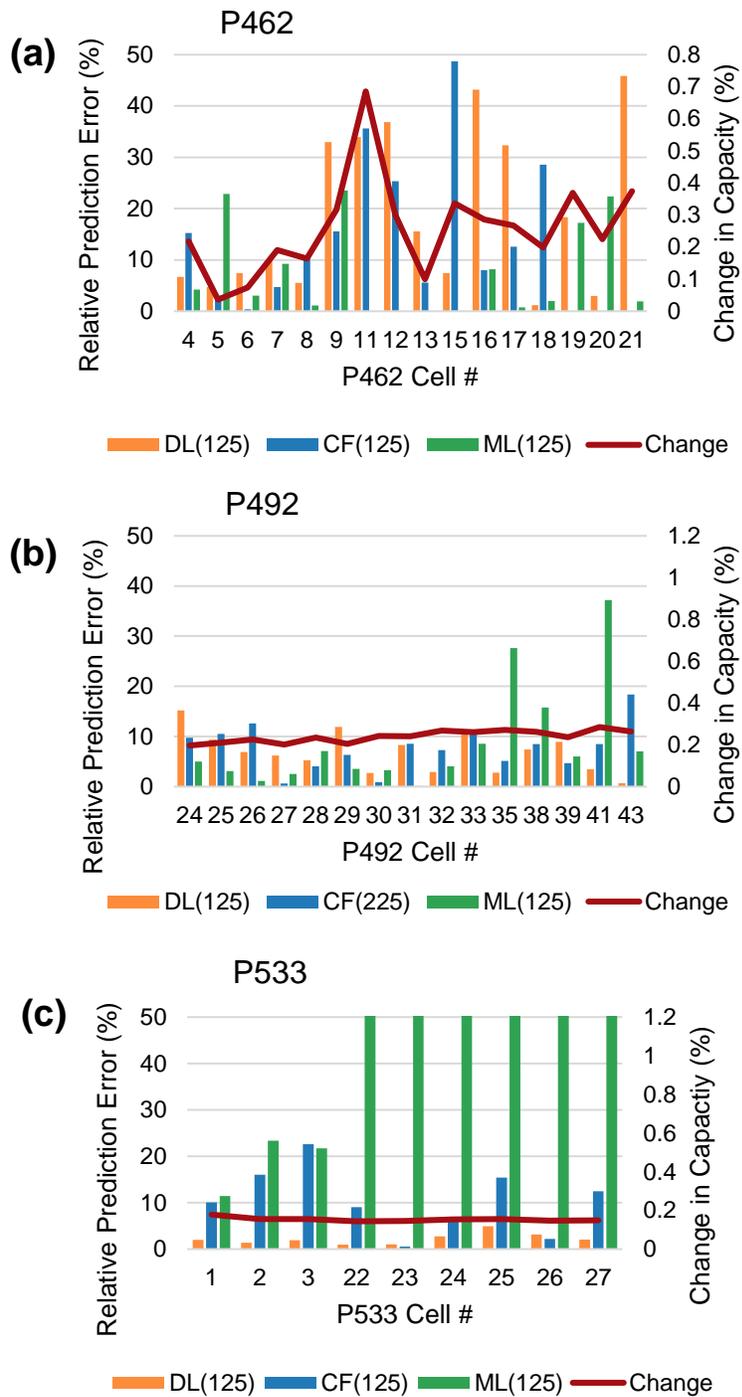
Fig. S1. Summary of relative prediction error in capacity loss for cells of the three test packs within this study. Please also see legend for Figure 5 for further details.
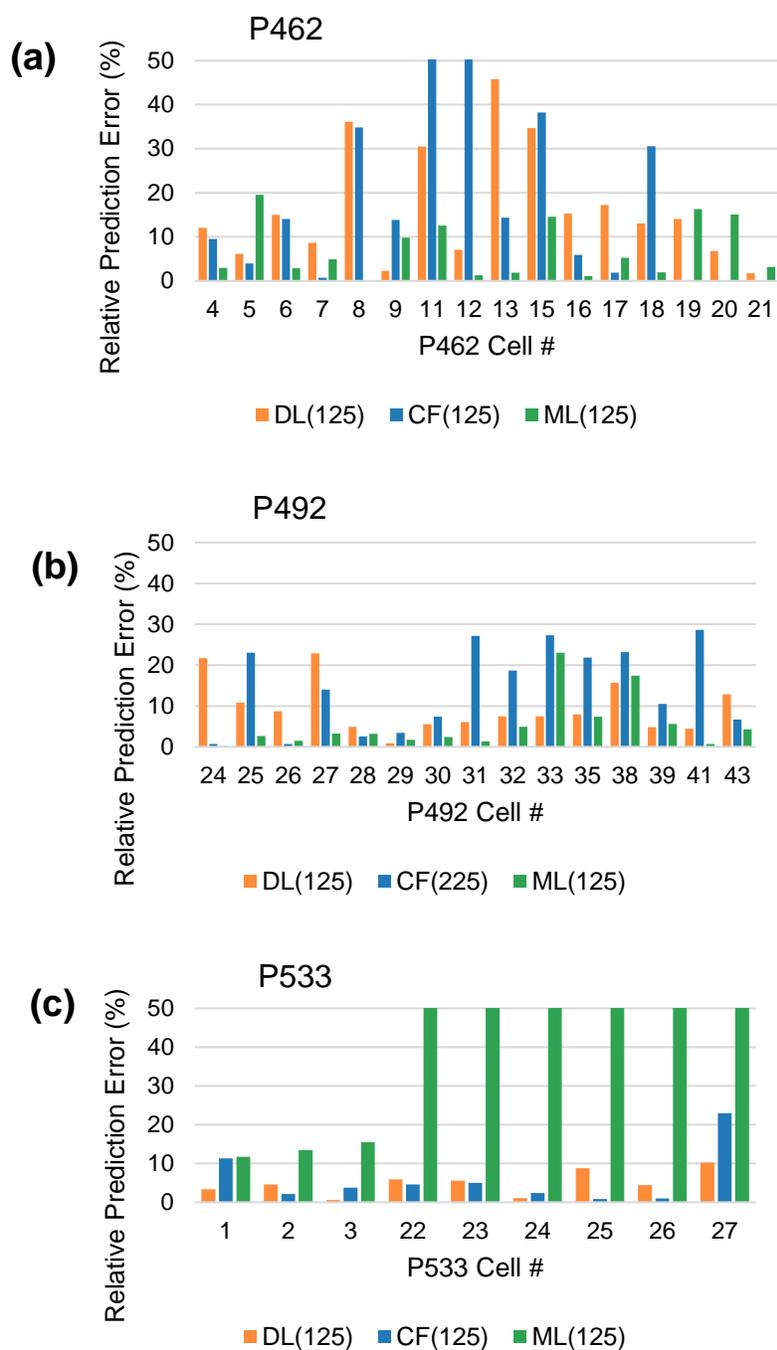
(a) P462

(b) P492

(c) P533

Fig. S2. Summary of relative prediction error in LLI for cells of the three test packs within this study. Please also see legend for Figure 6 for further details.
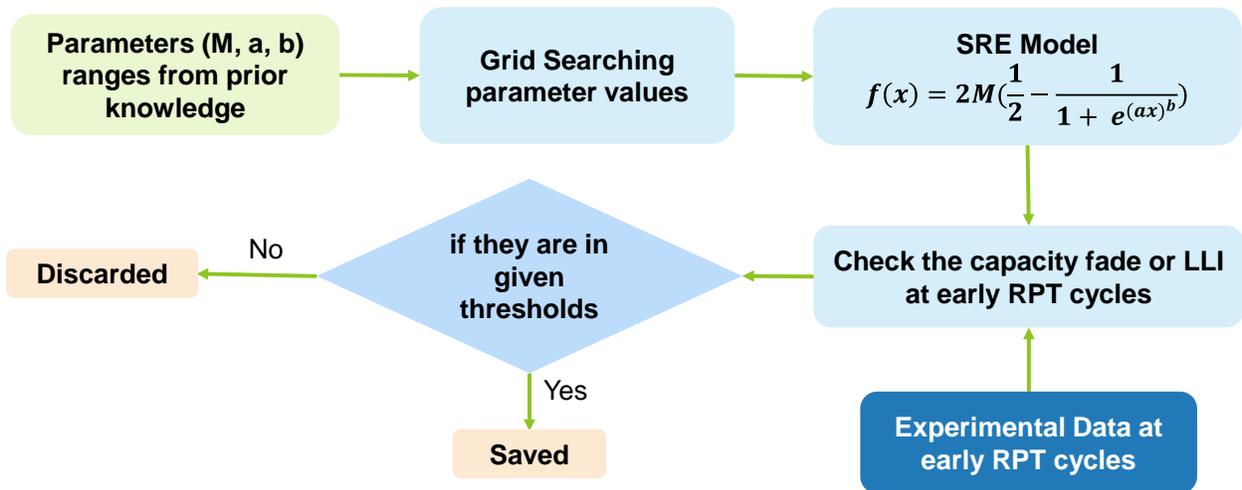
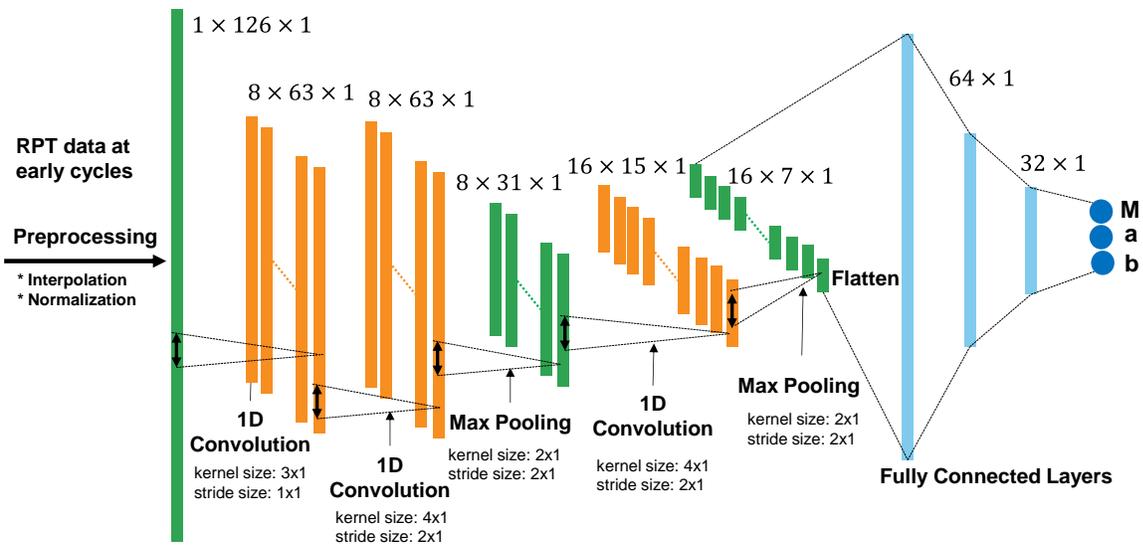**Fig. S3.** Diagram of experimental data informed synthetic data generation process.

Fig. S4. DL architecture with 1-D convolutional neural network for SRE parameters predictions.
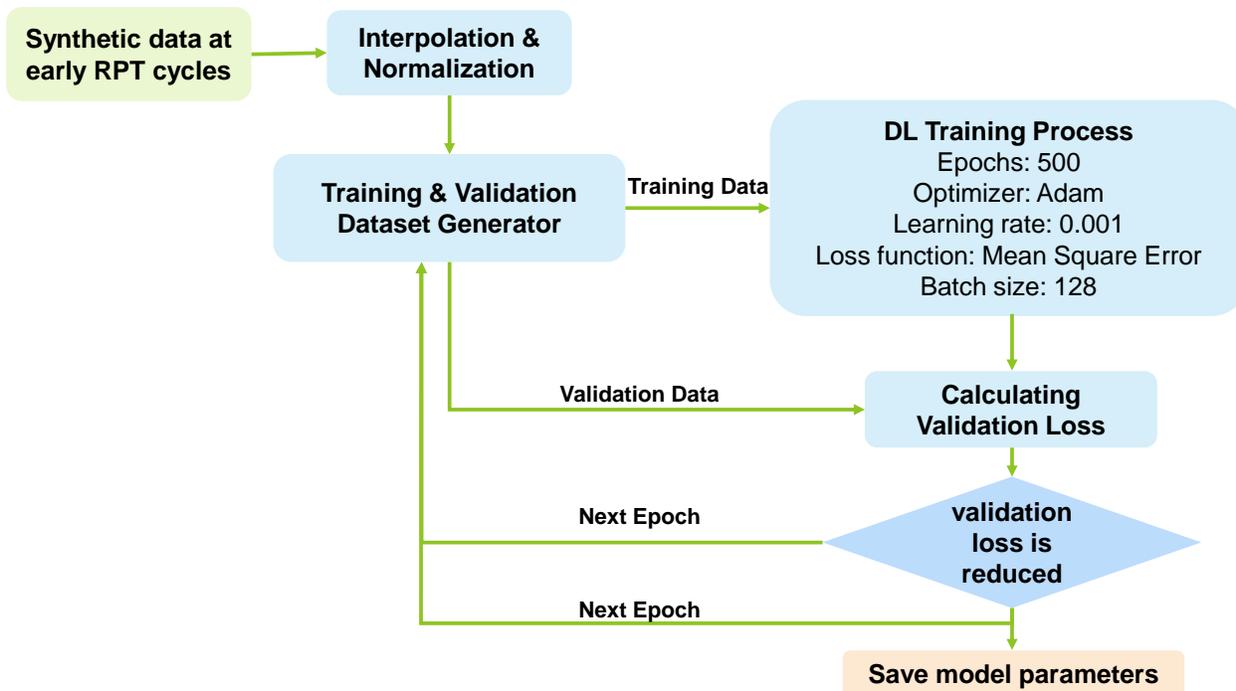
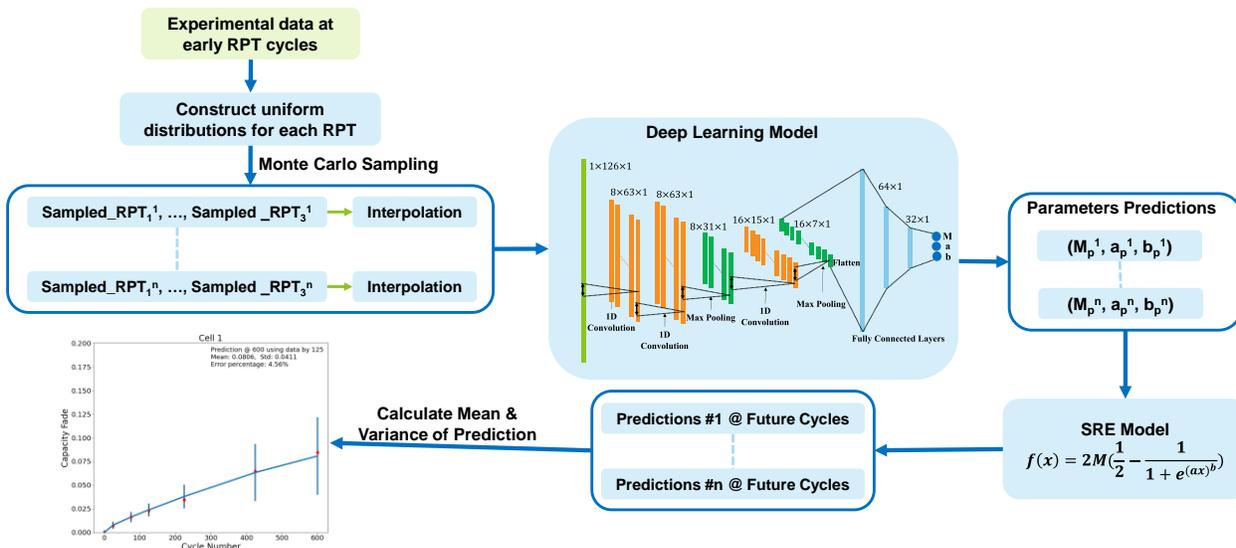**Fig. S5.** Diagram to illustrate the workflow of the DL training process using SRE synthetic data.

Fig. S6. Monte Carlo based DL prediction framework for capacity fade or LLI prediction using early cycle experimental data.

Table S1. Summary of cell design metrics for the three battery packs investigated.

	Moderate loading cells ($L_{moderate}$)	Low loading cells (L_{low})
Pack number	P462, P533	P492
Cathode design		
Type	Toda NMC 532	
Area, cm^2	Single layer pouch cell: 14.1 Coin cell for LAM testing: 1.27	
Coating Thickness, µm	71	62
Foil Thickness, µm	20	20
Coating Loading, mg/cm2	18.57	11.4
Porosity, %	35.6	33.1
Anode design		
Type	Superior Graphite SLC 1506T	
Area, cm^2	Single layer pouch cell: 14.9	
Coating Thickness, µm	70	57
Foil Thickness, µm	10	10
Coating Loading, mg/cm^2	9.38	6.38
Porosity, %	38.2	37.4
Separator design		
Type	Celgard 2320	
Area, cm^2	Single layer pouch cell: 15.5	
Thickness, µm	20	
Porosity, %	39	

Table S2. Summary of cell designs and cycling protocols.

	Moderate loading cells ($L_{moderate}$)	Low loading cells (L_{low})
Pack numbers	P462, P533	P492
Formation profiles	1. Tap charge to 1.5 V and hold for 15 min 2. Rest at OCV for 12 h 3. Charge 3 cycles at C/10 between 4.1V and 3.0 V and discharge 3 cycles at C/2 between 4.1 V and 3.0 V 4. Hold at 20% state of Charge for 6 h	
Aging profiles	1. Charge at designated charging rate profiles until 4.1 V. Hold at 4.1V until the total charging time reaches 10 min 2. Rest at OCV for 15 min 3. Discharge at C/2 4. Rest at OCV for 15 min	
C-rates (highest throughout the charging protocol)	1C, 4C, 6C, 7.5C, 9C	1C, 4C, 6C, 9C
Designated charging profiles and the C-rates in the profiles *Details of the profiles in Table S3	P462 Cells: CC-CV 4C: Cell 16, 17, 18 6C: Cell 04, 05, 06 2-Step 4C: Cell 19, 20, 21 6C: Cell 07, 08, 09 MS5 7.5C: Cell 13, 15 9C: Cell 11, 12 P533 Cells: CC-CV 1C: Cell 01, 02, 03 Intermittent profile A Cell 23, 24, 25 Intermittent profile B Cell 26, 27, 28	P492 Cells: CC-CV 1C: Cell 24 - 29 4C: Cell 30, 31 6C: Cell 32, 33, 35 9C: Cell 39, 41, 43
RPT intervals (cycle numbers where RPTs occurring)	P462 Cells: 0, 25, 50, 75, 100, 125, 175, 225, 275, 325, 375, and 450 P533 Cells: 0, 25. 75, 125, 225, 425, and 600	P492 Cells: 0, 25, 75, 125, 225, 325, 375, 450, and 600
Total number of life cycles	P462: 450 P533: 600	P492: 600

Table S3. Detailed descriptions of cycling profiles.

Designated charging profiles	Detailed descriptions
CC-CV	Charge at CC to 4.1V, then go to CV and continue charging until 60 min (if using 1C), 15 min (4C), 10 min (≥ 6C).
2-Step	A two-step charging protocol that charges at a CC until 4.1V is reached and then step down to a lower CC step until the total charge time reaches 10 min.[4,18]
MS5	A five-step charging protocol that maximizes the magnitude of CC charging portion by providing a stepwise charging profile. The c-rate ramps down in 5 steps until total charging time reaches 10 min.[4]
Intermittent profile A	A mixture of charging protocols throughout battery life cycles. 1 in every 5 cycles using the 7.5C Voltage Ramp profile (below), rest of the cycles using 1C CC-CV.
Intermittent profile B	A mixture of charging protocols throughout battery life cycles. Cycles 1-5 and 26-30 in each cycle life set using 7.5C Voltage Ramp profile (below), remainder using 1C CC-CV.
7.5C Voltage Ramp[17]	- Step 1: CC charge step, charge at 7.5C (240 mA) until 4.0V - Step 2: Increase voltage at 2 mV/10s until 4.1V - Step 3: Go to CV step@4.1V if 4.1V is reached Step 2 and continue charging until total 10 min charge time has reached.

Table S4. The total number of cycles per scheduled RPT.

Cell Type	RPT 1 Cycle #	RPT 3 Cycle #	RPT 4 Cycle #	EOL Cycle #
P462	25	100	125	450
P492	25	125	225	450
P533	25	125	225	425