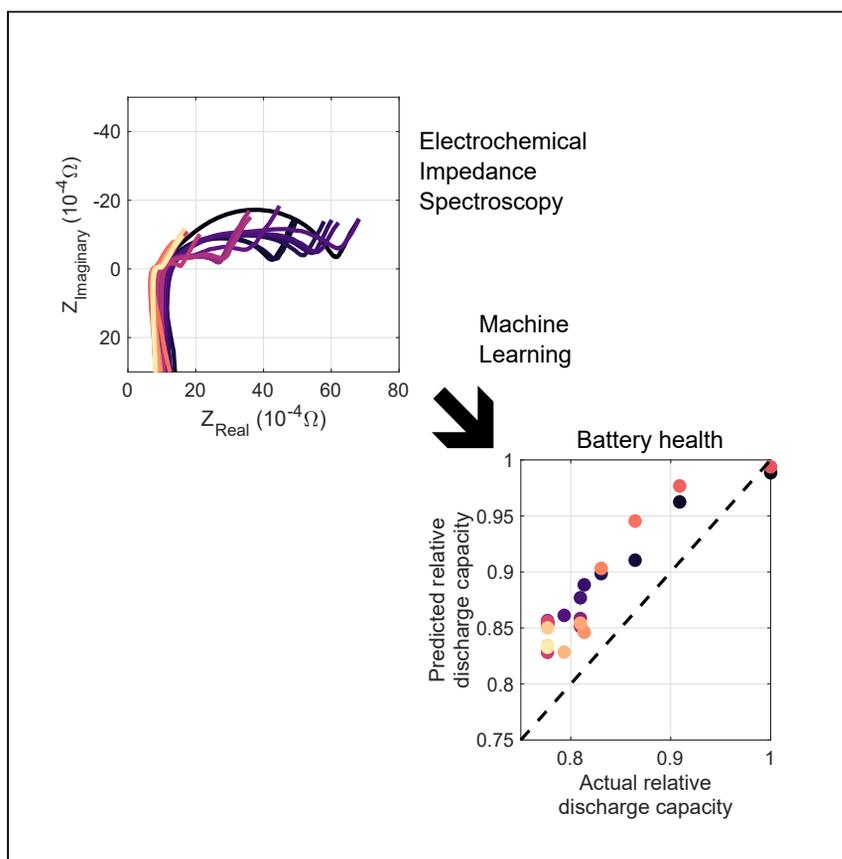


Article

Predicting battery capacity from impedance at varying temperature and state of charge using machine learning



Gaspar et al. demonstrate prediction of battery capacity using electrochemical impedance spectroscopy data recorded under varying conditions of temperature and state of charge. A variety of methods for featurization of impedance data are tested using several machine-learning model architectures to rigorously investigate the limits of using impedance to monitor battery health.

Paul Gaspar, Andrew Schiek, Kandler Smith, Yuta Shimonishi, Shuhei Yoshida

paul.gaspar@nrel.gov, pauljgaspar@gmail.com

Highlights

Machine learning is used to predict battery health from impedance spectroscopy

Impedance recorded under varying conditions can predict cell health with 1.9% error

Featurization methods for impedance are explored via machine-learning pipelines

Interrogation of models reveals key frequency ranges for monitoring battery health



Article

Predicting battery capacity from impedance at varying temperature and state of charge using machine learning

Paul Gasper,^{1,3,*} Andrew Schiek,¹ Kandler Smith,¹ Yuta Shimonishi,² and Shuheï Yoshida²

SUMMARY

Prediction of battery health from electrochemical impedance spectroscopy (EIS) data can enable rapid measurement of battery state in real-world applications without using additional sensors or time-consuming performance measurements. However, deconvoluting the effect of capacity, state of charge, and temperature on EIS response is complicated analytically. Here, various machine-learning models, such as linear, Gaussian process, random forest, and artificial neural network regression, are utilized to predict capacity from EIS using hundreds of capacity, direct current (DC) resistance, and EIS measurements recorded under varying conditions of health, temperature, and state of charge (SOC). Several feature extraction and selection methods from traditional electrochemical analysis and statistical modeling are explored using machine-learning pipelines. EIS data from just two frequencies can accurately predict capacity, and interrogation shows that the optimal set of frequencies is not usually intuitive. Best results are achieved with an ensemble model, which predicts battery capacity with a mean absolute error of 1.9% on data from unobserved cells.

INTRODUCTION

Accurate monitoring of battery states like temperature, state of charge (SOC), resistance, and capacity is crucial for ensuring the safety and reliability of lithium (Li)-ion battery energy storage systems used in electric vehicles or for stationary energy storage systems. Although certain states, like temperature, can be monitored using relatively cheap sensors, other states, like battery capacity, are measured using time-consuming diagnostic tests that may take anywhere from several hours to days,² making these methods infeasible for use in real-world applications. Battery capacity may be predicted online via forecasting life models, which may use reduced-order algebraic expressions,^{3–5} machine learning methods,^{6,7} or neural network approaches.⁸ Other methods, such as state observer algorithms, like Kalman filters, may track battery health by updating an internal model based on real-time observations.^{9,10} But directly monitoring battery health via diagnostic measurements may be more tolerant to model error or detect imminent battery failure because of external or internal events. Rapid diagnostic tests, such as direct-current (DC) internal resistance (DCIR) measurements,^{11,12} pseudo-random binary pulse tests,^{13–15} and electrochemical impedance spectroscopy (EIS)^{1,16} can be performed in only a few minutes or less and require a fraction of the energy and power compared with a full charge and discharge. The main obstacle preventing use of rapid diagnostic tests to diagnose battery health is the complex relationship between health metrics like discharge capacity and resistance; the resistance of a battery is highly sensitive

¹National Renewable Energy Lab, 15013 Denver West Parkway, Golden, CO, USA

²DENSO Corporation, Kariya, Aichi 448-866, Japan

³Lead contact

*Correspondence: paul.gasper@nrel.gov, pauljgasper@gmail.com
<https://doi.org/10.1016/j.xcrp.2022.101184>



to temperature and SOC, and the capacity and resistance of the battery may evolve non-monotonically as a battery is used. Thus, using resistance to monitor battery capacity requires accounting for the effect of temperature and SOC on resistance, effects that may change as the battery degrades.

The effect of temperature, SOC, and capacity on the EIS response of Li-ion batteries has been researched extensively. Waag et al.¹¹ described not only the effect of temperature, SOC, and capacity on EIS but also the effect of resting time as well as the effect of all of these on DCIR measurements. Analysis of the effect of battery state on the EIS measurement is often assisted using equivalent circuit models (ECMs). Heins et al.¹⁷ fit EIS data with an ECM derived by analysis of the distribution of relaxation times, qualitatively mapping ECM parameters versus capacity and SOC after exploring temperature dependence on a cell at beginning of life (BOL). Mc Carthy et al.¹⁸ reported similar relationships between ECM parameters and battery states as Heins et al.¹⁷ and analyzed correlations between battery state (health, temperature, SOC) to ECM parameters and EIS components at single frequencies: the real component (Z_{Real}), the imaginary component ($Z_{\text{Imaginary}}$), and phase angle ($\angle Z$) of the impedance. Gantenbein et al.¹⁹ developed quantitative relationships to describe the variation of ohmic, charge-transfer, and diffusion resistances versus temperature and SOC, essentially using EIS and voltage relaxation data to develop a complete battery model.

Time-resolved DC pulse data (voltage data collected frequently enough to capture dynamic changes) can be equivalent to impedance. Barai et al.¹² showed equivalence between DC and alternating current (AC) measurements as long as the timescales are the same, whereas other researchers have demonstrated that time-resolved current and voltage data from a series of binary DC pulses can be mathematically transformed to calculate an impedance spectrum.^{15,20} This implies that time-resolved DC measurements can be used in conjunction with or in replacement of EIS when the measurement contains current switching at varying frequencies. Practically speaking, AC-DC conversion hardware often introduces ripple currents,^{21,22} which could be leveraged to monitor impedance at specific frequencies in real-world battery modules. Impedance may also be directly recorded by battery management systems at the cell or string level²³ and used as input data for models predicting battery state. Other types of DC modulation can be used to monitor the entire impedance spectrum throughout the charge or discharge of a battery.²⁴ Other methods for recording impedance at the pack or module level have also been reported in the academic literature.^{25–27}

Although these studies help to explain the qualitative relationships between battery state and EIS response, they do not clearly identify useful strategies for making quantitative predictions. Luckily, there is substantial prior work, prompting several recent review articles on battery health detection,²⁸ relationships between battery health diagnostics like EIS and capacity,^{16,29} internal temperature measurement,^{28,30} and SOC monitoring.²⁸ One of the major topics in these reviews as well as in individual studies is the challenge of extracting useful features from the EIS data to predict a single battery state while trying to minimize or correct the influence of other states; e.g., predicting capacity while minimizing error induced by varying cell temperature. There are two obvious approaches for dealing with this challenge: select features that only correspond to a single state (sensitive to capacity but not to temperature; e.g., avoid the problem), or deconvolute the effects of each battery state on the feature (e.g., correct the problem). For example, Galeotti et al.³¹ used the ohmic resistance ($>10^4$ Hz) with a linear regression model to predict battery

capacity, based on the insight that the ohmic resistance was not very sensitive to SOC but still responsive to changes in capacity. Zhang et al.³² also fit EIS measurements with an ECM but attempted to correct for SOC, temperature, and capacity interactions by using the charge-transfer resistance, R_{CT} (characteristic frequency near 10^1 Hz), as an input to a non-linear function dependent on SOC, temperature, and R_{CT} :

$$q = \alpha \cdot \exp(\beta_1 + \beta_2 \cdot \text{SOC} + \beta_3 \cdot \text{SOC}^2 + \gamma \cdot T) \cdot R_{CT} \quad (\text{Equation 1})$$

where q is the relative discharge capacity; α , β_1 , β_2 , β_3 , and γ are fitting parameters; and SOC and T are the SOC and temperature. Other studies utilize domain expertise to pull out features in the Nyquist or Bode representations of the EIS data. Schuster et al.³³ used the values of Z_{Real} at the y axis intercept ($\sim 10^3$ Hz) and at the minimum of the low-frequency valley (often near 10^{-1} Hz), predicting the capacity of cells operating under calendar aging, mild cycle aging, and harsh cycle aging conditions using linear models, but they were not able to develop a model that extrapolated well from one test condition to another. La Rue et al.³⁴ used the minimum $Z_{\text{Imaginary}}$ value at frequencies of less than 1 rad/s (greater than 0.16 Hz) to extract the peak of the semi-circular feature present in most battery EIS measurements to estimate SOC at 0°C or 23°C, but they did not account for battery health. Du et al.¹⁵ used a high-dimensional lookup table consisting of the Z_{Real} and $Z_{\text{Imaginary}}$ values recorded from 1–31 Hz at 1-Hz intervals versus temperature, minimizing the city-block distance between measured impedance and the pre-recorded lookup values to predict temperature, which, they suggest, can account for SOC interactions. Zhu et al. use the maximum value of $|Z|$ from across the impedance spectra, finding that it is a useful predictor of SOC while being relatively insensitive to aging or temperature.²⁴

The most common feature extraction approach is to only use data from one or two frequencies of interest rather than extracting information from the entire EIS curve. Love et al.³⁵ chose a frequency qualitatively based on a statistical analysis of the EIS sensitivity to SOC and health, selecting a frequency of 316 Hz to detect cells that were overcharged while avoiding SOC interactions. Liebhart et al.³⁶ used a mathematical approach, estimating the partial derivatives of the impedance with respect to capacity and temperature to determine the frequency most sensitive to capacity but least sensitive to temperature, finally selecting the phase angle, $\angle Z$, at 30 Hz; they also attempted to correct for any bias induced by temperature at this frequency by assuming that the measured phase angle is a superposition of the capacity and the temperature effects, modeling the temperature component using a fourth-order polynomial,

$$\begin{aligned} \gamma_q &= \angle Z(30 \text{ Hz}) - \gamma_T \\ \gamma_T &= a + bT + cT^2 + dT^4 \end{aligned}$$

where $\angle Z$ is the measured phase angle at 30 Hz; γ_q and γ_T are the capacity and temperature components of $\angle Z$, respectively; T is the temperature in Kelvin; and a , b , c , and d are fitted parameters. The remaining capacity effect, γ_q , could then be used to predict capacity using a linear model or a lookup table. Mc Carthy et al.³⁷ addressed the opposite problem, predicting internal temperature from impedance while accounting for battery capacity and SOC effects by qualitatively analyzing the correlation between temperature, SOC, and battery capacity versus measurement frequency for the real, imaginary, and phase components of the impedance, choosing 200 Hz as the frequency most sensitive to temperature but least sensitive to capacity and SOC. Wang et al.³⁸ selected two frequencies by qualitative analysis, using $\angle Z$ at 79.4 Hz to measure temperature using a lookup table

and assuming that SOC is known via coulomb counting or some other method and then predicting capacity using the magnitude of the impedance, $|Z|$, at 10 Hz using a lookup table versus $|Z|$, temperature, and SOC. Caposciutti et al.³⁹ used qualitative analysis to propose that any frequency between 10^{-1} and 10^1 Hz could be used to predict battery capacity but without accounting for any temperature or SOC effects. Measurement frequency can also be selected to minimize estimation error when some underlying model for the sensitivity of the impedance versus temperature, SOC, or other experimental variables exists, as demonstrated by Beelen et al.⁴⁰ to estimate temperature with less than 1°C error across a wide range of temperatures and SOCs.

Because of the complexity of handling the convoluted effects of capacity, temperature, and SOC on EIS or resistance measurements, machine learning methods have been proposed in recent years to tackle the issue. Zhang et al.¹ recorded EIS data 9 times during each cycle on cells cycling with constant ambient temperatures of 25°C, 35°C, or 45°C, varying the EIS measurement current, temperature, SOC, and rest time, and then used the raw values Z_{Real} and $Z_{\text{Imaginary}}$ at each frequency as input features to a Gaussian process regression (GPR) model with an automatic relevance determination (ARD) kernel to adaptively weight each feature. They concluded that the critical frequencies for predicting capacity from EIS are 17.8 and 2.16 Hz; however, this only accounts for influence of temperature and capacity because they split the EIS measurements recorded throughout each cycle into separate datasets, avoiding the need to account for varying current, SOC, and rest time. La Rue et al.⁴¹ used the least absolute shrinkage and selection operator (LASSO) regularization method⁴² to determine features for predicting SOC for lithium iron-phosphate and graphite cells from the $Z_{\text{Imaginary}}$ data, finding that 4 frequencies spaced between 10^{-1} – 10^0 Hz accurately predicted SOC. Ran et al.¹⁴ utilized a K-means clustering algorithm to classify cells as healthy, aged, or end-of-life (EOL) using a variety of features extracted from a series of 3 charge/discharge pulses. The first component from principal-component analysis (PCA) of the current-voltage time-series data, DCIR calculated using the voltage difference after 10-s DC pulse duration at varying current, and the open-circuit potential of the battery as estimated by fitting of the final voltage relaxation curve with an exponential relaxation model. Various deep-learning approaches have also been suggested. Kim et al.⁴³ used a generative adversarial network (GAN) architecture to predict capacity using the data published by Zhang et al.,¹ utilizing a 2-dimensional latent state representation learned by the GAN to predict capacity via a GPR model instead of the using the raw EIS data. Rastegarpanah et al.⁴⁴ utilized a convolutional neural network to predict the capacity of aged Nissan LEAF batteries from impedance measurements conducted at varying SOCs, comparing the results with an artificial neural network (ANN) that uses ECM parameters as inputs and finding that the ANN model using raw impedance data as input performs better than the ANN using ECM parameters as input. Messing et al.⁴⁵ independently conducted a nearly identical experimental and modeling effort to predict SOC rather than capacity, also finding that an ANN performed better when using raw EIS data rather than ECM parameters as input.

One unstated drawback to essentially all studies conducted on battery state estimation via EIS or resistance measurements is a relative lack of data. Because of limited data, most of the summaries above ignored SOC, temperature, or health effects on the EIS during the experimental and modeling components of their work. Not only this, but most research in this area is based on data from fewer than 10 cells, with some developing detailed models based on data from only a single battery. No studies explore a wide variety of aging conditions. The reason for this lack of data

is probably simple: battery aging and performance characterization studies are expensive and time consuming, so conducting aging with a variety of aging conditions while simultaneously measuring battery capacity, DCIR, and EIS at varying SOC and temperature is cost, time, or channel prohibitive. Even for the highly cited work by Zhang et al.,¹ which shares data from over 20,000 distinct EIS measurements, the data are recorded from only 12 cells, each undergoing an identical aging procedure at only 3 distinct temperatures. This lack of variance in the aging trajectories and EIS measurement conditions hinders the development of modeling methods that could be useful in real-world applications and makes it impossible to develop a realistic understanding of the possibilities and limitations of using EIS or DCIR measurements to predict battery state.

This work shares data from a comprehensive aging study conducted by DENSO Corporation on large-format Li-ion prismatic cells intended for use in electric vehicles. The aging study comprises 32 cells aged for up to 500 days and 2,500-equivalent full cycles (final capacity between 98 and 65%) under varying conditions: calendar aging at varied temperature and SOC; cycle aging at varied temperature, average SOC, depth of discharge, duty cycle, and C rate; and drive cycle aging. A summary of the study and a capacity fade model developed using data from the first 300 days of aging has been published previously by Smith et al.⁴ Cell performance was monitored not only by capacity checks but also by DCIR and EIS measurements recorded at varying SOC and temperature with a total of 489 EIS measurements and 2,584 DCIR measurements. An exploration of this dataset reveals the challenge of using resistive health metrics like DCIR or EIS to monitor battery capacity. The dataset is then used to train machine-learning pipelines with linear, GPR, or random forest (RF) model architectures, each using a variety of feature extraction and/or selection algorithms, with the goal of predicting battery capacity using impedance data. Models of each architecture using the most promising features are then retrained using Bayesian hyperparameter optimization via cross-validation on the training set. An ANN is also trained using the raw impedance data as input for comparison with the other machine learning methods. The lowest error models achieve about 2% mean absolute error (MAE) on test data, learning to account for the influence of temperature and SOC on the impedance while still making relatively accurate predictions of capacity. The lowest-error models of each architecture are then combined into an ensemble model, which results in 1.9% MAE on the test set. Analysis of all results suggests that the most reliable approach for training a model to predict capacity from EIS is to perform an exhaustive search for the best combination of two frequencies; the frequencies that are most critical for monitoring capacity for these cells lie near 10^0 and 10^3 Hz, but the optimal frequencies vary by model and are not necessarily intuitive; two features with low correlation often perform better than two features with high correlation on unseen data. Finally, guidelines for data acquisition and model identification are presented, with an eye toward remaining challenges for monitoring battery state using EIS in real-world applications.

RESULTS AND DISCUSSION

Predicting capacity using resistance is difficult because both evolve non-monotonically as a battery degrades

In real-world scenarios, battery aging occurs under a wide variety of environmental conditions and use cases. Monitoring battery capacity via resistance or impedance would be simple if battery capacity and resistance were insensitive to the use, but this is not borne out by accelerated aging studies. The relationship between capacity and resistance is further complicated because capacity and resistance health metrics may vary non-monotonically during cell lifetime, dependent on the aging stress.⁴

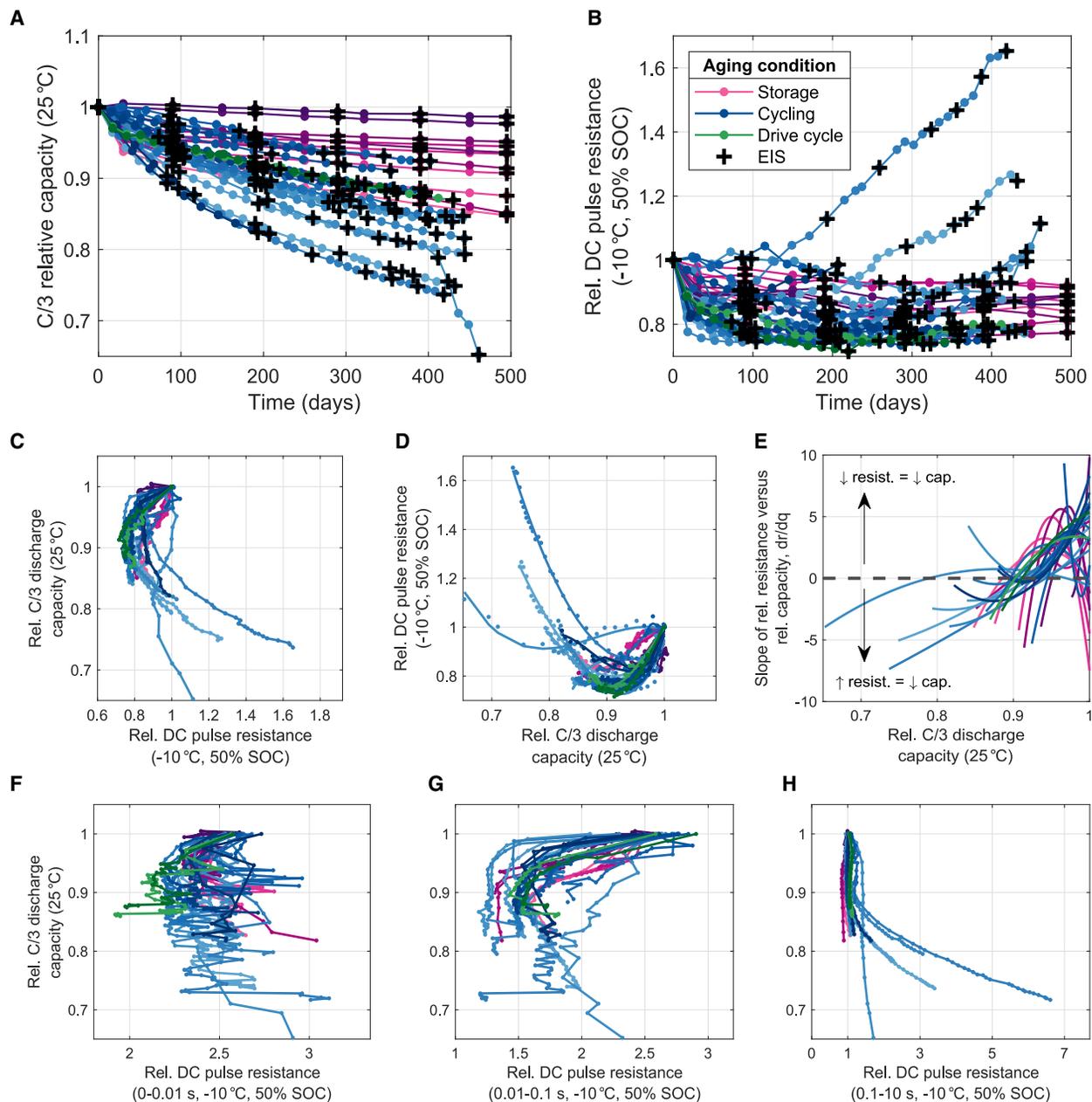


Figure 1. Comparison of discharge capacity and DC pulse resistance throughout battery lifetime

(A and B) Evolution of (A) relative C/3 discharge capacity at 25°C and (B) relative DCIR at -10°C and 50% SOC. Black markers denote capacity checks where EIS data were recorded at 50% SOC and -10°C and 25°C . Once during the lifetime of each cell, between 300 and 400 days, impedance was also recorded at 0°C and 10°C . For about 1/3 of the cells, further impedance measurements were recorded at varied temperatures and SOCs after completion of the aging tests.

(C–E) Relationships between relative capacity and resistance. (C) Relative capacity versus relative DCIR. (D) Data in (C) fit with a cubic polynomial. (E) The slope of the cubic fits in (D) versus relative capacity.

(F–H) Relative capacity versus DCIR, where DCIR is calculated using various time windows after the start of the DC current pulse. (F) 0–0.01 s time window. (G) 0.01–0.1 s time window. (H) 0.1–10 s time window.

The DCIR data in (B)–(E) are calculated using the entire 0–10 s time window.

The relative C/3 discharge capacity and relative DCIR at -10°C at 50% SOC from this aging study are plotted versus time in Figure 1; the top plots (Figures 1A and 1B) show relative discharge capacity and DCIR evolution from cells aging under

static storage or cycling conditions as well as cells aging under a worldwide harmonized light vehicle test protocol (WLTP) drive cycle. Capacity for all cells is, for the most part, monotonically decreasing, but some cells show a 1%–2% capacity increase near BOL (Figure 1A). This capacity increase likely occurs for all cells,⁵ but capacity was not measured with enough frequency early during cell life. DCIR varies positively and negatively throughout aging, decreasing by up to 30% before stabilizing for a long period (Figure 1B). Some cycling cells show a dramatic DCIR increase leading up to EOL. Details about the experimental procedures for cell aging and performance measurements are described in the [supplemental experimental procedures](#) and shown in [Figures S1](#) and [S2](#).

The decrease in resistance during early life may be surprising because other works constantly increasing resistance with decreasing capacity,^{46,47} but this behavior has been noted before⁵ and reported previously for this dataset.⁴ This initial decrease in resistance may be due to cracking of LiNiMnCoO₂ (NMC) positive electrode particles during the first 10–100 cycles, which exposes fresh surface area, decreasing cell resistance. Some studies may “equilibrate” cells prior to aging by cycling all cells in a test for a few tens of cycles, defining BOL as after cell equilibration, which would obfuscate this effect. However, subjecting cells to equilibration cycles is not achievable in real-world applications.

Cell capacities and resistances are directly compared in [Figure 1C](#). Considering the goal of predicting capacity using resistive health metrics, there are several key aspects:

- Early capacity loss is nearly always associated with decreasing resistance.
- In middle of life, capacity continues to decrease, and the resistance is stable and eventually increases.
- Although cell resistance is nearly always increasing as cells approach EOL, the onset of resistance increase has widely varying capacity values.
- A single value of DCIR may correspond to more than one capacity value throughout cell lifetime.

Although the value of capacity does not directly correspond to the value of resistance, it may be that changes to capacity map directly to changes in resistance. To study this, the relative resistance as a function of relative capacity was fit using a cubic function (Figure 1D). The slope was then calculated analytically (Figure 1E). For most cells, the inflection from decreasing to increasing resistance occurs around 90% relative capacity, but this varies from 95% to 80% relative capacity. Although the resistance increases with decreasing capacity for nearly all cells at EOL, the magnitude of that slope is not strongly associated with a specific capacity value. Thus, as shown in [Figures 1C–1E](#), not only is the value of resistance not a good predictor of capacity, but the change of resistance is also not a good measure of the change of capacity.

The time domain used for calculating DCIR also substantially affects the relationship between capacity and resistance. Very short time windows correspond to ohmic resistances and facile charge-transfer behaviors, middling duration time windows correspond to slower charge-transfer behaviors and porous electrode effects, and long time windows correspond to solid-state diffusion of Li and changing stoichiometry of the electrodes. DCIR is calculated by fitting Ohm’s law to a specific time window of the voltage response after beginning a discharge current pulse. In [Figures 1B–1E](#), the plotted values correspond to DCIR, which was calculated after

10 s (0–10 s) versus the discharge pulse current. In Figures 1F–1H, results are plotted using varying time windows: 0–0.01 s (Figure 1F), 0.01–0.1 s (Figure 1G), and 0.1–10 s (Figure 1H). At the shortest time window, 0–0.01s, the resistance for all cells is relatively stable within measurement accuracy, decreasing by perhaps 5% in early life and then only obviously increasing for a few cells as they approach very low capacity values. In the next time window, 0.01–0.1 s, the resistance decreases by up to 50% in early life before saturating or increasing. Finally, in the longest time window (0.1–10 s), resistance is nearly constant for most cells, with a few cycling cells showing dramatic increases. This variance in the relationship of DCIR and capacity throughout cell lifetime as a function of the time window clearly contains a substantial amount of information, but recording and visualizing these data is a complex challenge.^{11,12} For measuring and visualizing the evolution of time-dependent resistance, EIS is a more useful method, with many analysis techniques to draw from in the literature.

Electrochemical impedance is sensitive to battery health but also to temperature and SOC

Similar trends as those observed in the DCIR are reflected by the EIS data. EIS trends across various time domains are often described using the terms ohmic resistance, polarization resistance, and diffusion resistance, associated with the minimum value of Z_{Real} at high frequencies, the width of the semicircular feature at moderate frequencies, and the slope of the impedance of the diffusion tail (near 45° straight line observed at low frequencies) versus frequency, respectively. EIS measurements recorded at -10°C and 25°C for a storage cell and a cycling cell, respectively, are shown in Figure 2, colored by capacity. The predominant trends of these two cells, which are qualitatively representative of trends observed in the other 30 cells, include the following:

- Decreasing polarization resistance during early aging, obvious for EIS data recorded at -10°C (Figures 2A and 2C)
- Increasing polarization resistance with decreasing capacity, mild for cells with only slight degradation, such as calendar-aged cells (Figures 2A and 2B), and severe for others (Figures 2C and 2D)
- Increasing ohmic resistance over cell life, obvious for EIS data recorded at 25°C (Figures 2B and 2D)

Qualitatively, the EIS changes relatively little during the middle portion of cell life ($\sim 95\%$ – 85% relative capacity), suggesting that capacity prediction in this regimen will be challenging because the impedance may be more affected by cell-to-cell variation or small changes in the measurement conditions than by any degradation mechanism. However, nearer to EOL, several cells have dramatic changes in the magnitude and shape of the EIS response, reflected by substantial increases to DCIR in Figure 1B, so detecting EOL should be comparatively straightforward. The physiochemical roots of changes to the shape and magnitude of the EIS of NMC-Gr Li-ion batteries are well described in other studies^{11,17,19} and are not repeated here.

The prediction of battery capacity using EIS in real-world applications is confounded by the strong dependence of the EIS response of any given cell on the cell temperature and SOC. EIS behavior versus temperature and SOC is shown in Figures 2E–2H at BOL and after aging. EIS trends versus temperature (Figures 2E and 2F) remain relatively unchanged throughout lifetime; the effect of temperature on the magnitude of the impedance is exponential, whereas the effect of capacity on impedance

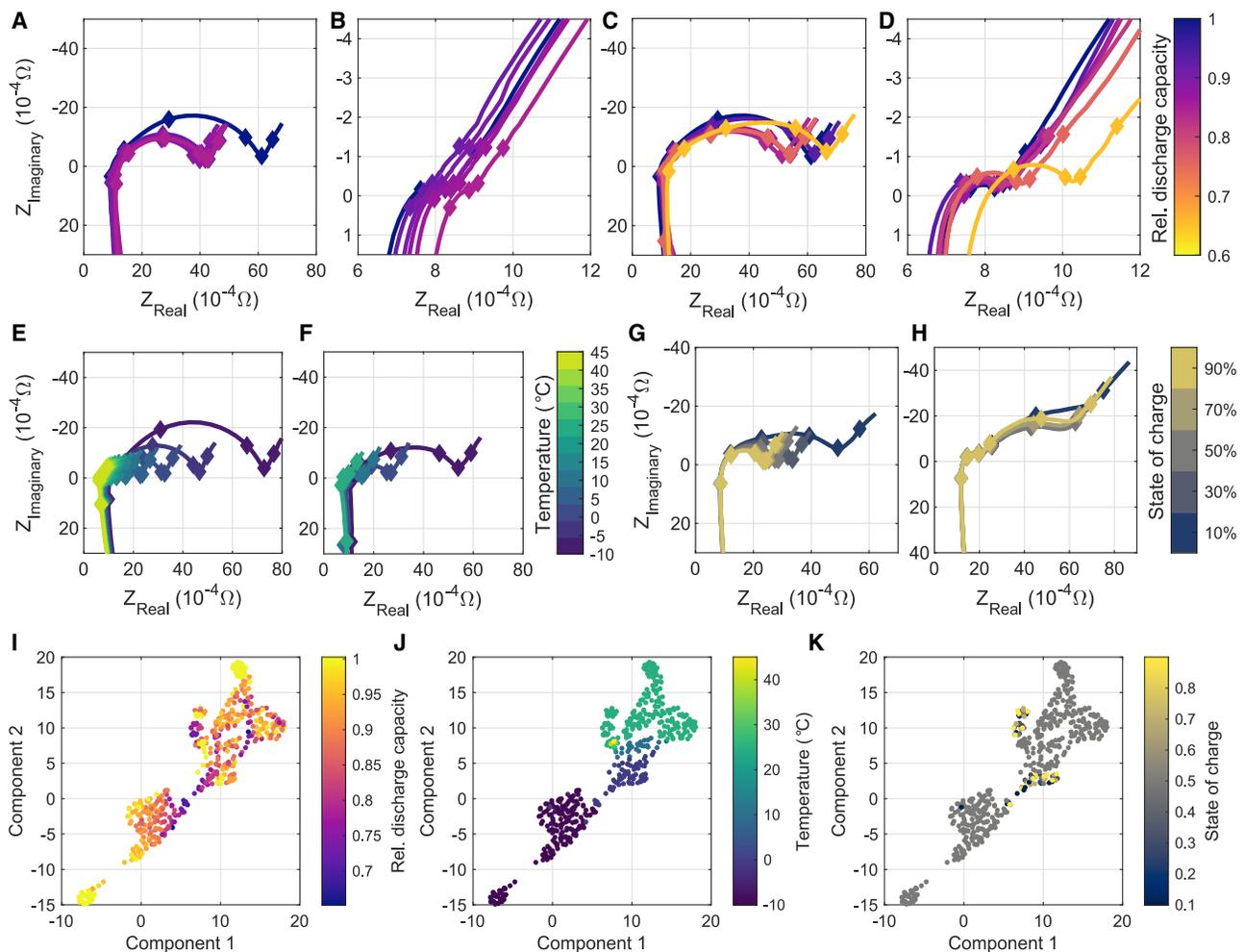


Figure 2. EIS sensitivity to discharge capacity, temperature, and SOC throughout battery lifetime

(A–D) EIS sensitivity to discharge capacity for a calendar aging cell at (A) -10°C and (B) 25°C and a cycling aging cell at (C) -10°C and (D) 25°C . Diamond markers correspond to frequency decades from 10^3 Hz (bottom left) to 10^{-2} Hz (top right).

(E and F) EIS sensitivity to temperature at 50% SOC at (E) BOL and (F) after 25% capacity loss.

(G and H) EIS sensitivity to SOC at 25°C at (G) BOL and (H) after 30% capacity loss.

(I–K) Two-component UMAP dimensionality reduction of the entire EIS dataset, colored by (I) relative capacity, (J) temperature, and (K) SOC.

is proportional, so impedance is always more sensitive to temperature than degradation state throughout cell life. In contrast, the effect of SOC on the magnitude of the impedance at BOL (Figure 2G) is similar to that of aging (Figures 2A and 2C), making qualitative distinction between SOC and aging effects difficult. Also, aging seems to affect the sensitivity of EIS to SOC because aged cells (Figure 2G) appear to be less sensitive to SOC than at BOL. This suggests that EIS may not be as effective for SOC estimation after substantial degradation, but there is not enough variation in the SOC of EIS measurements in this dataset to arrive at a clear conclusion on this.

The results in Figures 2A–2H seem to paint a clear picture of the relationships between EIS, capacity, temperature, and SOC, but it is very difficult to visualize data from multiple cells at a time without making them uninterpretable. To visualize high-dimensional data, unsupervised dimensionality reduction algorithms can be used. Dimensionality reduction attempts to map high-dimensional data to a lower

Table 1. Modeling pipelines explored in this work

Pipeline	Features (raw impedance values or extracted)	Feature selection
0	N/A (dummy model)	N/A (dummy model)
1A	$Z_{\text{Real}}, Z_{\text{Imaginary}}, Z , \angle Z$	N/A
1B		single frequency
1C		two frequencies
1D		correlation search ⁵³
1E		SISSO ⁵⁴
1F		embedded feature selection ^a
2A	statistical features	N/A
2B		correlation search
2C		SISSO
3A	PCA	N/A
3B	UMAP	N/A
4	graphical features	N/A
5	ECM parameters	N/A

All approaches were tested using linear, GPR, and RF model architectures. Pipeline 0 is a "dummy model" used as a baseline for evaluating model performance. SISSO refers to the "sure independence screening and sparsifying operator" algorithm.⁵⁴

^aEmbedded feature selection uses the LASSO⁴² algorithm for linear models and the automated relevance determination (ARD) form of the squared-exponential kernel for GPR models. RF models were not trained for pipeline 1F because no common embedded feature selection method exists.

dimensional space while preserving the relative variance between individual measurements. One such method is the universal manifold approximation and projection (UMAP) method;⁴⁸ other well-known methods include PCA⁴⁹ and t-distributed stochastic neighbor embedding (t-SNE).⁵⁰ UMAP is used to project all of the EIS data in this work from 138 dimensions (Z_{Real} and $Z_{\text{Imaginary}}$ data at 69 frequencies) to two. The result of this projection is shown in Figures 2I–2K, with points colored by capacity (Figure 2I), temperature (Figure 2J), and SOC (Figure 2K). Temperature seems to have the strongest effect because EIS measurements are distinctly clustered by measurement temperature. The effect of relative discharge capacity is not as distinct, but there are still notable trends; for instance, the values of the second UMAP component seem to be highly correlated with relative discharge capacity for EIS measured at -10°C . There is not enough variation of the EIS measurement SOC throughout aging to resolve the effect of SOC on the entire dataset.

Extracting features from EIS data to predict battery capacity

Many various methods to extract features from EIS data to predict battery state have been proposed in prior work. These strategies are usually designed to maximize the sensitivity of any predictive model to the desired target state (capacity, temperature, or SOC). For instance, Galeotti et al.³¹ fit EIS data with an ECM, using the fitted value of the ohmic resistance as a predictive feature. Other approaches, such as extracting statistical features, have shown success for battery lifetime prediction from charge and/or discharge data.⁵¹ Features can also be selected directly from the raw data, as in the many studies that use impedance values from just one or two frequencies rather than the entire measured spectrum.^{35–39} These various feature extraction and selection methods could also be applied concurrently, constructing what is often referred to as a modeling pipeline. Here, a wide variety of feature extraction and selection methods are explored to determine an effective approach for predicting capacity from impedance (Table 1). Because model architecture can also have a significant effect, this work trains linear, GPR, and RF regression algorithms for all pipelines. An ANN architecture was also tested, using the FastAI library to automatically handle defining of the network architecture and optimization of training

hyperparameters.⁵² As a baseline, all modeling results are compared with a “dummy model,” which simply uses the mean of the capacity from the training set to make predictions. In total, there are 2,615 pipelines trained for each model architecture, most of which are from the exhaustive search for the best two frequencies (2,346 pipelines). The rationale for each method reported in Table 1 is described throughout the rest of this section.

Pipelines 1B–1F explore various approaches for selecting impedance values from a small subset of frequencies as features. Using impedance from just 1 or 2 frequencies has been proposed for predicting cell safety,³⁵ capacity,^{36,38} and temperature.³⁷ This method is practical because recording impedance at a small number of frequencies can be done using simpler hardware than required for a fully detailed EIS measurement, can be conducted in a short amount of time, and has low energy and power requirements. The challenge lies in how to select one, two, or any number of frequencies from a larger set of available data. Figure S3 shows the correlation between the Z_{Real} and $Z_{\text{Imaginary}}$ impedance and the relative discharge capacity as a function of frequency for the training and test sets, considering the whole dataset as well as only EIS recorded at -10°C or 25°C . Considering all data, the correlation between impedance and capacity never exceeds about ± 0.4 . There are frequency ranges where the correlation between impedance and capacity is nearly ± 0.75 when the EIS measurement temperature is constant at -10°C or 25°C , but the magnitude or sign of these correlations is not constant with respect to temperature, so most of the correlation is lost when EIS temperature is allowed to vary.

Even high correlation does not guarantee good model performance, and this simple qualitative analysis does not give a straightforward way to select the best combination of two, three, or more frequencies. The best one, two, or more frequencies can be determined via an exhaustive search, as conducted by pipelines 1B and 1C for one and two frequencies, respectively. However, it is combinatorically complex to exhaustively search for larger sets. To select more possible features, feature selection algorithms can be used. One intuitive approach is to select features that have high correlation with the target (i.e., peaks in Figures S3A and S3B), and, when selecting more than one, to avoid redundancy with any prior features. A feature selection algorithm leveraging this intuition, referred to here as the “correlation search,” was proposed by Greenbank and Howey⁵³ and is used in this work (pipeline 1D). Another approach is to select a high correlation feature, train a model, and then select a new feature that is highly correlated with the residual error of the prior model; the feature selection algorithm SISSO⁵⁴ conducts this approach, considering sets of potential features rather than single features at each iteration (pipeline 1E). Finally, the best set of features can be directly optimized using embedded feature selection algorithms, such as LASSO,⁴² penalized regression for linear models, or ARD kernels for GPR models (pipeline 1F). LASSO is not guaranteed to find the optimal subset of features, and ARD kernels for GPR models may require substantially more computational resources than standard kernels.

Pipelines 2A–2C calculate common statistical measures from the raw EIS data. This approach worked well for predicting battery lifetime from voltage data in the work by Severson et al.⁵¹ (after a key feature engineering step) and has been repeated for other datasets since.^{55,56} This work considers statistics such as the interquartile range (IQR), mean absolute deviation (MAD), median absolute deviation (MdAD), mean, median, range, and variance. The correlation between these statistics and the relative capacity, calculated separately on Z_{Real} , $Z_{\text{Imaginary}}$, $|Z|$, and $\angle Z$ data, are shown in Figures S3C–S3F. Of all the statistics, only the MdAD of the phase angle

of $Z(\angle Z)$ has an absolute correlation greater than 0.4 for the training and test data. Pipelines 2B and 2C use the correlation search and SISO feature selection algorithms, respectively, to choose a subset of the 28 statistical features.

Pipelines 3A and 3B use component scores from the dimensionality reduction techniques PCA⁴⁹ and UMAP⁴⁸ as features for capacity prediction. PCA is deterministic; i.e., exactly repeatable, computationally fast, and very well understood. UMAP often results in qualitatively better clustering for data with non-linear relationships between variables in high-dimensional data but is not deterministic, so model performance may vary slightly even with the exact same dataset and hyperparameter values. For PCA models (pipeline 3A), the first 10 components were used, whereas for UMAP (pipeline 3B), a grid search was used to tune three hyperparameters.

Pipeline 4 uses graphical features from the Nyquist plots to predict capacity. Expert analysis of EIS data often uses notable shapes or points from Nyquist and/or Bode plots of the EIS data, tracking these features versus experimental variables like temperature or SOC¹¹ or using them as inputs for a predictive model.³³ Graphical features could include the x axis intercept or peak of a semicircular region on a Nyquist plot. These and other features of interest are shown in Figure S4 and used in pipeline 4. Although this approach is straightforward to implement on small datasets, there are challenges when using EIS recorded under a widely varying set of conditions. For instance, in this dataset, the magnitude of inductance and ohmic resistance are quite large compared with polarization and diffusion effects when EIS is recorded at 25°C (Figures 2B and 2D) and vice versa when EIS is recorded at -10°C (Figures 2A and 2C), and it is not possible to define logical rules that pick out changes in shape or slope that remain consistent across all 32 cells throughout aging and at varying measurements of temperature and SOC. Thus, data are selected by picking out a small number of obvious features and then simply grabbing other points spaced evenly throughout the frequency space.

Pipeline 5 uses fitted ECM parameter values to predict capacity. Prior studies have used the parameters of ECMs fit to EIS data as features for predicting battery state.^{31,32} However, as with the graphical approach, utilizing a single ECM across a dataset with substantial variance presents many difficulties. A not-so-secret but rarely mentioned aspect of fitting ECMs to EIS data is that the optimization often has many local minima, requiring researchers to tailor their fits by changing initial guesses or setting bounds to result in a “good” fit; e.g., a model that fits the data accurately with parameter values that vary according to physical intuition. This sort of procedure does not lend itself well to automation. In this work, a relatively simple ECM (Figure S4) was used, trading some accuracy for reliability, because all spectra in the dataset could be fit using the same set of initial guesses. ECM parameters were then used as features for model training by pipeline 5.

Predicting capacity using EIS recorded at varied temperature and SOC

The MAEs from training, cross-validation, and test sets from all 2,615 pipelines for the linear, GPR, and RF model architectures are reported as bivariate histograms in Figure 3; a well-performing model should have low training error and low error on unseen data (MAE_{Test} and MAE_{CV}), corresponding to the lower left corner of each plot. The baseline dummy model MAE is plotted as black lines. Model architecture seems to play a significant role in the relative effect of feature engineering. Linear models are quite sensitive to extraction and selection of features, with a wide range of MAEs, whereas the GPR and RF models seem to be intrinsically better regularized; i.e., they perform better when making predictions on unseen data than

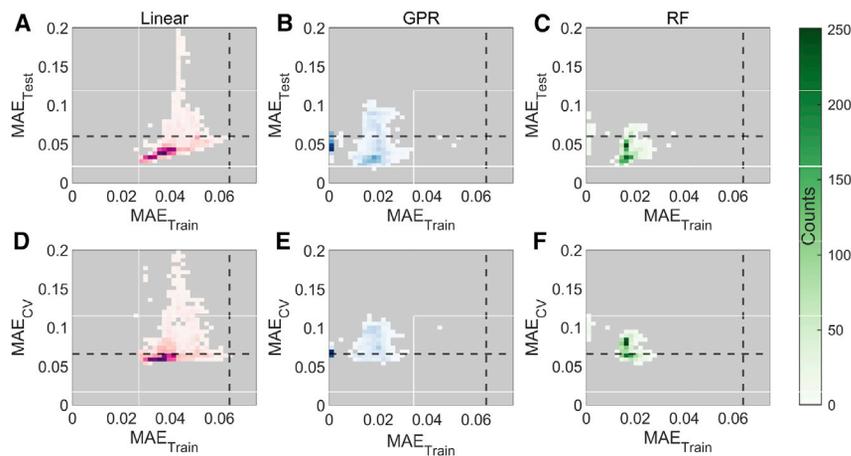


Figure 3. Performance of all machine-learning pipelines on training, cross-validation, and test sets
(A–C) Bivariate histograms of the MAE on the training set versus MAE on the test set for (A) linear, (B) GPR, and (C) RF model architectures. Black lines across each plot denote the performance of the baseline dummy model. A gray background color is used to clearly denote regions with zero counts.
(D–F) MAE on the training set versus cross-validation MAE for the (D) linear, (E) GPR, (F) and RF model architectures.

the linear models for almost all pipelines. However, there are also several GPR and RF pipelines that are clearly overfit to the data; that is, models where the training MAE is extremely small compared with the test MAE, and the test MAE is greater than that of the best-performing pipelines. Overall, most trained pipelines perform better than the dummy baseline on the training and test sets, but only a few perform better than baseline on the cross-validation set. There seems to be a clear boundary for model performance on unseen data, with models from each architecture approaching a minimum test MAE of about 2% and a minimum cross-validation MAE of about 5%. The ANN resulted in a test MAE of 2.73%, performing well but not exceptionally; neural networks, empirically, do not seem to perform better than traditional machine-learning models on tabular data⁵⁷ and also seem to perform worse on “small” datasets than traditional machine-learning models.⁵⁸

All model architectures benefit from extraction and/or selection of high-quality features. Figures 4, S5, and S6 compare the results of GPR, linear, and RF models from all pipelines described in Table 1. Using the raw data without any modification (pipeline 1A) performs surprisingly well on the test set for the RF model, but the GPR and linear models are overfit. Searching for the single best frequency (pipeline 1B) works decently but always performs worse than models using two frequencies. The exhaustive search for the best combination of two frequencies (pipeline 1C) is the most stable approach, always resulting in a model with high performance on the test and cross-validation sets. Selecting from more than two frequencies using the correlation search (pipeline 1D) or SISSO (pipeline 1E) occasionally results in high-performing models; a more detailed comparison of pipelines 1A–1E is reported later. Of the remaining model pipelines (1F–5), most struggle to reliably produce models with low test MAE and low cross-validation MAE. Only PCA or UMAP dimensionality reduction techniques (pipelines 3A and 3B) routinely result in models that perform better than baseline on the test and cross-validation splits. Use of statistical features (pipelines 2A–2C) usually results in average performance but is consistently outperformed by other pipelines. Pipelines 4 and 5, which use the graphical or ECM features that are routinely used in EIS analysis by domain experts, perform poorly.

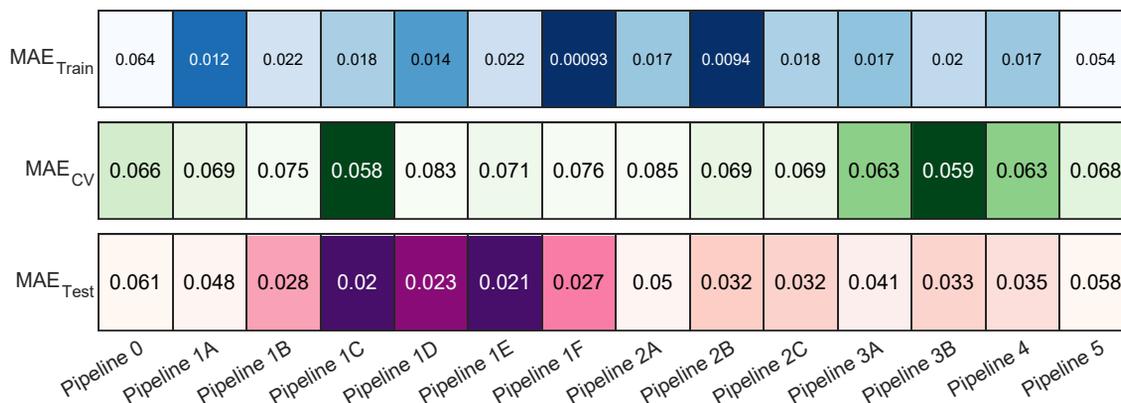


Figure 4. MAE of GPR models for the training, cross-validation, and test splits

For pipelines with more than one trained model, the results for the model with the lowest test error are shown in each column. Low MAE for each row is denoted by a dark background, and high MAE is denoted by a light background.

For pipelines selecting a small subset of frequencies (pipelines 1B–1E), the selected frequencies matter. Figure 5 shows MAE_{Test} and MAE_{Cross-validation} for single- and double-frequency models. For single-frequency models (Figure 5A), prediction errors for test and cross-validation splits are highest at frequencies less than 1 Hz, suggesting that there is too much noise in the low-frequency data to make accurate predictions or that the low-frequency EIS response has been affected by changing of the open-circuit voltage curve because of aging, introducing variance that cannot be accounted for by these models. Low-frequency impedance is usually dominated by solid-state diffusion resistance from the positive electrode material and cell capacitance effects, both of which may be sensitive to the exact stoichiometry of the electrodes, which drift over time during aging, leading to noise unrelated to the health of the cell.^{2,11,19} The minimum test MAE for all model architectures seems to occur near 2 Hz, with another minimum observed for some architectures between 10² and 10³ Hz. These frequencies (10⁰ Hz and 10²–10³ Hz) tend to be dominated by charge-transfer related processes, which may be substantially affected by cycling (Figure 2). Prediction errors again increase as the frequency approaches the maximum frequency of 3.5 × 10⁴ Hz.

Figures 5B–5D show contour maps of MAE_{Test} for the double-frequency models (pipeline 1C). Like the single-frequency case, models using only very low (<10⁻¹ Hz) or very high (>10³ Hz) perform poorly. Low-error regions for all architectures appear near the middle of the frequency range, between 10⁰ and 10³ Hz, and often extend away from the diagonal; that is, use frequencies that are not very close to one another. Using two frequencies spaced one or more decades apart may enable the model to detect and account for multiple degradation modes because the relationship between cell resistance and capacity changes depending on the time domain of the measurement (Figures 1G and 1H). Two frequencies spaced apart from one another may also help to correct for the effect of temperature or SOC.

Correlation search (pipeline 1D) and SISO (pipeline 1E) feature selection algorithms are used to search for subsets of frequencies larger than 2. A comparison of frequencies selected by pipelines 1B–1E is made in Figures S7–S9 for each model architecture. Selected frequencies are also compared with the correlation of those features with the relative discharge capacity. The optimal single frequency or two frequencies are non-intuitive; i.e., they do not necessarily have high correlation to

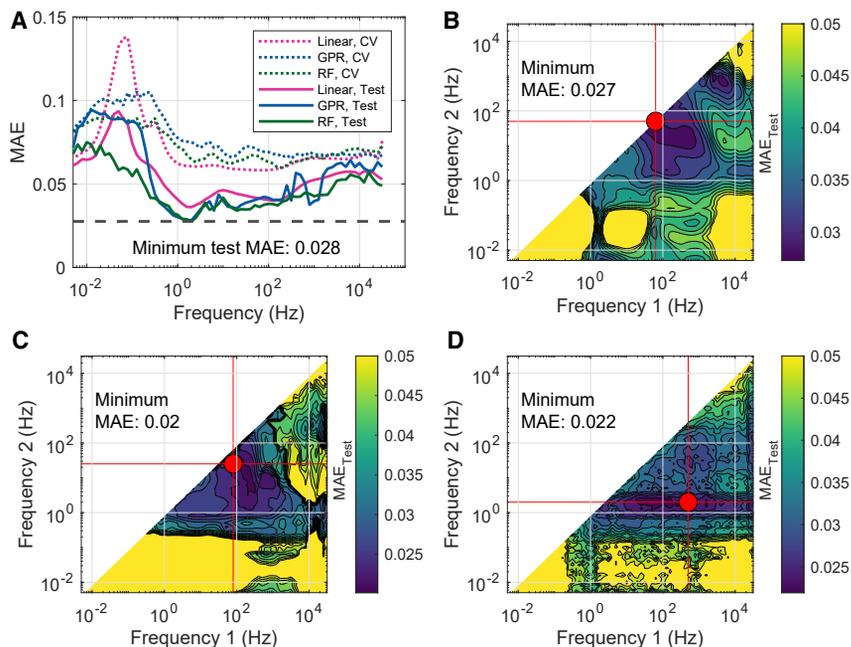


Figure 5. Importance of selected frequencies for single- and double-frequency models for all model architectures

(A) MAE of single-frequency models versus frequency. Test MAE is reported using a solid line, and cross-validation MAE is reported using a dashed line. Linear model results are shown in pink, GPR model results in blue, and RF model results in green.

(B–D) Contour plots of MAE_{Test} versus frequencies for double-frequency models using (B) linear, (C) GPR, and (D) RF model architectures. Note the change of the color bar scale for each axis. A red marker and lines denote the location of the minimum error point on each contour.

the relative discharge capacity. The correlation search and SISSO feature selection algorithms utilize frequencies in the range of 10^0 – 10^3 Hz, qualitatively agreeing with the analysis shown in Figure 5. SISSO and the correlation search select at least one magnitude-related feature (Z_{Real} or $|Z|$) as well as at least one phase-related feature ($Z_{Imaginary}$ or $\angle Z$). SISSO also prefers fitting fewer features from the explored range (2–10 features), selecting features from just 3 frequencies, suggesting that data from only a small subset of key frequencies are enough for predicting cell health.

Retraining best performing pipelines with model hyperparameter optimization and weights

Pipelines 1A, 1C, and 1E were further tuned by retraining using Bayesian hyperparameter optimization. For pipeline 1A, various model hyperparameters may help regularize model performance; for instance, use of ridge regression for linear models. For pipeline 1C, the two best-performing double-frequency sets were hyperparameter tuned. Pipeline 1E was tuned using the optimal SISSO feature selection hyperparameters identified in the initial search. Hyperparameter optimization was conducted via the MATLAB function “bayesopt” for all tunable model hyperparameters over 30 iterations; see supplemental experimental procedures for details. Hyperparameter optimization for linear and RF model architectures was conducted with and without weighting the training data. Weighted regression is used to help address the relatively high prevalence of EIS data that were measured at -10°C and 25°C compared with data measured at 0°C and 10°C by giving each row of the training data a weight so that all temperatures contribute equally to the calculated loss. This imbalance of

Table 2. Performance of various modeling pipelines on test data

Model	Pipeline	Defaults	Hyperparameter optimized	Hyperparameter optimized (weighted)
Baseline	0	6.05%	–	–
Linear	1A	28.95%	*2.26%	3.04%
	1C (best)	2.73%	2.72%	2.64%
	1C (second best)	2.74%	2.63%	2.63%
	1E	3.04%	3.04%	–
GPR	1A	4.79%	4.14%	–
	1C (best)	*2.01%	2.11%	–
	1C (second best)	2.02%	2.36%	–
	1E	2.14%	2.15%	–
RF	1A	3.07%	2.66%	2.97%
	1C (best)	2.19%	2.07%	*2.00%
	1C (second best)	2.25%	2.21%	2.09%
	1E	2.61%	2.36%	2.29%
ANN	1A	–	2.73%	–

Test MAE of linear, GPR, and RF models using pipelines 1A, 1C, and 1E with default model parameters, after 30 iterations of hyperparameter optimization, and after hyperparameter with weights to address dataset imbalance. ANN model performance is also reported. The best-performing model for each model architecture is denoted with a preceding asterisk.

the data is an issue because real-world battery systems would not predominantly be used at temperatures of -10°C and 25°C ; rather, there would be a more uniform distribution of temperatures. Other methods for handling dataset imbalance include undersampling overrepresented classes or oversampling under-represented classes.⁵⁹

Table 2 reports the MAE on the test set of linear, GPR, and RF with and without hyperparameter optimization and weighting as well as the ANN performance; various ANN hyperparameters were tuned during model training. Overall, hyperparameter optimization tends to improve model performance for all architectures. Weighting of the linear and RF models again tends to slightly improve performance. The best linear model uses all raw impedance data and ridge regression to regularize model performance, and the best RF model is weighted and uses a minimum leaf size of 2 and a maximum number of 355 splits and randomly samples all 8 input features (2 frequencies, 4 features per frequency) when constructing each split. Although these approaches improve results generally, the best overall models from each architecture are not necessarily discovered after hyperparameter optimization or weighting; the best GPR model uses MATLAB defaults. These pipelines are referred to in the next section simply as the best linear, best GPR, and best RF models.

Interrogation of best performing models

The best performing linear, GPR, and RF models are interrogated to analyze their underlying behavior and point out important differences between each model architecture. The simplest model to interrogate is the best linear model because the coefficient for each feature, β , can be directly plotted as a function of frequency (Figure 6). On average, the Z_{Real} and $\angle Z$ features are weighted more heavily than the $Z_{\text{Imaginary}}$ or $|Z|$ features. The weights of features sometimes seem intuitive; for instance, values of Z_{Real} near 10^3 Hz have a large negative correlation to relative discharge capacity and also have a relatively large negative coefficient value, but the largest feature weights actually occur in regions with low correlation to relative discharge capacity.

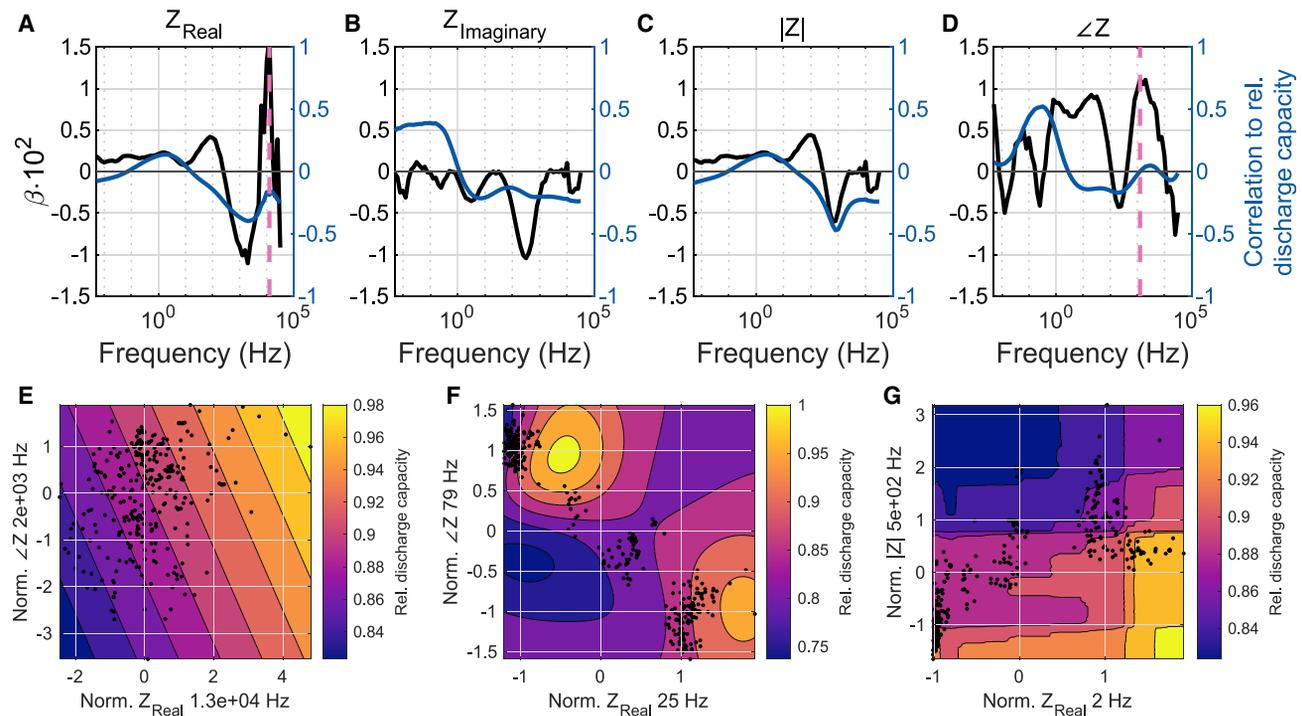


Figure 6. Interrogation of best-performing linear, GPR, and RF models

(A–D) Coefficients (β) for each feature for the hyperparameter-optimized linear model using all impedance features. All y axes are matched to aid visual comparison. The secondary y axis on each plot, in blue, reports the correlation between each feature and the relative C/3 discharge capacity. The pink vertical lines in (A) and (D) denote features used for bivariate partial dependence analysis in (E).

(E–G) Bivariate partial dependence plots of the (A) best linear, (B) best GPR, and (C) best RF models. Because only two features can be shown, highly sensitive features from each were selected: (E) the Z_{Real} and $\angle Z$ features with the largest coefficient values and (F and G) the two most sensitive features of 8, as determined by univariate partial dependence analysis. Black dots correspond to the values of features in the training set.

Interrogating non-parametric model architectures such as GPR and RF requires a different approach. One method to investigate the sensitivity of a model prediction to the values of the input features is partial dependence analysis. Partial dependence is a model-agnostic analysis method that interrogates the relationship between a selected set of input features and the model response by varying those features across the observed range while holding other features constant. For the best GPR and RF models, there are 8 input features (Z_{Real} , $Z_{\text{Imaginary}}$, $|Z|$, and $\angle Z$ at two frequencies), so univariate partial dependence was conducted on all features to determine the two features that have the highest average sensitivity, and then bivariate partial dependence was conducted on these. Plots from the univariate partial dependence analysis are reported in [Figures S10](#) and [S11](#) for the best GPR and best RF models. For the best linear model, the Z_{Real} and $\angle Z$ features with the largest coefficient values were used as an example.

For the best linear model ([Figure 6E](#)), the partial dependence analysis reveals that the underlying model structure is simply a plane, and the diagonal contour lines denote that the model is response sensitive to both features. This is only a 2D “slice” of the model; the actual model is a 276-dimensional surface, with the exact values for each coefficient plotted in [Figure 6](#). For the best GPR model ([Figure 6F](#)), the analysis reveals that the GPR has learned non-monotonic, non-linear relationships between the impedance and the relative discharge capacity. However, the GPR still results in a smooth function because GPR models utilize Bayesian methods to balance accuracy with simplicity and penalize functions that do not vary smoothly. The best

Table 3. Average and maximum errors compared for the best models and the ensemble model

Model	MAE _{Train}	MAE _{CV}	MAE _{Test}	MaxAE _{Train}	MaxAE _{CV}	MaxAE _{Test}
Baseline	6.4%	6.6%	6.1%	24.5%	24.5%	12.0%
Linear	2.6%	6.0%	2.3%	21.8%	27.6%	8.2%
GPR	1.8%	5.8%	2.0%	*8.4%	29.5%	11.4%
RF	*1.2%	6.4%	2.0%	15.6%	*22.8%	11.3%
Ensemble	1.6%	*5.6%	*1.9%	11.3%	26.6%	*7.2%

MAE and MaxAE for the training, cross-validation (CV), and test splits for the dummy baseline, best linear, best GPR, and best RF models as well as an ensemble model that predicts the average value of the three best models. The lowest value in each column is denoted with a preceding asterisk.

GPR model appears to have learned two local maxima for the predicted capacity, related to data recorded at -10°C (bottom right) and 25°C (top left), and there is less variation between these two regimens. For the best RF model (Figure 6G), the contours vary in discrete steps. These discrete steps may help the RF model to reduce error when the underlying relationship is not smooth, which would not be possible to model using a linear or GPR model but could also lead to unexpected predictions when the unseen data are outside of the distribution of the training data. Further interrogation of these models is shown in Figure S12, which compares the distributions of predicted relative discharge capacity values to the true values to examine systematic errors related to the true capacity value or the EIS measurement temperature.

The performance of the best three models is shown in Table 3, which reports the MAE and maximum absolute error (MaxAE) of each model on the training, cross-validation, and test splits. One more model is reported: an ensemble model, which predicts the average value of predictions by the best linear, GPR, and RF models. By taking the average of many models, each with their own systematic bias, an ensemble of models may be more accurate than any one individual model when model errors counteract one another. The ensemble model has the lowest MAE on the cross-validation and test splits of any model tested here, as well as the lowest MaxAE on the test split.

Predictions made by the ensemble model on the training, cross-validation, and test splits are shown in Figure 7. The performance on the training and test splits is excellent, with all predictions following the trends of the observed capacity fade across cells aged under varying load (resting for calendar cells, constant-current cycling, or cycling using a drive cycle) at different temperatures, average SOCs, and cycling protocols. Predictions are relatively uniformly distributed around the parity line, indicating low systematic error, although there is a clear tendency of the training set to underpredict capacity near BOL and overpredict capacity near EOL. Performance on the cross-validation split is substantially worse, and most predictions are noisily distributed between 85% and 95% relative discharge capacity, but some specific cells appear to be well predicted because the density of points is higher near the parity line. The cross-validation split may be inherently harder because all models perform substantially worse on the cross-validation split (Figure 3). Real-world performance of the ensemble model would likely lie somewhere between the performance of the test and cross-validation splits.

Guidelines for real-world battery monitoring using impedance

Monitoring battery state in the real-world using impedance relies first and foremost on acquisition of high-quality measurements, capacity and EIS, recorded under a widely varying set of battery states. Although the dataset presented in this work is

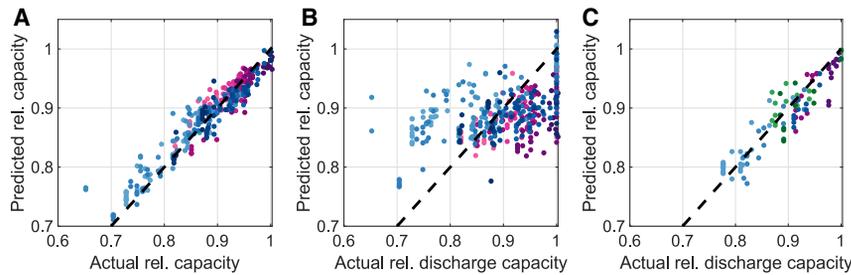


Figure 7. Predictive performance for the ensemble model using the best linear, GPR, and RF models

(A) Parity plot for the training set. Markers are colored by battery cell, with pink tones denoting calendar-aged cells, blue tones denoting cycle-aged cells, and green tones denoting cells tested using a WLTP drive cycle, as in Figure 1. The black dashed line on each graph, referred to as the parity line, denotes a perfect prediction.

(B) Parity plot for the cross-validation splits.

(C) Parity plot for the test split.

substantial, there are several limitations that future efforts may address. This dataset does not fully explore the effect of SOC on the EIS measurement throughout cell lifetime, and extreme values of SOC could result in poor predictions from models trained on these data. EIS measurement of temperature and SOC are often conducted at a single value throughout the lifetime of cells during lab testing so that the trends can be visualized clearly, but varying the EIS measurement of SOC and temperature throughout aging is probably the most efficient way to acquire a balanced dataset; visualizing data trends at constant states (temperature, SOC, or health) could then be done using a generative model.⁴³ This work also does not explore the effect of relaxation time on the EIS, which can substantially affect the EIS response;¹¹ this lack of data may negatively affect model performance on data recorded during field use. Recording EIS multiple times as the battery relaxes after cycling and/or after setting the battery SOC would mitigate this issue. Capacity and impedance data recorded during electric vehicle maintenance could be used to increase the amount of training data over time.

Just as the training data should be recorded under varying conditions to develop more extensible data-driven models, so should the target variables. Here, EIS is used only to predict battery performance for a single type of use; specifically, the relative discharge capacity at a C/3 rate and 25°C ambient temperature between a constant set of voltage limits. However, if the training data are available, then data-driven models may be trained to learn not just the state of the battery but also predict battery response to varying future loads by incorporating information on the operation of the battery into the predictive model. Initial work demonstrating this type of data-driven approach has been published by Jones et al.⁶⁰ This type of approach may especially help data-driven estimation of complex problems, such as state of power, which depends not only on cell health but the load on the battery.

After data acquisition, model identification needs to be performed, considering a wide variety of possible models and the practical limitations the model may impose on real-world implementation. From the variety of modeling pipelines explored here, the most reliable way to accurately predict battery capacity from EIS is to record data at as many frequencies as possible, covering a critical range between 10^0 and 10^3 Hz at a minimum and then selecting 2–4 key frequencies using an exhaustive search or a feature selection algorithm, investigating model performance

via train/cross-validation/test splits and model interrogation techniques. Impedance at these key frequencies may then be measured using cheaper electronics in real-world devices. Component values from dimensionality reduction algorithms such as PCA or UMAP also show good performance when used as features for predicting capacity, but these methods require the entire EIS spectrum and, thus, cannot be readily implemented using cheaper hardware. Methods for transforming the time-resolved DC response into an impedance spectrum exist,^{15,20} but these introduce additional complexity.

One key challenge identified in this work is accurate capacity prediction between 95%–85% relative discharge capacity; during this regimen, the resistance of most batteries in this dataset is very stable, resulting in low prediction accuracy. Combining this approach with other degradation modeling efforts, such as reduced-order degradation models,^{4,5} and creating an ensemble of various model types may help address systematic errors in any one individual approach for predicting battery health.

Overall, the outlook for using rapid EIS or time-resolved DC measurements to diagnose battery health seems promising, but more efforts are required to implement this approach in practice. The dataset provided here, which includes raw EIS data recorded at varying battery health and measurement conditions, can hopefully spur further efforts to use machine learning methods for analyzing battery data. More than 10,000 models were trained in this work to explore the effect of model architecture, feature selection, feature extraction, and hyperparameter optimization on model performance. Interesting work utilizing machine learning methodologies can include not only performance prediction but also be used to analyze trends in large datasets or for synthetic data generation, such as use of GANs to augment experimental measurements with representative synthetic data.

EXPERIMENTAL PROCEDURE

Resource availability

Lead contact

Inquiries regarding this work can be sent to P.G. (Paul.Gasper@nrel.gov, pauljgasper@gmail.com).

Materials availability

No new materials were created for this study. Tested batteries are not available.

Data and code availability

All original code and required data were deposited at www.github.com/NREL/battery_capacity_from_eis and are publicly available as of the date of publication. Some of the methods shown in this work have been replicated in Python using the dataset from Zhang et al.¹ and are available at https://github.com/battery-data-commons/mrs-sp22-tutorial/tree/main/predict_capacity_from_eis.

Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

Data acquisition and modeling

Details regarding data acquisition methods and modeling are provided in the [supplemental experimental procedures](#). Data acquisition methods include information on battery aging test procedures and battery performance measurement

procedures. Modeling methods include detailed descriptions of each pipeline and a description of “sklearn”-style machine learning modeling pipeline implementation in MATLAB.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.crp.2022.101184>.

ACKNOWLEDGMENTS

The National Renewable Energy Laboratory is operated by the Alliance for Sustainable Energy under contract DE-AC3608GO28308 for the U.S. Department of Energy. Support for this work was provided by DENSO Corporation under agreement FIA-19-01897. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

AUTHOR CONTRIBUTIONS

P.G., conceptualization, methodology, software, formal analysis, data curation, visualization, writing – original draft, writing – review & editing; A.S., software, data curation, writing – review & editing; K.S., conceptualization, supervision, project administration, writing – review & editing; Y.S., investigation, resources, writing – review & editing; S.Y., conceptualization, investigation, resources, project administration, writing – reviewing & editing.

DECLARATION OF INTERESTS

Y.S. and S.Y. are employees and shareholders of DENSO Corporation.

Received: June 2, 2022

Revised: September 15, 2022

Accepted: November 14, 2022

Published: December 14, 2022

SUPPORTING CITATIONS

The following references appear in the supplemental information: Mistry et al. (2021).⁶¹

REFERENCES

1. Zhang, Y., Tang, Q., Zhang, Y., Wang, J., Stimming, U., and Lee, A.A. (2020). Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning. *Nat. Commun.* 11, 1706. <https://doi.org/10.1038/s41467-020-15235-7>.
2. Pastor-Fernández, C., Yu, T.F., Widanage, W.D., and Marco, J. (2019). Critical review of non-invasive diagnosis techniques for quantification of degradation modes in lithium-ion batteries. *Renew. Sustain. Energy Rev.* 109, 138–159. <https://doi.org/10.1016/j.rser.2019.03.060>.
3. 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>.
4. Smith, K., Gasper, P., Colclasure, A.M., Shimonishi, Y., and Yoshida, S. (2021). Lithium-ion battery life model with electrode cracking and early-life break-in processes. *J. Electrochem. Soc.* 168, 100530. <https://doi.org/10.1149/1945-7111/ac2ebd>.
5. Smith, K., Saxon, A., Keyser, M., Lundstrom, B., Cao, Z., and Roc, A. (2017). Life prediction model for grid-connected Li-ion battery energy storage system. In 2017 American Control Conference (ACC) (IEEE), pp. 4062–4068. <https://doi.org/10.23919/ACC.2017.7963578>.
6. Richardson, R.R., Osborne, M.A., and Howey, D.A. (2017). Gaussian process regression for forecasting battery state of health. *J. Power Sources* 357, 209–219. <https://doi.org/10.1016/j.jpowsour.2017.05.004>.
7. Richardson, R.R., Birkel, C.R., Osborne, M.A., and Howey, D.A. (2019). Gaussian process regression for in situ capacity estimation of lithium-ion batteries. *IEEE Trans. Ind. Inf.* 15, 127–138. <https://doi.org/10.1109/TII.2018.2794997>.

8. Li, W., Sengupta, N., Dechent, P., Howey, D., Annaswamy, A., and Sauer, D.U. (2021). Online capacity estimation of lithium-ion batteries with deep long short-term memory networks. *J. Power Sources* 482, 228863. <https://doi.org/10.1016/j.jpowsour.2020.228863>.
9. Plett, G.L. (2004). Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs. *J. Power Sources* 134, 262–276. <https://doi.org/10.1016/j.jpowsour.2004.02.032>.
10. Zhang, D., Couto, L.D., Gill, P.S., Benjamin, S., Zeng, W., and Moura, S.J. (2021). Thermal-enhanced adaptive interval estimation in battery packs with heterogeneous cells. *IEEE Trans. Contr. Syst. Technol.* 1–14. <https://doi.org/10.1109/TCST.2021.3091108>.
11. Waag, W., Käbitz, S., and Sauer, D.U. (2013). Experimental investigation of the lithium-ion battery impedance characteristic at various conditions and aging states and its influence on the application. *Appl. Energy* 102, 885–897. <https://doi.org/10.1016/j.apenergy.2012.09.030>.
12. Barai, A., Uddin, K., Widanage, W.D., McGordon, A., and Jennings, P. (2018). A study of the influence of measurement timescale on internal resistance characterisation methodologies for lithium-ion cells. *Sci. Rep.* 8, 21. <https://doi.org/10.1038/s41598-017-18424-5>.
13. Locorotondo, E., Cultrera, V., Pugi, L., Berzi, L., Pierini, M., and Lutzemberger, G. (2021). Development of a battery real-time state of health diagnosis based on fast impedance measurements. *J. Energy Storage* 38, 102566. <https://doi.org/10.1016/j.est.2021.102566>.
14. Ran, A., Zhou, Z., Chen, S., Nie, P., Qian, K., Li, Z., Li, B., Sun, H., Kang, F., Zhang, X., et al. (2020). Data-driven fast clustering of second-life lithium-ion battery: mechanism and algorithm. *Adv. Theor. Simul.* 3, 2000109. <https://doi.org/10.1002/adts.202000109>.
15. Du, X., Meng, J., Peng, J., Zhang, Y., Liu, T., and Teodorescu, R. (2022). Sensorless temperature estimation of lithium-ion battery based on broadband impedance measurements. *IEEE Trans. Power Electron.* 1. <https://doi.org/10.1109/TPEL.2022.3166170>.
16. Pastor-Fernández, C., Uddin, K., Chouchelamane, G.H., Widanage, W.D., and Marco, J. (2017). A comparison between electrochemical impedance spectroscopy and incremental capacity-differential voltage as Li-ion diagnostic techniques to identify and quantify the effects of degradation modes within battery management systems. *J. Power Sources* 360, 301–318. <https://doi.org/10.1016/j.jpowsour.2017.03.042>.
17. Heins, T.P., Schlüter, N., Ernst, S.T., and Schröder, U. (2020). On the interpretation of impedance spectra of large-format lithium-ion batteries and its application in aging studies. *Energy Technol.* 8, 1900279. <https://doi.org/10.1002/ente.201900279>.
18. Mc Carthy, K., Gullapalli, H., Ryan, K.M., and Kennedy, T. (2022). Electrochemical impedance correlation analysis for the estimation of Li-ion battery state of charge, state of health and internal temperature. *J. Energy Storage* 50, 104608. <https://doi.org/10.1016/j.est.2022.104608>.
19. Gantenbein, S., Weiss, M., and Ivers-Tiffée, E. (2018). Impedance based time-domain modeling of lithium-ion batteries: Part I. *J. Power Sources* 379, 317–327. <https://doi.org/10.1016/j.jpowsour.2018.01.043>.
20. Weddle, P.J., Kee, R.J., and Vincent, T. (2018). A stitching algorithm to identify wide-bandwidth electrochemical impedance spectra for Li-ion batteries using binary perturbations. *J. Electrochem. Soc.* 165, A1679–A1684. <https://doi.org/10.1149/2.0641809jes>.
21. Zhang, L., Fan, W., Wang, Z., Li, W., and Sauer, D.U. (2020). Battery heating for lithium-ion batteries based on multi-stage alternative currents. *J. Energy Storage* 32, 101885. <https://doi.org/10.1016/j.est.2020.101885>.
22. Goldammer, E., Gentejohann, M., Schlüter, M., Weber, D., Wondrak, W., Dieckerhoff, S., Gühmann, C., and Kowal, J. (2022). The impact of an overlaid ripple current on battery aging: the development of the SiCWell dataset. *Batteries* 8, 11. <https://doi.org/10.3390/batteries8020011>.
23. Hertz, J. (2022). Active balancing innovation maximizes Li-ion battery potential. <https://www.batterytechnology.com/battery-management-systems/active-balancing-innovation-maximizes-li-ion-battery-potential>.
24. Zhu, X., Hallemans, N., Wouters, B., Claessens, R., Lataire, J., and Hubin, A. (2022). Operando odd random phase electrochemical impedance spectroscopy as a promising tool for monitoring lithium-ion batteries during fast charging. *J. Power Sources* 544, 231852. <https://doi.org/10.1016/j.jpowsour.2022.231852>.
25. Gong, Z., Liu, Z., Wang, Y., Gupta, K., da Silva, C., Liu, T., Zheng, Z.H., Zhang, W.P., van Lammeren, J.P.M., Bergveld, H.J., et al. (2018). IC for online EIS in automotive batteries and hybrid architecture for high-current perturbation in low-impedance cells. In 2018 IEEE Applied Power Electronics Conference and Exposition (APEC) (IEEE), pp. 1922–1929. <https://doi.org/10.1109/APEC.2018.8341280>.
26. Kersten, A., Kuder, M., Han, W., Thiringer, T., Lesnicar, A., Weyh, T., and Eckerle, R. (2020). Online and on-board battery impedance estimation of battery cells, modules or packs in a reconfigurable battery system or multilevel inverter. In IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society (IEEE), pp. 1884–1891. <https://doi.org/10.1109/IECON43393.2020.9254515>.
27. Lee, Y.-D., Park, S.-Y., and Han, S.-B. (2015). Online embedded impedance measurement using high-power battery charger. *IEEE Trans. Ind. Appl.* 51, 498–508. <https://doi.org/10.1109/TIA.2014.2336979>.
28. Kennedy, T., Gullapalli, H., Ryan, K.M., and Mc Carthy, K. (2021). Review—use of impedance spectroscopy for the estimation of Li-ion battery state of charge, state of health and internal temperature. *J. Electrochem. Soc.* 168, 080517. <https://doi.org/10.1149/1945-7111/ac1a85>.
29. Iurilli, P., Brivio, C., and Wood, V. (2021). On the use of electrochemical impedance spectroscopy to characterize and model the aging phenomena of lithium-ion batteries: a critical review. *J. Power Sources* 505, 229860. <https://doi.org/10.1016/j.jpowsour.2021.229860>.
30. Li, D., Wang, L., Duan, C., Li, Q., and Wang, K. (2022). Temperature prediction of lithium-ion batteries based on electrochemical impedance spectrum: a review. *Int. J. Energy Res.* 7905. <https://doi.org/10.1002/er.7905>.
31. Galeotti, M., Cinà, L., Giammanco, C., Cordiner, S., and Di Carlo, A. (2015). Performance analysis and SOH (state of health) evaluation of lithium polymer batteries through electrochemical impedance spectroscopy. *Energy* 89, 678–686. <https://doi.org/10.1016/j.energy.2015.05.148>.
32. Zhang, Q., Huang, C.-G., Li, H., Feng, G., and Peng, W. (2022). Electrochemical impedance spectroscopy based state of health estimation for lithium-ion battery considering temperature and state of charge effect. *IEEE Trans. Transport. Electrification*, 1. <https://doi.org/10.1109/TTE.2022.3160021>.
33. Schuster, S.F., Brand, M.J., Campestrini, C., Gleissenberger, M., and Jossen, A. (2016). Correlation between capacity and impedance of lithium-ion cells during calendar and cycle life. *J. Power Sources* 305, 191–199. <https://doi.org/10.1016/j.jpowsour.2015.11.096>.
34. La Rue, A., Weddle, P.J., Ma, M., Hendricks, C., Kee, R.J., and Vincent, T.L. (2019). State-of-Charge estimation of LiFePO₄-Li₄Ti₅O₁₂ batteries using history-dependent complex-impedance. *J. Electrochem. Soc.* 166, A4041–A4046. <https://doi.org/10.1149/2.0221916jes>.
35. Love, C.T., Virji, M.B.V., Rocheleau, R.E., and Swider-Lyons, K.E. (2014). State-of-health monitoring of 18650 4S packs with a single-point impedance diagnostic. *J. Power Sources* 266, 512–519. <https://doi.org/10.1016/j.jpowsour.2014.05.033>.
36. Liebhart, B., Diehl, S., and Endisch, C. (2020). Sensitivity analysis of battery cell aging estimators based on impedance spectroscopy regarding temperature compensation. In 2020 IEEE Conference on Control Technology and Applications (CCTA) (IEEE), pp. 801–806. <https://doi.org/10.1109/CCTA41146.2020.9206395>.
37. Mc Carthy, K., Gullapalli, H., and Kennedy, T. (2022). Real-time internal temperature estimation of commercial Li-ion batteries using online impedance measurements. *J. Power Sources* 519, 230786. <https://doi.org/10.1016/j.jpowsour.2021.230786>.
38. Wang, X., Wei, X., Dai, H., and Wu, Q. (2015). State estimation of lithium ion battery based on electrochemical impedance spectroscopy with on-board impedance measurement system. In 2015 IEEE Vehicle Power and Propulsion Conference (VPPC) (IEEE), pp. 1–5. <https://doi.org/10.1109/VPPC.2015.7353021>.
39. Caposciutti, G., Bandini, G., Marracci, M., Buffi, A., and Tellini, B. (2021). Li-ion batteries state of health analysis via electro-chemical impedance spectroscopy. In 2021 IEEE International Workshop on Metrology for Automotive (MetroAutomotive) (IEEE), pp. 36–41. <https://doi.org/10.1109/MetroAutomotive50197.2021.9502724>.

40. Beelen, H.P.G.J., Raijmakers, L.H.J., Donkers, M.C.F., Notten, P.H.L., and Bergveld, H.J. (2016). A comparison and accuracy analysis of impedance-based temperature estimation methods for Li-ion batteries. *Appl. Energy* 175, 128–140. <https://doi.org/10.1016/j.apenergy.2016.04.103>.
41. La Rue, A., Weddle, P.J., Kee, R.J., and Vincent, T.L. (2020). Feature selection for state-of-charge estimation of LiFePO₄-Li₄Ti₅O₁₂ batteries via electrochemical impedance. In 2020 American Control Conference (ACC) (IEEE), pp. 231–236. <https://doi.org/10.23919/ACC45564.2020.9147727>.
42. Kukreja, S.L., Löfberg, J., and Brenner, M.J. (2006). A least absolute shrinkage and selection operator (LASSO) for nonlinear system identification. *IFAC Proc. Vol. 39*, 814–819. <https://doi.org/10.3182/20060329-3-AU-2901.00128>.
43. Kim, S., Choi, Y.Y., and Choi, J.-I. (2022). Impedance-based capacity estimation for lithium-ion batteries using generative adversarial network. *Appl. Energy* 308, 118317. <https://doi.org/10.1016/j.apenergy.2021.118317>.
44. Rastegarpanah, A., Hathaway, J., and Stolkin, R. (2021). Rapid model-free state of health estimation for end-of-first-life electric vehicle batteries using impedance spectroscopy. *Energies* 14, 2597. <https://doi.org/10.3390/en14092597>.
45. Messing, M., Shoa, T., Ahmed, R., and Habibi, S. (2020). Battery SoC estimation from EIS using neural nets. In 2020 IEEE Transportation Electrification Conference & Expo (ITEC) (IEEE), pp. 588–593. <https://doi.org/10.1109/ITEC48692.2020.9161523>.
46. Mussa, A.S., Liivat, A., Marzano, F., Klett, M., Philippe, B., Tengstedt, C., Lindbergh, G., Edström, K., Lindström, R.W., and Svens, P. (2019). Fast-charging effects on ageing for energy-optimized automotive LiNi₁/3Mn₁/3Co₁/3O₂/graphite prismatic lithium-ion cells. *J. Power Sources* 422, 175–184. <https://doi.org/10.1016/j.jpowsour.2019.02.095>.
47. Naumann, M., Spingler, F.B., and Jossen, A. (2020). Analysis and modeling of cycle aging of a commercial LiFePO₄/graphite cell. *J. Power Sources* 451, 227666. <https://doi.org/10.1016/j.jpowsour.2019.227666>.
48. McInnes, L., Healy, J., and Melville, J. (2020). UMAP: uniform Manifold approximation and projection for dimension reduction. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1802.03426>.
49. Abdi, H., and Williams, L.J. (2010). Principal component analysis. *WIREs Comput. Stat.* 2, 433–459. <https://doi.org/10.1002/wics.101>.
50. van der Maaten, L., and Hinton, G. (2008). Visualizing Data using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2605.
51. 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>.
52. Howard, J., and Gugger, S. (2020). Fastai: a layered API for deep learning. *Information* 11, 108. <https://doi.org/10.3390/info11020108>.
53. Greenbank, S., and Howey, D. (2021). Automated feature extraction and selection for data-driven models of rapid battery capacity fade and end of life. *IEEE Trans. Ind. Inf.* 18, 2965–2973. <https://doi.org/10.1109/TII.2021.3106593>.
54. Ouyang, R., Curtarolo, S., Ahmetcik, E., Scheffler, M., and Ghiringhelli, L.M. (2018). SISSO: a compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates. *Phys. Rev. Mater.* 2, 083802. <https://doi.org/10.1103/PhysRevMaterials.2.083802>.
55. Paulson, N.H., Kubal, J., Ward, L., Saxena, S., Lu, W., and Babinec, S.J. (2022). Feature engineering for machine learning enabled early prediction of battery lifetime. *J. Power Sources* 527, 231127. <https://doi.org/10.1016/j.jpowsour.2022.231127>.
56. Sulzer, V., Mohtat, P., Lee, S., Siegel, J.B., and Stefanopoulou, A.G. (2021). Promise and challenges of a data-driven approach for battery lifetime prognostics. In 2021 American Control Conference (ACC) (IEEE), pp. 4427–4433. <https://doi.org/10.23919/ACC50511.2021.9483312>.
57. Grinsztajn, L., Oyallon, E., and Varoquaux, G. (2022). Why Do Tree-Based Models Still Outperform Deep Learning on Tabular Data? arXiv. <https://doi.org/10.48550/arXiv.2207.08815>.
58. Sendek, A.D., Ransom, B., Cubuk, E.D., Pellouchoud, L.A., Nanda, J., and Reed, E.J. (2022). Machine learning modeling for accelerated battery materials design in the small data regime. *Adv. Energy Mater.* 12, 2200553. <https://doi.org/10.1002/aenm.202200553>.
59. Fernandez, A., Garcia, S., Herrera, F., and Chawla, N.V. (2018). SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. *J. Artif. Intell. Res.* 61, 863–905. <https://doi.org/10.1613/jair.1.11192>.
60. Jones, P.K., Stimming, U., and Lee, A.A. (2022). Impedance-based forecasting of lithium-ion battery performance amid uneven usage. *Nat. Commun.* 13, 4806. <https://doi.org/10.1038/s41467-022-32422-w>.
61. Mistry, A., Verma, A., Sripad, S., Ciez, R., Sulzer, V., Brosa Planella, F., Timms, R., Zhang, Y., Kurchin, R., Dechent, P., et al. (2021). A minimal information set to enable verifiable theoretical battery research. *ACS Energy Lett.* 6, 3831–3835. <https://doi.org/10.1021/acscenergylett.1c01710>.