

Article

Battery Stress Factor Ranking for Accelerated Degradation Test Planning Using Machine Learning

Saurabh Saxena ^{1,*}, Darius Roman ², Valentin Robu ^{2,3}, David Flynn ² and Michael Pecht ¹

¹ Center for Advanced Life Cycle Engineering, University of Maryland, College Park, MD 20742, USA; pecht@umd.edu

² Smart Systems Group, School of Engineering & Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK; dvr1@hw.ac.uk (D.R.); V.Robu@hw.ac.uk (V.R.); D.Flynn@hw.ac.uk (D.F.)

³ Centrum Wiskunde & Informatica, Science Park 123, 1098 XG Amsterdam, The Netherlands

* Correspondence: saxenas@terpmail.umd.edu

† Current address: Argonne National Laboratory, 9700 S Cass Ave, Lemont, IL 60439, USA.

Abstract: Lithium-ion batteries power numerous systems from consumer electronics to electric vehicles, and thus undergo qualification testing for degradation assessment prior to deployment. Qualification testing involves repeated charge–discharge operation of the batteries, which can take more than three months if subjected to 500 cycles at a C-rate of 0.5C. Accelerated degradation testing can be used to reduce extensive test time, but its application requires a careful selection of stress factors. To address this challenge, this study identifies and ranks stress factors in terms of their effects on battery degradation (capacity fade) using half-fractional design of experiments and machine learning. Two case studies are presented involving 96 lithium-ion batteries from two different manufacturers, tested under five different stress factors. Results show that neither the individual (main) effects nor the two-way interaction effects of charge C-rate and depth of discharge rank in the top three significant stress factors for the capacity fade in lithium-ion batteries, while temperature in the form of either individual or interaction effect provides the maximum acceleration.

Keywords: lithium-ion batteries; cycle life; temperature; C-rate; accelerated testing; machine learning



Citation: Saxena, S.; Roman, D.; Robu, V.; Flynn, D.; Pecht, M. Battery Stress Factor Ranking for Accelerated Degradation Test Planning Using Machine Learning. *Energies* **2021**, *14*, 723. <https://doi.org/10.3390/en14030723>

Received: 11 January 2021

Accepted: 28 January 2021

Published: 30 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Lithium-ion batteries are used as energy storage devices in a variety of applications, including portable electronics, grid storage, electric vehicles, and marine energy systems [1,2]. The market for Li-ion batteries is expected to register a compound annual growth rate (CAGR) of approximately 22% during the forecast period (2019–2024) [3]. The increasing demand is also enabled by the continuous decline in battery prices, with the volume weighted average battery pack cost falling by 85% from 2010 to 2018, reaching an average of \$176/kWh [4]. Batteries are critical subsystems that provide primary and/or standby energy within products and systems, and their performance degrades during their lifetime due to various degradation mechanisms resulting in capacity and power fade [5,6].

Cycle testing of Li-ion batteries is conducted to qualify a battery population according to the capacity and power requirements for its targeted application. However, testing under normal operating conditions can be prohibitively time-consuming. For example, 500 cycles at discharge and charge C-rates of 0.5C can take approximately 95 days (more than three months) to conclude. If the testing is conducted to reach a predefined failure threshold or end of life (i.e., a drop to 80% capacity in many cases), then the testing time can exceed even four months because state-of-the-art commercial Li-ion batteries can maintain more than 80% of capacity for more than 500 cycles. A survey [7] collating feedback from professionals from a broad spectrum of industrial sectors, including battery cell producers and battery pack and component developers, revealed time-to-market and reliability as their primary concerns. Due to the costs and associated delays in accessing the market, the battery

industry is increasingly interested in utilizing accelerated testing procedures. Accelerated test methods are also desirable for qualification of second-life Li-ion batteries where the supply is expected to surpass 200 gigawatt-hours per year by 2030 and applications require less-frequent battery cycling [8].

Accelerated testing is conducted by elevating certain stress factors to increase the degradation rate and precipitate failure earlier than under normal operation. The cycling operation of batteries under a qualification process is usually characterized by five primary stress factors: ambient temperature, discharge C-rate, charge C-rate, constant voltage charge cut-off C-rate, and depth of discharge (DOD) [9–11]. Some or all these stress factors can be selected to accelerate the battery degradation process. The effects of these stress factors on battery degradation are not uniform, however, and some factors can have a disproportionate impact on degradation. Additionally, multifactorial and interdependent stress factors may work in combination (interaction effects) to accelerate certain failure modes. It is not usually feasible to design a test matrix with many factors due to limited testing resources and the time needed to conduct such tests. Therefore, to achieve an optimal time reduction with the fewest stress factors, a key challenge is to identify the most impactful stress factors and utilize them for accelerated testing and modeling.

Within the literature, there is a significant database of experimental studies that have investigated the effects of individual stress factors on Li-ion battery degradation during cycling operation. However, there is limited research on the battery cycle life degradation under the influence of multiple (three or more) stress factors simultaneously. It is important to understand the individual and coupled effects of these stressors on battery degradation. The following studies have provided some initial analysis on the influence of multiple and interacting stressors: Diao et al. [12] conducted a full factorial design of experiment (DOE) to study the effects of temperature, discharge C-rate, and charge current cut-off on Li-ion batteries and found that only temperature influenced the capacity fade rate. Wang et al. [13] studied the cycle life of Li-ion batteries under three stress factors, including DOD, temperature, and discharge rates. They found that the capacity loss was strongly affected by time and temperature, whereas the effect of DOD was less important at a 0.5C discharge rate. Schimpe et al. [14] studied the cycle aging of commercial Li-ion batteries under temperature and charge and discharge C-rates stress factors and with and without constant voltage charging step. They showed that current rates did not show any effects on capacity loss at high temperature (55 °C). However, these studies lacked analysis to separate the individual and interaction effects of stress factors and did not consider all the five stress factors that characterize the cycling operation of batteries.

Cui et al. [15] conducted orthogonal experiments to determine the key stress factor for capacity loss in commercial Li-ion batteries cycled at shallow (20% to 30%) DOD. The stress factors in their study included temperature, discharge C-rate, end-of-charge voltage, and DOD. They identified temperature—followed by discharge rate and DOD—as the most impactful stress factor for battery capacity loss. The orthogonal design did not include the end-of-charge voltage stress factor, the effects of which were investigated separately. They did not investigate the interaction effects of stress factors.

Prochazka et al. [16] used a D-optimization method to reduce the full factorial design of four factors, including current, temperature, state of charge (SOC), and the change of SOC with five levels each ($4^5 = 1024$ tests) to 46 tests. They then implemented a linear regression model with individual, quadratic, and interaction effects of stress factors to conclude that three out of four factors were significant for the batteries they tested. These two studies [15,16] chose four stress factors, which alone cannot constitute the full cycle profile of batteries, leading to an incomplete ranking of individual and/or interaction effects of all the stress factors on battery capacity fade.

Su et al. [17] used seven principal factors—including the charge/discharge currents during the constant current regime, the charge/discharge cut-off voltages and the corresponding durations during the constant voltage regime, and the ambient temperature in a reduced orthogonal DOE (18 tests)—with three different levels for each factor. They

then conducted analysis of variance (ANOVA) to identify significant stress factors and their ranking. However, they chose stress factors that were not practical to define the conventional constant current/constant voltage (CCCV) charge (limited by charge cut-off current and not time) and constant current discharge cycling profile used for the battery qualification purpose. They also chose more than two levels of stress factors for screening purposes, which unnecessarily increased the number of tests for a constant number of desired effects and then did not consider the interaction effects of the stress factors. They recommended the determination of interaction effects of stress factors on battery degradation for future research.

This paper presents a study to address the limitations of current state of the art and provide guidance in the accelerated battery test planning. This study combines design of experiment for data generation and two machine learning techniques, least absolute shrinkage and selection operator (LASSO) and random forests (RF) for the stress factor effects analysis and ranking. The data generation work included the implementation of a multi-stress half-fractional DOE consisting of two levels of stress factors to reduce the number of tests to train the models. To the best of our knowledge, this is the first multifactorial experimental study that investigates the significance and ranking of individual (main) and two-way interaction effects of all five stress factors (temperature, charge current, discharge current, constant voltage charge cut-off current, and DOD) required to constitute a typical cycling operation used in the battery qualification testing.

The remainder of the paper is structured as follows: Section 2 describes the half-fractional factorial DOE and experimental procedure. Section 3 discusses the machine learning techniques used in this paper. Section 4 presents the two case studies and discusses the analysis of the effects of stress factors. Section 5 presents the conclusions.

2. Design of Experiment

A DOE was used to conduct the cycle testing of batteries with five different stress factors: ambient temperature, discharge C-rate, charge C-rate, constant voltage charge cut-off current, and DOD. Two levels for each selected stress factor were chosen to cover the entire range of the stress factor. The purpose of this study was to screen the stress factors and rank them in terms of their importance for battery capacity fade acceleration; therefore, two levels were sufficient for each stress factor. The DOE presented in this section was not intended for the accelerated testing of batteries, but rather served as a preliminary step to determine impactful stress factors that could be used for planning the accelerated testing of batteries. The DOE for the testing was as follows:

Factors : $k = 5$, $X_1 = T$, $X_2 = \text{Discharge C-rate}$, $X_3 = \text{Charge Cut-off}$, $X_4 = \text{Charge C-rate}$, $X_5 = \text{DOD} (= \frac{I \cdot t}{Q_N})$

where k is the number of stress factors, I is the constant discharge current, t is the discharge time, Q_N is battery capacity after N cycles, and T is the ambient temperature.

Levels : $L_1 = 2$, $L_2 = 2$, $L_3 = 2$, $L_4 = 2$, $L_5 = 2$

A full factorial DOE is the most comprehensive as it considers all the individual and higher-order interaction effects. A full factorial design (2^k) for five stress factors will result in $N = L_1 \cdot L_2 \cdot L_3 \cdot L_4 \cdot L_5 = 32$ test cases. Third- or higher-order interaction effects of stress factors are usually insignificant in most practical cases and thus can be confounded with individual and two-way interaction effects. A half-fractional factorial design can preserve individual and two-way interaction effects of stress factors while reducing the number of test cases, which can be later extracted using statistical analysis and machine learning. In half-fractional design, one of the stress factors is represented as a logical multiplication of the remaining four or a lower number of stress factors. Table 1 shows the process of representing DOD (X_5) as a function of the remaining four stress factors in the proposed design. The design in Table 1 is resolution V design, where the individual effects are aliased with at least fourth-order effects and the second-order (two-way) interaction effects

are confounded with third-order effects. To understand the repeatability of results, three samples per test case have been considered.

Table 1. Half-fractional factorial test matrix.

Experiment No.	Temperature (°C) (X1)	Discharge C-Rate (X2)	Charge Cut-Off (X3)	Charge C-Rate (X4)	DOD (X5 = X1·X2·X3·X4)
1	25 (L)	0.5C (L) *	0.01C (H) *	0.8C (L)	0.5 (L)
2	25 (L)	0.5C (L)	0.01C (H)	1.2C (H)	1 (H)
3	25	0.5C	0.2C (L)	0.8C	1
4	25	0.5C	0.2C	1.2C	0.5
5	25	2C/1.3C ^ (H)	0.01C	0.8C	1
6	25	2C/1.3C	0.01C	1.2C	0.5
7	25	2C/1.3C	0.2C	0.8C	0.5
8	25	2C/1.3C	0.2C	1.2C	1
9	55 (H)	0.5C	0.01C	0.8C	1
10	55	0.5C	0.01C	1.2C	0.5
11	55	0.5C	0.2C	0.8C	0.5
12	55	0.5C	0.2C	1.2C	1
13	55	2C/1.3C	0.01C	0.8C	0.5
14	55	2C/1.3C	0.01C	1.2C	1
15	55	2C/1.3C	0.2C	0.8C	1
16	55	2C/1.3C	0.2C	1.2C	0.5

* H—High (+1), L—Low (−1). ^ Based on the manufacturers' datasheets, two different high levels of discharge C-rates have been used in the two case studies.

The ranges of stress factors in the DOE were selected with a focus on portable electronic application and electrode-electrolyte interfacial side-reaction degradation mechanisms (such as solid electrolyte interphase [SEI] layer). Battery capacity loss due to low-temperature-based transport limitations and lithium plating was not considered as part of this study. Factor ranges were chosen carefully to prevent/minimize abusive or unsafe conditions such as thermal decomposition of the SEI layer or electrolyte. Based on the literature, these mechanisms do not occur until 100 °C [18]. The battery surface temperature in the testing never reached those levels.

The first factor in the DOE was ambient temperature with a range of 25 °C to 55 °C. In the range, 25 °C represents a typical room-temperature condition and 55 °C (which is lower than the battery datasheet specification of 60 °C) was considered as the maximum (accelerated level) limit. This was based on battery manufacturers' suggestions to prevent any abusive degradation mechanisms, considering that battery temperature can further rise due to the charge-discharge operations. The second stress factor was the battery discharge C-rate. The range 0C to 0.5C represents a typical normal range of continuous discharge rate for portable electronic applications such as smartphones and laptops. Discharge rates up to 2C are also expected, but only for short durations; hence, a discharge C-rate range of 0.5C to 2C was considered. Many battery manufacturers usually prescribe 2C as the maximum continuous discharge current.

The third stress factor was charge cut-off current during the constant voltage charging step. Reducing the charge cut-off current increases the constant voltage charging time but ensures that the open circuit voltage of the battery reaches closer to the end-of-charge voltage. Although time-consuming, C/100 (0.01C) is the lowest practical charge cut-off current that is used sometimes in battery testing. C/20 (0.05C), C/10 (0.1C), and C/5 (0.2C) are other examples of charge cut-off current levels that are typically used to charge a battery almost fully. Thus, two levels of C/5 and C/100 were used for charge cut-off. The charge cut-off 0.01C represents a high level (in Table 1) because it corresponds to "more" charging of the battery. Charge C-rate was the fourth stress factor in the design space. Charge C-rates close to 0.8C are used in portable electronic devices and have been prescribed in battery datasheets. Thus, 0C to 0.8C can be considered as the normal operating range. The charge C-rate stress factor at increased levels can accelerate unsafe mechanisms, such as lithium plating and dendrites. Thus, the maximum limit of charge C-rate was fixed to 1.2C

based on suggestions from the battery manufacturers. DOD, the fifth factor, represents the amount of battery capacity being utilized in each cycle. DOD at 100% represents the full cycle condition and the maximum limit. Considering that the entire 0–100% range represents the normal operation and the factor effects are independent of the factor range (if there are no new mechanisms), 50% was selected as a lower limit of DOD in the testing.

All the test cases described in Table 1 involved continuous cycling of a Li-ion battery under the specified stress factors to characterize the capacity fade trend. The cycling profile included charge/discharge operations. The charging operation was conducted using a CCCV standard charge algorithm. The battery was charged using the prescribed constant charge C-rate up to the end-of-charge voltage (4.4 V) followed by the “top-up” using constant voltage charging until the charging current dropped below the prescribed charge cut-off. The discharge operation was performed using the prescribed constant discharge C-rate until the prescribed DOD was achieved. DOD of 100% was defined as the discharge to 3 V at 0.5C discharge current rate. For high discharge current tests (2C/1.3C), therefore, the batteries were further discharged to 3 V at 0.5C after hitting the 3 V threshold at the prescribed high level (2C/1.3C) of discharge C-rate. Rest times after discharge and after charge were not considered as stress factors in this study and were kept fixed at 10 min for the testing.

The cell characterization testing was conducted at the beginning of cycling testing (Table 1) to set up the baseline and intermediately between the cycling tests for comparison with baseline characteristics. The characterization testing included the measurements of “true” battery discharge capacity. This true value of discharge capacity was defined at a standard condition for comparison across different tests in Table 1. Discharge capacity measurement tests were conducted by cycling the battery at 25 °C ambient temperature using a standard full charge/discharge cycle. The battery was charged at 0.8C constant current up to the end-of-charge voltage (4.4 V) followed by the constant voltage charging until the charging current dropped below $C/20$. Following the 10 min of rest after the charging, the battery was discharged at $C/2$ constant current up to the end-of-discharge voltage (3 V). After discharge, a 10-min rest period was provided before charging the battery for the next cycle. For the initial capacity measurements of fresh cells, five cycles were conducted and the average value of the discharge capacities was used as the initial capacity. For the capacity measurements at prescribed intervals and at the end of the cycling testing, two cycles were conducted and the discharge capacity from the second cycle was used as the real capacity measurement. The intermediate characterization of discharge capacity was conducted at fixed intervals of 100 cycles. These discharge capacities measured at every 100 cycles were used for adjusting the discharge time required to achieve 50% DOD at a prescribed discharge C-rate in Table 1. The cycle testing based on the DOE in Table 1 was continued until 300 cycles (approximately 2 months) or a 20% drop in discharge capacity measurements, whichever occurred earlier.

3. Algorithms for Stress Factor Ranking

Two machine learning algorithms, LASSO regression and RF, were considered for evaluating the effects of the stress factors on battery capacity fade and for ranking them in terms of their impact based on the experimental data in this study. It is important to note that these algorithms have been used in this study only for stress factor (effects) ranking before accelerated testing and not for building predictive capacity fade and life models [19,20]. The algorithms, therefore, were not applied to the other (unseen) test datasets for prediction purposes, instead the algorithms used all of the experimental data generated as part of this study.

LASSO is specifically designed to model linear relationships with the added benefit of pruning the model weights to better fit the data. The pruning function, also referred to as regularization, also acts as a predictor variable (feature) ranking property [21]. While linear regression can lead to overfitting in small datasets, LASSO is more suitable due to its regularization property. RF, on the other hand, is a non-parametric, non-linear ensemble

model consisting of an ensemble of decision trees [22]. Given its decision tree ensemble construction, RF, is able to measure relative importance between predictor variables by looking at how much the tree node of a particular predictor variable reduces the loss function, typically measured as impurity [22].

The task of identifying the highest contributing predictor variables for prediction is termed “predictor variable (feature) selection” and is generally applied in datasets where the number of predictor variables exceeds the order of hundreds. The primary goal of predictor variable selection is a reduction in computational time and memory. However, the present work made use of predictor variable selection for stress factor ranking in order of importance to capacity fade. Traditionally, supervised predictor variable selection is grouped into three broad categories: filter, wrapper, and embedded methods [23–25]. The focus here was on embedded methods, in particular LASSO and RF due to (1) superior performance over filter methods, (2) better interpretability when compared to wrapper methods, and (3) use of the predictor performance as the objective function to evaluate the variable subset. The problem can be specified mathematically as follows:

$$Y = f(X) + \varepsilon \quad (1)$$

$$X = (X1, X2, \dots, X5, X1.X2, X1.X3, \dots, X4.X5) \quad (2)$$

where Y is the response variable; X is the matrix of predictor variables; and ε is the random error term, which is independent of X and has zero mean. The purpose of the algorithms is to learn the unknown function f and understand how Y is affected by X . In this study, predictor variables (features) refer to individual stress factors and their two-way interactions.

The average capacity fade rate was used as the response variable Y in Equation (1). It was calculated by dividing the total drop in capacity (ΔAh) with the cumulative discharge (Ah), which is defined as the total amount of charge (Ah) delivered by the battery during the discharge operations of the entire testing. Cumulative discharge was selected in the denominator of Equation (3) in place of number of cycles because the term “cycle” is not well defined for the tests in Table 1. For example, a cycle with 50% DOD cannot be directly compared with a cycle with 100% DOD. Cumulative discharge also provides a good sense of testing (discharge) time for a given C-rate.

$$Y = \frac{\Delta \text{Capacity (Ah)}}{\text{Cumulative Discharge (Ah)}} \quad (3)$$

The estimation of f requires training data generated by the cycle testing of batteries as per the DOE described in Table 1. There are two types of machine learning approaches: parametric and non-parametric. In this paper, both of these approaches have been demonstrated. The LASSO regression falls into the parametric category, where an assumption about the functional form of f is made in advance and then training data are used to estimate the coefficients. Non-parametric approaches like RF do not make an explicit assumption about the functional form of f .

$$f(X) = \beta_0 + \beta_1 X1 + \beta_2 X2 + \dots + \beta_5 X5 + \beta_{12} X1.X2 + \dots + \beta_{45} X4.X5 + \varepsilon \quad (4)$$

A linear model for five stress factors can be described by Equation (4), where $\beta_0, \beta_1, \dots, \beta_5, \beta_{12}, \dots, \beta_{45}$ are the regression coefficients (effect sizes) that need to be estimated. The most conventional algorithm to estimate these coefficients is least squares, which tries to minimize the residual sum of squares (RSS), defined as follows:

$$RSS = \sum_{i=1}^n (y_i - \widehat{f(x_i)})^2 = \sum_{i=1}^n (y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_{i1} - \widehat{\beta}_2 x_{i2} - \dots - \widehat{\beta}_5 x_{i5} - \widehat{\beta}_{12} x_{i1}.x_{i2} - \dots - \widehat{\beta}_{45} x_{i4}.x_{i5})^2 \quad (5)$$

where \widehat{f} is an estimate of function f ; $\widehat{\beta}_0, \widehat{\beta}_1, \dots, \widehat{\beta}_5, \widehat{\beta}_{12}, \dots, \widehat{\beta}_{45}$ are estimates of $\beta_0, \beta_1, \dots, \beta_5, \beta_{12}, \dots, \beta_{45}$ respectively; and n is the number of training observations. LASSO regression, a regularization technique developed for linear regression problems, has been widely used in the literature as a predictor variable selection algorithm due to its simplicity and ease of interpretation [21,26]. The key assumption is that the best possible prediction rule is sparse; that is, only a few of the coefficients are different from zero. Coefficients different from zero are directly proportional to the variable's importance. The higher the coefficient at a nonzero weight, the more important the variable for the prediction. In LASSO regression, sparsity is achieved by adding a term that penalizes the loss based on convex relaxation via the l_1 norm and a regularization term λ . The loss function formula thus becomes:

$$\min \sum_{i=1}^n (y_i - \widehat{f}(x_i))^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (6)$$

Equation (6) represents the summation of RSS (Equation (5)) and a l_1 norm penalty term. λ represents the regularization term. The higher the regularization term, the more the weights of the model shrink to zero, thus revealing the important predictor variables. The value of λ was chosen based on the 10-fold cross-validation approach. In this approach, the training data are randomly divided into 10 parts or folds, and the first fold is treated as the validation set while the remaining nine folds are used for training purposes. This process is repeated 10 times to choose each one of those 10 folds as the validation set, and the mean squared error is calculated for each validation set. The final mean squared error (MSE) is calculated as follows:

$$MSE = \frac{1}{10} \sum_{i=1}^{10} MSE_i \quad (7)$$

A value of λ that minimizes the MSE can be chosen. However, to use a slightly larger value of λ , which allows a stricter penalty on coefficients, this study used a value of λ corresponding to the MSE, which was higher by one standard deviation from the minimum MSE. A non-parametric bootstrap method [27] was further used to calculate the probability of each coefficient in the regression model being zero using a 1000-sample size. In the bootstrap method, a training sample of size n —the same as that of the actual training data—was randomly sampled with replacement. For this sample, the 10-fold cross-validation was conducted to select λ corresponding to MSE at one standard deviation and calculate corresponding estimates of coefficients. This process was repeated 1000 times, and then the probability of zero was calculated for each coefficient.

$$\text{Bootstrap Probability of } (\beta_j = 0) = \frac{1}{1000} (\text{number of times } \beta_j = 0) \quad (8)$$

The non-parametric machine learning method, RF, has also been implemented for variable ranking as a comparison to the bootstrapped LASSO method. Randomized-based ensemble techniques such as RF have only recently been used as a predictor variable selection method [28]. In ensemble learning such as RF, a collection of single regression models (decision trees) are trained, and the output of the ensemble regression is obtained by aggregating or averaging the outputs of the single models. In other words, the generated RF is nothing but a collection of constructed decision trees, sequentially conducting binary splits of the data during training in order to produce a homogeneous subset [22].

RF can handle interactions between variables, distinguishing relevant from irrelevant variables even when the number of variables is much larger than the number of samples [22,29]. The RF algorithm requires substantial hyperparameter tuning; therefore, instead of the common grid search, a random search approach was adopted. Exponential distributions were considered for each of the following hyperparameters: maximum number of predictor variables considered for splitting the node, maximum tree depth, minimum

number of samples placed in a node before splitting, and minimum number of samples in a leaf node. Empirically and theoretically randomly chosen trials, no matter the distribution from which they are chosen, are more efficient for hyperparameter optimization than trials on a grid [30], hence the adoption of the technique.

Due to the randomization effect inherent to RF, the masking effect that plagues the decision tree algorithm—by which one predictor variable might never occur in any split because it leads to splits that are slightly worse—is greatly decreased, especially when the forest comprises decision trees on the order of hundreds or even thousands. The number of decision trees is therefore an important parameter that in some cases requires tuning. For instance, if computational power is not a concern, the bigger the number of decision trees the better the prediction. In the present work, the number of trees was restricted to 70 to prevent overfitting. Variable importance was calculated as the decrease in node impurity weighted by the probability of reaching that node, where the probability was simply the number of samples that reach the node, divided by the total number of samples. In mathematical terms, variable importance values over all trees in the forest φ_m can be calculated by:

$$Imp(x_p) = \frac{1}{M} \sum_{m=1}^M \sum_{t \in \varphi_m} p(t) \Delta i(s_t, t) \quad (9)$$

where $p(t) \Delta i(s_t, t)$ is the impurity increase, typically measured via mean decrease accuracy, mean increase error, or variance for regression; p is the number of predictor variables; t is the node number; m is the number of decision trees in the forest; and $p(t)$ is the proportion of samples reaching node t over the total number of samples [31,32]. The present work adopts mean decrease accuracy for predictor variable selection.

4. Results and Discussion

This section presents the two case studies where the cycle testing results of the graphite/LiCoO₂ batteries from two manufacturers (a total of 96 cells) are discussed and the algorithms from Section 3 are applied to the test data for ranking the stress factors for battery capacity fade.

4.1. Case Study 1

This case study used the Li-ion batteries from Manufacturer 1 with a minimum capacity of 2720 mAh. These batteries were tested as per Table 1. The discharge capacity measurements from the testing are presented in Figure 1 to show the degradation trend of the batteries. The Y-axis in the figure represents the discharge capacity (Ah). The X-axis shows the cumulative discharge of the batteries in Ah. The legends have been shortened to keep the figures clear. A legend of “T1S1” denotes Test 1 and Sample 1.

Figure 2 shows the average capacity fade rate for all the test cases, which would serve as the response variable Y (Equation (1)). The batteries show the highest capacity fade rate for 55 °C and 2C test conditions (Tests 13–16). These batteries reached below 80% of their initial capacities, much below the 100 Ah, suggesting an accelerated degradation. On the other hand, the 25 °C and 0.5C test cases (Tests 1–4) show the least capacity fade rate, with batteries not even reaching 90% of initial capacities after 500 Ah of cumulative discharge. There is some overlap among test cases 5–12 in terms of their degradation trends and average capacity fade rate.

The rate capability of batteries from Manufacturer 1 were limited; hence, at 2C discharge current the batteries could barely provide 50% of the rated capacity. This issue was exacerbated by aging, and during the testing it was difficult to control the DOD factor for these batteries. Among the five stress factors used in the testing, DOD is the most difficult to control due to its dependence on the discharge current, rate capability, and even temperature. Due to the continuously varying nature of DOD for the batteries under DOD = 0.5 and discharge C-rate = 2C test conditions, it is difficult to statistically analyze the effects of this stress factor on capacity fade. For DOD = 1 and discharge C-rate = 2C test conditions, a DOD of 1 was maintained by further discharging to 3 V at 0.5C after

hitting the 3 V threshold at the 2C level of discharge C-rate. For Manufacturer 1, only cases corresponding to DOD = 1 (eight test conditions) have been considered, leaving only four stress factors and six of their two-way interactions.

$$f(X) = \beta_0 + \beta_1 X_1 + \dots + \beta_4 X_4 + \beta_{12} X_1 \cdot X_2 + \dots + \beta_{34} X_3 \cdot X_4 + \varepsilon \quad (10)$$

Ignoring the coefficients related to DOD (Table 2), 11 coefficients (Equation (10)) need to be estimated from eight test conditions, leading to infinite solutions. LASSO can deal with the condition where the number of predictor variables (11) is higher than the number of test conditions (8), as it shrinks some of the coefficient estimates to zero. As described in Section 4, a 10-fold cross-validation has been used to choose the value of λ corresponding to the MSE value one standard deviation away from minimum MSE and estimate the coefficients. A non-parametric bootstrap method with 1000 samples was used to calculate the probability of each coefficient being zero. All the predictor variables were standardized using the z-score method prior to model fitting. Figure 3 shows box plots of all 10 coefficient estimates generated from bootstrap LASSO. The red line in the figure shows the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points that are not considered outliers, and the outliers are shown using the + symbol.

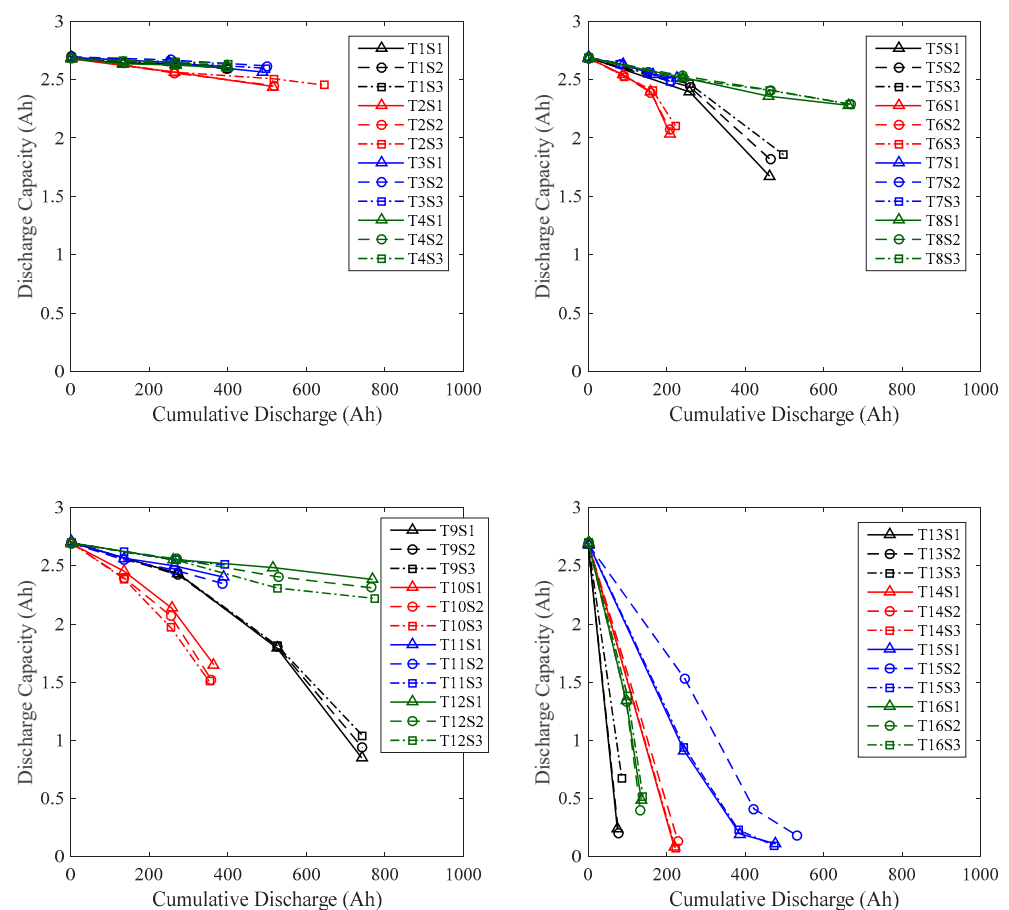


Figure 1. Variation of discharge capacity with cumulative discharge (Ah) for Manufacturer 1.

After dropping the DOD variable as mentioned earlier and choosing the conditions with DOD = 1, the eight test conditions represent resolution IV design where the two-way interaction of any two stress factors will be confounded with the two-way interaction of the remaining two stress factors. Since LASSO allows the inclusion of all the interactions together even with fewer test conditions, it can automatically identify which of the two-way interactions are more useful than the others. The two-way interactions of discharge

C-rate, charge current cut-off, and charge C-rate (β_{23} , β_{24} , β_{34}) are very close to zero and hence are not useful in describing the response Y . Figure 4 shows the ordering of coefficients in terms of their bootstrap probability of zero. The variables with the least probability are more useful for describing the response Y . The probability ranking also correlates with the medians of absolute effect size estimates (Figure 3) with the most useful variable temperature showing the largest absolute effect size estimate median. Temperature, discharge C-rate, and their two-way interaction are the first three ranked effects as per LASSO.

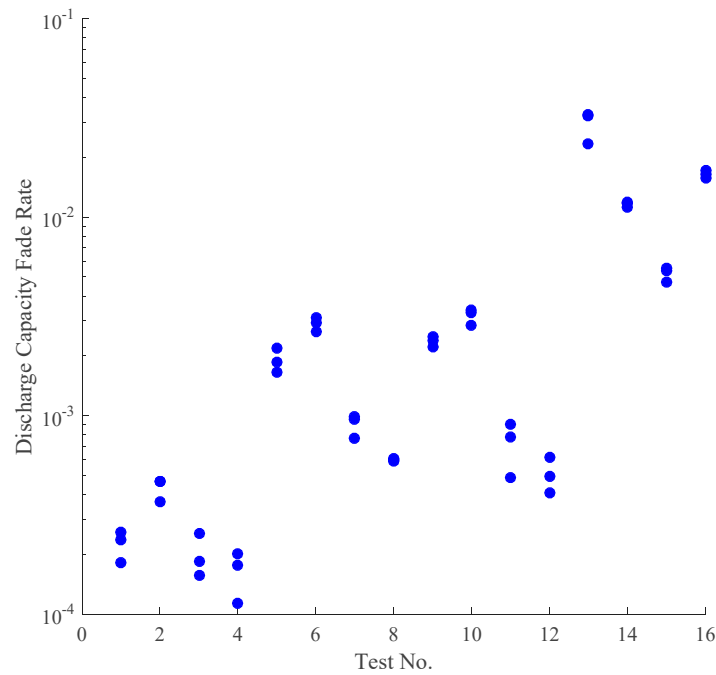


Figure 2. Capacity fade rate (Y) for batteries from Manufacturer 1 for Tests 1–16.

Table 2. List of predictor variables and corresponding model coefficients.

Predictor Variable	Coefficient/Effect Size
Intercept	β_0
Temperature	β_1
Discharge C-rate	β_2
Charge Current Cut-off	β_3
Charge C-rate	β_4
DOD	β_5
Temperature \times Discharge C-rate	β_{12}
Temperature \times Charge Current Cut-off	β_{13}
Temperature \times Charge C-rate	β_{14}
Temperature \times DOD	β_{15}
Discharge C-rate \times Charge Current Cut-off	β_{23}
Discharge C-rate \times Charge C-rate	β_{24}
Discharge C-rate \times DOD	β_{25}
Charge Current Cut-off \times Charge C-rate	β_{34}
Charge Current Cut-off \times DOD	β_{35}
Charge C-rate \times DOD	β_{45}

Similarly, when training and deploying the RF algorithm on Manufacturer 1 data, the results in Figure 5 are obtained. In agreement with the LASSO predictor, RF confirms that discharge C-rate and temperature are the two most significant covariates for capacity fade under life cycle testing. Individual effects of charge current cut-off value and charge current C-rate are ranked as the two least significant predictor variables. A higher charge C-rate (than the range considered in this study) might contribute to a higher acceleration

and may alter the ranking; however, high charge rates are expected to induce the additional degradation mechanism of lithium plating.

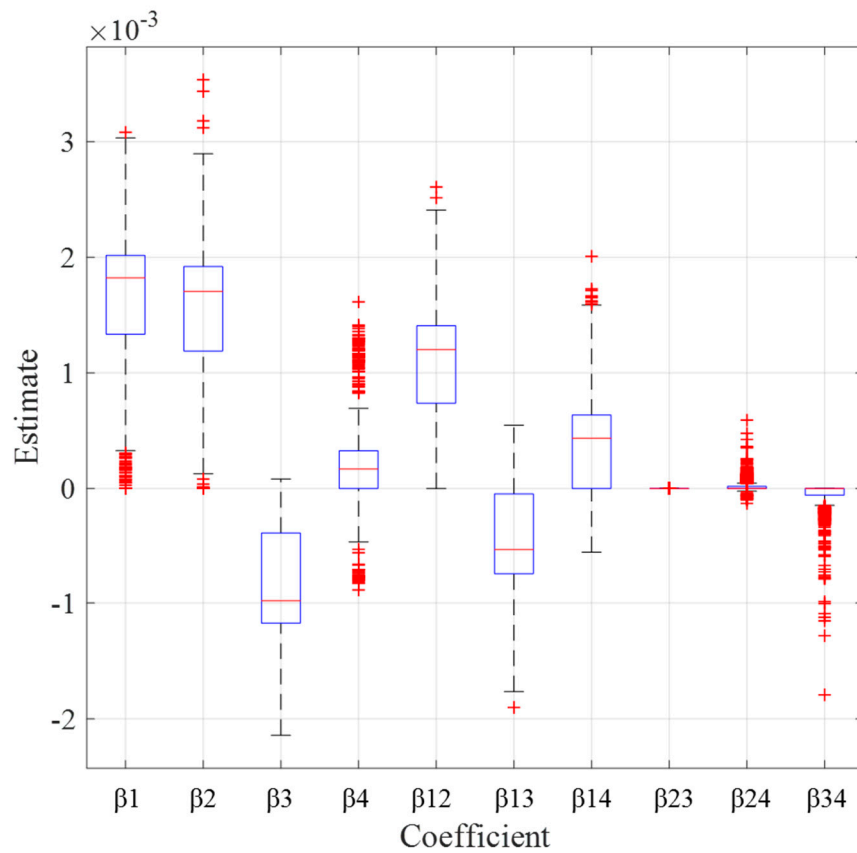


Figure 3. Box plots of (individual and interaction) effect size estimates generated from bootstrap least absolute shrinkage and selection operator (LASSO) for Manufacturer 1.

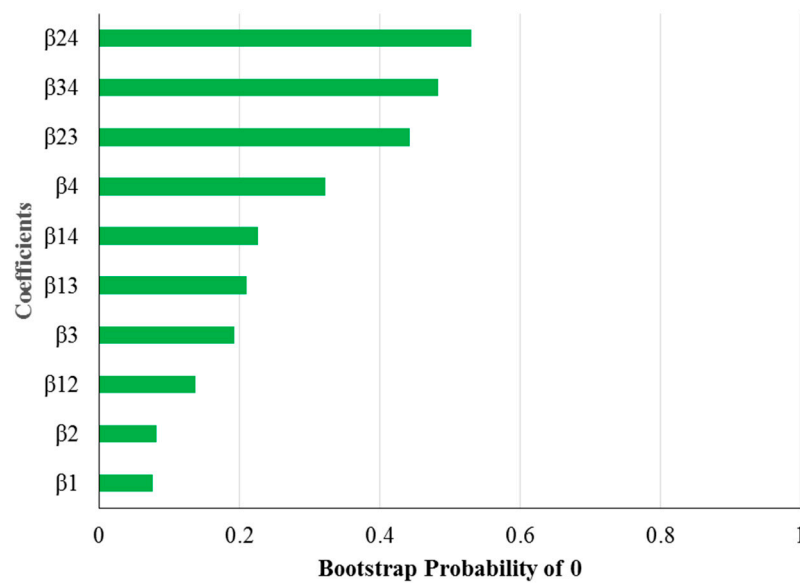


Figure 4. Ordering of (individual and interaction) effect sizes in terms of their bootstrap probability of zero for Manufacturer 1.

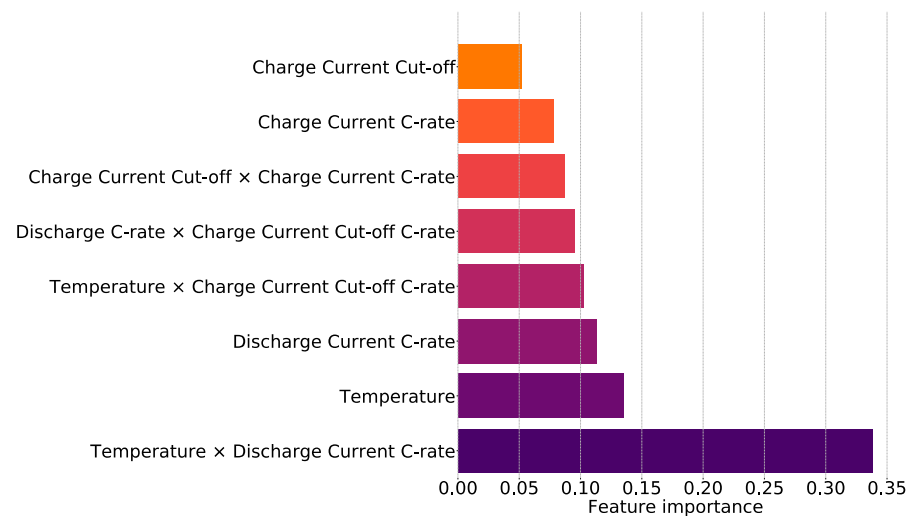


Figure 5. Random forest (RF) results for predictor variable (feature) ranking in decreasing order of importance for Manufacturer 1.

Based on the analysis from two algorithms, temperature and discharge C-rate are considerably more significant in accelerating capacity fade of batteries from Manufacturer 1. In addition, a combination of high temperature and high discharge C-rate would result in further acceleration due to the significant interaction effect of these two variables.

4.2. Case Study 2

The Li-ion batteries from Manufacturer 2 with a minimum capacity of 4450 mAh were used for this case study. Batteries were tested under conditions described in Table 1 with one change in the high level of discharge C-rate. A high discharge rate of 1.3C was used instead of 2C to follow the manufacturer-specified upper discharge C-rate limit. The capacity fade results are presented in Figure 6. Data for all three samples under each test case are plotted. The batteries from Manufacturer 2 show a large capacity fade rate for 55 °C and 1.3C test conditions (Tests 13–15) in Figure 6. Testing for samples 2 and 3 under test 15 was stopped 100 cycles earlier than sample 1 due to a test equipment problem. Surprisingly, batteries under test cases 9 and 10 at 55 °C and 0.5C test conditions showed more capacity fade than that of Test 16 under 55 °C and 1.3C conditions. Test 9 had higher stress in terms of charge cut-off current and DOD as compared to Test 16, while Test 10 had higher stress in terms of only charge cut-off as compared to Test 16. Sample 1 of the 25 °C and 2C case (Test 8) showed a higher capacity fade rate compared to the other two samples, possibly indicating poor quality. The average capacity fade rate for each test case and sample is plotted in Figure 7. The capacity fade rate is much higher for test cases at 55 °C (Tests 9–16). Unlike batteries from Manufacturer 1, these batteries were able to maintain the DOD values of 0.5 closely except for the two samples under Test 13. However, even for those two samples, the DOD shifts away from 0.5 only toward the end of testing in the region of the sharp capacity drop in Figure 6. Therefore, it is reasonable to assume constant DOD for all test cases and use it as a predictor variable for analyzing the data of the batteries from Manufacturer 2.

The inclusion of DOD leads to five individual effects and 10 two-way interaction effects of stress factors, which should be estimated from the 16 test conditions described in Table 1. While fitting the model, it was observed that $\ln(Y)$ provided a better fit for a linear model; hence, for these batteries, $\ln(Y)$ was subsequently used as the response variable for LASSO regression. All the predictor variables were standardized using the z-score method prior to model fitting.

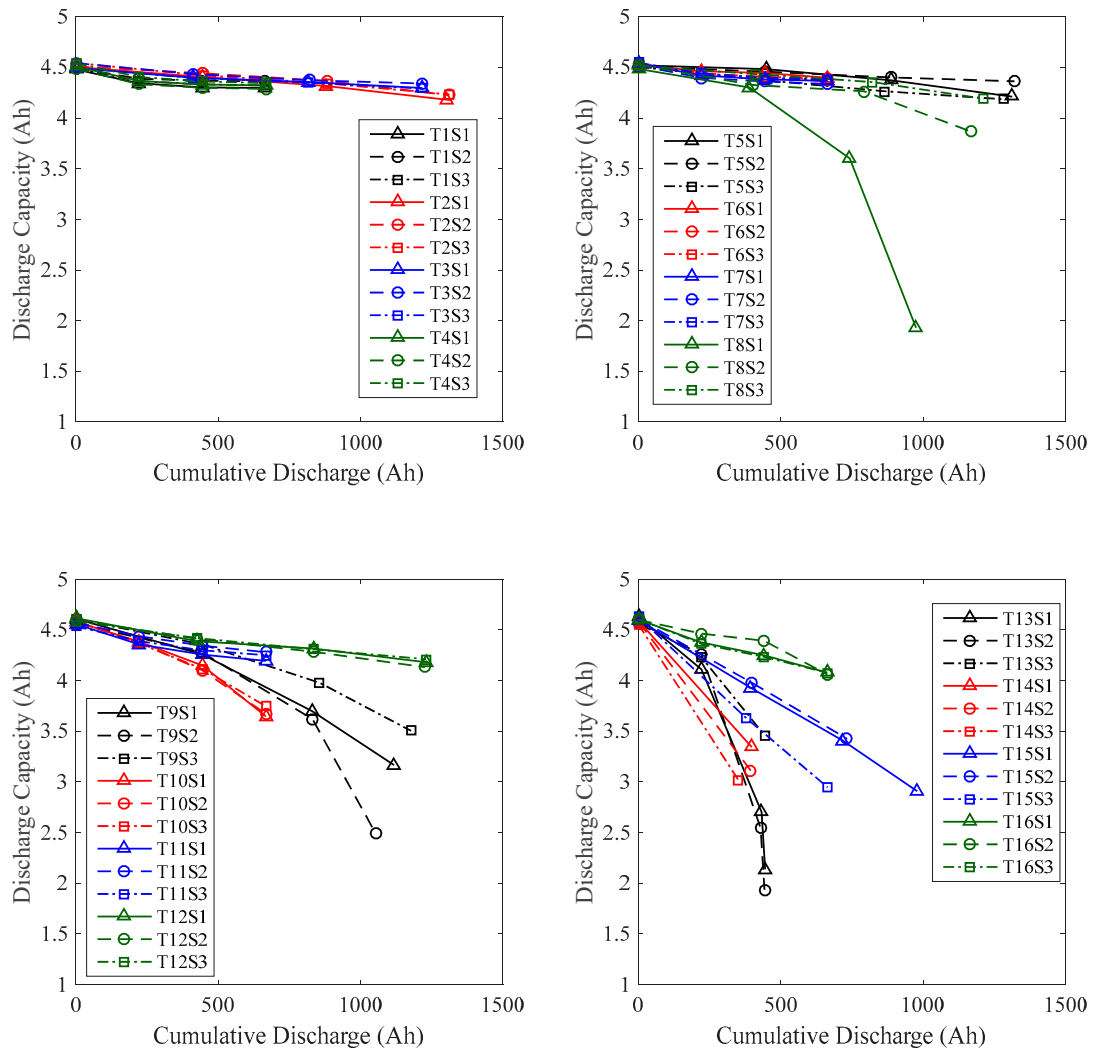


Figure 6. Variation of discharge capacity with cumulative discharge (Ah) for Manufacturer 2.

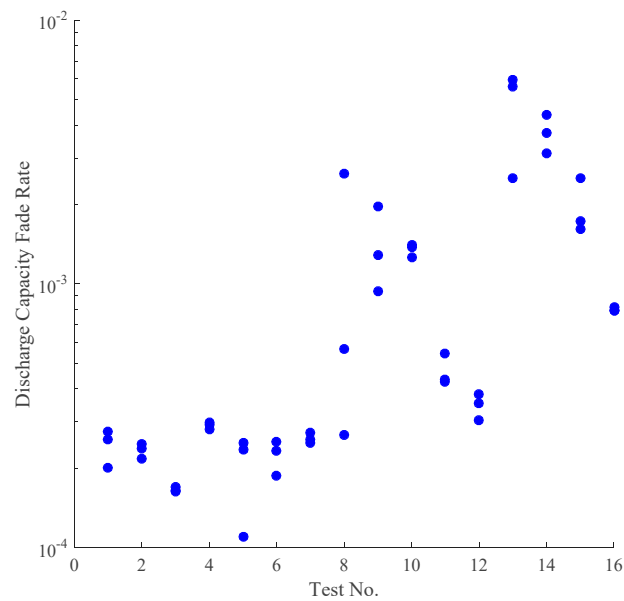


Figure 7. Capacity fade rate (Y) for batteries from Manufacturer 2 for Tests 1–16.

LASSO regression analysis was performed with all 15 individual and two-way interaction effects-based predictor variables using data from all 16 test conditions mentioned in Table 1. Figure 8 shows the box plots of all 15 coefficient estimates generated from bootstrap LASSO using λ corresponding to one standard deviation MSE in a 10-fold cross-validation. The median estimate is roughly zero for coefficients β_4 , β_5 , β_{15} , β_{24} , β_{34} , β_{35} , and β_{45} ; hence, LASSO automatically removes the predictor variables not useful for describing response $\ln(Y)$. Figure 9 shows the ordering of predictor variables using a bootstrap probability of zero. As per LASSO, the three most influential predictor variables are temperature, interaction of temperature and charge current cut-off, and discharge C-rate.

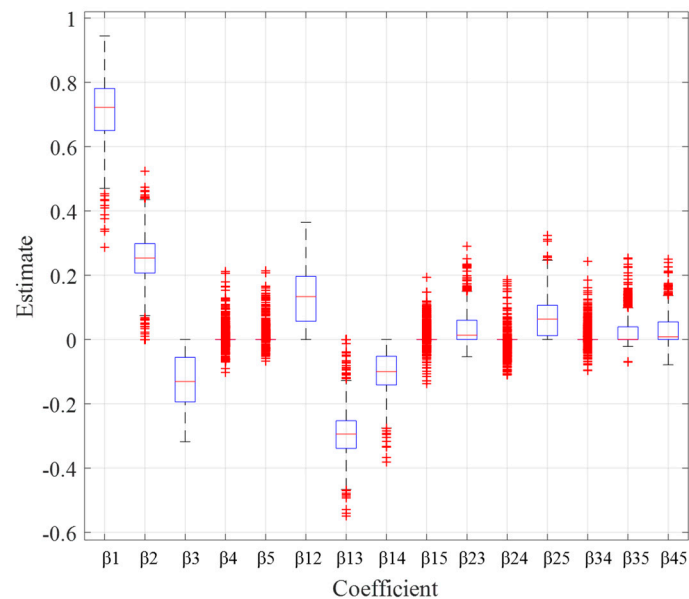


Figure 8. Box plots of (individual and interaction) effect size estimates generated from bootstrap LASSO for Manufacturer 2.

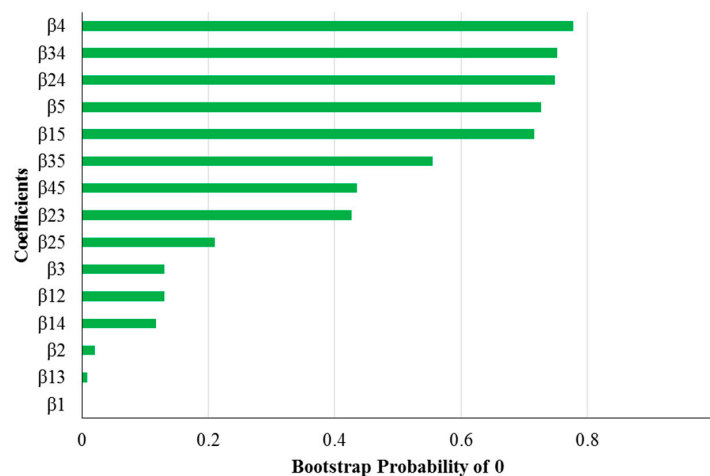


Figure 9. Ordering of (individual and interaction) effect sizes in terms of their bootstrap probability of zero for Manufacturer 2.

The RF algorithm was also deployed on data from Manufacturer 2, and the results are illustrated in Figure 10. The RF algorithm concludes that the three most influential predictor variables are interaction between temperature and discharge C-rate, interaction between discharge C-rate and charge current cut-off, and individual effect of temperature; however, the algorithm ranks interaction between temperature and discharge C-rate as far more important than all other covariates.

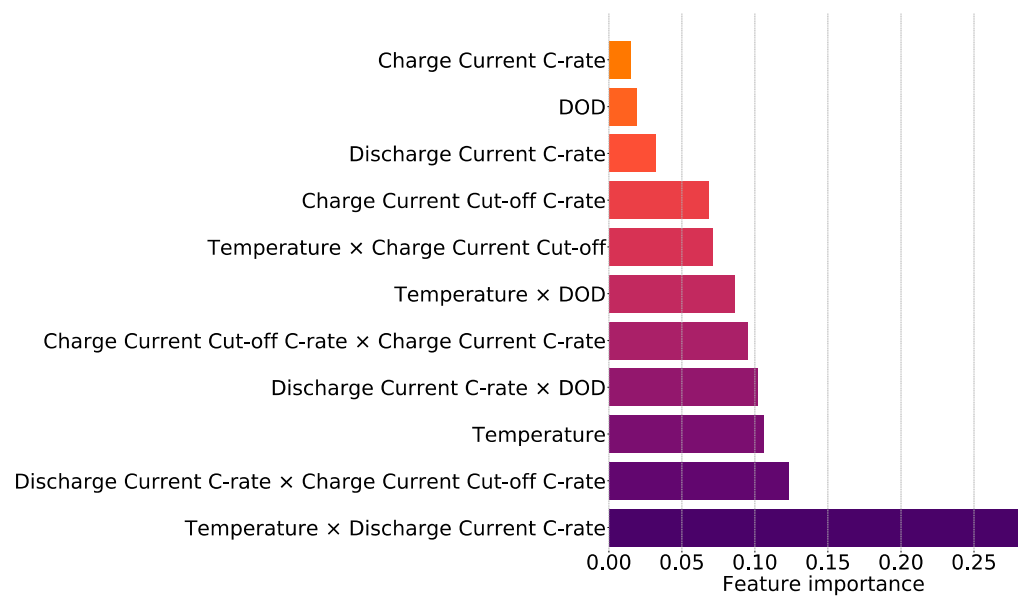


Figure 10. RF results for predictor variable (feature) ranking in decreasing order of importance for Manufacturer 2.

The RF algorithm also indicates that individual covariates such as charge current C-rate and DOD are the least important predictor variables; however, their interactions certainly contribute to the onset of degradation. In conclusion, Manufacturer 2's batteries were less sensitive to individual covariate values; rather, acceleration of degradation was governed by the interaction of temperature, discharge rates, and cut-off current values. RF also shows the importance of temperature, discharge C-rate, and charge cut-off current in terms of interactions; however, there is again some inconsistency in the ranking of interaction variables and their constituent individual effect variables compared to the LASSO algorithm.

The ranking of individual and interaction effects of stress factors obtained by LASSO is related with the effect size (β) estimates (Figures 3 and 8). The importance of a variable obtained by RF (Figures 5 and 10), on the other hand, is a measure of relative improvement in predictive accuracy (during cross-validation on validation set) due to the inclusion of that variable. LASSO, a parametric method, is well-suited for model interpretation and variable ranking. Thus, ranking obtained by LASSO is critical for this study considering its focus on the effect sizes. RF, a non-parametric method, is more suited for prediction tasks and provides variable importance as an extra benefit.

The differences in LASSO ranking of individual and interaction effects for the two manufacturers will not affect the selection order of the factors for accelerated test planning. Temperature should be the preferred factor for both the manufacturers for one-factor-based accelerated test design. For two-factor-based accelerated test design, temperature and discharge C-rate can be preferred factors for both manufacturers—the combination of effects sizes β_2 and β_{12} is roughly equivalent to combination of β_3 and β_{13} for Manufacturer 2 (Figure 8). Temperature, discharge C-rate, and charge cut-off current will be the preferred factors for three-factor-based accelerated test design.

5. Conclusions

This study determines and ranks the individual and two-way interaction effects of five stress factors on the battery capacity fade to identify most effective stress factors for accelerated battery test planning. This work also demonstrates an application of machine learning in ranking the battery stress factors and their interactions. Using a half-fractional factorial design with 16 tests, the capacity fade of 96 Li-ion batteries was studied under five stress factors (ambient temperature, discharge C-rate, charge C-rate, DOD, and constant

voltage charge cut-off current) that constitute a typical cycling profile for batteries in the qualification testing. The half-fractional factorial design was sufficient to preserve the individual (main) and two-way interaction effects of stress factors.

Machine learning methods, LASSO and random forest, were implemented for ranking the effects of stress factors. Temperature, discharge C-rate, and constant voltage charge cut-off current stress factors in the form of either their individual effects or interactions were always among the top four ranked effects for the capacity fade of batteries from the two manufacturers. Individual effects of charge C-rate and DOD were found to be least significant for the battery capacity fade. The study shows that charge cut-off current during the constant voltage charging phase, a relevant parameter during battery charging that has been under represented in the battery literature, was more dominant for battery degradation compared to the charge current used during the constant current charging phase.

The stress factor rank list obtained from this study can be used for planning accelerated Li-ion battery qualification testing with only top-ranked factors. However, the ranking should be read in context with the ranges of stress factors chosen in the experimental design, as harsher levels of an insignificant factor of charge C-rate may have the potential to trigger additional degradation mechanisms such as lithium plating. The selection of only two levels of each stress factor in the design of experiments, although sufficient for ranking purposes, limits the determination of non-linear effects of stress factors. While the implementation of RF in this study has been limited to the factor ranking, it can also be used for data-driven modeling to predict battery capacity fade in future work.

Author Contributions: S.S. designed the experiments, processed experimental data, and implemented LASSO; D.R. implemented random forest method and assisted with data analysis; V.R., D.F., and M.P. assisted with the problem statement, provided the technical feedback, and supervised the study; All the authors wrote and edited the paper. All authors have read and agreed to the published version of the manuscript.

Funding: The authors are grateful to the UK EPSRC Centre for Doctoral Training in Embedded Intelligence under grant reference EP/P001173/1 for financially supporting the machine learning modeling work. The authors would like to thank the more than 150 companies and organizations that support research activities at the Center for Advanced Life Cycle Engineering at the University of Maryland annually. Furthermore, the authors would like to acknowledge the support of the Lloyds Register Foundation under grant AtRI_100015.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Couraud, B.; Norbu, S.; Andoni, M.; Robu, V.; Gharavi, H.; Flynn, D. Optimal Residential Battery Scheduling with Asset Lifespan Consideration. In Proceedings of the 2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), The Hague, The Netherlands, 26–28 October 2020; pp. 630–634. [CrossRef]
2. Tang, W.; Roman, D.; Dickie, R.; Robu, V.; Flynn, D. Prognostics and Health Management for the Optimization of Marine Hybrid Energy Systems. *Energies* **2020**, *13*, 4676. [CrossRef]
3. The Market for Lithium-Ion Battery is Expected to Register a CAGR of Approximately 22%, during the Forecast Period (2019) Markets Insider. Available online: <https://markets.businessinsider.com/news/stocks/the-market-for-lithium-ion-battery-is-expected-to-register-a-cagr-of-approximately-22-during-the-forecast-period-2019-1028690618> (accessed on 8 January 2021).
4. A Behind the Scenes Take on Lithium-Ion Battery Prices | BloombergNEF. Available online: <https://about.bnef.com/blog/behind-scenes-take-lithium-ion-battery-prices/> (accessed on 8 January 2021).
5. Hendricks, C.; Williard, N.; Mathew, S.; Pecht, M. A failure modes, mechanisms, and effects analysis (FMMEA) of lithium-ion batteries. *J. Power Sources* **2015**, *297*, 113–120. [CrossRef]
6. Andoni, M.; Tang, W.; Robu, V.; Flynn, D. Data analysis of battery storage systems. *CIREC Open Access Proc. J.* **2017**, *2017*, 96–99. [CrossRef]
7. Voltaiq Survey on Battery Industry Finds Analytical Challenges and Resource Constraints as Major Obstacles to Product Development—Green Car Congress. Available online: <https://www.greencarcongress.com/2019/06/20190626-voltaiq.html> (accessed on 8 January 2021).

8. Engel, H.; Hertzke, P.; Siccardo, G. Electric Vehicles, Second Life Batteries, and Their Effect on the Power Sector | McKinsey. 2019. Available online: <https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/second-life-ev-batteries-the-newest-value-pool-in-energy-storage> (accessed on 25 January 2021).
9. Specification Approval Sheet. Available online: http://www.tenergy.com/30005_datasheet.pdf (accessed on 25 January 2021).
10. Smith, K.; Saxon, A.; Keyser, M.; Lundstrom, B.; Cao, Z.; Roc, A. Life prediction model for grid-connected Li-ion battery energy storage system. In Proceedings of the American Control Conference, Seattle, WA, USA, 24–26 May 2017; pp. 4062–4068. [CrossRef]
11. Lithium-Ion Battery Data Sheet. September 2010. Available online: <https://www.ineltro.ch/media/downloads/SAAltem/45/45958/36e3e7f3-2049-4adb-a2a7-79c654d92915.pdf> (accessed on 25 January 2021).
12. Diao, W.; Saxena, S.; Pecht, M. Accelerated cycle life testing and capacity degradation modeling of LiCoO₂-graphite cells. *J. Power Sources* **2019**, *435*, 226830. [CrossRef]
13. Wang, J.; Liu, P.; Hicks-Garner, J.; Sherman, E.; Soukiazian, S.; Verbrugge, M.W.; Tataria, H.; Musser, J.W.; Finamore, P. Cycle-life model for graphite-LiFePO₄ cells. *J. Power Sources* **2011**, *196*, 3942–3948. [CrossRef]
14. Schimpe, M.; Von Kuepach, M.E.; Naumann, M.; Hesse, H.C.; Smith, K.; Jossen, A. Comprehensive Modeling of Temperature-Dependent Degradation Mechanisms in Lithium Iron Phosphate Batteries. *J. Electrochem. Soc.* **2018**, *165*, A181–A193. [CrossRef]
15. Cui, Y.; Du, C.; Yin, G.; Gao, Y.; Zhang, L.; Guan, T.; Yang, L.; Wang, F. Multi-stress factor model for cycle lifetime prediction of lithium ion batteries with shallow-depth discharge. *J. Power Sources* **2015**, *279*, 123–132. [CrossRef]
16. Prochazka, W.; Pregartner, G.; Cifrain, M. Design-of-Experiment and Statistical Modeling of a Large Scale Aging Experiment for Two Popular Lithium Ion Cell Chemistries. *J. Electrochem. Soc.* **2013**, *160*, A1039–A1051. [CrossRef]
17. Su, L.; Zhang, J.; Wang, C.; Zhang, Y.; Li, Z.; Song, Y.; Jin, T.; Ma, Z. Identifying main factors of capacity fading in lithium ion cells using orthogonal design of experiments. *Appl. Energy* **2013**, *160*, A1039–A1051. [CrossRef]
18. Maleki, H.; Deng, G.; Anani, A.; Howard, J. Thermal Stability Studies of Li-Ion Cells and Components. *J. Electrochem. Soc.* **1999**, *146*, 3224–3229. [CrossRef]
19. Tang, W.; Andoni, M.; Robu, V.; Flynn, D. Accurately Forecasting the Health of Energy System Assets. In Proceedings of the IEEE International Symposium on Circuits and Systems, Florence, Italy, 27–30 May 2018. [CrossRef]
20. He, W.; Williard, N.; Osterman, M.; Pecht, M. Prognostics of lithium-ion batteries based on Dempster-Shafer theory and the Bayesian Monte Carlo method. *J. Power Sources* **2011**, *196*, 10314–10321. [CrossRef]
21. Nie, F.; Huang, H.; Cai, X.; Ding, C. Efficient and Robust Feature Selection via Joint 2,1-Norms Minimization. In Proceedings of the Advances in Neural Information Processing Systems 23 (NIPS 2010), Vancouver, BC, Canada, 6–11 December 2010.
22. Hapfelmeier, A.; Ulm, K. A new variable selection approach using Random Forests. *Comput. Stat. Data Anal.* **2013**, *60*, 50–69. [CrossRef]
23. Guyon, I.; Elisseeff, A. An Introduction to Variable and Feature Selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.
24. Miao, J.; Niu, L. A Survey on Feature Selection. *Procedia Comput. Sci.* **2016**, *91*, 919–926. [CrossRef]
25. Saeys, Y.; Inza, I.; Larrañaga, P. A review of feature selection techniques in bioinformatics. *Bioinformatics* **2007**, *23*, 2507–2517. [CrossRef] [PubMed]
26. Tibshirani, R. Regression Shrinkage and Selection Via the Lasso. *J. R. Stat. Soc. Ser. B* **1994**, *58*, 267–288. [CrossRef]
27. Hastie, T.; Tibshirani, R.; Wainwright, M. *Statistical Learning with Sparsity: The Lasso and Generalizations*; CRC Press, LLC: Boca Raton, FL, USA, 2015.
28. Archer, K.J.; Kimes, R.V. Empirical characterization of random forest variable importance measures. *Comput. Stat. Data Anal.* **2008**, *52*, 2249–2260. [CrossRef]
29. Nguyen, T.T.; Huang, J.Z.; Nguyen, T.T. Unbiased feature selection in learning random forests for high-dimensional data. *Sci. World J.* **2015**, *2015*, 471371. [CrossRef] [PubMed]
30. Bergstra, J.; Bengio, Y. Random Search for Hyper-Parameter Optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
31. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
32. Louppe, G.; Wehenkel, L.; Suter, A.; Geurts, P. Understanding variable importances in forests of randomized trees. In Proceedings of the Advances in Neural Information Processing Systems 26 (NIPS 2013), Lake Tahoe, NV, USA, 5–10 December 2013.