Expert Systems With Applications 197 (2022) 116817

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

State-of-health estimation of Li-ion batteries in the early phases of qualification tests: An interpretable machine learning approach

Gyumin Lee [a], Juram Kim [b], Changyong Lee [c,*,1]

[a] School of Business Administration, Ulsan National Institute of Science and Technology, Republic of Korea
[b] Center for R&D Investment and Strategy Research, Korea Institute of Science and Technology Information, Republic of Korea
[c] Graduate School of Management of Technology, Sogang University, Republic of Korea

ARTICLE INFO

Keywords:
State-of-health estimation
Qualification test
Li-ion battery
Interpretable machine learning
SHapley Additive exPlanation method

ABSTRACT

Reducing the time and cost associated with lithium-ion (Li-ion) battery qualification tests is critical to developing electronic devices and establishing their quality assurance policies. In this study, we develop an interpretable machine learning model for estimating the future state-of-health (SOH) of Li-ion batteries in the early phases of qualification tests. First, a window-moving technique is used to extract the statistical features that represent battery capacity-fading behaviors over certain cycles. Second, a machine learning model is developed to estimate a battery's future SOH value at a certain cycle. Third, the performance and reliability of the machine learning model are assessed using multiple experiments with varying forecast horizons for SOH estimation. Finally, the SHapley Additive exPlanation (SHAP) method is applied to the model to identify which statistical features are important when estimating a battery's SOH value. The experimental results confirm that the proposed approach can reduce the time required for qualification tests to 100 cycles, i.e., less than a month in practice, with less than a 5% mean absolute percentage error (MAPE) and a 0.002 mean squared error (MSE). The results of model interpretation by SHAP demonstrate that the changes in the SOH values of Li-ion batteries are more important than the values themselves to the SOH estimation. Moreover, the SOH degradation trends near the 100th cycle during the qualification tests are proved to have a significant impact on the future SOH values of the batteries.

1. Introduction

Lithium-ion (Li-ion) batteries have been widely used across industries from portable electronics and electric vehicles to power grids and energy storage systems because of their high energy and power density, high efficiency, long life cycles, and low maintenance requirements (Huang et al., 2019). Regardless of applications, Li-ion batteries exhibit a loss of their capacity with time, as a result of repeated charging and discharging cycles. The capacity fade is caused by various failure mechanisms such as solid electrolyte interface growth, electrode porosity reduction, and electrode particle fracture (Hendricks et al., 2015; Vetter et al., 2005).

A Li-ion battery is considered to have reached its end-of-life (EOL) if its state-of-health (SOH) of the battery, which is defined as the ratio of the battery's residual capacity to its initial value, drops below a certain threshold (e.g., 0.8) (Cannarella & Arnold, 2014; Le & Tang, 2011; Spotnitz, 2003). Qualification tests on Li-ion batteries comprise a series of charging and discharging steps to monitor the SOH degradation of the batteries. Although the qualification tests can be accelerated by controlling various factors that influence the rate of SOH degradation such as operating temperature (Zhang et al., 2020), cell-to-cell variation (Wang et al., 2021), and C-rate loading (Saxena et al., 2019; Zhang et al., 2021), the accelerated testing takes several months in practice as it requires hundreds of cycles to determine the remaining useful life (RUL) of the batteries (Johnson & White, 1998; Saxena et al., 2018). In practice, the qualification tests are considered a major bottleneck in developing electronic devices and establishing quality assurance policies (Lee et al., 2019; Pecht et al., 2016). Therefore, industrial practitioners have called for systematic methods that reduce the time and cost associated with the qualification tests on Li-ion batteries.

Curve-fitting approaches have been widely used to model the capacity-fading behavior of Li-ion batteries during qualification tests. In this regard, various parameter estimation methods that fit prescribed curves to capacity fade data from the batteries and extrapolate the

* Corresponding author at: Graduate School of Management of Technology, Sogang University, 35 Baekbeom-ro, Mapo-gu, Seoul 04107, Republic of Korea.
E-mail address: changyong@sogang.ac.kr (C. Lee).
[1] ORCID ID: 0000-0002-1636-9481.

https://doi.org/10.1016/j.eswa.2022.116817
Received 11 August 2021; Received in revised form 1 March 2022; Accepted 2 March 2022
Available online 5 March 2022

curves beyond the range have been employed (Barbarisi et al., 2006; Zhang et al., 2009; He et al., 2011; Junping et al., 2009; Matsushima, 2009; Olivares et al., 2013; Sun et al., 2011; Xiong et al., 2017; Yu et al., 2015). For instance, Matsushima (2009) proved that the capacity fade is proportional to the square root of the time the battery has been used. Although the curve-fitting approaches have proved useful for observing the overall capacity fade trend of Li-ion batteries, they depend on unrealistic assumptions regarding the prescribed capacity fade curves, considering that the capacity-fading behaviors of the batteries are nonlinear and heterogeneous across individual batteries (Lee et al., 2020). Furthermore, these approaches struggle to estimate future SOH values of Li-ion batteries in the early phases of the qualification test, when the inflection points on which the capacity fade of Li-ion batteries accelerates are not yet observed (Saxena et al., 2018; Zhang & Lee, 2011).

Recent literature presents a trend towards developing machine learning models that address the nonlinearity and heterogeneity of the capacity-fading behaviors of Li-ion batteries, which can be classified into two categories. One approach adopts classification models to identify unhealthy batteries in the early phases of qualification tests. For instance, Saxena et al. (2018) presented a one-class support vector machine (SVM) that identifies anomalous batteries. Lee et al. (2020) developed a capacity-fading behavior analysis method that distinguishes unhealthy batteries from healthy batteries based on a local outlier factor, kernel density estimation, and a hidden Markov model. The other approach uses regression models to estimate the future SOH values of Li-ion batteries based on the temporal characteristics observed from capacity fade data. For instance, Widodo et al. (2011) developed a prognostic framework for battery health based on an SVM and a relevance vector machine using the time-series features extracted using a sample entropy algorithm. Li et al. (2020) introduced a prognostic framework for predicting the SOH and RUL of Li-ion batteries based on a variant long short-term memory (LSTM) shared by multiple cells without mutual interference. Razavi-Far et al. (2019) presented an integrated imputation-prediction scheme for estimating the RUL of Li-ion batteries with missing observations based on extreme learning machines. Chaoui & Ibe-Ekeocha (2017) proposed an autoregressive architecture based on recurrent neural networks for SOH and state-of-charge (SOC) estimation.

However, while machine learning models have proved quite useful for identifying unhealthy batteries and estimating SOH values of the batteries, the results of prior studies cannot be easily generalized and deployed in practice for the following reasons: First, the sample size of the capacity fade data used in previous studies is too small. Second, the experiments they conducted only cover a limited context of analysis regarding the forecast horizons for the SOH estimation. Finally, the prior studies do not effectively assist experts in decision-making because they rely solely on black-box models whose internal mechanisms are not revealed. In practice, the key to the successful deployment of machine learning models is interpretability, which makes the models' behaviors and predictions understandable to humans and further transfers learning into a knowledge base (Adadi & Berrada, 2018). For instance, if machine learning models provide different outcomes from those of experts, such inconsistencies should be reconciled by understanding and verifying how the models arrived at their conclusions. When used for SOH estimations, interpretability allows us to assess the feasibility and reliability of the analysis results and provides insight into methods for improving the machine learning models and further qualification testing procedures.

Considering these issues, we develop an interpretable machine learning model for estimating future SOH values of Li-ion batteries in the early phases of qualification tests. First, a window-moving technique is used to extract the statistical features (i.e., moving average, moving first-order difference, and moving variance) that represent a battery's capacity-fading behavior over certain cycles. Second, a machine learning model is developed to estimate the SOH value of the batteries at a certain cycle. Third, the performance and reliability of the developed machine learning model are assessed through multiple experiments with varying forecast horizons for the SOH estimation. Finally, the SHapley Additive exPlanation (SHAP) is applied to the machine learning model to identify which statistical features are important when estimating a battery's SOH value.

The developed approach employs six representative machine learning models (i.e., random forest (RF), gradient boosting machine (GBM), SVM, k-nearest neighbor (KNN), multilayer perceptron (MLP), and gated recurrent unit (GRU)) and the capacity fade data of Li-ion batteries obtained from actual qualification tests. The analysis results confirm that the developed approach can reduce the time required for qualification tests to 100 cycles, i.e., less than a month in practice, with less than a 5% mean absolute percentage error (MAPE) and a 0.002 mean squared error (MSE). The results of model interpretation by SHAP demonstrate that the changes in the SOH values of Li-ion batteries are more important than the values themselves to the SOH estimation. Moreover, the SOH degradation trends near the 100th cycle during the qualification tests are proved to have a significant impact on the future SOH values of the batteries. As a summary, the key contributions of this study are presented, as follows:.

- Over 300 capacity fade data samples from Li-ion batteries obtained from real-world qualification tests by a commercial automobile company are used. Furthermore, a synthetic minority oversampling technique (SMOTE) is used to complement the small sample size regarding the batteries that exhibit rarely observed capacity fade trends. This strategy provides the machine learning model with an ample sample size of data so that the model can be fully trained for SOH estimation.
- Multiple experiments are conducted with different combinations of forecast origin and forecast target cycles to examine the practicality of the proposed approach in diverse contexts. Through the experiments, the ability of the developed machine learning model to predict the results for actual qualification tests on Li-ion batteries is thoroughly investigated.
- The use of the SHAP for examining the contributions of statistical features to the SOH estimation can be helpful when addressing the inconsistencies between the forecasting outcomes from industrial practitioners and machine learning models; this method can also be used to improve the qualification testing procedures on Li-ion batteries.
- The results of the SOH estimation and model interpretation show that the developed approach can serve as a practical method for reducing the duration of qualification tests on Li-ion batteries, as well as will help improve the efficiency and reliability of the qualification testing procedures.

The remainder of this paper is organized as follows: Section 2 explains the methodology. Section 3 presents the experimental settings employed in this study. Section 4 reports the experimental results and discusses the performance and reliability of the proposed interpretable machine learning model. Finally, Section 5 concludes with the limitations of the current study and suggests future research directions.

2. Methodology

The proposed approach consists of four steps, as illustrated in Fig. 1. First, in the feature extraction step, the SOH values of Li-ion batteries are calculated from capacity fade data, the forecast horizons for SOH estimation are determined, and then the statistical features are extracted. Second, in the SOH estimation step, a machine learning model is trained to estimate future SOH values of the batteries. Third, the performance and reliability of the developed machine learning model are examined by establishing performance metrics and conducting five-fold stratified cross-validation. Finally, SHAP is applied to understand the machine

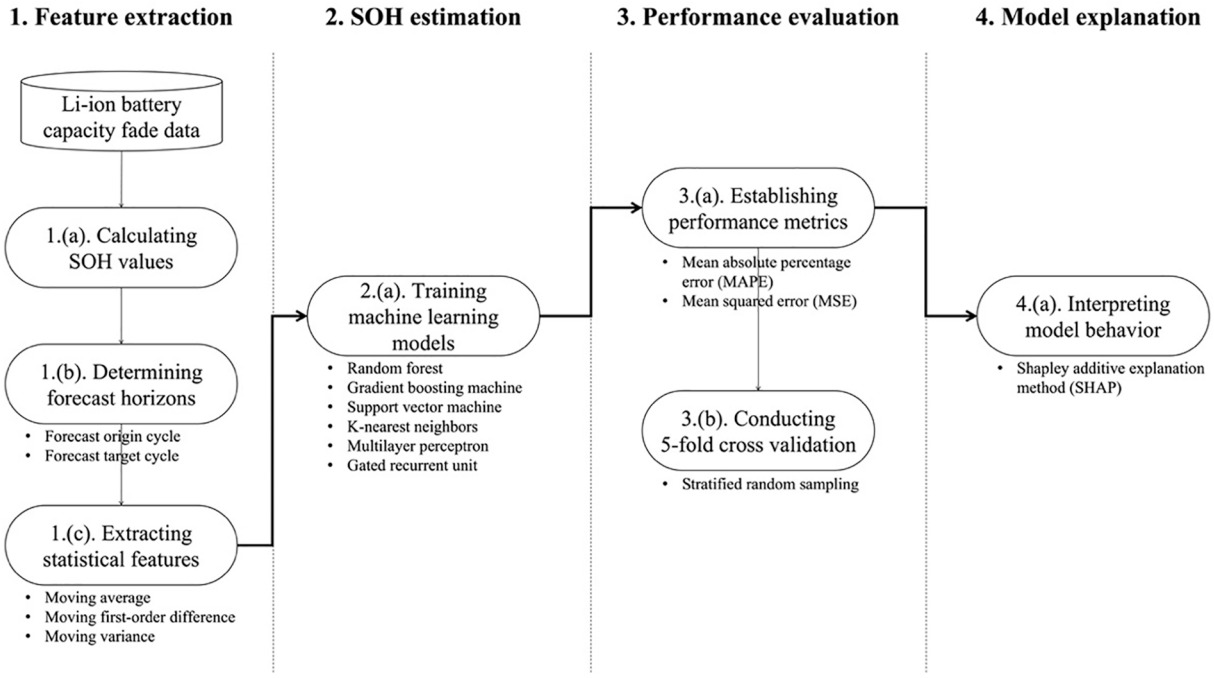

Fig. 1. The overall process of the proposed approach.

learning model's behaviors associated with the SOH estimation, as well as identify the contribution of the statistical features to the estimation.

2.1. Feature extraction

SOH value of a Li-ion battery at a certain cycle is defined as the rate of change in the battery capacity values from the first cycle to the cycle of interest as formulated in (1), where $C(0)$ and $C(t)$ represent the battery capacity at the first cycle and cycle t, respectively (Lee et al., 2020):.

$$SOH(t) = \frac{C(t)}{C(0)} \tag{1}$$

The forecast horizon for the SOH estimation is determined by setting a forecast origin cycle and forecast target cycle. In other words, using the accumulated SOH values from the beginning of the capacity fade data to the forecast origin cycle (e.g., 100th cycle), a machine learning model estimates the SOH value at the forecast target cycle (e.g., 700th cycle). However, the direct use of the SOH values accumulated until the forecast origin cycle for the future SOH estimation may not provide enough information regarding changes in the SOH degradation trend. Therefore, we define three types of statistical features: moving average ($movAVG(t)$), moving first-order difference ($movFOD(t)$), and moving variance ($movVAR(t)$). These features capture the slight changes and fluctuations in SOH values over cycles. Specifically, the statistical features for a certain cycle t are calculated by moving a prescribed window from the beginning to the forecast origin cycle as formulated in (2)–(4), where t indicates the index of the charging and discharging cycle and W represents the size of the prescribed window (Ma et al., 2019):.

$$movAVG(t) = \frac{\sum_{j=t}^{t+W-1} SOH(j)}{W} \tag{2}$$

$$movFOD(t) = \frac{\sum_{j=t}^{t+W-1} |SOH(j+1) - SOH(j)|}{W} \tag{3}$$

$$movVAR(t) = \frac{\sum_{j=t}^{t+W-1} \{SOH(j) - movAVG(t)\}^2}{W} \tag{4}$$

2.2. SOH estimation

To estimate the future SOH values of batteries, the machine learning model accepts the statistical features extracted from the capacity fade data up to the forecast origin cycle as inputs and estimates the SOH value at the forecast target cycle. The model is then trained to reduce the error between the actual SOH values and estimated SOH values of the battery samples in the dataset. In the proposed approach, six major machine learning algorithms are employed as the model for SOH estimation, as summarized below:.

- RF: An ensemble machine learning model based on a multitude of decision trees. Each tree is created using a randomly selected subset of the input variables as well as subsamples of the training datasets. The bagging technique is then used to predict the value of the output variable (Ayvaz & Alpay, 2021; Ho, 1995). After all of the trees are fully grown, RF performs a regression by averaging the prediction results from the trees, as formulated in (5):

$$\hat{f} = \frac{1}{B}\sum_{b=1}^{B} f_b(x'), \tag{5}$$

where $\hat{f}$ represents the final estimation, x' is a new data sample, f_b is an individual decision tree, and B is the number of decision trees.

- GBM: An ensemble machine learning model based on multiple weak estimators (Friedman, 2001; Moscatelli et al., 2020). This model first builds an imperfect baseline model $F_0(x)$, which estimates the value of the output variable, and then adds a series of residual estimators

$h_m(x)$ in the direction of improving the estimation accuracy, as described in (6)-(7):

$$F_0(x) = \underset{\gamma}{\operatorname{argmin}} \sum_{i-1}^{n} L(y_i, \gamma), \tag{6}$$

$$F_m(x) = F_{m-1}(x) + \underset{h_m}{\operatorname{argmin}} \left[\sum_{i-1}^{n} L(y_i, F_{m-1}(x_i) + h_m(x_i)) \right], \tag{7}$$

where γ represents the residual, L is a loss function, and F_m is the updated estimation model at the m^{th} iteration.

- SVM: A supervised machine learning model based on constructing a hyperplane that functions as a regression curve in a high-dimensional space (Hastie et al., 2009; Widodo et al., 2011). This model is trained by solving a quadratic optimization problem involving the hyperplane. The objective function of the optimization problem in the case of regression is defined as follows:

$$minimize \frac{1}{2}||w||^2, \ subjectto |y_i - \langle w, x_i \rangle - b| \le \varepsilon, \tag{8}$$

where w is the normal vector for the hyperplane, b is an intercept, and ε is a free parameter that serves as a threshold.

- KNN: A non-parametric supervised learning model based on the feature similarities between an object and its neighbors (Altman, 1992; Maleki et al., 2021). When it comes to the regression analysis, this model estimates the output variable value of a new data sample by calculating the average of the output variable values of the k nearest neighbors. The feature similarities can be calculated using the Euclidean and Manhattan measures, as expressed in (9)-(10):

$$Euclidean\ distance = \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2}, \tag{9}$$

$$Manhattan\ distance = \sum_{i=1}^{k} |x_i - y_i|, \tag{10}$$

where k denotes the number of input variables and x_i and y_i are the i^{th} input variables of two data samples x and$y$..

- MLP: A supervised machine learning model that makes predictions using feed-forward computations based on multiple layers of nodes (Hastie et al., 2009; Paliwal & Kumar, 2009). In each layer, the latent feature vector for a data sample is calculated by applying weights and biases and using a nonlinear activation function, as shown in (11):

$$o_l = \phi\left(\sum_i w_i x_i + b\right), \tag{11}$$

where i is the index of the data sample, l is the index of the layer, w_i indicates the weights, b represents the bias, and ϕ is the activation function. This model is trained using a back-propagation technique that minimizes the errors between the predicted output and the actual values.

- GRU: A derivative of recurrent neural networks primarily used for the estimation of time-series data based on recurrently connected hidden layers (Cho et al., 2014; Niu et al., 2020). This model adopts a gating mechanism that remedies the gradient vanishing problem, as described in (12)-(15):

$$update\ gate: z_t = \phi\left(W^{zx} x_t + W^{zh} h_{t-1}\right) \tag{12}$$

$$reset\ gate: r_t = \phi\left(W^{rx} x_t + W^{rh} h_{t-1}\right) \tag{13}$$

$$current\ memory\ content: \ c_t = tanh\left(W^{cx} x_t + W^{ch} r_t h_{t-1}\right) \tag{14}$$

$$memory\ state: \ h_t = (1 - z_t) h_{t-1} + z_t c_t \tag{15}$$

where t denotes a timestamp, x_t is the input vector at t, W denotes the parameter matrices, and ϕ is the activation function. In each time step, the output value is calculated based on the current input data and the state information from the previous time step.

2.3. Performance evaluation

A stratified five-fold cross-validation scheme is used to evaluate the performance of the machine learning model. Random sampling is applied within each stratum after the battery samples in the whole dataset were divided into homogeneous strata based on their SOH values at the forecast target cycle. For each validation fold, one group is used as a test dataset, and the remaining four groups are used as a training dataset. Specifically, the machine learning model is trained to learn the relationships between the statistical features extracted from the capacity fade data up to the forecast origin cycle and the actual SOH values at the forecast target cycle using the training dataset. Next, the model estimates the future SOH values at the forecast target cycle for the battery samples in the test set. These estimation results are used to evaluate the performance of the machine learning model.

To assess the performance of the developed machine learning model for the SOH estimation, the MAPE and MSE are employed (de Myttenaere et al., 2016; Puchalsky et al., 2018). These are defined as:.

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\left| SOH_t^i - \widehat{SOH_t^i} \right|}{SOH_t^i} * 100 \right), \tag{16}$$

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(SOH_t^i - \widehat{SOH_t^i} \right)^2, \tag{17}$$

where n is the number of test samples and SOH_t^i and $\widehat{SOH_t^i}$ are the actual and estimated SOH values, of the i^{th} test sample at the forecast target cycle t, respectively. The performance of the trained machine learning model is assessed by aggregating the performance evaluation metrics on the test dataset for each fold of the cross-validation process.

2.4. Model explanation

This study adopts SHAP to reveal the internal mechanisms of the machine learning model for estimating the SOH values of Li-ion batteries. SHAP is a unified framework that explains a machine learning model's behaviors for prediction using the Shapley value, which is the average marginal contribution of a feature value based on cooperative game theory (Lundberg & Lee, 2017; Smith & Alvarez, 2021). Specifically, SHAP assigns an importance value to each feature for each prediction based on additive feature attribution methods that comply with a set of desirable properties (i.e., local accuracy, missingness, and consistency). The additive feature attribution methods have an explanation

model g that is a linear function of binary variables, as shown in (18):.

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i z_i', \tag{18}$$

where $z' \in \{0,1\}^M$ is a coalition vector that represents whether the i^{th} feature is present (=1) or absent (=0), M denotes the number of features, and $\phi_i \in \mathbb{R}$. In this scheme, the explanation model uses simplified inputs x' that is mapped to the original inputs x by a mapping function $h_x(x') = x$, and it is built and trained to explain the original prediction model f, by ensuring $g(z') \approx f(h_x(z'))$, where $z' \approx x'$. Given that SHAP measures the importance of a feature by comparing the differences in model predictions with or without the feature, the SHAP value ϕ_i (i.e., the importance value of a feature) is calculated as a weighted average of all possible differences, as shown in (19):.

$$\phi_i = \sum_{S \subseteq \{x_1, \cdots, x_m\} \setminus \{x_i\}} \frac{|S|!(M - |S| - 1)!}{M!} [f_x(S \cup \{x_i\}) - f_x(S)], \tag{19}$$

where S is a subset of the features except the i^{th} feature, and $f_x(S \cup \{x_i\})$ and $f_x(S)$ are the predictions by the model with and without the i^{th} feature. It should be noted that the SHAP becomes an additive feature attribution method only when $\phi_0 = f_x(\varnothing)$, which indicates the baseline prediction without any feature. Depending on the properties SHAP satisfies, the prediction for each data sample by the original model equals the sum of the SHAP values of all features for the sample. Consequently, the SHAP values can be regarded as the contribution of each feature to predictions by the model.

3. Experimental settings

3.1. Capacity fade data

The capacity fade data of 379Li-ion batteries are used as the initial dataset in this study. The dataset was obtained from real-world qualification tests performed by a commercial automobile company. The capacities of the batteries were measured up to 900 cycles, which corresponds to approximately five months under accelerated stress conditions or 2–4 years in use conditions; here, the qualification tests were carried out under consistent temperature conditions for all batteries. The capacity values were recorded at various lengths for each battery sample since different batteries reached their EOL at different cycles. For example, testing on batteries that exhibited aberrantly rapid capacity-fading behaviors was stopped when they reached their EOL regardless of whether 900 cycles had elapsed, whereas testing on other batteries that exhibited gradual capacity-fading behaviors continued until the 900th cycle. Thus, the number of available data samples varies depending on the forecast target cycle as the machine learning model requires the actual SOH values of batteries at the forecast target cycle. Fig. 2 presents part of the battery samples in our dataset, which are distinguished by line style according to their capacity-fading behaviors.

3.2. Experimental condition selection

During a qualification test, the rate of SOH degradation of a Li-ion battery accelerates after an inflection point, which is usually located between 200 and 250 cycles. By conducting a series of experiments with various forecast target cycles, therefore, we examine the reliability and practicality of the proposed approach in various analysis contexts. The time required to collect the capacity fade data (i.e., the forecast origin cycle) should be carefully considered for using the machine learning model to the SOH estimation in practice. If the duration of data collection is too short, the machine learning model cannot be provided with adequate information. By contrast, if the duration of data collection is too long, the test time of the Li-ion batteries is not sufficiently reduced. Therefore, various forecast origin cycles and forecast target cycles are selected for the experimental conditions.

Given these issues, multiple experiments are conducted under various combinations of the forecast horizons, as listed in Table 1. Consequently, every machine learning model is applied to the dataset with four forecast origin cycles between 100 and 250 (i.e., between approximately a month to two months under accelerated stress condi-

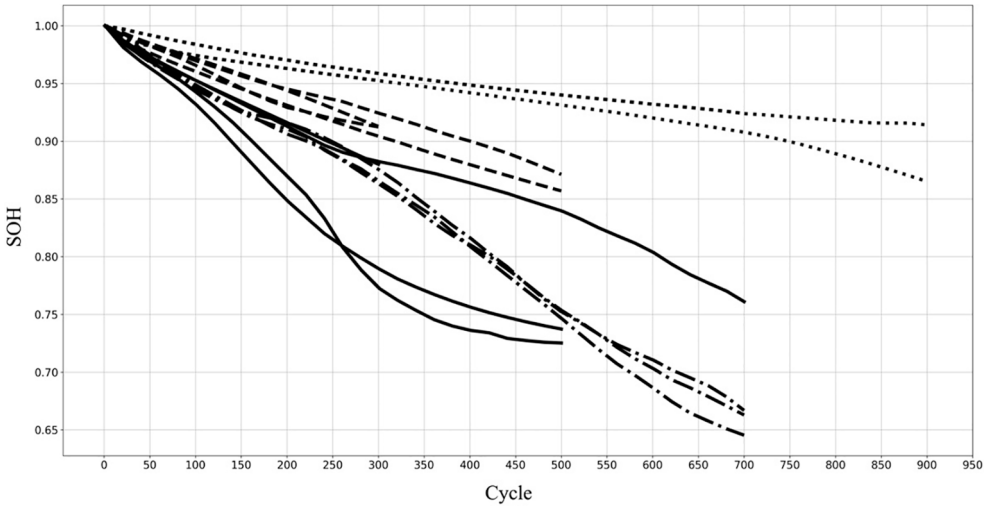

Fig. 2. Examples of the recorded SOH values for the battery samples.

Table 1
Experimental conditions.

Factor	Conditions
Forecast origin cycle	100, 150, 200, 250
Forecast target cycle	300, 500, 700
Machine learning model	RF, GBM, SVM, KNN, MLP, GRU

tions) and three forecast target cycles between 300 and 700 (i.e., between approximately two and a half months to four months under accelerated stress conditions) for the experiments. Considering that the specified forecast target cycles indicate general criteria for judging the EOL of batteries (i.e., 300th, 500th, and 700th cycle), the experiments with these conditions allow us to investigate the ability of the machine learning models to provide estimates for the batteries' life expectancy in the early phases of the qualification tests.

3.3. Data oversampling

The initial dataset is found to be imbalanced in terms of the number of available battery samples according to the SOH value at each forecast target cycle. In the dataset, most of the battery samples' SOH values are located near 0.7–0.8 at the 700th cycle, whereas battery samples exhibiting abnormal capacity fade trends (e.g., an SOH value higher than 0.9 at the 700th cycle or lower than 0.8 before the 300th cycle) are rarely observed. This disparity implies that the SOH estimation will be biased toward a specific range of SOH values.

For this reason, we employ SMOTE to complement the small sample size regarding the batteries that exhibit rarely observed capacity fade trends. SMOTE is a data oversampling method that creates new synthetic observations using linear interpolation of data samples from a minority group. The method is implemented as follows (Chawla et al., 2002): (1) a training sample p_i in the minority group is selected and its K nearest neighbors in the same group are determined, where K is a user-defined number that determines the minimal cluster size; (2) for each neighbor q_k in the set of nearest neighbors, a new sample is created based on a linear interpolation between p_i and q_k; and finally, (3) the above steps are repeated for each sample in the training dataset. Through data oversampling, the developed approach reflects various types of capacity-fading behaviors of Li-ion batteries.

To apply SMOTE, the dataset is divided into six subsets based on the range of the SOH value at each forecast target cycle, as follows: [0,0.7), [0.7,0.75), [0.75,0.8), [0.8,0.85), [0.85,0.9), and [0.9,1.0]. Then, SMOTE is performed to generate synthetic samples for the minority subsets that contain a relatively small number of original samples. Here, we set the minimum number of data samples in each subset, including newly created samples, to 20 to avoid the overfitting problems caused by generating too many synthetic samples. Table 2 reports the number of available data samples for each forecast target cycle before and after conducting data oversampling.

Table 2
Number of available data samples for each forecast target cycle.

Forecast target	Before oversampling	After oversampling
300	379	401
500	214	237
700	111	151

3.4. Hyperparameter selection

The hyperparameters used for the proposed approach should be carefully searched to achieve the best predictive performance by each machine learning model. To find the optimal set of hyperparameters for each machine learning model, pilot experiments of SOH estimation by all machine learning models are conducted under a specific experimental condition: the 100th cycle as the forecast origin cycle and the 700th as the forecast target cycle. The pilot experiments are designed to investigate the predictive performance of each machine learning model with various combinations of hyperparameter values picked from a specified hyperparameter set based on a random search algorithm; here, we use Optuna (Akiba et al., 2019), which is an open-source hyperparameter optimization framework, to conduct the experiments automatically. Optuna performs 10 trials using different combinations of hyperparameters to identify the optimal set of hyperparameters for each machine learning model that produces results with the smallest MAPE. Table 3 presents the hyperparameter sets used for the pilot experiments for each machine learning model.

Based on the performance evaluation metrics obtained from the pilot experiments, the optimal hyperparameters for each model were determined as follows: 10 trees and $n_{features}$ as the maximum number of features for the RF; 500 trees, $n_{features}$ as the maximum number of features, and 0.04 learning rate for the GBM; the radial basis function kernel and 1.0 as the regularization parameter for the SVM; seven neighbors for the KNN; three hidden layers with 256, 128, and 64 neurons, a batch size of 32, and a 0.0005 learning rate for the MLP; and one hidden layer with 64 neurons, a batch size of 32, and a 0.0003 learning rate for the GRU.

Table 3
Specified hyperparameter set for each machine learning model.

Machine learning model	Hyperparameter	Category/Range
RF	Number of trees	{10, 50, 100, 500}
	Maximum number of features	$\{\sqrt{n_{features}}, \log_2(n_{features}), n_{features}\}$
GBM	Number of trees	{10, 50, 100, 500}
	Maximum number of features	$\{\sqrt{n_{features}}, \log_2(n_{features}), n_{features}\}$
	Learning rate	[0.01, 0.1)
SVM	Kernel	{"linear", "poly", "rbf" }
	Regularization parameter	(0, 1]
KNN	Number of neighbors	[2, 10]
MLP	Number of hidden layers	[2, 6]
	Number of neurons	$\left\{\frac{256}{2^{i-1}} \mid i = 1, 2, \cdots, n_{layers}\right\}$
	Learning rate	[0.0001, 0.01)
	Batch size	{16, 32, 64}
GRU	Number of hidden layers	{1, 2, 3}
	Number of neurons	{16, 32, 64}
	Learning rate	[0.0001, 0.01)
	Batch size	{16, 32, 64}

4. Results

4.1. SOH estimation

We applied the developed approach to our dataset by varying the forecast origin cycles and forecast target cycles for the chosen experimental conditions. First, for each dataset, the three types of statistical features were extracted from the capacity fade data samples by moving windows of ten cycles from the beginning of the data to the forecast origin cycle. As the proposed approach employed three statistical features, $3*(forecastorigincycle-10)$ input features were used for SOH estimation for each data sample. For example, if the forecast origin cycle was set to the 100th cycle, $3*90 = 270$ features were obtained. Second, the dataset was divided into five groups based on the stratified five-fold cross-validation scheme. Third, SMOTE was performed only for the training set to remove the effects of oversampling from the performance evaluation. Finally, the machine learning models were trained using the extracted statistical features and the actual SOH values from the training sets, and then the models estimated the SOH values of the battery samples in the test sets. As an example, Table 4 presents the SOH estimation results from RF after setting the 100th cycle as the forecast origin cycle and the 700th cycle as the forecast target cycle.

We examined the performance of the proposed approach using the MAPE and MSE as performance evaluation metrics, as reported in Table 5. The performance evaluation metrics for the various forecast horizons confirm that the proposed approach can capture future SOH degradation trends using the data collected in the early phases of qualification tests. It is noteworthy that the proposed approach is robust across various contexts in terms of the period of estimation considering that an MAPE under 10% can be interpreted as "highly accurate predictions" (Lewis, 1982). Moreover, all of the models achieved an MSE of less than 0.02, which indicates that the proposed approach can estimate the actual SOH value with little deviation considering the range of SOH value is between 0 and 1. Overall, there is a clear tendency: the shorter the distance between the forecast origin cycle and forecast target cycle, the better the predictive performance of the machine learning model. This result occurs because the shorter forecast horizons provide the machine learning models with information regarding more recent trends in the batteries' SOH values. It should be noted that some machine learning models produced their best performance for estimating the SOH values at 700th cycle when the forecast origin cycle was set to the 150th cycle. This result supports our contention that machine learning models are effective in estimating the future SOH values of Li-ion batteries in the early phases of qualification tests, given that the inflection points are usually located between 200 and 250 cycles.

In terms of the comparison among the machine learning models, the

Table 4
A sample of the SOH estimation results based on RF (forecast origin = 100th cycle and forecast target = 700th cycle).

Battery Sample	Extracted statistical features (input) …	*movAVG*(*t*)	…	*movFOD*(*t*)	…	*movVAR*(*t*)	…	Estimated SOH	Actual SOH
M-1-1		0.9903		0.000617		0.000003		0.7398	0.7610
M-1-2		0.9891		0.000649		0.0000034		0.7715	0.7814
M-1-3	…	0.9912	…	0.000641	…	0.0000034	…	0.7255	0.7067
M-1-5		0.996		0.000391		0.0000011		0.8142	0.7930
M-1-6		0.9956		0.000367		0.0000011		0.8040	0.8007
…		…		…		…		…	…

Table 5
Performance evaluation metrics for machine learning models.

(a) MAPE results of the performance evaluation

Forecast origin	Forecast target	Machine learning model RF	GBM	SVM	KNN	MLP	GRU
100	300	2.26%	**2.12%**	5.57%	2.66%	3.09%	2.99%
	500	**3.64%**	3.66%	5.27%	4.19%	4.29%	4.97%
	700	**4.52%**	4.63%	6.29%	6.19%	13.16%	5.96%
150	300	1.89%	**1.83%**	5.76%	2.01%	2.28%	2.09%
	500	**3.01%**	3.08%	5.14%	3.64%	3.71%	4.37%
	700	3.99%	**3.77%**	6.01%	5.49%	11.5%	6.34%
200	300	1.30%	**1.28%**	6.67%	1.70%	1.88%	1.51%
	500	**3.00%**	3.09%	4.88%	3.38%	3.53%	3.75%
	700	**3.80%**	3.85%	5.73%	5.11%	6.19%	5.91%
250	300	0.88%	**0.65%**	6.92%	1.18%	1.68%	0.85%
	500	2.59%	**2.46%**	4.94%	3.16%	3.49%	3.38%
	700	**4.09%**	4.13%	5.63%	4.57%	5.77%	5.67%

(b) MSE results of the performance evaluation

Forecast origin	Forecast target	Machine learning model RF	GBM	SVM	KNN	MLP	GRU
100	300	**0.0009**	**0.0009**	0.0032	0.0012	0.0017	0.0014
	500	**0.0014**	0.0015	0.0028	0.0018	0.0019	0.0025
	700	**0.0019**	0.0020	0.0034	0.0035	0.0183	0.0031
150	300	**0.0007**	**0.0007**	0.0033	**0.0007**	0.0009	**0.0007**
	500	**0.0010**	0.0011	0.0027	0.0013	0.0015	0.0019
	700	0.0016	**0.0014**	0.0031	0.0027	0.0147	0.0034
200	300	**0.0003**	**0.0003**	0.0040	0.0005	0.0006	**0.0003**
	500	**0.0010**	0.0011	0.0024	0.0012	0.0016	0.0015
	700	**0.0014**	**0.0014**	0.0029	0.0024	0.0032	0.0032
250	300	0.0002	**0.0001**	0.0042	0.0003	0.0004	0.0002
	500	0.0008	**0.0007**	0.0025	0.0011	0.0014	0.0012
	700	**0.0016**	**0.0016**	0.0027	0.0020	0.0035	0.0027

Table 6
Distribution of MAPE and MSE from RF for each SOH range (forecast origin = 100th cycle and forecast target = 700th cycle).

Range of actual SOH value at the forecast target cycle	MAPE	MSE	# of samples in the range
0.90–	0.27%	0.0000	2
0.85–0.90	4.06%	0.0020	12
0.80–0.85	3.25%	0.0009	22
0.75–0.80	3.81%	0.0013	43
0.70–0.75	4.65%	0.0017	26
–0.70	8.08%	0.0049	6

ensemble models (i.e., RF and GBM) outperformed the others and achieved less than 5% MAPE and 0.002 MSE under all experimental conditions. Indeed GRU, which has advantages in managing many features and long sequences with its deep architecture and has exhibited outstanding performance on time-series problems in recent years, is likely the most advanced model among the employed machine learning models. However, our experimental results demonstrate that the ensemble models achieve better predictive performance than GRU for the SOH estimation. This result occurs because the deep architectures of GRU have a much larger number of parameters than others, which may be inefficient for this univariate time-series prediction task in terms of the model complexity. It is also noteworthy that RF performed well for a relatively shorter forecast origin cycle (i.e., the 100th), whereas GBM performed better for a longer forecast origin cycle (i.e., the 250th). Therefore, considering that the proposed approach focuses on reducing the time required for qualification tests on Li-ion batteries by estimating the future SOH values, RF was identified as the best-performing model.

To provide a more detailed analysis, changes in the MAPE and MSE depending on the actual SOH values of each battery sample were examined. Table 6 presents the distribution of the MAPE and MSE produced by RF for each range of actual SOH values, with the 100th cycle as the forecast origin cycle and the 700th cycle as the forecast target cycle. The results demonstrate that RF showed a high predictive performance over all ranges of actual SOH values as it achieved less than 10% MAPE and 0.005 MSE even in the worst case. These outcomes show that the proposed approach can serve as a practical method for reducing the duration of qualification tests on Li-ion batteries.

4.2. Model interpretation

To understand the behaviors and predictions of the best-performing model, SHAP was applied to the SOH estimation results from RF. However, calculating the exact SHAP values is challenging because of its exponential computational cost, which occurs because it accounts for every possible subset of features. For this reason, we used Tree SHAP, a variation of SHAP for tree-based models, because it reduces the computational complexity considerably by tracking the subsets of features that flow into each leaf of the tree (Lundberg et al., 2018, 2020). We obtained the SHAP values for each statistical feature at each charging and discharging cycle for every battery sample.

The explanations based on the SHAP values allow us to examine the statistical features and cycles that are important when estimating the SOH values of batteries as well as provide insight into how the qualification testing procedure can be improved. The model's predictions were interpreted at the local and global levels. First, for each battery sample, the contribution of each statistical feature to the SOH estimation results was visualized by the SHAP values at the local level. Fig. 3 provides explanations for the SOH estimation results from RF trained with the 100th cycle as the forecasting origin cycle and the 700th cycle as the forecasting target cycle. In the figure, the SHAP values of the features that increase the estimation are represented by a red bar, while those that decrease the estimation are represented by a blue bar. In addition, the base value indicates the model's baseline prediction, which is the mean value of all predictions. For example, for the battery sample 'M-1-7', whose SOH at the 700th cycle (i.e., the forecasting target cycle) was 0.7905, RF estimated its SOH as 0.7914, and the feature *movFOD* at the 56th and 89th cycles contributed most to the estimated value by the model. Specifically, the first-order difference in SOH values at the 56th cycle causes the model to estimate that the sample's SOH is higher than that of the baseline prediction, while the first-order difference at the 89th cycle causes the estimation to be lower than the baseline prediction. For another battery sample 'T1-1-13_1′, whose SOH at the 700th cycle is 0.6629, RF estimated its SOH as 0.6615, and the statistical features at the majority of cycles influence the model in the direction of lowering the estimated value. Similar to the above example, the *movFOD* at the 56th and 89th cycles contributed most to the estimated value by the model. However, in this case, the *movFOD* at the 56th cycle drives the estimated value to be decreased. This observation indicates that certain features at certain cycles are more important than others.

We then investigated how the features affect the model's predictions based on the trends of their SHAP values at the global level. In particular, we compared the SHAP values of the features across all experimental conditions. Fig. 4 presents the distributions of the SHAP values of the top ten statistical features sorted by their absolute SHAP values. In the plots, the data samples are represented as colored dots from blue to red in order of the original values of the statistical features. Overall, there exists a clear tendency: the lower the values of *movFOD* and *movVAR*, the higher the estimated SOH values, while the opposite tendency exists for *movAVG*. In addition, for the large forecast horizons (i.e., the difference between the forecast origin cycle and forecast target

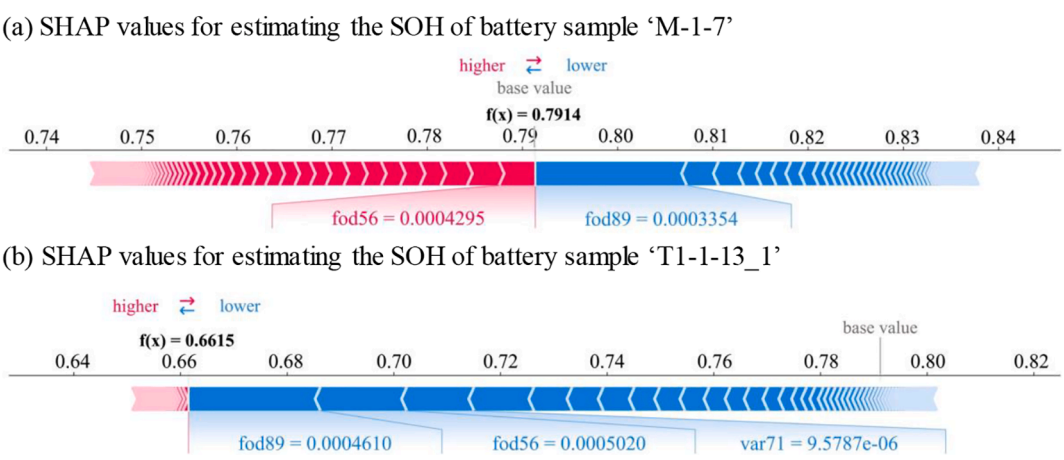

Fig. 3. Examples of local explanation for SOH estimation by RF (forecast origin = 100th cycle and forecast target = 700th cycle). (a) SHAP values for estimating the SOH of battery sample 'M-1-7'. (b) SHAP values for estimating the SOH of battery sample 'T1-1-13_1′.

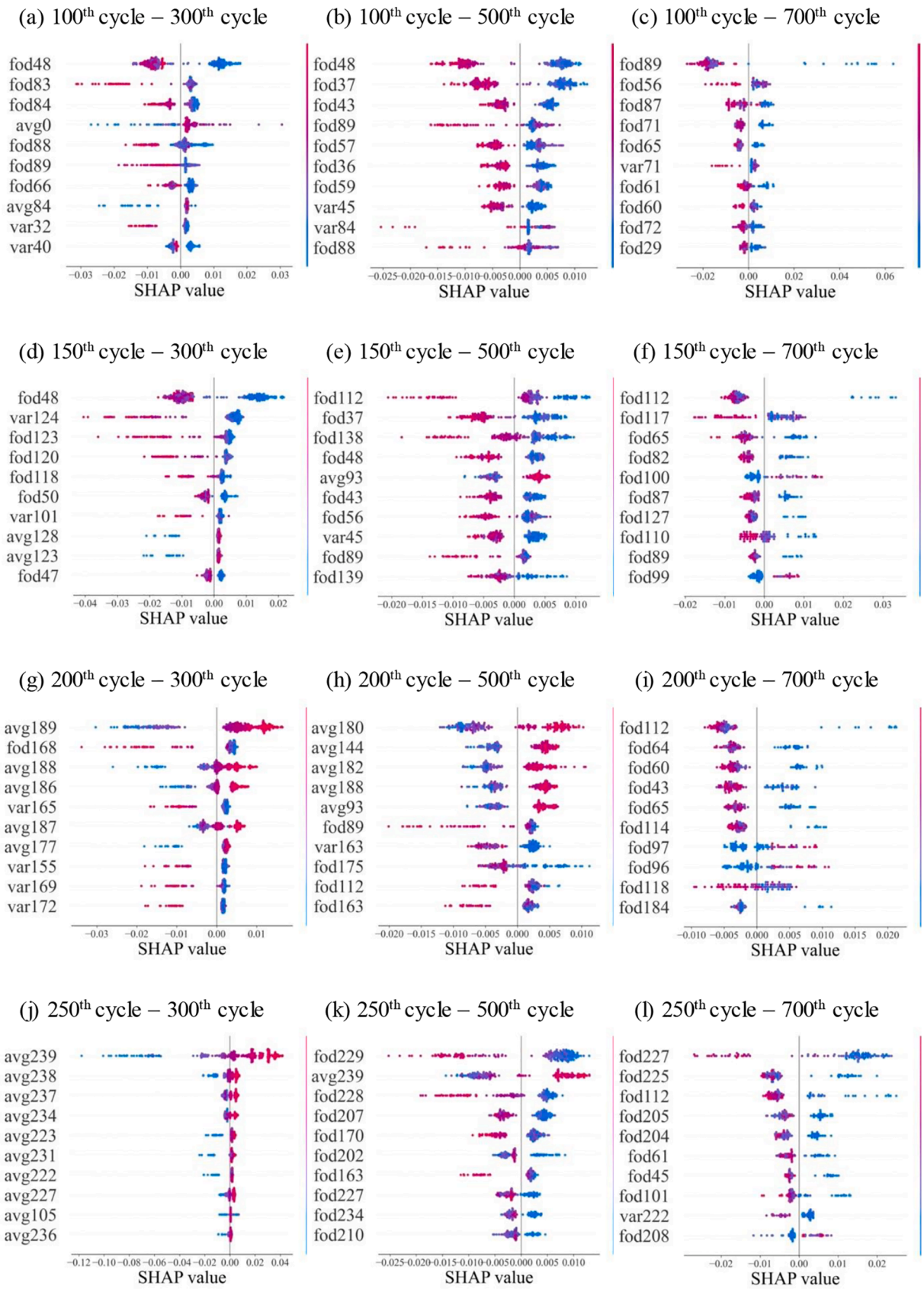

Fig. 4. Examples of the global explanations for SOH estimation by RF. (a) 100th cycle – 300th cycle. (b) 100th cycle – 500th cycle. (c) 100th cycle – 700th cycle. (d) 150th cycle – 300th cycle. (e) 150th cycle – 500th cycle. (f) 150th cycle – 700th cycle. (g) 200th cycle – 300th cycle. (h) 200th cycle – 500th cycle.(i) 200th cycle – 700th cycle. (j) 250th cycle – 300th cycle. (k) 250th cycle – 500th cycle. (l) 250th cycle – 700th cycle.

Table 7
Mean absolute SHAP values for every ten cycles (forecast target = 700th cycle).

Cycles	100 → 700			150 → 700			200 → 700			250 → 700		
	movAVG	*movFOD*	*movVAR*	*movAVG*	*movFOD*	*movVAR*	*movAVG*	*movFOD*	*movVAR*	*movAVG*	*movFOD*	*movVAR*
0–9	**0.000149**	0.000250	0.000134	**0.000335**	0.000158	0.000087	**0.000189**	0.000111	0.000073	0.000147	0.000146	0.000047
10–19	0.000065	0.000293	**0.000244**	0.000046	0.000093	0.000074	0.000038	0.000124	0.000117	0.000015	0.000094	0.000044
20–29	**0.000080**	0.000549	0.000116	0.000046	0.000177	0.000084	0.000020	0.000191	0.000052	0.000022	0.000090	0.000041
30–39	0.000036	0.000289	0.000091	0.000025	0.000182	0.000076	0.000027	0.000232	0.000045	0.000023	0.000099	0.000046
40–49	0.000034	0.000414	0.000090	0.000034	0.001064	0.000058	0.000013	0.000563	0.000028	0.000015	0.000241	0.000060
50–59	0.000052	0.000383	0.000118	0.000029	0.000202	0.000137	0.000017	0.000131	0.000062	0.000016	0.000074	0.000047
60–69	0.000043	**0.000900**	0.000138	0.000025	**0.001203**	0.000075	0.000030	0.000361	0.000054	0.000014	0.000379	0.000018
70–79	0.000054	0.000887	**0.000188**	0.000031	0.000471	**0.000152**	0.000036	0.000323	0.000174	0.000066	0.000256	0.000037
80–89	0.000079	**0.002824**	0.000099	0.000024	0.000310	0.000052	0.000023	0.000358	0.000043	0.000019	0.000153	0.000029
90–99				0.000027	0.000890	**0.000289**	0.000017	0.000507	**0.000254**	0.000021	0.000239	**0.000094**
100–109				0.000030	0.000707	0.000141	0.000027	**0.001286**	0.000094	0.000022	0.000198	0.000054
110–119				0.000034	**0.001098**	0.000100	0.000027	0.000441	0.000100	0.000019	0.000389	0.000034
120–129				0.000027	0.000491	0.000036	0.000024	0.000193	0.000071	0.000037	0.000137	0.000025
130–139				**0.000057**	0.000403	0.000064	0.000029	0.000319	0.000135	0.000021	0.000072	0.000031
140–149							0.000047	0.000071	0.000014	0.000020	0.000049	0.000056
150–159							0.000031	0.000082	0.000021	0.000012	0.000066	0.000021
160–169							0.000050	0.000118	0.000083	0.000033	0.000129	0.000031
170–179							0.000031	**0.001078**	**0.000178**	0.000024	0.000248	0.000035
180–189							**0.000279**	0.000526	0.000088	0.000009	0.000397	0.000055
190–199										0.000175	0.000148	**0.000094**
200–209										0.000131	**0.000926**	0.000055
210–219										**0.000213**	0.000277	0.000038
220–229										**0.000369**	**0.002105**	0.000045

cycle), *movAVG* at any cycle does not have a greater impact than other features. This result indicates that the changes in SOH values are more important than the SOH values themselves. In terms of the number of charging and discharging cycles, for higher forecast target cycles, the statistical features at earlier cycles have more impact than those at later cycles, as described in Fig. 4(g)-(i). For example, when the forecast origin cycle is set to the 200th cycle and the forecast target cycle is set to the 700th cycle, the SHAP values for the features before 100 cycles are higher than those after 100 cycles. On the contrary, when the forecast target cycle is set to the 300th or 500th cycle, the features before 100 cycles are more included in the top ten features than those after 100 cycles.

To provide a more detailed examination of the tendencies, we obtained average (absolute) SHAP values for each statistical feature for every ten cycles, as reported in Table 7. In the table, the average SHAP values were calculated for the SOH estimation results by setting the forecasting target cycle to the 700th. The values marked in bold represent the first and second highest average SHAP values for each statistical feature. Accordingly, the statistical features between 60 and 100 cycles were found to be important in SOH estimation. In particular, the *movFOD* near the 100th cycle presents higher SHAP values than others. This result demonstrates that the SOH degradation trends near the 100th cycle have a considerable impact on future SOH values. However, when the forecast origin cycle is set to the 250th cycle, the statistical features near the 200th cycle exhibit high SHAP values. These findings are consistent with the results of prior studies, which indicated that the inflection point of the SOH degradation trend is usually located between 200 and 250 cycles. From the analysis results, we can conclude that the qualification tests on Li-ion batteries should be conducted at least up to 100 cycles since the capacity-fading behavior near the 100th cycle was proved the most significant when estimating the future SOH values.

5. Conclusions

This study developed an interpretable machine learning model for estimating future SOH values of Li-ion batteries in the early phases of qualification tests. The developed approach showed that it can reduce the time required for qualification tests to 100 cycles (i.e., before the inflection point is observed), with less than 5% MAPE and 0.002 MSE. Moreover, the results of model interpretation using SHAP demonstrated that the moving first-order difference and moving variance calculated from the ten SOH values around the 70th cycle during qualification tests are significant factors when estimating the SOH values of Li-ion batteries. Therefore, the developed approach was expected to serve as a starting point for developing more effective qualification testing procedures.

While many machine learning-based SOH estimation methods often require other variables related to the battery such as state of charge, voltage, temperature, and internal resistances, the developed approach used only the capacity fade data from the batteries to forecast future SOH values and provide explanations for the prediction process. The developed approach demonstrated the potential to reduce the time and efforts required to obtain suitable datasets for model training, and it was also expected to be generalized and deployed to any estimation tasks using univariate time-series data in practice.

However, the present study had certain limitations that should be addressed in future research. First, the performance of the developed approach may be improved by adopting more advanced deep learning models such as convolutional neural networks, autoregressive models based on LSTM, and graph neural networks. Second, the predictive performance of the developed approach for SOH estimation was not tested under various temperature conditions. This limitation indicates that the applicability of the developed approach was not fully examined to be deployed in practice where temperature has a significant influence on the SOH of Li-ion batteries. It is, therefore, necessary to conduct experiments using additional data collected under various temperature conditions in future research. Third, the developed approach directly used the features extracted from the SOH values' statistical measures; however, there exist other advanced methods that can generate more informative features for machine learning models from the univariate time-series data such as recurrence plots (Eckmann et al., 1987) and the Markov transition matrix (Aalen & Johansen, 1978). Finally, sensitivity

analyses are needed to properly set the user-defined parameters used in the proposed approach, such as the size of the window, which are used to extract statistical features as well as the number of samples that will be generated by the data oversampling technique.

CRediT authorship contribution statement

Gyumin Lee: Software, Validation, Formal analysis, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Juram Kim:** Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Changyong Lee:** Conceptualization, Methodology, Investigation, Supervision, Writing – review & editing, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Aalen, O. O., & Johansen, S. (1978). An empirical transition matrix for non-homogeneous Markov chains based on censored observations. *Scandinavian Journal of Statistics, 5*(3), 141–150.

Adadi, A., & Berrada, M. (2018). Peeking Inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2623–2631). https://doi.org/10.1145/3292500.3330701

Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician, 46*(3), 175. https://doi.org/10.2307/2685209

Ayvaz, S., & Alpay, K. (2021). Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time. *Expert Systems with Applications*, *173*(September 2020), 114598. doi: 10.1016/j.eswa.2021.114598.

Barbarisi, O., Vasca, F., & Glielmo, L. (2006). State of charge Kalman filter estimator for automotive batteries. *Control Engineering Practice, 14*(3 SPEC. ISS.), 267–275. https://doi.org/10.1016/j.conengprac.2005.03.027

Cannarella, J., & Arnold, C. B. (2014). State of health and charge measurements in lithium-ion batteries using mechanical stress. *Journal of Power Sources, 269*, 7–14. https://doi.org/10.1016/j.jpowsour.2014.07.003

Chaoui, H., & Ibe-Ekeocha, C. C. (2017). State of charge and state of health estimation for lithium batteries using recurrent neural networks. *IEEE Transactions on Vehicular Technology, 66*(10), 8773–8783. https://doi.org/10.1109/TVT.2017.2715333

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16*(2), 321–357. https://doi.org/10.1613/jair.953

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, 1724–1734. doi: 10.3115/v1/d14-1179.

de Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean Absolute Percentage Error for regression models. *Neurocomputing, 192*, 38–48. https://doi.org/10.1016/j.neucom.2015.12.114

Eckmann, J.-P., Kamphorst, S. O., & Ruelle, D. (1987). Recurrence plots of dynamical systems. *Europhysics Letters (EPL), 4*(9), 973–977. https://doi.org/10.1209/0295-5075/4/9/004

Friedman, J. (2001). Friedman, J., 2001. Greedy function approximation : A gradient boosting machine. The Annals of Statistics 29, 1189-1232. *The Annals of Statistics*, *29*(5), 1189–1232.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction. In *Springer Science & Business Media.* Springer Science & Business Media. doi: 10.1007/b94608.

He, W., Williard, N., Osterman, M., & Pecht, M. (2011). Prognostics of lithium-ion batteries based on Dempster-Shafer theory and the Bayesian Monte Carlo method. *Journal of Power Sources, 196*(23), 10314–10321. https://doi.org/10.1016/j.jpowsour.2011.08.040

Hendricks, C., Williard, N., Mathew, S., & Pecht, M. (2015). A failure modes, mechanisms, and effects analysis (FMMEA) of lithium-ion batteries. *Journal of Power Sources, 297*, 113–120. https://doi.org/10.1016/j.jpowsour.2015.07.100

Ho, T. K. (1995). Random decision forests. *Proceedings of 3rd International Conference on Document Analysis and Recognition*, 278–282. https://ieeexplore.ieee.org/abstract/document/598994/.

Huang, L., Zhang, Z., Wang, Z., Zhang, L., Zhu, X., & Dorrell, D. D. (2019). Thermal runaway behavior during overcharge for large-format Lithium-ion batteries with different packaging patterns. *Journal of Energy Storage, 25*(March), Article 100811. https://doi.org/10.1016/j.est.2019.100811

Johnson, B. A., & White, R. E. (1998). Characterization of commercially available lithium-ion batteries. *Journal of Power Sources, 70*(1), 48–54. https://doi.org/10.1016/S0378-7753(97)02659-1

Junping, W., Jingang, G., & Lei, D. (2009). An adaptive Kalman filtering based State of Charge combined estimator for electric vehicle battery pack. *Energy Conversion and Management, 50*(12), 3182–3186. https://doi.org/10.1016/j.enconman.2009.08.015

Le, D., & Tang, X. (2011). Lithium-ion battery state of health estimation using Ah-V characterization. *Proceedings of the Annual Conference of the Prognostics and Health Management Society 2011, PHM 2011*, 362–367. doi: 10.36001/phmconf.2011.v3i1.2073.

Lee, C., Jo, S., Kwon, D., & Pecht, M. G. (2020). Capacity-fading behavior analysis for early detection of unhealthy Li-Ion batteries. *IEEE Transactions on Industrial Electronics, 68*(3), 2659–2666. https://doi.org/10.1109/tie.2020.2972468

Lee, J., Kwon, D., & Pecht, M. G. (2019). Reduction of Li-ion battery qualification time based on prognostics and health management. *IEEE Transactions on Industrial Electronics, 66*(9), 7310–7315. https://doi.org/10.1109/TIE.2018.2880701

Lewis, C. D. (1982). *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting.* Butterworth Scientific.

Li, P., Zhang, Z., Xiong, Q., Ding, B., Hou, J., Luo, D., … Li, S. (2020). State-of-health estimation and remaining useful life prediction for the lithium-ion battery based on a variant long short term memory neural network. *Journal of Power Sources, 459*(March), Article 228069. https://doi.org/10.1016/j.jpowsour.2020.228069

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., … Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence, 2*(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9

Lundberg, S. M., Erion, G. G., & Lee, S.-I. (2018). Consistent Individualized Feature Attribution for Tree Ensembles. *ArXiv Preprint*, *2*. http://arxiv.org/abs/1802.03888.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *2017-Decem*(Section 2), 4766–4775.

Ma, C., Zhai, X., Wang, Z., Tian, M., Yu, Q., Liu, L., … Yang, X. (2019). State of health prediction for lithium-ion batteries using multiple-view feature fusion and support vector regression ensemble. *International Journal of Machine Learning and Cybernetics, 10*(9), 2269–2282. https://doi.org/10.1007/s13042-018-0865-y

Maleki, N., Zeinali, Y., & Niaki, S. T. A. (2021). A k-NN method for lung cancer prognosis with the use of a genetic algorithm for feature selection. *Expert Systems with Applications*, *164*(September 2020), 113981. doi: 10.1016/j.eswa.2020.113981.

Matsushima, T. (2009). Deterioration estimation of lithium-ion cells in direct current power supply systems and characteristics of 400-Ah lithium-ion cells. *Journal of Power Sources, 189*(1), 847–854. https://doi.org/10.1016/j.jpowsour.2008.08.023

Moscatelli, M., Parlapiano, F., Narizzano, S., & Viggiano, G. (2020). Corporate default forecasting with machine learning. *Expert Systems with Applications, 161*, Article 113567. https://doi.org/10.1016/j.eswa.2020.113567

Niu, T., Wang, J., Lu, H., Yang, W., & Du, P. (2020). Developing a deep learning framework with two-stage feature selection for multivariate financial time series forecasting. *Expert Systems with Applications, 148*, Article 113237. https://doi.org/10.1016/j.eswa.2020.113237

Olivares, B. E., Cerda Muñoz, M. A., Orchard, M. E., & Silva, J. F. (2013). Particle-filtering-based prognosis framework for energy storage devices with a statistical characterization of state-of-health regeneration phenomena. *IEEE Transactions on Instrumentation and Measurement, 62*(2), 364–376. https://doi.org/10.1109/TIM.2012.2215142

Paliwal, M., & Kumar, U. A. (2009). Neural networks and statistical techniques: A review of applications. *Expert Systems with Applications, 36*(1), 2–17. https://doi.org/10.1016/j.eswa.2007.10.005

Pecht, M., Shibutani, T., Kang, M., Hodkiewicz, M., & Cripps, E. (2016). A fusion prognostics-based qualification test methodology for microelectronic products. *Microelectronics Reliability, 63*, 320–324. https://doi.org/10.1016/j.microrel.2016.04.002

Puchalsky, W., Ribeiro, G. T., da Veiga, C. P., Freire, R. Z., & dos Santos Coelho, L. (2018). Agribusiness time series forecasting using Wavelet neural networks and metaheuristic optimization: An analysis of the soybean sack price and perishable products demand. *International Journal of Production Economics, 203*(June), 174–189. https://doi.org/10.1016/j.ijpe.2018.06.010

Razavi-Far, R., Chakrabarti, S., Saif, M., & Zio, E. (2019). An integrated imputation-prediction scheme for prognostics of battery data with missing observations. *Expert Systems with Applications, 115*, 709–723. https://doi.org/10.1016/j.eswa.2018.08.033

Saxena, S., Kang, M., Xing, Y., & Pecht, M. (2018). Anomaly detection during lithium-ion battery qualification testing. In *2018 IEEE International Conference on Prognostics and Health Management.* https://doi.org/10.1109/ICPHM.2018.8448735

Saxena, S., Xing, Y., Kwon, D., & Pecht, M. (2019). Accelerated degradation model for C-rate loading of lithium-ion batteries. *International Journal of Electrical Power and Energy Systems*, *107*(January 2018), 438–445. doi: 10.1016/j.ijepes.2018.12.016.

Smith, M., & Alvarez, F. (2021). Identifying mortality factors from Machine Learning using Shapley values – a case of COVID19. *Expert Systems with Applications*, *176*(June 2020), 114832. doi: 10.1016/j.eswa.2021.114832.

Spotnitz, R. (2003). Simulation of capacity fade in lithium-ion batteries. *Journal of Power Sources, 113*(1), 72–80. https://doi.org/10.1016/S0378-7753(02)00490-1

Sun, F., Hu, X., Zou, Y., & Li, S. (2011). Adaptive unscented Kalman filtering for state of charge estimation of a lithium-ion battery for electric vehicles. *Energy, 36*(5), 3531–3540. https://doi.org/10.1016/j.energy.2011.03.059

Vetter, J., Novák, P., Wagner, M. R., Veit, C., Möller, K. C., Besenhard, J. O., … Hammouche, A. (2005). Ageing mechanisms in lithium-ion batteries. *Journal of Power Sources, 147*(1–2), 269–281. https://doi.org/10.1016/j.jpowsour.2005.01.006

Wang, Q., Wang, Z., Zhang, L., Liu, P., & Zhang, Z. (2021). A novel consistency evaluation method for series-connected battery systems based on real-world operation data. *IEEE Transactions on Transportation Electrification, 7*(2), 437–451. https://doi.org/10.1109/TTE.2020.3018143

Widodo, A., Shim, M. C., Caesarendra, W., & Yang, B. S. (2011). Intelligent prognostics for battery health monitoring based on sample entropy. *Expert Systems with Applications, 38*(9), 11763–11769. https://doi.org/10.1016/j.eswa.2011.03.063

Xiong, R., Zhang, Y., He, H., Zhou, X., & Pecht, M. G. (2017). A double-scale, particle-filtering, energy state prediction algorithm for lithium-ion batteries. *IEEE Transactions on Industrial Electronics, 65*(2), 1526–1538. https://doi.org/10.1109/TIE.2017.2733475

Yu, Z., Huai, R., & Xiao, L. (2015). State-of-charge estimation for lithium-ion batteries using a Kalman filter based on local linearization. *Energies, 8*(8), 7854–7873. https://doi.org/10.3390/en8087854

Zhang, F., Liu, G., & Fang, L. (2009). Battery state estimation using Unscented Kalman Filter. *2009 IEEE International Conference on Robotics and Automation*, *3*, 1863–1868. doi: 10.1109/ROBOT.2009.5152745.

Zhang, J., & Lee, J. (2011). A review on prognostics and health monitoring of Li-ion battery. *Journal of Power Sources, 196*(15), 6007–6014. https://doi.org/10.1016/j.jpowsour.2011.03.101

Zhang, L., Fan, W., Wang, Z., Li, W., & Sauer, D. U. (2020). Battery heating for lithium-ion batteries based on multi-stage alternative currents. *Journal of Energy Storage, 32* (February), Article 101885. https://doi.org/10.1016/j.est.2020.101885

Zhang, L., Hu, X., Wang, Z., Ruan, J., Ma, C., Song, Z., Dorrell, D. G., & Pecht, M. G. (2021). Hybrid electrochemical energy storage systems: An overview for smart grid and electrified vehicle applications. *Renewable and Sustainable Energy Reviews*, *139* (October 2020), 110581. doi: 10.1016/j.rser.2020.110581.