Interpretable Battery Cycle Life Range Prediction Using Early Cell Degradation Data

Huang Zhang, Yang Su, Faisal Altaf, Torsten Wik, *Member, IEEE*, and Sébastien Gros

***Abstract*— Battery cycle life prediction using early degradation data has many potential applications throughout the battery product life cycle. For that reason, various data-driven methods have been proposed for point prediction of battery cycle life with minimum knowledge of the battery degradation mechanisms. However, managing the rapidly increasing amounts of batteries at end of life with lower economic and technical risk requires prediction of cycle life with quantified uncertainty, which is still lacking. The interpretability (i.e., the reason for high prediction accuracy) of these advanced data-driven methods is also worthy of investigation. Here, a physics-informed quantile regression forest (QRF) model, having the advantage of not assuming any specific distribution of cycle life, is introduced to make cycle life range prediction with uncertainty quantified as the width of the prediction interval (PI), in addition to point predictions with high accuracy. The hyperparameters of the QRF model are optimized with a proposed alpha-logistic-weighted criterion so that the coverage probabilities associated with the PIs are calibrated. The interpretability of the final QRF model is explored with two global model-agnostic methods, namely, permutation importance and partial dependence plot.**

***Index Terms*— Cycle life early prediction, interpretable machine learning, lithium-ion (Li-ion) battery, prediction interval (PI), quantile regression forest (QRF).**

I. Introduction

LITHIUM-ION (Li-ion) batteries have become the main choice of energy storage in electric vehicles (EVs), with a rapidly growing market that is spurred by governmental policies and subsidies with the aim of enhancing energy sustainability and carbon emission reduction [1], [2]. However, Li-ion batteries degrade with time due to both calendar aging and cyclic aging, which leads to a deterioration of their performance [3]. Understanding these aging processes and providing a reliable cycle life prediction of Li-ion batteries based on early degradation data would enable many new possibilities throughout the battery life. We give five examples of such possibilities here. First, the total driven distance of an EV battery can be translated into a number of equivalent full cycles [2]. Reliable cycle life prediction of Li-ion batteries using early degradation data would facilitate automotive companies to quickly adjust their warranty policy for new batches of Li-ion batteries from suppliers while greatly reducing time and cost of long aging experiments. Second, warranty and pricing based on prediction of cycle life in a battery second life application as energy storage strongly affect how the second life battery market will evolve in the future, and reducing the uncertainty associated with cycle life prediction will reduce the cost of battery deployment [4]. Third, accurate and reliable cycle life prediction with high accuracy also facilitates predictive maintenance by reducing the sudden failure rate and the maintenance costs of battery-based applications [5]. Fourth, an early prediction model can also be combined with a design parameter optimization algorithm to identify high-cycle-life charging protocols [6]. Finally, accurate prediction of the battery life with early degradation data is of crucial importance for improving the battery development and manufacturing processes [7].

Unfortunately, accurate battery cycle life early prediction using relatively little degradation data that cover a limited range of lifetime is challenging because the degradation process of Li-ion batteries is highly nonlinear with negligible capacity fade at early cycles and is influenced by not only the operating conditions but also variances due to imperfect manufacturing tolerances. These factors contribute to the complexity of battery cycle life prediction [8]. This complexity and the importance of battery cycle life early prediction with high accuracy have made this an intense research area. Throughout the literature, the prediction methods can be generally divided into three categories—model-based methods, data-driven methods, and hybrid methods.

The model-based methods, in turn, can be roughly classified into three categories. In the first one, a physics-based model, such as an electrochemical model (EM), is incorporated into a recursive filter framework, such as the extended Kalman filter [9] or a particle filter [10], in which internal parameters are updated from measured data. However, computational complexity in terms of high memory requirements and long computation time inevitably limits their applicability in real-time battery management system (BMS). In the second category, empirical models are identified based on cell characterization

Manuscript received 29 July 2022; revised 14 October 2022; accepted 20 November 2022. Date of publication 5 December 2022; date of current version 13 June 2023. This work was supported by Volvo AB and Swedish Energy Agency under Project 45540-1. *(Corresponding author: Huang Zhang.)*

Huang Zhang is with the Department of Electromobility, Volvo Group Trucks Technology, 405 08 Gothenburg, Sweden, and also with the Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden (e-mail: huangz@chalmers.se).

Yang Su is with UMR ECOSYS, INRAE-UVSQ, Université Paris-Saclay, 78850 Thiverval-Grignon, France (e-mail: yang.su@inrae.fr).

Faisal Altaf is with the Department of Electromobility, Volvo Group Trucks Technology, 405 08 Gothenburg, Sweden (e-mail: faisal.altaf@volvo.com).

Torsten Wik is with the Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden (e-mail: torsten.wik@chalmers.se).

Sébastien Gros is with the Department of Engineering Cybernetics, Faculty of Information Technology, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway (e-mail: sebastien.gros@ntnu.no).

Digital Object Identifier 10.1109/TTE.2022.3226683

data from lab experiments. To improve their accuracy in onboard vehicle applications, model parameters can be adapted by the onboard BMS using measurement and state estimation data in a Bayesian filtering framework. This can include a range of Kalman filters [11], dual fractional-order extended Kalman filters [12], and particle filters [13]. The prediction accuracy of empirical models with a recursive filter highly depends on the fit model. In addition to these two model-filter-based models, a third category of semiempirical models has also been developed to capture the direct relationship between the operating conditions and the battery state-of-health (SoH), by interpolating and fitting experimental data. The only difference between the empirical and semiempirical models is that the former does not use any physical relation in the model structure, whereas the latter uses some level of physical insights in the model formulation. Most semiempirical models in the literature study the battery calendar aging and cyclic aging separately and then combine both to make predictions under various operating conditions [14], [15] [16]. Although it is easier to implement semiempirical models than the model-filter-based methods described, semiempirical models are open-loop approaches where the model parameters are determined by data fitting.

Data-driven methods for battery cycle life prediction are generally black-box models developed based on machine learning or deep learning approaches to capture the mapping between inputs (e.g., features extracted from incremental capacity curves [17]) and desired outputs (e.g., SoH). These methods can be either nonprobabilistic or probabilistic. Nonprobabilistic data-driven methods include autoregression (AR)-based models [18], [19], artificial neural network (ANN) [20], [21] [22], and support vector machine (SVM) [23], [24]. Despite the high accuracy of these nonprobabilistic data-driven methods on cycle life point prediction, they are unfortunately not able to provide any uncertainty estimate of their predictions. The uncertainty level of predictions can enable a system or a user to make risk-informed decisions [25]. For this reason, probabilistic data-driven methods, such as Gaussian process regression (GPR) [26], [27] and relevance vector machine (RVM) [28], [29], can be a better choice, as they have the ability to output a probability density function (pdf) and predict both the cycle life and the associated confidence interval.

Hybrid methods aim at leveraging the advantages of several different models. With the hybrid data-driven and model-based approach, a physics-based model is incorporated into a recursive filter framework (e.g., particle filter [30], [31] and Kalman filter [32]), and the model parameters are identified and updated with measurements. Remaining useful life (RUL) is obtained by projecting the estimated internal state to the future until a predefined end-of-life (EOL) threshold is reached. The data-driven model in this hybrid approach has been used for estimating the battery internal state from measurement data [30], extrapolating the measurements beyond the range of currently available measurements [30], [32] and as a replacement of a degradation model [31] in the physics-based model prediction case. The aforementioned studies have shown that hybrid methods have the potential to improve prediction accuracy further in comparison with one single data-driven model. However, they are difficult to use in online applications in a BMS due to its high computational complexity.

As illustrated by Severson et al. [5], cycle life for battery cells does not follow a normal distribution, which is a presumption of many probabilistic data-driven methods (e.g., GPR and RVM) that provide uncertainty information associated with the predictions. While advanced data-driven methods offer high prediction accuracy of battery cycle life despite minimum knowledge of the battery degradation mechanisms, interpretability of machine learning models is still under-explored in the literature. Extracting relevant battery aging knowledge from a machine learning or deep learning model in terms of underlying relationships, either in data or learned by the model, can provide valuable insights into battery aging. These insights can then be used to guide discoveries of aging mechanisms, improvements of battery manufacturing, and development of fast-charging protocols.

This work tries to tackle the aforementioned problems by introducing a quantile regression forest (QRF) model for reliable cycle life range prediction of Li-ion battery cells. The prediction intervals (PIs) are constructed by using a quantile regression method that estimates quantiles of the response variable given values of the input variables [33], [34]. The advantages of the QRF model over other probabilistic models are that asymmetric PIs can be estimated without assuming any specific distribution (e.g., Gaussian) of the output variable, i.e., the cycle life. In addition, the QRF model is a nonparametric model, which means that the number of parameters automatically adapts to the complexity of the training data.

The novelty and contributions are summarized as follows.

1) This work proposes the first application of QRF model to provide battery cycle life point prediction and uncertainty range prediction. It is shown that the QRF model not only provides point prediction with high accuracy but also cycle life range prediction with high probability without assuming any specified distribution for the cycle life. The performance of the proposed QRF model is demonstrated on a public dataset that includes various operating conditions in terms of realistic charging current profiles. Its point prediction performance is benchmarked to the elastic net model, whose exceptional early prediction performance was successfully demonstrated by Severson et al. [5] and its range prediction performance is compared with two popular probabilistic models, i.e., GPR and RVM.
2) An alpha-logistic-weighted (ALW) criterion is proposed for optimizing hyperparameters of the QRF model and its effectiveness of improving the coverage probability of the final QRF model is demonstrated. The proposed criterion can also be used for optimizing hyperparameters of other regression models that are capable of providing range predictions.
3) To statistically interpret the width of battery cycle life range prediction, two hypothesis tests are conducted. As a result, there is sufficient evidence in the first test

that width of range prediction is highly correlated with absolute mean prediction error at a significance level of 0.05, which suggests that width of range prediction may provide more information for decision-making under uncertainties than we get from point predictions alone; there is also sufficient evidence in the second test that the width of range prediction is highly correlated with six input features at a significance level of 0.05, which suggests that the width of the range prediction is mainly affected by values of these six input features.

4) To interpret the final QRF model for cycle life prediction and reveal the underlying relationships in data learned by the QRF model, permutation importance and partial dependence plot (PDP) are employed as model-agnostic methods to rank individual feature importance and quantitatively show the marginal effect each feature has on the predicted battery cycle life. Subsequently, an electrochemical interpretation is given to support what has been revealed by these two model-agnostic methods.
5) In an application of selecting the high-cycle-life charging protocol, the expected battery cycle life (EBCL) of a charging protocol can be determined with consideration of both point predictions and the uncertainty associated with the predictions. It is demonstrated that the final QRF model facilitates decision-making to select the high-cycle-life charging protocol that reduces the occurrence of unacceptably short cycle life.

II. Theoretical Background

A. Quantile Regression Forest

The random forest regression models approximate the conditional mean by a weighted sum over all the observations. Instead of averaging over all observations in every leaf of every tree, one could use all observations from each tree to construct an empirical cumulative distribution function of the response variable. Therefore, QRFs use full information of all the observations via combining quantile methods and random forest. For simplicity, first, we denote the τth quantile of Y given $\boldsymbol{X} = \boldsymbol{x}$ by $q_\tau(Y|\boldsymbol{X} = \boldsymbol{x})$, where $\boldsymbol{X}$ is the input random variable, possibly high-dimensional, and Y is the real-valued output random variable. The conditional cumulative distribution function $F(y|\boldsymbol{X} = \boldsymbol{x})$ is defined as the probability of Y smaller than y given $\boldsymbol{X} = \boldsymbol{x}$, i.e.,

$$F(y|\boldsymbol{X} = \boldsymbol{x}) = P(Y \leq y|\boldsymbol{X} = \boldsymbol{x}). \tag{1}$$

For a continuous conditional cumulative distribution function $F(y|\boldsymbol{X} = \boldsymbol{x})$, as defined above, the τth quantile $q_\tau(Y|\boldsymbol{X} = \boldsymbol{x})$ is defined such that the probability of Y less than or equal to $q_\tau(Y|\boldsymbol{X} = \boldsymbol{x})$ is equal to τ for a given $\boldsymbol{X} = \boldsymbol{x}$, i.e.,

$$q_\tau(Y|\boldsymbol{X} = \boldsymbol{x}) = \inf\{y : F(y|\boldsymbol{X} = \boldsymbol{x}) \geq \tau\}. \tag{2}$$

During the inference, for an input $\boldsymbol{X} = \boldsymbol{x}$, the leaf of the tth regression tree that contains $\boldsymbol{x}$ is denoted by $l(\boldsymbol{x}, \boldsymbol{\theta}_t)$, where $\boldsymbol{\theta}_t$ is the parameter vector that determines how the tth tree is grown, for example, input variables that are considered in each node split. The weight $w_i(\boldsymbol{x}, \boldsymbol{\theta}_t))$ from each tree is calculated by

$$w_i(\boldsymbol{x}, \boldsymbol{\theta}_t) = \frac{\mathbb{1}_{\{\boldsymbol{x}_i \in l(\boldsymbol{x}, \boldsymbol{\theta}_t)\}}}{\#\{j : \boldsymbol{x}_j \in l(\boldsymbol{x}, \boldsymbol{\theta}_t)\}}, \quad i = 1, \ldots, N \tag{3}$$

where N is the total number of the observations and $\mathbb{1}_{\{\boldsymbol{x}_i \in l(\boldsymbol{x}, \boldsymbol{\theta}_t)\}}$ is an indicator function equal to 1 if $\boldsymbol{x}_i \in l(\boldsymbol{x}, \boldsymbol{\theta}_t)$ and otherwise equal to 0. $\#\{j : \boldsymbol{x}_j \in l(\boldsymbol{x}, \boldsymbol{\theta}_t)\}$ denotes the total number of observations that are in the leaf $l(\boldsymbol{x}, \boldsymbol{\theta}_t)$.

Therefore, the weight from the whole random forest is defined as the average of $w_i(\boldsymbol{x}, \boldsymbol{\theta}_t)$ over all the T regression trees grown, which reads as

$$w_i(\boldsymbol{x}) = \frac{1}{T} \sum_{t=1}^{T} w_i(\boldsymbol{x}, \boldsymbol{\theta}_t). \tag{4}$$

Furthermore, the constructed conditional cumulative distribution function of the QRF model is expressed as

$$F(y|\boldsymbol{X} = \boldsymbol{x}) = E(\mathbb{1}_{\{Y \leq y\}}|\boldsymbol{X} = \boldsymbol{x}) \tag{5}$$

where $\mathbb{1}_{\{Y \leq y\}}$ is an indicator function and equal to 1 if $Y \leq y$ and otherwise equal to 0. $E(\mathbb{1}_{\{Y \leq y\}}|\boldsymbol{X} = \boldsymbol{x})$ can be approximated by the weighted sum over the observations of $\mathbb{1}_{\{Y \leq y\}}$. Thus, an empirical conditional probability function $\hat{F}$, given $\boldsymbol{X} = \boldsymbol{x}$, can be obtained as

$$\hat{F}(y|\boldsymbol{X} = \boldsymbol{x}) = \sum_{i=1}^{N} w_i(\boldsymbol{x}) \mathbb{1}_{\{y_i \leq y\}} \tag{6}$$

where the weights $w_i(\boldsymbol{x})$ are the same as defined in (4) and the indicator function $\mathbb{1}_{\{y_i \leq y\}}$ determines whether the weight will be counted or not, depending on the condition $y_i \leq y$.

Finally, the estimated τth quantile $\hat{q}_\tau(Y|\boldsymbol{X} = \boldsymbol{x})$ is obtained by replacing $F(y|\boldsymbol{X} = \boldsymbol{x})$ in (2) with $\hat{F}(y|\boldsymbol{X} = \boldsymbol{x})$ in (6), i.e.,

$$\hat{q}_\tau(Y|\boldsymbol{X} = \boldsymbol{x}) = \inf\{y : \hat{F}(y|\boldsymbol{X} = \boldsymbol{x}) \geq \tau\}. \tag{7}$$

In addition, based on the standard random forest grown in the QRF model, the conditional mean of Y given $\boldsymbol{X} = \boldsymbol{x}$ can also provide point predictions by

$$\hat{f}(\boldsymbol{x}) = \sum_{i=1}^{N} w_i(\boldsymbol{x}) y_i \tag{8}$$

where the weights $w_i(\boldsymbol{x})$ are defined in (4).

B. Prediction Interval

The PIs can be constructed from the conditional quantiles estimated by the QRF model. Specifically, the $(1 - \alpha) \times 100\%$ PIs for output variable Y, given $\boldsymbol{X} = \boldsymbol{x}$, is constructed by

$$\hat{I}(\boldsymbol{x}) = [\hat{q}_{\alpha/2}(Y|\boldsymbol{X} = \boldsymbol{x}), \hat{q}_{1-\alpha/2}(Y|\boldsymbol{X} = \boldsymbol{x})]. \tag{9}$$

For example, the 95% PI for the output Y is estimated by $\hat{I}(\boldsymbol{x}) = [\hat{q}_{.025}(Y|\boldsymbol{X} = \boldsymbol{x}), \hat{q}_{.975}(Y|\boldsymbol{X} = \boldsymbol{x})]$, which should be interpreted as follows: given $\boldsymbol{X} = \boldsymbol{x}$, a new observation of output Y is in the interval $\hat{I}(\boldsymbol{x})$ with a probability of 95%.

C. Permutation Importance

In order to understand the underlying battery degradation process, the goal of battery cycle life prediction should not only be limited to learn a regression function $\hat{f}$ that is capable of making battery cycle life predictions with high accuracy but also to identify input variables from feature engineering that are the most important for the prediction accuracy of the learned model. A tool-like variable importance can be helpful for identifying which input variables that are the most important and therefore should be measured with high precision [35].

The concept of variable importance was first introduced by Breiman [36] for random forests, in which variable importance of an input variable is measured by the decrease in predication accuracy when the values of this input variable are randomly permuted/shuffled in out-of-bag samples and then dropping the out-of-bag samples down the corresponding trees. Louppe et al. [37] characterized an alternative measure of variable importance based on the mean decrease impurity (MDI). Impurity is quantified by the splitting criterion of the decision trees (e.g., mean squared error for continuous outputs). However, variable importance based on the MDI method favors high cardinality input variables over low cardinality input variables, such as binary variables or categorical variables with a small number of possible categories. Furthermore, variable importance based on the MDI method can only be calculated on the training set during growth of trees. Therefore, there are possibilities that the MDI method gives high importance to input variables that may not be predictive on unseen data when the model is overfitted [37]. With the aim of mitigating those limitations of variable importance based on the MDI method, permutation importance is developed as a generalized model-agnostic method for measuring the importance of an input variable by calculating the increase of prediction error after permuting values of the input variable [35]. An input variable is considered to be important if shuffling its value leads to an increase in the model error and vice versa. Permutation importance can be computed on both the training set and the test set, which makes it possible to identify features that contribute the most to the generalized prediction power of the fit model. The permutation importance algorithm below describes how the measure is calculated.

Algorithm 1 Permutation Importance Algorithm

Input: Learned model $\hat{f}$, training or test set $\mathcal{G}$, and error score function s (e.g., R^2 score)
Compute the reference error score s_0 of the learned model $\hat{f}$ on dataset $\mathcal{G}$
for each input variable $j = 1, 2, \ldots, p$ **do**
  for each repetition $m = 1, 2, \ldots, M$ **do**
    Randomly shuffle the values in the column corresponding to input variable j in the dataset G to generate a corrupted version of the dataset, $\widetilde{\mathcal{G}}$
    Compute the error score $s_{j,m}$ of the learned model $\hat{f}$ on corrupted version of the dataset $\widetilde{\mathcal{G}}$
  end for
  Compute importance Im_j for input variable j defined as $\mathrm{Im}_j = s_0 - \frac{1}{M}\sum_{m=1}^{M} s_{j,m}$
end for
Sort input variables in descending order of variable importance.

D. Partial Dependence Plot

After the most important input variables have been identified, the next step is to understand the dependence of the approximation $\hat{f}(\boldsymbol{X})$ on the joint values of the input variables [38].

Consider the subvector $\boldsymbol{X}_{\mathcal{S}}$ of length $\ell < p$ of the input vector $\boldsymbol{X} = (X_1, X_2, \ldots, X_p)^T$, indexed by $\mathcal{S} \subset \{1, 2, \ldots, p\}$. Let $\mathcal{C}$ be the complement set, with $\mathcal{S} \cup \mathcal{C} = \{1, 2, \ldots, p\}$, and $\boldsymbol{X}_{\mathcal{C}}$ the corresponding subvector. In principle, the approximation $\hat{f}(\boldsymbol{X})$ depends on all of the input variables, i.e., $\hat{f}(\boldsymbol{X}) = \hat{f}(\boldsymbol{X}_{\mathcal{S}}, \boldsymbol{X}_{\mathcal{C}})$. However, if the variables in $\boldsymbol{X}_{\mathcal{S}}$ do not have strong interactions with those in $\boldsymbol{X}_{\mathcal{C}}$, then the average, or partial dependence, of $\hat{f}(\boldsymbol{X})$ on $\boldsymbol{X}_{\mathcal{S}}$ is approximately

$$\hat{f}_{\mathcal{S}}(\boldsymbol{X}_{\mathcal{S}}) = E_{X_{\mathcal{C}}}[\hat{f}(\boldsymbol{X})] = E_{X_{\mathcal{C}}}[\hat{f}(\boldsymbol{X}_{\mathcal{C}}, \boldsymbol{X}_{\mathcal{S}})]. \tag{10}$$

In practice, the partial dependence function $\hat{f}_{\mathcal{S}}(\boldsymbol{X}_{\mathcal{S}})$ can therefore be estimated by

$$\bar{f}_{\mathcal{S}}(\boldsymbol{X}_{\mathcal{S}}) = \frac{1}{N}\sum_{i=1}^{N} \hat{f}(\boldsymbol{X}_{\mathcal{S}}, \boldsymbol{x}_{i\mathcal{C}}) \tag{11}$$

where $\{\boldsymbol{x}_{1\mathcal{C}}, \boldsymbol{x}_{2\mathcal{C}}, \ldots, \boldsymbol{x}_{N\mathcal{C}}\}$ are the values of $\boldsymbol{X}_{\mathcal{C}}$ occurring in the training set and N is the total number of samples in the training set.

E. Performance Evaluation Metrics

1) Evaluating Point Prediction Quality: There are various performance evaluation metrics that can be used to evaluate the accuracy of point predictions. The most common ones are root mean squared error (RMSE), mean absolute percentage errors (MAPEs), and coefficient of determination (R^2), which are defined as follows:

$$\mathrm{RMSE}(y_i, \hat{y}_i) = \sqrt{\frac{1}{N_T}\sum\nolimits_{i=1}^{N_T}(y_i - \hat{y}_i)^2} \tag{12}$$

$$\mathrm{MAPE}(y_i, \hat{y}_i) = \frac{1}{N_T}\sum_{i=1}^{N_T}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{13}$$

$$R^2(y_i, \hat{y}_i) = 1 - \frac{\sum_{i=1}^{N_T}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N_T}(y_i - \bar{y})^2} \tag{14}$$

where N_T is the number of samples to be evaluated (i.e., all samples in the test set), $\hat{y}_i = \hat{f}(\boldsymbol{x}_i)$ denotes the estimated mean cycle life predicted by the model, y_i denotes the corresponding observed cycle life, and $\bar{y} = (1/N_T)\sum_{i=1}^{N_T} y_i$ is the average cycle life for a total of N_T samples in the test set.

2) Evaluating Range Prediction Quality: Two commonly used metrics for range predictions are PI coverage probability (PICP) and mean PI width (MPIW) [39].

The PICP shows the percentage of output values covered between the lower and upper bounds of the PIs and as such assesses the calibration of the range predictions. A larger PICP

means that more output values will fall in the constructed PIs and vice versa. The PICP is defined as

$$\text{PICP} = \frac{1}{N_T}\sum_{i=1}^{N_T} c_i \times 100\% \tag{15}$$

where $c_i = 1$, if the output value y_i is covered by the interval from the lower bound L_i to the upper bound U_i of the constructed PI; otherwise, $c_i = 0$.

In a practical application, such as the battery cycle life prediction in the present work, the width of constructed PIs is of equal importance as the coverage probability since it does not make much sense to have PIs with high coverage probability and large width at the same time. For a narrower PI, the prediction is clearly more informative than for a wider PI. Therefore, there is also a need to assess the sharpness of range predictions, which can be done via MPIW defined as

$$\text{MPIW} = \frac{1}{N_T}\sum_{i=1}^{N_T} |U_i - L_i|. \tag{16}$$

Theoretically, it is desirable to have PIs with a PICP value close to their nominal coverage (e.g., 95%) and a small MPIW value.

Both PICP and MPIW only assess PIs from one aspect. A comprehensive measure of both coverage probability and width of PIs at the same time is the averaged interval score (AIS), proposed in [40], which is defined as

$$\text{AIS} = \frac{1}{N_T}\sum_{i=1}^{N_T}((U_i - L_i) + \frac{2}{\alpha}(L_i - y_i)\mathbb{1}_{\{y_i < L_i\}} + \frac{2}{\alpha}(y_i - U_i)\mathbb{1}_{\{y_i > U_i\}}). \tag{17}$$

The AIS defined above rewards narrow PIs and penalizes intervals missed by the observation with a weight depending on α, as defined in (9). The AIS is used for standard QRF training.

III. Methodology and Problem Formulation

A. Feature Engineering and Selection

Generally, feature engineering can be divided into two categories—manual feature engineering based on domain knowledge [41], [42] and automatic feature engineering, such as autoencoders [43], and restricted Boltzmann machine [44].

In this work, manual feature engineering based on battery domain knowledge is adopted. More specifically, 33 features were extracted from data of the first 100 cycles, at which point most batteries have not yet exhibited any significant capacity degradation. The 33 features are divided into five groups and listed in Table I.

To reduce the computation time and improve the performance of learned models, a feature selection method is employed, which selects an effective subset of features by reducing irrelevant and redundant features [45]. For this, a random-forest-based recursive feature elimination with cross validation (RF-RFE-CV) is employed, which selects a subset of features by recursively removing features with the least importance in the current feature set. Only the training set is used for feature selection to avoid introducing optimistically biased performance estimates. As a result, 12 features were automatically selected as a feature subset and then fed into the QRF and elastic net models for battery cycle life prediction.

B. Problem Formulation

The battery cycle life early prediction problem can be formulated as a regression problem with the goal of learning a mapping f from a random input vector (a term we will use interchangeably with features) $\boldsymbol{X} = (X_1, X_2, \ldots, X_p)^T$ in the space $\mathcal{X} \subseteq \mathbb{R}^p$ to a random output (a term we will use interchangeably with response) variable Y in the space $\mathcal{Y} \subseteq \mathbb{R}^+$, $f : \mathcal{X} \to \mathcal{Y}$, given a training set $\mathcal{D} = \{\boldsymbol{x}_i, y_i\}_{i=1}^N$, where N is the number of assumed independent and identically distributed samples in the training set. In the present case, $\boldsymbol{x}_i \in \mathcal{X}$ represents p features extracted from the first 100 cycles and $y_i \in \mathcal{Y}$ is the observed battery cycle life.

To learn the mapping function f, the conditional mean minimizing the expected squared error loss, with the assumption that the squared error loss function is symmetric around zero, is used

$$E(Y|\boldsymbol{X} = \boldsymbol{x}) = \underset{f(\boldsymbol{x})}{\arg\min}\, E\{(Y - f(\boldsymbol{x}))^2 | X = \boldsymbol{x}\}. \tag{18}$$

In practice, the approximation of the conditional mean is achieved by minimization of a squared error-type loss function over the training set $\mathcal{D}$, and the resulting learned regression function is denoted as $\hat{f}$.

The conditional mean only reveals one aspect of the conditional distribution of a response variable Y and gives no information about the uncertainty associated with the predicted conditional mean. In our case, though, we are interested to find the range of predicted battery cycle life in which the battery will reach its EOL with high probability. We propose QRF to handle this case, yielding both a point prediction as well as its uncertainty. More specifically, its point prediction is provided by the estimated conditional mean, given $\boldsymbol{X} = \boldsymbol{x}$, i.e.,

$$\hat{f}(\boldsymbol{x}) = \sum_{i=1}^{N} w_i(\boldsymbol{x}) y_i \tag{19}$$

where the weights $w_i(\boldsymbol{x})$ are defined in (4).

The corresponding range prediction, provided by a 95% PI given $\boldsymbol{X} = \boldsymbol{x}$, is given by

$$\hat{I}(\boldsymbol{x}) = [\hat{q}_{.025}(Y|\boldsymbol{X} = \boldsymbol{x}), \hat{q}_{.975}(Y|\boldsymbol{X} = \boldsymbol{x})]. \tag{20}$$

C. Proposed PI Evaluation Criterion

From a decision-making perspective, having point predictions with a low RMSE and range predictions with a high PICP together with a low MPIW is preferable. However, there is a tradeoff between maximizing PICP and minimizing MPIW. The AIS defined in (17) only assesses PIs over each validation sample in the training set without consideration of whether a preassigned nominal coverage probability $(1 - \alpha) \times 100\%$ is satisfied or not on the training set. Therefore, it is expected to propose a comprehensive measure that includes both two

TABLE I
33 FEATURES IN FIVE GROUPS

Groups	Features
Charge-related	Average charge time for the first 5 cycles [5].
Discharge voltage curve-related	Minimum, variance, skewness, and kurtosis of difference of the discharge voltage curve between cycle 100 and cycle 10 (i.e., $\Delta Q_{100-10}(V)$) [5]. Amplitude and position shift of the highest peak in the discharge incremental capacity curve between cycle 10 and cycle 100 (i.e., $dQdV_{100-10}$).
Capacity-related	Slope of the linear fit to the capacity fade curve from cycle 2 to cycle 100 [5]. Intercept of the linear fit to capacity fade curve from cycle 2 to cycle 100 [5]. Discharge capacity at cycle 2 [5]. Discharge capacity at cycle 100 [5]. Difference between maximum discharge capacity within the first 100 cycles and discharge capacity at cycle 2 [5].
Temperature-related	Minimum, variance, skewness, and kurtosis of difference in the discharge cell temperature, as a function of voltage, between cycle 100 and cycle 10 (i.e., $\Delta T_{100-10}(V)$). Minimum, maximum, mean, and variance of discharge cell temperature as a function of voltage at cycle 10 (i.e., $T_{10}(V)$). Minimum, maximum, mean, and variance of discharge cell temperature as a function of voltage at cycle 100 (i.e., $T_{100}(V)$). Difference in minimum, maximum, mean, and variance of discharge cell temperature, as a function of voltage, between cycle 10 and cycle 100.
Internal resistance-related	Minimum internal resistance from cycle 2 to cycle 100 [5]. Maximum internal resistance from cycle 2 to cycle 100. Internal resistance at cycle 2 [5]. Internal resistance at cycle 100. Difference in internal resistance between cycle 100 and cycle 2 [5].

properties of PIs. More importantly, the calibration property is prioritized rather than the sharpness property of the PIs. Thus, higher penalties should be given in the case of unsatisfactory nominal coverage probability $(1-\alpha)\times 100\%$ that is usually preassigned by a decision maker. To address such problems, we propose an ALW criterion based on the work by Khosravi et al. [39], which reads as

$$\text{ALW} = \frac{\text{MPIW}}{\sigma(\alpha, \text{PICP})} \tag{21}$$

where $\sigma(\cdot)$ is the logistic function defined as

$$\sigma(\alpha, \text{PICP}) = \frac{1}{1+e^{-\frac{1}{\alpha}(\text{PICP}-(1-\alpha))}} \tag{22}$$

where α is the same as in (9). The terms $(1/\alpha)$ and $1-\alpha$ determine the growth rate and the midpoint of the logistic curve, respectively. By inserting (22) into (21), (21) can be rewritten as

$$\text{ALW} = \text{MPIW}(1+e^{-\frac{1}{\alpha}(\text{PICP}-(1-\alpha))}). \tag{23}$$

The ALW defined above rewards lower MPIW and exponentially penalizes unsatisfactory PICP that is lower than nominal coverage probability $(1-\alpha)\times 100\%$ with a weight depending on α.

D. Hyperparameter Optimization

In this work, the objective of hyperparameter optimization is to find an optimal set of hyperparameters for the QRF model that minimizes the value of AIS, defined in (17) or the proposed ALW defined in (23), given a training set at each train–test split. The leave-one-out cross-validation (LOO-XVE) method [46] is adopted for evaluating the averaged performance of a QRF model given a set of hyperparameters. The LOO-XVE is suitable for small datasets, as in the present work, where prediction performance also outweighs computational cost at the training stages. Optuna [47], a Bayesian hyperparameter optimization framework, is used to search for the optimal set of hyperparameters for the QRF model given the training set. The final QRF model with the optimal set of hyperparameters is learned on the whole training set and is then evaluated on the test set at each train–test split.

E. Proposed Expected Battery Cycle Life of a Charging Protocol

In an application of selecting the high-cycle-life charging protocol, the EBCL of a charging protocol can be calculated by averaging over all predicted mean cycle lives of cells charged with this charging protocol [6]

$$\text{EBCL} = \frac{1}{B}\sum_{i=1}^{B}\hat{f}(\boldsymbol{x}_i) \tag{24}$$

where B denotes the total number of battery cells charged with this charging protocol.

However, it will be a difficult selection decision to be made when the expected battery cycle lives for two charging protocols calculated using EBCL are very close to each other. Therefore, we propose the EBCL range (EBCLR) that helps facilitate decision-making in this case

$$\text{EBCLR} = |U-L| \tag{25}$$

where U and L are the upper bound and lower bound, respectively, of the EBCLR prediction. Considering the relatively small dataset used in this work, we adopt the median approach to combine PIs for each charging protocol [48]

$$L = \text{Median}(L_1, \ldots, L_B);\ U = \text{Median}(U_1, \ldots, U_B) \quad (26)$$

where $L_1, \ldots, L_B$ are the lower bounds of battery cycle life range predictions of a group of B cells that are charged with the same charging protocol and $U_1, \ldots, U_B$ are the upper bounds of battery cycle life range predictions of a group of B cells that are charged with the same charging protocol.

IV. Experiments and Results

A. Battery Dataset

The battery dataset used in the present work is originally from the work of Toyota Research Institute in collaboration with Stanford University and MIT [5]. An early prediction model developed in their work [5] was later used for selecting high-cycle-life charging protocols [6]. There are 124 lithium iron ferrous phosphate (LFP)/graphite cells in this dataset with a nominal capacity of 1.1 Ah. The 124 cells are from three different test batches (i.e., the "2017-05-12" batch, the "2017-06-30" batch, and the "2018-04-12" batch) with batch date referring to the date the batch started. All the cells are tested at a constant temperature of 30 °C in an environmental chamber. The cells are charged with a one- or two-step fast-charging protocol and identically discharged at a 4 C-rate. Cells are charged from 0% to 80% state-of-charge (SoC) with one of 72 charging protocols, for example, a charging protocol "5.6C(36%)-4.3C" consists of a 5.6 C charging step from 0% to 36% SoC, followed by a 4.3 C step from 36% to 80% SoC. It is assumed that the Li-ion batteries reach their EOL when their discharge capacity has decreased to 80% of their initial nominal capacity. Time-series voltage, current, and cell temperature were continuously measured during cycling. The internal resistance was measured per cycle during charging at 80% SoC by averaging ten pulses of ±3.6C with a pulsewidth of around 30 ms.

B. Train–Test Split

There are 72 different charging protocols in this battery dataset with nominal charging time from 0% to 80% SoC ranging from 9 to 13.3 min. For the purpose of reducing the possibly large sampling error due to the small dataset used in this work, the stratified random sampling method [49] is employed to randomly split the dataset into a training set that contains 80% of the total dataset (99 samples) for optimizing model hyperparameters and learning the final model and a test set that contains 20% of the total dataset (25 samples) for evaluating the performance of the final model. At each split, equal ratios of fast-charged (i.e., less than 10.5 min) cells, medium-charged (i.e., between 10.5 and 11.7 min) cells, and slow-charged (i.e., greater than 11.7 min) cells are maintained in the training and test sets. Moreover, in order to reduce the random effect of the selected split, stratified random sampling is repeated five times, and then, the results of five train–test splits are averaged.

C. Performance Evaluation and Results

For a fair comparison, the optimal sets of hyperparameters for the QRF model and other benchmark models (i.e., elastic net regression model, GPR, and RVM) are obtained using the hyperparameter optimization method described in Section IV-B, given the same training set at each train–test split. The optimal sets of hyperparameters for the QRF model and the other models are reported in Table II.

The final QRF model and other models with their optimal sets of hyperparameters are learned on the same training set and then evaluated using the same test set at each train–test split, with evaluation metrics for point predictions or range prediction. The results of five train–test splits are averaged and reported in Tables III and IV.

To compare the point prediction and range prediction results between the QRF model and other models, the performance improvement in percentage is calculated with elastic net model as a benchmark in the point prediction comparison and with QRF using AIS defined in (17) as PI evaluation criterion in LOO-XVE as a benchmark in the range prediction comparison. In terms of point prediction, it can be seen from the results in Table III that the QRF outperforms the elastic net regression model, GPR, and RVM, evaluated by all three performance measures, i.e., RMSE, MAPE, and R^2. In terms of range prediction, it can be seen from the results in Table IV that even though the QRF model using the proposed ALW as PI evaluation criterion in LOO-XVE has an 11.2% worse MPIW value than the QRF model using the AIS as PI evaluation criterion in LOO-XVE, it has a 4.9% better PICP value that is much closer to the nominal coverage probability 95%, and its overall evaluation of PIs via AIS is 10.3% better than those by the QRF model using AIS. Notably, the GPR model has the best PICP value over all other models but is not able to compete with the QRF model measured by MPIW or AIS.

In summary, these results suggest that the QRF model is capable of providing at least 20% higher point prediction accuracy than the elastic net model whose effectiveness was demonstrated in the work of Severson et al. [5]. A possible explanation would be that, as a nonparametric model, the QRF model is more flexible than parametric models (e.g., the elastic net model) to extract complex patterns in the battery data without necessarily incurring severe overfitting. Moreover, by using the proposed ALW as PI evaluation criterion in LOO-XVE, the range prediction performance by the QRF models has indeed improved: higher coverage probability that is closer to the preassigned nominal 95% coverage probability guarantees higher reliability of the final QRF model. Even though GPR provides the best coverage probability, the QRF model outperforms GPR in terms of width of PIs evaluated by MPIW and comprehensive measure of both coverage probability and width of PIs by AIS. The final QRF model using the proposed ALW as PI evaluation criterion in LOO-XVE is the one used for later analyses.

D. Correlation Analysis and Results

To statistically interpret the width of battery cycle life range prediction, hypothesis tests using Pearson's correlation

TABLE II
OPTIMAL MODEL HYPERPARAMETERS

Models	Model hyperparameters	Optimal values 5 train-test splits
Elastic net	α - the relative weight of the L1 and L2 penalties; λ - the regularization parameter.	$\alpha_1 = 0.42$, $\lambda_1 = 0.0572$; $\alpha_2 = 1.00$, $\lambda_2 = 0.0313$; $\alpha_3 = 0.01$, $\lambda_3 = 0.0218$; $\alpha_4 = 0.01$, $\lambda_4 = 0.0333$; $\alpha_5 = 0.57$, $\lambda_5 = 0.0945$.
GPR (Sum of a radial basis function kernel and a white noise kernel)	σ_l - the length scale; σ_f^2 - the signal variance; σ_n^2 - the noise variance.	$\sigma_{l_1} = 8.1393$, $\sigma_{f_1}^2 = 976.3646$, $\sigma_{n_1}^2 = 17508.5982$; $\sigma_{l_2} = 8.7730$, $\sigma_{f_2}^2 = 983.8729$, $\sigma_{n_1}^2 = 26238.2269$; $\sigma_{l_3} = 10.5000$, $\sigma_{f_3}^2 = 989.1063$, $\sigma_{n_3}^2 = 20682.1931$; $\sigma_{l_4} = 8.1573$, $\sigma_{f_4}^2 = 947.8936$, $\sigma_{n_4}^2 = 13712.8706$; $\sigma_{l_5} = 7.2395$, $\sigma_{f_5}^2 = 787.9068$, $\sigma_{n_5}^2 = 20715.8807$.
RVM (Radial basis function kernel)	γ - the kernel scale; β - the inverse noise variance.	$\gamma_1 = 0.01002$, $\beta_1 = 0.00003$; $\gamma_2 = 0.01003$, $\beta_2 = 0.00003$; $\gamma_3 = 0.01009$, $\beta_3 = 0.00003$; $\gamma_4 = 0.01189$, $\beta_4 = 0.00004$; $\gamma_5 = 0.01397$, $\beta_5 = 0.00003$.
QRF+AIS	n_{tree} - the number of trees; m_{try} - the number of random features in each split; l_{node} - the minimum number of samples at a leaf node.	$n_{tree_1} = 213$, $m_{try_1} = 9$, $l_{node_1} = 1$; $n_{tree_2} = 1393$, $m_{try_2} = 12$, $l_{node_2} = 2$; $n_{tree_3} = 1776$, $m_{try_3} = 12$, $l_{node_3} = 3$; $n_{tree_4} = 110$, $m_{try_4} = 6$, $l_{node_4} = 1$; $n_{tree_5} = 191$, $m_{try_5} = 11$, $l_{node_5} = 4$.
QRF+ALW	n_{tree} - the number of trees; m_{try} - the number of random features in each split; l_{node} - the minimum number of samples at a leaf node.	$n_{tree_1} = 1184$, $m_{try_1} = 8$, $l_{node_1} = 4$; $n_{tree_2} = 982$, $m_{try_2} = 8$, $l_{node_2} = 6$; $n_{tree_3} = 1072$, $m_{try_3} = 12$, $l_{node_3} = 6$; $n_{tree_4} = 1753$, $m_{try_4} = 8$, $l_{node_4} = 1$; $n_{tree_5} = 854$, $m_{try_5} = 10$, $l_{node_5} = 4$.

TABLE III
BATTERY CYCLE LIFE POINT PREDICTION PERFORMANCE

Models	Point prediction evaluation			Performance improvement (%)		
	RMSE (cycles)	MAPE (%)	R^2	RMSE (cycles)	MAPE (%)	R^2
Elastic net	196	20.2	0.70	/	/	/
GPR	216	14.0	0.63	-10.2%	-30.6%	-10.0%
RVM	226	16.0	0.60	-15.3%	-20.8%	-14.3%
QRF+AIS	142	10.9	0.85	27.6%	46.0%	20.0%
QRF+ALW	158	12.0	0.81	19.4%	40.6%	15.7%

TABLE IV
BATTERY CYCLE LIFE RANGE PREDICTION PERFORMANCE

Models	Range prediction evaluation			Performance improvement (%)		
	PICP (%)	MPIW (cycles)	AIS (cycles)	PICP (%)	MPIW (cycles)	AIS (cycles)
GPR	96.0	677	984	6.7%	-54.6%	-50.9
RVM	94.4	731	1602	4.9%	-66.9%	-145.7%
QRF+AIS	90.0	438	652	/	/	/
QRF+ALW	94.4	487	585	4.9%	-11.2%	10.3%

coefficient are conducted. More specifically, we determine whether there is a linear correlation between width of range prediction and variables under investigation. The first step is to define the null hypothesis (i.e., H_0—the width of range prediction does not linearly correlate with the variable under investigation) and the alternative hypothesis (i.e., H_1—width of range prediction linearly correlates with the variable under investigation). A p-value less than a significance level of 0.05 allows to reject the null hypothesis, which indicates that the width of range prediction highly correlates with the variable under investigation [50].

The first hypothesis test is performed to decide whether there is a linear correlation between the width of range prediction and the absolute mean prediction error (i.e., the distance between the point prediction and the observation) on the test set over five train–test splits. The second hypothesis

TABLE V
HYPOTHESIS TEST RESULTS OF CORRELATION BETWEEN WIDTH OF RANGE PREDICTION AND INPUT FEATURES

Input features	Pearson's correlation coefficient values	p-values	Decisions
var_dQ_100_10	-0.563	1e-11	rejected H_0
minimum_dQ_100_10	-0.545	1e-10	rejected H_0
IR_2	-0.279	0.0016	rejected H_0
variance_dT_100_10	-0.275	0.0019	rejected H_0
intercept_lin_fit_2_100	-0.239	0.0072	rejected H_0
maximum_IR_2_100	-0.215	0.0160	rejected H_0
peak_amplitude_dQdV_100_10	-0.172	0.0551	Failed to reject H_0
diff_maximum_T_100_10	-0.166	0.0643	Failed to reject H_0
peak_position_dQdV_100_10	-0.155	0.0848	Failed to reject H_0
diff_mean_T_100_10	-0.136	0.1299	Failed to reject H_0
minimum_T_10	0.129	0.1521	Failed to reject H_0
diff_IR_100_2	-0.008	0.9255	Failed to reject H_0

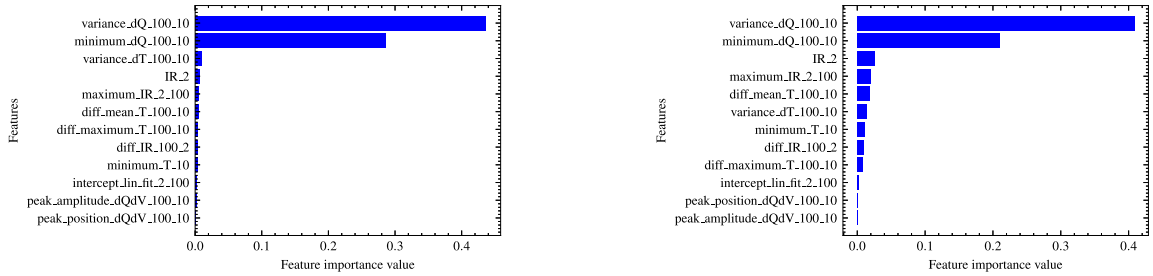


Fig. 1. Permutation importance ranking of 12 selected features. (a) Permutation importance ranking on the training set. (b) Permutation importance ranking on the test set.

test is performed to decide whether there is a linear correlation between the width of range prediction and each of the 12 input features on the test set in the five train–test splits.

For the first hypothesis test, Pearson's correlation coefficient is found to be 0.71 with a sufficiently low p-value that allows to reject the null hypothesis (p-value < 0.05). We, therefore, conclude that there is sufficient evidence that the width of range prediction and the absolute mean prediction error are highly correlated.

For the second hypothesis test, Pearson's correlation coefficient values and corresponding p-values between the width of range prediction and the 12 input features are listed in Table V, where the results are sorted in ascending order of their p-values. It can be seen from the results in Table V that in total, six input features are highly correlated with width of range prediction and that the variance and minimum of difference of the discharge voltage curve between cycles 100 and 10 (i.e., $\Delta Q_{100-10}(V)$) have the largest effect on the predicted range.

E. Computational Aspects

The whole experiment in this work runs on a laptop with an Intel Core i5 CPU and 16-GB memory. The hyperparameter optimization for the QRF model during the training stage takes approximately 10 h at one train–test split. Then, the final QRF model with the optimal set of hyperparameters is learned on the training set. At the inference stage, it takes the final QRF model approximately 150 ms to predict the cycle life of a battery cell given its input feature realization from its early degradation data.

For the real application of the final QRF model, we must first investigate the viability of embedding the final QRF model on a BMS, for example, memory requirement of a final QRF model implementation on a BMS. If the final QRF model is heavy, then we will prefer to implement a large part of it in the cloud, whereas preprocessing of data and feature engineering can be done onboard so that we do not have to send high-frequency data to the cloud via vehicle's telemetry gateway. Moreover, the main purpose of this model is to perform early prediction of cycle life. Since aging is a slow process, the final QRF model does not need to run in real time. We may run it periodically or in an event-triggered manner in the cloud based on preprocessed data from onboard BMS.

V. MODEL INTERPRETATION

Permutation importance is computed on the training set, as shown in Fig. 1(a). The two most important features identified are related to the $\Delta Q_{100-10}(V)$ curve, i.e., the discharge capacity change as a function of voltage between cycles 100 and 10, which was selected in the cycle life prediction model developed by Severson et al. [5]. This $\Delta Q(V)$ curve is of great interest because the curve itself and its derivatives contain rich information for degradation diagnosis [51], [52] [53]. Based on the training set, the most important feature is the variance of $\Delta Q_{100-10}(V)$ curve (see [5] for the definition of this feature), which means that

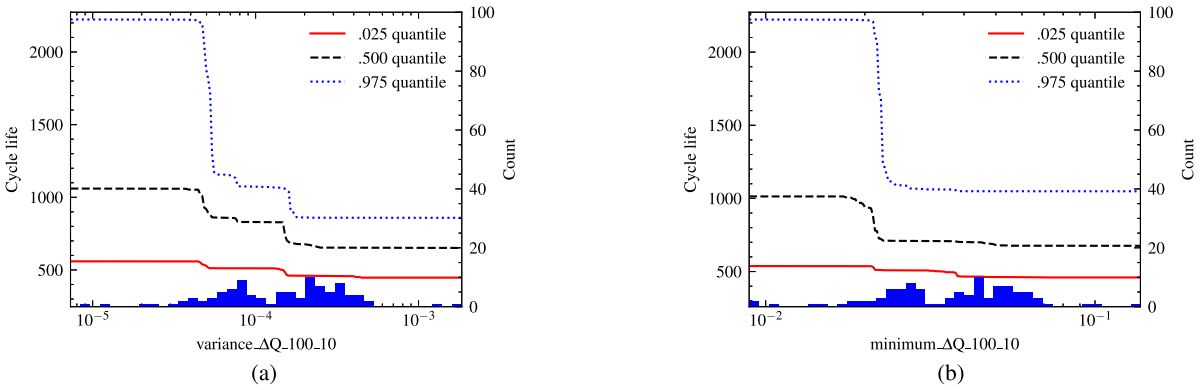


Fig. 2. PDPs for the cycle life prediction with respect to variance of $\Delta Q_{100-10}(V)$ and minimum of $\Delta Q_{100-10}(V)$. Histogram at the bottom shows observations of the feature, with scale to the right. (a) Variance of $\Delta Q_{100-10}(V)$. (b) Minimum of $\Delta Q_{100-10}(V)$.

the QRF model relies on this feature the most for making predictions. However, measuring the feature importance on the training set on which the QRF model is trained is not as informative as that on the unseen data. If the QRF model is overfitted, the feature importance measured on the training set may mislead us to believe that the wrong features are important. Therefore, the feature importance is also measured on the test set. It is shown in Fig. 1(b) that the most importance feature is still the variance of $\Delta Q_{100-10}(V)$ curve, which means that this feature does indeed contribute the most to the prediction performance of the QRF model.

In order to further illustrate how the most important feature affects the predicted cycle life, a 1-D PDP is computed on the training set (see Fig. 2). A lower bound of the predicted battery cycle life as a function of variance of $\Delta Q_{100-10}(V)$ is provided by the.025 quantile curves, while the 0.975 quantile curve provides an upper bound of the predicted battery cycle life as a function of variance of $\Delta Q_{100-10}(V)$. The median value of the predicted battery cycle life is provided by the 0.50 quantile curves. The histogram on the x-axis shows the distribution of the observations of the variance of $\Delta Q_{100-10}(V)$ in the training set. The quantile curves flatten out in the regions of sparse observations of the variance of $\Delta Q_{100-10}(V)$ in the training data, thus not providing much information. In the region of dense distribution of variance of $\Delta Q_{100-10}(V)$ in the training set, battery cycle life rapidly decreases when the variance of $\Delta Q_{100-10}(V)$ increases from 10^{-5} to 10^{-3}, which indicates that a small increase of variance of $\Delta Q_{100-10}(V)$ during discharge has a large effect on battery degradation rate. The physical meaning of the variance of $\Delta Q_{100-10}(V)$ is associated with the dependence of discharge energy dissipation on voltage. The variance of $\Delta Q_{100-10}(V)$ reflects the degree of nonuniformity in the discharge energy dissipation with voltage [5]. Thus, the larger the value of the variance of $\Delta Q_{100-10}(V)$, the larger the degree of nonuniformity in the discharge energy dissipation under galvanostatic conditions, which is consistent with the monotonic relationship between the variance of $\Delta Q_{100-10}(V)$ and cycle life [Fig. 2(a)]. The second most important feature is the minimum of $\Delta Q_{100-10}(V)$, for which battery cycle life decreases with increasing value of the minimum of $\Delta Q_{100-10}(V)$, but not to the same extent as for the variance of $\Delta Q_{100-10}(V)$.

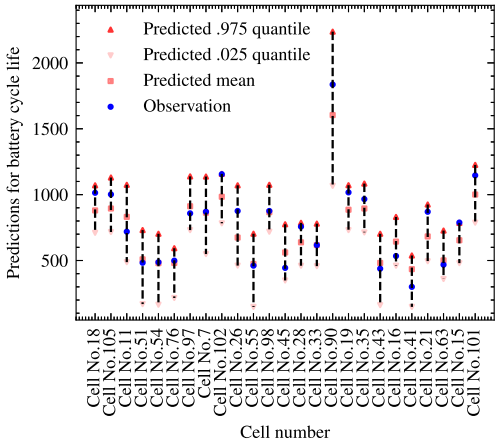


Fig. 3. 95% PIs and mean predictions by QRF.

The PIs to the left in Fig. 2(a) and on both ends (i.e., left and right) of the x-axis in Fig. 2(b) are larger than those in the middle of the x-axis. The reason is that observations are lacking on both ends of the x-axis, and therefore, the learned QRF model is not confident to make predictions on these two zones, which leads to larger prediction uncertainty represented by the width of PIs.

Furthermore, we would like to point out that data uncertainties are not considered in the present work, and it is assumed that all measurements are accurate and taken as if they were true.

The PDP results further illustrate how the variance of $\Delta Q_{100-10}(V)$ and the minimum of $\Delta Q_{100-10}(V)$ affect the predictions of battery cycle life quantitatively. Similarly, PDPs can be computed for all other features used as inputs to the QRF model on the training set.

Severson et al. [5] rationalized highly predictive features extracted from early cycle discharge voltage curves (i.e., the variance of $\Delta Q_{100-10}(V)$ and the minimum of $\Delta Q_{100-10}(V)$) by experimentally investigating degradation modes that do not lead to immediate capacity fade but are manifested in the discharge voltage curves. They found out that the loss of active material of the delithiated negative electrode contributes to a shift in the discharge voltage curve, with no change in capacity fade at early cycles. At a high number of cycles, loss of active material of the delithiated negative electrode

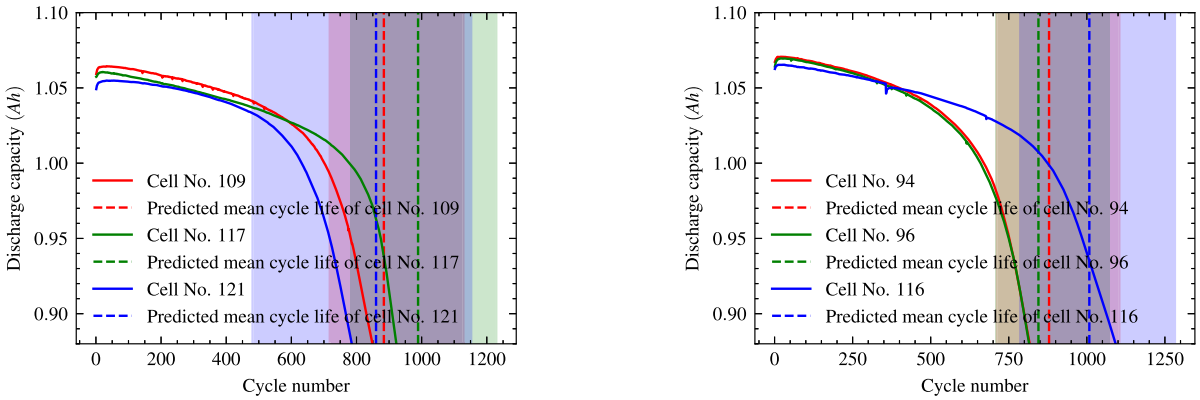


Fig. 4. Selected two groups of cells. (a) Predicted mean and 95% PI of Cell No.109, No.117, and No.121. (b) Predicted mean and 95% PI of Cell No.94, No.96, and No.116.

TABLE VI
TWO GROUPS OF CELLS WITH TWO CHARGING PROTOCOLS

Batch date	**2018-04-12**					
Charging protocols	**5.6C(36%)-4.3C**			**5.6C(19%)-4.6C**		
Cell number	**No.109**	**No.117**	**No.121**	**No.94**	**No.96**	**No.116**
Observed cycle life (cycles)	850	923	786	817	816	1093
Predicted mean cycle life (cycles)	884	989	860	879	845	1008
The lower bound of range prediction (cycles)	718	785	481	712	711	787
The upper bound of range prediction (cycles)	1128	1229	1152	1104	1069	1282
True battery cycle life of a charging protocol (cycles)	853			909		
Expected battery cycle life of a charging protocol (cycles)	911			911		
Expected battery cycle life range of a charging protocol (cycles)	434			392		

induces lithium plating, which irreversibly accelerates the capacity loss. This degradation behavior is consistent with the high feature importance of variance and minimum of $\Delta Q(V)$, as shown in Fig. 1. Throughout the literature, this degradation behavior is widely observed at various ambient temperatures (e.g., 23 °C [54], [55], 30 °C [56], and 45 °C [57]) when the negative electrode capacity is larger than that of the positive electrode, as in case of LFP cells that we used in this work. Therefore, if these two features [i.e., minimum and variance of $\Delta Q_{100-10}(V)$] are extracted from early degradation data under different ambient temperatures other than 30 °C, then the resulting learned QRF model based on this training dataset may still provide accurate cycle life point and range prediction.

VI. APPLICATION CASES

Based on predicted quantiles, PIs of battery cycle life can be constructed. To examine the prediction performance of the final QRF model by using it in a more intuitive way, we show in Fig. 3 the 95% PIs and the mean predictions made by QRF on the test set at one split. There are 24 out of 25 observed cycle life samples within the PIs resulting in 96% PICP, close to nominal 95% coverage probability. This indicates that the constructed PIs with 95% coverage probability exhibit good coverage probability of the observed cycle life values, which is required for a reliable battery cycle life range prediction.

The upper bound of the 95% PI is the 97.5% quantile prediction of cycle life, which means that battery cycle life may exceed the upper bound with a probability of around 2.5%. Correspondingly, the lower bound of the 95% PI is the 2.5% quantile prediction of battery cycle life, which means that battery cycle life may fall below the lower bound with a probability of around 2.5%. The lower bound of the PIs would facilitate conservative decision-making, while the upper bound would facilitate optimistic decision-making on battery applications.

To demonstrate how the learned QRF model facilitates decision-making when the EBCL that is calculated using (24) under two different charging protocols are very close to each other, two groups of cells are selected where each group contains three cells from the same batch date and charged with the same charging protocol (see Table VI).

The first group contains cell No.109, No.117, and No.121. All of them are from the batch date "2018-04-12" and charged with the charging protocol "5.6C(36%)-4.3C." As one can see in Fig. 4(a), three cells exhibit different capacity fade curves. The true battery cycle life of this charging protocol is calculated as the averaged observed cycle lives of three cells, i.e., 853 cycles. The second group contains cell No.94, No.96, and No.116. All of them are from the batch date "2018-04-12" and charged with the charging protocol "5.6C(19%)-4.6C." As one can see in Fig. 4(b), cell No.94 and

No.96 exhibit very similar capacity degradation curve, while the cell No.116 exhibits longer cycle life. The true battery cycle life of this charging protocol is calculated as the averaged observed cycle lives of three cells, i.e., 909 cycles. Therefore, based on the true cycle lives of two charging protocols, the charging protocol "5.6C(19%)-4.6" should be selected as the high-cycle-life charging protocol. The EBCL of the charging protocol "5.6C(36%)-4.3C," calculated using (24), is equal to 911 cycles, while the EBCL of the charging protocol "5.6C(19%)-4.6C," calculated using (24), is also equal to 911 cycles. With point prediction alone, it is very difficult to select the high-cycle-life charging protocol out of two charging protocols as they have the same expected battery cycle lives. However, with additional range prediction, the EBCLR of each charging protocol can be calculated using (25), i.e., 434 cycles for the charging protocol "5.6C(36%)-4.3C" and 392 cycles for the charging protocol "5.6C(19%)-4.6C." Therefore, the second charging protocol (i.e., "5.6C(19%)-4.6C") is chosen here as the preferred among two due to lower uncertainty (and hence lower economic risk) around the EBCL.

In the two examples given above, it can be seen how the discrepancy between the predicted mean cycle life and the observed cycle life of a cell agrees with the width of range prediction. The widths of range prediction may provide more information for decision-making (for instance, when selecting high-cycle-life charging protocol) under uncertainty associated with cycle life prediction than we get from single-point predictions alone. Interestingly, the differences in observed cycle life within each group can be quite remarkable [see Fig. 4(a) and (b)], considering that they are expected to have almost the same aging process. The reason why cells from the same batch and charged with the same charging protocols can have a very different cycle life is most likely due to production-related factors (e.g., the variance of material properties and process parameters). Apparently, in this case, the cell-to-cell variations caused by the production process are quite significant, and these variations are indeed captured by the model via the extracted features.

VII. Conclusion

In this present article, we have proposed a QRF model for Li-ion battery cycle life prediction using early degradation data. To the best of our knowledge, it is the first time that the QRF model is introduced to battery cycle life range prediction even though it has been used in predictions of drug effect, crop yield, and so on. The proposed PI evaluation criterion in LOO-XVE can be used for optimizing hyperparameters of other regression models that are capable of providing range predictions so that a high PICP value, which satisfies a preassigned nominal coverage probability, as well as low MPIW value can be available for decision-making. Two global model-agnostic methods were employed to interpret the final QRF model, and they can be easily employed also for other advanced data-driven methods. These interpretation techniques can reveal underlying battery aging mechanisms and help find features that have the highest predictive power for cycle life prediction with data-driven methods. There are, however, several improvements that can be made in future work. First, the battery dataset used for training and testing is relatively small with cells tested at an ambient temperature of 30 °C. A larger battery dataset with cells tested at ambient temperatures other than 30 °C is desired for both validating prediction performance of the proposed QRF model and effectiveness of the proposed two interpretation techniques that may reveal the underlying degradation process of cells tested at other ambient temperatures. Second, permutation importance and PDP are the two methods used in the present work to implicitly interpret the model. Additional interpretation techniques can be introduced in the future for further interpretation of other advanced data-driven models. Third, toward online application of the final QRF model on a realistic BMS, several aspects need to be investigated, including, for example, computational efficiency and memory footprint of the final QRF model for real-time embedded applications. Fourth, it would be very interesting to investigate the robustness of the QRF model with respect to both varying operating conditions and measurement noise in the data. Fifth, our proposed method does not consider calendar aging that happens during any dedicated resting periods for long time (e.g., battery storage and vehicle parking). Since the calendar aging impacts the early capacity fade of a cell, it is reasonable to expect that a QRF model trained using calendric and cyclic aging data may still show good prediction performance. However, this needs to be tested and verified in our future work. Finally, the hybrid data-driven method is also an important category, and one possibility of using a second data-driven model to extrapolate values of important features can be investigated so that even earlier or less degradation data is needed for battery cycle life prediction with high accuracy and reliability. Another possibility of using the physics-based model for plausibility check and rationalization of prediction output from the data-driven model can also be investigated. This may help interpret confidence in the predictions by the data-driven model.

Acknowledgment

The authors would like to thank Xixi Liu from the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden, for her constructive suggestions on the structure of the manuscript, and Che-Tsung Lin from the Department of Electrical Engineering, Chalmers University of Technology, for the discussion on hyperparameter optimization.

References

[1] M. A. Hannan, M. S. H. Lipu, A. Hussain, and A. Mohamed, "A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: Challenges and recommendations," *Renew. Sustain. Energy Rev.*, vol. 78, pp. 834–854, Oct. 2017.

[2] G. Zubi, R. Dufo-López, M. Carvalho, and G. Pasaoglu, "The lithium-ion battery: State of the art and future perspectives," *Renew. Sustain. Energy Rev.*, vol. 89, pp. 292–308, Jun. 2018.

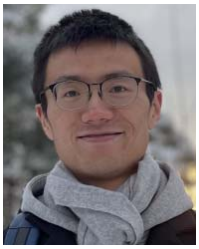
[3] J. Vetter *et el.*, "Ageing mechanisms in lithium-ion batteries," *J. Power Sources*, vol. 147, nos. 1–2, pp. 269–281, 2005.

[4] E. Martinez-Laserna et al., "Battery second life: Hype, hope or reality? A critical review of the state of the art," *Renew. Sustain. Energy Rev.*, vol. 93, pp. 701–718, Oct. 2018.

[5] A. K. Severson et al., "Data-driven prediction of battery cycle life before capacity degradation," *Nature Energy*, vol. 4, no. 5, pp. 383–391, May 2019.

[6] P. M. Attia et al., "Closed-loop optimization of fast-charging protocols for batteries with machine learning," *Nature*, vol. 578, no. 7795, pp. 397–402, Feb. 2020.
[7] P. Fermín-Cueto et al., "Identification and machine learning prediction of knee-point and knee-onset in capacity degradation curves of lithium-ion cells," *Energy AI*, vol. 1, Aug. 2020, Art. no. 100006.
[8] T. Baumhöfer, M. Brühl, S. Rothgang, and D. U. Sauer, "Production caused variation in capacity aging trend and correlation to initial cell performance," *J. Power Sources*, vol. 247, pp. 332–338, Feb. 2014.
[9] M. Huang, M. Kumar, C. Yang, and A. Soderlund, "Aging estimation of lithium-ion battery cell using an electrochemical model-based extended Kalman filter," in *Proc. AIAA Scitech Forum*, Jan. 2019, p. 0785.
[10] C. Lyu, Q. Lai, T. Ge, H. Yu, L. Wang, and N. Ma, "A lead-acid battery's remaining useful life prediction by using electrochemical model in the particle filtering framework," *Energy*, vol. 120, pp. 975–984, Feb. 2017.
[11] Y. Chang, H. Fang, and Y. Zhang, "A new hybrid method for the prediction of the remaining useful life of a lithium-ion battery," *Appl. Energy*, vol. 206, pp. 1564–1578, Nov. 2017.
[12] X. Hu, H. Yuan, C. Zou, Z. Li, and L. Zhang, "Co-estimation of state of charge and state of health for lithium-ion batteries based on fractional-order calculus," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10319–10329, Nov. 2018.
[13] L. Zhang, Z. Mu, and C. Sun, "Remaining useful life prediction for lithium-ion batteries based on exponential model and particle filter," *IEEE Access*, vol. 6, pp. 17729–17740, 2018.
[14] E. Sarasketa-Zabala, E. Martinez-Laserna, M. Berecibar, I. Gandiaga, L. M. Rodriguez-Martinez, and I. Villarreal, "Realistic lifetime prediction approach for Li-ion batteries," *Appl. Energy*, vol. 162, pp. 839–852, Jan. 2016.
[15] M. Schimpe, M. E. von Kuepach, M. Naumann, H. C. Hesse, K. Smith, and A. Jossen, "Comprehensive modeling of temperature-dependent degradation mechanisms in lithium iron phosphate batteries," *J. Electrochem. Soc.*, vol. 165, no. 2, p. A181, 2018.
[16] J. de Hoog et al., "Combined cycling and calendar capacity fade modeling of a nickel-manganese-cobalt oxide cell with real-life profile validation," *Appl. Energy*, vol. 200, pp. 47–61, Aug. 2017.
[17] C. She, L. Zhang, Z. Wang, F. Sun, P. Liu, and C. Song, "Battery state of health estimation based on incremental capacity analysis method: Synthesizing from cell-level test to real-world application," *IEEE J. Emerg. Sel. Topics Power Electron.*, early access, Sep. 14, 2022, doi: 10.1109/JESTPE.2021.3112754.
[18] B. Long, W. Xian, L. Jiang, and Z. Liu, "An improved autoregressive model by particle swarm optimization for prognostics of lithium-ion batteries," *Microelectron. Rel.*, vol. 53, no. 6, pp. 821–831, 2013.
[19] Y. Zhou and M. Huang, "Lithium-ion batteries remaining useful life prediction based on a mixture of empirical mode decomposition and ARIMA model," *Microelectron. Rel.*, vol. 65, pp. 265–273, Oct. 2016.
[20] J. Wu, C. Zhang, and Z. Chen, "An online method for lithium-ion battery remaining useful life estimation using importance sampling and neural networks," *Appl. Energy*, vol. 173, pp. 134–140, Jul. 2016.
[21] X. Hu, S. E. Li, and Y. Yang, "Advanced machine learning approach for lithium-ion battery state estimation in electric vehicles," *IEEE Trans. Transport. Electrific.*, vol. 2, no. 2, pp. 140–149, Jun. 2016.
[22] Y. Zhang, R. Xiong, H. He, and M. G. Pecht, "Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 5695–5705, Jul. 2018.
[23] A. Nuhic, T. Terzimehic, T. Soczka-Guth, M. Buchholz, and K. Dietmayer, "Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods," *J. Power Sources*, vol. 239, pp. 680–688, Oct. 2013.
[24] T. Qin, S. Zeng, and J. Guo, "Robust prognostics for state of health estimation of lithium-ion batteries based on an improved PSO–SVR model," *Microelectron. Rel.*, vol. 55, nos. 9–10, pp. 1280–1284, Aug./Sep. 2015.
[25] M. Kläs and A. M. Vollmer, "Uncertainty in machine learning applications: A practice-driven classification of uncertainty," in *Proc. Int. Conf. Comput. Saf., Rel., Secur.* Cham, Switzerland: Springer, 2018, pp. 431–438.
[26] R. R. Richardson, M. A. Osborne, and D. A. Howey, "Battery health prediction under generalized conditions using a Gaussian process transition model," *J. Energy Storage*, vol. 23, pp. 320–328, Jun. 2019.
[27] D. Liu, J. Pang, J. Zhou, Y. Peng, and M. Pecht, "Prognostics for state of health estimation of lithium-ion batteries based on combination Gaussian process functional regression," *Microelectron. Rel.*, vol. 53, no. 6, pp. 832–839, Jun. 2013.
[28] D. Wang, Q. Miao, and M. Pecht, "Prognostics of lithium-ion batteries based on relevance vectors and a conditional three-parameter capacity degradation model," *J. Power Sources*, vol. 239, pp. 253–264, Oct. 2013.
[29] D. Liu, J. Zhou, D. Pan, Y. Peng, and X. Peng, "Lithium-ion battery remaining useful life estimation with an optimized relevance vector machine algorithm with incremental learning," *Measurement*, vol. 63, pp. 143–151, Mar. 2015.
[30] L. Liao and F. Köttig, "A hybrid framework combining data-driven and model-based methods for system remaining useful life prediction," *Appl. Soft Comput.*, vol. 44, pp. 191–199, Jul. 2016.
[31] J. Wei, G. Dong, and Z. Chen, "Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5634–5643, Jul. 2018.
[32] X. Zheng, H. Wu, and Y. Chen, "Remaining useful life prediction of lithium-ion battery using a hybrid model-based filtering and data-driven approach," in *Proc. 11th Asian Control Conf. (ASCC)*, Dec. 2017, pp. 2698–2703.
[33] R. Koenker and G. Bassett, Jr., "Regression quantiles," *Econometrica, J. Econ. Soc.*, vol. 46, no. 1, pp. 33–50, 1978.
[34] R. Koenker, *Quantile Regression*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[35] A. Fisher, C. Rudin, and F. Dominici, "All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously," *J. Mach. Learn. Res.*, vol. 20, no. 177, pp. 1–81, Jan. 2019.
[36] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[37] G. Louppe, L. Wehenkel, A. Sutera, and P. Geurts, "Understanding variable importances in forests of randomized trees," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 26, 2013, pp. 431–439.
[38] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001.
[39] A. Khosravi, S. Nahavandi, and D. Creighton, "Construction of optimal prediction intervals for load forecasting problems," *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1496–1503, Aug. 2010.
[40] T. Gneiting and A. E. Raftery, "Strictly proper scoring rules, prediction, and estimation," *J. Amer. Stat. Assoc.*, vol. 102, no. 477, pp. 359–378, Jan. 2012.
[41] P. Y. Guo, Z. Cheng, and L. Yang, "A data-driven remaining capacity estimation approach for lithium-ion batteries based on charging health feature extraction," *J. Power Sources*, vol. 412, pp. 442–450, Feb. 2019.
[42] N. Williard, W. He, M. Osterman, and M. Pecht, "Comparative analysis of features for determining state of health in lithium-ion batteries," *Int. J. Prognostics Health Manag.*, vol. 4, no. 1, pp. 1–7, Oct. 2020.
[43] M. Yousefi-Azar, V. Varadharajan, L. Hamey, and U. Tupakula, "Autoencoder-based feature learning for cyber security applications," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, May 2017, pp. 3854–3861.
[44] X. Bi and H. Wang, "An enhanced high-order Boltzmann machine for feature engineering," *Eng. Appl. Artif. Intell.*, vol. 78, pp. 37–52, Feb. 2019.
[45] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Comput. Elect. Eng.*, vol. 40, no. 1, pp. 16–28, Jan. 2014.
[46] J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY, USA: Springer, 2017.
[47] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2019, pp. 2623–2631.
[48] Y. Grushka-Cockayne and V. R. R. Jose, "Combining prediction intervals in the M4 competition," *Int. J. Forecasting*, vol. 36, no. 1, pp. 178–185, Jan. 2020.
[49] Z. Reitermanova et al., "Data splitting," in *Proc. WDS*, vol. 10, 2010, pp. 31–36.
[50] J. A. Rice, *Mathematical Statistics and Data Analysis*. Boston, MA, USA: Cengage Learning, 2006.
[51] I. Bloom, L. K. Walker, J. K. Basco, D. P. Abraham, J. P. Christophersen, and C. D. Ho, "Differential voltage analyses of high-power lithium-ion cells. 4. Cells containing NMC," *J. Power Sources*, vol. 195, no. 3, pp. 877–882, 2010.
[52] C. R. Birkl, M. R. Roberts, E. McTurk, P. G. Bruce, and D. A. Howey, "Degradation diagnostics for lithium ion cells," *J. Power Sources*, vol. 341, pp. 373–386, Feb. 2017.
[53] M. Berecibar et al., "Online state of health estimation on NMC cells based on predictive analytics," *J. Power Sources*, vol. 320, pp. 239–250, Jul. 2016.

[54] D. Anseán et al., "Fast charging technique for high power $LiFePO_4$ batteries: A mechanistic analysis of aging," *J. Power Sources*, vol. 321, pp. 201–209, Jul. 2016.
[55] D. Anseán et al., "Operando lithium plating quantification and early detection of a commercial $LiFePO_4$ cell cycled under dynamic driving schedule," *J. Power Sources*, vol. 356, pp. 36–46, Jul. 2017.
[56] E. Sarasketa-Zabala, F. Aguesse, I. Villarreal, L. M. Rodriguez-Martinez, C. M. López, and P. Kubiak, "Understanding lithium inventory loss and sudden performance fade in cylindrical cells during cycling with deep-discharge steps," *J. Phys. Chem. C*, vol. 119, no. 2, pp. 896–906, Jan. 2015.
[57] M. Safari and C. Delacourt, "Aging of a commercial graphite/$LiFePO_4$ cell," *J. Electrochem. Soc.*, vol. 158, no. 10, p. A1123, 2011.



Huang Zhang received the B.Eng. degree in mechanical design, manufacturing, and automation from the Harbin Institute of Technology, Harbin, China, in 2015, the M.Sc. degree in electrical engineering from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2017, and the M.Sc. degree in energy engineering and management from the Instituto Superior Técnico, Lisbon, Portugal, in 2017. He is currently pursuing the Ph.D. degree with the Automatic Control Group, Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden.

His research interests include uncertainty estimation in machine learning, deep learning, and optimal control. His current research applications include battery state-of-health degradation diagnostics and prognostics, battery lifetime prediction, and technoeconomic analysis of second-life batteries.

Yang Su received the B.Eng. degree in nuclear engineering from the Harbin Institute of Technology, Harbin, China, in 2016, the B.Sc. degree in nuclear engineering from the Grenoble Institute of Technology, Grenoble, France, in 2016, the M.Sc. degree in electrical engineering from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2018, the M.Sc. degree in master energy environment: science technology and management from the École Polytechnique, Palaiseau, France, in 2018, and the Ph.D. degree from INRAE/AgroParisTech, Université Paris-Saclay, Thiverval-Grignon, France, in 2021.

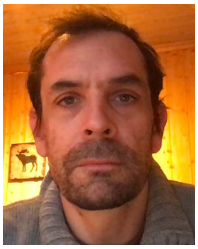
He is currently a Post-Doctoral Researcher with INRAE/UVSQ, Université Paris-Saclay. His main research areas are in big data and machine learning, climate change, and ecosystem.

Faisal Altaf received the B.Sc. degree in mechatronics engineering from the National University of Sciences and Technology (NUST), Islamabad, Pakistan, in 2004, the M.Sc. degree in electrical engineering from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2011, and the Ph.D. degree in control systems from the Chalmers University of Technology, Gothenburg, Sweden, in 2016.

He has held several positions in industry and academia since 2004, where he has been involved in a wide range of research and product development activities related to mechatronics and control systems. His main focus has been on modeling, estimation, and optimal control design for automotive battery management systems (BMS) since 2011. From 2016 to 2021, he worked as a Lead Control System Engineer at NEVS, Västergötland, Sweden, and a Principal Researcher and a Control System Architect at Volvo Group Trucks Technology, Gothenburg. He is currently working as a Chief Engineer at the Electromobility Department, Volvo Group Trucks Technology, where he is leading research and development on advanced BMS technologies for a range of heavy-duty electric vehicles. His current research interests are at the intersection of traction batteries, control engineering, power electronics, embedded systems, data analytics, and system engineering with a special focus on model- and learning-based controls for automotive applications.

Torsten Wik (Member, IEEE) received the M.Sc. degree in chemical engineering, the Licentiate of Engineering degree in control engineering, and the Ph.D. and Docent degrees from the Chalmers University of Technology, Gothenburg, Sweden, in 1994, 1996, 1999, and 2004, respectively.

From 2005 to 2007, he worked as a Senior Researcher in control system design at Volvo Group Trucks Technology, Gothenburg. In 2007, he returned to the Chalmers University of Technology as an Associate Professor, where he is currently a Full Professor and the Head of Automatic Control at the Department of Electrical Engineering. His main research areas are optimal control, process control, and environmentally motivated control applications. For a bit more than a decade, he has led a growing group of researchers on battery management systems.



Sébastien Gros received the Ph.D. degree from the École polytechnique fédérale de Lausanne (EPFL), Lausanne, Switzerland, in 2007.

After a journey by bicycle from Switzerland to the Everest base camp in full autonomy, he joined the Research and Development Group, Strathclyde University, Glasgow, U.K., focusing on wind turbine control in 2010. In 2011, he joined KU Leuven, Leuven, Belgium. He joined the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden, in 2013, where he became an Associate Professor in 2017. In 2019, he moved to the Norwegian University of Science and Technology (NTNU), Trondheim, Norway, where he is currently the Head of the Department of Engineering Cybernetics. His research interests include numerical methods, optimal control, model predictive control (MPC), Markov decision processes, reinforcement learning, learning for MPC, and stochastic and multiagent systems. His research applications include smart buildings and smart houses, energy management, smart grid, and smart traffic.