



# Early prediction of battery degradation in grid-scale battery energy storage system using extreme gradient boosting algorithm



Chico Hermanu Brillianto Apribowo <sup>a,b</sup>, Sasongko Pramono Hadi <sup>a</sup>, Francisco Danang Wijaya <sup>a</sup>, Mohammad Isnaeni Bambang Setyonegoro <sup>a</sup>, Sarjiya <sup>a,c,\*</sup>

<sup>a</sup> Department of Electrical and Information Engineering, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia

<sup>b</sup> Department of Electrical Engineering, Universitas Sebelas Maret, Surakarta, 57126, Indonesia

<sup>c</sup> Center for Energy Studies, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia

## ARTICLE INFO

### Keywords:

Battery energy storage system  
Battery degradation  
Remaining useful life  
Extreme gradient boosting algorithm  
Hyperparameter tuning

## ABSTRACT

The growth of battery energy storage systems (BESS) is caused by the variability and intermittent nature of high demand and renewable power generation at the network scale. In the context of BESS, Lithium-ion (Li-ion) battery occupies a crucial position, although it is faced with challenges related to performance battery degradation over time due to electrochemical processes. This battery degradation is a crucial factor to account for, based on its potential to diminish the efficiency and safety of electrical system equipment, thereby contributing to increased system planning costs. This implies that the health of battery needs to be diagnosed, particularly by determining remaining useful life (RUL), to avoid unexpected operational costs and ensure system safety. Therefore, this study aimed to use machine learning models, specifically extreme gradient boosting (XGBoost) algorithm, to estimate RUL, with a focus on the temperature variable, an aspect that had been previously underemphasized. Utilizing XGBoost model, along with fine-tuning its hyperparameters, proved to be a more accurate and efficient method for predicting RUL. The evaluation of the model yielded promising outcomes, with a root mean square error (RMSE) of 90.1 and a mean absolute percentage error (MAPE) of 7.5 %. Additionally, the results showed that the model could improve RUL predictions for batteries within BESS. This study significantly contributed to optimizing planning and operations for BESS, as well as developing more efficient and effective maintenance strategies.

## 1. Introduction

Approximately 80 % of the world's energy supply is derived from fossil fuels, including coal, oil, and natural gas. The combustion of these fuels is a significant contributor to greenhouse gas emissions (GHG), especially carbon dioxide (CO<sub>2</sub>), a significant driver of climate change [1]. In response, there has been a collaborative global effort to increase the utilization of renewable energy sources (RES) as a critical strategy to support Net Zero Emission (NZE) initiatives [2]. This shift towards renewable energy not only aims to address the pressing environmental concerns but also aligns with the broader international commitments, such as the 2015 Paris Agreement and the United Nations' Sustainable Development Goals (SDG), which underline the need for sustainable and environmentally friendly energy solutions [3].

Electricity is very important in achieving the world's sustainable development goals. The reliable, sustainable, and clean production of

energy has a positive impact on the global economy and development and also benefits the environment. In the industrialized era, the increase in fossil fuel combustion and its decreasing availability have compelled the industry to turn towards renewable energy (RE), including photovoltaic (PV) and wind power [4]. However, there are continuous challenges in electrical grids relying on renewable energy (RE) and ensuring the reliability of energy supply. The variations and uncertainties in RE sources, especially during peak generation periods, can lead to changes in power flow directions, causing voltage spikes in the grid [5]. Differences between energy demand and thermal generator output further increase the peak-to-average demand ratio, significantly affecting total energy costs [6,7]. In this scenario, adopting battery energy storage systems (BESS) technology serves as a practical solution to solve these challenges.

To increase the integration of RE, BESS proves to be an ideal technology, aiding power system in managing fluctuations and interruptions

\* Corresponding author. Department of Electrical and Information Engineering, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia.

E-mail address: [sarjiya@ugm.ac.id](mailto:sarjiya@ugm.ac.id) (Sarjiya).

in grid-scale and high-demand generation [8,9]. BESS is typically connected to the network through converter power electronics, facilitating rapid and flexible control. Furthermore, it finds application in various network functions, including voltage and frequency support, load shifting, transmission and distribution interruption, and peak shaving [10,11]. Large-scale BESS can store energy from renewable sources, transforming the power grid into an adaptable and flexible system [12, 13]. At the grid level, BESS efficiently manages energy by meeting both high and low-demand storage needs. Moreover, battery, being faster than many other energy storage devices, can be strategically installed in diverse locations for various purposes [14].

The phenomenon of electromigration (EM), comprising material migration induced by an electric field, has been investigated for a century, initially discovered by French scientist M. Gerardin [15]. EM has been recognized as a major contributor to the failure of integrated circuits (ICs) interconnects since the advent of packaged ICs in the late 1960s. This phenomenon significantly impacts the reliability of electronic system, particularly in modern power electronics applications such as computer-controlled machinery, telecommunications, and mechatronics [16]. Advances in nanomaterials, exemplified by environmentally-friendly synthesis methods such as Nd<sub>2</sub>Sn<sub>2</sub>O<sub>7</sub> nanostructures and the development of high-efficiency photocatalysts namely CoFe<sub>2</sub>O<sub>4</sub>@SiO<sub>2</sub>@Dy<sub>2</sub>Ce<sub>2</sub>O<sub>7</sub> nanocomposites, have made substantial progress in improving environmental sustainability and energy storage efficiency [17,18]. The synthesis of nanocrystalline neodymium zirconate and heterojunctions of WO<sub>3</sub>-SrTiO<sub>3</sub> further shows the significance of advanced materials for environmental management and efficient energy applications [19,20]. Additionally, the enhanced storage capacity of lanthanum zirconium oxide ceramics signifies innovations in energy storage technologies [21]. In the context of BESS, EM phenomenon can have profound implications. Various factors, including variations in current densities, temperature distributions, material changes, manufacturing tolerances, and existing damages, must be considered to ensure the efficient and cost-effective design of this technology. These studies collectively describe the critical role of material science in advancing the efficiency, reliability, and sustainability of modern power electronics and BESS.

BESS configurations can contain lithium-ion (Li-Ion) battery or other battery technology panels, depending on network and microgrid voltage, current, and energy requirements [22]. Although lithium battery is widely used in BESS applications in numerous industrial system and network, unexpected failures can lead to catastrophic occurrence, such as performance degradation, operational problems, and disasters. Therefore, accurately and swiftly predicting the health of lithium battery is crucial [23]. According to Cardoso et al. battery age also impacts the cost of power system planning. The study suggested that considering battery degradation limits could reduce the annual total savings for photovoltaic and storage system from 5 to 12 % [24].

When examining the degradation in BESS using lithium-ion battery, a crucial mechanism to consider is the development of the solid electrolyte interface (SEI) layer. This layer is obtained from side reactions between the electrolyte and anode, creating an exceedingly thin passivation layer on the graphite anode particles, typically a few nanometers thick. In addition, the SEI layer, formed during the initial usage cycle, contributes to permanent capacity loss [25,26]. It is crucial to be aware that the phenomenon of lithium plating (LP) also affects battery performance, primarily triggered by factors including 1) the chemical composition of the electrolyte, 2) the capacity ratio between the anode and cathode, as well as 3) operational conditions namely low temperatures and high charging rates. Another crucial fact to be aware of is that the unmodified graphite electrodes at the anode are especially susceptible to LP, and it is irreversible [27]. Previously, LP can evolve into dendritic structures that might damage the cell separator, resulting in a series of internal short circuits. This scenario leads to localized heating in the cell and can potentially cause thermal runaway, battery failure, and fire outbreak.

**Fig. 1** shows the relationship between battery degradation models and optimal energy system planning. Typically, battery degradation models serve as constraints in optimization planning, and also influence the balance of grid power distribution. Simultaneously, the penetration of renewable energy system (RES) and the annual increase in load, in accordance with national RE regulations, are considered in the system inputs. This planning affects the life cycle of BESS, which may need replacement after optimization in a specific year.

Based on the above discussion, battery degradation needs to be considered, with factors including state of charge (SOC), depth of discharge (DOD), life cycle, and temperature. In addition, the degradation impacts the total cost of planning, capacity (electricity and energy), and the optimal location of BESS. Investigation of battery degradation models is expected to enhance the flexibility of conventional and RES generators, increase system load factors, and reduce system planning costs due to intermittency.

Several elements in the usage process, including chemicals and environmental conditions, contribute to battery degradation. When battery fails to meet specific energy and capacity requirements, it is considered to have reached the end of life (EOL). Therefore, battery degradation should be considered for safety reasons, in order to prevent explosions [28]. Damage to battery system also results in performance degradation or damage to battery system equipment, thereby leading to higher repair costs. It is also important to be aware that battery degradation significantly impacts system costs, state of health (SOH), and remaining useful life (RUL), thereby serving as critical measures of battery health and performance. SOH reveals aging rates, total reduced capacity, and increased resistance, while RUL shows the current time and EOL [23]. The main factors contributing to BESS capacity degradation include nonlinear voltage, current, temperature, fast charging, slow discharge, and extended battery life [29]. Due to battery inherent nonlinearity, estimating RUL of BESS system is challenging. The machine learning RUL estimation method proves more flexible and nonlinear compared to a method based on a failure physical model.

Early prediction of SOH and RUL is critical areas of exploration aimed at improving battery reliability and lifespan. The impact of battery degradation is primarily evident in changes in battery capacity and efficiency. In general, the capacity, energy, and power of both SOH and RUL are influenced by battery age [30]. In addition, the degradation of lithium battery performance is dependent on charging behavior, discharge, and temperature fluctuations [31]. SOH diagnostics show a decline in performance, prompting preventive measures to avoid potential accidents [32]. Accurate SOH estimation is crucial for predicting RUL as a capacity indicator. Previous literature [30] showed the necessity for precise and accurate predictions of SOH and RUL to overcome limitations imposed by battery degradation, improve performance, and optimize battery operation.

**Fig. 2** shows the relationship between battery degradation modeling, RUL, and SOH. According to a previous study [30], battery degradation process model was developed in the combined framework of SOH and RUL. The framework addresses factors contributing to battery degradation and failure in SOH assessment model. Diagnosing and estimating SOH aids in modeling RUL battery by determining the time or cycles required to reach 80 %. To model battery degradation accurately, precise calculations of SOH and RUL are essential.

Methods for early prediction and estimation are broadly categorized into three groups, namely direct measurement, modeling, and data as shown in **Fig. 3**. Direct measurement method, such as open circuit voltage to calculate capacity and electronic circuit impedance spectroscopy to measure impedance, are used for direct measurement of battery capacity or impedance [34]. Model-based SOH estimation use electrochemical processes or similar electrical circuits to simulate mathematical or stochastic models representing battery degradation events. These models estimate battery capacity or resistance by determining model parameters. Various algorithms, including Kalman filtering, double sliding mode observers, and particle filtering (PF) [35],

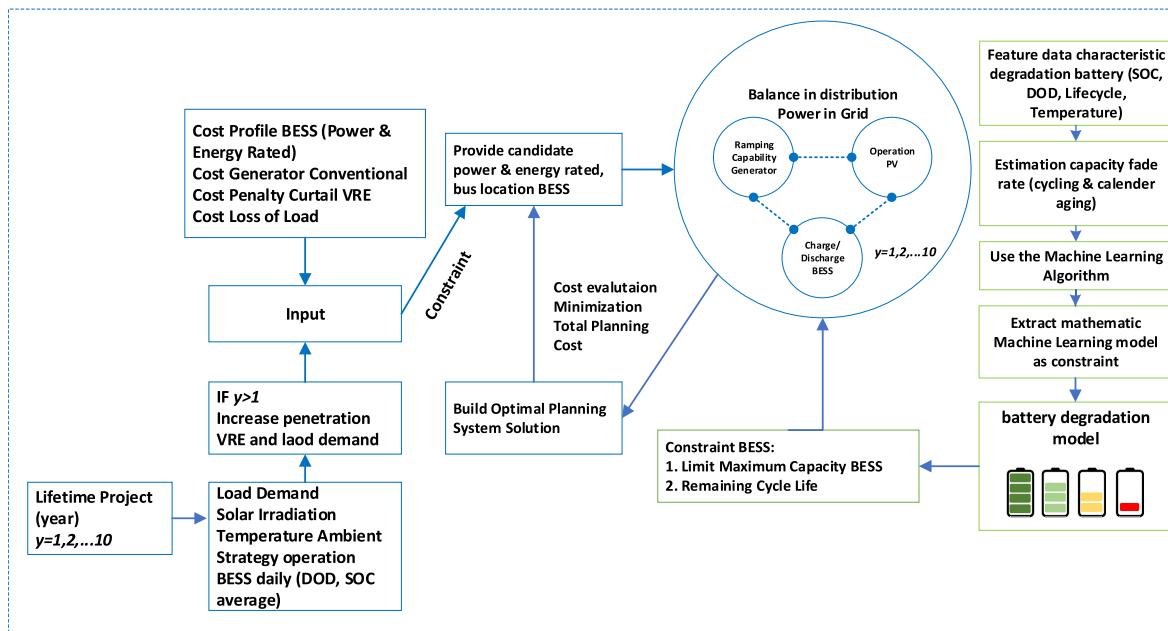


Fig. 1. Framework of battery degradation in optimal power system planning.

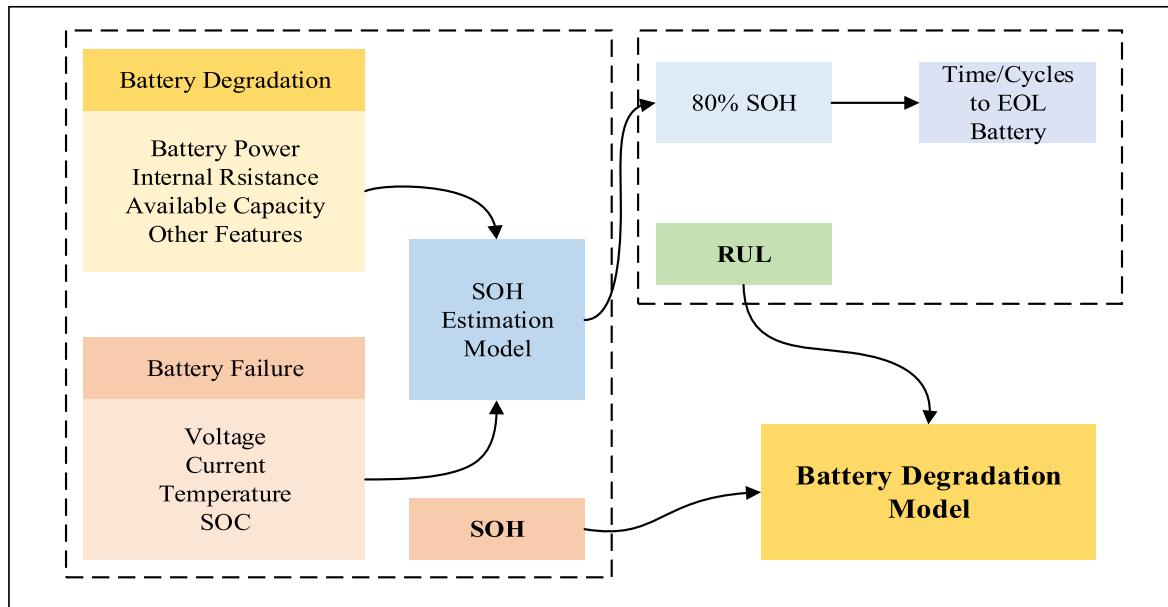


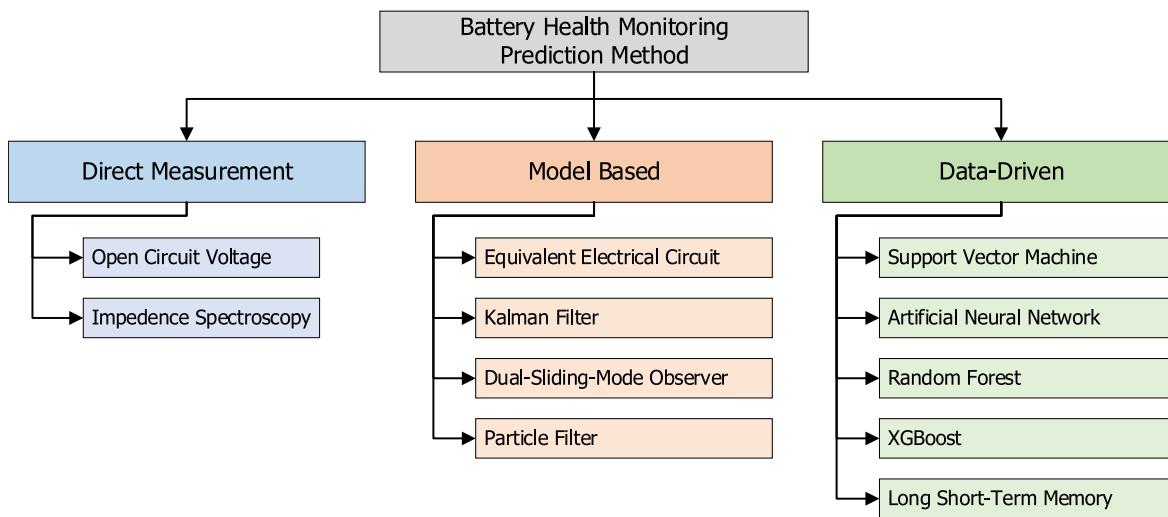
Fig. 2. Relationship battery degradation model [33].

[36], are used. In comparison to direct measurements and model-based methods, data-based methods do not require a detailed understanding of failure mechanisms. Rather, training samples are provided with specific data to extract input and output information, offering accurate predictions of degradation trends [37]. Data-driven approaches are preferred for the versatility, and health estimation and prediction techniques based on data are gaining popularity in academic and corporate settings [38]. These technologies treat battery as black box rather than relying on real mathematical models. Machine learning approaches, such as support vector machines (SVM), relevance vector machines (RVM), Bayesian approaches, artificial neural networks (ANN), and repeat networks, map the relationship between health indicators and battery SOH using historical data [39].

Data-driven approaches strongly affirm the widespread acceptance and endorsement of the data-driven strategy and battery application

analysis community [40]. This method is used to craft new models for estimating SOH [41,42], SOC [43,44], and internal resistance (IR) [45, 46]. Additionally, it is used for predicting RUL during internal resistance (IR), cycle degradation [47,48], calendar aging [49], electronic impedance spectroscopy data [50]. Leveraging machine learning technologies, data-driven methods offer adaptability and are widely accepted in academia and business [38]. Machine learning techniques were prominent data-driven tools due to the flexibility and nonlinear matching capabilities. The majority of machine learning methods shown in Table 1 lead to several conclusions.

Numerous studies estimated battery RUL using linear regression, with the linear regression elastic net model being a common method. This model can incorporate movements of non-important characteristics and reduce characteristic coefficients to enhance estimation results [64]. Some explorations, including [51,52], recorded high accuracy in



**Fig. 3.** Classification of battery RUL estimation method.

battery RUL estimation using neural network algorithms, with errors of 6.7 % and 0.5 %, respectively. However, the feature sets in these models often neglect temperature, a crucial factor in battery capacity degradation. Other studies, such as [54,55], include temperature as a feature, but cycle life and capacity are not considered. K.A. Severson, among others, uses the linear regression elastic net model for RUL estimation with a new dataset [47], yielding an error of 9.6 %. Despite its high accuracy, linear regression has the disadvantage of challenging the interpretation of intercept coefficients, which, when not handled carefully, may lead to misinterpretations. Furthermore, previous studies often do not apply hyperparameter tuning to optimize parameters for machine learning models, which can enhance the accuracy of RUL prediction.

XGBoost machine learning model is a highly flexible method that allows parallel computational processes to expedite estimation and effectively handle large datasets. In addition, XGBoost finds extensive applications in various industries, including machinery, energy system, and industrial infrastructure [65]. To address the gap in RUL estimation using LFP battery dataset, this study implements XGBoost model. The model incorporates temperature considerations into the feature to accurately predict cycle life through early RUL battery prediction. The expected result is to inspire the optimization of the service life of BESS in power system grids.

This study aims to predict BESS degradation through early RUL battery prediction, accounting for temperature. RUL estimation is generated using XGBoost machine learning algorithm. The limitations include exclusively conducting early RUL prediction using LFP battery data, evaluated with root mean square error (RMSE) and mean absolute percentage error (MAPE), and not applying this degradation model to BESS.

In summary, the novelty and contribution of the study can be outlined as follows:

- This study enhances early predictions for RUL using data-driven methods to support BESS optimization in power system. The novelty lies in the application of XGBoost machine learning algorithm and hyperparameter tuning.
- Additional feature sets include temperature when training XGBoost model for improved RUL estimation.
- Feature selection results representing battery degradation are expected to be used for BESS optimization as early RUL prediction for the remaining service life of battery.

The subsequent section of this paper includes Section 2, which

describes the relationship between BESS and battery degradation prediction. Section 3 presents the materials and methodology used in this work and Section 4 describes the case study. Moreover, Section 5 summarizes the results and discussions, and Section 6 concludes the paper.

## 2. Prognosis of degradation in BESS

BESS can simply be defined as a storage solution for energy networks. Based on usage, battery could experience reduced capacity, efficiency, and safety issues. Typically, BESS was connected to battery management system (BMS) responsible for gathering crucial battery health information. BMS collected important data, including current, voltage, power, temperature, and cycles. By using this data, battery forecasts such as SOH and RUL could be calculated. SOH and RUL were the commonly used parameters for predicting battery degradation, influenced by battery capacity, energy, and energy generation. Specifically, SOH represented the proportion of battery capacity used to calculate total aging, with a new battery typically starting at 100 %. However, RUL provided information about battery life. Replacing battery with 0 % SOH and damaged ones was crucial to ensuring system security and avoiding safety problems.

In simple terms, RUL equation can be written as shown in Equation (1):

$$RUL(t) = t - t_{eol} \quad (1)$$

Where  $t$  and  $t_{eol}$  represents the  $t$ -th cycle and the number of remaining cycles when the battery reaches EOL. Calculating battery RUL was challenging due to various variables including the current health condition, historical data, and potential failures. Typically, RUL prognosis relied on battery health, hence, a standardized framework was not considered the ideal model for estimating RUL due to data restrictions, model complexity, and technology limitations. RUL prediction methods were typically characterized as physics, mathematics, data, or hybrid models [66].

A general two-stage approach was used to determine the prognosis of BESS RUL, as shown in Fig. 4. In the offline model prediction phase, historical battery parameters were needed to develop battery lifetime model. Historical battery data, including SOC, DOD, current, voltage, charging and discharge time, internal resistance, temperature, capacity, and cycle number, were extracted and normalized. Studies typically chose combinations of these historical data to predict battery degradation. The lifetime battery model was obtained using specific methods and algorithms, and this model was used in the online model development phase for real-time battery RUL prediction. The required data for

**Table 1**  
Recent work on battery RUL method.

Ref.	Algorithm	Feature List						Error (%)
		Neural Network	Support Vector Machine	Gaussian/Bayesian	Linear Regression	XGBoost	Dataset	
[47]		✓					LFP Battery (APR18650M1A)	9.1
[51]	✓						CALCE & NASA Battery	2
[52]	✓	✓					NASA Battery	6.7
[53]	✓	✓	✓				LFP Battery (APR18650M1A)	0.5
[54]	✓	✓	✓	✓			NASA Battery	0.002
[55]	✓	✓	✓	✓	✓		NASA Battery	0.6
[56]		✓	✓	✓	✓		NASA and CALCE Battery	0.009
[57]								
[58]							NASA Battery	3
[59]							LFP Battery (APR18650M1A)	12.5
[60]							NASA Battery	3.2
[61]							NASA Battery	1.88
[62]							LFP Battery (APR18650M1A)	6.7
[63]							LFP Battery (APR18650M1A)	0.5
This Paper								

battery life prediction model were collected from sensors or BMS and normalized.

### 3. Framework estimation of RUL

#### 3.1. XGBoost

XGBoost is a class of machine learning algorithms used to model classification and regression prediction problems. Furthermore, it is a well-known tree-based ensemble model noted for the high performance and outstanding prediction capabilities. The gradient boost of XGBoost creates new models from previous model errors and adds the remaining models to make final predictions [58]. XGBoost is a boosting method, comprising a collection of decision trees whose next tree structure depends on the previous tree. Using boosting method, the model sequentially trains decision trees to capture complex data patterns, creating a single stable and robust tree by combining weak trees. During each stage of the training procedure, a new tree is constructed to reduce the loss function by compensating for the prediction residues left over from the previous simple tree. XGBoost model can be expressed as follows:

$$\hat{y}_i = \emptyset(x_i) = \sum_{k=1}^K f_k(x_i) \quad (2)$$

In Equation (2),  $\hat{y}_i = \emptyset(x_i)$  represents the combination of several decision trees (up to K trees) to obtain the final prediction  $\hat{y}_i$  for the input  $x_i$ . Each tree contributes to its  $f_k(x_i)$  prediction [57].

#### 3.2. Hyperparameter tuning

Hyperparameter tuning is a method for selecting parameters used in machine learning models to achieve optimal results. The process of hyperparameter tuning is divided into two methods, namely grid search and random search, as shown in Fig. 5.

The grid search method systematically explored various parameter combinations, assessing each model generated. A typical drawback associated with this method is the time-consuming nature of searches, particularly for grid models. Meanwhile, random search methods arbitrarily test patterns from a parameter list, offering a relatively high chance of finding optimal parameters. The time required is also less, depending on the number of repetitions, as the search process is random.

In hyperparameter tuning, one set included cross-validation (CV), a method that validated data in training and test sets. It is crucial to be aware that CV divided training sets into CV subsets known as folds. For example, in a 5-fold CV model seen in Fig. 6, each test comprised using one fold to evaluate the model, and the remaining folds for training. This process was repeated several times, and the average performance of each fold determined the final validation metric. The hyperparameter adjustment parameter, an iteration, referred to the number of repetitions in the sample parameter search. This study used the random search method due to the high possibility to find optimal parameters, reducing the time required for the machine learning model.

#### 3.3. RUL estimation

Machine learning estimation generally consisted of three steps, namely data pre-processing, training, and estimation. These steps are described with XGBoost RUL-estimation flow chart shown in Figs. 7 and 8. The data collection and processing phase aimed to clean and normalize raw battery data, preparing input and target data functions for RUL estimation. In this context, the data was divided into training, validation, and test sets.

In the training step, machine learning models were trained with optimal parameter selection for RUL estimation. Hyperparameter tuning, using the random search method and CV, helped determine appropriate parameter values. The final step comprised evaluating the

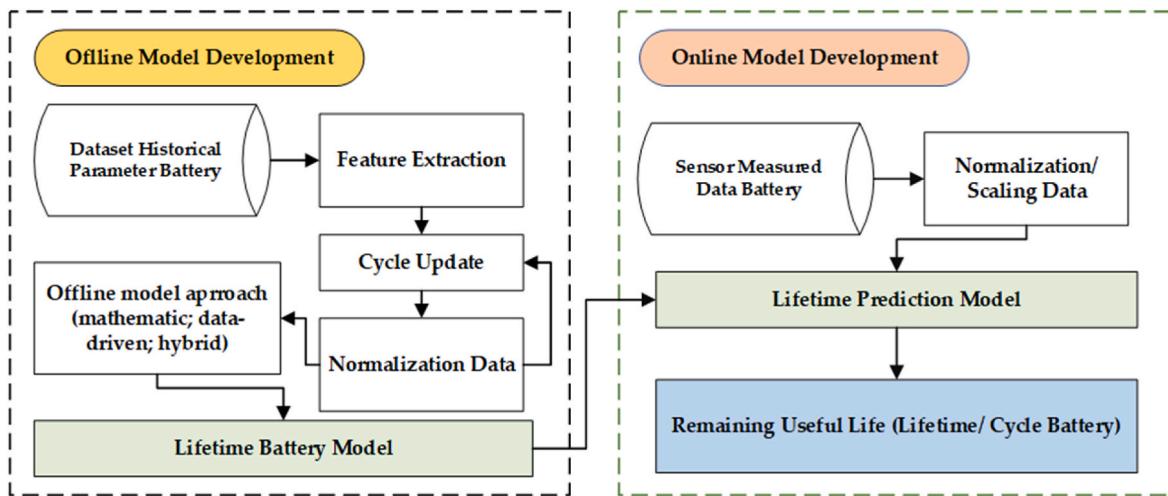


Fig. 4. Diagram of RUL prognosis on BESS [68].

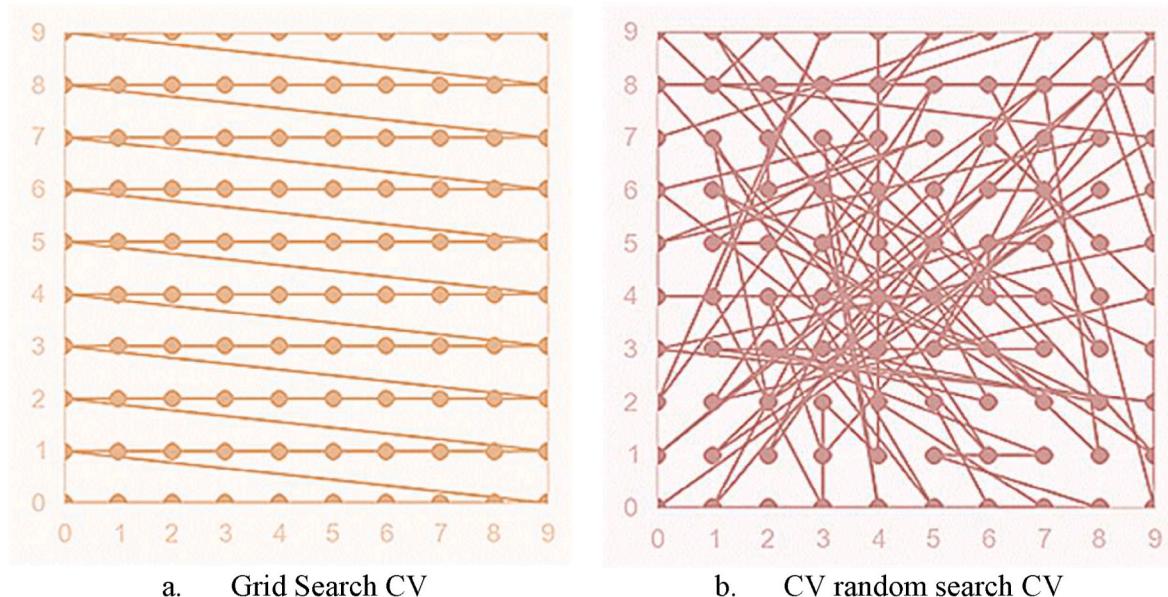


Fig. 5. Hyperparameter tuning method.

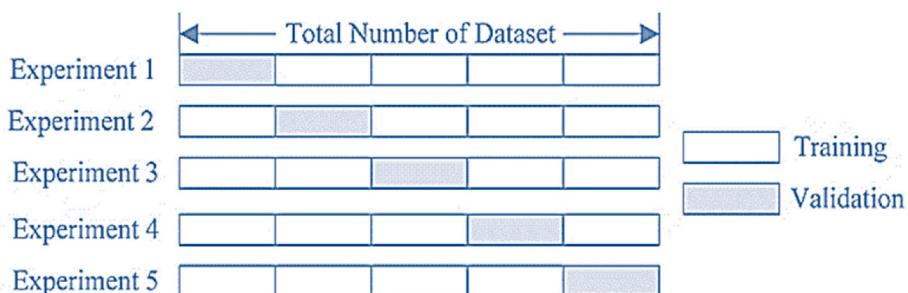


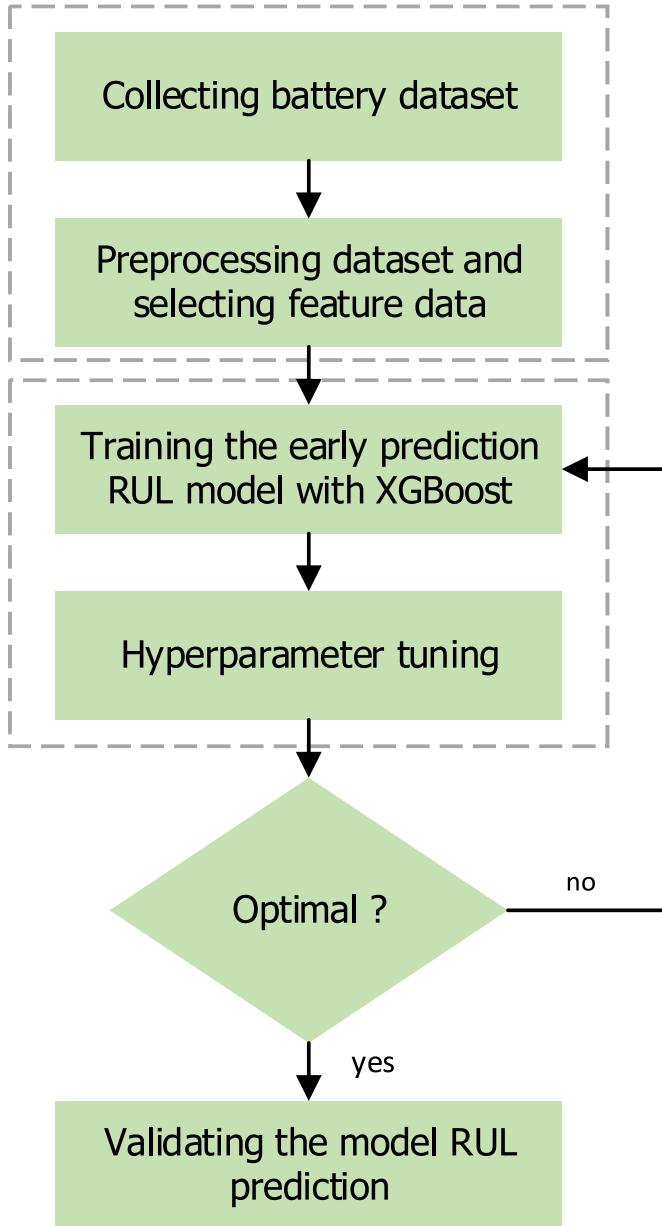
Fig. 6. Cross validation.

estimation model with XGBoost using test data.

#### 3.4. Evaluate the estimation accuracy

In this study, MAPE and RMSE were used to evaluate RUL estimation. Specifically, MAPE accurately measured prediction errors, while RMSE

determined the average square difference between predicted and actual values. Both metrics assessed the precision and accuracy of capacity prediction, with lower RMSE and MAE values indicating higher performance and better consistency between predicted and actual capacity levels [57].



**Fig. 7.** Structure of early prediction model for RUL.

$$MAPE (\%) = \frac{100}{K} \sum_{k=1}^K \frac{|l(k) - \hat{l}(k)|}{l(k)} \quad (3)$$

$$RMSE = \sqrt{\frac{1}{K} \sum_{k=1}^K (l(k) - \hat{l}(k))^2} \quad (4)$$

In Equations (3) and (4),  $l(k)$  is the actual values,  $\hat{l}(k)$  is the estimated values, and  $K$  is the number of data.

#### 4. Case study

To estimate battery RUL, 124 battery datasets were partitioned into training, validation, and test sets using the A123 system (APR18650M1A) [47]. XGBoost model enabled accurate cycle life predictions, with assessments based on standard metrics, namely RMSE and MAPE, which were common tools for evaluating performance in this context. The primary objective was to increase the accuracy of early RUL

estimation.

The dataset comprised measurements from 124 LFP battery cells with nominal capacities of 1.1 Ah and nominal voltages of 3.3 V, indicating varied charge and discharge patterns. Manufactured by A123 System (APR18650M1A), battery went through horizontal recycling [47].

The testing setup featured a cylindrical fixture on a 48-channel Arbin LBT potentiostat in a convection temperature chamber set to 30 °C. The dataset was divided into three “batches,” each containing 48 cells, identified by the batch date or the test initiation date, with each batch showing some deviation.

When constructing RUL estimation model with XGBoost, battery feature set data was clarified and normalized. Statistical analysis of battery characteristics facilitated the selection and application of the desired feature set. Ensuring equal weight for the data, the first 100 life cycles of each battery cell were considered based on statistical data regarding changes in voltage discharge capacity. All 124 battery datasets were used to model machine learning, with the data partitioned into training data (77 cells), validation data (23 cells for developing machine learning models), and test data (24 cells for model evaluation), using random partitioning. Other algorithms such as random forest, SVM, and linear regression (elastic net) were applied to compare the evaluation results of XGBoost algorithm in estimating RUL.

**Table 2** shows the scenario list of the feature model used in this study, comprising three scenarios to compare the accuracy results of RUL estimation.

- In the variance model, the feature list included only the variance of  $\Delta Q_{a-b}(V)Q_{a-b}(V)$ , depicting the difference in capacity between the two cycles. Specifically, cycles 100 and 10 were analyzed in the  $\Delta Q_{100-10}(V)$  section.
- The discharge model incorporated six features related to battery discharge capability.
- In the full model, nine features were used, offering a wide range of variations that significantly influenced battery degradation.

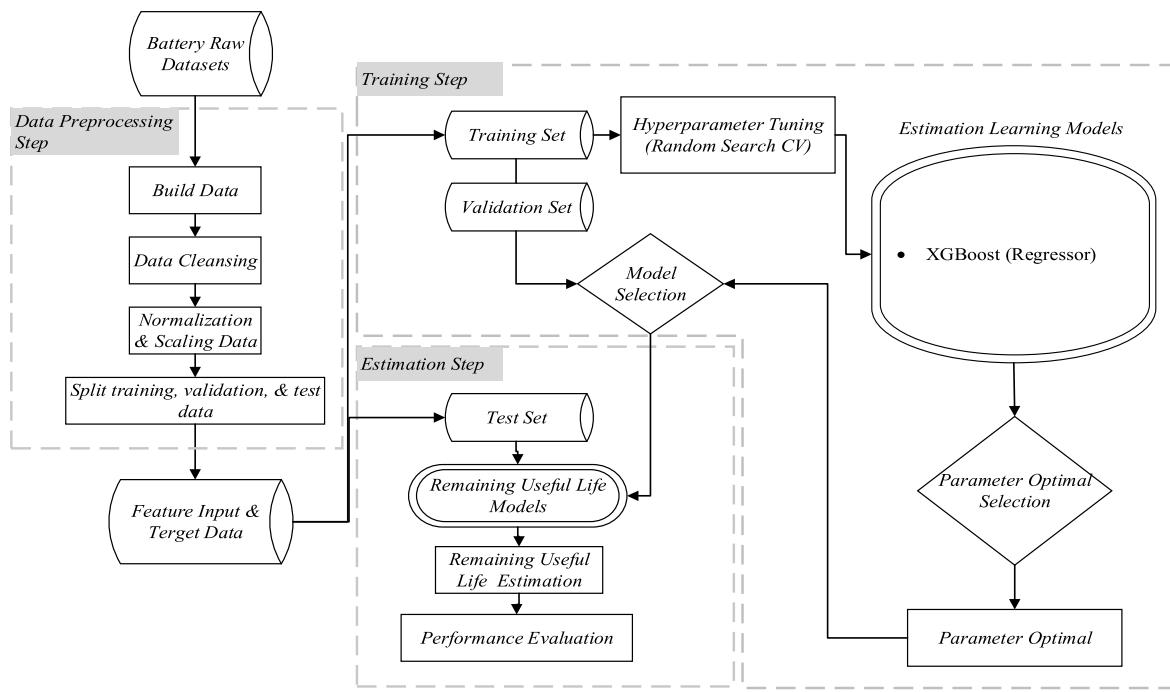
#### 4.1. Battery characteristic

**Figs. 9 and 10** show the relationship between the discharge capacity of 124 LFP cells and battery cycles, as well as changes in battery voltage curve relative to discharge capacity. As battery cycles increased, the total capacity of battery diminished. However, the initial 100 cycles were unsuitable for estimating battery degradation due to the linearity of discharge capacity. In this context, changes in overvoltage provided a clear and analyzable indicator. According to Severson et al. the transformation of  $\Delta Q(V)$  was intriguing because its derivatives offered valuable data for efficiently diagnosing degradation [47].  $\Delta Q(V)$  was a compelling data source derived from the discharge voltage capacity curve as a function of voltage in a specific cycle. This transformation was particularly interesting, as the voltage curve and its derivatives effectively contributed to degradation diagnosis, providing comprehensive insights into the electrochemical evolution of individual cells during cycling. Furthermore, it captured crucial information about battery performance and degradation patterns, enhancing the predictive capability of the model for early RUL prediction:

$$\Delta Q_{b-a}(V) = Q_b - Q_a \quad (5)$$

In Equation (5),  $\Delta Q(V)$  represented the change in the voltage curve, where  $a$  was EOL cycle and  $b$  was early life cycle. This equation could be redeveloped for input data features, incorporating statistical data on the voltage curve between changes in two cycles.

Statistical data included minimum, variance, skewness, and kurtosis. Specifically, variance measured the variability of the mean, skewness assessed the asymmetry in value distribution, and kurtosis showed the



**Fig. 8.** Flowchart estimation RUL battery with XGBoost algorithm.

**Table 2**  
Scenario feature list.

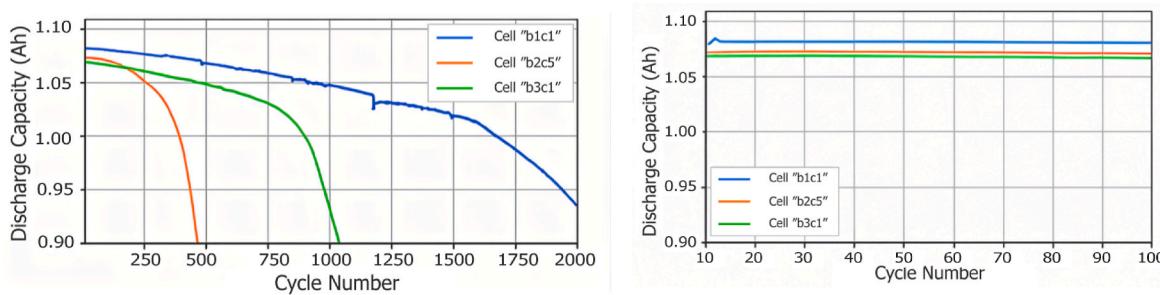
Feature Model	Type Characteristic
1. Variance	Variance of $\Delta Q_{a-b}(V)$
2. Discharge	Minimum of $\Delta Q(V)$ Variance of $\Delta Q_{a-b}(V)$ Skewness of $\Delta Q_{a-b}(V)$ Kurtosis of $\Delta Q_{a-b}(V)$ Discharge Capacity <i>Difference between max discharge capacity and capacity in cycle 2</i> Minimum of $\Delta Q(V)$
3. Full	Variance of $\Delta Q_{a-b}(V)$ <i>Slope of the Linear fit to the capacity fade curve</i> <i>Intercept of the Linear fit to capacity fade curve</i> Discharge Capacity Average charge time Minimum internal resistance <i>Difference between Internal resistance</i> Temperature

degree of sharpness. The data were selected due to the ability to reveal clear trends for prediction or estimation.

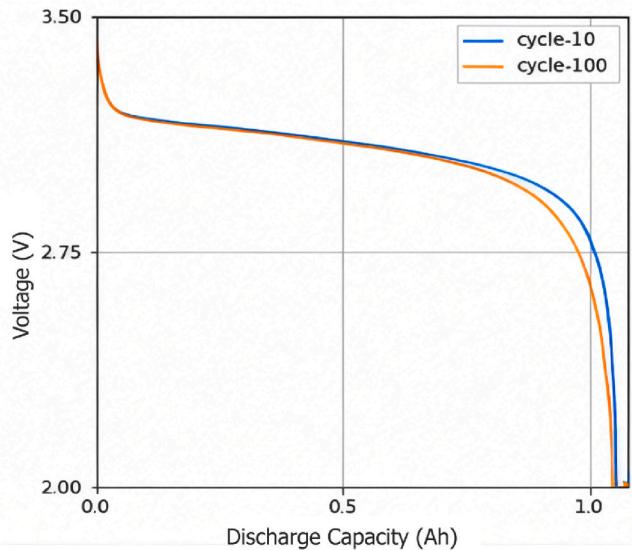
Fig. 11 shows the changes in the derivative of the discharge voltage curve capacity,  $\Delta Q(V)$ , from cycle 10 to 100. In this aspect, capacity

degradation was evident, specifically when compared with Figs. 8 and 9. For a new battery cell, the voltage peak started at 3.2 V, gradually decreasing to 2 V by cycle 2000, along with a decrease in the capacity of the discharge voltage curve  $\Delta Q(V)$ . To validate the usability of this discharge voltage curve as input data features, determining the correlation value was essential. The correlation value ( $p$ ) of the variance with battery cycle life was  $-0.89$ , showing a high correlation. The negative sign signified an opposite slope, but with this high correlation value, the discharge voltage curve data became crucial for input data features in early RUL prediction model.

Statistical data on variations in voltage discharge capacity, including discharge capacity, charge time, internal resistance, and temperature, could be used as input data characteristics. These statistics were selected because the values impacted changes in the battery as it declined, potentially enhancing the accuracy of the estimating model. This exploration used the total dataset of 124 batteries to model machine learning. Subsequently, the data were partitioned randomly into training, validation, and test data comprising 77 cells, 23 cells used for developing machine learning performance models, and 24 cells for model evaluation, respectively.



**Fig. 9.** Characteristic battery discharge.



**Fig. 10.** Voltage through capacity discharge at the 10th and 100th cycles.

## 5. Result and discussion

### 5.1. Performance evaluation

During the experiment, the machine learning model used hyperparameter tuning to estimate RUL. This hyperparameter tuning method used random search cross-validation (CV). The settings of CV fold and the number of iterations had a significant impact on the training results of RUL estimation. To optimize the evaluation results, CV fold settings for each algorithm were varied. In this study, the evaluation results of XGBoost algorithm were compared with the random forest and SVM algorithms. The specified parameters for calculating the total number of combinations of parameters were as follows :

$$NC = CV \cdot i \quad (6)$$

In Equation (6),  $NC$  is the total number of combinations,  $CV$  is the cross-

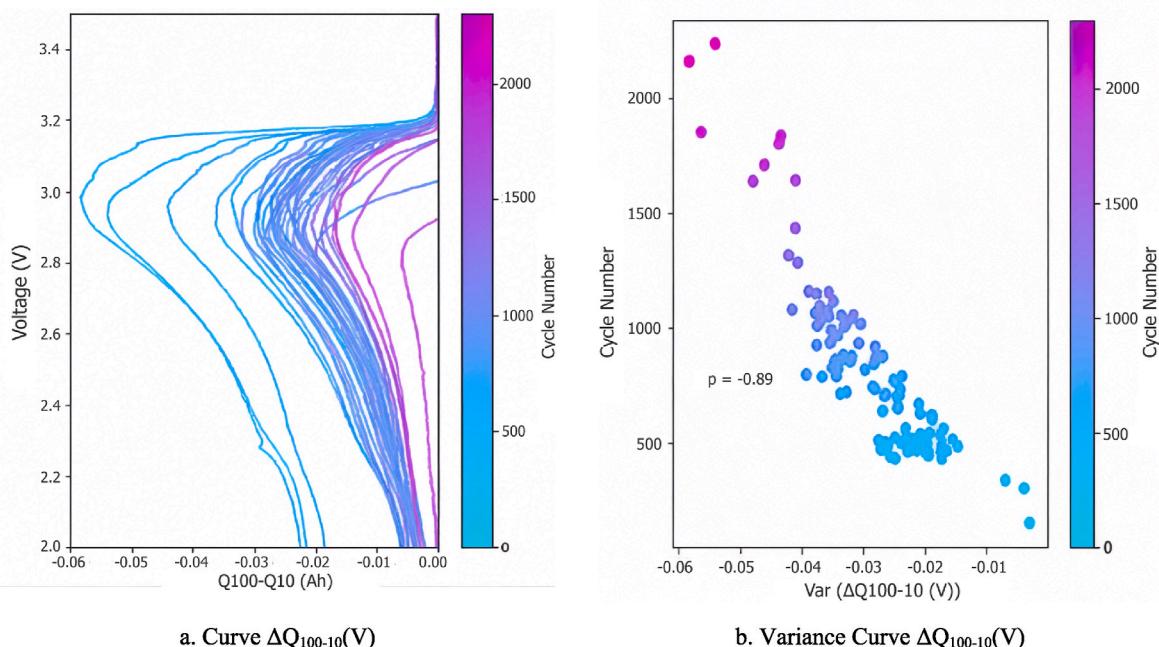
validation number, and  $i$  is the iterations number.

The random forest algorithm achieved optimal evaluation results during hyperparameter tuning with 15-fold cross-validation and 200 iterations. Due to the 15-fold cross-validation, the training data, consisting of 77 cells, was further divided into 5 datasets, which were used for validation, while the remaining 72 served as training data. The values of the 5 validation datasets varied randomly according to the fold, from fold 1 to fold 15. Moreover, the total combination of parameters determined by the random forest was 3000 combinations. Each experiment conducted through hyperparameter tuning yielded a cross-validation score accuracy, representing the accuracy value calculated for each fold in CV. The accuracy value was the average value of each fold based on the list of parameters applied in the random search method to obtain the right parameters.

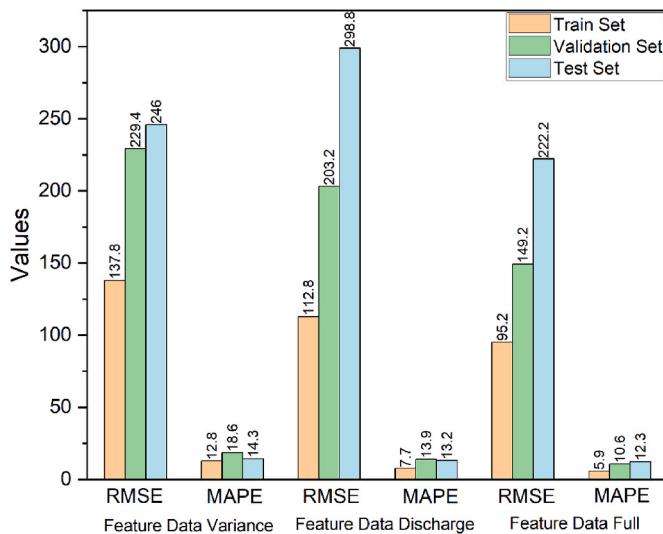
The evaluation results of RUL estimation from the random forest algorithm were shown in Fig. 12. MAPE was 13.2 % lower than the variance feature in the discharge feature but had a higher RMSE. Meanwhile, the full model feature showed the most optimal evaluation results with MAPE of 12.3 % and RMSE of 222.2. The optimal parameters to achieve the evaluation results of the full model feature were shown in Table 3, with CV accuracy value of 92 %.

The random forest algorithm comprised a substantial number of trees (495), generally beneficial for learning complex patterns but posing a potential overfitting risk. The use of 'log2' in feature selection implied consideration of the logarithm base 2 of the total number of features, a common approach for regression problems. Concerning the parameter max\_depth, a tree with considerable depth (220) could capture intricate patterns but introduced the risk of overfitting. Both min\_samples\_split and min\_samples\_leaf parameters were set to 2, allowing trees to grow with minimal constraints on the sample count at nodes. The 'bootstrap' parameter was set to 'True,' showing the use of bootstrap samples during tree construction, which was consistent with the standard Random Forest methodology.

SVM algorithm achieved optimal evaluation results during hyperparameter tuning with 13-fold cross-validation and 500 iterations, resulting in a total combination of parameters determined by SVM of 6500 combinations. The evaluation results of RUL estimation from SVM algorithm were shown in Fig. 13. Typically, MAPE was 15.6 % lower than the discharge feature in the variance feature. The discharge feature



**Fig. 11.** Curve and variance  $\Delta Q_{100-10}(V)$ .



**Fig. 12.** Evaluation results of RUL estimation in various scenarios with the random forest algorithm.

**Table 3**  
List of parameters optimal applied for random forest algorithm.

Parameter	Values
n_estimators	495
max_features	log2
max_depth	220
min_sample_split	2
min_sample_leaf	2
bootstrap	True

showed high values for both MAPE and RMSE, while the full model feature demonstrated the most optimal evaluation results with MAPE and RMSE of 9.5 % and 150.6, respectively. The optimal parameters to achieve the full model feature evaluation results were shown in [Table 4](#), with a CV accuracy value of 93 %.

SVM algorithm used a radial basis function (RBF) kernel, a common choice for handling non-linear problems. The regularization parameter, C, was set to 232, showing lower bias and higher variance, which

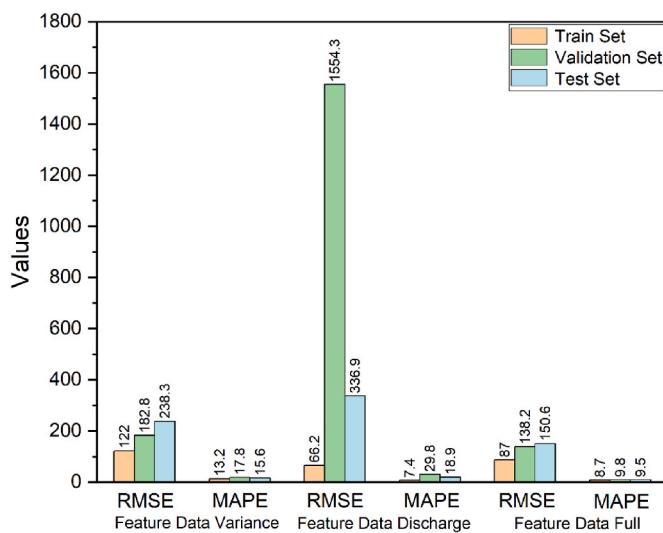
**Table 4**  
List of parameters optimal applied for SVM algorithm.

Parameter	Values
kernel	rbf
C	232
epsilon	0.026
gamma	0.838

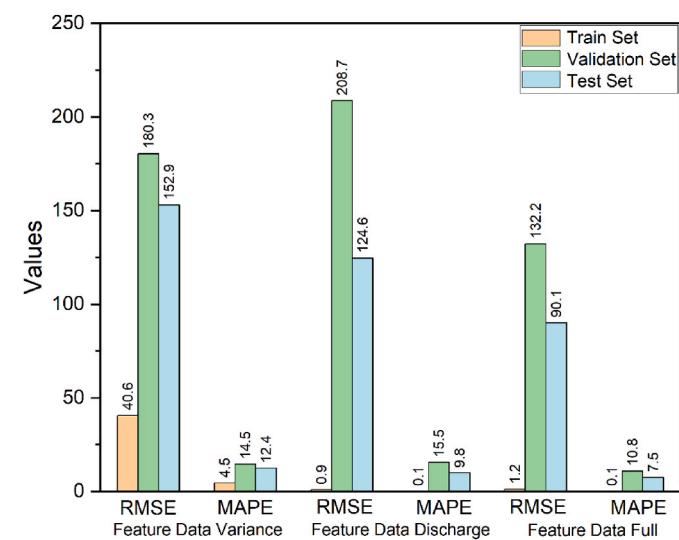
resulted in potential overfitting. The epsilon parameter specified the epsilon-tube, where no penalty was associated with points predicted in a certain distance from the actual value in the training loss function. Moreover, the specific value of 0.026 suggested fine-tuning during parameter optimization. The gamma parameter defined the influence of a single training example, with a low value signifying ‘far’ and a high value ‘close.’ The relatively high value of 0.838 could lead to a model capturing noise in the training data.

In XGBoost experiment, the model selection used was the tree booster regressor, as opposed to the linear and dart models, provided that the data features owned and applied were not in the form of classification and were not linear. Furthermore, optimal evaluation results were obtained during hyperparameter tuning with 15-fold cross-validation and 300 iterations, resulting in a total combination of parameters determined by SVM of 4500 combinations. The evaluation results of RUL estimation were shown in [Fig. 14](#). In the variance feature, MAPE was 12.4 %, while in the discharge, the value was 9.8 %. Furthermore, the full model feature showed the most optimal evaluation results with a MAPE of 7.5 % and RMSE of 90.1. The optimal parameters to achieve the full model feature evaluation results were shown in [Table 5](#) with CV accuracy value of 99 %. Although the training data had a MAPE of 0.1 %, the validation and test data showed higher values. The difference in MAPE between the validation and test data was only 3.3 %. This showed that the algorithm did not experience overfitting, which would lead to a decrease in accuracy when estimating with new data.

In XGBoost algorithm, the learning\_rate was set at 0.12, representing a moderate pace for updating the model weights. The max\_depth parameter was configured with a value of 15, allowing the model to learn complex patterns but causing a risk of overfitting. The min\_child\_weight was set to 1, showing increased sensitivity to splits that partition the data. With a subsample value of 0.6, a significant proportion (60 %) of the data was used for tree growth, aiding in preventing overfitting. The colsample\_bytree was set at 0.4, implying that only 40 %



**Fig. 13.** Evaluation results of RUL estimation in various scenarios with SVM algorithm.



**Fig. 14.** Evaluation results of RUL estimation in various scenarios with XGBoost algorithm.

**Table 5**

List of parameters optimal applied for XGBoost algorithm.

Parameter	Values
learning_rate	0.12
max_depth	15
min_child_weight	1
subsample	0.6
colsample_bytree	0.4
n_estimators	173
objectives	reg:squarederror

of the features were used per tree, serving to prevent overfitting and introduce randomness. The model considered 173 trees (n\_estimators), striking a balance between learning complex relationships and avoiding overfitting. The objective function was reg:squarederror, suitable for regression tasks aiming to minimize the difference between predicted and actual values.

Based on the three scenarios in the three applied algorithms, there was a lower average error in the discharge feature model compared to the variance model. However, variance model of SVM algorithm had a lower value, implying that a large number of combinations in hyper-parameter settings do not always directly result in optimal outcomes. Generally, in machine learning models, adding data features related to the target estimation increases the estimation accuracy. This is evidenced by the full feature data, where it had the lowest error in all three experiments with different algorithms compared to other features.

In this study, SVM and random forest algorithms, in addition to XGBoost, were used to estimate RUL for the prognosis of BESS degradation. The graph in Fig. 15 showed the observed data and estimated RUL using the full model data features for all three algorithms. It should be acknowledged that in random forest algorithm, the estimation results diverged from the diagonal line around cycle 1500, while in SVM, the deviation occurred around cycle 1700. However, XGBoost consistently showed estimation points closely related to the diagonal line across all cycles.

After examining Fig. 15, it became evident that as estimation points became closer to the diagonal line, the performance of algorithm was better, signifying higher accuracy. This observation was in line with the evaluation results shown in Table 6, where RMSE for XGBoost was smaller compared to random forest and SVM. According to MAPE results, the accuracy rates were 87.7 % for random forest, 90.5 % for SVM, and 92.5 % for XGBoost. RMSE result for XGBoost was the lowest at 90.1, showing a higher prognosis accuracy compared to other algorithms. Severson's elastic net model showed RMSE of 214 and MAPE of 9.1 %. Functioning as a linear regression model combining L1 and L2 regularization terms, elastic net performed well in the presence of

**Table 6**

Test evaluation comparison model to estimate RUL.

No	Model	RMSE	MAPE (%)
1	Linear Regression (elastic net) (Severson et al.) [48]	214	9.1
2	Random Forest	222.2	12.3
3	SVM	150.6	9.5
4	XGBoost Regressor	90.1	7.5

feature correlations or when predictors outnumbered observations but were less suited for capturing complex nonlinear relationships in the data.

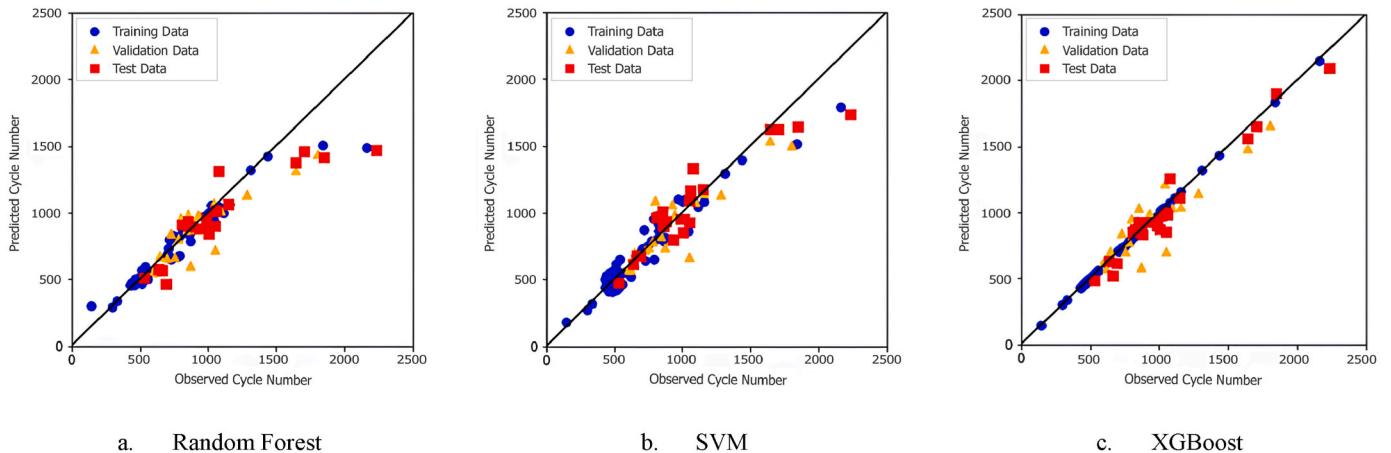
Based on the observation of RUL estimation with random forest model, the estimation results tended to underfit or deviate from the actual life cycle value. However, it showed positive result when using bagging algorithms and reduced overfitting, when battery life cycle exceeded 1500 cycles. Furthermore, in SVM, despite the advantage in high dimensions and flexible kernel options, the estimation results deviated from the actual life cycle value when the battery life cycle exceeded 1500 cycles. In this context, the trend remained accurate in RUL estimation using XGBoost.

XGBoost offered high flexibility, allowing for parallel computational processes to expedite estimation. Furthermore, it showed regularization similar to linear regression and was effective in handling large datasets. This is the reason XGBoost was compared with the linear regression model of Severson et al. using the same input feature set. A comparison of the model evaluation results with the test data showed that XGBoost showed the highest accuracy and could improve accuracy by 1.6 % in MAPE compared to Severson et al. model.

## 5.2. Discussion

Battery aging significantly impacted power system costs, with key factors including voltage, current, temperature, fast charging, slow discharge, and extended battery life. This aging led to capacity degradation in BESS, consequently increasing costs in power system planning. The main contributors to battery degradation were the increased layer of SEI on the graphite anode and LP phenomenon caused by high charge/discharge rates and temperature. Addressing this, prognosis of BESS degradation comprised estimating RUL, using data collected from BESS with BMS as the input feature set.

Capacity and voltage generally showed a significant correlation with battery degradation but during the initial 100 cycles, degradation in battery voltage was not evident. Therefore, analyzing battery degradation required the use of the discharge voltage curve capacity,  $\Delta Q(V)$ , which was crucial for gaining insights into battery degradation. Additionally, battery temperature, charge/discharge time, discharge

**Fig. 15.** Graph of observed data and estimated data of RUL.

capacity, and internal resistance factors also contributed to battery degradation.

This study applied three different algorithms for early RUL prediction, namely RF, SVM, and XGBoost. The results were comprehensively analyzed and compared with other explorations, and it was observed that RF showed an RMSE of 222.2 and a MAPE of 12.3 %. Despite being a versatile machine learning method, capable of handling regression and classification tasks through the construction of numerous decision trees, the higher RMSE and MAPE suggested that the model may not have been as effective for this specific task as some other models. Meanwhile, SVM, with RMSE of 222.2 and MAPE of 12.3 % using an RBF kernel, excelled in capturing complex nonlinear patterns but fell short of the performance achieved by XGBoost. The effectiveness of SVM is dependent on selecting the right kernel and accurately setting hyperparameters.

Severson elastic net model yielded RMSE of 214 and MAPE of 9.1 %. Despite the proficiency in handling feature correlations and scenarios with an abundance of predictors through L1 and L2 regularization terms, model was not well-suited for capturing intricate nonlinear relationships in the dataset. XGBoost showed the best performance with RMSE of 90.1 and MAPE of 7.5 %. Praised for the predictive power, XGBoost used an ensemble of decision trees to effectively capture underlying patterns and relationships in the data, as shown by the relatively low RMSE and MAPE compared to other models.

Even though metrics alone may not reveal some results, ensemble methods such as random forest and XGBoost required meticulous tuning to prevent overfitting. The well-tuned XGBoost showed effective generalization to the test data, while SVM and Random Forest models revealed higher error rates, potentially showing overfitting or underfitting. Outperformed by more complex models, Linear Regression, even with elastic net regularization, suggested that the relationship between features and RUL was probably nonlinear and intricate. The lower MAPE of XGBoost also showed robustness in terms of percentage errors, making the predictions more reliable in relative terms. Finally, battery degradation was closely related to battery life, hence, estimating battery RUL could open new opportunities in battery production, usage, and optimization in order to improve the efficiency of RE and maintain a stable operating system in the future power grid.

## 6. Conclusion

In conclusion, BESS evolved with the increasing penetration of RE, making the technology crucial for managing fluctuations and irregularities in RE generation, as well as meeting high demands on the grid scale, and supporting power system. Large-scale BESS enabled the storage of energy from renewable sources, contributing to the development of a flexible and adaptive electricity grid. Depending on the voltage, current, and power requirements of grid and microgrid applications, BESS could consist of lithium-ion battery or other battery technology panels. However, unexpected battery failures could lead to severe performance damage, operational malfunctions, and potentially catastrophic consequences.

Battery degradation was influenced by the growth of SEI layer and LP phenomena. The primary factor in battery degradation was the decreasing peak capacity of battery. In early RUL prediction, using A123 system (APR18650M1A) for battery datasets, the capacity change of the discharge voltage curve Q100-10(V) s, in battery life cycles showed a correlation ( $p$ ) of  $-0.89$ . This implied that using variable data on changes in discharge voltage capacity as input data was crucial. Furthermore, incorporating functional data with statistical information on changes in discharge capacity, internal resistance, charging time and temperature could enhance the accuracy of RUL estimation through machine learning.

XGBoost showed high flexibility and could execute computational processes in parallel, resulting in increased estimation. In addition, it possessed a linear regression-like regulation and was efficient in handling extensive datasets. It should be remembered that the utilized

algorithms in this study were random forest, SVM, and XGBoost. To improve prediction outcomes, hyperparameter tuning was implemented. Based on the evaluation results of XGBoost (regressor), the average absolute error value was 7.5 %, showing a 92.5 % accuracy, and RMSE of 90.1. After comparing the evaluation results of other models with test data, XGBoost model proved to be the most accurate, surpassing the accuracy of the elastic net (linear regression) by 1.6 %. Subsequently, the prognosis of BESS degradation through RUL estimation could be applied to optimize BESS model in the power system, considering applied battery degradation as part of those constraints.

## Credit author statement

Francisco Danang Wijaya: Supervision, Writing – review & editing. Sasongko Pramono Hadi: Supervision, Writing – review & editing. Sarjiya Sarjiya: Conceptualization, Funding acquisition, Resources, Supervision, Validation, Writing – review & editing. Mokhammad Isnaeni Bambang Setyonegoro: Supervision, Writing – review & editing. Chico Hermanu Brillianto Apribowo: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft.

## 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.

## Data availability

Data will be made available on request.

## Acknowledgments

The authors are grateful to the Center for Higher Education Funding – Ministry of Education, Culture, Research, and Technology of Indonesia (BPPT – Kemendikbudristek RI), the Indonesia Endowment Fund for Education – Ministry of Finance Republic of Indonesia (LPDP – Kemenkeu RI), and Universitas Sebelas Maret, for funding this investigation with Research Grant: Penelitian Disertasi Doktor (PDD-UNS) and contract number 228/UN27.22/PT. 01.03/2023.

## References

- [1] A. Hassan, S.Z. Ilyas, A. Jalil, Z. Ullah, Monetization of the environmental damage caused by fossil fuels, Environ. Sci. Pollut. Res. Int. 28 (17) (May 2021) 21204–21211, <https://doi.org/10.1007/S11356-020-12205-W>.
- [2] M. Murshed, Can regional trade integration facilitate renewable energy transition to ensure energy sustainability in South Asia? Energy Rep. 7 (Nov. 2021) 808–821, <https://doi.org/10.1016/J.EGRYR.2021.01.038>.
- [3] H.E. Armstrong, The Paris Agreement, Nature 127 (3207) (2015) 600–601, <https://doi.org/10.1038/127600a0>.
- [4] S.B. Wali, et al., Techno-economic assessment of a hybrid renewable energy storage system for rural community towards achieving sustainable development goals, Energy Strategy Rev. 50 (2023), 101217, <https://doi.org/10.1016/J.ESR.2023.101217>.
- [5] H. Hao, D. Wu, J. Lian, T. Yang, Optimal Coordination of Building loads and energy storage for power grid and end user services, IEEE Trans. Smart Grid 9 (5) (Sep. 2018) 4335–4345, <https://doi.org/10.1109/TSG.2017.2655083>.
- [6] F.X.N. Soelami, et al., “Energy management modeling for microgrid system in a Smart Build,” Jurnal Nasional Teknik Elektro dan Teknologi Informasi, Articles 9 (4 SE) (Dec. 2020), <https://doi.org/10.22146/jnteti.v9i4.488>.
- [7] X. Wang, F. Li, Q. Zhang, Q. Shi, J. Wang, Profit-oriented BESS siting and sizing in deregulated distribution systems, IEEE Trans. Smart Grid 14 (2) (2023) 1528–1540, <https://doi.org/10.1109/TSG.2022.3150768>.
- [8] M. Victoria, K. Zhu, T. Brown, G.B. Andresen, M. Greiner, The role of storage technologies throughout the decarbonisation of the sector-coupled European energy system, Energy Convers. Manag. 201 (Dec. 2019), 111977, <https://doi.org/10.1016/J.ENCONMAN.2019.111977>.
- [9] C.H.B. Apribowo, Sarjiya, S.P. Hadi, F.D. Wijaya, Integration of battery energy storage system to increase flexibility and penetration renewable energy in Indonesia: a brief review, ICPERE 2022 - 5th International Conference on Power

- Engineering and Renewable Energy, Proceedings (2022), <https://doi.org/10.1109/ICPERE56870.2022.10037530>.
- [10] Y. Shi, B. Xu, D. Wang, B. Zhang, Using battery storage for peak shaving and frequency regulation: joint optimization for superlinear gains, *IEEE Trans. Power Syst.* 33 (3) (May 2018) 2882–2894, <https://doi.org/10.1109/TPWRS.2017.2749512>.
  - [11] C.H.B. Apribowo, Sarjiya, S.P. Hadi, F. Danang Wijaya, Optimal sizing and siting of battery energy storage systems with retired battery, *ICT-PEP 2022 - International Conference on Technology and Policy in Energy and Electric Power: Advanced Technology for Transitioning to Sustainable Energy and Modern Power Systems*, Proceedings (2022) 327–332, <https://doi.org/10.1109/ICT-PEP57242.2022.9988958>.
  - [12] R.S. Go, F.D. Munoz, J.P. Watson, Assessing the economic value of co-optimized grid-scale energy storage investments in supporting high renewable portfolio standards, *Appl. Energy* 183 (Dec. 2016) 902–913, <https://doi.org/10.1016/J.APENERGY.2016.08.134>.
  - [13] C.H.B. Apribowo, S.P. Hadi, F.D. Wijaya, M. Isnaeni, B. Setyonegoro, Optimal Allocation of Vanadium Redox Flow Battery Storage Systems with Integrated Variable Renewable Energy," *2023 15th International Conference on Information Technology and Electrical Engineering, ICITEE*, 2023, pp. 375–380, <https://doi.org/10.1109/ICITEE59582.2023.10317686>.
  - [14] V. Krishnan, T. Das, Slow dynamics model of compressed air energy storage and battery storage technologies for automatic generation control, *Energy Systems* 7 (2) (2015) 271–295, <https://doi.org/10.1007/S12667-015-0157-5>, 2015 7:2.
  - [15] P.S. Ho, T. Kwo, Electromigration in metals, *Rep. Prog. Phys.* 52 (3) (1989) 301, <https://doi.org/10.1088/0034-4885/52/3/002>. Mar.
  - [16] I. Mahariq, S. Berjokzina, Experimental realization of electromigration at high power for copper wires, *Journal of Energy Systems* 3 (4) (2019) 158–167, <https://doi.org/10.30521/jes.616982>.
  - [17] S. Zinatloo-Ajabshir, M. Salavati-Niasari, Facile synthesis of nanocrystalline neodymium zirconate for highly efficient photodegradation of organic dyes, *J. Mol. Liq.* 243 (2017) 219–226, <https://doi.org/10.1016/j.molliq.2017.08.050>.
  - [18] S. Zinatloo-Ajabshir, M. Salavati-Niasari, Preparation of magnetically retrievable CoFe2O4@SiO2@Dy2Ce207 nanocomposites as novel photocatalyst for highly efficient degradation of organic contaminants, *Compos. B Eng.* 174 (May) (2019), 106930, <https://doi.org/10.1016/j.compositesb.2019.106930>.
  - [19] S. Zinatloo-Ajabshir, M.S. Morassaei, M. Salavati-Niasari, Eco-friendly synthesis of Nd 2 Sn 2 O 7-based nanostructure materials using grape juice as green fuel as photocatalyst for the degradation of erythrosine, *Compos. B Eng.* 167 (2019) 643–653, <https://doi.org/10.1016/j.compositesb.2019.03.045>. February.
  - [20] G. Hosseini-zadeh, S.M. Sajjadi, L. Mostafa, A. Yousefi, R.H. Vafaei, S. Zinatloo-Ajabshir, Synthesis of novel direct Z-scheme heterojunction photocatalyst from WO3 nanoparticles and SrTiO3 nanoparticles with abundant oxygen vacancies, *Surface. Interfac.* 42 (2023), 103349, <https://doi.org/10.1016/j.surfint.2023.103349>.
  - [21] S. Zinatloo-Ajabshir, M.H. Esfahani, C.A. Marjerrison, J. Greedan, M. Behzad, Enhanced electrochemical hydrogen storage performance of lanthanum zirconium oxide ceramic microstructures synthesized by a simple approach, *Ceram. Int.* 49 (23) (2023) 37415–37422, <https://doi.org/10.1016/j.ceramint.2023.09.067>.
  - [22] D. Jimenez, D. Ortiz-Villalba, A. Perez, M.E. Orchard, Lithium-ion battery degradation assessment in microgrids, *IEEE International Autumn Meeting on Power, Electronics and Computing, ROPEC*, 2018, <https://doi.org/10.1109/ROPEC.2018.8661410>, 2018, Mar. 2019.
  - [23] B. Gou, Y. Xu, X. Feng, State-of-Health estimation and remaining-useful-life prediction for lithium-ion battery using a hybrid data-driven method, *IEEE Trans. Veh. Technol.* 69 (10) (Oct. 2020) 10854–10867, <https://doi.org/10.1109/TVT.2020.3014932>.
  - [24] G. Cardoso, T. Brouhard, N. DeForest, D. Wang, M. Heleno, L. Kotzur, Battery aging in multi-energy microgrid design using mixed integer linear programming, *Appl. Energy* 231 (Dec. 2018) 1059–1069, <https://doi.org/10.1016/J.APENERGY.2018.09.185>.
  - [25] J. Urquiza, P. Singh, A review of health estimation methods for Lithium-ion batteries in Electric Vehicles and their relevance for Battery Energy Storage Systems, *J. Energy Storage* 73 (Dec. 2023), 109194, <https://doi.org/10.1016/J.EST.2023.109194>.
  - [26] M.B. Pinson, M.Z. Bazant, Theory of SEI formation in rechargeable batteries: capacity fade, accelerated aging and lifetime prediction, *J. Electrochem. Soc.* 160 (2) (Dec. 2013) A243–A250, <https://doi.org/10.1149/2.044302JES/XML>.
  - [27] V. Agubra, J. Fergus, Lithium ion battery anode aging mechanisms, *Materials* 6 (4) (2013) 1310–1325, <https://doi.org/10.3390/MA6041310>. Mar. 2013.
  - [28] Y. Choi, S. Ryu, K. Park, H. Kim, Machine learning-based lithium-ion battery capacity estimation exploiting multi-channel charging profiles, *IEEE Access* 7 (2019) 75143–75152, <https://doi.org/10.1109/ACCESS.2019.2920932>.
  - [29] H. Rauf, M. Khalid, N. Arshad, Machine learning in state of health and remaining useful life estimation: theoretical and technological development in battery degradation modelling, *Renew. Sustain. Energy Rev.* 156 (Mar. 2022), 111903, <https://doi.org/10.1016/J.RSER.2021.111903>.
  - [30] H. Rauf, M. Khalid, N. Arshad, Machine learning in state of health and remaining useful life estimation: theoretical and technological development in battery degradation modelling, *Renew. Sustain. Energy Rev.* 156 (Mar. 2022), 111903, <https://doi.org/10.1016/J.RSER.2021.111903>.
  - [31] H. Meng, Y.F. Li, A review on prognostics and health management (PHM) methods of lithium-ion batteries, *Renew. Sustain. Energy Rev.* 116 (Dec. 2019), 109405, <https://doi.org/10.1016/J.RSER.2019.109405>.
  - [32] D. Liu, J. Pang, J. Zhou, Y. Peng, M. Pecht, Prognostics for state of health estimation of lithium-ion batteries based on combination Gaussian process functional regression, *Microelectron. Reliab.* 53 (6) (Jun. 2013) 832–839, <https://doi.org/10.1016/J.MICROREL.2013.03.010>.
  - [33] H. Rauf, M. Khalid, N. Arshad, Machine learning in state of health and remaining useful life estimation: theoretical and technological development in battery degradation modelling, *Renew. Sustain. Energy Rev.* 156 (Mar. 2022), 111903, <https://doi.org/10.1016/J.RSER.2021.111903>.
  - [34] S.M.R. Islam, S.Y. Park, Precise online electrochemical impedance spectroscopy strategies for Li-ion batteries, *IEEE Trans. Ind. Appl.* 56 (2) (Mar. 2020) 1661–1669, <https://doi.org/10.1109/TIA.2019.2958555>.
  - [35] X. Hu, H. Yuan, C. Zou, Z. Li, 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.* 67 (11) (Nov. 2018) 10319–10329, <https://doi.org/10.1109/TVT.2018.2865664>.
  - [36] M. Cacciato, G. Nobile, G. Scarella, G. Scelba, Real-time model-based estimation of SOC and SOH for energy storage systems, *IEEE Trans. Power Electron.* 32 (1) (Jan. 2017) 794–803, <https://doi.org/10.1109/TPEL.2016.2535321>.
  - [37] J. Zhang, J. Lee, A review on prognostics and health monitoring of Li-ion battery, *J. Power Sources* 196 (15) (2011) 6007–6014, <https://doi.org/10.1016/j.jpowsour.2011.03.101>.
  - [38] X. Hu, J. Jiang, D. Cao, B. Egardt, Battery health prognosis for electric vehicles using sample entropy and sparse Bayesian predictive modeling, *IEEE Trans. Ind. Electron.* 63 (4) (2016) 2645–2656, <https://doi.org/10.1109/TIE.2015.2461523>.
  - [39] H. Chaoui, C.C. Ibe-Ekeocha, State of charge and state of health estimation for lithium batteries using recurrent neural networks, *IEEE Trans. Veh. Technol.* 66 (10) (Oct. 2017) 8773–8783, <https://doi.org/10.1109/TVT.2017.2715333>.
  - [40] M. Aykol, et al., Perspective—combining physics and machine learning to predict battery lifetime, *J. Electrochem. Soc.* 168 (3) (2021), 030525, <https://doi.org/10.1149/1945-7111/abec55>.
  - [41] D. Roman, S. Saxena, V. Robu, M. Pecht, D. Flynn, Machine learning pipeline for battery state-of-health estimation 3 (2021), <https://doi.org/10.1038/s42256-021-00312-3>. Nat Mach Intell.
  - [42] T. Qin, S. Zeng, J. Guo, Robust prognostics for state of health estimation of lithium-ion batteries based on an improved PSO-SVR model, *Microelectron. Reliab.* 55 (9–10) (2015) 1280–1284, <https://doi.org/10.1016/j.microrel.2015.06.133>.
  - [43] E. Chemali, P.J. Kollmeyer, M. Preindl, A. Emadi, State-of-charge estimation of Li-ion batteries using deep neural networks: a machine learning approach, *J. Power Sources* 400 (2018) 242–255, <https://doi.org/10.1016/j.jpowsour.2018.06.104>.
  - [44] Y.-W. Chang, Performance measurement of the fourth party logistics providers, *iBusiness* 5 (2) (2013) 7–10, <https://doi.org/10.4236/ib.2013.52b002>.
  - [45] K. Liang, Z. Zhang, P. Liu, Z. Wang, S. Jiang, Data-driven ohmic resistance estimation of battery packs for electric vehicles, *Energies* 12 (24) (2019) 24, <https://doi.org/10.3390/en12244772>.
  - [46] J. Remmlinger, M. Buchholz, M. Meiler, P. Bermreuter, K. Dietmayer, State-of-health monitoring of lithium-ion batteries in electric vehicles by on-board internal resistance estimation, *J. Power Sources* 196 (12) (2011) 5357–5363, <https://doi.org/10.1016/j.jpowsour.2010.08.035>.
  - [47] K.A. Severson, et al., Data-driven prediction of battery cycle life before capacity degradation, *Nat. Energy* 4 (5) (2019) 383–391, <https://doi.org/10.1038/s41560-019-0356-8>.
  - [48] X. Tang, K. Liu, X. Wang, F. Gao, J. MacRo, W.D. Widanage, Model migration neural network for predicting battery aging trajectories, *IEEE Transactions on Transportation Electrification* 6 (2) (2020) 363–374, <https://doi.org/10.1109/TTE.2020.2979547>.
  - [49] K. Liu, X. Hu, Z. Wei, Y. Li, Y. Jiang, Modified Gaussian process regression models for cyclic capacity prediction of lithium-ion batteries, *IEEE Transactions on Transportation Electrification* 5 (4) (2019) 1225–1236, <https://doi.org/10.1109/TTE.2019.2944802>.
  - [50] Y. Zhang, W. Zhang, Y. Feng, J. Ma, Promoted CO2 electroreduction over indium-doped SnP3: a computational study, *J. Energy Chem.* 48 (2020) 1–6, <https://doi.org/10.1016/j.jechem.2019.12.025>.
  - [51] Y. Wu, Y.W. Li, Remaining useful life prediction of lithium-ion batteries using neural network and bat-based particle filter, *IEEE Access* 7 (2019) 54843–54854.
  - [52] L. Ren, J. Dong, X. Wang, Z. Meng, L. Zhao, M.J. Deen, A data-driven auto-CNN-LSTM prediction model for lithium-ion battery remaining useful life, *IEEE Trans. Ind. Inf.* 17 (5) (May 2021) 3478–3487, <https://doi.org/10.1109/TII.2020.3008223>.
  - [53] J. Xia, et al., Historical data-independent remaining useful life prediction method based on dual-input deep learning neural network, *J. Energy Storage* 72 (Nov. 2023), 108427, <https://doi.org/10.1016/j.est.2023.108427>.
  - [54] B. Chinomona, C. Chung, L.K. Chang, W.C. Su, M.C. Tsai, Long short-term memory approach to estimate battery remaining useful life using partial data, *IEEE Access* 8 (2020) 165419–165431, <https://doi.org/10.1109/ACCESS.2020.3022505>.
  - [55] K. Liu, Y. Shang, Q. Ouyang, W.D. Widanage, A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery, *IEEE Trans. Ind. Electron.* 68 (4) (2021) 3170–3180, <https://doi.org/10.1109/TIE.2020.2973876>.
  - [56] Y. Zhang, R. Xiong, H. He, Z. Liu, “A LSTM-RNN method for the lithium-ion battery remaining useful life prediction,” *2017 Prognostics and System Health Management Conference, PHM-Harbin 2017 - Proceedings* (2017) 1–4, <https://doi.org/10.1109/PHM.2017.8079316>.
  - [57] S. Safari, Y.-C. Byun, S. Ko, A Novel Approach for Predicting Remaining Useful Life and Capacity Fade in Lithium-Ion Batteries Using Hybrid Machine Learning, *IEEE Access*, 2023, <https://doi.org/10.1109/ACCESS.2023.3329508>, 1–1.
  - [58] P. Jun, et al., A Data-Driven RUL Prediction Method Enhanced by Identified Degradation Model for Lithium-Ion Battery of EVs, vol. 2019, ECCE. IEEE Energy Conversion Congress and Exposition, 2019.

- [59] S. Paradis, M. Whitmeyer, Pay Attention: Leveraging Sequence Models to Predict the Useful Life of Batteries, vol. 2019, 2019 [Online]. Available: <http://arxiv.org/abs/1910.01347>.
- [60] J. Liu, Z. Chen, Remaining useful life prediction of lithium-ion batteries based on health indicator and Gaussian process regression model, *IEEE Access* 7 (2019) 39474–39484, <https://doi.org/10.1109/ACCESS.2019.2905740>.
- [61] K. Park, Y. Choi, W.J. Choi, H.Y. Ryu, H. Kim, LSTM-based battery remaining useful life prediction with multi-channel charging profiles, *IEEE Access* 8 (2020) 20786–20798, <https://doi.org/10.1109/ACCESS.2020.2968939>.
- [62] L. Ren, L. Zhao, S. Hong, S. Zhao, H. Wang, L. Zhang, Remaining useful life prediction for lithium-ion battery: a deep learning approach, *IEEE Access* 6 (2018) 50587–50598, <https://doi.org/10.1109/ACCESS.2018.2858856>.
- [63] D. Lin, et al., Early prediction of remaining useful life for grid-scale battery energy storage system, *J. Energy Eng.* 147 (6) (2021) 1–8, [https://doi.org/10.1061/\(asce\)ey.1943-7897.0000800](https://doi.org/10.1061/(asce)ey.1943-7897.0000800).
- [64] M.-F. Ng, J. Zhao, Q. Yan, G.J. Conduit, Z.W. Seh, Predicting the state of charge and health of batteries using data-driven machine learning, *Nat. Mach. Intell.* 2 (2020) 161–170, <https://doi.org/10.1038/s42256-020-0191-4>.
- [65] A. Prakash, J. Thangaraj, S. Roy, S. Srivastav, J.K. Mishra, Model-aware XGBoost method towards optimum performance of flexible distributed Raman amplifier, *IEEE Photon. J.* 15 (4) (2023), <https://doi.org/10.1109/JPHOT.2023.3286272>.
- [66] C.H.B. Apribowo, S. Sarjiya, S.P. Hadi, F.D. Wijaya, Optimal planning of battery energy storage systems by considering battery degradation due to ambient temperature: a review, challenges, and new perspective, *Batteries* 8 (12) (2022) 290, <https://doi.org/10.3390/BATTERIES8120290>. *Page 290*.