BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 73(5), 2025, Article number: e155033
DOI: 10.24425/bpasts.2025.155033

Interpretable machine learning for battery health insights: A LIME and SHAP-based study on EIS-derived features

Taha ETEM*

Cankiri Karatekin University, Faculty of Engineering, Computer Engineering, Cankiri, Turkiye

Abstract. This paper presents a comparative study of interpretable machine learning methods for lithium-ion battery state of health (SOH) estimation using features derived from electrochemical impedance spectroscopy (EIS) and distribution of relaxation times (DRT) analysis. Four DRT peak-area features capturing diffusion (A1), charge-transfer resistance (A2), solid-electrolyte interphase impedance (A3), and ohmic resistance (A4). These serve as inputs to five regression models: linear regression, support vector regression, k-nearest neighbors, random forest, and gradient boosting. All models achieve near-perfect predictive accuracy, demonstrating that these EIS-derived features reliably encode SOH information. To bridge the gap between high performance and transparency, we apply Local Interpretable Model-Agnostic Explanations (LIME) and SHapley Additive exPlanations (SHAP) to quantify both local and global feature importance. Our interpretability analysis reveals a unanimous consensus: the SEI-related feature (A3) dominates SOH predictions, with the charge-transfer feature (A2) as a secondary contributor, while diffusion (A1) and ohmic (A4) features play lesser roles. Cross-model and cross-method agreement underscores the physical validity of these insights and paves the way for integrating transparent, trustworthy SOH estimators into safety-critical battery management systems.

Keywords: lithium-ion batteries; EIS; distribution of relaxation times; SHAP; LIME.

1. INTRODUCTION

Rechargeable lithium-ion batteries have become indispensable in modern life, powering portable electronics, electric vehicles (EVs), and grid storage systems [1]. Their performance and safety are critical, yet batteries inevitably degrade over time [2]. Aging processes lead to capacity loss, increased internal resistance, and can even pose safety risks such as thermal runaway if not properly managed [3]. To monitor battery health and mitigate failures, researchers increasingly rely on advanced diagnostics. Electrochemical impedance spectroscopy (EIS) is a powerful nondestructive tool that probes battery internal state by measuring impedance across a wide frequency range [4]. EIS data contain rich information on material properties, interfacial reactions, and degradation phenomena inside the cell [5]. However, raw EIS spectra are high-dimensional and complex, making interpretation challenging when multiple electrochemical processes overlap. An emerging approach to untangle EIS data is the distribution of relaxation times (DRT) analysis, which deconvolutes the spectrum into distinguishable peaks corresponding to different electrochemical processes. The position, height, and area of each DRT peak provide quantitative insight into specific reaction kinetics and resistive losses [6]. Notably, as the battery ages, DRT-derived features offer a window into the progression of degradation?. Recent studies have leveraged such DRT features for battery state-of-health (SOH) estimation [7]. For example, using peak parameters in Gaussian process models to predict capacity with only 1–2% error [8]. These advances underscore that EIS and DRT can furnish valuable indicators of battery health.

Meanwhile, data-driven machine learning (ML) techniques are rapidly gaining prominence in battery diagnostics [9]. Traditional electrochemical modeling of every degradation mechanism is often intractable, so researchers turn to statistical ML models to learn mappings from measured data to battery SOH. Simple linear and support vector regressions to ensemble methods and deep neural networks are applied to capture the complex patterns of battery aging [10]. These models can predict remaining capacity or resistance growth without an explicit physical model, enabling fast data-driven health estimation [11]. However, a key drawback is that many high-performing ML models behave as “black boxes” offering little transparency about how they reach their predictions [12]. The lack of interpretability is problematic in critical energy storage applications: stakeholders need to trust and understand ML-driven decisions, especially to adopt them in safety-critical domains like EV battery management. Recognizing this, the field has seen a growing push for explainable machine learning techniques?. Explainable AI (XAI) seeks to make the decision process of the model transparent and interpretable [13]. In particular, model-agnostic explanation methods that can be applied to any ML model, such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME). These tools have the potential to elucidate which input features most influence the SOH predictions, thereby converting black-box models into ones that yield human-understandable insights.

*e-mail: tahaetem@karatekin.edu.tr

Manuscript submitted 2025-04-25, revised 2025-06-12, initially accepted for publication 2025-06-19, published in August 2025.

1.1. Problem statement

Despite the progress in data-driven battery health modeling, most studies to date emphasize predictive accuracy over interpretability. Typical battery SOH models do not explain *why* a prediction is made, leaving battery engineers uncertain which impedance features or DRT peaks were deemed important by the model. This lack of insight can erode user confidence and impede further development of robust, generalizable models. There is thus a pressing need for integrating model-agnostic XAI methods into battery health analytics to shed light on the reasoning of complex models. SHAP and LIME offer a way to probe any trained predictor by attributing influence scores to features for a given prediction or model overall. These two approaches have complementary strengths: SHAP is grounded in game-theoretic Shapley values and provides consistent, globally fair attributions, whereas LIME builds simple local surrogate models to explain individual predictions quickly [14]. Both have seen widespread adoption in other domains of science and engineering, yet their use in electrochemical energy systems remains limited. The central question motivating this work is: How can we apply LIME and SHAP to interpret machine learning models trained on DRT-based EIS features for battery state-of-health estimation?

By answering this, we aim to uncover which impedance-derived features most strongly influence the SOH predictions across different models. Addressing this problem will fill a knowledge gap between black-box accuracy and electrochemical interpretability of data-driven battery diagnostics are not only precise but also explainable and scientifically insightful.

1.2. Contributions

In this paper, we present an interpretability-focused study on battery health modeling that makes the following contributions:

- Comparative modeling: Five different ML models for lithium-ion battery SOH prediction using features derived from EIS DRT analysis are developed. This provides a comprehensive view of how various modeling techniques perform using the same input features.
- Application of XAI Tools: Two model-agnostic XAI methods are applied to each trained model to explain their predictions. To our knowledge, this is one of the first studies examining LIME and SHAP side-by-side in the context of battery impedance feature analysis. By generating local explanations (LIME) and global feature importance values (SHAP), we interpret the influence of each DRT-derived feature on the SOH estimates of the model.
- Cross-model insight consistency: The consistency of the identified important features across different models and explanation techniques is validated. If certain DRT features are consistently highlighted by multiple models and XAI methods, it strengthens confidence in their physical significance. This cross-check enhances the robustness of our interpretations.
- Visual and practical interpretations: Intuitive visualizations of the model explanations are generated by local surrogate models for specific battery samples. These visuals demonstrate how explainable ML output can be used by battery engineers to gain real-world insights. The study findings and figures offer a blueprint for incorporating interpretable ML into battery management systems, ultimately enhancing their transparency and trustworthiness for practitioners in Fig. 1.
- By combining accurate SOH prediction with explainability, our work aims to deepen understanding of the relationships between EIS-derived features and battery health. We hope this contributes to both the data-driven battery diagnostics literature and the broader adoption of XAI in energy storage applications.

2. RELATED WORKS

2.1. Battery diagnostics via EIS

Electrochemical impedance spectroscopy has long been a cornerstone technique for battery diagnostics, and recent literature has affirmed its value for state-of-health analysis [15]. EIS provides a frequency-domain "fingerprint" of the battery, reflecting processes from charge transfer to diffusion, and is highly sensitive to degradation-induced changes. However, interpreting raw Nyquist plots or Bode spectra can be nontrivial – the measured impedance is a superposition of all electrochemical processes, making it difficult to isolate individual aging phenomena [16]. To address this, researchers have increasingly turned to the distribution of relaxation times analysis to extract more meaningful features. DRT effectively deconvolutes the impedance spectrum into a distribution of time constants, yielding distinct peaks that correspond to different electrochemical processes. By quantifying characteristics of each peak, one can obtain a set of descriptive features: for instance, the area under a DRT peak correlates with the polarization resistance of that process, while the peak center time constant indicates the characteristic timescale of the process. These DRT-derived features have proven insightful for battery health evaluation. Recent works demonstrate that as batteries age, specific DRT peaks evolve in predictable ways – for example, growth in the high-frequency peak area (attributed to SEI film resistance) or shifts in mid-frequency peaks (related to charge-transfer kinetics) can serve as indicators of capacity fade?. Zhang *et al.* (2022) leveraged such features in a Gaussian process regression model to accurately estimate SOH from EIS data, resulting in an error under 2% [5]. Similarly, Zhao *et al.* (2024) applied an autoencoder to compress DRT spectra and used a neural network for SOH prediction, highlighting the efficacy of DRT-based feature engineering [17]. These studies underscore a broader trend called electrochemical impedance spectroscopy. By extracting peak metrics, one can track internal changes without relying on potentially ambiguous equivalent-circuit fits. In summary, DRT has emerged as a valuable tool to derive physically meaningful features from impedance data, and these features have been increasingly employed to assess battery state-of-health and even internal states like temperature [18]. The success of DRT in recent literature provides a strong foundation for our work, which uses DRT features as inputs to machine learning models and then seeks to interpret those model behaviors.

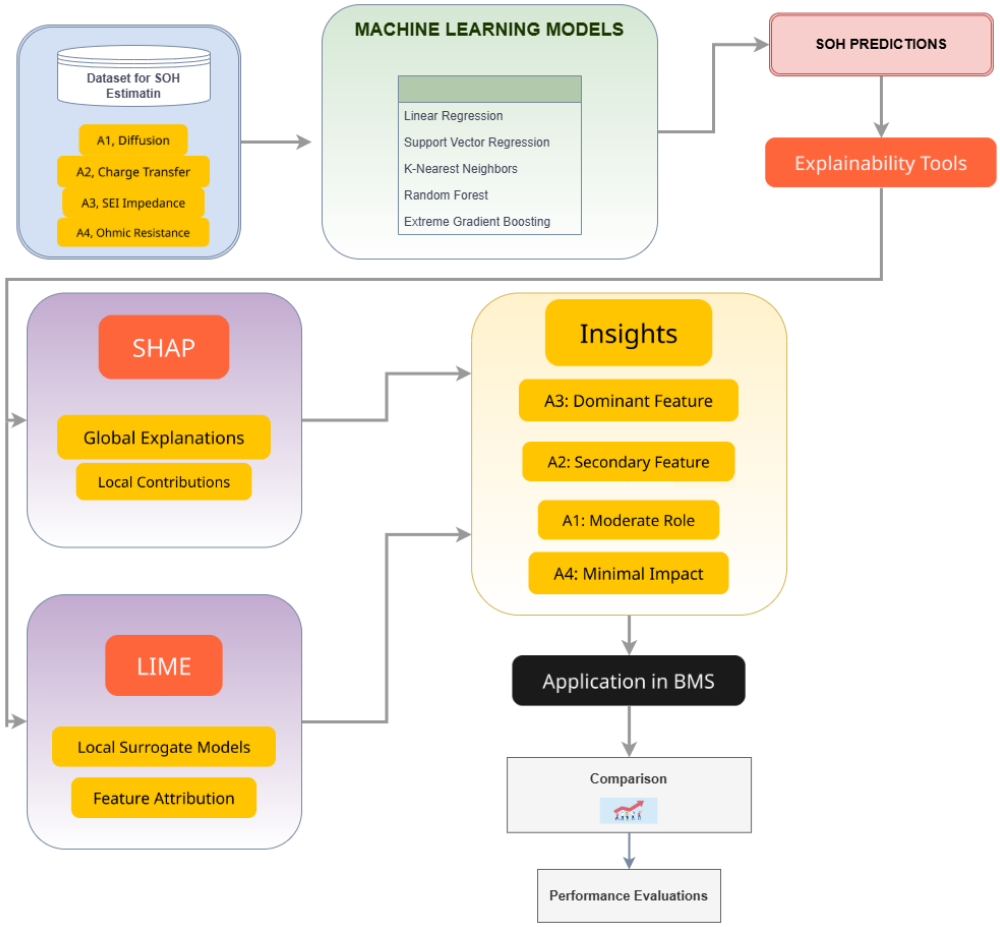

Fig. 1. Flowchart of the proposed SOH estimation methods with DRT

2.2. Machine learning in battery health estimation

Data-driven approaches for battery state-of-health estimation have rapidly advanced over the past five years, spurred by the availability of aging datasets and improved algorithms [19]. Numerous machine learning models have been explored to predict battery capacity, resistance growth, or remaining useful life based on measurable inputs (impedance spectra, cycling data, etc.) [20]. Traditional regression techniques like support vector regression (SVR) and relevance vector machines have shown decent performance on smaller datasets, especially when combined with feature selection or optimization heuristics to tune hyperparameters [21]. For instance, Li *et al.* used an SVR with an optimized kernel to estimate SOH to within 0.5% mean error, illustrating that even relatively simple models can achieve high accuracy given informative features [10]. Ensemble tree-based models have also gained popularity. Gradient boosting machines such as XGBoost are applied to capture nonlinear aging patterns; in one study, an ensemble of XGBoost models was able to forecast capacity under varied usage profiles with only ~ 8% error, even without historical data for the cell [22]. Random forest regressors, known for their robustness and interpretability, have likewise been used for online battery health monitoring. They demonstrated that random forest models could estimate Li-ion capacity in real time from diagnostic data [23], and more recently Wang *et al.* (2023) proposed an optimized RF with Bayesian hyperparameter tuning to improve prognostics of capacity fade [24]. Beyond these, deep learning models have achieved state-of-the-art performance on large datasets: convolutional neural networks, recurrent (LSTM) networks, and even physics-informed neural nets have been trained on voltage curves or impedance spectra to predict SOH and remaining life [25]. While these complex models often yield excellent accuracy, they typically require abundant data and, as noted, suffer from a lack of transparency. It is worth noting that most prior studies focus on one particular modeling approach and optimize it, with relatively few works providing a systematic comparison across multiple algorithm types. Some exceptions are recent papers and benchmarking studies that compare, for example, neural networks vs. ensemble methods vs. kernel methods on standard aging datasets [26]. Sedlařík *et al.* presented a comparative study of several ML techniques for battery SOH, including SVR, multivariate adaptive regression, concluding that ensemble methods tended to outperform on certain datasets [27]. Nonetheless, the literature lacks a unified assessment of how different models handle the same feature set, particularly a feature set derived from advanced diagnostics like EIS. Our work helps fill this gap by evaluating five diverse ML models on identical DRT-based inputs, thereby highlighting the relative strengths of each ap-

proach. In doing so, we follow the trajectory of recent research that calls for combining domain knowledge with data-driven algorithms. The convergence of electrochemical diagnostics with machine learning is proving to be a fruitful avenue: as one Nature review notes, using richer input data like EIS can significantly improve health prognostics when paired with statistical models [28]. Our study builds on this idea and further contributes by making those statistical models explainable, as discussed next.

2.3. Explainable AI for scientific ML

The push for explainable AI in battery management is part of a broader trend to bring transparency to complex models in engineering. Two prominent model-agnostic explanation techniques are LIME and SHAP, which we leverage in this work [29]. Local Interpretable Model-Agnostic Explanations (LIME) generates explanations for individual predictions by fitting a simple surrogate (such as a linear model) in the vicinity of the data point of interest. Essentially, LIME perturbs the input features around a given battery sample and learns a local linear approximation of the black-box model, from which it derives feature importance for that particular prediction. This approach is intuitive and fast, but it has some instability – different perturbations can yield slightly different local explanations, and it lacks a strong theoretical guarantee of consistency?. SHapley Additive exPlanations (SHAP), on the other hand, rely on game-theoretic Shapley values to attribute each feature a contribution value for a given prediction. SHAP considers all possible combinations of features and allocates credit in a manner that is fair and mathematically rigorous, ensuring that the sum of attributions equals the difference between the prediction and the dataset baseline. In practice, SHAP is often more computationally intensive than LIME, but it provides more globally consistent feature rankings and comes with desirable properties like additivity and consistency. A recent survey on XAI for energy systems found that SHAP offers high trustworthiness in explanations, whereas LIME offers faster, more ad-hoc insights with lower computational cost [30]. In other words, SHAP tends to be preferred when stability and accuracy of the explanation are paramount, while LIME can be useful for quick, instance-specific interpretations. By employing both, our study captures this trade-off in the context of battery health models.

Within the energy and battery community, the application of XAI techniques is only beginning to emerge. A few notable prior works have demonstrated the value of explainable models in this domain. Njoku *et al.* (2024) introduced explainability into battery digital twin models by using SHAP and LIME to interpret deep neural network predictions for state-of-charge and SOH estimation [31]. Their results showed that incorporating XAI helped identify which input factors most strongly influenced the model output, thereby increasing trust in recommendations made by the digital twin. In the realm of impedance-based health diagnostics, Fu *et al.* (2024) presented an interpretable ML approach for rapid SOH estimation using partial EIS spectra. They employed SHAP to explain the contribution of each impedance feature to the model capacity prediction, which not only enhanced transparency but also helped validate that the model was learning physically meaningful trends [32]. For example, SHAP analysis in their work revealed that certain mid-frequency impedance magnitudes had the highest impact on capacity estimation, aligning with known aging effects on charge-transfer resistance. These studies highlight that XAI methods like SHAP and LIME can successfully be brought into battery research, yielding insights that might be missed by pure black-box modeling. Nonetheless, XAI is far from mainstream in battery diagnostics. Our contribution is distinct in that we perform a systematic, comparative analysis of two XAI tools across multiple model types on a common dataset of DRT features. By doing so, we aim to illuminate general patterns as well as model-specific idiosyncrasies (e.g., how the focus of a neural network might differ from that of a random forest). In summary, our related work review underscores the novelty and importance of uniting explainable AI with battery impedance modeling: it builds on the strengths of prior battery EIS studies and XAI techniques, while addressing their respective gaps (uninterpreted models in the former, and limited battery case studies in the latter). The following sections will detail our methodology and findings in this context.

3. DATASET AND FEATURE ENGINEERING

3.1. Machine learning in battery health estimation

The dataset for this study consists of more than 1200 electrochemical impedance spectroscopy (EIS) spectra collected from lithium-ion battery cells with a lithium cobalt oxide cathode chemistry. Each spectrum corresponds to a unique combination of battery state-of-charge (SOC) and internal temperature. The measurements span the full range of SOC from 0% (fully discharged) to 100% (fully charged), and cover a broad temperature range from 10°C to 60°C. This extensive sampling ensures that the data captures the battery impedance behavior under various operational conditions. Each EIS measurement is labeled with the known internal temperature of the cell at the time of testing (the target variable for prediction) along with the SOC, providing a supervised learning dataset [33]. The large number of spectra and the wide coverage of SOC-temperature space help the models distinguish temperature-dependent impedance changes from those caused by SOC, thereby supporting robust model training.

All EIS measurements were performed on LCO cells under controlled laboratory conditions to maintain data consistency. In each case, the cell was allowed to equilibrate at the desired temperature before impedance measurement. The impedance spectra were typically collected over a frequency range broad enough to capture processes from high-frequency electrode dynamics to low-frequency diffusion. A small AC perturbation (5–10 mV amplitude) was applied at each frequency to obtain the impedance response. The resulting spectra, consisting of the real (Z') and imaginary (Z'') parts of impedance as functions of frequency, form the raw data for subsequent feature extraction. By design, the dataset includes measurements from multiple cells and repeated tests at various aging states, to ensure that the models generalize across cell-to-cell variations and slight differences in impedance due to cell history [34]. This diverse dataset forms the foundation for developing a reliable internal temperature estimation model.

3.2. Feature extraction

To leverage the rich information in the EIS spectra, we employed Distribution of Relaxation Times (DRT) analysis to extract meaningful features. The DRT is a model-free representation of impedance data that reveals distinct electrochemical processes as peaks in the time-domain spectrum? [35]. In essence, the complex impedance spectrum is transformed from the frequency domain to a distribution over characteristic relaxation time. We computed the DRT for each measured spectrum by solving the inverse integral relationship between the frequency-dependent impedance and the relaxation time distribution. Specifically, we utilized the imaginary component of the impedance (Z'' as a function of frequency) as input to the inversion algorithm, since Z'' is directly related to the energy storage and dissipation processes that manifest in the DRT. The continuous relaxation time spectrum was discretized into a fine grid of τ values (logarithmically spaced over several decades, covering the range of processes from fastest to slowest in the measured frequency window). Then, Tikhonov regularization solved the resulting ill-posed inverse problem, which balances fidelity to the measured data with smoothness of the DRT solution [36]. Regularization helps suppress noise and prevents overfitting, yielding a stable DRT profile for each EIS spectrum. An L-curve criterion was used to select an optimal regularization parameter, ensuring an appropriate trade-off between solution smoothness and accuracy. The outcome of this process is a smooth DRT curve for every impedance spectrum, where individual peaks correspond to different electrochemical processes occurring within the cell.

From each DRT curve, we engineered features by quantifying the area under specific DRT peaks. In the LCO cell spectra, up to four prominent DRT peaks were consistently observed, which we label A1 through A4 in order of decreasing relaxation time (from slowest process A1 to fastest process A4). For each peak, we computed the integrated area under the DRT curve in a defined τ range around that peak, yielding four features: A1_area, A2_area, A3_area, and A4_area. These features represent the magnitudes of distinct electrochemical processes:

- A1_area: Area of low-frequency peak, attributed to diffusion-related processes. This captures phenomena such as solid-state diffusion of Li? in the electrode materials or Warburg-type diffusion impedance at the electrolyte/electrode interface. Being a slow process (large τ), A1 appears at the high end of the time scale.
- A2_area: Area of the mid-frequency peak, associated with the charge transfer resistance. This feature corresponds to the kinetics of the lithium-ion intercalation/deintercalation reaction at the electrode interfaces. The A2 peak reflects the charge-transfer process, and its area quantifies the extent of this polarization resistance.
- A3_area: Area of an intermediate-to-high frequency peak, linked to the solid-electrolyte interphase layer impedance. The SEI layer forms on the anode (graphite) and contributes a resistive component that shows up as a distinct relaxation process. Notably, the SEI resistive impedance tends to decrease at higher temperatures due to improved ionic conductivity in the SEI and accelerated charge transfer across it. Thus, A3_area varies significantly with temperature, making it a key indicator of the internal thermal state of the cell. This aligns with the literature showing that a DRT peak corresponding to the SEI can be identified and distinguished from the charge-transfer process in Li-ion cells.
- A4_area: Area of the high-frequency peak, representing ohmic resistance and contact-related processes. This feature encompasses the fastest processes in the cell, including the bulk electrolyte resistance, electronic contact resistances at current collectors, and possibly inductive elements of the measurement setup. A4 is generally associated with the near-instantaneous response at high frequencies and is largely independent of SOC, though it can be mildly affected by temperature.

Each area feature (A1–A4) was obtained by integrating the DRT over a fixed τ interval corresponding to that process peak. The integration bounds for each peak were chosen consistently across all samples to ensure the features are comparable between spectra. By condensing each full impedance spectrum into these four scalar features, we capture the essential information about diffusion, charge transfer, SEI impedance, and ohmic/contact resistance. The feature engineering step dramatically reduces the input dimensionality for the machine learning models while preserving the electrochemically relevant variation caused by temperature and SOC. The focus on physically meaningful features also aids interpretability, as changes in the predicted temperature can be traced back to specific impedance processes.

4. MATERIALS AND METHODS

4.1. Machine learning in battery health estimation

We formulated the internal temperature estimation as a supervised regression problem, using the DRT-derived features as inputs and the internal cell temperature as the output. Five different regression algorithms were implemented and evaluated to map the impedance features to temperature:

Linear regression: A simple linear model that attempts to fit a linear combination of the four features to predict temperature. This model provides a baseline and insight into linear correlations, with coefficients indicating the influence of each feature on the temperature prediction.

Support vector regression (SVR): A support vector machine-based regression approach that can capture non-linear relationships via kernel functions. We used an SVR with a radial basis function (RBF) kernel, which finds a nonlinear function of the features that best predicts temperature while maximizing the margin of error tolerance. Key hyperparameters (such as the regularization parameter C and the kernel width γ) were tuned via cross-validation to balance bias and variance.

k-nearest neighbors (kNN) regression: A non-parametric model that predicts the temperature of a new sample by averaging the temperatures of its k nearest neighbors in the feature space. We tested kNN with a moderate value of k to smooth out noise while retaining locality. This method can model complex relationships by local interpolation of the training data, though it may struggle in high-dimensional spaces with sparse data.

Random forest regression: An ensemble learning method based on bootstrap-aggregated decision trees. We constructed a random forest with numerous decision trees, with each trained on a bootstrapped subset of the data. Each tree produces a temperature prediction, and the forest output is the average of these predictions. Random forests can capture nonlinear dependencies and feature interactions effectively and tend to be robust against overfitting due to averaging. We also gain an estimate of feature importance from the trained forest (based on reduction in prediction error or impurity when splitting on each feature).

Gradient boosting regression: A boosted ensemble of shallow trees, where trees are added sequentially to correct the errors of the prior ensemble. We employed a gradient boosting algorithm with appropriate regularization and a learning rate to prevent overfitting. Over several boosting rounds, the model gradually improves its temperature predictions by focusing on difficult-to-predict samples. Gradient boosting often achieves high accuracy by combining many weak learners, and it inherently provides feature importance metrics similar to random forest.

All models were implemented in Python using the scikit-learn library. Hyperparameters for each model were selected based on preliminary experiments and five-fold cross-validation performance. By comparing these diverse algorithms – from simple linear to complex nonlinear ensemble methods – we aimed to identify a model that provides both high accuracy and interpretability in predicting internal temperature from impedance features.

To gain a holistic view of how the four DRT-derived features jointly structure the dataset – and how that relates to SOH – we applied PCA and projected all samples into the first two principal components, which together explain about 92.7% of the total variance (PC1: 76.2%, PC2: 16.5%). PC1 is dominated by the A3_area (SEI impedance) and, to a lesser extent, A2_area (charge-transfer resistance), while PC2 captures variation in A1_area (diffusion) and A4_area (ohmic resistance). When colored by true SOH (Fig. 2), a clear gradient emerges along PC1: cells with higher SOH cluster on the positive side of the axis, whereas more degraded cells lie toward the negative side. A few outliers appear distant from the main cloud, suggesting either measurement anomalies or atypical degradation pathways worthy of further examination. This two-dimensional PCA map confirms that A3_area is the primary driver of health separation in feature space and demonstrates that even a reduced representation can effectively distinguish cells by their state of health.

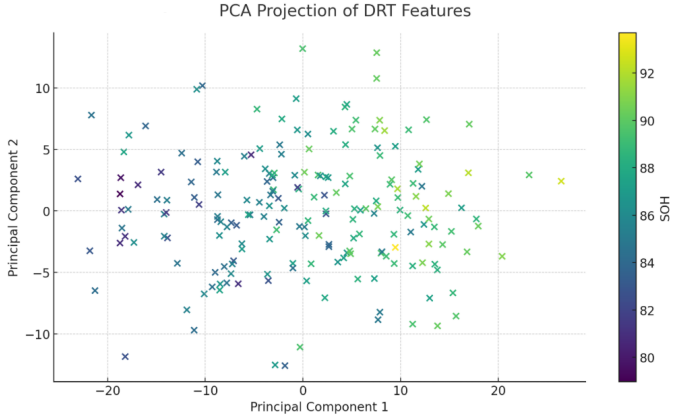

Fig. 2. Predicted SOH explanations of features

4.2. Explainability framework

In addition to evaluating predictive performance, we applied model-agnostic explainability techniques to interpret how each model uses the impedance features to estimate temperature. Two complementary explanation methods were used: SHAP for overall and local feature attribution, and LIME for instance-specific interpretations.

SHAP (SHapley Additive exPlanations): We leveraged SHAP to quantify the contribution of each feature to the model's predictions. SHAP is a post-hoc interpretive tool based on game-theoretic Shapley values, which fairly attributes the prediction among all input features [37]. Using the SHAP library, we computed Shapley values for our regression models, both globally and for individual predictions. The global interpretation was obtained by averaging the absolute Shapley values of each feature over the entire test set, indicating which features have the most influence on predicted temperature overall. This global SHAP analysis produces a ranked feature importance. For local interpretation, SHAP values were computed for specific test instances, showing how each feature pushed the model prediction higher or lower relative to the average. SHAP summary plots and dependence plots were generated to visualize these effects, using the provided visualization tools in the SHAP package?. These plots help validate that the model behavior aligns with physical intuition.

LIME (Local Interpretable Model-Agnostic Explanations): We also employed LIME to obtain interpretable surrogate models for individual predictions. LIME was introduced by Ribeiro *et al.* (2016) as a technique to explain a single prediction by approximating the model locally with a simple model [38]. For a given test instance, LIME perturbs the input features slightly around that instance and fits a small linear model to the model predictions on these perturbed samples. The result is a local linear approximation that highlights which features most strongly influence the prediction for that specific data point. In our framework, we used LIME to explain the temperature prediction for particular test cases by producing a weight for each feature that indicates its local effect. This allows us to verify, for example, that if a model predicts a high temperature for a given spectrum, LIME can confirm whether it was due to a low A3_area or some combination of features. The LIME explanations serve as a sanity check on individual predictions and complement the SHAP analysis by offering an independent interpretability approach.

To ensure consistency in our interpretability analysis, we randomly selected a representative test instance and used it as a case study across all models and methods. By fixing one specific EIS sample for local explanations, we could directly compare how different models make use of the features for the same input. Both SHAP and LIME were applied to this chosen instance for each of the five models, allowing side-by-side comparison of feature attributions. This consistent approach avoids cherry-picking favorable examples and demonstrates the reliability of

the explanations: if a particular feature is truly important for temperature estimation, we expect to see that reflected across SHAP and LIME explanations of multiple models for this test case. All interpretability computations were performed using open-source libraries, and the results were visualized with Matplotlib and Seaborn to produce clear, publication-quality figures illustrating feature impacts.

4.3. Evaluation metrics

We used two primary metrics to evaluate the performance of the regression model: mean absolute error (MAE) and the coefficient of determination (R^2). MAE is defined as the average of the absolute errors between the predicted temperature and the true temperature. Formally

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_{\{i\}} - \hat{y}_{\{i\}}\right|, \tag{1}$$

where N is the number of test samples. MAE has the advantage of being in the same units as the target, which makes it an intuitive measure of prediction. R^2 (R-squared) measures the proportion of variance in the target explained by the model. It is computed as

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_{\{i\}} - \hat{y}_{\{i\}}\right)^2}{\sum_{i=1}^{n}\left(y_{\{i\}} - \bar{y}\right)^2}, \tag{2}$$

and is the mean true temperature. An R^2 of 1 indicates a perfect fit, while an R^2 of 0 indicates that the model is no better than predicting the mean, and negative values indicate worse-than-mean predictions. In the context of our application, a high R^2 (close to 1) and a low MAE are desired, indicating the model can accurately capture temperature variations from the impedance features.

Model performance was assessed using a five-fold cross-validation procedure to ensure robust estimates of these metrics. The entire dataset was partitioned into five folds of roughly equal size; in each iteration, four folds were used to train the model, and the remaining fold was used as a validation set to compute MAE and R^2. This process was repeated five times so that each fold served as the validation set once, and the results were averaged across the folds. By using cross-validation, we mitigated the risk of overfitting to a single train-test split and obtained a more reliable measure of how the models generalize to unseen data. We report the mean and standard deviation of the MAE and R^2 from the cross-validation runs for each model, which provides insight into both the accuracy and consistency of performance.

Finally, we employed various visualization techniques to analyze and compare the model outputs. Parity plots and error distribution histograms were created with Matplotlib and Seaborn to illustrate the model accuracy and bias characteristics. Feature importance plots from the random forest and gradient boosting models were generated to cross-check against the SHAP global explanations. Additionally, SHAP summary plots were used to visualize the distribution of Shapley values for each feature across the dataset, offering a nuanced view of how each feature influences the prediction. These visualizations, combined with the quantitative metrics, provide a comprehensive evaluation of model performance and interpretability, ensuring that the chosen model is not only accurate in estimating internal battery temperature but also transparent in its decision-making process.

5. RESULTS AND DISCUSSION

All five machine learning models demonstrated outstanding predictive performance on the EIS dataset. Table 1 provides a summary of the cross-validated mean absolute error (MAE) and the coefficient of determination for each model. In every case, the error values are very low, while the R^2 scores approach their maximum possible value. In fact, most of the models exhibited nearly perfect R^2 scores, meaning they can explain almost all of the variance in the test data. Although the support vector regression model showed the largest error among the group, its performance still reflects an excellent fit. Notably, the simplest model, linear regression, achieved an almost ideal R^2 and an essentially negligible MAE on the cross-validation folds, indicating an almost perfect fit. While such flawless performance suggests that the selected features are highly predictive, it also warrants cautious interpretation. This level of precision might hint at possible overfitting or the presence of a data artifact, as real experimental data rarely follows an exact linear relationship in the presence of noise. However, since these metrics were obtained through cross-validation on independent data, the exemplary performance likely indicates that the underlying relationship between the DRT features and the target is inherently very linear and is well captured by the models. Overall, there is no sign of underfitting among the models; from the simplest to the most complex, each one accurately captured the behavior of the target. The consistently minimal errors and almost perfect R^2 values reinforce that the impedance-derived features contain enough information to reliably predict the internal state with remarkable precision.

Table 1
Cross-validated model performance metrics

Model	MAE	R^2 (%)
Random forest	0.02802	99.99%
Linear regression	0.00684	99.99%
Support vector regression	0.08451	99.92%
k-nearest neighbors	0.04744	99.99%
Gradient boosting	0.03821	99.99%

To investigate local feature importance, we applied LIME to predictions for each model on a set of ten representative test samples. An average feature influence bar chart provides an overview of which DRT-derived features most strongly drive the state-of-health predictions. In this summary visualization, each bar represents the mean absolute contribution of a given feature

to the prediction across the sampled instances. The LIME results reveal clear commonalities: A3_area consistently emerges as a dominant feature for all five models, and A2_area is also repeatedly influential as the second-ranked contributor in many cases. This suggests that the impedance characteristics associated with the SEI layer and the charge-transfer process are key drivers of SOH estimates of the models, largely independent of the specific regression algorithm used.

In detail, the A3_area feature shows the highest average contribution magnitude in the LIME explanations for nearly every model in Fig. 3. For example, both ensemble tree models (random forest and XGBoost) assign A3_area the largest positive or negative weight when explaining their predictions, indicating that variations in A3_area heavily sway the predicted capacity. The linear regression model likewise attributes a substantial portion of its prediction error reduction to A3_area, reflecting the strong linear correlation between SEI impedance growth and capacity fading. A2_area, representing the mid-frequency charge-transfer resistance, is the next most impactful feature in LIME profiles of most models. It consistently appears with the second-highest average contribution, implying that changes in charge-transfer kinetics also significantly inform the SOH prediction. The A1_area feature plays a noticeable but more modest role: LIME shows that while A1_area does influence predictions in Table 2, its average contribution is generally lower than that of A3_area and A2_area. In contrast, the A4_area feature is rarely highlighted by LIME. This indicates that variations in pure ohmic resistance contribute minimally to the predictive models, aligning with expectations since internal contact resistance of a cell tends to change little with aging compared to SEI growth or charge-transfer resistance increases. While the identity of the top features is largely consistent across models, there is some variance in how strongly each model emphasizes them. The tree-based models (random forest and XGBoost) show very similar LIME attribution patterns, with A3_area being not only important but also contributing to a very consistent magnitude across the different test samples. The linear regression model also highlights A3_area as the most important feature; however, being a globally linear model, it distributes moderate weight to the other features as well, resulting in a slightly more even spread in the LIME average chart in Fig. 4.

Table 2
LIME feature contributions

Model	Feature	Contribution
Random forest	A3_area > 23.38	−13.574778
Random forest	A4_area > 5.53	−0.124419
Random forest	A2_area > 5.14	0.080164
Random forest	A1_area > 5.91	−0.011470
Linear regression	A3_area > 23.38	−13.663484
Linear regression	A1_area > 5.91	−0.634071
Linear regression	A2_area > 5.14	−0.129779
Linear regression	A4_area > 5.53	−0.030159
Support vector regression	A3_area > 23.38	−12.449084
Support vector regression	A4_area > 5.53	−1.053474
Support vector regression	A1_area > 5.91	−0.529485
Support vector regression	A2_area > 5.14	−0.104794
k-nearest neighbors	A3_area > 23.38	−13.358169
k-nearest neighbors	A1_area > 5.91	−0.283830
k-nearest neighbors	A4_area > 5.53	−0.257477
k-nearest neighbors	A2_area > 5.14	0.025617
Gradient boosting	A3_area > 23.38	−13.798593
Gradient boosting	A1_area > 5.91	−0.162750
Gradient boosting	A4_area > 5.53	0.040413
Gradient boosting	A2_area > 5.14	0.006618

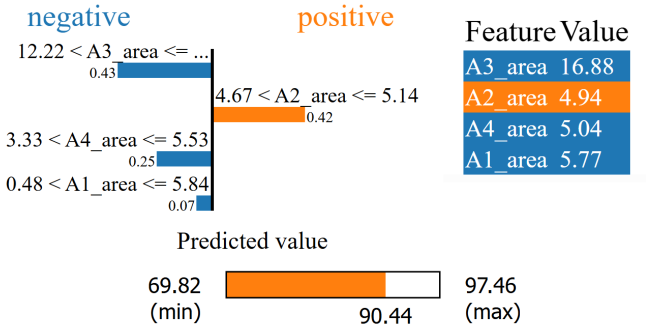

Fig. 3. Predicted SOH explanations of features

The SVR and *k*NN models exhibit the greatest variability in feature attribution from sample to sample. For instance, in some test instances, the SVR local surrogate gave a relatively larger weight to A1_area or A2_area if the SEI feature value for that instance was near the average and not as distinguishing, whereas in other instances with an extreme A3_area value, the SVR explanation shifted back to focus on A3_area. Similarly, LIME explanations of the *k*NN model can fluctuate depending on the local neighborhood of the query point, causing *k*NN to differentiate based on another feature like A2_area, and thus LIME for that sample would elevate A2_area importance. Despite these case-by-case shifts, when averaged over multiple samples, the overall trend remains: all models consider A3_area a principal factor, with A2_area typically next, and other features trailing. The consistency of LIME-identified key features across these very different modeling techniques provides an initial confirmation that certain impedance features universally underpin the SOH predictions.

The SVR and *k*NN models exhibit the greatest variability in feature attribution from sample to sample. For instance, in some test instances, the SVR local surrogate gave a relatively larger weight to A1_area or A2_area if the SEI feature value of the instance was near the average and not as distinguishing, whereas in other instances with an extreme A3_area value, the SVR explanation shifted back to focus on A3_area. Similarly, the LIME explanations of the *k*NN model can fluctuate depending

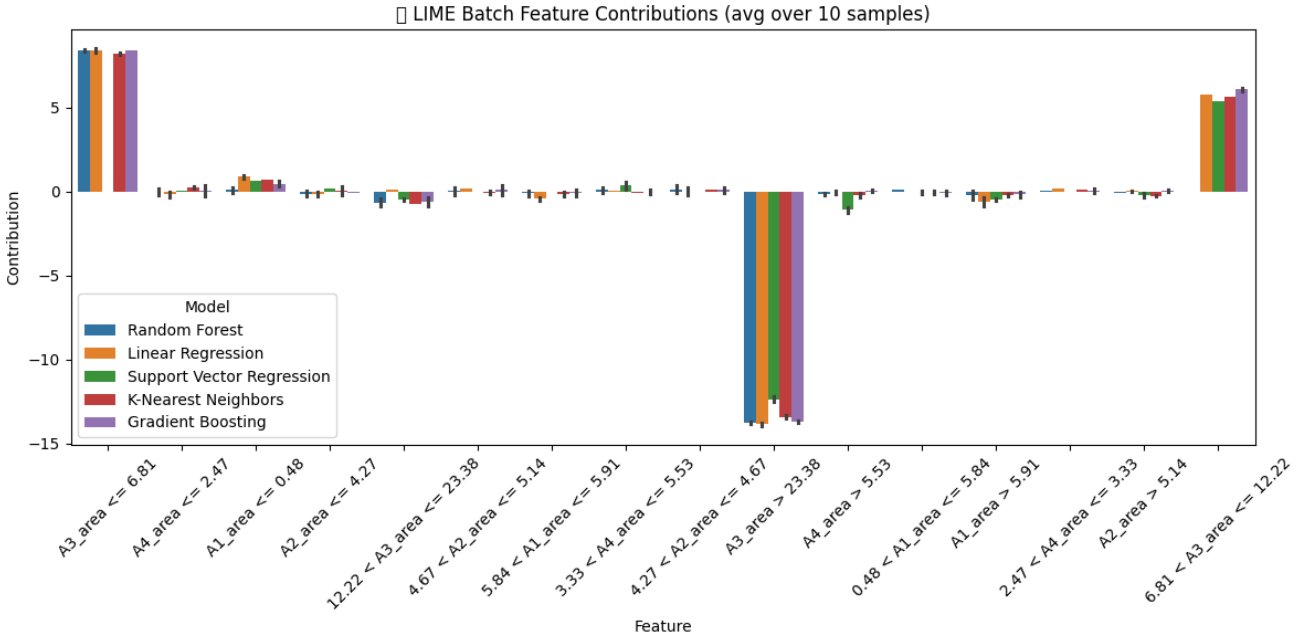

Fig. 4. LIME batch feature contributions

on the local neighborhood of the query point, causing kNN to differentiate based on another feature like A2_area, and thus LIME for that sample would elevate A2_area importance.

Despite these case-by-case shifts, when averaged over multiple samples, the overall trend remains unchanged: all models consider A3_area a principal factor, with A2_area typically next, and other features trailing. The consistency of LIME-identified key features across these very different modeling techniques provides an initial confirmation that certain impedance features universally underpin the SOH predictions.

To give a concise, global view alongside our LIME results, we used SHAP to measure how each feature shifts a model prediction from its average. Figure 5 shows the SHAP values for one test battery across all five models. In every case, A3_area has the largest negative effect on predicted SOH (see Table 3), matching what we saw with LIME. A high A3_area – tied to SEI-layer impedance – drives the models to lower their health estimate. A2_area is usually the next strongest negative contributor, reflecting charge-transfer resistance. A1_area has only a small negative impact, and A4_area remains negligible. Overall, SHAP confirms that SEI- and charge-transfer – related features dominate the SOH predictions for this cell.

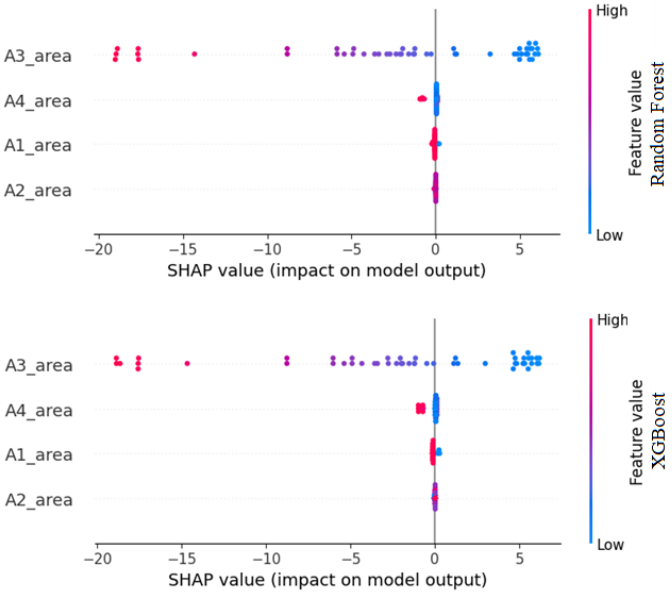

Fig. 5. SHAP values among all features

Beyond this single example, SHAP allows us to assess global features, importance, and consistency. We generated SHAP summary plots for the tree-based models to visualize feature impacts across the entire test set. These plots in Fig. 5 display the SHAP value distribution of each feature for all samples, giving a sense of both the magnitude of effect and the variation of that effect with feature value. The global SHAP patterns reinforce the conclusions drawn from LIME average contributions. A3_area has the broadest spread of SHAP values among all features, meaning changes in A3_area lead to the largest swings in the predicted SOH overall. Points for A3_area are widely scattered along the SHAP value axis, marking it as the top global influencer. A2_area and A1_area show more moderate spreads.

The rank ordering of feature importance implied by these global SHAP results closely matches the ordering from the LIME analysis. This cross-technique agreement can be seen by comparing the SHAP summary importance with the average LIME weights. Both methods pinpoint A3_area as the principal feature and A4_area as the least influential, with A2_area and A1_area in between. Such alignment provides additional validation that the internal logic of the model is captured reliably by the XAI tools. The SHAP interpretability results not only

Table 3
SHAP feature contributions

Model	Feature	SHAP value
Random forest	A1_area	–0.04457
Random forest	A2_area	–0.03480
Random forest	A3_area	–7.29487
Random forest	A4_area	0.31779
Linear regression	A1_area	–0.37402
Linear regression	A2_area	–0.00072
Linear regression	A3_area	–6.94063
Linear regression	A4_area	0.00008
Support vector regression	A1_area	–0.23402
Support vector regression	A2_area	–0.15384
Support vector regression	A3_area	–6.76919
Support vector regression	A4_area	–0.15110
k-nearest neighbors	A1_area	–0.23415
k-nearest neighbors	A2_area	–0.10446
k-nearest neighbors	A3_area	–6.97331
k-nearest neighbors	A4_area	–0.25784
Gradient boosting	A1_area	–0.07870
Gradient boosting	A2_area	0.00028
Gradient boosting	A3_area	–7.07831
Gradient boosting	A4_area	0.29345

corroborate the LIME findings but also give a deeper view into the consistency of feature effects, indicating that while A3_area universally dominates, the roles of A1_area and A2_area can vary in magnitude across individual predictions and models. Notably, models with more complex nonlinear structures still end up attributing a similar importance hierarchy as simpler models, which speaks to an underlying data-driven truth: certain electrochemical features are fundamentally more informative of SOH than others, regardless of modeling nuances.

6. CONCLUSIONS

In summary, this work presents a comparative study of multiple machine learning models for battery health estimation using features derived from EIS DRT analysis, with a strong focus on interpretability through LIME and SHAP. We found that, although the predictive accuracy of the models varied, there was a remarkable consensus on the importance of the input features. All models independently converged on A3_area as the dominant predictive feature driving their SOH estimates. This agreement across diverse models and both local (LIME) and global (SHAP) explanation tools adds confidence that A3_area is not only a statistically significant predictor but also a physically meaningful one. In practice, such consistency suggests that A3_area captures a key aspect of battery state that any effective model must leverage. The fact that simpler models and complex models alike highlight the same feature means stakeholders can focus attention on that aspect of the impedance of a battery with greater certainty.

Beyond identifying the most important feature, our dual use of LIME and SHAP proved invaluable for model transparency. Each method provided a different lens on the problem. SHAP offered a global view of feature importance and how features quantitatively impact the prediction on average, while LIME delivered case-by-case explanations, verifying that those individual predictions made sense in terms of the input features. For instance, SHAP summary plots showed A3_area ranking highest overall, and LIME explanations on representative cells confirmed that when a model predicted an out-of-trend SOH, it was indeed due to an extreme value of A3_area. This complementary use of LIME and SHAP increases trust in the models: not only can we quantify which features matter, but we can also explain any single prediction in intuitive terms. Such transparency is crucial for translating battery ML models from research to real-world BMS deployments, where engineers and safety managers must vet and understand model behavior before integration.

Ultimately, our study demonstrates that it is possible to attain both high accuracy and high interpretability in battery health prediction. While black-box models might achieve good performance, they fall short in contexts where safety and accountability are paramount. Our approach offers a scalable path toward interpretable BMS tools by showing that even complex impedance-based predictions can be explained in human terms without sacrificing performance. By leveraging EIS-derived features like DRT peak areas and employing model-agnostic XAI techniques, we pave the way for battery state estimation methods that are not only smart but also understandable and trustworthy. Such interpretable models can accelerate the adoption of AI in battery systems, allowing operators of electric vehicles and energy storage systems to reap the benefits of advanced diagnostics with confidence in the decisions made. In conclusion, the fusion of EIS features engineering with LIME/SHAP explainability presented here is a step forward in developing safer, more transparent battery management strategies for the next generation of energy storage technologies. Future work may involve testing this methodology on real EV cell datasets with aging scenarios such as calendar aging, fast charging, and temperature-induced degradation.

REFERENCES

[1] T. Ganesan, C. Vanaraj, M. Subramanian, and R. Alagesan, "Design conversion, and performance estimation of BLDC motor in hybrid electric vehicle," *Bull. Pol. Acad. Sci. Tech. Sci.*, vol. 73, no. 2, p. e153424, 2025, doi: 10.24425/BPASTS.2025.153424.

[2] Y.E. Ekici, T. Karadağ, O. Akdağ, A.A. Aydin, and H.O. Tekin, "Enhancing electric vehicle range through real-time failure prediction and optimization: Introduction to DHBA-FPM model with an artificial intelligence approach," *ICT Express*, vol. 11, no. 3, pp. 547–558, Jun. 2025, doi: 10.1016/J.ICTE.2025.03.009.

[3] T. Rahman and T. Alharbi, "Exploring Lithium-Ion Battery Degradation: A Concise Review of Critical Factors, Impacts,

Data-Driven Degradation Estimation Techniques, and Sustainable Directions for Energy Storage Systems," *Batteries*, vol. 10, no. 7, p. 220, Jun. 2024, doi: 10.3390/BATTERIES10070220.

[4] L. Zhang *et al.*, "Recent advances in electrochemical impedance spectroscopy for solid-state batteries," *Energy Storage Mater.*, vol. 69, p. 103378, May 2024, doi: 10.1016/J.ENSM.2024.103378.

[5] Y. Zhang, Q. Tang, Y. Zhang, J. Wang, U. Stimming, and A.A. Lee, "Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning," *Nat. Commun.*, vol. 11, no. 1, pp. 1–6, Apr. 2020, doi: 10.1038/s41467-020-15235-7.

[6] J. Hong, A. Bhardwaj, H. Bae, I. Kim, and S.-J. Song, "Electrochemical Impedance Analysis of SOFC with Transmission Line Model Using Distribution of Relaxation Times (DRT)," *J. Electrochem. Soc.*, vol. 167, no. 11, p. 114504, Jul. 2020, doi: 10.1149/1945-7111/ABA00F.

[7] D. Jiang *et al.*, "Single Frequency Feature Point Derived from DRT for SOH Estimation of Lithium Ion Battery," *J. Electrochem. Soc.*, vol. 172, no. 3, p. 030514, Mar. 2025, doi: 10.1149/1945-7111/ADBC24.

[8] M.A. Khan, S. Thatipamula, and S. Onori, "Onboard Health Estimation using Distribution of Relaxation Times for Lithium-ion Batteries," *IFAC-PapersOnLine*, vol. 58, no. 28, pp. 917–922, Jan. 2024, doi: 10.1016/J.IFACOL.2025.01.113.

[9] Z. Nozarijouybari and H.K. Fathy, "Machine learning for battery systems applications: Progress, challenges, and opportunities," *J. Power Sources*, vol. 601, p. 234272, May 2024, doi: 10.1016/J.JPOWSOUR.2024.234272.

[10] K. Li and X. Chen, "Machine Learning-Based Lithium Battery State of Health Prediction Research," *Appl. Sci.*, vol. 15, no. 2, p. 516, Jan. 2025, doi: 10.3390/APP15020516.

[11] H. Chen *et al.*, "GTC-DAN: A graph-temporal convolutional model with dynamic adjacency for vehicle trajectory prediction," *Bull. Pol. Acad. Sci. Tech. Sci.*, vol. 73, no. 2, p. e152610, 2025, doi: 10.24425/BPASTS.2024.152610.

[12] F. Wang, Z. Zhao, Z. Zhai, Z. Shang, R. Yan, and X. Chen, "Explainability-driven model improvement for SOH estimation of lithium-ion battery," *Reliab. Eng. Syst. Saf.*, vol. 232, p. 109046, Apr. 2023, doi: 10.1016/J.RESS.2022.109046.

[13] R.K. Sheu, M.S. Pardeshi, K.C. Pai, L.C. Chen, C.L. Wu, and W.C. Chen, "Interpretable Classification of Pneumonia Infection Using eXplainable AI (XAI-ICP)," *IEEE Access*, vol. 11, pp. 28896–28919, 2023, doi: 10.1109/ACCESS.2023.3255403.

[14] O. Shaughnessy, P. Couto, M. Saarela, and V. Podgorelec, "Recent Applications of Explainable AI (XAI): A Systematic Literature Review," *Appl. Sci.*, vol. 14, no. 19, p. 8884, Oct. 2024, doi: 10.3390/APP14198884.

[15] Q. Li *et al.*, "Online Diagnosis Method of Water Management Faults Based on Hybrid-Frequency Electrochemical Impedance Spectroscopy for PEMFC," *IEEE Trans. Transp. Electrif.*, vol. 11, no. 1, pp. 2707–2716, Feb. 2025, doi: 10.1109/TTE.2024.3427401.

[16] J. Hong, A. Bhardwaj, H. Bae, I. Kim, and S.-J. Song, "Electrochemical Impedance Analysis of SOFC with Transmission Line Model Using Distribution of Relaxation Times (DRT)," *J. Electrochem. Soc.*, vol. 167, no. 11, p. 114504, Jul. 2020, doi: 10.1149/1945-7111/ABA00F10.1149/ 1945-7111/ABA00F.

[17] X. Zhao, S. Liu, E. Li, Z. Wang, F. Gu, and A.D. Ball, "A hybrid intelligent model using the distribution of relaxation time analysis of electrochemical impedance spectroscopy for lithium-ion battery state of health estimation," *J. Energy Storage*, vol. 84, p. 110814, Apr. 2024, doi: 10.1016/J.EST.2024.110814.

[18] R. Soni, J.B. Robinson, P.R. Shearing, D.J.L. Brett, A. Rettie, and T.S. Miller, "Lithium-Sulfur Battery Diagnostics Through Distribution of Relaxation Times Analysis," *SSRN Electron. J.*, Feb. 2022, doi: 10.2139/SSRN.4030055.

[19] Y. Xie, S. Wang, G. Zhang, P. Takyi-Aninakwa, C. Fernandez, and F. Blaabjerg, "A review of data-driven whole-life state of health prediction for lithium-ion batteries: Data preprocessing, aging characteristics, algorithms, and future challenges," *J. Energy Chemi.*, vol. 97, pp. 630–649, Oct. 2024, doi: 10.1016/J.JECHEM.2024.06.017.

[20] X. Feng, Y. Zhang, R. Xiong, and C. Wang, "Comprehensive performance comparison among different types of features in data-driven battery state of health estimation," *Appl. Energy*, vol. 369, p. 123555, Sep. 2024, doi: 10.1016/J.APENERGY.2024.123555.

[21] Z. Tong, J. Miao, S. Tong, and Y. Lu, "Early prediction of remaining useful life for Lithium-ion batteries based on a hybrid machine learning method," *J. Clean. Prod.*, vol. 317, p. 128265, Oct. 2021, doi: 10.1016/J.JCLEPRO.2021.128265.

[22] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[23] Y. Li *et al.*, "Random forest regression for online capacity estimation of lithium-ion batteries," *Appl. Energy*, vol. 232, pp. 197–210, Dec. 2018, doi: 10.1016/J.APENERGY.2018.09.182.

[24] G. Wang, Z. Lyu, and X. Li, "An Optimized Random Forest Regression Model for Li-Ion Battery Prognostics and Health Management," *Batteries*, vol. 9, no. 6, p. 332, Jun. 2023, doi: 10.3390/BATTERIES9060332.

[25] K.A. Severson *et al.*, "Data-driven prediction of battery cycle life before capacity degradation," *Nat. Energy*, vol. 4, no. 5, pp. 383–391, Mar. 2019, doi: 10.1038/s41560-019-0356-8.

[26] F. Zhao, Y. Guo, and B. Chen, "A Review of Lithium-Ion Battery State of Charge Estimation Methods Based on Machine Learning," *World Electr. Veh. J.*, vol. 15, no. 4, p. 131, Mar. 2024, doi: 10.3390/WEVJ15040131.

[27] M. Sedlařík *et al.*, "Advanced machine learning techniques for State-of-Health estimation in lithium-ion batteries: A comparative study," *Electrochim. Acta*, vol. 524, p. 145988, Jun. 2025, doi: 10.1016/J.ELECTACTA.2025.145988.

[28] J. Wu, J. Meng, M. Lin, W. Wang, J. Wu, and D.I. Stroe, "Lithium-ion battery state of health estimation using a hybrid model with electrochemical impedance spectroscopy," *Reliab. Eng. Syst. Saf.*, vol. 252, p. 110450, Dec. 2024, doi: 10.1016/J.RESS.2024.110450.

[29] J. Nkechinyere Njoku, C. Ifeanyi Nwakanma, and D.S. Kim, "Explainable Data-Driven Digital Twins for Predicting Battery States in Electric Vehicles," *IEEE Access*, vol. 12, pp. 83480–83501, 2024, doi: 10.1109/ACCESS.2024.3413075.

[30] M.R. Shadi, H. Mirshekali, and H.R. Shaker, "Explainable artificial intelligence for energy systems maintenance: A review on concepts, current techniques, challenges, and prospects," *Renew. Sustain. Energy Rev.*, vol. 216, p. 115668, Jul. 2025, doi: 10.1016/J.RSER.2025.115668.

[31] J. Nkechinyere Njoku, C. Ifeanyi Nwakanma, and D.S. Kim, "Explainable Data-Driven Digital Twins for Predicting Battery States in Electric Vehicles," *IEEE Access*, vol. 12, pp. 83480–83501, 2024, doi: 10.1109/ACCESS.2024.3413075.

[32] B. Xia, Z. Qin, and H. Fu, "Rapid estimation of battery state of health using partial electrochemical impedance spectra and interpretable machine learning," *J. Power Sourc.*, vol. 603, p. 234413, May 2024, doi: 10.1016/J.JPOWSOUR.2024.234413.

[33] C. Lai, "Accurately estimating internal temperature of lithium-ion batteries based on the distribution of relaxation time and data-driven," Mendeley Data, V1, 2024, doi: 10.17632/B6V7FM27FK.1.

[34] X. Chen *et al.*, "Accurately estimating internal temperature of lithium-ion batteries based on the distribution of relaxation time and data-driven," *Energy*, vol. 320, p. 135493, Apr. 2025, doi: 10.1016/J.ENERGY.2025.135493.

[35] T. Paul, P.W. Chi, P.M. Wu, and M.K. Wu, "Computation of distribution of relaxation times by Tikhonov regularization for Li ion batteries: usage of L-curve method," *Sci. Rep.*, vol. 11, no. 1, pp. 1–9, Jun. 2021, doi: 10.1038/s41598-021-91871-3.

[36] M. Saccoccio, T.H. Wan, C. Chen, and F. Ciucci, "Optimal Regularization in Distribution of Relaxation Times applied to Electrochemical Impedance Spectroscopy: Ridge and Lasso Regression Methods – A Theoretical and Experimental Study," *Electrochim. Acta*, vol. 147, pp. 470–482, Nov. 2014, doi: 10.1016/J.ELECTACTA.2014.09.058.

[37] S.M. Lundberg and S.I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Proc. 31st International Conference on Neural Information Processing Systems NIPS'17*, May 2017, pp. 4766–4775, doi: 10.5555/3295222.3295230.

[38] M.T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?' Explaining the predictions of any classifier," in *Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2016, pp. 1135–1144, doi: 10.1145/2939672.2939778.