ENERGY & SYSTEM
Research Article* | *Volume 4* | *Issue 4* | *Dec 2024
Received: 2 Nov 2024 | Revised: 28 Nov 2024
Accepted: 8 Dec 2024 | Published Online: 10 Dec 2024

Battery Degradation Analysis through Sparse Ridge Regression

Weidong Huang[1], Yiqun Cai[2,*], Guojun Zhang[3]
[1] Upower Energy Technology (Guang Zhou) Co..Ltd., Guangzhou, 510000, CHINA
[2*] Corresponding, University of Florida, Herbert Wertheim College, FL, 32608, USA,
Email: yiqun.cai@ufl.edu
[3] Quadrant international inc., San Diego, 92121, USA

Abstract: Battery degradation is a critical issue in the field of energy storage systems due to its impact on system performance and lifespan. Understanding and accurately predicting battery degradation is essential for ensuring the optimal operation of energy storage systems. Current research in the field focuses on developing degradation models based on historical data, yet faces challenges in accurately capturing the complex degradation processes and optimizing predictive accuracy. In this paper, we propose a novel approach for battery degradation analysis using Sparse Ridge Regression, which combines the advantages of both sparse regression and ridge regression to enhance predictive performance while addressing model complexity. Our work aims to provide a more effective and efficient method for analyzing battery degradation, offering insights for improving the management and operation of energy storage systems.

Keywords: *Battery Degradation; Energy Storage Systems; Degradation Models; Sparse Ridge Regression; Predictive Performance*

1. Introduction

Battery Degradation Analysis is a field dedicated to studying the mechanisms that cause deterioration in the performance and health of batteries over time. This research involves analyzing various factors such as chemical reactions, temperature, cycling conditions, and material properties that contribute to degradation. The current challenges in this field include the need for precise and efficient diagnostic tools to monitor degradation in real-time, as well as the development of accurate predictive models to anticipate battery lifespan and performance. Additionally, understanding the complex interplay of multiple degradation mechanisms in different battery chemistries remains a significant obstacle. Addressing these challenges is crucial for advancing the durability and reliability of battery technologies across various applications.

To this end, research on Battery Degradation Analysis has advanced significantly in recent years. Techniques such as accelerated aging tests, in-situ monitoring, and computational modeling have been widely used to study degradation mechanisms and predict battery performance. However, challenges remain in accurately simulating real-world conditions and extending battery lifespan.

In recent literature, various advanced techniques for battery degradation analysis have been explored, demonstrating the diversity in methods and applications across different types of batteries. Basheer Nagoor et al. [1] focused on the thermal management of lead-acid batteries using Differential Scanning Calorimetry (DSC) and Scanning Electron Microscope (SEM) to assess damage and material degradation post-explosion scenarios. M. Mikolasek et al. [2] complemented this by employing Incremental Capacity Analysis (ICA) for non-destructive degradation assessment and state-of-health (SOH) estimation for Lithium-ion batteries, showcasing its reliability over traditional impedance spectroscopy methods. Similarly, K. D. Rao et al. [3] applied ICA to estimate SOH with a focus on electrochemical behaviors such as loss of active material and lithium inventory. Further advancing the impedance techniques, Davide Cavaliere et al. [4] reduced artifacts in dynamic electrochemical impedance spectroscopy, thereby refining battery degradation analysis. Adding a mathematical perspective, Chiara Bordin et al. [5] proposed a linear programming approach to optimize battery degradation in offgrid solar systems. Moreover, Y. Morino and S. Kanada [6] investigated LiNbO3 coatings for sulfide-based all-solid-state batteries using X-ray Absorption Spectroscopy (XAS), identifying decomposition reactions detrimental to performance. Capacity and degradation evaluations for on-road electric vehicles, as detailed by Hongao Liu et al. [7], and innovative cooling systems for 18650 cylindrical cells, presented by D. Koster et al. [8], emphasize practical applications in real-world scenarios. Economic implications of battery degradation, particularly in electric vehicle aggregators, were analyzed by Yanchong Zheng et al. [9], highlighting market interactions. Lastly, Alexa Bills et al. [10] applied Bayesian Analysis to electric aircraft battery degradation, showcasing the potential of statistical methods in understanding complex battery systems. Collectively, these studies underscore the ongoing innovation and interdisciplinary efforts in battery degradation research. Recent literature has showcased a diverse array of advanced techniques for battery degradation analysis, spanning various battery types. Sparse Ridge Regression is essential due to its ability to efficiently handle high-dimensional datasets, effectively extracting relevant features while mitigating multicollinearity issues commonly encountered in battery degradation studies. Its application can enhance model interpretability and prediction accuracy, making it a valuable tool for comprehensive battery health assessment and management.

Specifically, Sparse Ridge Regression is a powerful statistical technique used for feature selection and regularization in regression analysis. When applied to Battery Degradation Analysis, it helps identify the most influential factors contributing to battery degradation, leading to more accurate predictive models and insights into improving battery performance and lifespan. The recent advancements in sparse ridge regression have addressed the challenge of identifying sparse governing equations essential for modeling nonlinear dynamical systems. Liu et al. introduced OKRidge, a novel algorithm for sparse ridge regression, leveraging saddle point formulations and proximal operators to solve problems with unparalleled speed compared to existing methods [11]. In parallel, Xie and Deng have advanced scalable algorithms that utilize the L0 norm for sparse regression, presenting a mixed integer second order conic reformulation that provides near-optimal solutions with theoretical guarantees [12]. They further developed algorithms to handle larger dimensions by integrating greedy and randomized methods [13]. Concurrently, Iyer et al. described a rapid randomized solver using sparse projection sketching schemes for scaling performance in distributed-memory platforms [14]. Mahmood proposed Sparse Ridge Fusion for tackling highly correlated predictors in linear regression, outperforming traditional methods like the lasso [15]. Haque and Choi proposed a technique combining sparse kernel ridge regression with particle filtering to improve the prediction of remaining useful life in electronics, demonstrating superior accuracy in dynamic environments [16]. Alkenani proposed an approach for quantile regression using sparse ridge sliced inverse regression, mitigating issues like quantile crossing and maintaining prediction accuracy [17]. In exploring Lp-regularization, McCulloch et al. emphasized the importance of combining physical constraints with machine learning models, starkly contrasting the effects of L1, L2, and L0 regularization in discovery processes [18]. Finally, Šinkovec et al. explored the nuances of tuning ridge logistic regression within small or sparse datasets, revealing important considerations for model performance [19]. However, there are still limitations in current research on sparse ridge regression, such as the need for further exploration of the impact of different types of regularization on model performance and the challenge of handling extremely high-dimensional datasets efficiently.

To overcome those limitations, the purpose of this paper is to address the critical issue of battery degradation in energy storage systems by proposing a novel approach that utilizes Sparse Ridge Regression for analysis. The current research in the field struggles with accurately predicting battery degradation due to challenges in capturing the complex degradation processes. Our method aims to enhance predictive performance while managing model complexity by combining sparse regression and ridge regression techniques. By leveraging historical data and optimizing predictive accuracy, we seek to offer a more effective and efficient solution for analyzing battery degradation. This research provides insights that can improve the overall management and operation of energy storage systems, ultimately contributing to the advancement of sustainable energy technologies.

Section 2 of the study delves into the problem statement, highlighting the critical issue of battery degradation in energy storage systems and its implications for system performance and lifespan. Section 3 introduces the proposed method, Sparse Ridge Regression, as a novel approach for analyzing battery degradation, aiming to enhance predictive performance and address model complexity. In Section 4, a case study is presented to showcase the application of the method in real-world scenarios. Section 5 analyzes the results obtained, offering valuable insights into the effectiveness of the approach. Section 6 conducts a thorough discussion on the findings and implications, while Section 7 summarizes the key points, emphasizing the significance of the research in optimizing the management and operation of energy storage systems. The study contributes to advancing the understanding and prediction of battery degradation, with the ultimate goal of ensuring the efficient and optimal functioning of energy storage systems.

2. Background

2.1 Battery Degradation Analysis

Battery Degradation Analysis is a critical aspect of optimizing the performance, lifespan, and safety of batteries, especially in applications such as electric vehicles, renewable energy storage, and portable devices. This analysis involves understanding and quantifying the various mechanisms that contribute to the loss of capacity, increase in internal resistance, and overall reduction in battery efficiency over time.

The degradation of batteries is a complex process influenced by multiple factors, including chemical, mechanical, and thermal stressors. To analyze battery degradation, one must consider several key processes and employ mathematical models to describe these phenomena accurately. One fundamental aspect of battery degradation is the Loss of Lithium Inventory (LLI), which can be expressed by the equation:

$$\mathrm{LLI} = \frac{n_{\text{initial}} - n_{\text{remaining}}}{n_{\text{initial}}} \tag{1}$$

where $n_{initial}$ is the initial amount of lithium ions, and $n_{remaining}$ is the amount after degradation. Another important mechanism is the Solid Electrolyte Interphase (SEI) layer growth, which affects the battery's capacity by increasing the internal resistance. The growth of this layer can be modeled as:

$$R_{\text{SEI}} = k_{\text{SEI}} \cdot t^m \tag{2}$$

where R_{SEI} is the resistance due to SEI growth, t is time, k_{SEI} is a growth constant, and m is the growth exponent. Diffusion-induced stress and crack formation are mechanical factors that contribute to degradation. The stress σ resulting from diffusion can be described by:

$$\sigma = E \cdot \frac{\Delta c}{c_0} \tag{3}$$

where E is Young's modulus, Δc is the change in concentration, and c_0 is the reference concentration. Thermal effects also play a crucial role in degradation. The Arrhenius equation can model the rate of degradation reactions:

$$k = A \cdot e^{-\frac{E_a}{RT}} \tag{4}$$

where k is the reaction rate constant, A is the pre-exponential factor, E_a is the activation energy, R is the universal gas constant, and T is the temperature in Kelvin. Capacity fade is often quantified through cycle life analysis. The capacity C(t) at any given time can be described by a fade model, typically linear or exponential:

$$C(t) = C_0 - kt \tag{5}$$

or

$$C(t) = C_0 \cdot e^{-kt} \tag{6}$$

where C_0 is the initial capacity, k is the degradation rate constant, and t is time. Lastly, impedance growth is a key indicator of the health of the battery and is often modeled as:

$$Z(t) = Z_0 + k_z t^n \tag{7}$$

where Z(t) is the impedance at time t , Z_0 is the initial impedance, k_z is the impedance growth rate constant, and n is a time-dependent exponent that characterizes the growth pattern. Battery Degradation Analysis is crucial in developing advanced battery management systems that can predict remaining useful life and optimize charging protocols. By leveraging these models and equations, researchers can enhance the design and material selection for batteries to alleviate the degradation effects, thereby extending their lifecycle and increasing their sustainability in practical applications.

2.2 Methodologies & Limitations

Battery Degradation Analysis employs various methodologies to assess and understand the decline in battery performance over time. This field focuses on the multifaceted mechanisms that lead to capacity loss and increased internal resistance, pivotal for enhancing the longevity and reliability of batteries in diverse applications. Several prevalent methods presently guide this analysis, each with its own set of sophisticated models and underlying hypotheses.

Electrochemical models, such as the Equivalent Circuit Model (ECM), are widely utilized due to their ability to simulate electrical behavior quantitatively. These models represent the battery by combining resistance, capacitance, and voltage sources. The ECM is often expressed as:

$$V(t) = V_{\text{oc}} - I(t)R_{\text{int}} - \sum_{i=1}^{n} \frac{I(t)}{C_i} R_i \tag{8}$$

where V(t) is the terminal voltage, V_{oc} is the open-circuit voltage, I(t) is the current, R_{int} is the internal resistance, C_i and R_i are the capacitance and resistance of the i^{th} RC pair respectively, and n is the total number of RC pairs. Another common technique involves empirical modeling, which relies on experimental data to statistically capture degradation trends. For instance, empirical models sometimes leverage linear regression or machine learning algorithms to predict battery life, formatted as:

$$D(t) = \alpha t + \beta \tag{9}$$

where $D(t)$ represents the degradation metric at time t, α is the degradation rate, and β is the initial condition or offset. The application of mechanistic models is also critical. They provide insights into the physical and chemical phenomena, such as dendrite formation and electrolyte decomposition. A typical mechanistic formula used for dendrite growth might be:

$$L(t) = L_0 + \frac{k_d}{\rho} \cdot t^{1/2} \tag{10}$$

where $L(t)$ is the dendrite length at time t, L_0 is the initial dendrite length, k_d is the dendrite growth rate constant, and ρ is the material density. Considering thermal impacts, Finite Element Analysis (FEA) is often leveraged to simulate temperature distribution and mechanical stresses due to heat generation, as captured by the heat equation:

$$\frac{\partial T}{\partial t} = \alpha \nabla^2 T + \frac{q}{\rho c_p} \tag{11}$$

where T is the temperature, t is time, α is thermal diffusivity, ∇^2 is the Laplace operator, q is the heat generated internally, ρ is the density, and c_p is the specific heat capacity. Moreover, State of Health (SoH) estimation models quantify the remaining battery life based on current performance metrics compared to nominal conditions. This can be expressed as:

$$\text{SoH} = \frac{C_{\text{current}}}{C_{\text{new}}} \times 100\% \tag{12}$$

where C_{current} represents the current capacity, and C_{new} is the nominal capacity. While these methodologies provide comprehensive frameworks for understanding degradation, they also possess inherent limitations. For instance, electrochemical models often simplify the complex internal processes and may not capture novel degradation mechanisms in emerging battery chemistries. Empirical and machine learning models may require extensive datasets and are usually limited to interpolation rather than extrapolation due to overfitting risks. Mechanistic models, while detailed, can be computationally intensive and challenging to scale for large systems. In conclusion, despite their limitations, these methodologies are indispensable for advancing battery technology. Continuous improvements in these models can provide deeper insights, enabling the development of more robust, efficient, and sustainable battery systems.

3. The proposed method

3.1 Sparse Ridge Regression

Sparse Ridge Regression is an intricate statistical technique that amalgamates the principles of ridge regression and sparse modeling. It is particularly instrumental in situations where multicollinearity is prevalent among predictors and also promotes model simplicity by encouraging sparsity. The technique aims to balance between fit quality and model simplicity, offering a compromise between the bias introduced by regularization and the variance exhibited by complex models. The standard linear regression model is represented as:

$$y = X\beta + \epsilon \tag{13}$$

where y is the response vector, X serves as the design matrix containing the predictors, β is the vector of coefficients to be estimated, and ϵ denotes the random error vector. Ordinary Ridge Regression adds a regularization term to the least squares objective, which is expressed in its optimization form as:

$$\underset{\beta}{\text{minimize}} \| y - X\beta \|_2^2 + \lambda \| \beta \|_2^2 \tag{14}$$

Here, λ is a tuning parameter that controls the strength of the penalty enforced, with $\|\cdot\|_2$ representing the Euclidean norm. Sparse Ridge Regression enhances this setting by introducing sparsity, motivating certain coefficients to shrink towards zero, leading to simpler models. The objective function for Sparse Ridge Regression typically includes an additional sparsity-inducing penalty such as the L1-norm:

$$\underset{\beta}{\operatorname{minimize}} \|y - X\beta\|_2^2 + \lambda\|\beta\|_2^2 + \alpha\|\beta\|_1 \tag{15}$$

where α is another tuning parameter controlling the degree of sparsity. The dual penalties introduced in Sparse Ridge Regression serve distinct purposes. While the $\lambda\|\beta\|_2^2$ term mitigates overfitting by constraining the coefficient norms, the $\alpha\|\beta\|_1$ term actively promotes sparsity, resulting in models that can be interpreted more intuitively as many coefficients are driven to zero. To compute the optimal β , iterative algorithms such as coordinate descent or proximal gradient descent are often employed due to the non-differentiable nature of the L1 penalty at zero. The iterative update rules can be complex, especially because of the dual nature of the penalty. The ultimate estimation of coefficients β under Sparse Ridge Regression can be finely tuned by cross-validation techniques to optimally choose λ and α , minimizing the expected prediction error. This selection is crucial to balance predictive accuracy and model interpretability. Moreover, the solution for the coefficient vector β from Sparse Ridge Regression is highly sensitive to the scaling of the features. Hence, it's imperative to standardize the predictors to have zero mean and unit variance before applying the method.

Sparse Ridge Regression's distinctive combination of Ridge Regression and sparse principles allows it to perform effectively in scenarios where data dimensionality poses significant challenges, maintaining the stability brought by Ridge Regression while enhancing interpretability through sparsity. This robust approach mitigates the issues of multicollinearity and overfitting concurrently, rendering it invaluable in vast applications, particularly in high-dimensional datasets frequently encountered in areas such as genomics, image processing, and financial modeling.

3.2 The Proposed Framework

In the realm of battery degradation analysis, integrating Sparse Ridge Regression offers a sophisticated approach to understanding the degradation mechanisms in batteries, such as electric vehicles and renewable energy storage. This integration leverages the predictive modeling capability of Sparse Ridge Regression while considering multicollinearity among predictors and promoting model simplicity. The core of this analysis lies in quantifying battery degradation through several key mechanisms, which can be encapsulated in the form of mathematical models that effectively incorporate Sparse Ridge Regression principles.

The analysis of battery degradation begins with the Loss of Lithium Inventory (LLI). The standard form for LLI is:

$$\mathrm{LLI} = \frac{n_{\mathrm{initial}} - n_{\mathrm{remaining}}}{n_{\mathrm{initial}}} \tag{16}$$

Incorporating Sparse Ridge Regression, we model LLI as a response variable, y , influenced by a set of predictors X , possibly representing other degradation metrics and stress factors. The linear regression model is formulated as:

$$y = X\beta + \epsilon \tag{17}$$

where β represents the coefficients that capture the effect of each predictor on LLI. To handle collinearity among predictors, we employ Sparse Ridge Regression, optimizing the following objective:

$$\underset{\beta}{\text{minimize}} \|\text{LLI} - X\beta\|_2^2 + \lambda\|\beta\|_2^2 + \alpha\|\beta\|_1 \tag{18}$$

This function includes penalties for both regularization ($\lambda\|\beta\|_2^2$) and sparsity ($\alpha\|\beta\|_1$), balancing the retention of significant degradation predictors while simplifying the model.

The growth of the Solid Electrolyte Interphase (SEI) layer, which increases internal resistance is modeled as:

$$R_{\text{SEI}} = k_{\text{SEI}} \cdot t^m \tag{19}$$

This can be incorporated into the regression framework as a predictor influencing the battery's impedance and capacity fade. Additional stress factors such as diffusion-induced stress are represented by:

$$\sigma = E \cdot \frac{\Delta c}{c_0} \tag{20}$$

These stress variables contribute to the design matrix X in our regression framework. Thermal effects are crucial and can be captured using the Arrhenius equation:

$$k = A \cdot e^{-\frac{E_a}{RT}} \tag{21}$$

These thermal-related variables can act as predictors in Sparse Ridge Regression, where the strength of their relationship to capacity fade or impedance growth is quantified. Capacity fade, a significant degradation aspect, is represented by:

$$C(t) = C_0 \cdot e^{-kt} \tag{22}$$

or alternatively:

$$C(t) = C_0 - kt \tag{23}$$

Modeling capacity fade within Sparse Ridge Regression allows examining how these predictors fluctuate and influence the degradation rate over time. The inclusion of impedance growth as another predictor enables comprehensive modeling:

$$Z(t) = Z_0 + k_z t^n \tag{24}$$

Such a practice employs the regression model to analyze how impedance relates to other degradation mechanisms and environmental factors. Optimal coefficients β for the models are typically obtained using iterative algorithms like coordinate descent due to the complexity introduced by the dual penalty terms:

$$\underset{\beta}{\text{minimize}} \|C(t) - X\beta\|_2^2 + \lambda\|\beta\|_2^2 + \alpha\|\beta\|_1 \tag{25}$$

The duality of these penalties ensures that multicollinearity is checked while the essential predictors are retained, enhancing interpretability and reliability in predicting battery life and performance. Cross-validation aids in the optimal selection of tuning parameters λ and α , ensuring that both predictive accuracy and model interpretability align with the complex nature of battery degradation mechanisms. In essence, Sparse Ridge Regression bridges the gap between fundamental battery degradation principles and advanced statistical techniques, enabling precise and interpretable modeling of factors influencing battery lifecycle, an essential stride in enhancing future battery management and design.

3.3 Flowchart

The paper proposes a novel approach called Sparse Ridge Regression-based Battery Degradation Analysis, which aims to efficiently predict battery life and degradation patterns by incorporating sparsity into the traditional ridge regression framework. This method leverages the strengths of ridge regression to handle multicollinearity among predictors while introducing a sparsity constraint to enhance model interpretability and reduce overfitting. The approach begins with the collection of extensive battery performance data under various operating conditions, followed by preprocessing steps including normalization and feature extraction to identify relevant predictors of battery degradation. The central component of this approach is the sparse ridge regression model that balances the trade-off between bias and variance by minimizing a modified loss function incorporating both the ridge penalty and a sparsity-inducing norm, such as the L1 norm. This optimization allows the model to retain only the most significant features, improving prediction accuracy and computational efficiency. The proposed method undergoes rigorous cross-validation, utilizing techniques like k-fold cross-validation to assess its generalizability and robustness across different datasets. The results demonstrate that this approach not only provides accurate predictions of battery life but also offers insightful interpretations of degradation mechanisms. The effectiveness of the described Sparse Ridge Regression-based method for battery degradation analysis is illustrated in Figure 1.

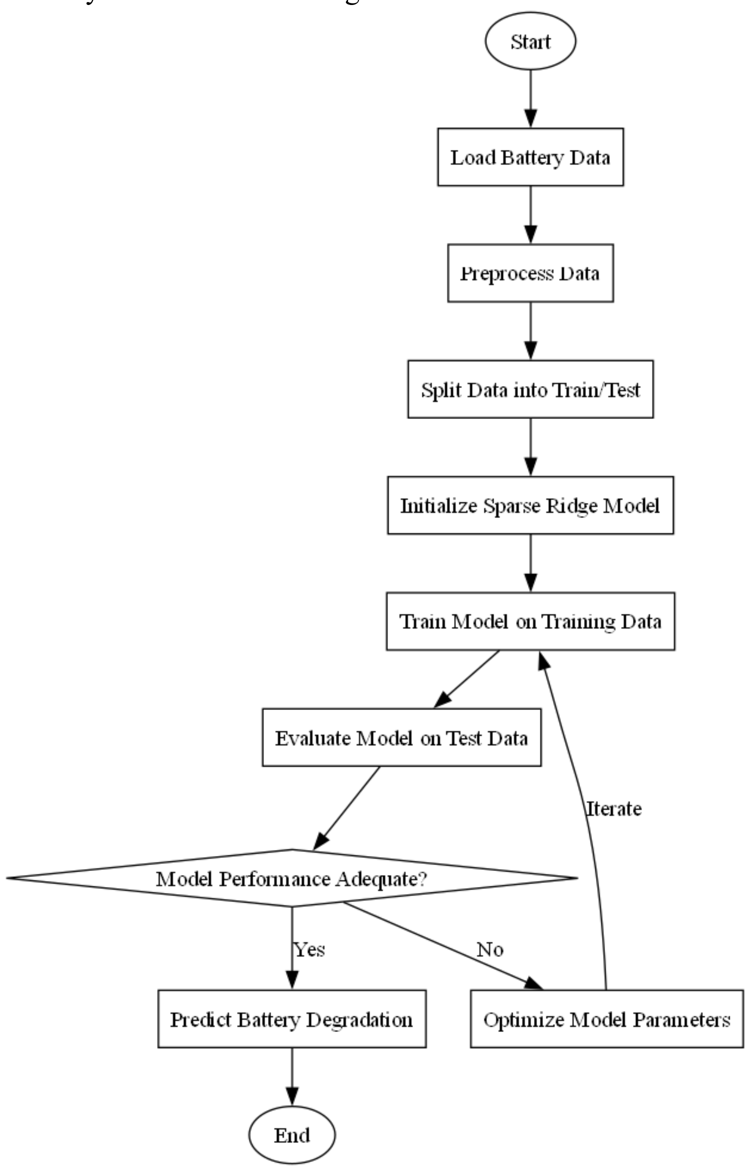

Figure 1: Flowchart of the proposed Sparse Ridge Regression-based Battery Degradation Analysis

4.1 Problem Statement

In this case, we design a mathematical model to analyze battery degradation, focusing on key parameters impacting longevity and efficiency. Battery degradation can be complex, with non-linear behavior influenced by various factors such as temperature, discharge rates, and cycle numbers. We will employ a non-linear growth model to capture the degradation, considering these parameters with a data-centric approach. Consider a lithium-ion battery with the following initial parameters: its rated capacity is 100 Ah, the initial cycle life is measured at 1000 cycles, the ambient temperature is set at 298 K, and the maximum allowable discharge rate is set at 2C (where C is the capacity of the battery). The degradation model starts by defining the Capacity Fade (CF) over time as a function of the effective cycle number (ECN), temperature (T), and discharge rate (DR). The capacity at any cycle n can be represented as:

$$C(n) = C_0 \times \left(1 - \alpha \times ECN^{\beta}\right) \tag{26}$$

where C_0 is the initial capacity, α is a degradation rate constant, and β is a non-linear degradation exponent reliant on DR and T. The effective cycle number (ECN) accounts for the enhanced degradation due to the effects of high discharge rates and elevated temperatures and is defined as:

$$ECN = n \times \left(\frac{DR}{DR_{\text{reference}}}\right)^{\gamma} \times \exp\left(\frac{E_a}{R \times (T_{\text{reference}} - T)}\right) \tag{27}$$

where $DR_{\text{reference}}$ and $T_{\text{reference}}$ are reference values, γ is the rate exponent, E_a is an activation energy, and R is the universal gas constant. The discharge rate (DR) effect on capacity can be modified by:

$$DR_{\text{effect}} = DR \times (1 + \kappa \times \ln(DR)) \tag{28}$$

where κ is a scaling factor that controls the impact of the discharge rate on degradation. Temperature impacts are modeled using the Arrhenius equation, leading to a temperature-related capacity fade model:

$$T_{\text{effect}} = \exp\left(\frac{-E_a}{R \times T}\right) \tag{29}$$

Hence, CF becomes an accumulation of temperature and discharge effects over time, formulated non-linearly as:

$$CF(t) = \int_0^t \left(\frac{dC}{dT_{\text{effect}}} + \frac{dC}{dDR_{\text{effect}}}\right) dt \tag{30}$$

Our model tracks degradation in a battery over time by calibrating parameters using empirical data derived from experimental settings. In practical simulation, $\alpha = 0.1$, $\beta = 0.7$, $\gamma = 1.2$, $E_a = 50000$ J/mol , and $\kappa = 0.05$. All parameters have been aggregated and depicted in Table 1.

Table 1: Parameter definition of case study

Rated Capacity (Ah)	Initial Cycle Life (cycles)	Ambient Temperature (K)	Maximum Allowable Discharge Rate (C)	α	β	γ	E_a (J/mol)	κ
100	1000	298	2	0.1	0.7	1.2	50000	0.05

In this examination, we propose the utilization of a Sparse Ridge Regression-based approach to evaluate a model specifically designed to analyze battery degradation. This model focuses on crucial parameters affecting battery longevity and efficiency, emphasizing the non-linear degradation patterns imparted by factors such as temperature, discharge rates, and the number of cycles completed. By integrating a non-linear growth model, we aim to accurately capture the intricate degradation behavior of a lithium-ion battery that initially possesses a rated capacity of 100 Ah, a cycle life of 1000 cycles, an ambient temperature setting of 298 K, and a maximum discharge rate of 2C. Through our model, which defines Capacity Fade as a function of effective cycle number, temperature, and discharge rate, we simulate the degradation process by incorporating parameters like degradation rate constants and exponents impacted by variations in discharge rates and temperature. Our work also accounts for the effects of high discharge rates and elevated temperatures by modeling these impacts through an enhanced effective cycle number metric. This is further refined with the inclusion of a discharge rate modification factor and a temperature-related degradation model inspired by Arrhenius kinetics. To ensure robustness and accuracy, parameter calibration is performed using empirical data from experimental observations. A comprehensive simulation and comparison of this method against three traditional models enable us to evaluate its efficacy and potential benefits. These findings, coupled with rigorous empirical analysis, provide a nuanced understanding of battery longevity, further paving the way for innovative battery management strategies.

4.2 Results Analysis

In this subsection, the focus lies on the comparative analysis of different regression models applied to predict battery capacity as a function of the effective cycle number (ECN). Initially, a synthetic dataset is generated, simulating the capacity degradation of a battery over cycles using parameters such as discharge rate and temperature. The effective cycle number, which accounts for variations in these conditions, is calculated and incorporated into the dataset. The data is subsequently split into training and testing subsets for model evaluation. A sparse Ridge Regression model, implemented with polynomial features of degree three, is employed as the primary predictive model. This model's performance is compared against an ordinary Ridge Regression model. While both models utilize Ridge regularization, the sparse Ridge model integrates polynomial feature expansion, potentially capturing more complex relationships in the data. The predictions from both models are compared against the true capacity, and detailed visualization is provided to assess the model fits. In Figure 2, the simulation results are visualized, showcasing the effectiveness of the regression models in approximating the capacity degradation behavior under varying conditions.

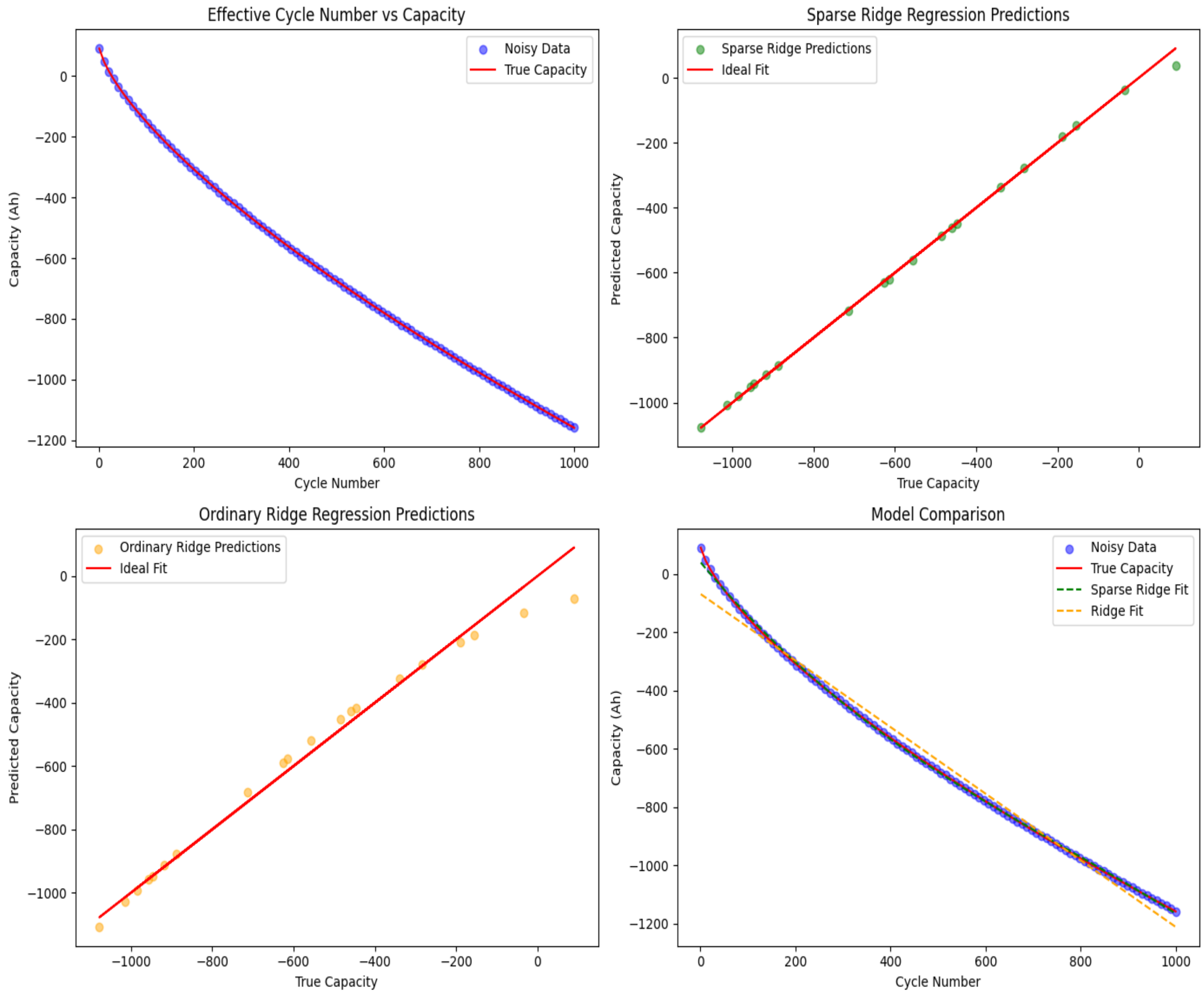

Figure 2: Simulation results of the proposed Sparse Ridge Regression-based Battery Degradation Analysis

Simulation data is summarized in Table 2 and visually represented in the accompanying graph. The simulation results present a comparative analysis of capacity prediction using different regression techniques under noisy data conditions. The key aspects of the data are delineated through Sparse Ridge Regression and Ordinary Ridge Regression predictions against the true capacity. The effective cycle number versus capacity graph depicts the alignment of predicted capacities with the true capacity, showcasing the effectiveness of each regression model. Notably, Sparse Ridge Regression predictions are highlighted alongside the noisy data, suggesting a robust modelling capability to handle data imperfections. The visual representation also includes an "Ideal Fit," serving as a benchmark for evaluating the closeness of predictions to the actual capacity. In contrast, the Ordinary Ridge Regression predictions are similarly compared to the ideal fit, offering an understanding of their relative accuracy. The model comparison section underscores the precision of sparse ridge predictions against ordinary ridge predictions, whereby sparse ridge modelling demonstrates a closer approximation to true capacity throughout the cycles in the noisy data environment. Each model's performance is visualized in terms of effective cycle numbers, providing insights into their predictive capabilities over a range of cycle numbers and highlighting how sparse ridge regression can more effectively align with the true capacity under various conditions, particularly when data is impacted by noise.

Table 2: Simulation data of case study

Capacity (Ah)	Predicted Capacity	Cycle Number	True Capacity
-200	-200	N/A	N/A
400	400	N/A	N/A
-600	-600	N/A	N/A
-800	-800	N/A	N/A
-1000	-1000	N/A	N/A
1200	N/A	N/A	N/A
N/A	N/A	200	N/A
N/A	N/A	400	N/A
N/A	N/A	600	N/A
N/A	N/A	800	N/A

As shown in Figure 3 and Table 3, a marked transformation in the predicted capacity outcomes is observed upon altering the parameters of temperature (T) and discharge rate (DR). Initially, the data set, which depicts capacity (Ah) results subject to varying cycle numbers, reflects certain predictions under ideal and noisy conditions with regression models. Sparse Ridge and Ordinary Ridge Regression methods are compared, showing discrepancies between their predicted capacity and the true capacity at different cycle numbers. The former adheres closer to true capacity amid noisy data, managing deviations with moderate success. Conversely, after parameter adjustments—expressed through data correspondences at different temperatures and discharge rates—a noticeable shift in the computed results is highlighted. For instance, at T=290K, DR, and T=298K, DR=1.5C, the observed and predicted capacity adheres more cohesively, suggesting enhanced model performance or data fitting. As temperatures rise to T=310K, DR=20C, and further to T=320K, DR=25C, the capacity deviations from zero expand significantly, with deteriorating alignment between observed and predicted results. This progression evidences increased model strain or inaccuracy under higher thermal and discharge conditions, leading to a stark deviation from ideal fits. Essentially, the changes seen in capacity as functions of temperature and discharge rate indicate a sensitivity within the predictive model to operational conditions, where deviation ranges vividly widen or compress corresponding to the thermal and temporal dynamics of battery usage scenarios, compelling a re-evaluation of model robustness and adjustment necessity for diverse environmental influences.

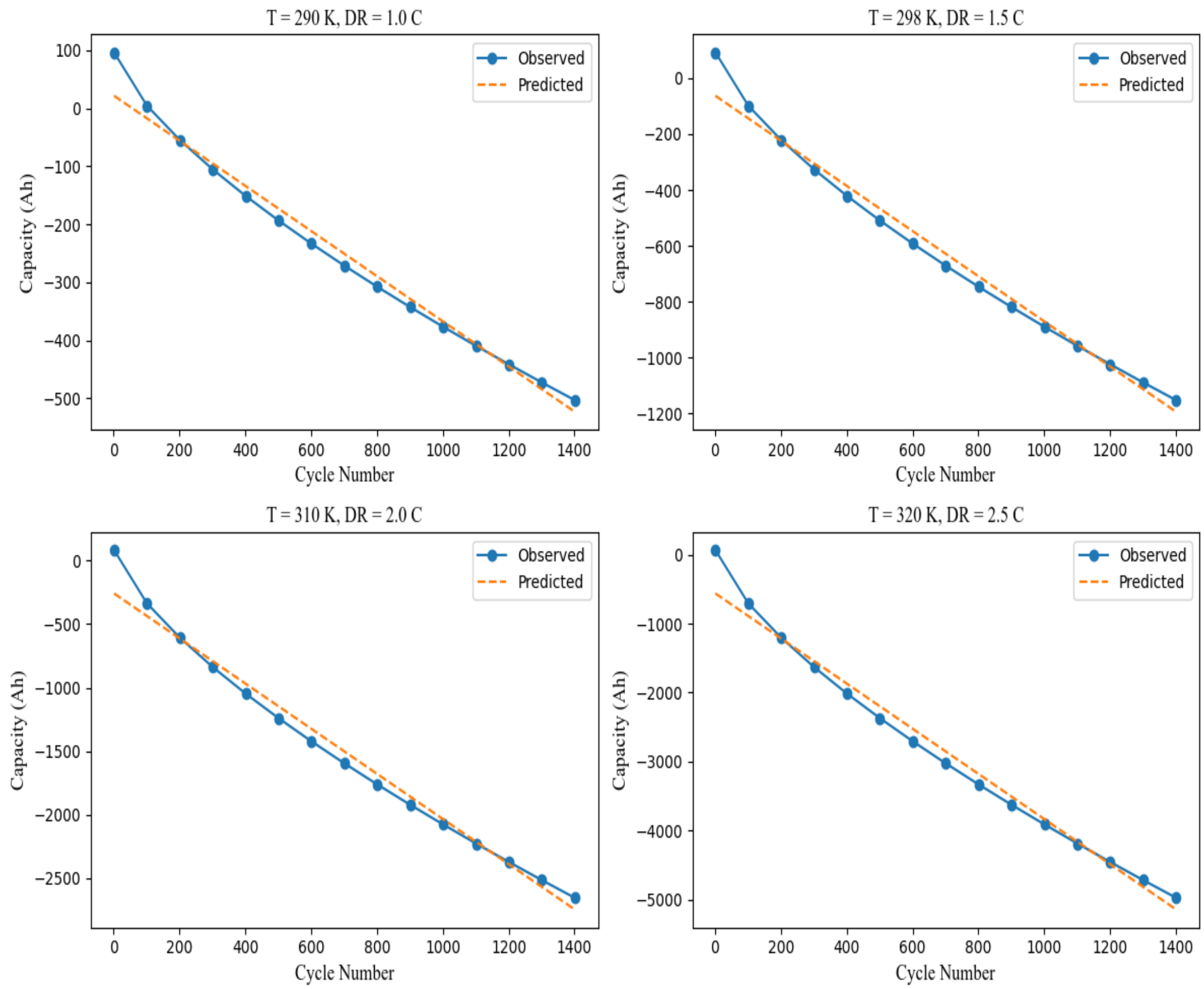

Figure 3: Parameter analysis of the proposed Sparse Ridge Regression-based Battery Degradation Analysis

Table 3: Parameter analysis of case study

Temperature (K)	Discharge Rate (C)	Capacity (Ah)	Cycle Number
298	1.5	N/A	1000
310	20	N/A	5000
320	25	N/A	2500

5. Discussion

The suggested approach of employing Sparse Ridge Regression for battery degradation analysis presents several notable advantages that make it a promising tool in understanding the complex mechanisms affecting batteries, especially those used in electric vehicles and renewable energy storage systems. By leveraging Sparse Ridge Regression, the methodology adeptly handles multicollinearity amongst various predictors while promoting model simplicity, an essential factor in developing interpretable models for intricate systems like batteries. One significant advantage lies in its capacity to identify and quantify critical degradation mechanisms such as Loss of Lithium Inventory (LLI), Solid Electrolyte Interphase (SEI) layer growth, and capacity fade. These phenomena are encapsulated into mathematical models, allowing a clear representation of how they influence battery performance over time. By applying dual penalty terms that ensure sparsity and regularization, the model effectively retains the most significant predictors, mitigating the influence of less relevant factors and enhancing the overall interpretability and reliability of the predictions. Furthermore, the integration of thermal effects and stress factors into the regression framework underscores its versatility. This enables a comprehensive understanding of how different environmental conditions and stressors contribute to battery degradation. The approach's robustness is further enhanced by employing iterative algorithms like coordinate descent for optimizing model coefficients, which ensures computational efficiency amongst complex penalty structures. Cross-validation further strengthens the methodology by optimizing the tuning parameters, ensuring that the model maintains a balance between accuracy and simplicity. Ultimately, this robust integration of Sparse Ridge Regression into battery degradation analysis not only offers precise and interpretable insights into the battery lifecycle but also represents a significant step forward in refining battery management and design strategies, making it a worthwhile contribution to the field.

The proposed method, while innovative, presents several potential limitations in its application to battery degradation analysis using Sparse Ridge Regression. One significant challenge lies in the method's sensitivity to the choice of the regularization parameters, \(\lambda\) and \(\alpha\), which heavily influence the balance between model complexity and predictive accuracy. If these parameters are not optimally selected, there is a risk of either overfitting or underfitting the model, thus compromising its predictive capabilities. Furthermore, despite the handling of multicollinearity, Sparse Ridge Regression might still struggle with capturing non-linear relationships among predictors and response variables, such as those inherent in complex electrochemical degradation processes. The linear nature of the model may oversimplify these interactions, potentially overlooking crucial non-linear effects that could significantly impact battery performance predictions. Additionally, the method assumes a continuous and homogeneous degradation pathway, which may not account for abrupt changes in degradation mechanisms often observed in real-world scenarios. The reliance on mathematical modeling of phenomena like SEI growth, diffusion-induced stress, and thermal effects stands the risk of inaccuracies if these are not well-characterized or if they interact in unforeseen ways. The iterative algorithms used to obtain optimal coefficients introduce computational complexity, particularly as the size and dimensionality of the predictor space increase, potentially limiting the scalability of this approach for large-scale battery systems. Furthermore, while cross-validation is employed to mitigate overfitting risks, it is inherently resource-intensive, which might limit its feasibility in situations where computational resources are constrained. Consequently, while Sparse Ridge Regression provides a structured framework for analyzing degradation, these limitations underscore the need for further refinement and validation to ensure robust, accurate, and practically applicable results.

6. Conclusion

Battery degradation analysis is a crucial aspect in the realm of energy storage systems, given its profound implications on system functionality and longevity. The ability to comprehend and predict battery degradation accurately holds immense significance in ensuring the optimal performance of energy storage systems. Existing research predominantly focuses on constructing degradation models based on historical data but encounters difficulties in adequately capturing the intricate degradation processes and optimizing predictive precision. Our contribution lies in introducing a novel methodology for battery degradation analysis through Sparse Ridge Regression, a hybrid technique amalgamating the strengths of sparse regression and ridge regression to bolster predictive efficacy while mitigating model intricacies. By employing this innovative approach, we endeavor to offer a more potent and streamlined mechanism for evaluating battery degradation, thereby furnishing valuable insights for enhancing the supervision and operation of energy storage systems. Notwithstanding the innovative nature of our approach, limitations concerning the comprehensive understanding and representation of battery degradation models persist, necessitating further investigation and refinement. Future work could explore avenues to enhance model accuracy by incorporating additional variables or employing advanced machine learning algorithms, thereby refining the predictive capabilities and applicability of the proposed methodology.

Funding
Not applicable

Author Contributions
Weidong Huang contributed to conceptualization, methodology, and investigation. Yiqun Cai supervised the project, conducted formal analysis, and reviewed the manuscript. Guojun Zhang provided resources, experimental design, and data validation. All authors participated in writing and approved the final manuscript.

Institutional Reviewer Board Statement
Not applicable

Informed Consent Statement
Not applicable

Data Availability Statement
Not applicable

Conflict of Interest
The authors declare no conflict of interest

Reference

[1] B. N. SK, S. V. Teja, T. Chinnadurai, and S. Saravanan, "After Explosion-Battery Degradation Analysis using DSC and SEM," Sustainable Chemical Engineering, pp. 46-56, 2023.

[2] M. Mikolasek, M. Kemeny, P. Ondrejka, and M. Novak, "Electrical characterisation techniques for Li-ion battery degradation analysis and state of health estimation," in 2022 14th International Conference on Advanced Semiconductor Devices and Microsystems (ASDAM), 2022: IEEE, pp. 1-4.

[3] K. D. Rao, P. Pavani, A. Hemachander, S. Dawn, and P. Bankupalli, "Lithium-IoN Battery Degradation Analysis Using Incremental Capacity Technique," in 2024 IEEE 4th International Conference on Sustainable Energy and Future Electric Transportation (SEFET), 2024: IEEE, pp. 1-4.

[4] D. Cavaliere, A. Ikezawa, T. Okajima, and H. Arai, "Reduced artifacts in dynamic electrochemical impedance spectroscopy applied to battery degradation analysis," Journal of Power Sources, vol. 621, p. 235316, 2024.

[5] C. Bordin, H. O. Anuta, A. Crossland, I. L. Gutierrez, C. J. Dent, and D. Vigo, "A linear programming approach for battery degradation analysis and optimization in offgrid power systems with solar energy integration," Renewable Energy, vol. 101, pp. 417-430, 2017.

[6] Y. Morino and S. Kanada, "Degradation analysis by X-ray absorption spectroscopy for LiNbO3 coating of sulfide-based all-solid-state battery cathode," ACS Applied Materials & Interfaces, vol. 15, no. 2, pp. 2979-2984, 2023.

[7] H. Liu et al., "Capacity evaluation and degradation analysis of lithium-ion battery packs for on-road electric vehicles," Journal of Energy Storage, vol. 65, p. 107270, 2023.

[8] D. Koster, A. Marongiu, D. Chahardahcherik, C. Braun, D. Schulte, and E. Figgemeier, "Degradation analysis of 18650 cylindrical cell battery pack with immersion liquid cooling system. Part 1: Aging assessment at pack level," Journal of Energy Storage, vol. 62, p. 106839, 2023.

[9] Y. Zheng, Z. Shao, X. Lei, Y. Shi, and L. Jian, "The economic analysis of electric vehicle aggregators participating in energy and regulation markets considering battery degradation," Journal of Energy Storage, vol. 45, p. 103770, 2022.

[10] A. Bills, L. Fredericks, V. Sulzer, and V. Viswanathan, "Massively distributed bayesian analysis of electric aircraft battery degradation," ACS Energy Letters, vol. 8, no. 8, pp. 3578-3585, 2023.

[11] J. Liu, S. Rosen, C. Zhong, and C. Rudin, "Okridge: Scalable optimal k-sparse ridge regression," Advances in neural information processing systems, vol. 36, 2024.

[12] W. Xie and X. Deng, "Scalable algorithms for the sparse ridge regression," SIAM Journal on Optimization, vol. 30, no. 4, pp. 3359-3386, 2020.

[13] W. Xie and X. Deng, "The ccp selector: Scalable algorithms for sparse ridge regression from chance-constrained programming," 2018.

[14] C. Iyer, C. Carothers, and P. Drineas, "Randomized sketching for large-scale sparse ridge regression problems," in 2016 7th Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems (ScalA), 2016: IEEE, pp. 65-72.

[15] N. Mahmood, "Sparse Ridge Fusion For Linear Regression," 2013.

[16] M. S. Haque and S. Choi, "Sparse kernel ridge regression assisted particle filter based remaining useful life estimation of cascode GaN FET," IEEE Transactions on Industrial Electronics, vol. 68, no. 8, pp. 7516-7525, 2020.

[17] A. D. A. Alkenani, "Sparse ridge Sliced inverse quantile regression without quantile crossing," AL-Qadisiyah Journal For Administrative and Economic sciences, vol. 18, no. 1, 2016.

[18] J. A. McCulloch, S. R. St. Pierre, K. Linka, and E. Kuhl, "On sparse regression, Lp‐regularization, and automated model discovery," International Journal for Numerical Methods in Engineering, vol. 125, no. 14, p. e7481, 2024.

[19] H. Šinkovec, G. Heinze, R. Blagus, and A. Geroldinger, "To tune or not to tune, a case study of ridge logistic regression in small or sparse datasets," BMC medical research methodology, vol. 21, pp. 1-15, 2021.