Project Proposal

CSYS5990A - Advacned Machine Learning

Data-driven modelling for Lithium-Ion battery life cycle for capacity degradation.

Muhammad Adeel Arif, Alireza Kashani Lotfabadi

Team: Alpha-Grads (Group 5)

1. Introduction

With the growing demand for improved resilience, sustainability, and predictability in clean energy generation, energy storage devices have become a key focus for the research community. Various models—stochastic, analytical, empirical, semi-empirical, and optimization-based—explore the performance and reliability of these devices, with lithium-ion (Li-ion) batteries being particularly popular due to their high energy density (compact size), deep discharge capacity, high efficiency, low maintenance, and fast charging capabilities.

Manufacturers need accurate predictions of battery discharge performance across different designs and sizes to manage production effectively. Meanwhile, end users are interested in reliable battery life estimates, especially for off-grid applications or when integrated with renewable energy sources [1]. Therefore, developing reliable models that can account for non-linear behaviors and sub-derived parameters to predict capacity degradation over multiple charge-discharge cycles is critical.

Modeling the non-linear behavior of Li-ion batteries is challenging due to the influence of factors such as ambient conditions (temperature and humidity), chemical resistance at the anodes (which increases with progressive discharge), and electrical non-linearities. Traditional methods often struggle with generalization and consistency across different battery makes, models, and sizes.

In contrast, machine learning techniques and data-driven models are becoming increasingly popular for capturing these non-linear battery behaviors, which deterministic models often fail to address [2]. With the growing availability of datasets, researchers are not only improving battery degradation prediction. Still, they are also working on batch classification and categorization based on size and make, offering a more robust understanding of battery performance.

2. Problem Definition and Algorithm

Conventional analytical, empirical, and optimization-based techniques have been effective in predicting the capacity degradation of Li-ion batteries. However, they struggle with estimating degradation under fast-charging conditions due to the non-linear nature of the degradation process, which is further complicated by thermal coupling and mechanical heterogeneity within the cells. These factors are difficult to model accurately using traditional methods. This project aims to develop data-driven models for predicting capacity degradation in Lithium Iron Phosphate (LFP) batteries, a specific type of Li-ion battery known for its enhanced energy density, thermal stability, and longer lifespan. We will focus on modeling degradation under fast-charging conditions, where traditional models fall short. Our approach will analyze features such as discharge voltage curves, internal resistance, and temperature fluctuations from early battery cycles to accurately predict the overall cycle life. We will employ machine learning techniques, specifically Random Forest and Long Short-Term Memory (LSTM) networks, to capture the complex non-linear behaviors and interactions within the battery. One of the key predictive features identified is the variance in voltage curves over the early cycles, which has proven to be highly correlated with capacity degradation, especially during fast charging. The robustness of our models will be evaluated using battery samples with varying rated life cycles and discharge capacities (Ah) to ensure accuracy and generalizability. By leveraging these data-driven approaches, this project aims to provide more accurate predictions of battery lifecycle management, helping manufacturers optimize production and enabling end-users to rely on more predictable battery performance in off-grid and renewable energy applications.

3. Dataset

The dataset for this project consists of 124 commercial **LFP/graphite cells** (A123 Systems, model APR18650M1A, 1.1 Ah nominal capacity), which were cycled under 72 distinct fast-charging conditions. These batteries demonstrated a wide range of cycle lives, from 150 to 2,300 cycles, with end-of-life defined as reaching 80% of nominal capacity. While all cells were discharged under identical conditions, the charging conditions were varied, generating a diverse dataset that captures different degradation behaviors. The experiments were conducted in a temperature-controlled environment (30° C), with continuous measurements of **voltage**, **current**, **internal resistance**, and **cell temperature** recorded throughout the cycling process. This resulted in a comprehensive dataset of approximately 96,700 charge-discharge cycles, making it one of the most extensive datasets available for lithium-ion battery research.

Our dataset has a labeled output in terms of battery voltage, which estimates the capacity degradation as described in Severson et al [3]. The features available for analysis are as follows:

1. **Discharge Capacity**: The amount of charge (typically measured in ampere-hours, Ah) a battery delivers during a discharge cycle. This represents the usable energy drawn from the battery.
2. **Charge Capacity**: The amount of charge (Ah) the battery can store during a charging cycle. This represents how much energy is stored in the battery.
3. **Discharge Energy**: The total energy (typically measured in watt-hours, Wh) delivered by the battery during the discharge cycle. This is a key metric for determining how much usable power the battery provides.
4. **Charge Energy**: The total energy (Wh) absorbed by the battery during the charging process. It represents how much energy is being stored within the battery.
5. **DC Internal Resistance**: The internal resistance (usually measured in ohms) of the battery during discharge. Higher internal resistance can lead to inefficiencies and heat generation, affecting battery performance.
6. **Temperature Maximum**: The maximum temperature (usually in Â°C) reached by the battery during the charging or discharging cycle. Excessive heat can accelerate battery degradation.
7. **Temperature Average**: The average temperature (Â°C) of the battery during operation. Monitoring this helps ensure the battery operates within safe thermal limits.
8. **Temperature Minimum**: The minimum temperature (Â°C) the battery reaches during the cycle. Extreme temperatures, both high and low, can negatively affect battery health and performance.
9. **Energy Efficiency**: The ratio of the energy discharged to the energy charged, typically expressed as a percentage. It provides insight into how effectively the battery stores and releases energy.
10. **Charge Throughput**: The cumulative amount of charge (Ah) delivered into the battery over time, which reflects how much total energy has been processed during charging cycles.
11. **Energy Throughput**: The total amount of energy (Wh) that has passed through the battery during both charging and discharging. It reflects the overall energy handling of the battery over time.

12. **Charge Duration**: The amount of time (typically in minutes or hours) it takes for the battery to complete a charging cycle. Faster charging times are desirable but can contribute to increased wear and degradation.

13. **Time Temperature Integrated**: A metric that tracks how the battery's temperature changes over time, which is critical for understanding thermal dynamics and the effects on battery degradation.

4. Related Work

Our work builds upon the research conducted by Severson et al. [3], using the same publicly available dataset from Stanford University's battery research group. In their study, they applied a semi-empirical approach, using an Elastic Net regularized linear model [4] to estimate battery degradation, and logistic regression to classify batteries based on discharge data from 100 and 5 cycles. They compared their results across three model variants: Variance, Discharge, and Full models.

While their approach is straightforward and powerful, the use of a linear model may limit its ability to fully capture the varying dynamics of battery degradation. This is especially true when considering non-linear behaviors that evolve over different timescales—some changing rapidly in the short term and others more gradually over the long term. For this reason, we aim to explore more advanced algorithms, such as LSTM networks, which can model a deeper, more complex understanding of temporal dynamics. Additionally, if time permits, we plan to extend our work to include dynamic modeling of battery behavior using this dataset.

Su et al. explored various machine learning techniques such as linear regression, neural networks, and convolutional neural networks (CNNs) to predict the cycle life of lithium-ion batteries (LIBs) [5]. Their study demonstrated the capability of CNNs to extract hidden features from early cycling data, resulting in better prediction accuracy than models based on expert-extracted features. Our work differs by specifically targeting Lithium Iron Phosphate (LFP) batteries and addressing the challenges posed by fast-charging conditions, an area that Su et al. did not explore. We also employ machine learning models such as Random Forest and LSTM to capture non-linear battery behaviors specific to fast charging.

Singh et al. [6] built upon the model developed by Severson et al., comparing Gaussian Process Regression (GPR) and Elastic Net Regression (ENR) to predict battery cycle life. GPR was shown to outperform ENR, particularly for longer cycle life predictions. While Singh et al. concentrated on Gaussian Process Regression (GPR) and Elastic Net Regression (ENR), our approach leverages more advanced temporal models like LSTM, which are better suited for fast-charging dynamics.

References

[1] S. B. Peterson, J. Apt, and J. Whitacre, "Lithium-ion battery cell degradation resulting from realistic vehicle and vehicle-to-grid utilization," *Journal of Power Sources*, vol. 195, no. 8, pp. 2385–2392, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378775309017443 (document)

[2] A. Nuhic, T. Terzimehic, T. Soczka-Guth, M. Buchholz, and K. Dietmayer, "Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods," *Journal of power sources*, vol. 239, pp. 680–688, 2013. (document)

[3] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggedakis *et al.*, "Data-driven prediction of battery cycle life before capacity degradation," *Nature Energy*, vol. 4, no. 5, pp. 383–391, 2019. (document)

[4] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 67, no. 2, pp. 301–320, 2005. (document)

[5] L. Su, M. Wu, Z. Li, and J. Zhang, "Cycle life prediction of lithium-ion batteries based on data-driven methods," *ETransportation*, vol. 10, p. 100137, 2021. (document)

[6] A. Singh, C. Feltner, J. Peck, and K. I. Kuhn, "Data driven prediction of battery cycle life before capacity degradation," *arXiv preprint arXiv:2110.09687*, 2021. (document)