Integration of Battery Archive Degradation Data into QuESt Tool for Enhanced Technoeconomic Analysis

By: David Braese, Alexander Headley, and David Rosewater

Motivation

The objective of this project is to develop a Python-based integration with Sandia National Laboratories' QuESt tool to retrieve and analyze battery degradation data from the Battery Archive. This tool will allow users to input specific battery parameters, including chemistry, expected C-rate, and temperature, to obtain degradation rates applicable to their projected use case. The integration will provide upper and lower bounds on degradation impacts, enhancing QuESt's ability to project long-term project value

Background

Understanding the degradation of lithium-ion batteries under various conditions is critical for predicting economic value over time. As energy systems increasingly rely on battery storage, there is a growing need for tools that can incorporate detailed, condition-specific degradation data to improve technoeconomic analyses. The Battery Archive database maintained by Sandia National Laboratories provides publicly available degradation data under varying conditions of chemistry, charge-discharge rates (C-rates), and temperature. However, translating this wealth of data into actionable insights for technoeconomic modeling has been a challenge. This project aims to bridge this gap by developing a tool to retrieve and process data from the Battery Archive to identify degradation rates pertinent to the user's application based on published test data. This work will enable users to input specific parameters and incorporate degradation predictions into their analyses, improving the accuracy of long-term value assessments for energy storage projects.

Methods

Data Collection and Cleaning:

- Cycling data is collected from BatteryArchive.org. Outliers beyond three standard deviations from a rolling mean are removed, and missing values are linearly interpolated.

Capacity Metric Calculation:

- For each cycle, the percent of initial capacity is calculated using the discharge capacity data. This becomes the dependent variable in model training.

Standard Linear Regression:

- A linear regression model is trained with cycle index as the input to estimate capacity degradation over time.

End-of-Life (EOL) Optimized Model:

- This model fixes the intercept at the initial capacity and uses SciPy's optimizes the slope by minimizing the RMSE over the final 10 cycles. This ensures to a better prediction at EOL.

Model Training and Evaluation:

- Models are trained using 5-fold cross-validation for robustness. The model is then compared to a set of holdout data to determine its accuracy. Model performance is compared across chemistries using R² and RMSE metrics.

Results

EOL Model fit:

- This model significantly reduces RSME and increases R^2 compared to the standard linear regression, especially in the end-of-life region where degradation typically increases.
- Chemistries with more linear degradation show very similar performance between both models, while more nonlinear chemistries benefit much more from the new EOL model.

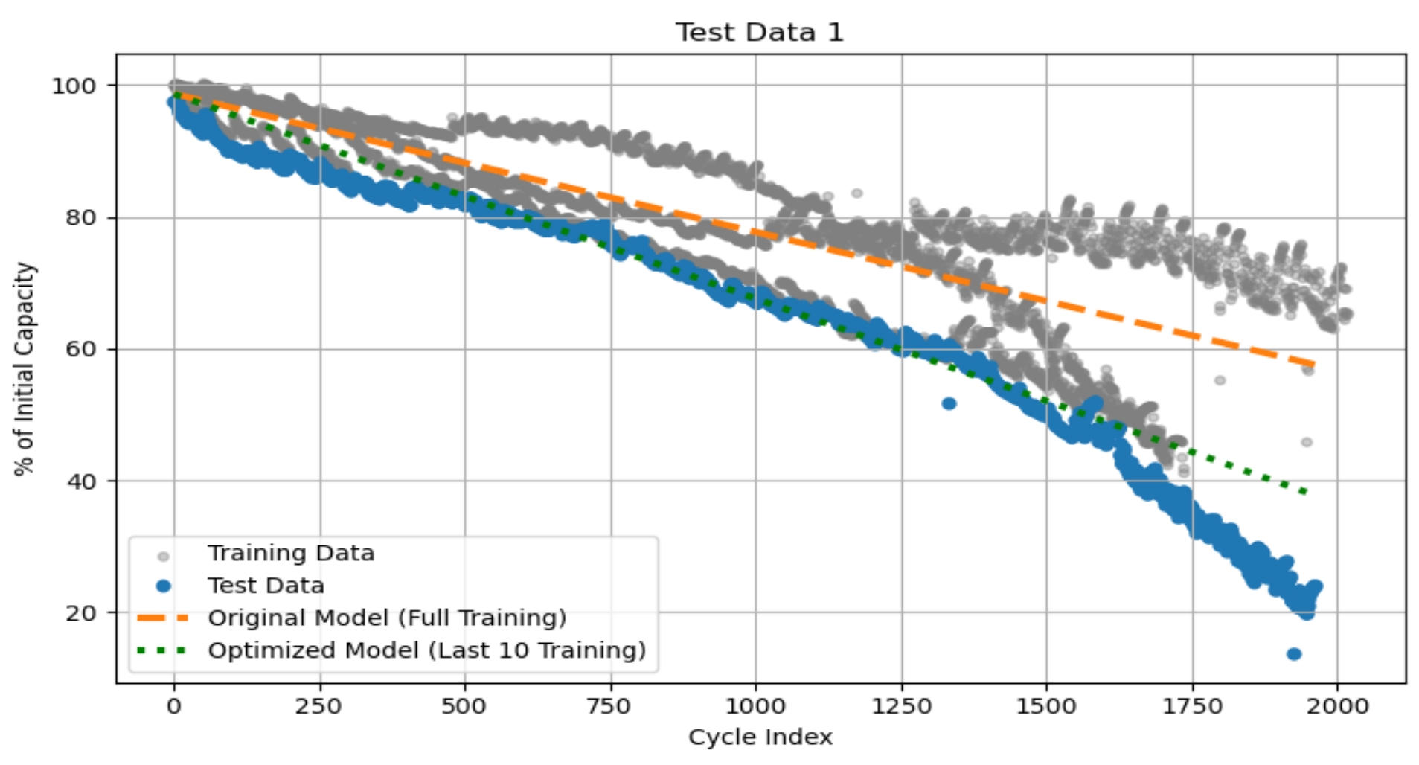
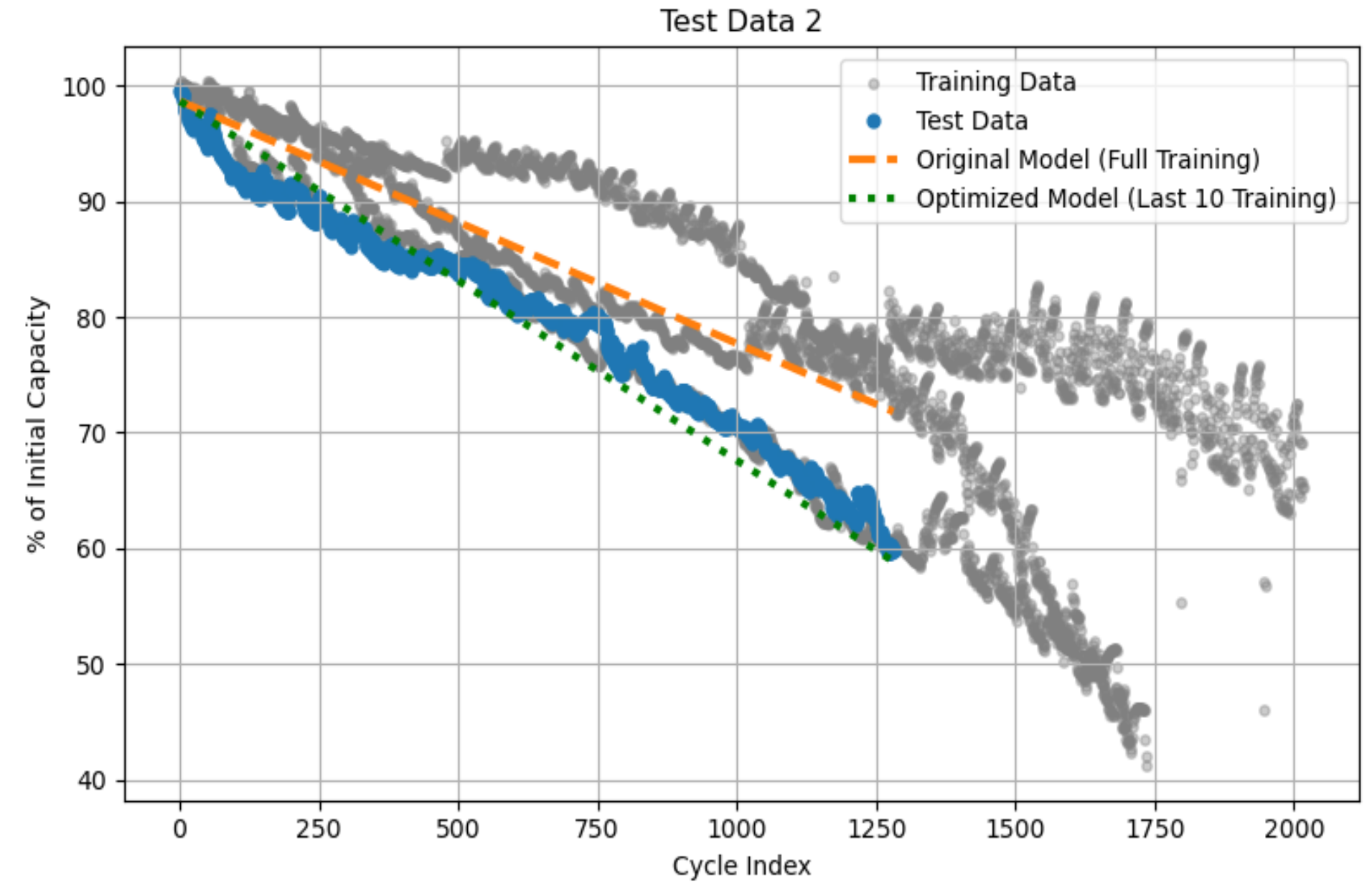


Figure 1: Shows both models on a set of holdout data. Cells are prism LCO alloys tested at 25 °C from 100%-0% SOC at 0.5-0.5 C rate.

Results (cont'd)

Figure 2: Shows both models on a set of holdout data. Cells are prism LCO alloys tested at 25 °C from 100%-0% SOC at 0.5-0.5 C rate.



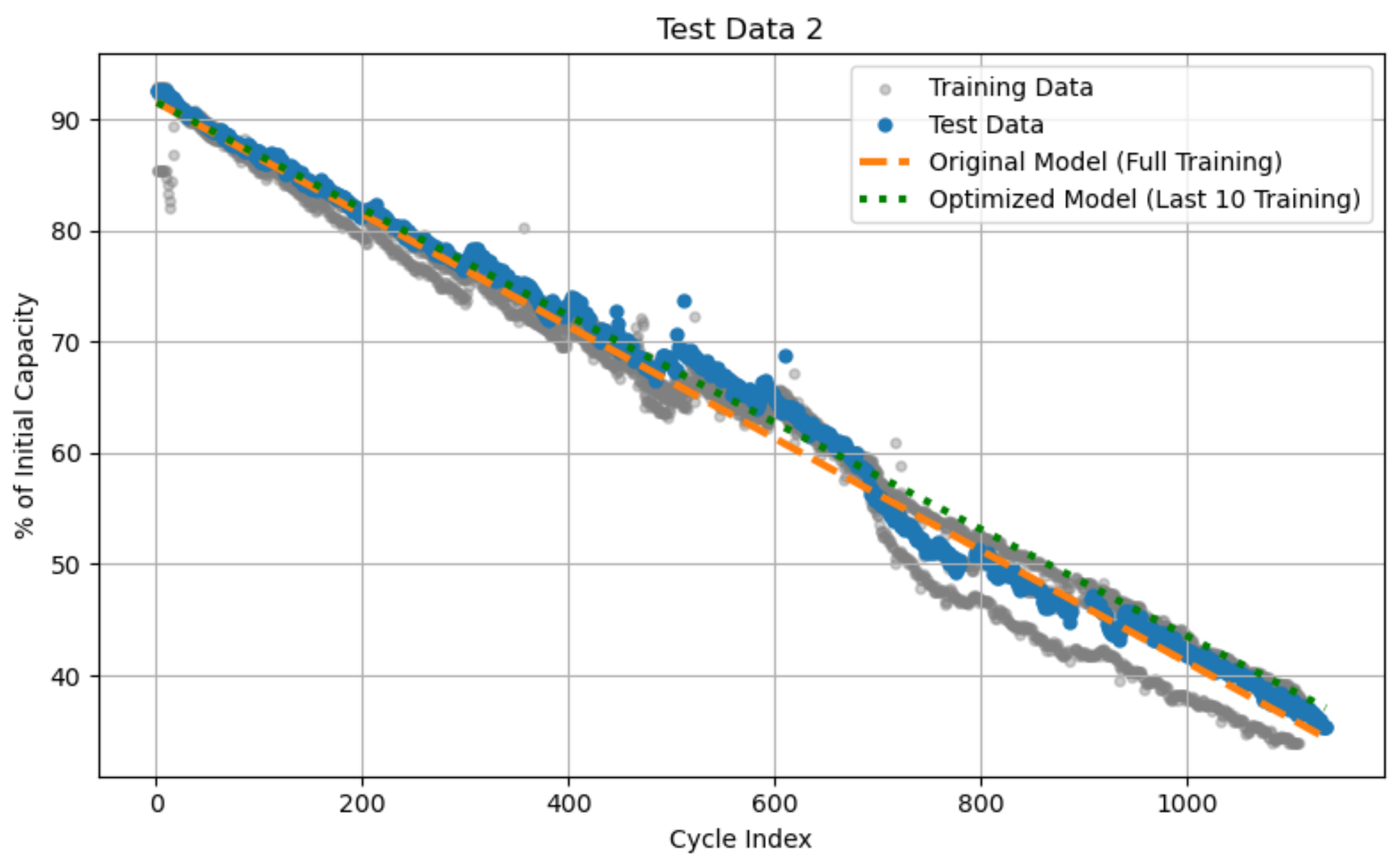
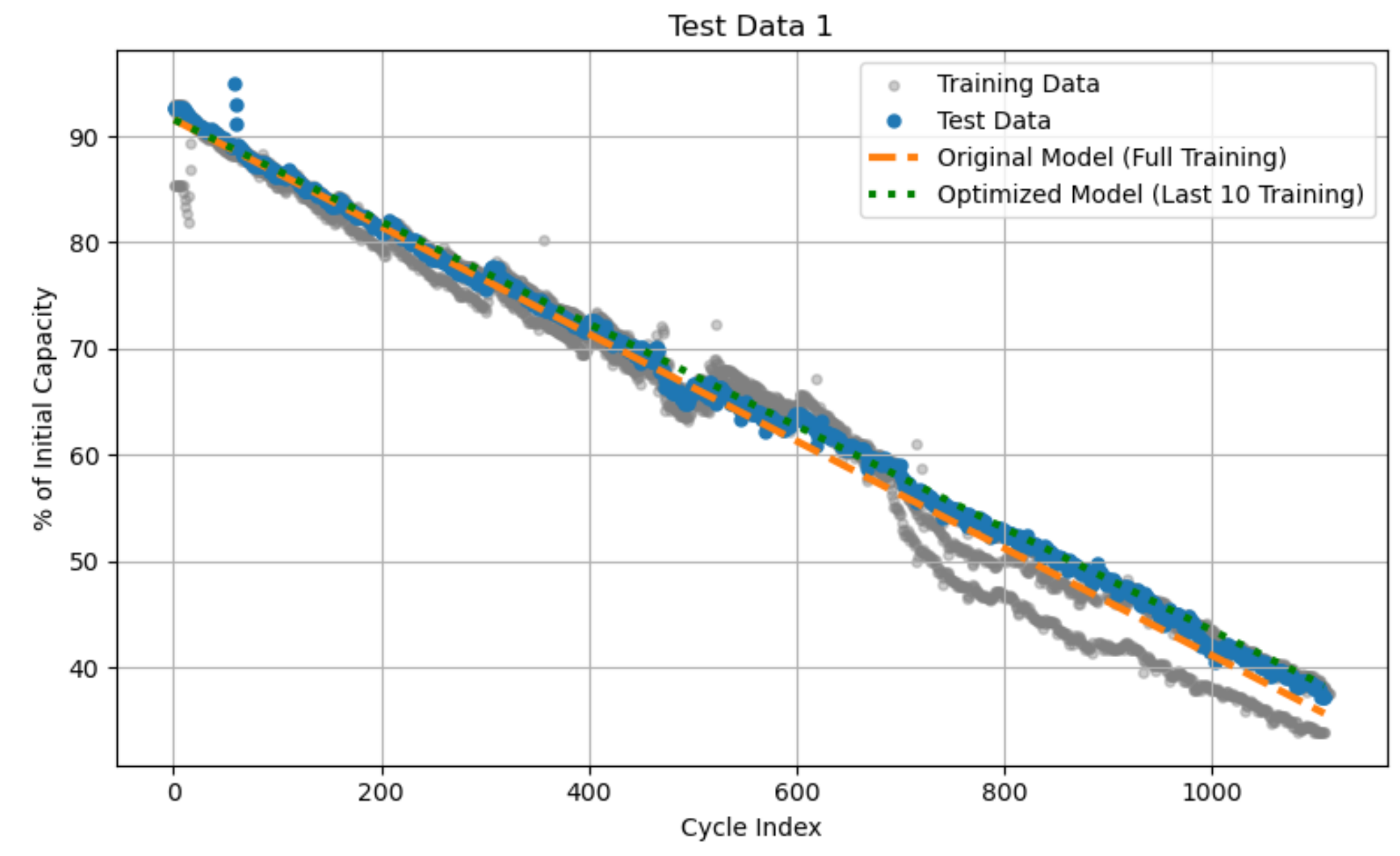


Figure 3: Shows degradation algorithm alongside its three sets of training data and a set of holdout test data. Cells are 18650 NMC-LCO alloys tested at 25 °C from 100%-0% SOC at 0.5-1.5 C rate. Start at 93% is due to noise at the beginning of the testing, the first valid measurements are at 93%.

Figure 4: Shows degradation algorithm alongside its three sets of training data and another different set of holdout test data. Cells are 18650 NMC-LCO alloys tested at 25 °C from 100%-0% SOC at 0.5-1.5 C rate. Start at 93% is due to noise at the beginning of the testing, the first valid measurements are at 93%.

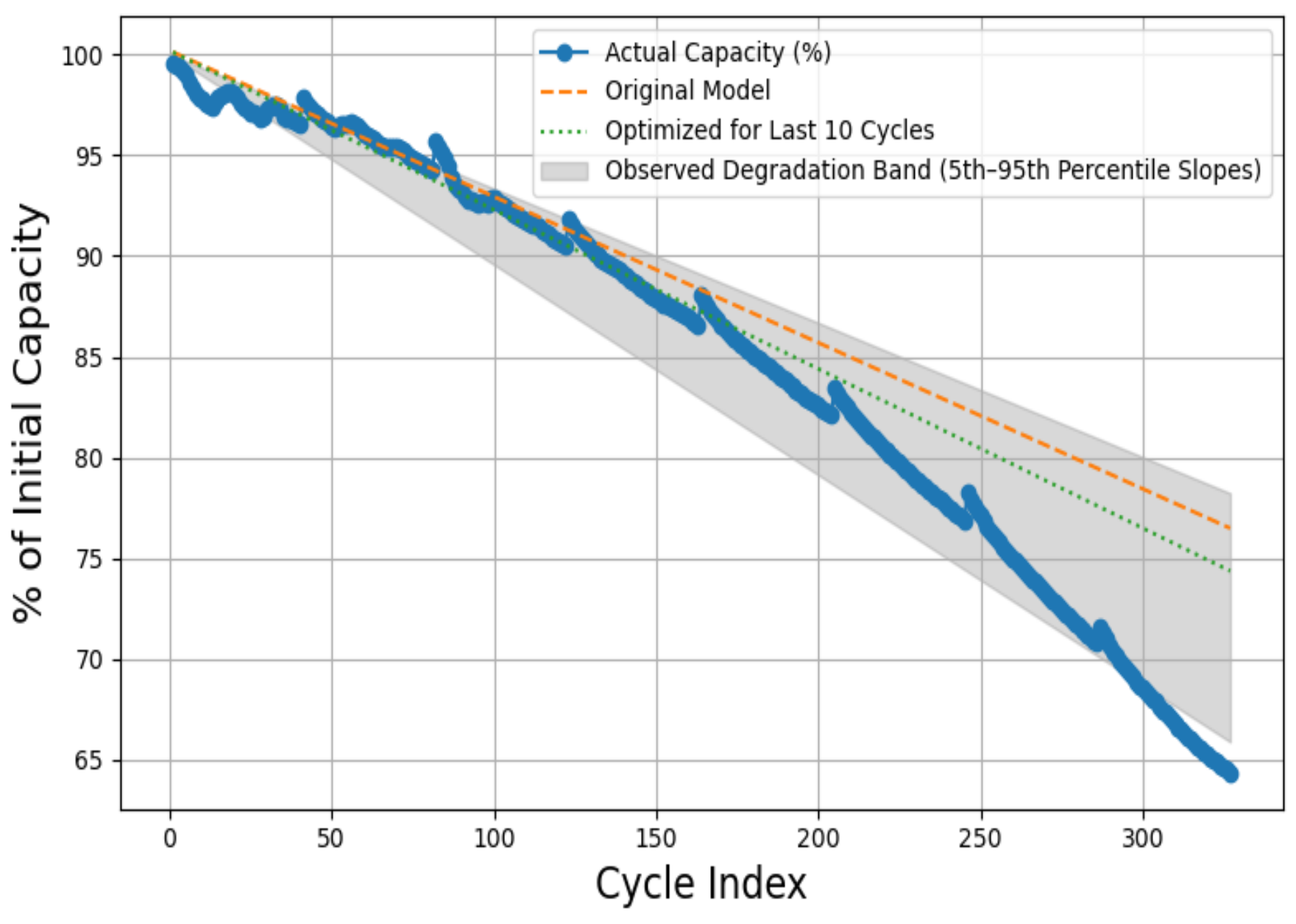


Discussion

The EOL model shows improved accuracy over standard regression by capturing late-stage degradation more effectively. Its simplicity and generalizability make it ideal for integration into QuESt 2.0, supporting more reliable long term technoeconomic battery analysis.

Future work

The next steps will be the inclusion of statistically derived confidence intervals to allow the user to best and worst case degradation for their system. Also, more rigorous testing on datasets that include temperature and discharge rate variations to see how these factors will affect the algorithms accuracy.



Contact and Acknowledgment:
For questions:
Dr. Alexander Headley, jheadley@memphis.edu. David Braese, dtbraese@memphis.edu

This material is based upon work supported by the U.S. Department of Energy, Office of Electricity (OE), Energy Storage (ES) Division.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525

References:

[1] Battery Archive, "Battery data repository," [Online]. Available: https://batteryarchive.org/. [Accessed: July 24, 2025].