

Analysis of Vehicle Battery Data for State of Health and Remaining Useful Life Prediction

Alejandro Gomez Rivas^a

^a*Fontys University of Applied Sciences, Eindhoven, Netherlands*

Abstract

The goal of the research is to create a model that estimates the state of health and predicts the remaining useful life (RUL) of a battery module used in a car. RUL prediction is important because if EV batteries have reached their End of Life, they are still usable for other applications. This report focuses on analyzing real-world datasets collected from different vehicles. It is important to determine which available battery data indicate battery degradation, since not all collected data correspond to battery aging. The data that indicates battery degradation can be used to train a model. The report starts with explaining the basics of a battery, then battery degradation is visualized during the real-world datasets analysis and recommendations are provided for further research. The main report contains the conclusions and further detailed information is provided in the appendix.

Keywords: Lithium-ion battery, State of Health (SoH), Remaining Useful Life (RUL), State of Charge (SoC), Battery degradation, Data driven approach

Contents

1	Introduction	3	9	Data Driven Models	20
1.1	Research Objective	3	9.1	ML Ensemble Models	20
1.2	Research Questions	3	9.2	Deep Learning Models	20
1.3	Scope and Structure of the Report	3	10	Recommendation	20
2	Understanding Batteries	4	10.1	Analyze More Data Using the Same HIs and Additional HIs	20
2.1	Battery Working Principles	4	10.2	Selecting Data Driven Model	20
2.2	Battery Definitions	5	11	Conclusion	20
2.3	Battery Degradation	6	Appendix A	Plan of Approach for SoH and RUL Prediction	24
2.4	Battery State Indicators	7	Appendix B	Collected Battery Data	31
2.5	Battery SoH Estimation	8	Appendix C	Carry Data Analysis	36
2.6	Methods to Estimate the Capacity	8	Appendix D	Daf Data Analysis	47
2.7	Health Indicators	8	Appendix E	Hysteryale Data Analysis	61
3	Remaining Useful Life Predictions Approach	9	Appendix F	Health Indicators	66
3.1	Battery Dataset Collection	9	Appendix G	Overview Data Driven Models	80
3.2	Data Preprocessing for Analysis	9	Appendix H	Model selection: Traditional Regres- sion Models	82
3.3	Data Analysis	9	Appendix I	Model Selection: ML Regression Models	83
3.4	Health Indicator Selection	9	Appendix J	Model Selection: Probabilistic Re- gression	84
3.5	Data Preprocessing for Modeling	9	Appendix K	Model Selection: Ensemble Data- Driven Models	84
3.6	Model Selection	9	Appendix L	Model Selection: Deep learning mod- els	84
3.7	Model Developing	9	Appendix M	Statistics	89
3.8	Prediction	10	Appendix N	Data Filtering and Interpolation Techniques	91
3.9	Validation	10	Appendix O	Python Code for Battery Analysis	92
4	Health Indicators Used in the Analysis	10			
4.1	Voltage SoC	10			
4.2	Voltage Cell Imbalance	10			
4.3	Voltage Parameters	10			
4.4	Voltage at Same SoC	10			
4.5	Temperature over Time	10			
4.6	Temperature and Weather Comparison	10			
4.7	SoC Temperature	10			
4.8	Ratio Distance SoC	10			
5	Carry Data Analysis	10			
5.1	Battery Information and Data	10			
5.2	Preprocessing	11			
5.3	Descriptive Analysis	11			
5.4	Health Indicator Analysis	11			
5.5	Conclusion Based on the Health Indicators	14			
6	Daf Data Analysis	14			
6.1	Battery Information and Data	14			
6.2	Preprocessing	14			
6.3	Descriptive Analysis	14			
6.4	Health Indicator Analysis	14			
6.5	Conclusion Based on the Health Indicators	18			
7	Hysteryale Data Analysis	18			
7.1	Battery Information and Data	18			
7.2	Preprocessing	18			
7.3	Health Indicator Analysis	19			
7.4	Conclusion Based on the Health Indicators	19			
8	Results Vehicle Analyses	20			

1. Introduction

We live in a demanding world where sustainability is one of the most important topics in the automotive industry (1). A vehicle battery consists of a lot of cells in a module, and every cell has a different battery life. Often, only a small number of cells in a module degrade faster than the others. If one cell is bad the total performance of the module is affected (2). If a module is then changed, it can affect the battery pack. If the pack is used and a new module is put into the pack that is different in State of Health (SoH), then the pack still won't work efficiently and the pack will age faster(3).

So it is important to know the SoH to predict the RUL. RUL prediction is important because if electric vehicle batteries have reached their End of Life (SoH < 80%), however the batteries are still usable for other applications (store renewable energy as an energy buffer) as shown in Figure 1 (4).

Therefore, this project focuses on analyzing and visualizing battery data from a mini truck and full electric vehicles. By determining the degradation of the battery module, a model can be created to predict the SoH of the battery, and the possibilities for a second life can be determined. So, with increasing demand for sustainable solutions, understanding and determining the battery health could be used to give the healthy cells a second life.

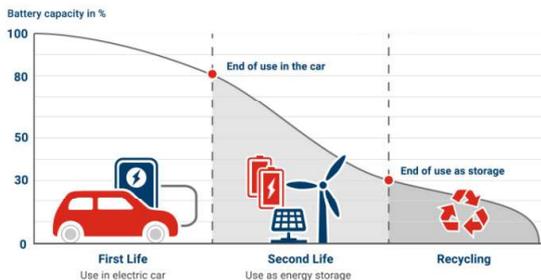


Figure 1: Battery life (4).

1.1. Research Objective

A method is needed to determine the SoH of the battery and predict the RUL. There is a lot of data of vehicle batteries available. So research is needed to find out what the available data contain, what does this data look like and how it can be related to battery degradation. After analyzing, the next step is to figure out how to select or create a data-driven method to process this data and turn it into a battery RUL prediction model.

1.2. Research Questions

The central research question guiding this project is:

- How can the RUL of battery cells be predicted using real-world vehicle data?

Additional sub-questions include:

- What does the available battery data contain, and how can it be analyzed?
- How can battery capacity be determined to estimate the SoH?

- How can the RUL of individual cells be accurately determined?
- How can the available data be related to battery aging patterns?
- How can a data-driven method be created to process the data and generate a battery RUL prediction?
- What are the differences between laboratory data and real-world data, and how do they affect prediction accuracy?

1.3. Scope and Structure of the Report

1.3.1. Scope

This report focuses on the analysis of real-world battery data from different vehicle datasets. The report contains the following topics: the basics of a battery are explained, the necessary steps to achieve the remaining useful life prediction, data analysis of different vehicles, the data is linked to the health indicators of battery degradation, a conclusion of the different data analyses, the further steps to achieve the remaining useful life prediction and a final conclusion. The appendix contains a detailed explanation of the topics discussed in the report.

1.3.2. Structure of the Report

The report is organized as follows:

1. **Chapter 1:** this chapter is the introduction, it provides background information, problem statement, research objectives, research questions, and scope.
2. **Chapter 2:** this chapter is about understanding batteries, it explains the battery topics to understand the further research.
3. **Chapter 3:** this chapter explains how to achieve remaining useful life predictions, it describes the steps and methods to achieve the remaining useful life prediction of a battery.
4. **Chapter 4:** this chapter explains the health indicators, it explains the health indicators used in different data analyses.
5. **Chapter 5:** this chapter analyzes the Carry dataset. The data is preprocessed, analyzed and linked to the health indicators.
6. **Chapter 6:** this chapter analyzes the Daf dataset. The data is preprocessed, analyzed and linked to the health indicators.
7. **Chapter 7:** this chapter analyzes the Hysteryale dataset. The data is preprocessed, analyzed and linked to the health indicators.
8. **Chapter 8:** this chapter draws a conclusion of the different data analyses. The conclusion is based on the health indicators.
9. **Chapter 9:** this chapter describes the potential data driven models for building a predictive model.
10. **Chapter 10:** this chapter describes the further steps. It contains an explanation of how to continue the research and provides recommendations for future research.
11. **Chapter 11:** this chapter is the conclusion. It contains a summary of findings during the research.