Exploring Battery Operational Trends in Hybrid-Electric Buses with Interpretable Machine Learning

Elena Toups

A Thesis

Submitted in partial fulfillment of the

Requirements for the degree of

Master of Science

University of Washington

2025

Committee:

Daniel Schwartz

Eric Stuve

Program authorized to offer degree:

Department of Chemical Engineering

University of Washington

Abstract

Exploring Battery Operational Trends in Hybrid-Electric Buses with Interpretable Machine Learning

Elena Toups

Chair of the supervisory committee:

Daniel Schwartz

Department of Chemical Engineering

While the advent of Lithium-ion batteries in transportation has provided significant improvements to the performance of vehicle energy storage systems, their long-term performance in application remains hard to predict. The growing accessibility of machine learning tools provides a promising route for developing data-driven approaches to battery prognostics. These approaches can leverage both physicochemical knowledge of degradation mechanisms in individual cells and operational trends for the full energy storage system. One withstanding drawback to developing these models is the current sparsity in open-source battery data spanning long timescales and in real-world applications. Introduced here is an operational Lithium-ion battery ($LiFePO_4$, A123) dataset from a fleet of 174 hybrid-electric buses (BAE Systems HybriDrive), collected during maintenance visits occurring between 2015 and 2019. The dataset contains information relating to the performance of the ESS, like voltage and temperature, but is unique in that it does not contain time-series data typically used to make inferences about the internal state. The dataset was explored using simple visualizations and interpretable machine learning methods to provide insights on battery usage and aging in a real-world transit application. Importantly, this work resulted in the extraction of strong discriminative features for determining declines in battery module performance, that are strongly correlated with time and maintenance replacements.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

To my advisor Dr. Daniel Schwartz for your support in this project, for an endless reservoir of great ideas and for applying the right amount of pressure for me to thrive, but not so much that I couldn't also ruminate.

To the students in the 2021 DIRECT program for the initial data preprocessing software, and my fellow capstone members for helping me develop it further.

To the incredibly diverse but like-minded clean energy community at the University of Washington for being a resilient and optimistic crowd, as well as for sharing my hatred for waste and passion for sustainability.

To all of the excellent professors I have had in the last two years – without y'all investing in the potential of students, the department would not be as wonderful and inclusive as it is.

To my fellow members of the EMI Lab for being a solid research group and for granting me the patience and grace to grow – sometimes clumsily – from an experimental chemist into something of a data scientist and engineer.

To Dr. David Beck for being a bright light and teaching me how to talk to my computer.

To the other class of 2025 (the undergraduates!) for being so outgoing and befriending me in Transport and Thermodynamics. Those courses take a community and y'all were a welcoming one.

To Chris Gould, Philip Romero and Ashley Fenton for being my first UW pals. Grad school would not have been fun or survivable without our late night problem solving, sushi lunch breaks and weekend runs.

To my parents, to whom I am endlessly grateful. You both instilled in me the value of an education and I am lucky to always have your support, even when I am navigating unchartered paths.

And to my siblings, thank you for being my first and best friends.

Chapter 1 INTRODUCTION

1.1 Transit Electrification

Globally, public transit agencies have begun improving urban air quality and environmental outcomes by transitioning vehicle fleets away from diesel engines in favor of battery-powered propulsion.[1][2] These transitions demonstrate a critical investment in the decarbonization of a sector that has long been a top global producer of greenhouse gas (GHG) emissions. In the United States in particular, transportation has contributed the singular largest tonnage of GHGs for several years, followed in close second by the energy and industrial sectors.[3] Transitions to hybrid-electric and battery-electric vehicle models has largely been enabled by the modern technological advancements in on-board energy storage systems.[4] For example, improvements in energy and power capacities afforded by Lithium-ion batteries have contributed to increased vehicle range.[5] This has granted more transit agencies the opportunity to consider fleet electrification, with the expectation of a cleaner environmental footprint and lower long term operational costs.[1] This marks an exciting era for energy storage research due to the large-scale deployment of rechargeable Lithium-ion batteries in a commercial setting, and an influx of new operational battery data to learn from.

Despite these developments, the electrification of public transit also introduces new challenges – in large part, arising from the complex aging mechanisms of batteries over time. This is because, as batteries remain in use, they experience physicochemical changes that result in a decline in performance. These aging mechanisms can compound over time and lead to rapid, nonlinear degradation regimes that are often difficult to predict.[6][7][8] This poses a challenge for transit programs who determine success by degree of ridership which is, in turn, largely dictated by the accessibility and regularity of routes.[9] These principles are easily compromised when vehicles breakdown during service from unexpected declines in battery performance, and even failure. These abrupt changes can lack prior identification and disrupt route scheduling, while often necessitating less fuel-efficient stand-ins, longer maintenance downtimes and costly tow servicing.[9][10] It is also worth noting that in cases where Li-ion batteries are operated until catastrophic failure, there is a risk of thermal runaway, posing significant safety concerns.[11] All of these challenges highlight the need for improved oversight on battery performance and aging in operation. This could enable more accurate designations of failure, including preemptive replacements of soon-to-fail battery modules during well informed maintenance visits.

Furthermore, elucidating the aging mechanisms of Li-ion batteries in high-power applications like transportation provides valuable insights for optimizing both primary service life and second-use applications. Mapping degradation patterns to operational conditions can help identify performance thresholds for transitioning batteries from their first use application to secondary uses, like stationary storage systems. [12][13] The engineering process for battery packs requires substantial upfront energy and resources, so maximizing their utilization can improve sustainability across the full lifecycle, moving this technology closer to true carbon neutrality. [14]

1.2 Data-Driven Battery Prognostics

As mentioned, when batteries age they undergo various degradation mechanisms that are rooted in physical and chemical phenomena occurring within the cells. These mechanisms, like lithium plating, active material loss, and dendritic formations, are driven by a variety of cycling conditions and can lead to losses in power and energy capacity.[15] How these mechanisms occur is best understood in small-scale, laboratory settings where advanced characterization and diagnostic techniques are at hand.[16] Some researchers have used semi-empirical models that use mechanistic information to determining the internal state of batteries in operation, but these models have had limited success.[17][18] This is partly because, in applied settings, batteries are in full cells engineered into battery packs, that degrade over larger timescales and in greater variations in environmental conditions. These factors have made it historically challenging to accurately predict battery lifetimes in application based only upon the degradation mechanisms occurring.[6]

One approach for bridging this gap is with data-driven models that are mechanism-agnostic, and instead use machine learning or statistical analysis for battery prognostics. These approaches are attractive because they are capable of fitting complex and nonlinear degradation trends.[8] A drawback, however, is that these models typically require large quantities of operational battery data, which is not always publicly available.[7]

1.3 Objectives

The first goal of this research is to contribute to the pool of open-source operational battery data, with tools for easily processing and parsing it. The next goal is to provide key insights into Li-ion usage and aging in a real-world application, for battery researchers. The last and over-arching goal is to enable the democratization of data-driven fleet prognostics, for end users like transit agencies.

Chapter 2 KING COUNTY METRO DATASET

2.1 INTRODUCTION

King County Metro (KCM) Transit Department is the authority public transit agency for King County, Washington that services a range of urban and suburban routes, including the metropolitan area of Seattle. KCM's modern transit program consists of electric trolleys, battery electric buses and the current largest fleet of hybrid electric buses in the country. Per their transition plan release documents, this bus fleet is expected to grow in the coming years as they transition to a fully battery-powered fleet by the year 2035.[19] This commitment, as well as the wealth of battery data provided by a large electric transit operation, makes KCM an ideal study case for the potential of battery-powered propulsion in a public transportation application. The specific dataset explored in this study was collected from an earlier subset of KCM's hybrid-electric bus fleet which is equipped with BAE Systems HybriDrive propulsion. These systems are series hybrid, meaning an energy storage system (ESS) entirely decouples a downsized diesel engine from the electric drive motor, allowing the engine to run at an optimal speed while affording the bus leaner fuel efficiencies and lower overall emissions. Another mechanism in place is regenerative braking, which recaptures some of the kinetic energy from the bus as it slows down and uses it to recharge the ESS.[20]

A cross-examination was completed to determine bus manufacturers and likely purchase dates, as detailed more in Chapter 2.3. This analysis places the majority of buses in this subset as being purchased sometime between 2012 and 2015. According to the dates of purchase for KCM, it can be reasonably thought that the bus models in this fleet are mostly New Flyer Industries Xcelsior XDE series, with 35', 40' and 60' lengths. It is also possible there are models of the 40' Orion VII manufactured by Daimler Buses North America also captured in the data as well. Provided the BAE Systems ESS would be standard across these models, with some generational updates, it is possible for there to be a mixture of these particular bus manufacturers present in this subset.

Nonetheless, the onboard ESS across vehicle models is consistent and comprised of a battery pack arranged hierarchically with eight 2.5Ah 26650 $LiFePO_4$ batteries (A123 Systems) in parallel constituting a submodule, 12 submodules in series comprising a module and 16 modules in series making a full pack. This arrangement of components is visualized in Figure 1. Maintenance visits occurring between 2015 and 2019 are the sole source of files in this dataset. Each file provides a singular snapshot in time of the status of the onboard battery pack. Some buses appear in more than one file, produced from the same or subsequent visits. Maintenance snapshots consist mostly of histogram variables that accumulate in time as the bus, and thus the battery pack, remain in operation. This time is divvyed up as seconds the pack or subcomponents spent within discrete ranges, or histogram bins, for each variable. Both the number of histogram bins and the bin ranges differ by variable and also by the level in the battery pack hierarchy; the packs have larger bin counts than modules. As an example, the voltage variable at the module-level has a bin each for collecting time spent lower than 2V and higher than 4V, as well as bins between 2V and 4V at 0.2V intervals. For operation-related metrics, this typically provides a normal-like distribution of time spent around the equilibrium states expected for the battery pack and each subcomponent.

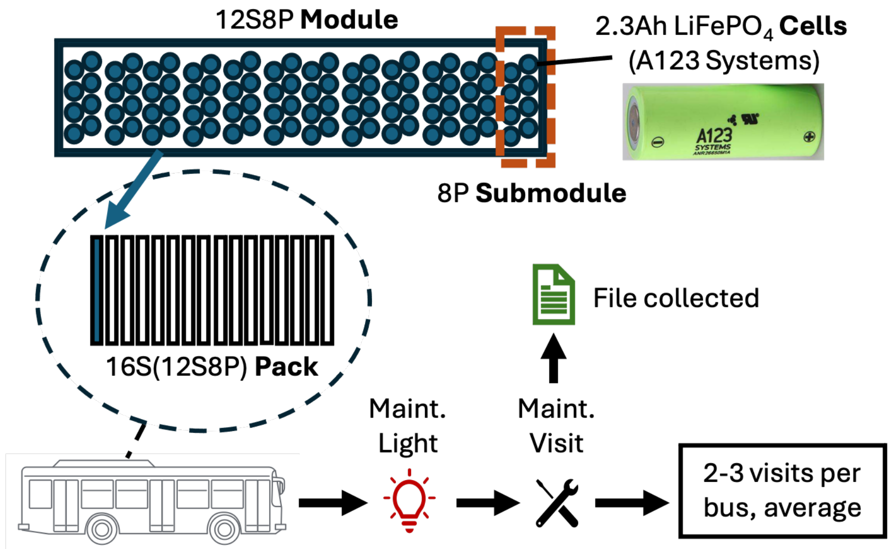


Figure 1. Data collection method and battery pack hierarchy.

2.2 Data Pre-Processing

A common difficulty of working with data from real-world operations is that it can lack the organization and structure necessary to easily parse and learn from it. This disorder can arise for a variety of reasons, like from flaws in the controls system, software updates or disruptions in the collection process. Whatever the cause may be, this disorder usually necessitates a pre-processing step for cleaning and organization before any desired results can be obtained. For this dataset, the first version of a software for pre-processing was designed in 2021, by a group in the Data Intensive Research Enabling Clean Technologies (DIRECT) program at the University of Washington. This rendition is still publicly available on GitHub.[21] The software was then re-adapted during a Data Science Capstone (CHEME 547) in 2023.[22] Since then, additional developments have been implemented as part of the research in this thesis. As the software stands, there are two main Python scripts that enable the parsing and exploration of the data: "etl.py" and "vis.py" – both named for their functionalities. The former is responsible for extracting, transforming and loading the raw data into directories that are ready for parsing. The later script assists in the exploration of the organized files by the creation of data structures and simple visualizations. These scripts together have laid the basis for further exploration using machine learning approaches, which were accomplished in tandem with an array of other tools, primarily in Jupyter notebooks. The full pipeline, from the reception of raw data files to the implementation of machine learning, is demonstrated below in Figure 2. The specific mechanisms of each script are further detailed in the following sections.

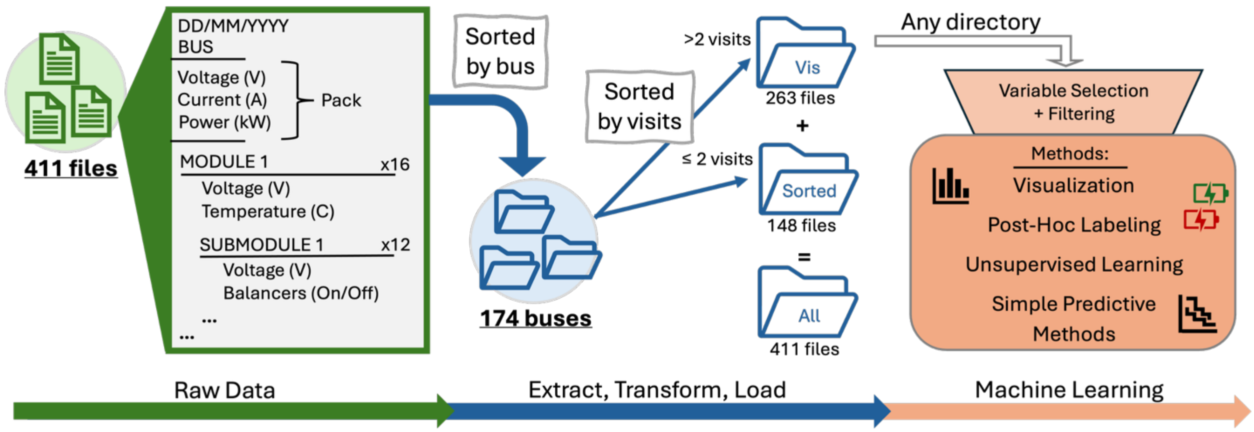


Figure 2. Full pipeline for data processing and learning.

2.2.1 *Extract, Transform, Load (**etl.py**)*

This Python script is designed to run only once, and is responsible for processing and organizing King County Metro data files, primarily CSVs, into a structured directory system. It accomplishes this with a series of functions capable of locating, reformatting and moving files. The reformatting process entails methods like removing extraneous rows and splitting unseparated data, so that formatting is consistent. It then examines features inside the files, like the bus and module serial numbers, the file size, and the row counts, to categorize them as containing only module voltage variables or the full set of expected variables. The files containing only voltage variables are conveniently sorted into a separate directory. Additionally, any files that were corrupted during their collection process and have unsalvageable data are sorted into another directory for incomplete files.

In total, there are around 411 useable files across the dataset. Once the extract, transform, load (etl.py) script is ran, these files are consistently formatted and contained in a hierarchy of folder directories as shown in Figure 2 above. The previously mentioned “incomplete” and “module voltages only” directories are extra subdirectories within the “Sorted” data directory. In general, it is convenient to have the files organized by individual buses so that a history can be painted for each bus across consecutive maintenance visits. This is also necessary for the post-hoc labeling of battery module replacements described later on. The script utilizes the serial numbers in the files to sort them by bus. It is also convenient to have a directory – “Vis” – that contains only bus folders with multiple visits so that, if desired, trends in the dataset can be explored for only buses with an understood history. The “Sorted” directory, thus, contains bus folders containing only 1 or 2 files, from only 1 or 2 maintenance visits. All useable files, sorted by bus and regardless of the number of visits, is included in the “All” directory.

*2.2.2 Visualization (**vis.py**)*

This Python script is designed to run with King County Metro (KCM) battery operational data that has been organized into directories using the previous script (etl.py). Its primarily purpose is to create 2-dimensional data structures with Pandas, called DataFrames, which are convenient starting points for visualizing the data and performing statistical analyses. This script also includes some functions for building visualizations from basic statistics, especially pertaining to the average distributions of operational variables in the dataset, some of which are explored in Chapter 2.3. Lastly, it includes a mechanism for labeling modules in a DataFrame as replaced during KCM maintenance visits, which is detailed in Chapter 2.5.

2.2.3 Filtering Before Learning

After warehousing the data, further processing was necessary to isolate usage trends and provide greater precision in the subsequent learning approaches. While all data collected from the sensors could generally be recognized as representing real conditions in the history of the battery packs and control systems, not all fragments provide equal insight into the conditions when the bus is in steady and routine operation. With careful logic, some of this unnecessary data can be filtered out to reduce unwanted noise in learning models. The exact boundaries for filtering are displayed in Table 2 and can be somewhat arbitrary, but nonetheless considerably improve model interpretability and performance.

Table 1. Baseline methodology for filtering operational battery data, before learning.

DATA DESCRIPTION	FILTERING LOGIC	OPERATION
Contains only zeroes.	No meaningful patterns captured	Removed.
Contains less than 100 hours of operational data.	Typically precedes formation of routine operation.	Removed.
"Duplicate files"; less than 1 day apart.	Bus likely remained under maintenance, no new operational data collected.	Any files with indication of module replacements are kept. Otherwise, only the last file is kept. All others removed.

2.3 AVERAGE OPERATIONAL TRENDS

This section explores operational trends apparent in the data, primarily by exploring what the distributions of collected operational variables look like, on average. Often what is displayed along the y-axes in these graphs is the "Average Normalized Time," which is obtained using the same process, as follows. First, for each individual (battery packs, modules, or submodules), the time collected in the variable histogram bins is normalized by the total time across the histogram bins for that individual. This results in a general profile of time spent across the variable bins for each individual. Then, these profiles are averaged together to obtain the average normalized distributions across all individuals in the dataset. These average profiles provide a good insight into how the energy storage system in the BAE Systems HybriDrive buses perform among KCM's fleet, *in general*.

2.3.1 Bus and Battery Pack

At the highest level of the ESS architecture for the KCM hybrid-electric buses is the full battery pack, which is 1-to-1 with each bus in the population. The expected performance output of each battery pack can be surmised from the architecture of the pack and the electrical speculations of an individual A123 26650 LFP cell.[23] Accordingly, each battery pack should have approximately 12 kWh of energy storage and is expected to have a nominal voltage around 633 V. As indicated by BAE Systems, the energy storage system supplies power during acceleration and hill climbing, and recharges during regenerative braking. This is because the system is designed to provide the average power for propelling the bus via the generator coupled to the engine, with the peak power being covered by the ESS. Acceleration and regenerative braking are, then, the two considerable modes of operation that charge and discharge the batteries, driving the voltage at least momentarily lower and higher, respectively.

As seen in Figure 3, the average normalized voltage distribution for the battery packs falls around the 630 V bin, which is what is expected. Additionally, there is a trend towards the pack spending time in the lowest adjacent histogram bin, at 620 V. This could be due to the bus propulsion system spending slightly more time with acceleration than regenerative braking, causing the lower voltage bin to dominate the charging/discharging behavior for the battery packs. This behavior could be in line with urban transit operation, which consists of regular and frequent stops along the route necessitating frequent acceleration from slow speeds. Overall, the average distributions at the pack level are interesting in terms of understanding these normal operational trends closer to the full energy storage system, but are somewhat removed from what is happening on the individual cell level. This is because, in total, there are 1,536 LFP cells per full battery pack, so whichever operational signs of performance decline and aging that are present at the individual cell-level are diluted for the full pack. Additionally, since replacements occur at the module level, the battery cells in any one module likely vary in age and in performance from cells in other modules, providing variable contributions to the pack's behavior.

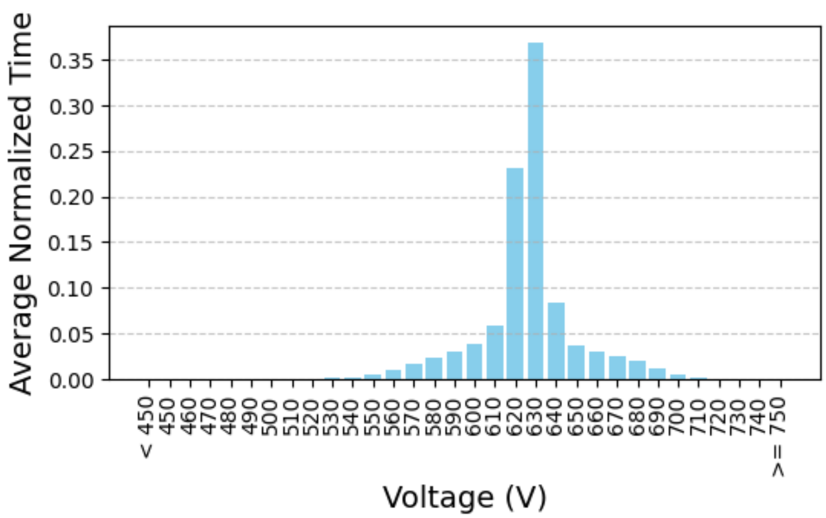


Figure 3. Average normalized histogram distribution of voltages for the full battery pack.

An important thing to take note of in these distributions is that the difference in time between maintenance files collected across variable histogram bins does not actually agree with the time passing between the time stamps located on each maintenance file. This could be due to the data collection method, whereby operational time for each histogram variable may be collected only when the bus and its battery management system are powered on. This would be a plausible assumption and would also explain why there is a consistent trend in the difference between this operational time and calendar time. It also provides a proxy for how much time the batteries in the ESS are spending between "operational aging" and "calendar aging", defined here as the time a battery spends actually in usage (charging/discharging) versus the real-time passed in a battery's lifetime, respectively. These specific metrics are discussed sometimes in literature, because the ratio of time spent between these two modes of aging can affect the degradation mechanisms occurring with usage. This could be due, in part, to some of the reportedly advantageous effects of battery rest occurring between charging and discharging protocols. In this research, it is also useful to understand the ratio of time spent between calendar aging and operational aging for the practical purpose of accurately reading and interpreting the timescales seen in the data.

For these above reasons, the average distribution of ratios for calendar time versus operational time passed between each maintenance file in the dataset was sought. This was calculated by finding an operational time difference and a calendar time difference between every pairwise set of maintenance files for a given bus and then finding the ratio between these time differences. Once all of the ratios were found, they were plotted in a histogram and fitted with a Gaussian distribution, as shown in Figure 4. The average ratio was found to be 0.34 with a standard deviation of 0.054, meaning the battery packs were likely spending around one third of their lifetimes in operation. In context, this equates to the buses in this population spending roughly 8 hours in operation per day on average.

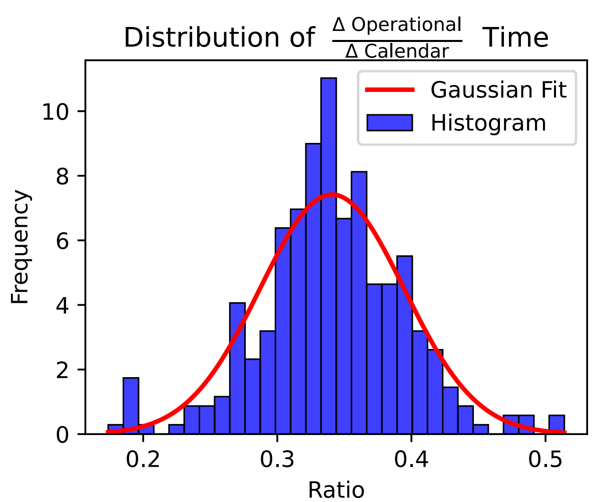


Figure 4. Distribution of ratios for change in operational/calendar time, using all piecewise comparisons of maintenance files.

Another interesting result from finding the average ratio of change between operational and calendar time is the ability to extrapolate back in time to estimate a start date for each bus. To demonstrate this, it is useful to look at a graph of operational time versus calendar time where the maintenance visits are plotted for each bus, as seen in Figure 5. Here, each point is a snapshot in time from when the bus came in for a maintenance visit and the points connected by a line belong to the same bus. Instances where the operational time drops back down to zero could be a symptom of the battery management system being restarted or entirely replaced. The slope of each line segment is effectively the ratio between the change in the operational and calendar time that was sought previously. Thus, starting with the first files in a series, the average ratio of 0.34 found across files can be used to extrapolate to an earlier calendar time that the bus likely began operations. This extrapolation was performed where there is visually a cluster of first files in the dataset to find an estimated start date for this cluster of buses to be around January 2015. Assuming the usage rates remained somewhat constant throughout the buses' lifetimes, this provides an idea of when these buses may have been purchased. These extrapolated start dates can be compared to online databases for bus purchase dates by King County Metro, like the one publicly-kept by the Canadian Public Transit Discussion Board.[24] Table 2 shows a snippet from this database containing the earliest purchased hybrid-electric buses that contained BAE Systems HybriDrive propulsion for King County Metro, also showing the purchase years. Using these insights along with information found online about the software versions from these years allows us to reasonably ascertain that most of the buses in this dataset are manufactured by New Flyer Industries, as part of their Xcelsior XDE series. Some of the buses with a much higher amount of accumulated operational time and that have untraceable software versions may potentially be older Orion VII models. A more thorough look at matching the buses from KCM's hybrid-electric bus fleet to the files in this dataset could provide valuable information on differences in expected loads experienced by the ESS. This is because variations in design parameters like bus length and engine, as well as differences in the route terrain, can affect the work required of the energy storage system during propulsion.

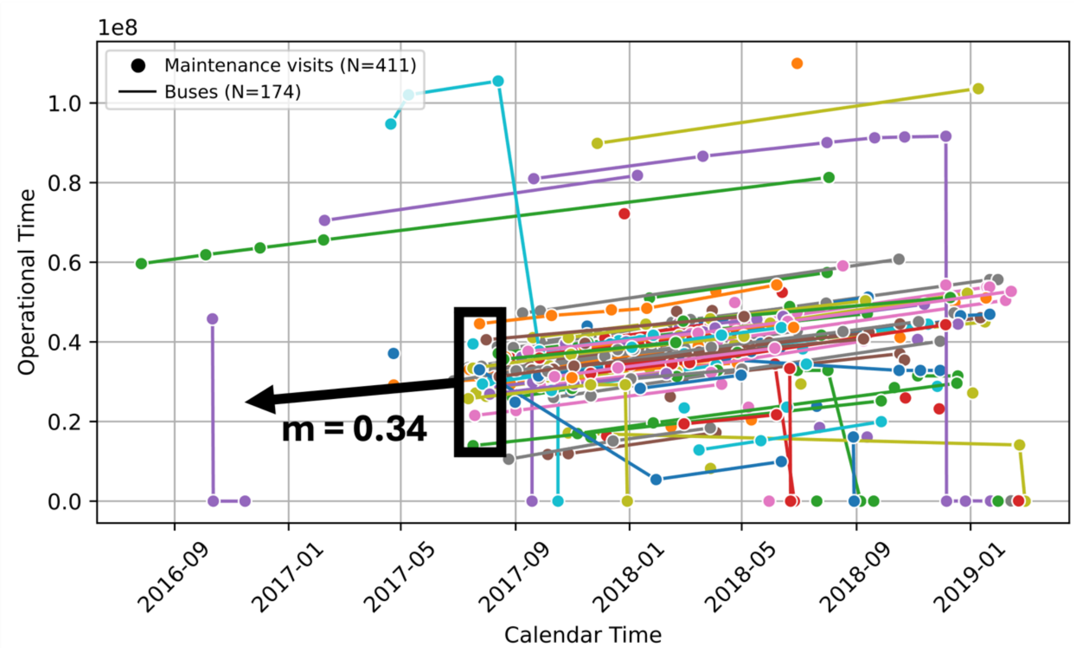


Figure 5. Operational time versus calendar time for each bus between maintenance visits.

Table 2. Compiled data on early hybrid-electric buses purchased by KCM, containing BAE Systems HybriDrive. [24]

FLEET NO.	AMT.	YEAR	MANUF.	MODEL	LENGTH	ENGINE
7001-7093	93	2010	OBI	Orion VII EPA10 HEV (07.501)"	40’	Cummins ISB6.7
7094-7199	106	2011-2012	OBI	Orion VII EPA10 HEV (07.501)"	40’	Cummins ISB6.7
3700-3759	60	2014	NFI	New Flyer Xcelsior XDE35	35’	Cummins ISB6.7
7200-7259	60	2014	NFI	New Flyer Xcelsior XDE40	40’	Cummins ISB6.7
6200-6219	20	2015	NFI	New Flyer Xcelsior XDE60	60’	Cummins ISL9
8000-8084	85	2015-2016	NFI	New Flyer Xcelsior XDE60	60’	Cummins ISL9
8100-8199	100	2016	NFI	New Flyer Xcelsior XDE60	60’	Cummins ISL9

2.3.2 Battery Modules

At the next level in the structure of the ESS is the modules, which are 16 in series and consist of 96 individual $LiFePO_4$ cells each. The nominal voltage of each module is around 39.6 V. An average normalized distribution of voltages, this time across all modules, is displayed in Figure 6. There is a tighter distribution here than there was for the full pack, and the greatest ratio of time is distributed in the 38.4 V bin, which is expected as it is the closest to the nominal. There is also a slight trend towards more time spent in the next adjacent lower bin than the higher one, which is consistent with the trend from the pack level. Again, this could be due to the modules spending more time in a quick discharging mode from frequent acceleration than charging from braking. Additionally, the average normalized temperature distributions for the modules, measured from two distinct sensors, is displayed in Figure 6. Sensor 1 has an average normalized temperature centered around room temperature (25 °C), while sensor 2 runs warmer around 35 °C on average. These distributions fall both within the stated operational range (-30 to 55 °C) for A123 $LiFePO_4$ batteries. The physical cause of the difference in average temperatures between sensor 1 and sensor 2 is unknown. The placement of battery sensors is typically optimized, as to increase the diagnostic capabilities while minimizing the number needed and their associated costs. Therefore, the sensors are likely placed in positions that capture unique measurements from one another. Commonly, when two sensors are used, one is placed on the perimeter of the battery component to capture the ambient or lowest temperature for the component, while the other is placed in a central location to capture the warmest accumulated temperature.[25] This provides insights on the full range of temperatures, and the ability to speculate on gradients across the pack. Figure 7 visualizes the distribution of average temperatures collected from sensor 1 and sensor 2 across each module location in the battery pack. It also graphs the average annual high and low temperatures in Seattle, for comparison. Both average module temperatures run warm, compared the Seattle temperatures. There is a slight visual temperature gradient across the pack, with the lower module numbers running a few degrees cooler on average. There is also a slight trend in alternating averages for the central most module temperatures, captured from sensor 1.

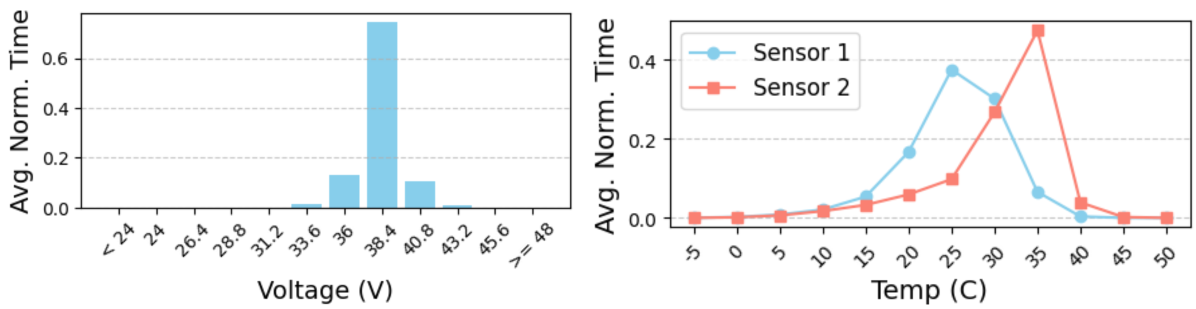


Figure 6. Average normalized distributions of the module operational variables: voltages (left) and temperatures from sensor 1 and sensor 2 (right).

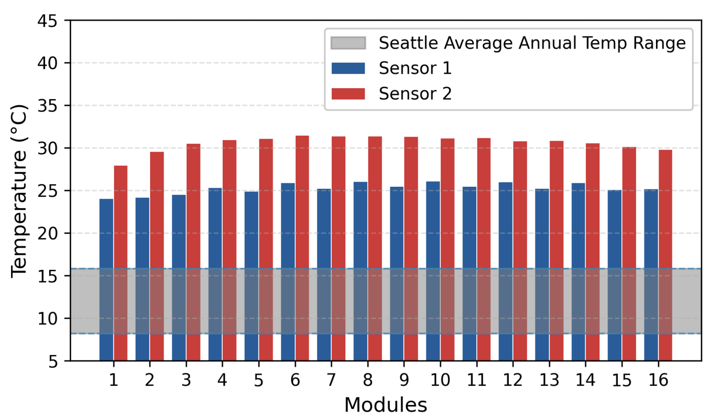


Figure 7. Average temperatures of each module by position, with a comparison to Seattle's average annual high and low temperatures.

2.3.3 Battery Submodules

The lowest level of the ESS where operational variables are captured is at the submodules. There are 12 submodules in series per module, and they are each comprised of 8 $LiFePO_4$ cells in parallel. This configuration produces a nominal voltage that is the same as for the individual cells (3.3 V). As shown in Figure 8, there are the same number of histogram bins for average normalized voltage distributions for each submodule location as there were for the modules, but the bins are now 1/12th of the magnitude. Also displayed is a heatmap visualization of the average normalized distributions for every submodule position across the dataset. There is a visual trend in alternating distributions across submodule positions, where some positions see a slightly lower amount of time spent in the nominal histogram bin. These variations in charging and discharging over timeframes and across submodules could be either the intention, or the consequence, of pack design. Overall, it is not a very significant trend, as can be seen by the difference in distributions between the two submodule positions with greatest disparity (submodule 2 and 7).

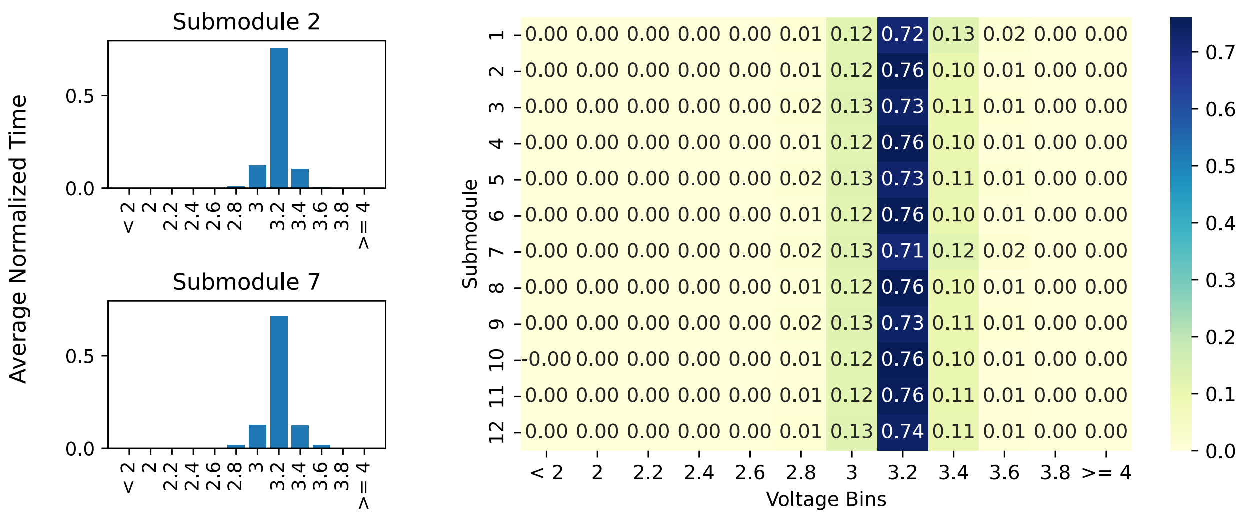


Figure 8. Visualizations of submodule voltages: a heatmap displaying average normalized time in voltage bins across all submodules in the dataset (left) and histograms of average voltage distributions for submodules with the greatest variance in distribution (right).

2.4 Labeling of Replacements

As batteries remain in operation, their capacities fade causing a decline in performance that can propagate throughout the other components in the ESS. For example, in the configuration of the BAE Systems ESS, if a single cell in a submodule begins to decline in performance or even completely fails, the other 7 cells in parallel will need to carry a greater amount of current to cover the same load. This can cause a charging imbalance that accelerates the aging of the healthy batteries and can affect the overall performance of the full pack. In order to prevent this, BAE Systems uses fault codes that are based on performance criteria for indicating failing batteries. These fault codes are responsible for triggering a maintenance light on the driver's dashboard, informing the driver to bring the bus in for maintenance when convenient. At these maintenance visits, failing or underperforming batteries in the ESS are replaced as whole modules. The fault codes that may indicate battery failure and are based upon operational metrics include a low and high voltage fault and a temperature out-of-range fault. The high and low voltage faults are triggered when a submodule has spent at least 10 consecutive seconds above the 3.95 V bin or below the 0.05 V bin. The temperature out-of-range fault is triggered when one or both of the temperature sensors are outside a band of +/- 2 standard deviations from the mean, and are greater than 125 °C or less than -35 °C for 60 seconds. These fault codes all capture operational behavior that is well outside of the normal expected ranges for these cells. These simple and extreme metrics provide some insight into what battery failure may look like in this dataset, but are not particularly sensitive to subtle changes in battery performance that may indicate battery degradation.

Additionally, it is not possible to exactly replicate the fault code criteria for replacements with this specific dataset for two main reasons. The first is that the fault codes use smaller histogram bins than the ones present in this dataset, capturing smaller deviations in voltages. Secondly, the dataset here lacks time series data and it is not possible to precisely determine when consecutive periods of time were spent in a discrete voltage range for binned time data. This necessitates an alternative approach for determining which battery modules were likely replaced by KCM during maintenance visits in this dataset, due to declines in performance or failure. It also creates a desire for more sensitive metrics for characterizing degradation than what is implemented by the fault codes.

Using the dataset at hand, a post-hoc method for labeling replaced battery modules in the dataset was implemented. This method relies on the existence of multiple consecutive maintenance files for each bus in the population. Since each of these files contains data including serial numbers for all 16 modules in the battery pack, it is possible to compare changes in the serial numbers across files to determine the replacement of a module has occurred. For these purposes, it was assumed that data files for a given bus are downloaded and reviewed at maintenance visits first, with any replacements of failure-indicated battery modules occurring afterwards. This is supported by the fact that files containing a new serial for a given module already have accumulated seconds in operation. So, modules in maintenance files that have not had a change in serial number in the next consecutive file are labeled “Healthy” and modules that have had a change in serial number in the next file are labeled “Swapped”. It is important to note that this method is insufficient for the labeling of buses that have not had subsequent maintenance visits, or are the last in a series of data files for a specific bus. Accordingly, the status of modules in these files are labeled “Unknown”, since replacements were not evidenced in the total span of time accounted for in this study. A simple visualization of this post-hoc labeling method for replaced battery modules is shown in Figure 9. These labels are used in the next chapter for easily interpretation of which modules were likely failing and replaced by KCM during a maintenance visit. They are not, however, used to inform the unsupervised learning method implemented, which will be discussed more in depth.

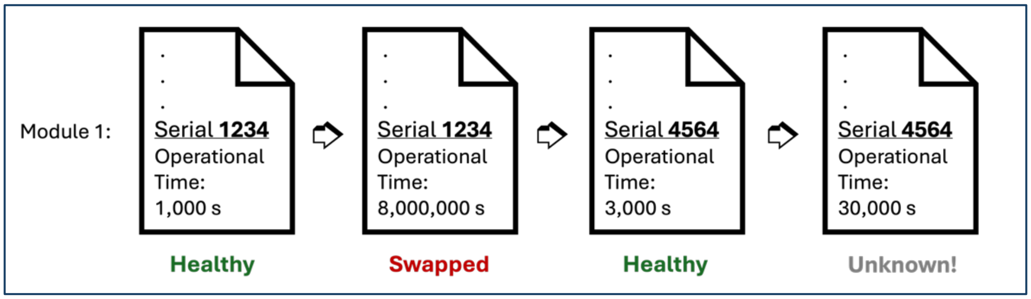
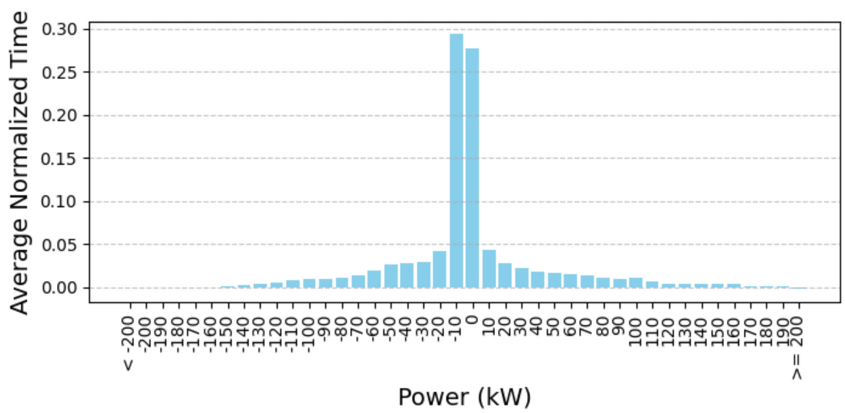


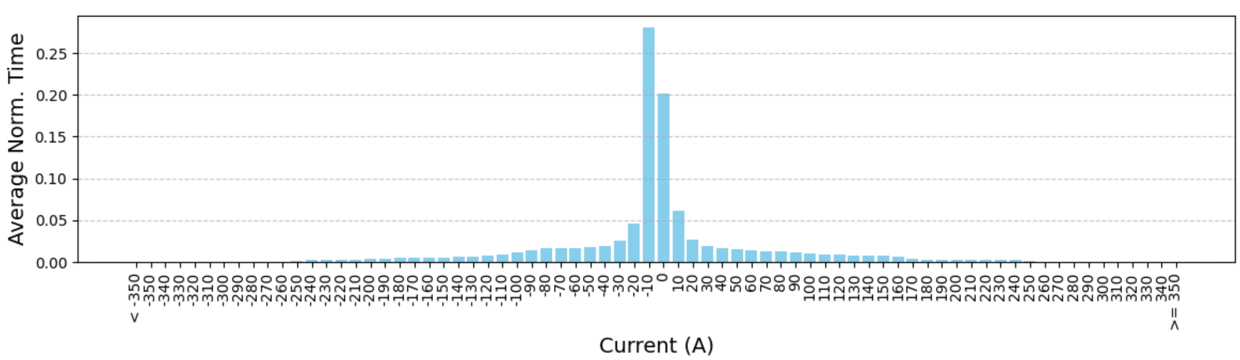
Figure 9. Visualization of post-hoc method for labeling module replacements by KCM, for buses with multiple consecutive maintenance visits.

2.5 Appendix

2.5.1 Bus and Battery Pack Supplemental Figures

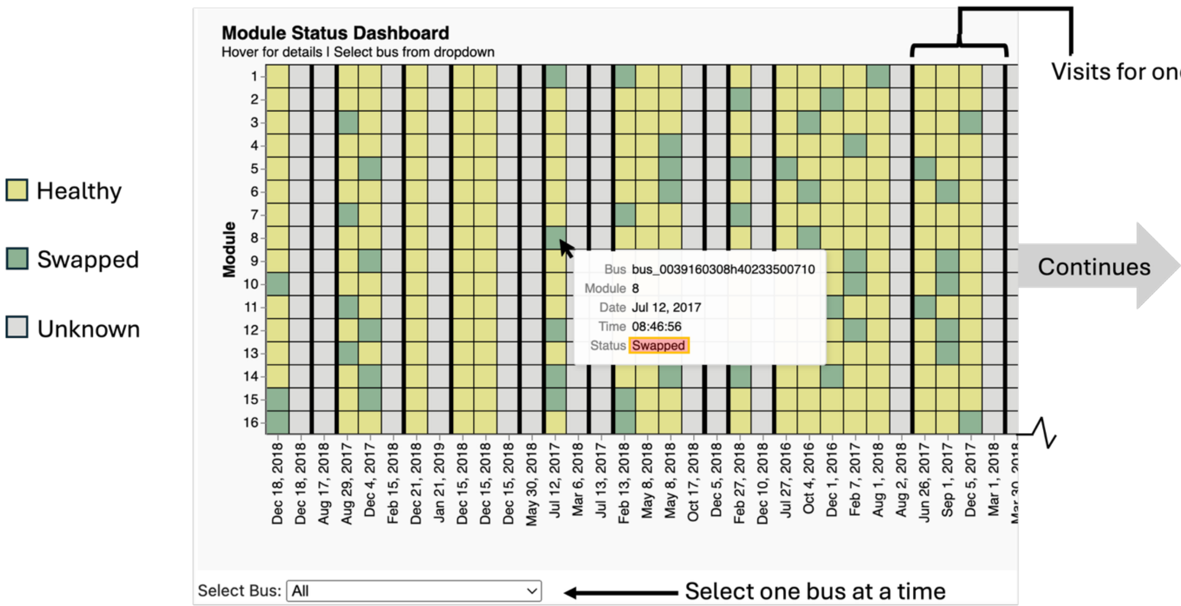
The average normalized power and current distributions for the battery pack are displayed:





2.5.2 Labeling of Replacements Supplemental Figures

An interactive visualization in the vis.py python script that allows the exploration of post-hoc labeling of module replacements by KCM at maintenance visits:



Chapter 3 EXPLORATORY ANALYSIS

3.1 INTRODUCTION

The previous chapter introduced an operational battery dataset from an earlier subset of King County Metro's hybrid-electric bus fleet. The primary method of exploring the dataset in that chapter was through visualizations of the average normalized profiles of the operational variables across the energy storage system hierarchy. These visualizations are generally effective for learning what the operational trends of the ESS components look like on average. However, it is necessary to explore how the components in the energy storage system vary significantly from these average distributions to understand the characteristics of battery aging and degradation in this dataset. To do so, this chapter explores an unsupervised machine learning method called principal component analysis that is applied on the battery usage variables of voltage and temperature. These are the two operational variables used in the fault codes of BAE Systems to most directly determine declines in battery performance and failure.

Principal component analysis (PCA) is a modern unsupervised machine learning technique based on the theoretical foundations of Karl Pearson, who postulated that multivariate datasets should contain a more fundamental and descriptive set of variables than the *a priori* ones. Pearson introduced the concept of best-fitting lines and planes in multidimensional space, laying the groundwork for dimensionality reduction.[26] Later, Harold Hotelling formalized PCA in 1933 by developing an algorithmic approach to extract principal components through eigenvalue decomposition of the covariance matrix, making it a potent statistical tool.[27] Unfortunately, at the time of conception, it was not practical to apply it to large datasets because of the sheer magnitude of calculations necessary. However, with the introduction of higher-level coding languages, greater processing power, and a massive influx of new data, statistical learning techniques like PCA are now widely accessible to anyone with a laptop. Thus, PCA has become a popular technique for transforming large, multivariate datasets into smaller and more interpretable ones. It achieves this by condensing the original variables into smaller, uncorrelated sets of linearly derived variables (principal components) that describe the greatest variance.[28] For this research, we implement PCA primarily as an exploratory method to elucidate trends that could lead to more efficient battery monitoring and fault detection. Another common usage of PCA is for feature extraction, where the extracted principal components are used as features in a prediction model. This later usage is not our primary focus for this thesis but will be explored in the next chapter by evaluating the extracted principal components in simple predictive models.

3.2 METHODOLOGY

To apply principal component analysis, the operational variables in the dataset are broken into subsets with n individuals described by p numerical variables, organized into an $n \times p$ data matrix $\mathbf{X} \in \mathbb{R}^{n,p}$ where each column vector x_j corresponds to the j^{th} variable. In this case, individuals represent either battery modules or battery submodules, and each column vector is actually a discrete bin from the variable histograms, such as the 2.0 V bin, in the case of voltages. The data matrices are always mean-centered first, so that the spread of the variance is measured around the mean. Classically, this approach seeks orthogonal linear combinations of the column vectors with maximum variance, given by $\mathrm{Var}(\mathbf{Xa}) = \mathbf{a}^{\mathrm{T}}\mathbf{Sa}$. In which, $\mathbf{S}$ is the $p \times p$ sample covariance matrix of the dataset and $\mathbf{a}$ is a vector of constants. Maximizing this variance results in the directions of the principal components. Typically, an additional constraint is imposed where $\boldsymbol{a}^T\boldsymbol{a} = 1$ so that this reduces to solving the eigenvalue problem:

$$\mathbf{Sa} = \lambda\mathbf{a} \tag{2}$$

where the eigenvectors $\mathbf{a}_1, \ldots, \mathbf{a}_{\mathrm{p}}$ of $\mathbf{S}$ define the directions (or principal axes) along which the data varies most, and the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_p$ represent the magnitude of variance captured along each axis. Since the eigenvectors are orthonormal, the principal components $\mathbf{Xa}_k$ are uncorrelated.

In practice, especially as designated in computational tools, PCA is performed via Singular Value Decomposition (SVD), rather than the eigen decomposition of the covariance matrix. This alternative approach for the mean-centered data matrix $\mathbf{X}$ is given by:

$$\mathbf{X} = \mathbf{U\Sigma V}^{\mathrm{T}} \tag{3}$$

where, $\mathbf{U}$ is an $n \times n$ orthogonal matrix whose columns are the left singular vectors, representing the principal components. While, $\mathbf{\Sigma}$ is a $n \times p$ rectangular diagonal matrix containing the singular values σ_k sorted in descending order, which scale the principal components. And $\mathbf{V}$ is a $p \times p$ orthogonal matric whose columns are the right singular values, or principal axes, defining the directions of maximum variance. The principal scores, which are the projections of the original data onto the principal axes, are given by: $\mathbf{Z} = \mathbf{XV} = \mathbf{U\Sigma}$. The captured variances of the principal components correspond to the eigenvalues λ_k of the covariance matrix and are derived from the singular values via: $\lambda_k = \sigma_k^2 / (n-1)$ where, again, σ_k are the singular values and n is the number of total individuals. This SVD-based approach for solving the principal components is more numerically stable and computationally efficient, especially for datasets with high dimensionality (where $p>>n$). In this work, PCA was implemented using a optimized SVD solver in the scikit-learn library in Python.

3.3 Features at the Module Level

3.3.1 Module Voltages

The methodology described for principal component analysis (PCA) was first applied to the voltage histograms for battery modules across all individual buses and files. This is a natural starting point for unsupervised exploration because of the extra layer of interpretability afforded by the post-hoc labeling of module replacements. Important to note, as before, this labeling scheme is not factored into the calculation of the principal components (PCs), and is only included as a visual clue for knowing modules were ultimately replaced by KCM in this dataset. Displayed on the left in Figure 10 is a graph representing the cumulative explained variance, as calculated, for each new PC. The first PC always lies along the axis of greatest variance, which in this case captures 75% of the total variance alone in this module voltage data. A general rule-of-thumb for data scientists is to use as many PCs as is necessary to capture at least 95% of the cumulative explained variance. Luckily, in this implementation, the first two PCs together capture this amount of variance, so the third with be disregarded as noise for now. Also in Figure 10, to the right, are the three loading values for each PC which describe the amounts of the original variables incorporated into each new principal component axes. The best interpretation of these new PCs is as new axes, on which the original data vary most. So, the scores of each individual (or each module currently) can be understood in terms of the original variables by inferencing which amount of the original variables are loaded as positive and negative values on the new PC axes. For example, negative values on the PC1 axis capture the behavior of modules that spend more time in the nominal voltage bin, while positive values capture moderate spread into the immediately adjacent voltage bins. While, the negative values in PC2 capture a slight lower spread from the nominal voltage bin and positive values capture the "wings" of more severe spread into the voltage bins twice removed from nominal.

Now that the principal component loadings for module voltages have been broken down, the original data can be explored in terms of the scores from the new PC axes. Figure 11 contains biplot histograms that display the scatter plots of module scores, in terms of PC1 and PC2. Each point on the graph represents a battery module and is color-coded by the labeling found via the post-hoc method for determining replacements at maintenance visits. Additionally, there are histograms that display the distributions in the graphs along each PC axis. These first two PCs were highly successful in separating the modules into distinctive clusters, by voltage profiles.

The bottom cluster follows a centered and moderate operational trend, correlating more strongly with "Healthy" modules. This cluster is centered in the bottom left quadrant with negative values of PC1 and PC2, which is characterized by a large portion of time spent at nominal voltages with some light spread and a slight lower voltage trend. This cluster also spreads into the bottom right quadrant, which is characterized by a light voltage spread and a slight lower voltage trend. Lastly, a small portion of this cluster also spreads into the upper left quadrant, which is characterized by significant time at nominal voltage with some more significant voltage spread. In conclusion, this cluster may span what the regular or routine operational trends of battery modules that are mostly healthy.

In comparison, the upper cluster primarily spans the upper two quadrants which are characterized by significant time spent in the heavier spread voltage bins. This upper cluster is a quantized step away from the healthier cluster, suggesting a significant variation seen in this population of battery modules. Additionally, the upper cluster is strongly correlated with battery replacements by King County Metro, apparent from the labeling scheme. For clarity purposes, the biplots are presented with and without the modules that are unable to be labeled. It is worth noting that there is a general trend of modules not being replaced in a maintenance visits if they were just replaced in that position on the same bus in the maintenance visit directly preceding it. Following this trend, most of the "Unknown" modules could be labeled "Healthy" if there was an immediate preceding replacement, however, that method was not followed here. So, unsurprisingly, a majority of the "Unknown" modules are actually "Healthy", as apparent in the biplots. This is because PCA scores the modules only on the voltage histogram trends and independent of the labeling, so it is able to successful group the "Unknown" modules in their appropriate clusters. Another thing to take note of in this biplot is a subtle further deviation between modules within the upper cluster. This suggests that failure may be bimodal, or that there is two distinctive groups within the battery modules, perhaps arising from different bus models or distinct differences in typical loads.

Finally, since PC2 showed strong discriminative abilities for differentiating between healthier and less healthy modules, the PC2 scores for each module were explored further. In Figure 12, the operational time spent in the most frequented voltage bins was plotted against the total operational time for each battery module. The battery modules were additionally labeled with their PC2 scores. A strong clustering is again apparent in these plots, specifically in the twice removed voltage bins for greater spread away from the nominal voltage. It can be seen that modules with greater total operational time have a larger voltage distribution spread, scoring higher in PC2, which differentiates them from younger modules. This analysis supports the hypothesis that this clustering demonstrates a decline in battery performance seen with aging.

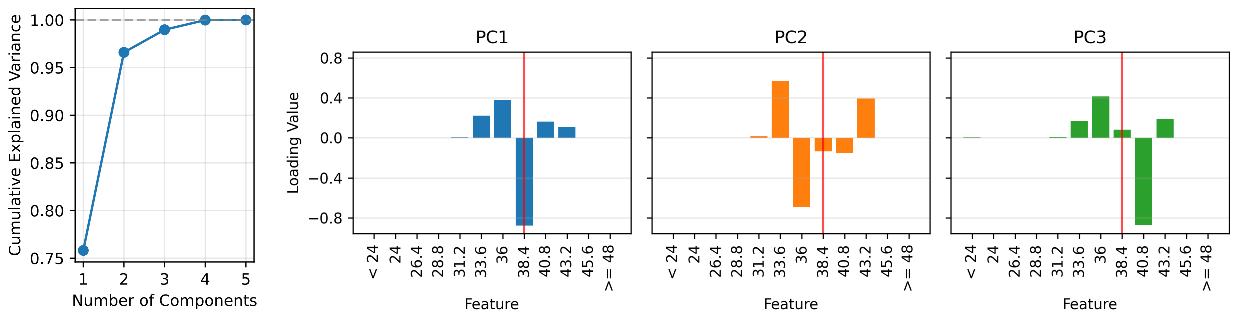


Figure 10. Metrics from principal component analysis on module voltages: the cumulative explained variance of each principal component (left) and the loading values of the original variables for the first three principal components (right). The nominal voltage bin is highlighted with a red vertical line.

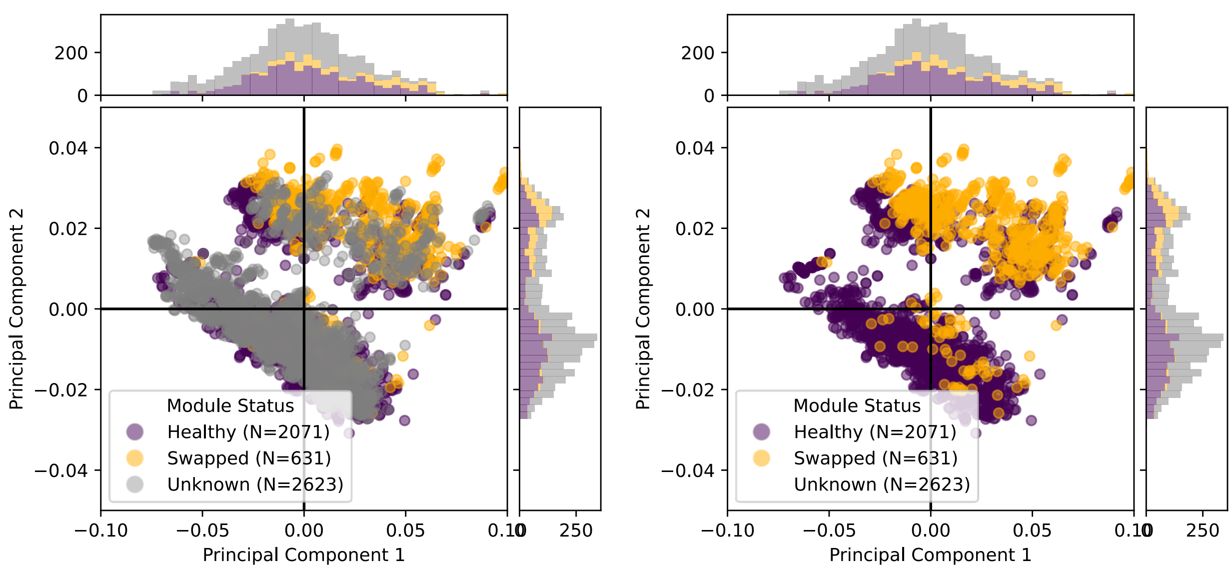


Figure 11. Biplots of battery modules across all maintenance visits scored with PC1 and PC2 calculated only from the original voltage variables: with all modules (left) and without unknown modules (right). Labeling is independent of the principal component analysis and displayed for interpretability purposes.

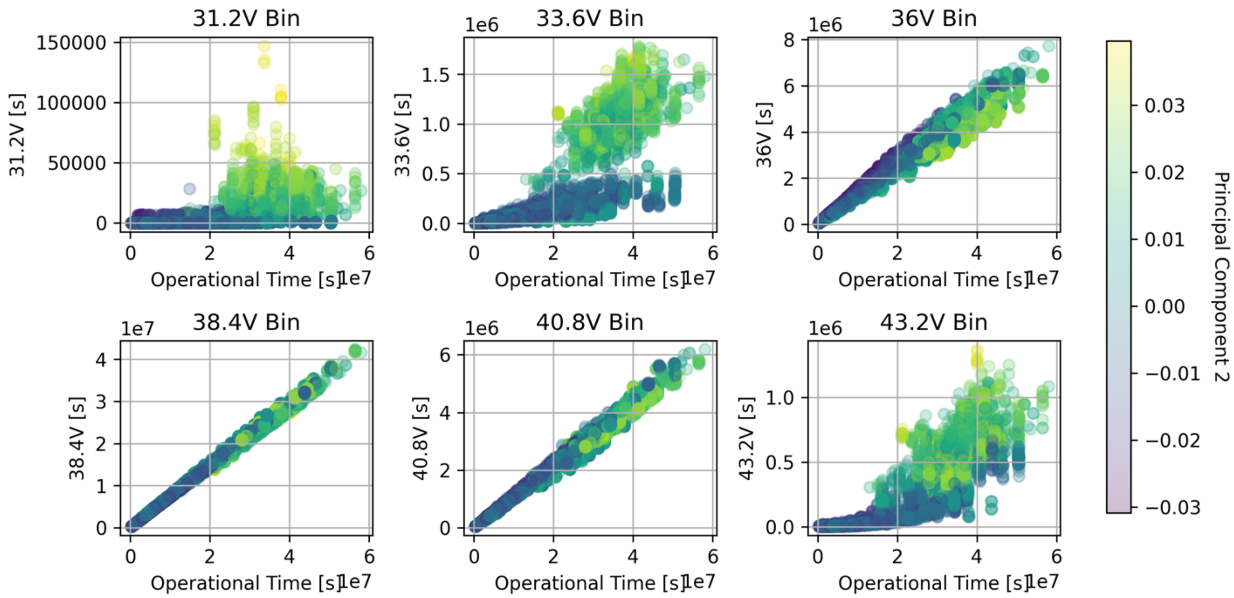


Figure 12. Graphs of operational time spent in most frequented voltage bins versus the total operational time across all voltage bins, for each battery module and from each maintenance visit. Battery modules are labeled with their principal component 2 scores, calculated only from the original voltage variables. Clustering is apparent in coloration and in space for battery modules with lower and higher amounts of operational time.

3.3.2 Module Temperatures, Sensor 2

Next, PCA was applied to the temperature histograms from the battery modules in this dataset. This was implemented separately for each sensor because the sensors are believed to have independent data streams located at different physical points on the battery modules. Also, since sensor 2 is thought to be centrally located on the battery module, it is more likely to be capturing the true thermal losses from normal operation. So, in this section, the PCA results from sensor 2 temperature data are discussed, with the results from sensor 1 temperature data included in the appendix.

As seen in Figure 13, the first two PCs are enough to capture around 98% of the cumulative explained variance so only these PCs will be discussed. The first PC captures temperatures above room temperature, with the positive axis capturing time the modules spend at a warm 30 °C and the negative axis capturing a hot 35 °C. The second PC captures time spent right at room temperature (25 °C) on the positive axis and in both the "warm" and "hot" bins for the negative axis. Figure 14 presents the biplots displaying the module scores on a scatter plot, in terms of PC1 and PC2 with the same labeling provided by the independent post-hoc method for determining replacements.

Again, there is a cluster that most of the modules labeled "Healthy" and "Unknown" fall inside of, which is centered in the bottom left quadrant. This cluster spans, from left to right along the lower hemisphere, between more time spent between hot and warm temperatures. Most of this healthy cluster is distributed with a slight negative PC1 value, entailing most of these modules spend more time at the hotter 35 °C.

In contrast, the upper cluster correlates with replacements and has positive PC2 values which are characterized by more time spent at room temperature (25 °C). One similarity the upper cluster has to the lower cluster is that it also sits mostly on the slightly negative side of PC1. This means modules in this “less healthy” cluster still show characteristics of spending more time at the hot 35 °C than at the warm 30 °C.

In conclusion, the PCs calculated from sensor 2 temperature data show a ability to discriminate between modules, in a similar way to the PCs based upon voltage data. The less healthy cluster, in this case, shows a unique defining characteristic of spending more time at room temperature than the normal warm/hot distribution of healthy modules.

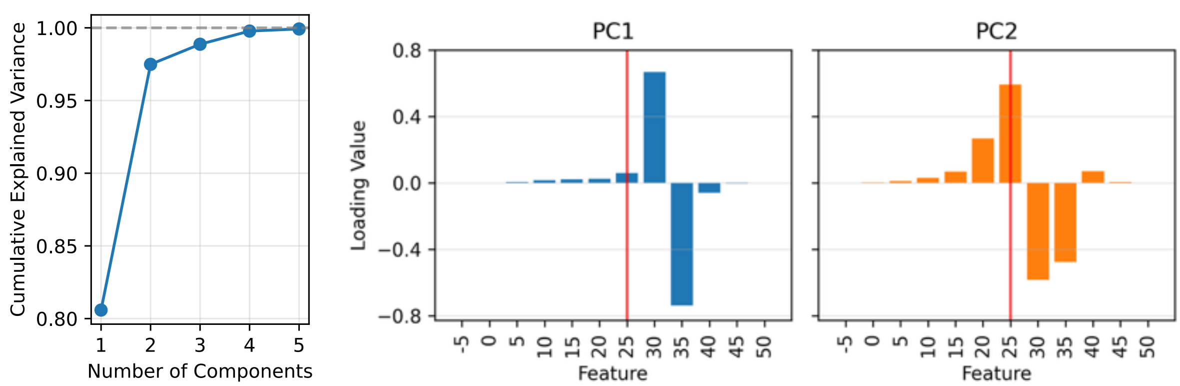


Figure 13. Metrics from principal component analysis on module temperatures from sensor 2: the cumulative explained variance of each principal component (left) and the loading values of the original variables for the first two principal components (right). The room temperature bin is highlighted with a red vertical line.

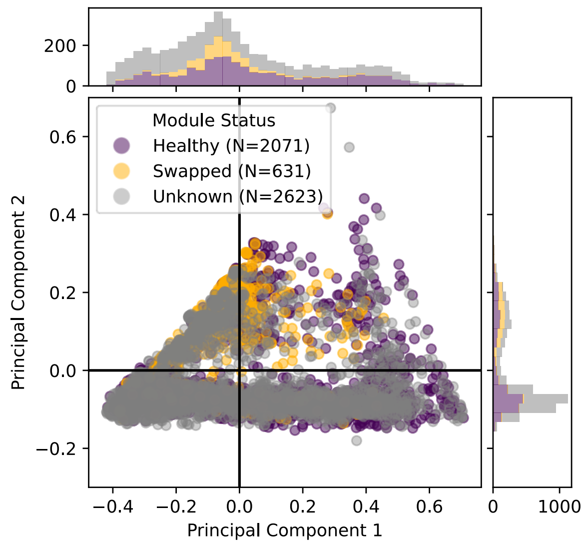


Figure 14. Biplots of battery modules across all maintenance visits scored with PC1 and PC2 calculated only from the original sensor 2 temperature variables: with all modules (left) and without unknown modules (right). Labeling is independent of the principal component analysis and displayed for interpretability purposes.

3.4 Appendix

3.4.1 Submodule Voltages Supplemental Figures

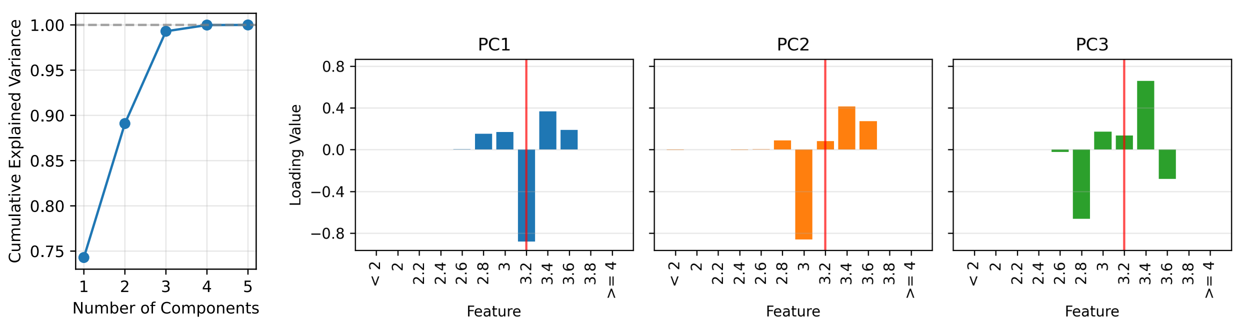


Figure 15. Metrics from principal component analysis on submodule voltages: the cumulative explained variance of each principal component (left) and the loading values of the original variables for the first three principal components (right). The nominal voltage bin is highlighted with a red vertical line.

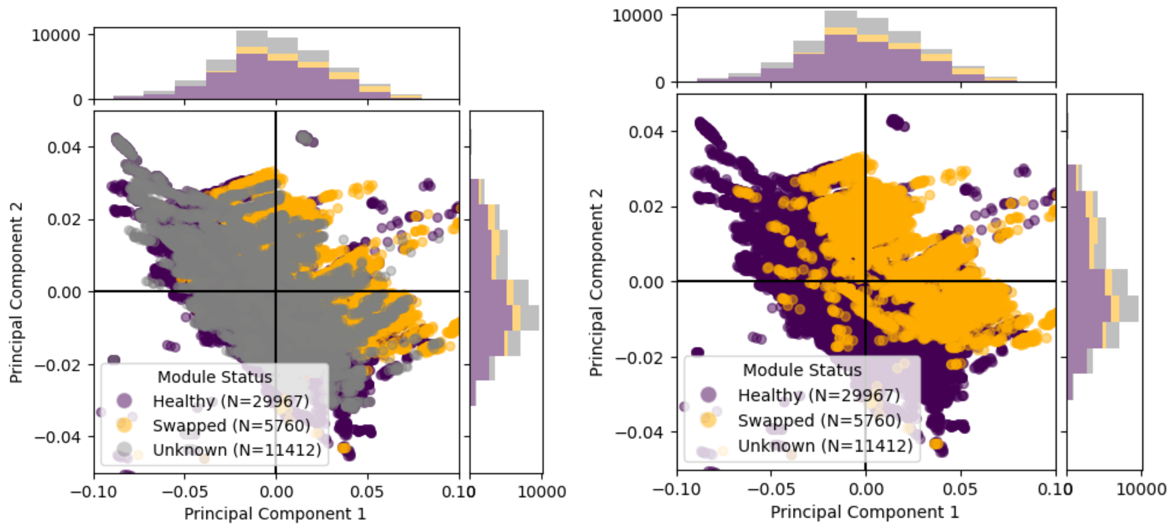


Figure 16. Biplots of battery submodules across all maintenance visits scored with PC1 and PC2 calculated only from the original voltage variables: with all submodules (left) and without unknown submodules (right). Labeling is independent of the principal component analysis and displayed for interpretability purposes.

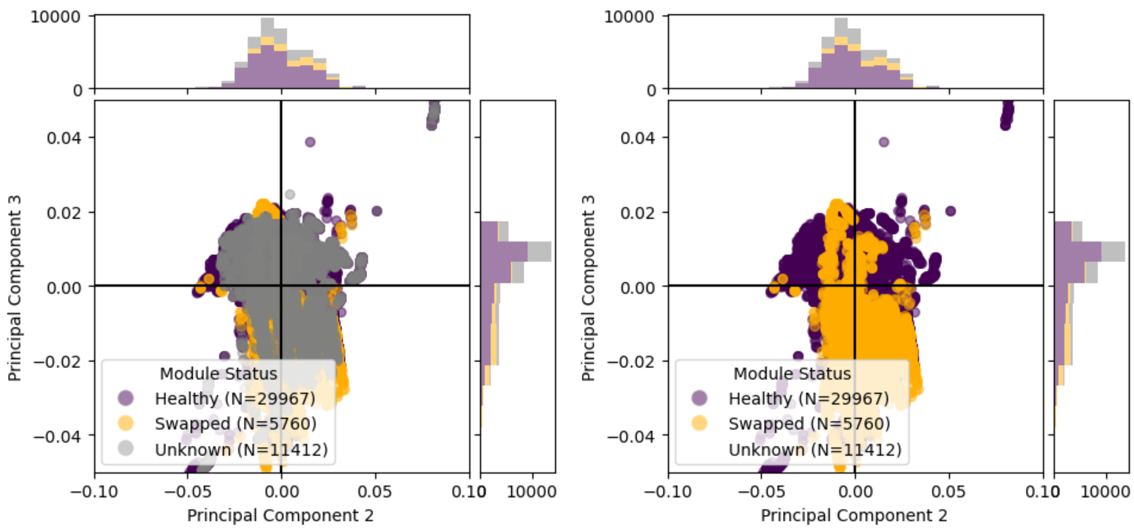


Figure 17. Biplots of battery submodules across all maintenance visits scored with PC3 and PC2 calculated only from the original voltage variables: with all submodules (left) and without unknown submodules (right). Labeling is independent of the principal component analysis and displayed for interpretability purposes.

3.4.2 *Module Temperatures, Sensor 1*

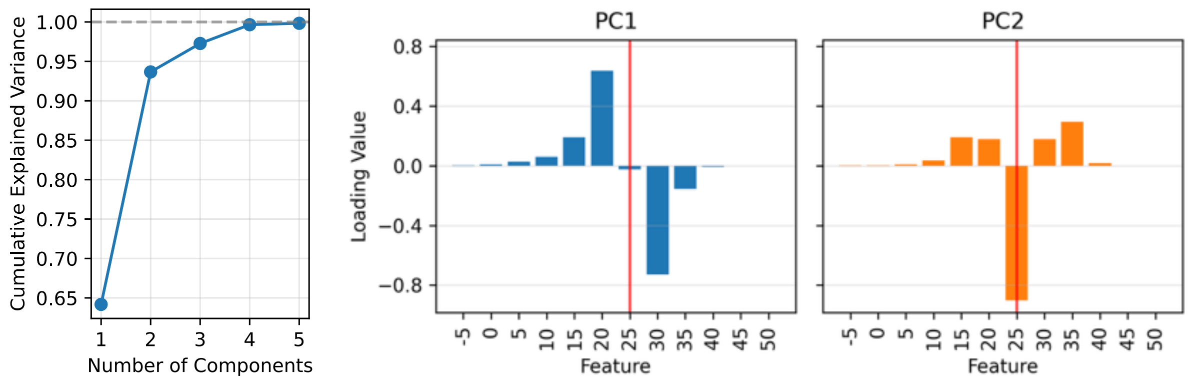


Figure 18. Metrics from principal component analysis on module temperatures from sensor 1: the cumulative explained variance of each principal component (left) and the loading values of the original variables for the first two principal components (right). The room temperature bin is highlighted with a red vertical line.

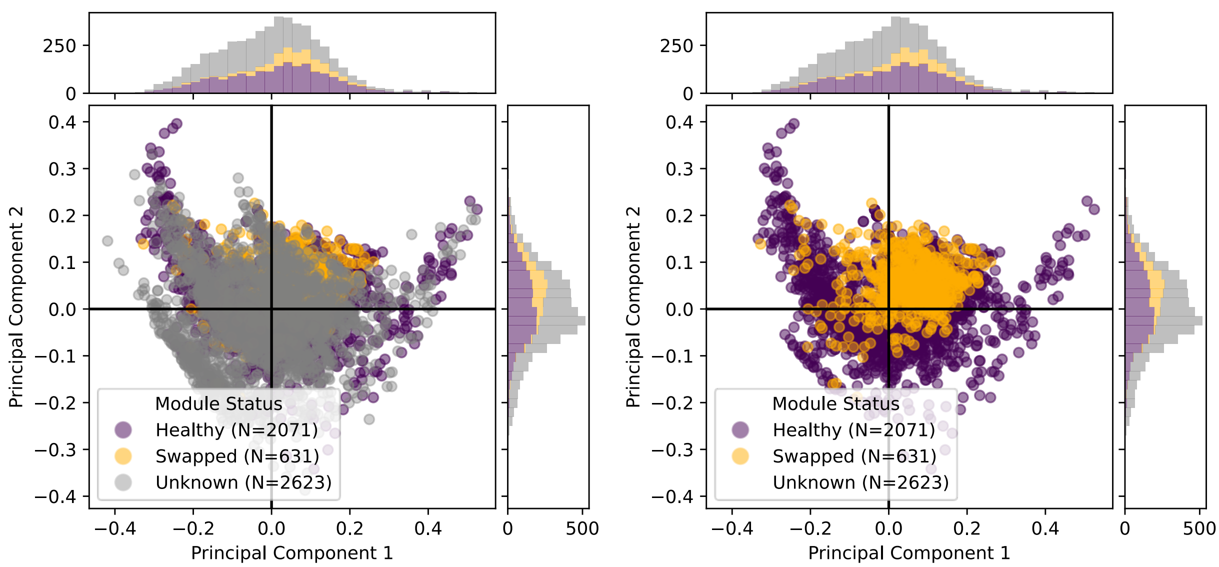


Figure 19. Biplots of battery modules across all maintenance visits scored with PC1 and PC2 calculated only from the original sensor 1 temperature variables: with all modules (left) and without unknown modules (right). Labeling is independent of the principal component analysis and displayed for interpretability purposes.

Chapter 4 SIMPLE PREDICTIVE METHODS

4.1 INTRODUCTION

In this chapter, the predictive power of the extracted PCs is evaluated by using the PC scores from module voltages as training features in linear regression models. This is implemented as another exploratory approach for understanding the discriminative abilities of the principal components found in the last chapter.

4.2 AGE OF MODULES

The first two principal components calculated from the voltage histogram data from battery modules across all files and buses (*3.3.1 Module Voltages*) showed to be the most discriminative in terms of separating modules into distinct clusters. Additionally, these module clusters exhibited strong correlations with operational age and maintenance replacements by King County Metro, as determined by post-hoc labeling. This supports the hypothesis that the clusters demonstrate operational trends linked to battery degradation after time and prolonged use. To quantify this relationship, Principal Component Regression (PCR) was implemented using scikit-learn, with the first two principal components (PC1 and PC2 scores) as features (X) and total operational time as the target (y). The data was split into 20% test and 80% training sets to evaluate the model performance without introducing bias. The regression results revealed a moderate linear relationship between the PC scores from voltage histograms and total operational time. The R^2 value was 0.549, which suggests that around 55% of the aging variance is described by the module voltages. Figure 15 displays the feature importances which are how much of each original voltage bin was used in prediction. Understandably, the voltage bins that were accounted for the most were the "wings" characterizing a large voltage spread, twice removed from the nominal voltage bin. These were the same bins included in the positive PC2 loading that showed a strong ability for separating modules into clusters. Unfortunately, these clusters are still seen in the parity plot in Figure 15, likely contributing to the merely moderate predictive ability of the PCR model. These clusters in the predictions could indicate a nonlinear voltage-time relationship, or distinctive modes of degradation in the module population.

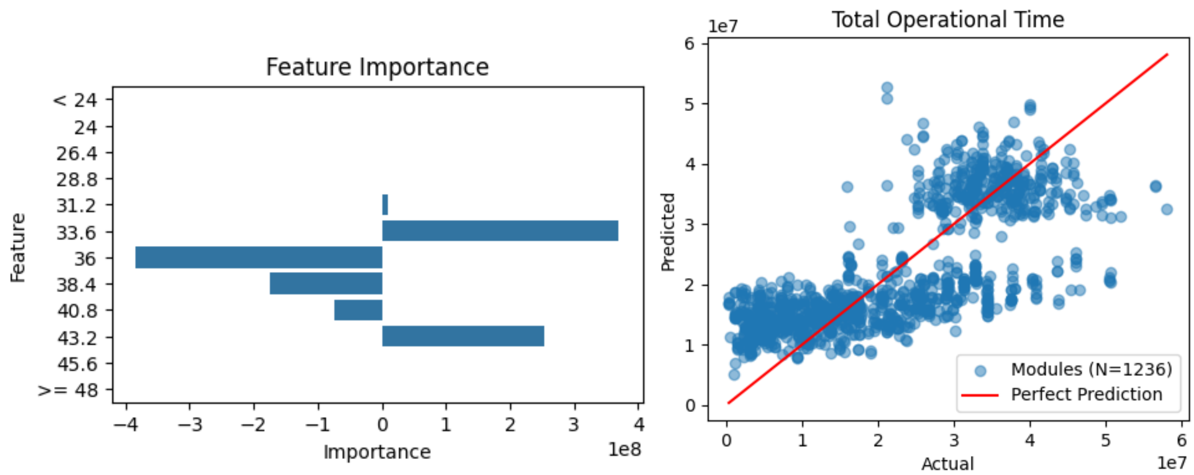


Figure 20. Predicting time to replacement with principal component regression: the feature importance of original voltage variables (left) and the parity plot of the model performance (right).

Chapter 5 CONCLUSIONS

5.1 Summary

This work focused on providing an open-source battery operational dataset from a transportation application, with easy-to-use tools and key insights into battery usage and aging over a large timescale. The accessibility of operational datasets, like this, is critical in enabling the development of data-driven battery prognostic models. These models may be able to bridge the gap in understanding between small-scale physicochemical degradation mechanisms and long-term battery usage and aging in application. Improving battery prognostics is an important step towards expediting the electrification of transportation, and reaching true carbon neutrality in this sector and beyond.

The second chapter of this thesis introduced the battery operational dataset used in this work, from a subset of hybrid-electric vehicles in the transit fleet of King County Metro. This included a description of the data collection method, the battery pack structure and metadata about the buses. The full pipeline for processing and learning from the data was discussed, with an introduction to the Python scripts for cleaning, organization, parsing and visualization. Some of these visualizations, pertaining to the average profiles of the onboard energy storage systems across the dataset, were considered with a discussion of regular operational trends. A cross-examination of operational time versus calendar time in the data for the ESS of each bus allowed a reasonable discernment of potential bus models and ages. Additionally, the fault code criteria for a battery module failure, per BAE Systems, was reviewed and an alternative post-hoc labeling method for replaced modules was introduced.

The third chapter focused on the implementation of an unsupervised machine learning method for unbiasedly condensing and exploring variations in the operational variables of temperature and voltage for battery modules. This method provided new variables, or principal components, with strong discriminative abilities for the battery modules in the dataset. In particular, the scoring of the battery modules with the new principal component values led to significant clustering that correlated strongly with operational aging and modules replacements by King County Metro. These findings support the hypothesis that there is battery degradation seen in this dataset from the apparent and identifiable trends associated with performance decline over time. In Chapter 4, the most discriminative of these principal components, extracted from module voltages, were evaluated as predictors of total operational time in a linear regression model. This model (R^2 =0.549) validated the relationship between the voltage variables and time, but did not provide strong predictions.

Beyond providing key insights into battery performance and aging, this work demonstrates the discriminative capabilities of metrics extracted from battery operational and usage data. Future work to improve these metrics in battery prognostics will need to explore methods for ground-truthing the trends established here, possibly with the acquisition of short spans of timeseries data and post-mortem analyses to make inferences about the internal states of the batteries in this application. An additional approach that could be taken is expanding the number of predictive variables, by extracting new ones for this dataset or by acquiring more recent data, to improve the fidelity of the machine learning outcomes. Ultimately, data-driven insights into battery usage and aging, like the ones provided here, hold strong promise for advancing the legitimacy and effectiveness of in-application battery prognostics.

Chapter 6 REFERENCES

[1] National Renewable Energy Laboratory (NREL), “BAE/Orion hybrid electric buses at New York City transit: A generational comparison (Technical Report No. NREL/TP-540-42217).” U.S. Department of Energy, Office of Energy Efficiency & Renewable Energy., Mar. 2008.

[2] Department for Transport, “Decarbonising Transport A Better: Greener Britain.” Great Minster House, 2021.

[3] EPA, “Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2022.” U.S. Environmental Protection Agency, EPA 430-R-24-004., 2024. [Online]. Available: https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks

[4] A. M. Ralls *et al.*, “The Role of Lithium-Ion Batteries in the Growing Trend of Electric Vehicles,” *Materials*, vol. 16, no. 17, p. 6063, Sep. 2023, doi: 10.3390/ma16176063.

[5] F. M. N. U. Khan, M. G. Rasul, A. S. M. Sayem, and N. K. Mandal, “Design and optimization of lithium-ion battery as an efficient energy storage device for electric vehicles: A comprehensive review,” *Journal of Energy Storage*, vol. 71, p. 108033, Nov. 2023, doi: 10.1016/j.est.2023.108033.

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

[7] M. Ansari, S. B. Torrisi, A. Trewartha, and S. Sun, “History-agnostic battery degradation inference,” *Journal of Energy Storage*, vol. 81, p. 110279, Mar. 2024, doi: 10.1016/j.est.2023.110279.

[8] P. Fermín-Cueto *et al.*, “Identification and machine learning prediction of knee-point and knee-onset in capacity degradation curves of lithium-ion cells,” *Energy and AI*, vol. 1, p. 100006, Aug. 2020, doi: 10.1016/j.egyai.2020.100006.

[9] S. Voronov, E. Frisk, and M. Krysander, “Data-Driven Battery Lifetime Prediction and Confidence Estimation for Heavy-Duty Trucks,” *IEEE Trans. Rel.*, vol. 67, no. 2, pp. 623–639, Jun. 2018, doi: 10.1109/TR.2018.2803798.

[10] R. Barnitt and K. Chandler, “New York City Transit (NYCT) Hybrid (125 Order) and CNG Transit Buses (NREL/TP-540-40125).” National Renewable Energy Laboratory (NREL), 2006. [Online]. Available: https://docs.nrel.gov/docs/fy07osti/40125.pdf

[11] S. Shahid and M. Agelin-Chaab, “A review of thermal runaway prevention and mitigation strategies for lithium-ion batteries,” *Energy Conversion and Management: X*, vol. 16, p. 100310, Dec. 2022, doi: 10.1016/j.ecmx.2022.100310.

[12] C. Xu *et al.*, “Electric vehicle batteries alone could satisfy short-term grid storage demand by as early as 2030,” *Nat Commun*, vol. 14, no. 1, p. 119, Jan. 2023, doi: 10.1038/s41467-022-35393-0.

[13] F. Aguilar Lopez, D. Lauinger, F. Vuille, and D. B. Müller, “On the potential of vehicle-to-grid and second-life batteries to provide energy and material security,” *Nat Commun*, vol. 15, no. 1, p. 4179, May 2024, doi: 10.1038/s41467-024-48554-0.

[14] H. Zhang *et al.*, “Life cycle environmental impact assessment for battery-powered electric vehicles at the global and regional levels,” *Sci Rep*, vol. 13, no. 1, p. 7952, May 2023, doi: 10.1038/s41598-023-35150-3.

[15] J. S. Edge *et al.*, “Lithium ion battery degradation: what you need to know,” *Phys. Chem. Chem. Phys.*, vol. 23, no. 14, pp. 8200–8221, 2021, doi: 10.1039/D1CP00359C.
[16] B. V. Vlijmen *et al.*, “Aging matrix visualizes complexity of battery aging across hundreds of cycling protocols,” *Energy Environ. Sci.*, p. 10.1039.D4EE05609D, 2025, doi: 10.1039/D4EE05609D.
[17] I. Bloom *et al.*, “An accelerated calendar and cycle life study of Li-ion cells,” *Journal of Power Sources*, vol. 101, no. 2, pp. 238–247, Oct. 2001, doi: 10.1016/S0378-7753(01)00783-2.
[18] M. Broussely, S. Herreyre, P. Biensan, P. Kasztejna, K. Nechev, and R. J. Staniewicz, “Aging mechanism in Li ion cells and calendar life predictions,” *Journal of Power Sources*, vol. 97–98, pp. 13–21, Jul. 2001, doi: 10.1016/S0378-7753(01)00722-4.
[19] “Moving to a Zero-Emission Bus Fleet: Transition Plan,” King County Metro Transit Department, May 2022.
[20] “SYSTEM MANUAL FOR HYBRIDRIVE ® Propulsion System.” BAE Systems, 2011.
[21] “KCM-DIRECT.” [Online]. Available: https://github.com/KCM-DIRECT
[22] “TheBattCave/TheBattmobile.” Accessed: May 17, 2024. [Online]. Available: https://github.com/TheBattCave/TheBattmobile
[23] “A123 Battery Specs.” [Online]. Available: https://www.buya123products.com/
[24] CPTDB Wiki, “King County Metro.” [Online]. Available: https://cptdb.ca/wiki/index.php/King_County_Metro
[25] J. Wang, D. Hu, H. Shen, T. Yang, and Y. Wang, “Optimization Methodology for Lithium-Ion Battery Temperature Sensor Placement Based on Thermal Management and Thermal Runaway Requirement,” in *2020 11th International Conference on Mechanical and Aerospace Engineering (ICMAE)*, Athens, Greece: IEEE, Jul. 2020, pp. 254–259. doi: 10.1109/ICMAE50897.2020.9178868.
[26] K. Pearson, “LIII. *On lines and planes of closest fit to systems of points in space*,” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 559–572, Nov. 1901, doi: 10.1080/14786440109462720.
[27] H. Hotelling, “Analysis of a complex of statistical variables into principal components.,” *Journal of Educational Psychology*, vol. 24, no. 6, pp. 417–441, Sep. 1933, doi: 10.1037/h0071325.
[28] M. Greenacre, P. J. F. Groenen, T. Hastie, A. I. D’Enza, A. Markos, and E. Tuzhilina, “Principal component analysis,” *Nat Rev Methods Primers*, vol. 2, no. 1, p. 100, Dec. 2022, doi: 10.1038/s43586-022-00184-w.