

Machine Learning and Telemetry Improves Outside Plant Power Resiliency for More Reliable Networks

A Technical Paper prepared for SCTE by

Stephanie Ohnmacht

Sr. Director, Product Management
Comcast
1800 Arch St, Philadelphia, PA 19403
Stephanie_Ohnmacht@comcast.com

Matthew Stehman

Senior Data Scientist
Comcast
1800 Arch St, Philadelphia, PA 19403
Matthew_Stehman@comcast.com

Table of Contents

Title	Page Number
1. Introduction.....	4
2. Identifying Power Supplies.....	4
2.1. Geolocation Match via Excel Algorithm	5
2.2. Algorithm 2.0	5
2.3. Machine Learning to Identify Power Supply through Open-Source Mapping Tools (Find My Power)	5
2.4. Manual Intervention.....	5
3. Asset Inventory Interface	6
3.1. Asset Inventory User Interface.....	6
3.2. Alarming Interface	7
3.3. Impacted Actives (nodes, cable modems) on Plant.....	7
4. Power Supply Telemetry	8
4.1. Telemetry Overview	9
4.2. Streaming Infrastructure and Real Time Eventing	10
4.3. Big Data Pipelines for Efficient Storage and Analysis	11
5. Modeling Power Supply Health	12
5.1. Modeling Objectives.....	13
5.2. Data Sets Used for Battery Runtime Prediction.....	14
5.3. Machine Learning for Battery Runtime Prediction	14
6. Conclusion.....	17
Acknowledgements	17
Abbreviations	18
Bibliography & References.....	18

List of Figures

Title	Page Number
Figure 1 - Architecture.....	4
Figure 2 - Geolocation Matching Algorithm.....	6
Figure 3 - Example Output from Find My Power	6
Figure 4 - Inventory/Alarming Tool Architecture	7
Figure 5 - PS Outage Impact Analysis Example.....	8
Figure 6 - Power Supply Overview	8
Figure 7 - Big Data Processing Pipeline Architecture	12
Figure 8 - Proactive vs Demand Maintenance.....	13
Figure 9 - Machine Learning Run Time Prediction Architecture.....	15
Figure 10 - Random Forest Performance. a.) Feature Importances, b.) Prediction Errors Compared to Baseline.....	16
Figure 11 - Example Random Forest Prediction of Battery Run Time	17

List of Tables

Title	Page Number
Table 1 - Relevant Power Supply Telemetry Metrics.....	9
Table 2 - List of Real Time Alerts.....	11

1. Introduction

As the network becomes more powerful through 10G technology, resiliency in our extensive power supply network is essential to ensure our customers are always connected. The stability of outside plant (OSP) power supplies (PS) is key to keeping the network online and serving customers. Maintaining accurate location and active telemetry information is vital to keeping the power supplies in optimal health. Machine learning is employed to analyze this massive amount of telemetry data and provide actionable insights related to operating conditions and the overall health of the power supplies and batteries.

The authors, Stephanie Ohnmacht & Matt Stehman, will present a multi-tier solution that was developed to address this at scale. The solution includes the integration of mainstream mapping technology, with machine learning (ML) routines to optimize location probabilities and use of big data pipelines and advanced data science techniques to inform proactive and demand maintenance activities. Predictive models are built on top of this large dataset to help detect long term variations in power supply performance as well as real-time performance evaluations during active outages. This paper will illustrate a proven approach to improving PS resiliency, leading to increased network reliability, and satisfying the requirements of a 10G network.

2. Identifying Power Supplies

Historically Comcast has had to maintain multiple information sources and databases to manage the network of power supplies. These separate databases included those used for network designers, technical operations, and asset inventory management. This disparate data provided an opportunity to find a solution to enable all systems to talk to each other with one unique key. Creating an innovative system to maintain PS information in the future. To start, a team was formed to begin a 4-step process to validate and correct location information for Comcast’s PS across databases.

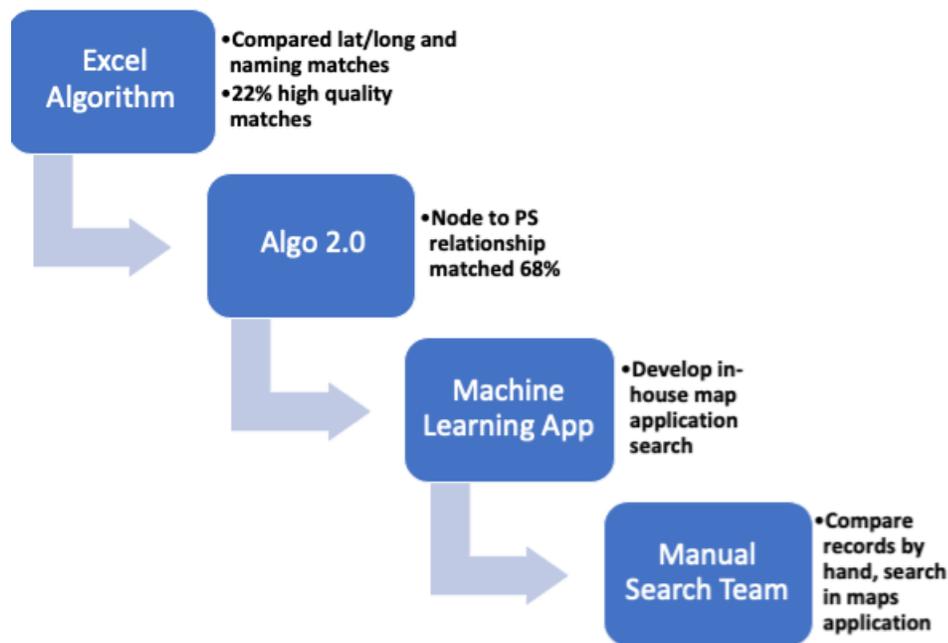


Figure 1 - Architecture

2.1. Geolocation Match via Excel Algorithm

The goal of the geolocation match was to begin to narrow down location discrepancies across databases. To determine a single PS referenced in multiple databases, the team developed an algorithm to compare the delta of linear feet between the two locations in the databases. If the PS locations were within 300 ft, then it was marked as the same location. In situations where the distance was greater than 300 ft, an index was applied to show a level of confidence of a match based on the linear feet differences. This level of geolocation matching resulted in 22% of the total population identified as having a high-quality match within the total population.

2.2. Algorithm 2.0

Once completing the first algorithm, the next phase of the analysis was to narrow down outstanding matches based on other known links within the databases. This approach focused on using node-to-power supply relationships to solve for additional location matches and increase level of confidence. This analysis increased matched records to 68% of the PS population.

2.3. Machine Learning to Identify Power Supply through Open-Source Mapping Tools (Find My Power)

A tool was developed that would identify PS in its natural environment (street or pole) using existing mapping applications to confirm the actual location. To initiate the project, power supply photos were taken in the field, then loaded into the application to begin training the system to identify different power supply types and installation environments. These examples were integral to the ability of the tool to accurately detect a PS in all situations. Once the right training model was developed, an open-source visual mapping tool was integrated into the tool that captured real street view imagery and provided associated latitude and longitude coordinates for the exact location of the identified power supply. After an attempted PS match, a level of confidence was determined by the tool and applied to the record which was then confirmed manually by a team.

This new system was able to find power supplies for 65% - 70% of the locations. The mapping tool did not have images for about 10% of the locations (given easements, landscape, and gated communities). About 45-50% of the power supplies predicted were correctly identified by the artificial intelligence and significantly boosted the team's productivity by 2X when verifying manually.

2.4. Manual Intervention

Once the team confirmed proposed matches from the machine learning tool, outstanding PS matches were set aside to be reviewed by field teams and updated.



Figure 2 - Geolocation Matching Algorithm

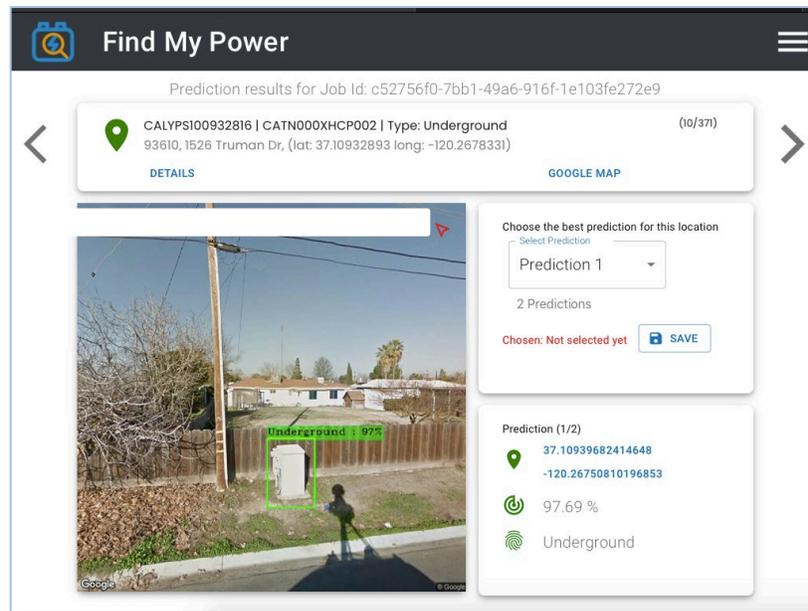


Figure 3 - Example Output from Find My Power

3. Asset Inventory Interface

Updated PS location information and node associations provided the opportunity to create a tool for management, field techs and operations called Power Supply Notebook (PSNB).

3.1. Asset Inventory User Interface

The team created a source of truth User Interface (UI) for all PS asset details. Including PS make/model/serial number, location, batteries, transponder, installation date, and maintenance records. For new PS the asset information is loaded once the PS is built and online. Ongoing updates and changes can be completed by administrators, technicians, business partners and the network help desk. Depending on the user's permission level, changes are added to a queue and approved by appropriate staff. All changes are tracked and recorded for historical purposes and investigation needs.

3.2. Alarming Interface

PS are polled every minute by in-house software. This data is monitored by the Correlation Engine (CE). As changes in telemetry flow through CE, anomalies are monitored to determine if an event is in progress. CE references all active devices on plant, downstream from the power supply and can correlate the impact of an event to number of devices impacted. Based on pre-determined threshold levels, by event type, an alert is sent to PSNB. PSNB is the presentation layer that displays current and historical alarms and events. Telemetry can be viewed by battery string, as well as the historical trends in battery voltage, input voltage, output current, temperature, output voltage, and output power. Events that need dispatching are sent to an automated dispatching system to be processed according to business rules. Alerts include PS on stand-by (on battery), PS loss of communication, and low battery.

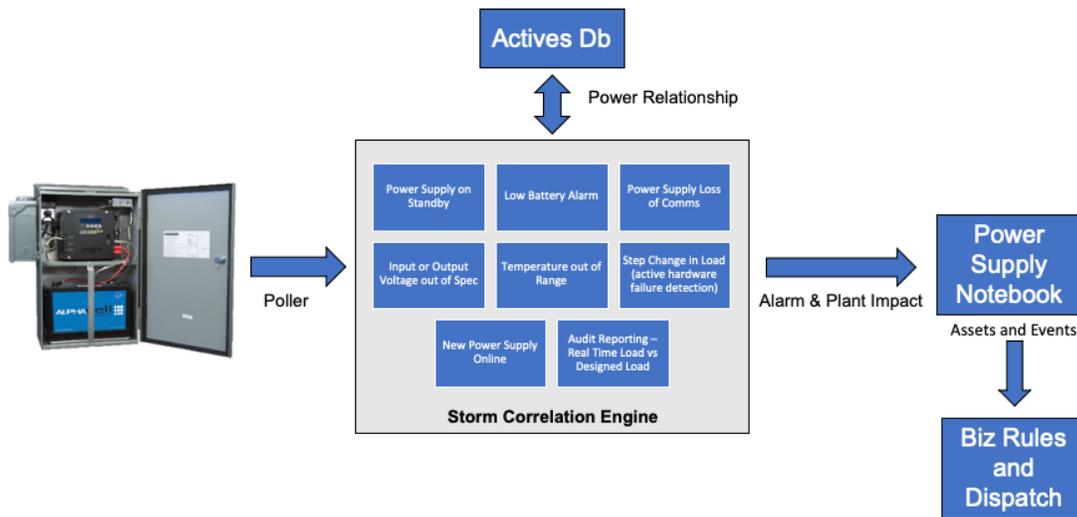


Figure 4 - Inventory/Alarming Tool Architecture

3.3. Impacted Actives (nodes, cable modems) on Plant

The data output from CE and passed to the PSNB UI, provides insightful data about outage impacts. In the past, it was assumed that any PS power outage affected all actives. Based on preliminary data, that is not the case. Data suggests that on average, 30-40% of customers stay online (have power) during a PS on discharge scenario. The geography of the power grids from the utility company and broadband provider do not overlap. As you can see in the example below, an outage from a PS only impacted a portion of the cable modems supported by that power supply. Empowered by this additional layer of information, teams can make more informed choices about how, when, and where to send a technician.

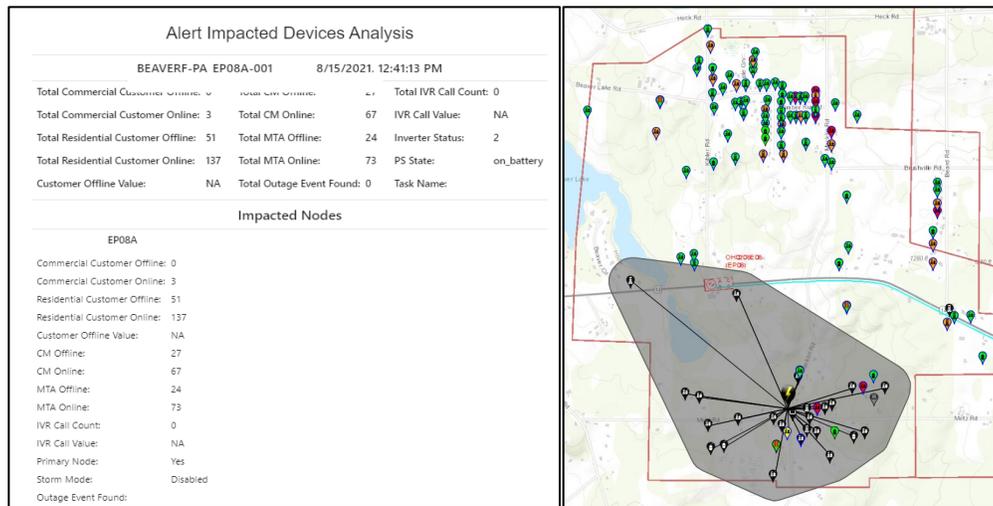


Figure 5 - PS Outage Impact Analysis Example

4. Power Supply Telemetry

The previous sections discussed the effort to build a source of truth that is foundational to the overall tracking and management of a power supply network. Once the global tracking and visibility for each power supply is established, live telemetry from individual power supplies can be incorporated to provide the status of each power supply as well as the entire network of power supplies. This section will discuss the telemetry gathered from each power supply in Comcast’s plant and how the raw data can be used to make informed decisions.

Each power supply is made up of several sub-components where each is responsible for various aspects of overall operation. Understanding the function of the different components as well as the various telemetry metrics captured is key to modelling the health of the power supplies. See Figure 6 for a simplified summary of the power supply with components relevant to this paper.

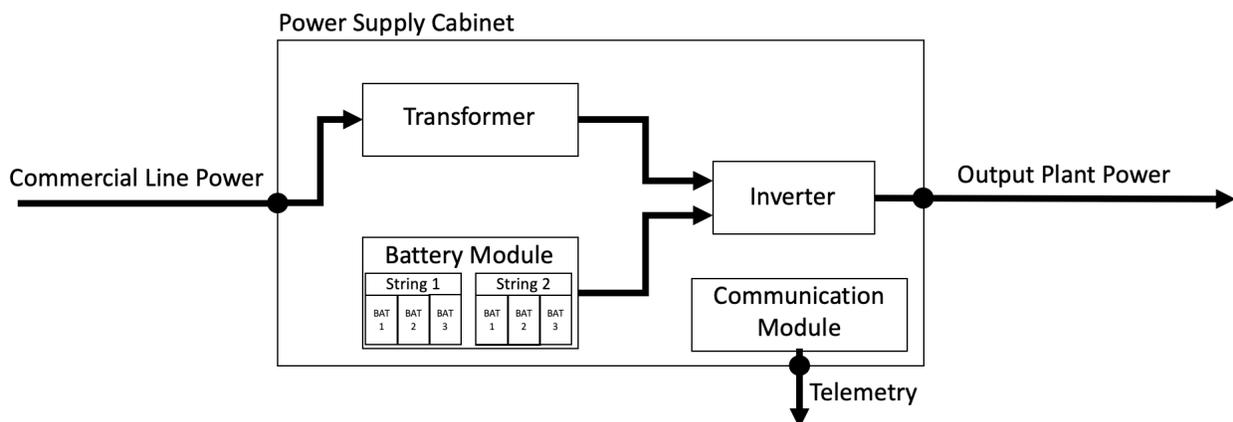


Figure 6 - Power Supply Overview

The transformer component is responsible for converting the commercial high voltage AC power down to operational range and format for the power supply. The battery module houses the strings of backup

batteries that are used to power the plant in case of a commercial power outage. The Inverter is used to convert the DC voltage from the backup batteries into AC as required by the elements of the RF plant (RF nodes, RPDs, amplifiers, etc..). The inverter also controls which power source to use in the event of a change in input voltages. The communications module is responsible for relaying all relevant information about the power supply through DOCSIS[®] protocols back to Comcast HQ for continuous monitoring.

The following sections will discuss the different telemetry metrics and data pipelines to consume the streaming data which allow for analysis and insights.

4.1. Telemetry Overview

The communications module is responsible for combining all measurements of the power supply state from the individual sub-components and packaging them into a telemetry stream that can be read through DOCSIS protocols. The PS telemetry is polled via simple network management protocol (SNMP) at two sampling intervals: every 60 seconds for high frequency engineering telemetry and hourly for configuration and system level information. The PS can also push alarms in between SNMP polls in the case of specific conditions being outside of acceptable values. Table 1 contains a simplified list of the telemetry available from the communications module.

Table 1 - Relevant Power Supply Telemetry Metrics

Frequency	Context	Telemetry Metrics	Description
Minute	Input to Power Supply	Voltage, Current, Power	Measurements of electrical input to the power supply transformer from commercial line power
	Power Supply Operation	Inverter Status	State of power supply: commercial input, battery backup/standby (outage), battery backup (self-test)
		Tamper Status	Indication of whether PS has been tampered with
		Internal Temperature	Environmental temperature of the power supply cabinet
		Alarm	Internal PS Alarms. Major: service impacting, Minor non-service impacting
	Battery Operation	Individual Voltages	Measured voltage at each battery terminal
		Total Voltage	Total voltage from all battery strings at the input to the inverter
		String Charge Current	Measured charging current at the input to each battery string
		String Discharge Current	Measured discharge current from each string of batteries
		String Float Charge Current	Measured floating charge current for each string of batteries

	Power Supply Output	Voltage, Current, Power	Measurements of electrical output from the power supply to the outside plant
Hourly	Power Supply Info	Vendor, Model Number, Serial Number	General information about the individual power supplies
		Configurations and supported functions	Overview of various configurations for each PS and which functionalities are supported
		System	Software versions, device ids, Ip/Mac addresses, etc....

The table indicates the general families of live telemetry captured from the communication modules of each power supply in the field. The minute polls tend to include more engineering specific information about voltages, currents, and statuses relevant to the functional operation of the power supply. The hourly polls contain more stable information useful for understanding the configurations and settings of the power supply.

From an operations standpoint, the power supply input telemetry gives an indication of the health of the underlying commercial power network. The inverter status is used to determine whether the PS is being powered from commercial power or back-up batteries. The output telemetry is useful for quantifying the size of the plant the power supply is powering and is a key feature in validating the power supply against design specifications. The output telemetry is also a key feature in outage runtime prediction models (will be discussed in Section 5), since how long a battery will last is largely dependent on how much current/power draw the plant requires. The PS temperature field is also useful in this regard because it is well known that batteries tend to perform worse in colder conditions due a resulting reduction in the chemical energy in each cell.

The individual battery metrics are very useful for monitoring batter health. For instance, inspection of the voltage discharge and charge curves can contain a lot of information about the health of the batteries. The durations and boundaries of these individual events contain a mass of information related to battery performance and health. Since voltage data is stored for each individual battery, a determination can be made about whether a specific battery may be performing sub-optimally as opposed to assuming the entire string of batteries needs to be addressed.

This section introduced the individual power supply telemetry metrics and gave some indication of how they can be used to monitor the health of the power supplies. The next sections will discuss the data pipelines built to consume and analyze the real-time telemetry data to provide more insights and help guide the field with resource allocation and maintenance activities.

4.2. Streaming Infrastructure and Real Time Eventing

The telemetry from each individual power supply is streamed through a real time alerting architecture where alarms and warnings can be generated based on a variety of conditions that may require action to be taken. The data is then landed in a short-term storage solution to support further deep dive analyses on recent events if needed. Due to the shear amount of telemetry being generated by the network of power supplies the raw data is only stored for a short time to minimize overhead. More information on how this challenge is handled will be provided in Section 4.3.

The streaming architecture allows for real time alerting and eventing to be performed as the raw telemetry is collected. The events are then landed in a front-end user interface for prioritization and determining the appropriate action. A general overview of the type of alarms currently implemented is given in Table 2.

Table 2 - List of Real Time Alerts

Context	Name	Description
Inverter Status	No Commercial Power	Power supply has no commercial power and is operating on battery
	Power Supply Running Test	Power supply is running a scheduled self-test to check battery performance
	Power Supply Test Failed	Power supply self-test did not last the intended duration (requires investigation)
Telemetry Variations	Output Current, Voltage, Power	Alert when measurement varies significantly from its nominal value
	High Temperature	Power supply temperature is reaching unsafe levels
	Low Battery Voltage	Power supply is operating on battery and voltages are reaching depletion levels
Communications	Loss of Communications	SNMP can no longer communicate with communications module
	PS Internal Alarm	Forwarding of internal power supply alarms

These alerts are meant to be an indication that a given power supply is currently experiencing an event of interest and requires an assessment. These alerts are currently logic based and developed from an in-depth knowledge of the operating conditions of the power supplies. A history of these alerts (including the duration the alert was active) are kept for analysis of long-term trends so analysts can evaluate trends and provide appropriate courses of action when needed.

These types of alerts are industry standard and allow for a live view of activity on the network. These alerts are only aware of a local time window around the present time stamp and thus are limited in their abilities to find complex patterns in the raw data. Thus, there is a large opportunity in this space to adopt a machine learning approach that can gain more complex insights from historical data in combination with real time data.

4.3. Big Data Pipelines for Efficient Storage and Analysis

In order to build predictive models that can learn from historical events, a historical database of training data is required. However, due to the size of Comcast’s power supply network which typically generates hundreds of billions of data points a day, it is cost prohibitive to store all the raw telemetry metrics for a long duration. To overcome this obstacle, Comcast has developed a staged strategy for identifying and storing events relevant to modeling the health of the power supplies. Figure 7 outlines the approach to only storing data that is useful for modeling.

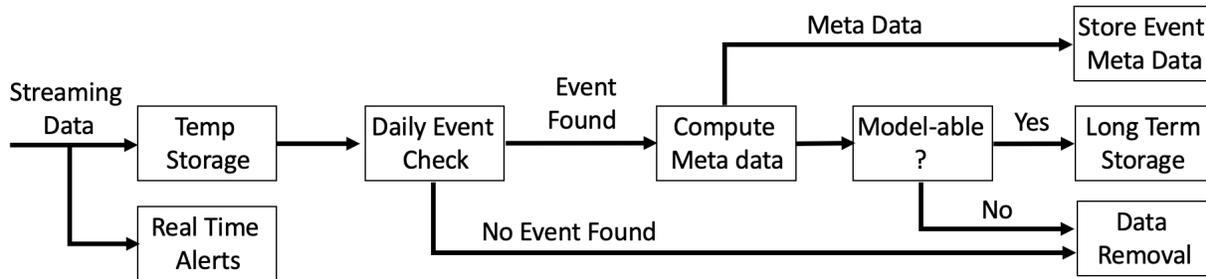


Figure 7 - Big Data Processing Pipeline Architecture

The big data pipelines consume the raw telemetry from the temporary storage solution discussed in Section 4.2. A daily job is then run that processes a single day’s worth of data for the entire footprint in one go. This was implemented because some meta data that is calculated for the individual events requires more information than a local time window would afford and thus event meta data cannot be calculated through the streaming event/alerting process discussed previously. The daily job processes the raw data and determines if the criteria for any model-able events were found. The set of events that are checked against are events that contain information related to the health of the power supply, namely: battery discharge and charge events. By looking at how batteries charge and discharge, one can infer the health of the batteries and the different approaches for this health prediction will be discussed in Section 5.1.

If an event check does determine that a model-able event has occurred for a given power supply, the entire set of raw data for the duration of that event is stored indefinitely in a long-term storage solution. All raw data outside of these model-able events are then removed as part of the standard temporary data retention policies. When the event is processed, there is a litany of meta data calculated that is stored to help describe the characteristics of the event which is placed in an event database that has no retention policy. This table is light weight since every event is only a single row in a table and the meta data allows for insights into the physics of the events computed from raw data. This approach results in storing only 0.25% of the original raw data for an equivalent period of time. Obviously, this number can change based on the number of events but by implementing this staged storage approach Comcast has significantly cut down on the cost of storing data while not missing anything relevant to long term modeling efforts.

The model-ability of an event is determined by another set of criteria on the event meta data, a few examples are event duration, data quality, etc. For example, while a power supply may need to go into battery mode for a split-second commercial power outage, that event might not be useful for modeling the health of power supplies due to the very short duration. The next section will discuss more about the modeling objectives and how the raw telemetry values can be used in building models that can deliver on those objectives.

5. Modeling Power Supply Health

The previous section introduced the power supplies and telemetry metrics as well as the data processing pipelines that consume the raw data to provide real time alerts and pre-process for future modeling efforts. This section will discuss the different objectives of predictive models in the power spaces as well as the different modeling approaches and some initial results from this relatively new initiative.

5.1. Modeling Objectives

In the cable domain, a positive customer experience is the main objective of all efforts. For outside plant power that means keeping the plant online and customers connected at all costs. Thus, a predictive model should use all available data to determine if a customer impacting event is imminent or the impact assessment of any live customer impacting events. These types of scenarios in the cable world are typically referred to as proactive maintenance and demand maintenance. Many initiatives exist in the RF network maintenance realm such as Wolcott, et al. (2016) and Wolcott, et al. (2018) and many practices have even become industry standard as documented in CableLabs PNM Best Practices Primer (2020). For the power supply domain most efforts have been focused on improving the engineering models related to the power supplies such as Anderson and Burgett (2014). The authors believe that there is a large opportunity to enhance the overall field of power supply management and maintenance with machine learning which can consider many more inputs outside of raw engineering data. Figure 8 showcases the distinction between proactive and demand maintenance.

	Customer Impact	Resolution Window
Customer Impact	Now: Experiencing Impact Future: Impact Imminent	(Now, Now) Demand Maintenance
Resolution Window	(Future, Future) Proactive Maintenance	Now: ASAP Future: Months, Weeks

Figure 8 - Proactive vs Demand Maintenance

Proactive maintenance is the ability to proactively detect and fix any potential issues on the network before a customer is impacted. For power supplies a simple but effective passive technique would be to do a standard workup on every power supply on a regular schedule, say annually. While this approach is generally good at keeping power supplies operational in the long term a more active approach can help address any potential issues in a more targeted fashion. This is where the telemetry-based models will be able to thrive. For example, by monitoring the charge and discharge curves during all events for each power supply, models can be built that attempt to catch any long-term degradations in power supply performance. This type of model can start to predict slow variations in the performance over many months of run time and help alert the field that certain power supplies may need to be visited sooner than others.

In a demand maintenance scenario, there is an active situation, and a model would need to predict the outcome of a given event to help the field teams allocate resources in real time. In this instance the data related to a specific event is heavily relied on in addition to any relevant historical data to predict and score potential outcomes of the current event. For example, say a severe weather event results in a loss of commercial power and the power supply needs to operate on battery backup to keep the plant and customers online. A predictive model in this case would need to estimate how long the batteries will run

until they will reach end of discharge. Based on the model outputs, the field teams can prioritize which power supplies may need a backup generator if the run time is expected to be less than the commercial power outage duration.

In either case, the output of the respective model will be consumed by field teams and prioritized and worked accordingly. Therefore, any additional information that can be supplied to the teams performing the work will help increase efficiency and add confidence to how the work will be prioritized. For this reason, the models that will be explored in this effort need to have some avenue for explainability of the given prediction. Examples of explainability not only include a level of confidence in each prediction but also elements like global and local feature importance to indicate what features the model believes to be most important in making a prediction.

Comcast is currently focusing on a demand maintenance use case using a machine learning based run time prediction model for power supplies operating on battery backup during commercial power outages. Comcast sees many potential avenues for using the aforementioned data sets and methodologies to enhance the outside plant power domain on many fronts, from long term budget planning and global resource allocations to using the power supply telemetry in other models related to access network root cause analysis.

5.2. Data Sets Used for Battery Runtime Prediction

For the goal of predicting how long the power supply can power the plant in the absence of commercial input power a machine learning model will need access to data that contains predictive power in this instance. It is apparent the telemetry related to the electrical circuitry discussion in Section 4.1 would be required to describe the physical principals occurring. In fact, the state-of-the-art modeling in the battery domain pretty much solely rely on the electrical data to make run time prediction, i.e., physics based and even more modern time series-based approaches.

However, as discussed in previous sections, Comcast has put in a significant amount of effort to create an accurate inventory of all the power supply assets which can provide extra enhancements to the modeling efforts. Items of particular interest are the battery manufacturer, age, type, install date and number of batteries in each power supply. This battery specific information will help enhance a predictive model using global patterns from various meta data fields related to the identity of the batteries and power supplies.

The final class of data under consideration as input to a machine learning model are features that can be engineered from the aforementioned data sources. The process of feature engineering has proven to be one of the most important aspects in developing high performing machine learning models. Features that are derived from the raw input data can produce significantly more accurate results than the raw features alone. While Comcast is still exploring the different approaches to feature engineering in this capacity, some immediately obvious features of interest for a run time prediction could be: information about the last discharge (test or outage) of a given power supply like time since, duration of, ending voltage, variance of individual battery voltages; other relevant statistics as well as historical information including number of discharge cycles in last n months, number of cold temperature cycles, time since last maintenance visit.

5.3. Machine Learning for Battery Runtime Prediction

The requirements for a run time prediction model are straightforward: when commercial input power is lost the model should predict how long the power supply can run on battery power before it can no longer power the plant. Figure 9 outlines the overall architecture of the desired model where the model takes

inputs from the aforementioned data sources and makes a run time prediction at the moment that the commercial power outage begins, as well as a confidence in the prediction and some form of explainability in what led the model to making that prediction. The 3 outputs combined will be the key aspect of how the field will prioritize and ultimately address predictions that indicate intervention will be needed prior to commercial power coming back online.

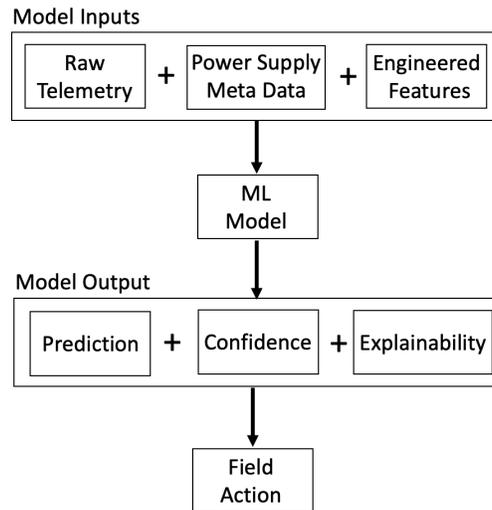


Figure 9 - Machine Learning Run Time Prediction Architecture

Given the variety of input features and the desire to have a form of model explainability, the initial machine learning architecture is a random forest (RF). Random forests have proven to excel in finding nonlinear patterns in data with limited initial assumptions and they also have the added benefit of the ability to add some sense of explainability. The results of the random forest model will be compared against a baseline physics-based model to determine the efficacy and potential enhancements over the baseline model that is implemented in production. More information on the baseline model can be found in Lin and Nispel (2022).

The random forest was trained on approximately 2500 examples where the power supplies operated on battery backup until it reached end of discharge and safely shuts down. The features used in this initial model include the initial conditions of the physical properties at the beginning of the event and some meta data on battery information (type, age, count). The model then makes run time predictions on a test data set of 500 examples compared against the predictions from the baseline model. The results of the random forest are shown in Figure 10 where Figure 10a shows what features the model found most important during model training and Figure 10b shows the distribution of individual prediction errors in minutes compared to the baseline model.

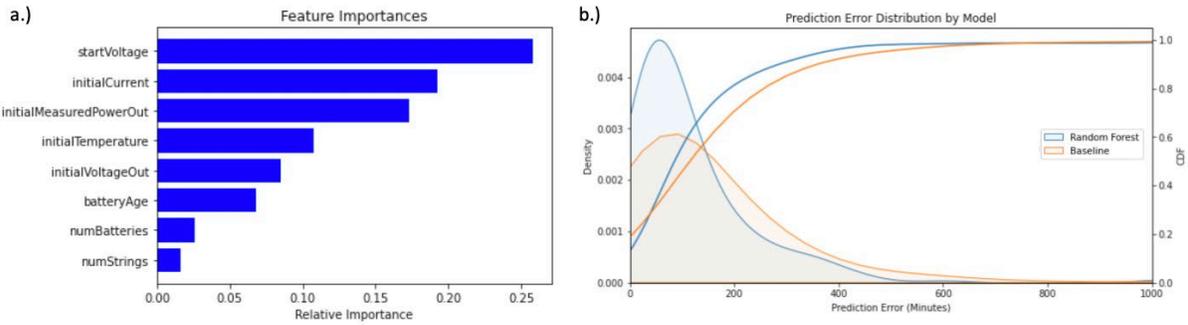


Figure 10 - Random Forest Performance. a.) Feature Importances, b.) Prediction Errors Compared to Baseline

As expected, the random forest model finds the most important features to be the average battery voltage in the first minutes of the event as well as the initial current draw and power consumption of the plant. These features combined essentially indicate the current capacity of batteries and how demanding it is to power the plant. The next group of features which are also found to be important are the ambient temperature at the beginning of the event, the output voltage of the power supply and the battery age in years. The model indicates that the number of batteries and the number of battery strings are marginally important for predicting the run time, likely due the majority of samples having a similar configuration. The random forest model tends to produce smaller errors compared to the physics-based baseline model as shown by the smaller tail in the distribution. As the model is enhanced with more features, the distribution of errors is expected to continue to shift lower indicating improved performance.

As discussed previously a random forest can also provide a measure of explainability for individual predictions. Figure 11 displays an example prediction from a particular sample. One can see that the output includes the prediction, confidence interval and a form of insight into the prediction. In this case the predicted run time was 270 minutes while the actual run time was 222 minutes. Along with the raw run time prediction, the 50% confidence interval is derived from the distribution of all the individual decision tree predictions. The width of this confidence interval gives an indication of how confident the model is in its raw prediction. Finally for a random forest, the raw prediction can be thought of as mean prediction plus a contribution from all the feature values for a particular sample. In the case below the model's mean prediction from training data is 239 minutes and features with blue bars increase the predicted duration over the mean and red decreases the predicted duration relative to the mean. In this case the most significant contribution is due to the battery age field, in this case the batteries are only 1 year old, so relatively new, which is why it is increasing the predicted duration. These types of supplementary information in addition to a run time prediction can help the field action and prioritize more efficiently.

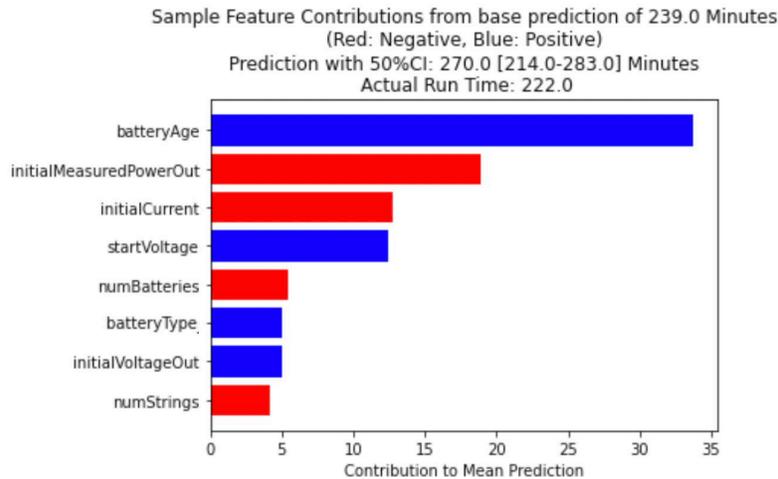


Figure 11 - Example Random Forest Prediction of Battery Run Time

This section presented an initial proof of concept model using random forests for predicting the battery run time for power supplies which performs on par if not slightly better than the current physics-based prediction model. More advanced features are being developed which are expected to significantly improve the predictive accuracy over the current models. Comcast hopes to use the results of this model (once fully validated) over the long term as an input to a proactive maintenance model. For example, this model could be run periodically during nominal conditions (in the absence of a commercial outage) to determine if a power supply would perform to spec in the event of an outage and therefore open a window to perform proactive maintenance on suspect power supplies.

6. Conclusion

Spending time and resources to go deep into the power supply environment and build a product to support and improve network powering reliability has the potential for huge resource savings. We now have visibility into customer impact during power outages by correlating backend systems and building powerful a front-end interface. The new systems and interfaces combined with real time power supply telemetry open the doors for machine learning models to significantly enhance the proactive and demand maintenance capabilities of field teams. Machine learning models can not only provide more accurate predictions but also allow for explainability of given predictions to further support planning efforts. These initiatives are rooted in providing an amazing customer experience and having a more resilient power supply network is helping Comcast continue towards that goal.

Acknowledgements

The authors would like to acknowledge our colleagues that contributed significantly to the success of this effort. Chris D’Andrea was instrumental to setting up the big data pipelines and exploratory data analysis of the events in preparation for machine learning models. Ilana Weinstein developed the first iteration of the random forest model and is continuing to make enhancements.

Cory Thompson was a key resource during the development of PSNB, CE and Find My Power to ensure the functionality, benefits, and accuracy were developed and working in production as designed. Alex Falcon led the strategy and prioritization of the tool user story, minimum viable product outcome and change management in the field.

Abbreviations

DOCSIS	Data Over Cable Interface Specifications
ML	machine learning
OSP	outside plant power
PS	power supply
PSNB	Power Supply Notebook
RF	random forest
SCTE	Society of Cable Telecommunications Engineers
SNMP	simple network management protocol
UI	User Interface

Bibliography & References

Anderson, R. & Burgett, J. (2014). *Improving Network Reliability Through Effective Power Management*. SCTE Cable-Tec Expo 2014, Denver, Co.

PNM Best Practices Primer: HFC Networks. CM-GL-PNM-3.1-V01-200506. Cable Labs, 2020.

Lin, K., & Nispel, M. (2022). *A New Model for Plant Power and Health Estimation*. SCTE Cable-Tec Expo 2022, Philadelphia, Pa.

Wolcott, L., Heslip, J., Thomas, B. & Gonsalves, R. (2016). *A Comprehensive Case Study of Proactive Network Maintenance*. SCTE Cable-Tec Expo 2016, Philadelphia, Pa.

Wolcott, L., O'Dell, M., Kuykendall, P., Gopal, V., Woodrich, J. & Pinckernell, N. (2018). *A PNM System Using Artificial Intelligence, HFC Network Impairment, Atmospheric and Weather Data to Predict HFC Network Degradation and Avert Customer Impact*. SCTE Cable-Tec Expo 2018, Atlanta, Ga.