WDRM V3 Documentation

Revision 1

Table of Contents

Executive Summary

This document describes the Wildfire Distribution Risk Model (WDRM). The WDRM is the primary model that PG&E (Pacific Gas and Electric) uses to understand how much wildfire risk is posed by the overhead distribution network, where the risk is concentrated, and why the risk exists. The WDRM's purpose is to inform PG&E's Wildfire Mitigation Plan (WMP) and the work plans developed by mitigation program teams to support PG&E's stand that, "Catastrophic wildfires will stop".

The purpose of this document is to describe the development of WDRM v3 (third version) by the Risk and Data Analytics (RaDA) team, explain significant model improvements over WDRM v2, and provide a high-level analysis of the model results. Asset location and attribute data, risk modeling approaches, modeled risk predictions, and mitigation risk reduction predictions will be described in detail.

The document addresses the following topics, summarized here for convenience:

The purpose of the WDRM and how is it used

The WDRM quantifies the risk of wildfire ignited by the overhead distribution grid to support wildfire mitigation planning and prioritization. It assesses expected wildfire risk over annual and multi-year timescales, given the conditions and assets associated with past grid events, including outages, PSPS damages and hazards, and ignitions, from 2015-2021. To support the development of work plans around specific categories of mitigations and allow model structures to reflect what is known about specific failure modes, sub-models are trained on well-defined non-overlapping categories of grid events – subsets of all events defined by shared cause, sub-cause, and assets involved. Sub-models are tuned to the asset and environmental attributes expected and empirically confirmed to best predict grid events within each subset. Subset-level estimates of ignition probability are multiplied by wildfire consequence values derived from simulated location/fuels/conditions-specific wildfire outcomes calibrated to the satellite record of historical wildfires in California to produce asset- and location-specific risk estimates. The resulting risk values can be aggregated across subsets to produce "composite risk" estimates that span specific groups of (or all) subsets, capturing total risk while providing planners the ability to drill into contributing causes and the assets involved. These results are available to planners via interactive maps that overlay grid asset data or via roll-up to circuit segments, the planning units of system hardening, defined as the segments of grid infrastructure protected by the same protective device.

Comparison of the improved version 3 to past models

The WDRM v2 consisted of two separate models built to support vegetation management and system hardening planning, focusing on vegetation-caused and conductor-involved ignitions, respectively and confined to the High Fire Threat Districts (HFTD)s. The WDRM v3 expands prediction coverage to the full PG&E distribution system and all categories of grid events, including third party, animal, support structures, transformers, vegetation caused, and equipment failure for several types of equipment and unifies these models into a single framework and set of results – via the sub-models and compositing described above. To maximize the information produced by the WDRM and the training data available to sub-models, v3 first trains on and predicts failures capable of producing ignitions (e.g. outages) and then estimates the probability of an ignition arising from the same cause as the failures in question. This later step is based on the ignitions subset categories of cause, sub-cause and asset type(s) and local environmental conditions, including vegetation types, fuel dryness, and wind. Ignition probability values are then multiplied by estimates of wildfire consequences of ignitions starting at all locations with grid assets and available fuels to produce location-specific risk values. The v3 consequence

calculations calibrate wildfire simulation outputs against historical outcome data on California wildfires from satellite observations and fire agency records. The WDRM v3 also incorporates improved model inputs, including Light Detection and Ranging (LiDAR) and satellite derived estimates of fall-in tree count (based on height and location), tree species data, improved and corrected ignition data, improved sources of event location data, improved tracking of which assets were involved in grid failures, and improved asset attributes, including changes over time. The net outcome of all these changes is improved predictive performance, greater coverage of events of all types, providing results for the full-service territory, and improved ability to quantify the contribution from specific causes, sub-causes, and asset types to overall risks – a feature necessary to plan well targeted mitigations.

WDRM v3 risk calculation, construction, and performance

The WDRM v3 defines risk as the product of the Likelihood of Risk Event (LoRE) and Consequence of Risk Event (CoRE). More precisely, the LoRE is the expected count of ignitions derived from the probability of an ignition of a given cause and sub-cause involving specific asset types and the CoRE is the expected consequence of an ignition, quantified by calibrated simulations. Risk can be calculated for specific subsets of grid event or aggregated across all subsets. If used to rank work locations and compared to work performed without ranking, the top 20% of probability of ignition predictions reduce the amount of work necessary to address an (out of sample) ignition from the historical data set by a factor of 2-4x, depending on the subset of events in question. In other words, the WDRM v3 can specify double to quadruple the risk mitigation for the same effort when compared to unranked work.

Calculation of the possible reduction in risk from mitigation efforts

The WDRM v3 defines mitigations via efficacy factors that determine the fraction of ignitions that are expected to be avoided by their implementation. For example, an efficacy of 0.75 would reduce ignitions by 75% of their unmitigated value (i.e. to 25% of the original). These efficacy factors are only applicable to the specific subset(s) of grid events they are capable of addressing and are expected to internalize achievable efficacy – for example, after accounting for private landowners who refuse to allow their trees to be cut under vegetation management efforts. Given efficacy values and the rules that govern their application to specific subsets, ignition probability estimates for relevant locations/assets are decreased by the amount dictated by the efficacy values. The mitigation effects are computed and applied everywhere to represent the outcome from theoretical maximal coverage of mitigation efforts. Planners are then instructed that only the locations where they perform the mitigation can be claimed as expected reduction and that multiple mitigation strategies cannot be applied to the same locations with additive outcomes. This pre-computation of potential mitigation aids in work planning and making tradeoffs between potential mitigation strategies.

To develop the underlying efficacy factors, RaDA worked with subject matter experts responsible for developing, planning, and implementing mitigations to elicit their subset-specific expected efficacy and the conditions under which they expect to achieve them. Some efficacy factors are defined directly at the subset level while others are aggregated to the subset level based on values provided at based on event- or asset-specific attributes.

New insights on wildfire risk

Risk is more strongly influenced by expected consequences than ignition probabilities. This is verified by examining the dramatic difference between the outcome of small and large (the vast majority with modest consequences each) vs. the destructive and catastrophic (comparatively rare but orders of magnitude more consequential) historical fires. By comparison, the range of ignition probability values is more constrained than the range of consequences. In addition, ignition probabilities tend not to be highest in locations where wildfire

consequence is highest. While this is fortunate for the rate of destructive wildfires it means that reducing the risk or occurrence of destructive wildfires may not be accompanied by a commensurate reduction in ignitions.

The top 20% of probability-ranked locations and assets host 2-4x the count of (out of sample) ignitions, depending on the subset in question. All else being equal, this suggests that the same amount of mitigation effort can address 2-4x the risk when prioritized vs. unprioritized.

No data set is perfect, yet event and asset data quality are very important drivers of model fits and performance. Hard work improving event cause and sub-cause, locations, and assets involved were all rewarded by model performance gains. This reinforces the value of ongoing work on data quality to be done in the future – and of models that are robust to uncertainty in outlying data labels and values.

Different event subsets have both qualitative/quantitative differences in location- or asset-specific based failure/ignition probabilities, and those patterns reflect what is understood about the causal processes that lead to grid events. For example, vegetation caused ignitions tend toward locations with more dense counts of fall in tree events, animal squirrel events are bifurcated between suburban and forest locations, and support structure failures are more likely with older equipment. These intuitive outcomes affirm that the machine learning methods have tuned their selected features and fit parameters in a manner that are both subset-specific and consistent with pre-existing knowledge – attributes that lend credibility to their finer scale and subtle predictions as well.

Ignitions can be predicted using raw outage probabilities but improved by passing outage probabilities through a probability of ignition given an outage model. This is because the probability of ignition given an outage model can account for subset-specific character of outages (i.e. more or less likely to cause a wire down or to occur during fire weather) and locations that experience more ignition conducive to wind, vegetation types, and fuel moisture.

Schedule for future improvements to the WDRM

Wildfire risk modeling has made marked improvement in the last year both at PG&E and other utilities and model developers. Nevertheless, it is still in a rapidly evolving phase. In support of PG&E's stand that catastrophic wildfires shall stop, PG&E will produce annual updates to the WDRM. As part of this continuous improvement, the next WDRM, v4 WDRM, is planned for finalization during Q1 of 2023. For this update planned improvements include consequence values that account for difficulty of suppression, and factors that influence the effectiveness of evacuation in advance of fires. The model will continue to account for PSPS hazards and damages as proxies for grid events that would have occurred with the power on and will add the ability to account for the effects on EPSS (protective device settings that make them more likely to trigger protective outages under fire conditions). Another area of focus is allowing for algorithmic specification of the circumstances under which specified mitigations are applicable, leading to a more precise application of mitigation effectiveness factors only to the degree to which they can be achieved – for example, the use of less combustible transformer fluids will be applicable only to transformers that lack them. Finally, the v4 WDRM will be based on 2022 updates to event, asset, weather, and vegetation/fuels data and will benefit from ongoing efforts to test out and incorporate new or improved model structures and covariates to improve model predictive performance.

1 Introduction

1.1 PG&E's Wildfire Risk Management Ecosystem

PG&E employs several risk models for Wildfire Risk Management. The models in PG&E's risk management ecosystem include:

- Planning models
  - Enterprise Risk Model
  - Wildfire Transmission Risk Model (WTRM)
  - Wildfire Distribution Risk Model (WDRM)
- Operational models for Public Safety Power Shut-off (PSPS) and Enhanced Power System Settings (EPSS)
  - Distribution OPW and IPW models
  - Transmission Operability Assessment

The outputs of the planning and operational models, along with input from Public Safety Specialists (PSS) and other Subject Matter Experts (SME), are used to form PG&E's Wildfire Mitigation Plan (WMP).

Figure 1 – PG&E Risk Management Ecosystem

The WDRM is the primary risk model that PG&E uses to prioritize mitigation work to reduce wildfires initiated by the distribution grid. The WDRM produces predictions of wildfire risk and estimates of risk reduction for potential mitigation activities. Mitigation program work planners use the WDRM results, along with their mitigation program work planning tools to develop work plans that systematically reduce wildfire risk while considering constraints such as budget allocation, human and equipment resource capacity, and regulatory commitments.

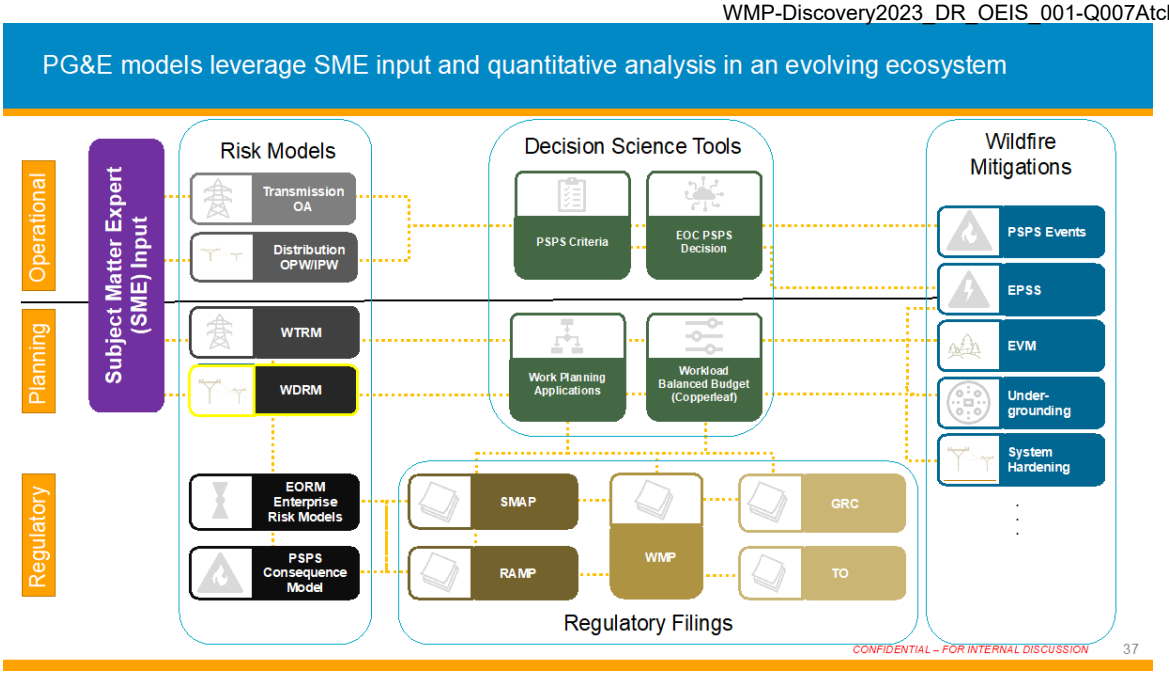


Figure 2 – Ecosystem View 2

Wildfire risk modeling, mitigation work planning, and mitigation work execution occur on overlapping time cycles. Figure 3 illustrates the expected coordination of the development and execution workflows.

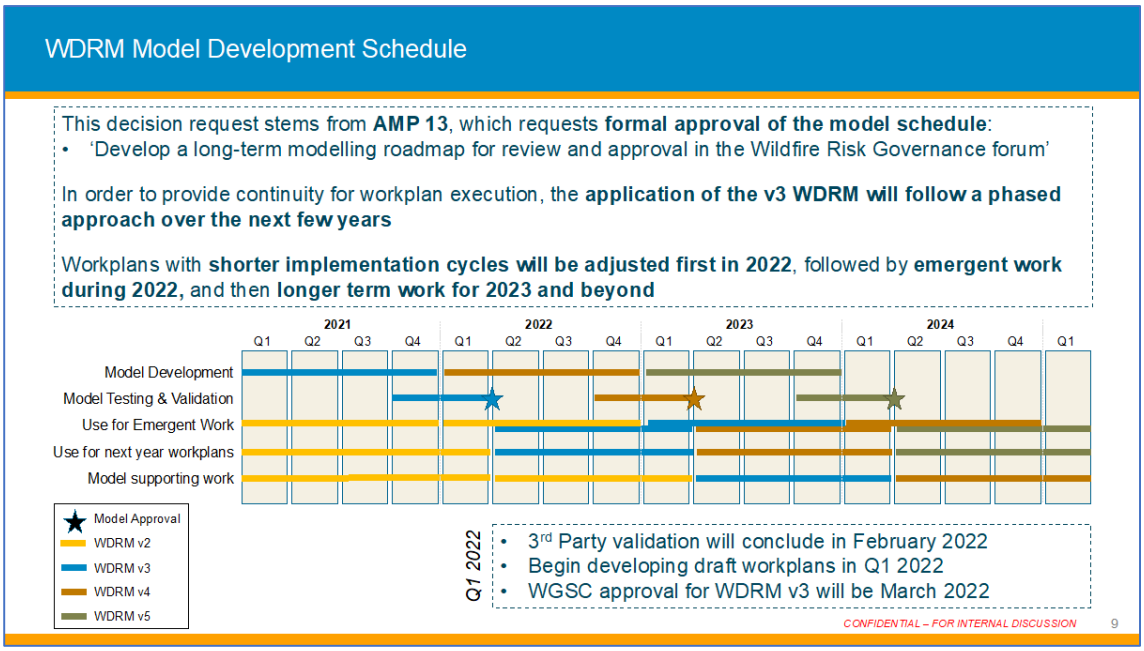


Figure 3 – WDRM Development Schedule

1.2 The Wildfire Distribution Risk Model (WDRM)

The WDRM was developed by PG&E's Risk and Data Analytics (RaDA) team to quantify the risk related to PG&E's overhead distribution equipment (also referred to as assets), which consist of overhead transformers, poles and other parts of support structures, lines (conductors), and line-related equipment such as interrupters.

The WDRM provides predictions of the where, why, and how much wildfire risk occurs during a wildfire season of June 1st through November 30th.

The WDRM estimates risk at geographic locations along the distribution network and for individual assets. The model makes a progression of predictions for each location or asset, first determining the likelihood of equipment failure or outage, then the likelihood of ignition given a failure or outage, and finally the expected consequence of an ignition. The total risk for a location or asset is calculated as the product of the probability of ignition and the consequence value. The total risk values are used by mitigation work planners to prioritize program mitigation efforts.

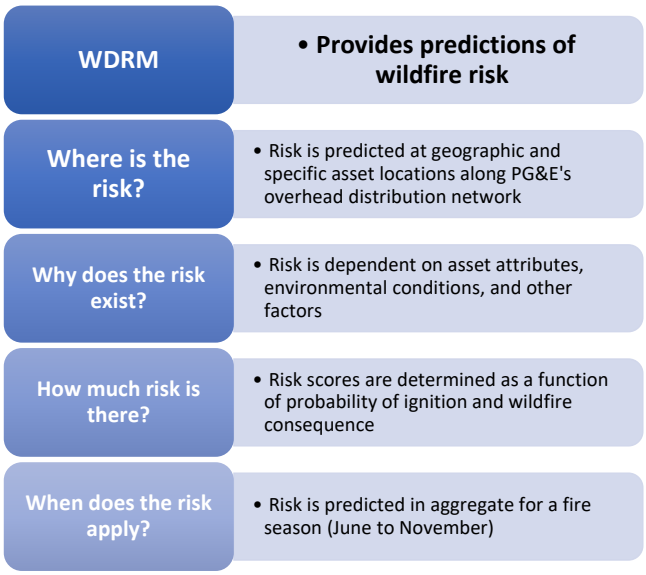


Figure 4 – WDRM Objectives

It is important to understand that since the WDRM is built as a support model for wildfire mitigation planning, the model therefore predicts risk as an aggregate value for an annual fire season. The model does not predict specific risk for specific, shorter time intervals within a fire season such as on a daily or monthly basis. The WDRM model results are relevant only for annual and multi-year planning programs. There are other models in the PG&E wildfire risk management ecosystem, such as the PSPS model, which are intended to inform on operational risk mitigation using near-term data and forecasts.

Through modeled relationships between wildfire risk and an array of environmental and asset attributes, the WDRM helps mitigation program work planners understand why risk exists at a location or asset, and, therefore, which action(s) may be most effective for reducing the wildfire risk.

Wildfire risk is estimated by first estimating its component parts, then multiplying them together. The two parts are:

1.2.1 Likelihood of a Risk Event (LoRE)

LoRE refers to the probability of a fire ignition event, or specifically the likelihood of an ignition occurring at a given location across the duration of a typical fire season. In WDRM v3, LoRE is predicted by 17 separate model subsets, each of which predicts LoRE for a specific combination of risk driver and assets. These 17 model subsets are grouped into three categories for the purposes of this document; contact from object models, conductor & other equipment-related models, and support structure & transformer-related models. Each of the 17 model subsets in WDRM v3 predicts LoRE in two stages: First by predicting the probability of an outage from a given category, and second, multiplying the probability of outage times the probability of an ignition resulting for an outage of that category. LoRE is described further under 3.2 *Calculating* LoRE (and expected count of events).

1.2.2 Consequence of a Risk Event (CoRE)

CoRE refers to the impact of an event (in terms of damage and hazard posed to the natural and built environment). The CoRE varies across the region based on simulated fire outcomes using detailed fuels, weather, and topography data. There is only one CoRE value for each 100m x 100m location along the grid (a.k.a. grid pixel) and the CoRE values are highest under location-specific conditions that simulate destructive fire outcomes (e.g., CoRE is generally higher at locations that are typically dry, windy, and have lots of burnable fuel). CoRE is described further under *Section Calculating CoRE*.

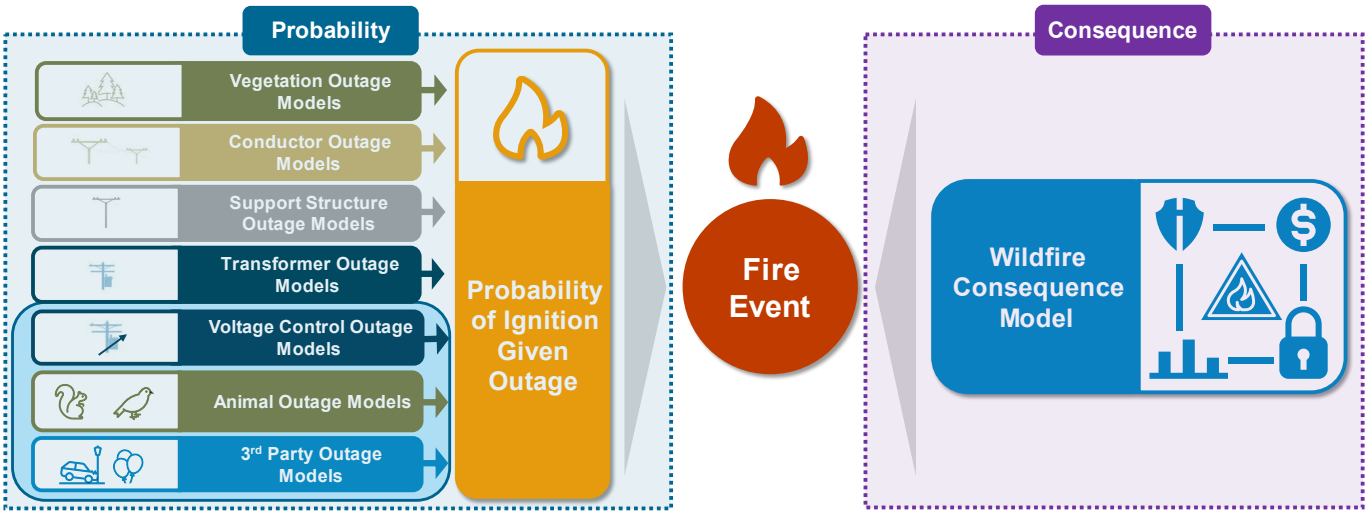


Figure 5 – Risk Approach Overview

1.3 Annual Program Planning for PG&E's Mitigation Programs

The WDRM primary purpose is to inform various PG&E wildfire risk mitigation plans, also known as Work Plans. Work plans are produced, then executed, by various programs, including:

- Enhanced Vegetation Management (EVM) – performs expanded vegetation trimming and removal in selected locations
- System Hardening – replaces existing assets with more resilient versions (e.g., stronger poles and wires) or moves assets to lower-risk locations (e.g., underground, different routes, etc.)
- Pole Replacement – replaces old and worn wood distribution poles
- Transformer Replacement – replaces old and overloaded distribution transformers
- Other inspection and repair programs – typically focused on finding and quickly fixing various overhead distribution asset issues

The output of the WDRM is a spatial map with several layers – each characterizing risk from different causes or assets. These layers of risk can be examined and compared individually, or they can be composited together to understand the full risk from overhead distribution lines at a particular location or asset.

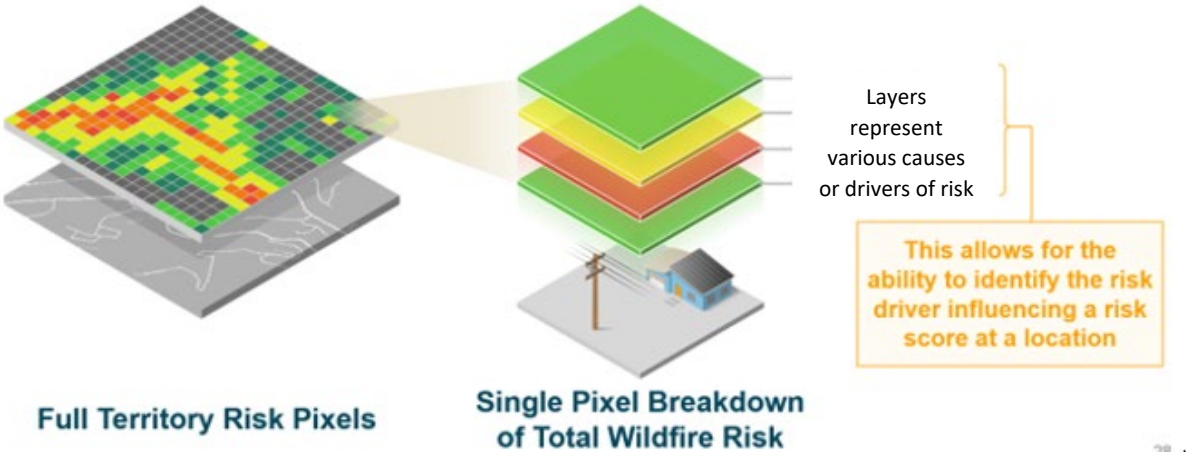


Figure 6 – Compositing of Risk Overview

Note that the risk varies across the geographic area depending on the characteristics of the environment, the weather, the vegetation, and the assets themselves. In Section 3, the RaDA team describes the modeling of each subset (or layer) and the compositing together of these layers to help inform PG&E's mitigation efforts.

1.4 Documentation Structure

The intent of this document is to lay out the RaDA team's approach to the third version (v3) of PG&E's WDRM. Source data, modeling approach, and resulting predictions are described in detail. This document also includes a description of the RaDA team's approach for predicting the *potential reduction of ris*k from mitigation efforts targeting specific categories of outages and ignitions.

While this document seeks to provide a lay reader with a comprehensive understanding of the WDRM, it is not intended to provide detailed mathematical or scientific description of all methods in use. Furthermore, while the application of WDRM predictions to work plan development is summarized, detailed description of how work plans are developed and executed is outside the scope of this document.

With this document:

- Section 2 provides a description of WDRM v3 improvements
- Section 3 presents the WDRM v3 development approach to risk prediction
- Section 4 discusses WDRM v3 predictive results and validation process
- Section 5 provides a schedule of future WDRM improvements
- Sections 6, 7, and 8 are appendices providing a more detailed discussion of covariate selection and tuning, special model topics, and methodological details

2 WDRM v3 Improvements

2.1 WDRM Version Evolution

Mitigating wildfire risk posed by utility assets is a rapidly evolving area of practice where improvements are achieved through the adoption of new data and methods. As such, PG&E's WDRM has evolved – and improved – over time.

The evolution and improvement of PG&E's WDRM from the 2019 WDRM v1 through the 2022 WDRM v3 is outlined in Figure below:

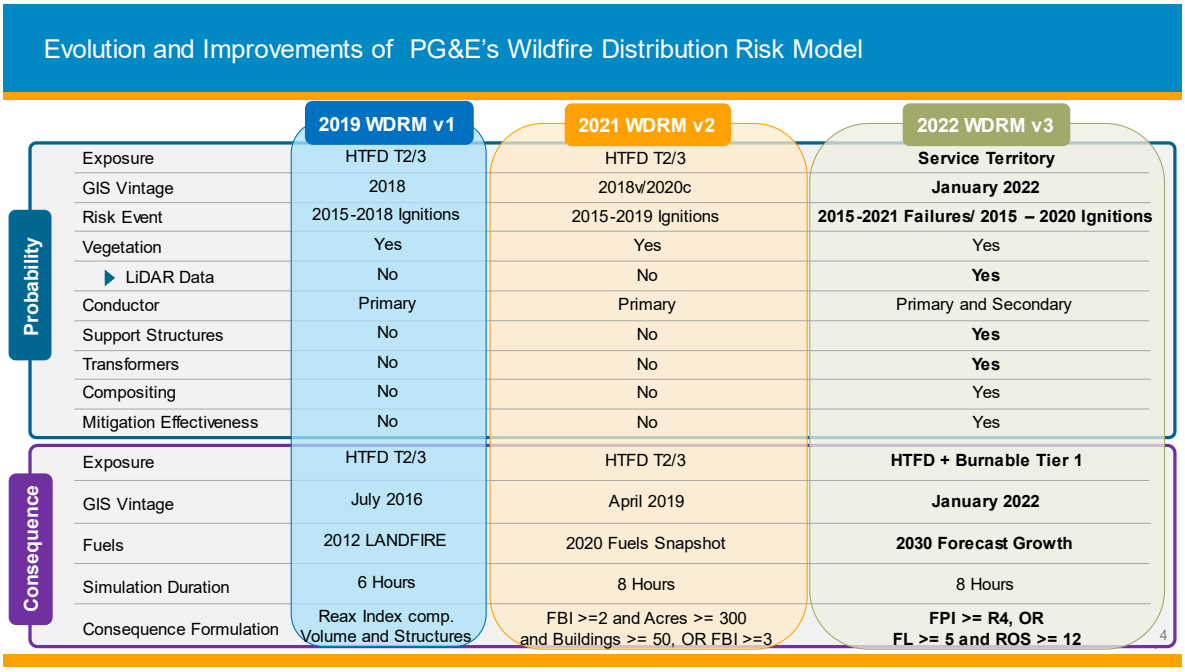


Figure 7 – WDRM Evolution and Improvement

Version 1 (v1) – 2017 Initial RAMP filing under the S-MAP agreement

The first-generation RAMP model for Enterprise Risk Management used probabilistic modeling. It was built in an Excel format (with an add-in feature called @Risk) and relied on Monte Carlo simulation to produce a baseline and mitigated Multi Attribute Risk Score (MARS[1]) and Risk Spend Efficiency[2] (RSE). The first version of the model used in the 2019 WMP predicted wildfire risk on a per-circuit basis and circuit segment basis, for High Fire Threat District (HFTD) only.

In 2019, PG&E started to prioritize circuits and circuit segments for wildfire risk mitigation using a probability of ignition from this v1 model, and wildfire consequence predictions from fire modeling software vendor REAX.

Mitigation work planned with the benefit of the v1 WDRM was performed during 2019, 2020, and 2021 depending on the work plan.

[1] The pre-cursor to multi-attribute value function (MAVF).
[2] Risk reduction divided by the cost of mitigation.

Version 2 (v2) – 2020-2021 Second RAMP filing in accordance with the S-MAP Settlement Agreement (Dec 2018)

The second generation of the WDRM took a meaningful step forward by using more advanced modeling, examining more sub-drivers with regards to ignitions, and using PG&E's version of the CPUC's multi-attribute value function (MAVF) to predict wildfire consequences, as required by the S-MAP Settlement Agreement. During the second year of the filing, the v2 model was improved to use more advanced algorithms and machine learning.

PG&E also improved the modeling to be more granular spatially and by risk driver. The spatial resolution of the model was improved from circuits in v1 to 100m x 100m areas (pixels) within HFTD in v2. In addition, the RaDA team modeled vegetation-caused ignitions separately from other ignitions along conductor lines to understand these two risk drivers. This resulted in two related but separate models (i.e., vegetation-caused and conductor-involved), each of which was used by a different wildfire risk mitigation program (EVM and System Hardening, respectively).

The second version was also upgraded to utilize physics-based fire simulation outputs provided by vendor Tecnosylva, mapped into MAVF fire size/severity tranches based on their simulated characteristics, to quantify wildfire consequence.

Version 3 (v3) – Modeling in 2021 for use in 2022+

Within the most recent model, which is the focus on this document, PG&E implements numerous improvements motivated by internal review as well as in reaction to feedback from public safety specialists, partners, interveners and a third-party review.

This version utilizes more-advanced machine-learning modeling techniques, incorporates newly available data, adds predictions of wildfire risk reduction when mitigating various sources of risk, expands to understand additional ignition sources and sub-drivers, and more.

This version also models several "causal pathways" to ignitions separately, allowing for the nature of these causes to inform the type of model structure and relevant covariates. This also allows for a more specific mapping between cause categories and the mitigations that address them.

2.2 v3 Improvement Highlights

The key improvements for WDRM v3 include:

- **Updated and expanded outage and ignition training event data**
    - PG&E has a rich set of ignition data, starting in 2015. The v3 model estimates both the probability of outage and probability of ignition from outages using outage and ignition data through 2020 (where outage data includes imputed data from damages incurred when the power was off for PSPS events) to enable more detailed dissection of data by cause and equipment involved without the loss of predictive power.
- **More accurate, covariate data inputs**
    - v3 includes more accurate inputs such as tree data from a combination of LiDAR and satellite data, updated weather data from PG&E's meteorology team, tree species data, predictions from PG&E's pole loading simulation program, etc.
- **An improved wildfire consequence model**
    - The wildfire consequence model received a major upgrade, using improved wildfire simulation data and calibration with historical fire records to provide improved predictions.

  - In line with an E3 validation report recommendation, the planning and operational versions of wildfire consequence were developed from a common core data set and framework. This has improved the coordination between the two models.
  - The fuels layer used in the Tecnosylva fire simulation was updated to replace current fire scars with a 2030 forecast.
- **New subset models to better characterize risks and inform work planning**
  - Version 3 of the model expands from 2 subset models in v2 to 17 subset models in v3. These new subset models are sub-divided based on primary/secondary voltage equipment, causes, and equipment types. It also includes a separation of branch, trunk, and "other" failures for vegetation, and a distinction between equipment failures due to internal causes (e.g., electrical overload) vs. external causes.
- **Expansion of the geographic area**
  - v3 expands beyond Tier 2 and 3 High Fire Threat Districts (HFTD) to predict wildfire consequence, outage probability, and ignition probability across PG&E's entire distribution system.
  - Note that some areas of the distribution system where wildfire ignitions are considered to be essentially impossible (e.g., in dense urban areas, in agricultural areas, in paved and otherwise non-burnable areas) are excluded from the simulations that quantify consequence and represented by a backstop value that is lower than the simulated values.
- **New data frame models applied to equipment failures**
  - Version 3 of the model includes asset-level data frame models of equipment for support structures (i.e., poles and cross arms) and transformers.
  - This model form is new in v3 – in v2, all failures were modeled with the assumption that risk was most correlated with environmental conditions like wind, moisture, etc. In v3, data frame models predict risk that is more typically correlated with the passage of time – like electrical overload posing a risk to transformers, which increases over time as the transformers are overloaded.
  - This was a recommendation from the E3 validation to develop and evaluate a wider range of model algorithms for each model subset.
- **Compositing of subset-specific probabilities and risk scores**
  - Version 3 of the model introduces the *compositing* of risk scores, i.e., separately estimating risks for each subset model corresponding to a specific causal pathway of ignitions, and then aggregating predictions across categories to assign total risks for all locations on the grid.
- **Prediction of risk reduction from program-specific mitigation activities**
  - To better-inform work planning teams seeking to optimize the risk reduction potential of their mitigation activities (like EVM and System Hardening), v3 computes mitigation predictions by systematically applying (i.e., "accurately keeping the books on") Subject Matter Expert (SME) estimates of the effectiveness of various mitigation activities under various circumstances. An estimate of risk reduction potential for each mitigation category is computed for all locations/assets in the distribution system.
  - This quantification of mitigation potential provides insights into their likely efficacy in different locations and facilitates tradeoffs between options.
  - This information feeds into a planning process whose goal is to focus work plans on locations/assets with highest risk reduction potential for a given mitigation option.

- **Automation of the modeling framework to ensure consistency and flexible configuration allowing for additional insights**
    - A modular configuration and automation framework is implemented to support the composite modeling of v3.
    - The RaDA team prioritized repeatable modeling runs driven by configuration of modular components and producing summary and detailed data products for every step in a run.
    - The composite modeling framework is implemented in Python and run in an AWS cloud environment.

2.3 Key Performance Metrics

The changes summarized above have improved the WDRM's ability to predict (and thus reduce) risk. The RaDA team uses a Receiver Operating Characteristic (ROC) curve and its Area Under the Curve (AUC) score as the key metrics of probability model performance. The team also uses a Top 20% concentration factor metric as a secondary metric focused on the forecasted yield of true positives in the top 20% of a model prioritization.

2.3.1 ROC and AUC Curves

The ROC curve is based on sorting the predicted probability (i.e., of experiencing an outage or ignition during a typical/future fire season) from highest to lowest for all locations/assets and then checking whether each asset/location, in order of probability, actually experienced an event from test data that is outside the event data used to train the model, (i.e., 25% of all outages or ignitions, depending on the model, randomly selected to be withheld from the training data for the purposes of testing model prediction accuracy). For every location or asset that experienced an event (true positive), the value of the y-axis is increased by one; and for everyone where an event did not occur (false positive) the value of the x-axis increased by one. This continues until all locations are evaluated (and therefore all events are accounted for). The x- and y-axis are typically normalized to the 0-1 fraction of their total counts.

In its normalized form, a straight diagonal line ROC has an area under the curve (AUC) of 0.5 and corresponds to a sort order that is no better than random guessing. The steeper the beginning of the ROC and the flatter the end, the more the sort order has concentrated to true positive predictions in the upper ranks. The highest possible AUC would correspond to a sort order that is 100% locations/assets that experienced an event until all have been accounted for and then all those that did not. This corresponds to an AUC of 1.0 (straight up from 0 to 1 – an unbroken streak of all true positives – and then straight over to 1). Across the various model components, the current AUC scores ranged from 0.65 to 0.85, thus indicating good to strong predictions.

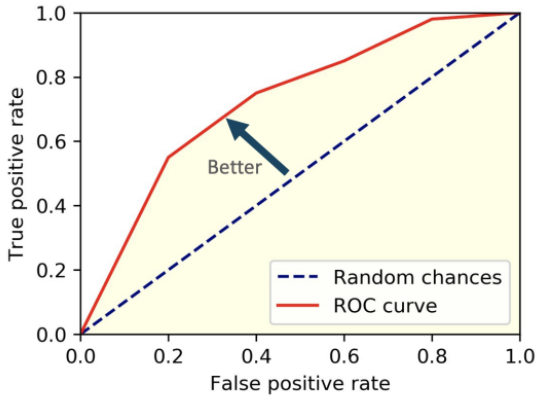


Figure 8 – Example ROC Curve with AUC Shaded

The interpretation of the ROC curves can be thought of in terms of the fraction of non-ignition locations you would need to affect (x-axis) to ensure that you affect some fraction of ignition locations (y-axis), thus avoiding an ignition (or outage for an outage model). The steeper the curve, the less likely that work will be performed that doesn't help avoid ignitions.

2.3.2 Top 20% Concentration Factor

This metric has been created by the RADA team to provide an improved insight on the anticipated yield of true positives in the first 20% of a model prioritization to the work plan development teams. As most work plans are prioritizing the top 80% of risk in the top 20% of locations this is a measure of how effective the model prioritization will be in directing work to locations where events will occur. Similar to the ROC-AUC this metric is developed in reference to a random selection of locations. As shown in figure 8, the Top 20% Concentration factor provides a measure of the predictive performance on the first portion of the ROC-AUC.

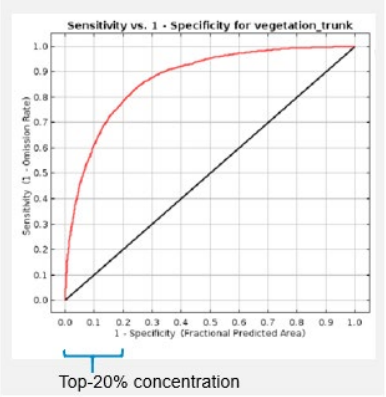


Figure 9 – Top 20% Concentration Factor Metric Example

2.3.3 v3 Performance

The full dataset that trained the overall model (i.e., "all modeled subsets") achieved an AUC score of 0.68. An AUC score of 0.68 is interpreted as a 68% chance that the model can distinguish between where ignitions are and are not likely to occur. This is for a much more extensive model than presented in v2.

While the added models to the System Hardening Composite makes a direct comparison between v2 and v3 difficult, the closest comparison is a comparison of the vegetation models. In v2 of the WDRM, the vegetation model had an overall AUC score of 0.73, while this year's model (which models trunks, branches, and other separately) improves over the v2 performance, with an AUC score for the vegetation composite of 0.78. While the AUC for the System Hardening composite drops, the v2 model focused solely on the conductor whereas the v3 considers several additional risk drivers while maintaining a satisfactory predictive capability.

The Top 20% Concentration Factor metric presents a similar pattern with the System Hardening still maintaining a factor greater than 2 and the vegetation composite improving to a concentration factor of 2.8.

Figure 10 – WDRM v2 vs. v3 Predictive Performance for System Hardening and Vegetation

Throughout this document, the AUC scores are presented as a reflection of each individual model's performance, as well as for the overall composites being used. The performance metrics for each sub-model (e.g., AUC for the ignition probability output predicting ignitions and the Top 20% Concentration Factor for the Ignition Probability modeling efforts) are presented along with the mean annual wildfire season ignitions. This table is ranked from highest to lowest AUC.

Table 1 – WDRM Model Subsets Ranked by AUC

subset_name	mean annual wildfire season ignitions	p(i) pred i AUC	p(i) pred i avg. precision	p(i) top 20% Concentration Factor
animal_squirrel	5.7	0.889	0.0011	4.00
support_structure_equipment_electrical	83	0.868	0.0047	3.90
vegetation_trunk	47	0.851	0.0016	3.69
voltage_control_equipment_type	14	0.857	0.0012	3.69
third_party_balloon	15	0.813	0.0009	3.39
vegetation_branch	58	0.775	0.0009	2.77
vegetation_other	26	0.755	0.0006	2.62
animal_other	15	0.742	0.0009	2.53
third_party_vehicle	38	0.725	0.0009	2.41
animal_bird	31	0.703	0.0003	2.36
primary_conductor	139	0.702	0.0023	2.33
other_equipment_type	45	0.670	0.0005	2.32
support_structure_equipment_cause	28	0.664	0.0000	2.13
secondary_conductor	31	0.663	0.0014	2.05
third_party_other	15	0.636	0.0010	1.87
transformer_equipment_cause	9	0.541	0.0000	1.13
transformer_equipment_leaking	0	nan	nan	nan

3 WDRM v3 Risk Prediction

3.1 Risk Model Overview

The two components of risk in the WDRM v3 are: (1) the Likelihood of a Risk Event (LoRE) expressed as the expected count of risk events per fire season and (2) the Consequence of a Risk Event (CoRE) given its location, where Risk = LoRE x CoRE. An overview of the methods described in this section is shown in the figure below. The risk events modeled by the WDRM are distribution-caused ignitions, with the probability of ignition, also known as LoRE, estimated separately for events grouped by various combinations of cause, sub-cause and equipment involved. Within each grouped subset of all events, the probability of outage is estimated first, with the result multiplied by the probability of an ignition given an outage to produce an estimate of LoRE. Subset-level LoRE values can then be aggregated across subsets and multiplied by CoRE to compute risk across causes and equipment types. Subset-level LoRE and risk values can also be used to support more granular mitigation planning where mitigation is understood to disrupt specific causes or protect specific equipment types.

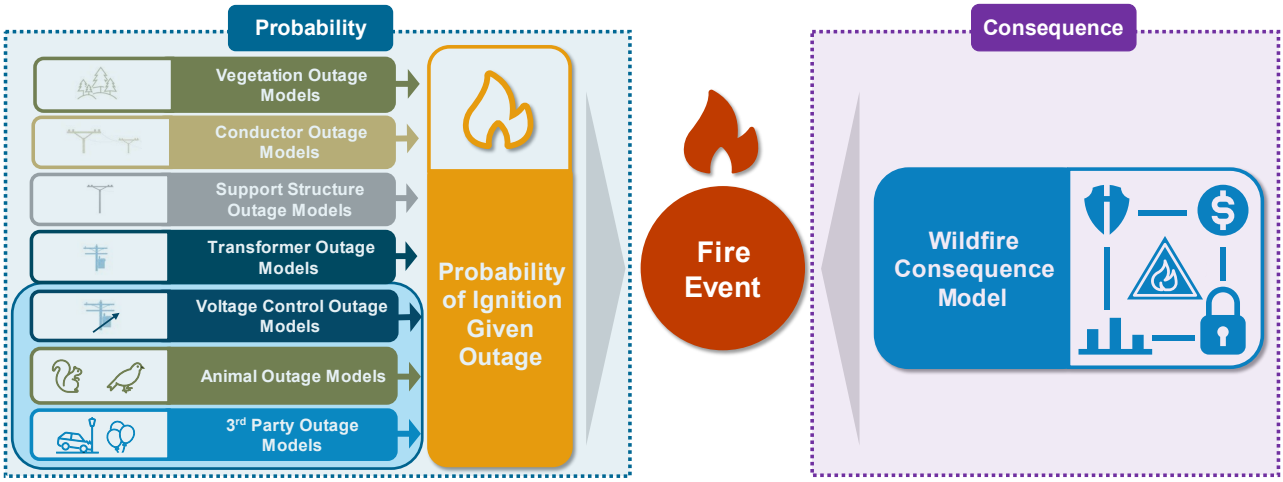


Figure 11 – WDRM Overview

3.1.1 Key Modeling Constructs – Modeling Categories and Subsets

For the purposes of this document, there are 3 WDRM modeling categories and 17 non-overlapping event subsets as shown in the table below. The 17 subsets are defined by categorical attributes of the events. The three Modeling Categories are higher level groupings of subsets and are used in this document to present high-level concepts, examples, and results without needing to document the details of all 17 subsets separately.

Table 2. WDRM Modeling Categories

Modeling Category	Category Subsets

Object Contact	• Risk caused by contact from foreign objects – 6 subsets
	  ○ Animal (squirrel)
	  ○ Animal (bird)
	  ○ Animal (other)
	  ○ Third party balloons
	  ○ Third party vehicles
	  ○ Third party other
	• Risk caused by vegetation (across all assets) – 3 subsets of outages
	  ○ Trunk-failure-caused – for all equipment types
	  ○ Branch-failure caused – for all equipment types
	  ○ Other types of vegetation failure caused – for all equipment types

Conductors & Other Equipment	• Risk caused by electrical or mechanical failures of equipment such as conductors, interrupters, or other equipment such as switches and fuses – 4 subsets
	  ○ Primary conductor
	  ○ Secondary conductor
	  ○ Voltage Control equipment type
	  ○ Other Equipment type

Support Structures & Transformer	• Risk caused by mechanical failures of non-electrical equipment such as poles and other components of support structures and electrical failure of transformers – 4 subsets
	  ○ Support structure– equipment electrical failure caused
	  ○ Support structure – equipment caused
	  ○ Transformer failure – equipment caused
	  ○ Transformer failure – equipment leaking caused

3.1.2 Risk Calculation Process

The one-page call out box below provides a summary of the model. Each of the bold statements is described in more detail within the sub-sections below, including:

- LoRE – Probability of Ignition
- CoRE – Consequence Value
- Wildfire Risk Scores

The high-level calculation process for LoRE, or probability of ignition, is:

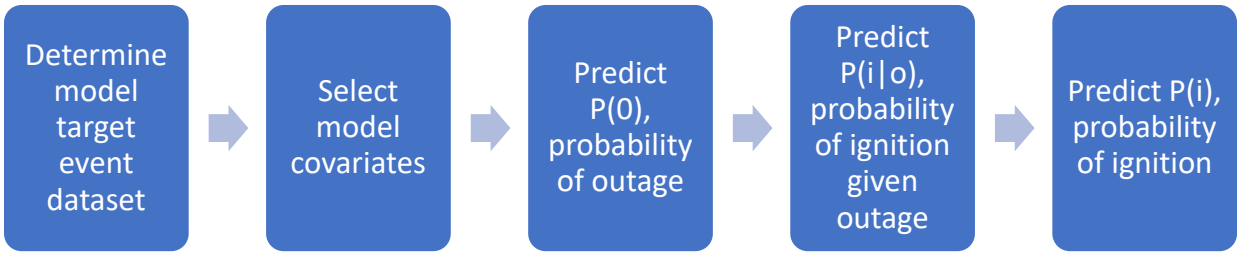


Figure 12 – LoRE Calculation Process

The high-level calculation process for CoRE, or Consequence value, is:

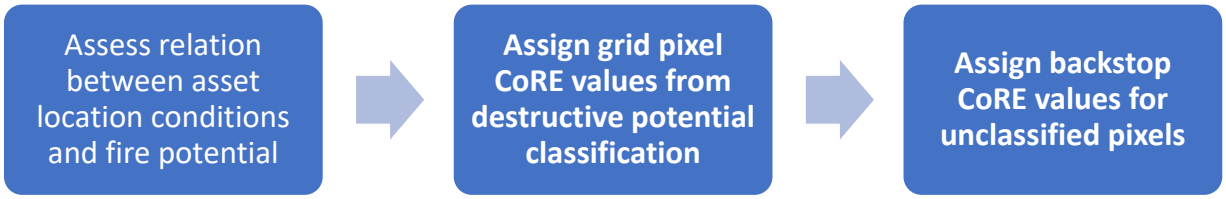


Figure 13 – CoRE Calculation Process

Using the results from LoRE and CoRE, the Wildfire Risk Score calculation process is:

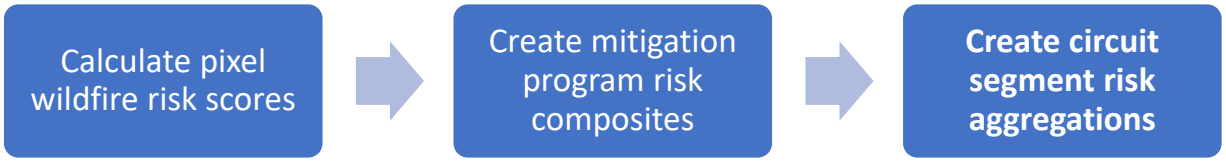


Figure 14 – Wildfire Risk Score Calculation Process

3.2 Model Development Process

The WDRM v3 is implemented primarily in Python, using a sequence of interconnected configurable computation tasks to complete a model run, store its results, and report out metrics of model performance. The key modeling steps are described in the following table.

Table 3 – WDRM Model Development Process

WDRM Model Development Process	
1	Define of the target set, which is the set of all events of type outage, ignition, and PSPS hazards and damages that are used to train the sub-models that comprise the WDRM v3. The target set is limited to events that occurred during the fire season months of June through November between 2015 and 2021 and is filtered to exclude events involving underground equipment, where outages were caused by wildfire or planned by the company, and the small portion of events without valid locations.
2	Identify subsets of the target set that share cause, sub-cause, and type of equipment involved. The subset manager (SSM) divides the target set into 17 non-overlapping subsets that span the target set.
3	Prepare the model covariate data.
4	Train the probability of ignition given an outage model using target set data.
5	Predict P(outage) and P(ignition|outage) for each subset, yielding estimates of P(ignition).
6	Determine the mitigation potential from various measures applied to each subset's risk.
7	Composite probabilities and risks from individual subsets into broader categories used for planning purposes.

Unless otherwise noted, 25% of training data is withheld from each model fit to be used to compute out of sample predictive performance metrics.

3.3 LoRE – Probability of Ignition

The sections below describe the data sets for models, the inputs (or covariates) to the LoRE models, and how the WDRM v3 estimates the probability of an ignition through two modeling steps, in which the probability of an outage for all assets or grid locations for each subset of outages and the probability an ignition is associated with an outage, given its characteristics, where P(ignition) = P(ignition|outage) x P(outage).

Throughout this description of calculating LoRE, this document refers to two different types of outage probability models: MaxEnt and Asset Attributes models. One or the other of these approaches are applied to all 17 subsets of outages. For subsets with outages driven by environmental determinants, such as vegetation caused outages, the WDRM v3 employs a MaxEnt model structure, with primarily spatially varying covariates resulting in grid pixel level estimates of P(outage). For modeling categories that relate to equipment failures due to internal attributes, such as transformers and support structures, the WDRM v3 employs Asset Attribute models fit via Random Forest to one row of data per-asset-year.

A third model type, Logistic Regression, is used to estimate the probability of ignitions associated with outages, given outage characteristics.

3.3.1 Model Target Event Dataset

The WDRM v3 draws on approximately 114,000 events in the target event dataset.[3] The definition of the events used, the target set, the source of this information, and the sub-setting of this information are described below.

For WDRM v3, the event data includes three types of events, which come from three different data sets.

- Outages and Forced Outages
    - Source: PG&E's Integrated Logging Information System, ILIS.
    - Outages are when electricity ceases to be delivered to customers. Detecting outages is done electronically and is automatically recorded.
- Hazards and Damages
    - Source: Post-PSPS Inspection Data.
    - These are issues classified as potential hazards or equipment damages identified during the inspection of de-energized equipment before power can be restored after a PSPS event.
- Ignitions
    - Source: PG&E's Historical Ignitions Data (2,500 CPUC reportable ignitions 2015-2021; 1,900 non-reportable ignitions).
    - This is the raw data behind CPUC reportable ignitions limited to reportable fire events that meet the following criteria:
        - A self-propagating fire of material other than electrical and/or communication facilities.
        - The resulting fire traveled greater than one linear meter from the ignition point.
        - The utility has knowledge that the fire occurred.
        - That caused damage to utility facilities and whose ignition is not associated with utility facilities are excluded from this reporting requirement (CPUC, 2014).

      Note that the ignition data set includes both CPUC reportable and non-reportable ignitions, occurring with or without an outage. Fires that caused damage to utility facilities and whose ignition is not associated with utility facilities were excluded.

[3] Note that the numbers do not include events categorized as "wildfire mitigation" or "underground events."

Collectively, the three types of events are described as failures. Failures refer to incidents where damage to the grid has occurred, or damage to the environment has occurred due to grid equipment operation even if no outage occurs. The failures data includes events that come from:

- Within the boundaries of PG&E's overhead distribution lines only
- 2015 through 2021
- Fire season only (June through November)

Not included in the target failure dataset are:

- Outages directly caused by wildfires
- Outages or ignitions caused by underground equipment

Events from within the fire season are selected to avoid training the model on events from causes that are not viable during the fire season, such as iced lines, snow loading, water damage, and water facilitated outages. Including such events would run the risk of training the WDRM to estimate wildfire risk in cases where there is none.

Categorical attributes of target set events, including primary vs. secondary voltage, cause, sub-cause, and the type of equipment involved, are used to define 17 non-overlapping subsets in the WDRM v3 are summarized in the following table.

Table 4 – Model Subset Characteristics

Subset	Voltage Category	Equipment Type	Cause	Sub-cause	Modeling Category	Model Type
vegetation_other	Any	Any	Vegetation	Other	Object Contact	MaxEnt
primary_conductor	Primary	conductor	Any	NA	Equipment	MaxEnt
vegetation_branch	Any	Any	Vegetation	Branch	Object Contact	MaxEnt
vegetation_trunk	Any	Any	Vegetation	Trunk	Object Contact	MaxEnt
animal_bird	Any	Any	Animal	Bird	Object Contact	MaxEnt
secondary_conductor	Secondary	Conductor	Any	NA	Equipment	MaxEnt
other_equipment_type	Any	Other	Any	NA	Equipment	MaxEnt
third_party_balloon	Any	Any	Third party	Balloon	Object Contact	MaxEnt
third_party_other	Any	Any	Third party	Other	Object Contact	MaxEnt
third_party_vehicle	Any	Any	Third party	Vehicle	Object Contact	MaxEnt
animal_squirrel	Any	Any	Animal	Squirrel	Object Contact	MaxEnt
voltage_control equipment_type	Any	Voltage Control	Any	NA	Equipment	MaxEnt
animal_other	Any	Any	Animal	Other	Object Contact	MaxEnt
support_structure equipment_cause	Any	Support Structure	Equipment	Structural	Support Structure/ Transformer	Asset Attribute
support_structure equipment_electrical	Any	Support Structure	Equipment	Electrical	Support Structure/ Transformer	Asset Attribute
transformer equipment_leaking	Any	Transformer	Equipment	Leaking	Support Structure/ Transformer	Asset Attribute
transformer equipment_cause	Any	Transformer	Equipment	Failure	Support Structure/ Transformer	Asset Attribute

The event counts, ignition counts, and ignitions per-outage rates for all 17 subsets are shown in the table below, sorted from highest to lowest event count. Note the substantial variation in ignition counts and ignitions per outage, demonstrating that different causal pathways leading to failures vary in their likelihood to cause ignitions.

Table 5 – WDRM v3 Target Dataset

Subset	Event Count	Ignition Count	Ignition per Outage
other_equipment_type	46,981	316	0.67%
primary_conductor	12,343	974	7.89%
transformer_equipment_cause	8,809	62	0.70%
third_party_vehicle	6,952	265	3.81%
vegetation_branch	6,912	406	5.87%
animal_bird	4,831	219	4.53%
support_structure_equipment_cause	4,631	194	4.19%
vegetation_trunk	4,388	329	7.50%
secondary_conductor	3,801	216	5.68%
animal_squirrel	3,694	40	1.08%
third_party_other	2,202	102	4.63%
third_party_balloon	2,127	103	4.84%
support_structure_equipment_electrical	2,096	582	27.77%
vegetation_other	1,655	184	11.12%
transformer_equipment_leaking	1,126	0	0.00%
animal_other	834	106	12.71%
voltage_control_equipment_type	502	99	19.72%
Totals	**113,884**	**4,197**	**3.69%**

Within the subsets, the data is filtered to look for complete cases, that is, cases that include all the event and covariate data needed to train the model – the remaining events provide the training event count. The asset attribute models withhold test event data from their official model runs, whereas the MaxEnt models train on all available data, with train/test splits supporting performance metrics computed as separate diagnostic model runs – for this reason there are test events withheld from the official MaxEnt fits and their corresponding test event counts are not applicable.

Subsets modelled using MaxEnt don't impute values missing from their raster-based covariates and drop some training events due to missing data. It is common for geo-tif data to be masked out over bodies of water or where the data set in question is otherwise deemed not applicable, but differences in the pixel size and projection between source tifs and grid pixels can cause some grid pixels to lack some covariate data. Subsets modelled using asset attributes impute missing values and thus do not drop any events.

The following table summarizes event training, test, and dropped counts for each of the model subsets.

Table 6 – Training Event Counts

Subset	Training Event Count	Test Event Count	Dropped Event Count
other_equipment_type	43,356	-	3,625
primary_conductor	11,100	-	1,243
transformer_equipment_cause	5,102	4,751	0
third_party_vehicle	6,549	-	403
vegetation_branch	6,640	-	272
animal_bird	4,489	-	342
support_structure_equipment_cause	1,363	518	0
vegetation_trunk	4,149	-	239
secondary_conductor	2,493	-	1,308
animal_squirrel	3,474	-	220
third_party_other	1,883	-	319
third_party_balloon	1,959	-	168
support_structure_equipment_electrical	854	284	0
vegetation_other	1,559	-	96
transformer_equipment_leaking	626	599	0
animal_other	742	-	92
voltage_control_equipment_type	445	-	57

The MaxEnt models and support structure asset attribute models were trained on 75% train and 25% test samples of subset specific events. The transformer asset attribute models used a monthly train test split described in 8.2.2.4 Model train, test, predict dataset splits.

3.3.2 Model Covariates

Model covariates refer to the data used as explanatory variables in the formulation of the models. The model covariates, also referred to as model features, become model inputs. Within subset models, the objective is to identify which covariate environmental conditions and asset attribute values are most common among ignition locations along the distribution grid. The WDRM uses these contributing covariates to create predictive models of probability for future ignitions across the grid. Covariates selected for modeling must have values available for every location at which an ignition probability prediction is to be made.

Model covariate descriptions, data sources, and model use can be found in Appendix: Pool of Covariates. Model covariates come from three general categories:

- Environmental, which are spatially dependent and have location-specific characteristics.
    - HFTD tier
    - Prevailing climate
    - Estimated LiDAR satellite strike tree counts
    - Tree species
    - Vegetation types
    - Topography
    - Soil characteristics
- Weather-dependent, influencing equipment failure rates and ignition viability.
    - Wind speed
    - Temperature
    - Fuel moisture
    - Precipitation totals
- Asset characteristics, used to consider physical states that degrade over time due to operation and environmental factors. Typically, an asset attribute state can be modified through mitigation.
    - Voltage type (primary or secondary)
    - Support structure age
    - Support structure height
    - Support structure circumference
    - Open maintenance tags
    - Structural loading estimate
    - Transformer age
    - Transformer manufacturer
    - Transformer electrical loading characteristics

The following table summarizes the environmental and weather-dependent data used by the WDRM.

Table 7 – Environmental and Weather Covariate Data Source Summary

Data Set	Vintage	Date Span	Description
CWHR	2014	n/a	California Wildlife Habitat Relationships (CWHR) version 9.0 - produced by California Department of Fish and Wildlife.
USSOILS	1995	n/a	USSOILS is an Arc 7.0 coverage containing hydrology-relevant information for 10,498 map units covering the entire conterminous United States. The coverage was compiled from individual State coverages contained in the October 1994 State Soil Geographic (STATSGO) database by the United States Geological Survey (USGS).
gridmet	2019	1979-2019	University of Idaho Gridded Surface Meteorological Dataset provides high spatial resolution (~4-km) daily surface fields of temperature, wind, precipitation, humidity, and radiation across the United States.
LANDFIRE	2016	n/a	LANDFIRE surface fuels model produced by United States Geological Survey (USGS).
NED	n/a	n/a	National Elevation Database (NED) - produced by United States Geological Survey (USGS).
NLCD	2016	n/a	National Land Cover Database (NLCD) - produced by United States Geological Survey (USGS).
PG&E meteorology	2021	2010-2020	POMMS data used as inputs to, or the outputs of, the Fire Potential Index (FPI) model
RTMA	Late 2019	2016-2018	The Real-Time Mesoscale Analysis (RTMA) is a high-spatial and temporal resolution analysis for near-surface weather conditions. Includes hourly analyses at 2.5 km for the contentinental United States - produced by National Oceanic and Atmospheric Administration (NOAA) and National Weather Service (NWS).
treemap	2014	n/a	A tree-level model of conterminous US forests circa produced by imputation of FIA (US Forest Service Inventory and Analysis) plot data.
Vegetation merged	2019	2019	High resolution, territory-wide vegetation data produced by merging PG&E-collected LiDAR data and Salo Sciences-processed satellite data.
WorldPop	2010 - 2020	n/a	WorldPop estimates number of people residing in each 100m grid cell.

3.3.3 Predicting Probability of Outage – P(o)

The determination of wildfire risk for an asset or at a specific grid location starts with predicting the probability of an outage – P(o). This section describes the calculation of the probability of an outage for the 17 model subsets. For simplicity, the modeling approach is described for the three general modeling categories described in the WDRM Modeling Categories table: Object Contact, Conductors & Other Equipment, and Support Structures & Transformers.

Object Contact and Conductors & Other Equipment subsets are modeled with MaxEnt, which uses a spatial modeling approach. Support Structures & Transformers subsets make use of asset attribute models.

The RaDA team then trained and tested the model selected for each subset. For each model, the RaDA team validated the model by computing model covariate importance and reviewing predictive performance metrics for subset models. This was used to diagnose and improve subset models. The type of model, the training data set, notable aspects of the models and the feature importance and AUC scores for each modeling category are described in the following sub-sections.

For each modeled subset, an iterative process was used by the RaDA team data scientists to find an optimal model for predicting P(o).

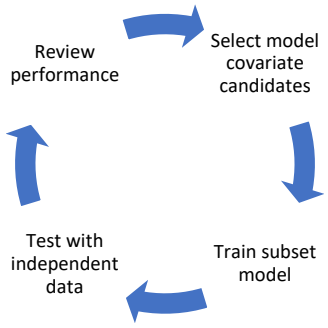


Figure 15 – Covariate Selection Process

3.3.3.1 Object Contact P(o) Models

Grid events due to contact from external objects fall into three causal categories:

- Vegetation
  - Typically branch or trunk failures that strike and/or exert structural loads on grid equipment
  - Other causes, such as vegetation growing into lines
- Animals
  - Faults due to spanning contact, usually birds or squirrels
  - Other causes including nests, snakes, and even slugs
- Third parties
  - Vehicle accidents that damage equipment or poles
  - Metallic balloons that create line shorts/drops
  - Other causes, such as outdoor project accidents, malfunctioning appliances, and gunshot damages

3.3.3.1.1 Vegetation P(o) Models

Vegetation-caused ignitions are caused by the presence of trees whose height and proximity to overhead distribution grid equipment make physical contact possible. The WDRM has three vegetation sub-models:

- Branch
- Trunk
- Other causes

The use of three vegetation subsets provides the Enhanced Vegetation Management (EVM) program team with the information needed to select appropriate mitigation actions to address vegetation risks.

The vegetation model subsets are trained on vegetation interactions with overhead distribution grid equipment. The model subsets predict the probability of vegetation-caused outages involving multiple equipment types, such as conductors, support structures, switches, fuses, transformers, etc. Predictions are made for 100m x 100m geographic pixel locations along the distribution grid infrastructure.

MaxEnt models are used for predicting vegetation-based outages. The MaxEnt model was selected for its ability to predict events that are primarily informed by spatial and environmental factors. The use of MaxEnt for vegetation models for WDRM v3 is consistent with the approach used for WDRM v2, but with one key difference. WDRM v2 used MaxEnt to directly predict the probability of ignition (P(i)), whereas WDRM v3 decouples the P(o) from P(i). See the Predict probability of ignition given outage – P(i|o) section of this document for more details on predicting of probability of ignition from an outage.

Events used for training the vegetation model subsets include outages, PSPS damages, PSPS hazards, and ignitions not known to be associated with an outage. The events were pulled from the fire seasons for 2015 through 2021. Details on events counts for model training can be found in the Training Event Counts table.

WDRM v3 highlights for the modeling the P(o) vegetation subsets:

- Tree mortality and susceptibility to branch drops is partially determined by habitat and biological characteristics, which available for modeling via broad spatial tree species designations. WDRM v3 used the California Wildlife Habitat Relationships (CWHR) 2014 data set as a proxy for tree species. CWHR is a predictive model that lists species predicted to occur at a given location under certain habitat conditions.
- LiDAR data was combined with satellite data to get better estimations of tree counts, tree density, and tree heights to characterize the fall-in trees near the distribution grid.
- Satellite derived tree data was used to supplement LiDAR data for full-service territory coverage, as LiDAR flights only took place only over the HFTD areas.
- Soil type data was introduced as a covariate.
- PG&E's internal meteorology data was adopted for v3 to capture weather conditions, including seasonal precipitation, prevailing fuel moisture, and prevailing winds.

The vegetation model subsets demonstrated strong performance. All three vegetation subsets predicted test event outages with AUCs greater than 0.83, as shown in the ROC plot to the right.

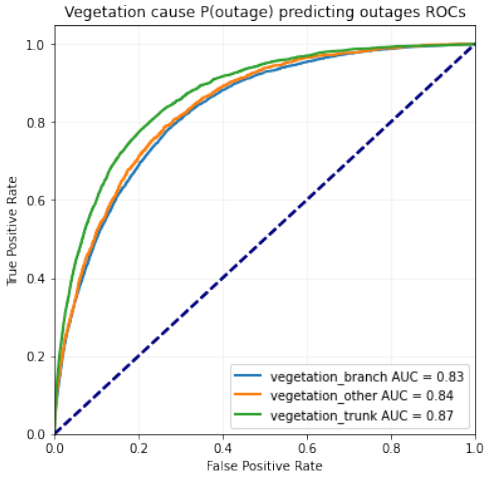


Figure 16 – Vegetation Subsets P(o) ROC Curves

Covariate analysis shows that tree characteristics and land use types tend to be the top performers for vegetation subsets.

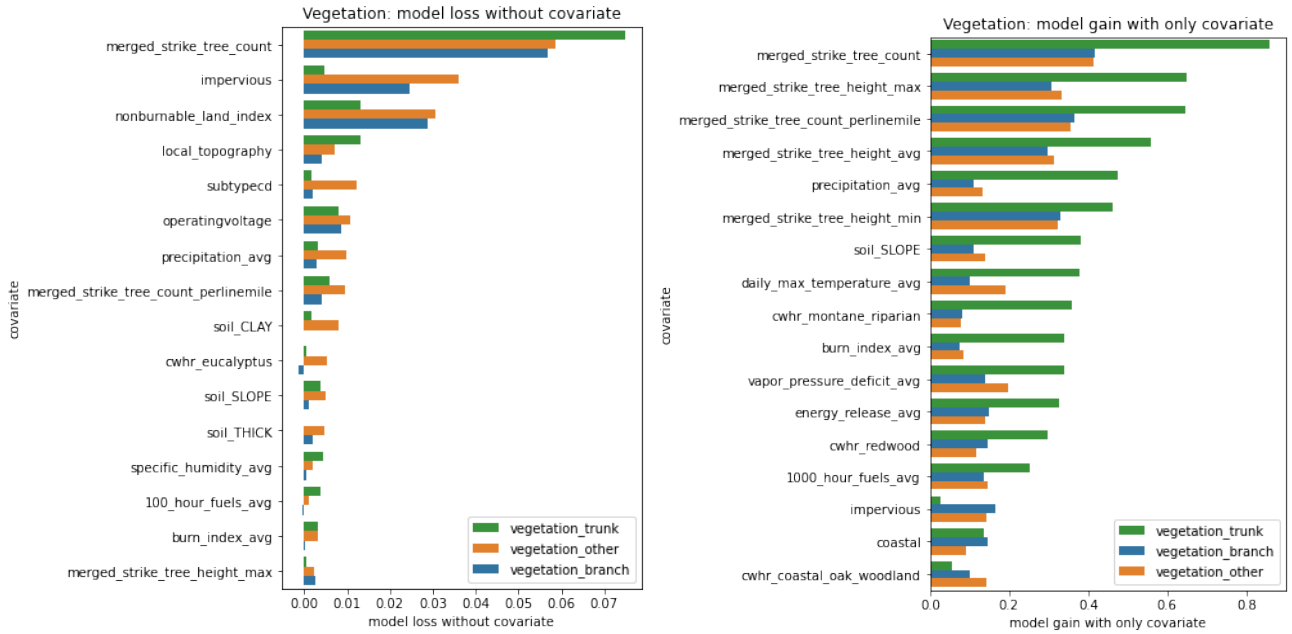


Figure 17 – Vegetation Subsets P(o) Covariate Analysis

3.3.3.1.2 Animal P(o) Models

Animal-related outages are caused by contact spanning phases or a phase to ground. Animal events are driven typically by animals with propensity to climb or perch atop poles and lines. The WDRM has three animal sub-models:

- Bird
- Squirrel
- Other causes

The use of three animal subsets provides mitigation teams with the information needed to select appropriate corrective actions to address animal risks.

The animal model subsets are trained on animal interactions with overhead distribution grid equipment. The model subsets predict the probability of animal-caused outages involving multiple equipment types, such as conductors, support structures, switches, fuses, transformers, etc. Predictions are made for 100m x 100m geographic pixel locations along the distribution grid infrastructure.

MaxEnt models are used for predicting animal-based outages. The MaxEnt model was selected for its ability to predict events that are primarily informed by spatial and environmental factors.

Events used for training the animal model subsets include outages, PSPS damages, PSPS hazards, and ignitions not known to be associated with an outage. The events were pulled from the fire seasons for 2015 through 2021. Details on events counts for model training can be found in the Training Event Counts table.

WDRM v3 highlights for the modeling the P(o) animal subsets:

- Preliminary results from a single animal subset model were reviewed by PG&E Subject Matter Experts (SMEs). On their advice, the original single animal model subset was the divided into three subsets. Subsequent analysis confirmed the predictive value of using three animal model subsets.
- Tree species and location were important covariates for differentiating between squirrel and bird causes for outages.
- Covariates that captured urban versus rural conditions were useful detecting habitat changes that enabled the model to differentiate between various types of animal damage.
- Anecdotally, it was observed that the spatial distribution for bird-related events was much broader than the spatial distribution for squirrels. This validated the creation of the bird and squirrel specific subsets.

While all animal model subsets demonstrated good performance, the squirrel-specific subset produced an exceptional result with an AUC of 0.88. The ROC performance plots for the three animal subsets generated against predicted test event outages are shown to the right.

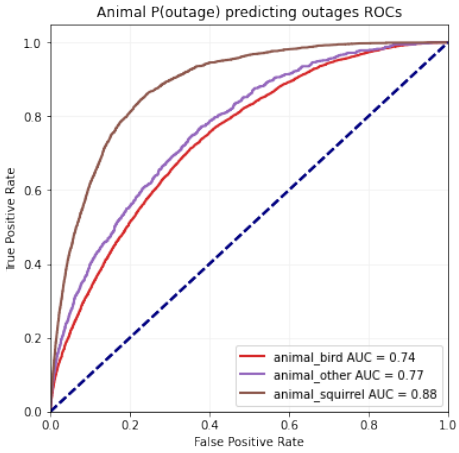


Figure 18 – Animal Subsets P(o) ROC Curves

Covariate analysis shows that tree characteristics, distribution line characteristics, and land use types tend to be the top performers for animal subsets.

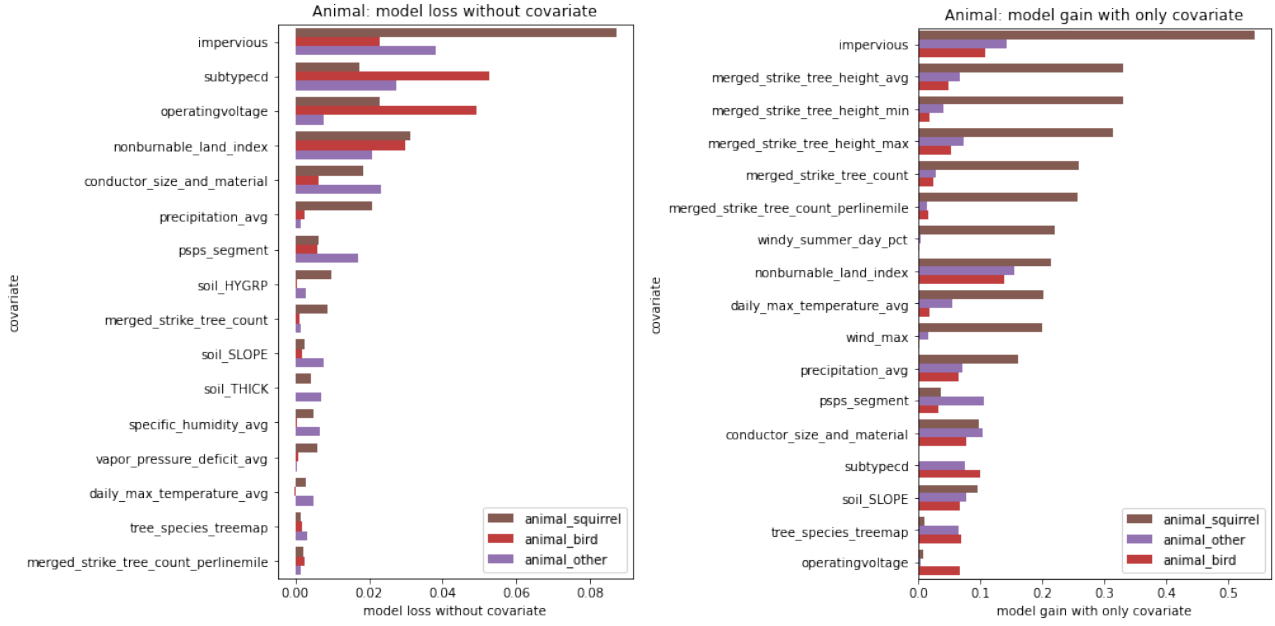


Figure 19 – Animal Subsets P(o) Covariate Analysis

3.3.3.1.3 Third Party P(o) Models

Third party-related outages are caused by human activity external to PG&E operations. The WDRM has three third-party sub-models:

- Vehicle
- Balloon
- Other causes

The use of three third party subsets provide mitigation teams with the information needed to select appropriate corrective actions to address human activity risks.

The third party model subsets are trained on human interactions with overhead distribution grid equipment. The model subsets predict the probability of third party-caused outages involving multiple equipment types, such as conductors, support structures, switches, fuses, transformers, etc. Predictions are made for 100m x 100m geographic pixel locations along the distribution grid infrastructure.

MaxEnt models are used for predicting third party-based outages. The MaxEnt model was selected for its ability to predict events that are primarily informed by spatial and environmental factors.

Events used for training the third party model subsets include outages, PSPS damages, PSPS hazards, and ignitions not known to be associated with an outage. The events were pulled from the fire seasons for 2015 through 2021. Details on events counts for model training can be found in the Training Event Counts table.

WDRM v3 highlights for the modeling the P(o) third party subsets:

- Preliminary results from a single third party subset model were reviewed by PG&E Subject Matter Experts (SMEs). On their advice, the original third party model subset was the divided into three subsets. Subsequent analysis confirmed the predictive value of using three model subsets.
- Line voltage and line characteristics were important covariates for understanding line spacing for balloon-related faults and support structure durability for withstanding vehicle impacts
- Covariates that captured urban versus rural conditions enabled the model to differentiate between high population density (balloons) and low population density (vehicular) events

While all third party model subsets demonstrated good performance, the balloon-specific subset produced an exceptional result with an AUC of 0.89. The ROC performance plots for the three third party subsets generated against predicted test event outages are shown to the right.

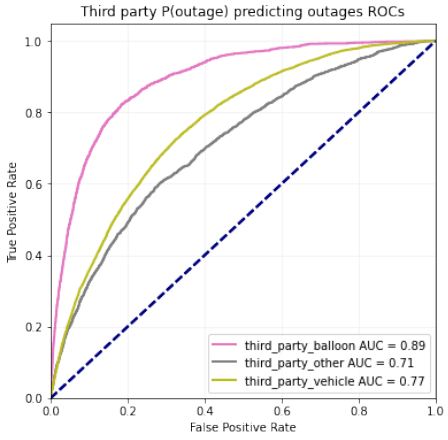


Figure 20 – Third Party Subsets P(o) ROC Curves

Covariate analysis shows that tree characteristics and land use types tend to be the top performers for third party subsets.

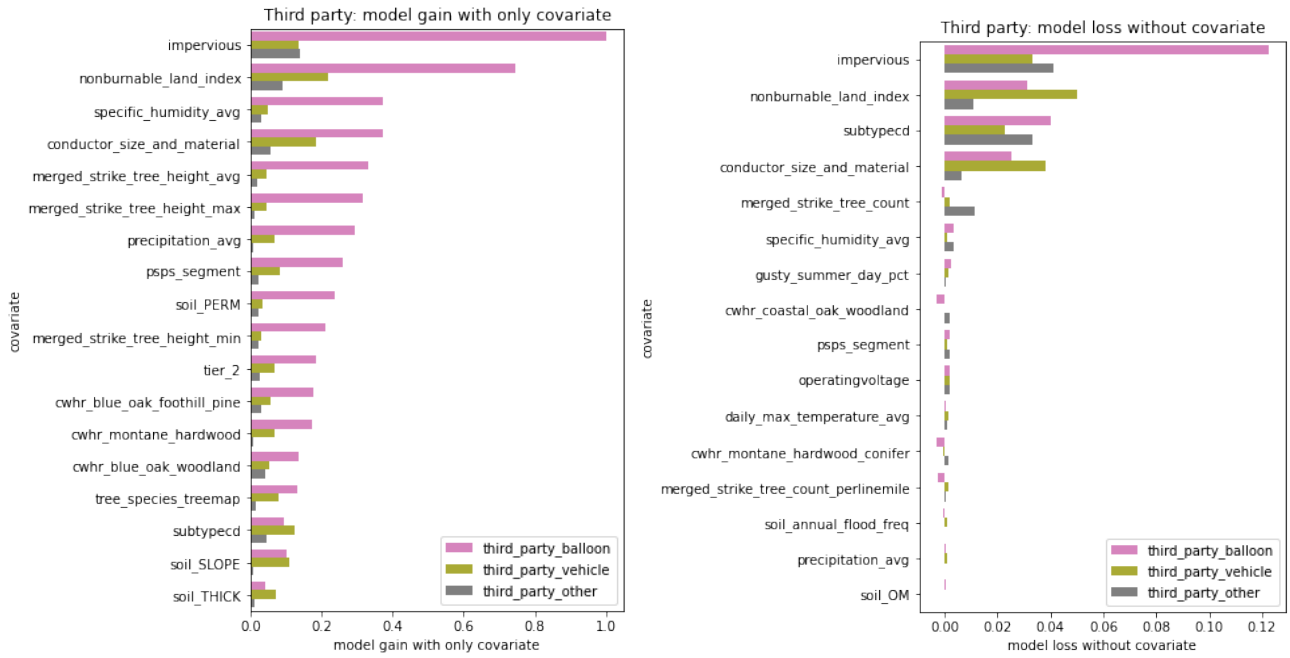


Figure 21 – Third Party Subsets P(o) Covariate Analysis

3.3.3.2 Conductor & Other Equipment P(o) Models

The conductor & other equipment subsets model equipment failures that are environmentally driven, typically due to factors contributing to gradual degradation or extreme conditions that result in premature failure. Four model subsets were developed for conductor & other equipment:

- Primary conductors and interrupters
- Secondary conductors
- Voltage control devices
- Other equipment

The use of four conductor & other equipment model subsets provides mitigation teams with the information needed to select appropriate corrective actions to address equipment risks.

The conductor & other equipment model subsets are trained on historical failures of overhead distribution grid equipment. Predictions are made for 100m x 100m geographic pixel locations along the distribution grid infrastructure.

MaxEnt models are used for predicting conductor & other equipment-based outages. The MaxEnt model was selected for its ability to predict events that are primarily informed by spatial and environmental factors. MaxEnt allows for a good understanding across the distribution system given that the event data is often uncertain regarding exactly which specific piece of equipment failed for any given outage. Therefore, alternative modeling approaches that require direct assignment to specific piece of equipment are not viable.

Events used for training the conductor & other equipment model subsets include outages, PSPS damages, PSPS hazards, and ignitions not known to be associated with an outage. The events were pulled from the fire seasons for 2015 through 2021. Details on events counts for model training can be found in the Training Event Counts table.

WDRM v3 highlights for the modeling the P(o) conductor & other equipment subsets:

- Preliminary modeling produced separate primary and secondary voltage subsets for each of conductors, interrupters, and other equipment types.
    - Analysis of the interrupter failures found that the devices were often damaged by an event but were not the event root cause. Line loading was typically the true cause.
    - Interrupter events were subsequently reclassified with other conductor events
- A relatively small group of highly ignition prone events involving voltage control devices was identified and removed from the other equipment category

While all conductor & other equipment model subsets demonstrated good performance, the voltage control device-specific subset produced an exceptional result with an AUC of 0.88. The ROC performance plots for the four conductor & other equipment subsets generated against predicted test event outages are shown to the right.

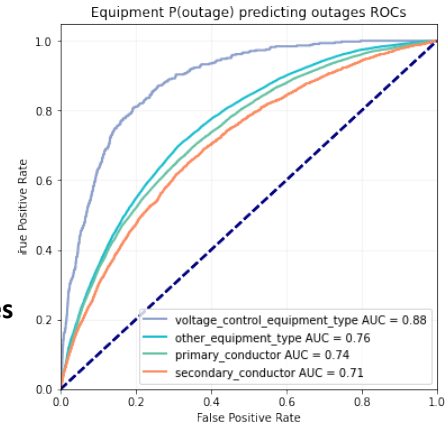


Figure 22 – Conductor and Other Equipment Subsets P(o) ROC Curves

Conductor size and material type, along with voltage, were found to be the most important covariates, especially for voltage control devices.

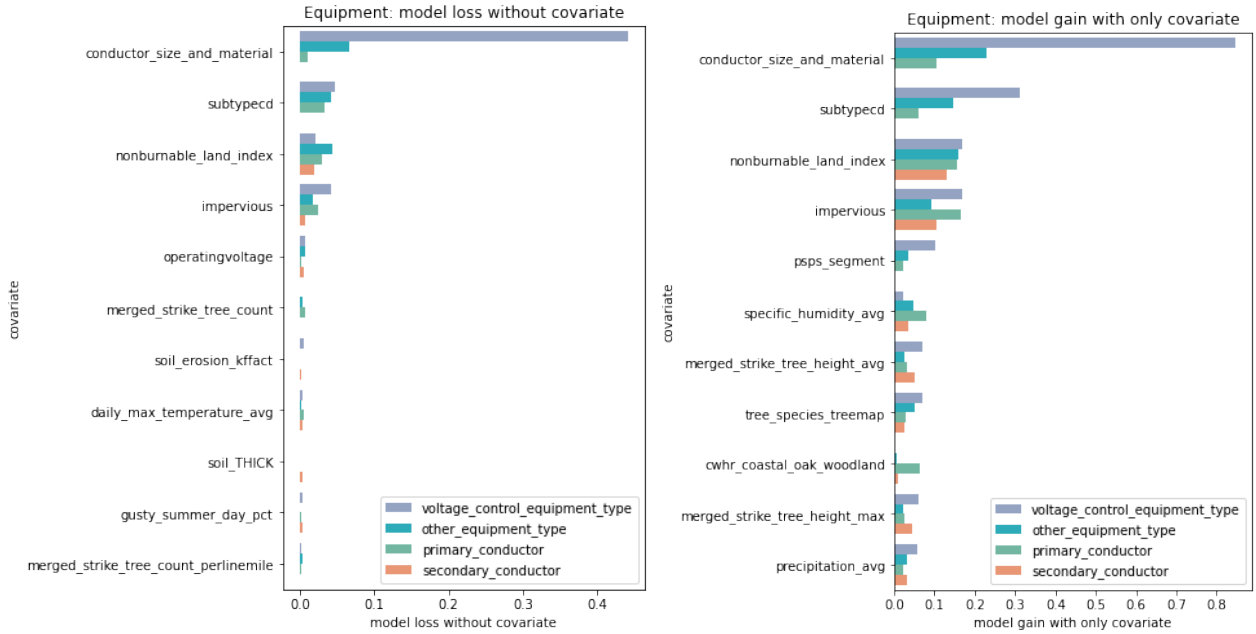


Figure 23 – Conductor & Other Equipment Subsets P(o) Covariate Analysis

3.3.3.3 Support Structures & Transformers P(o) Model

The support structures and transformers group of subsets predict the annual probability of individual support structure or transformer unit equipment failures along the distribution grid. Four model subsets were developed:

- Support structure equipment cause – pole and other structure failures
- Support structure equipment electrical – pole fire or electrical-related outage events
- Transformer equipment cause – transformer overload or failure
- Transformer equipment leaking – leaking transformer

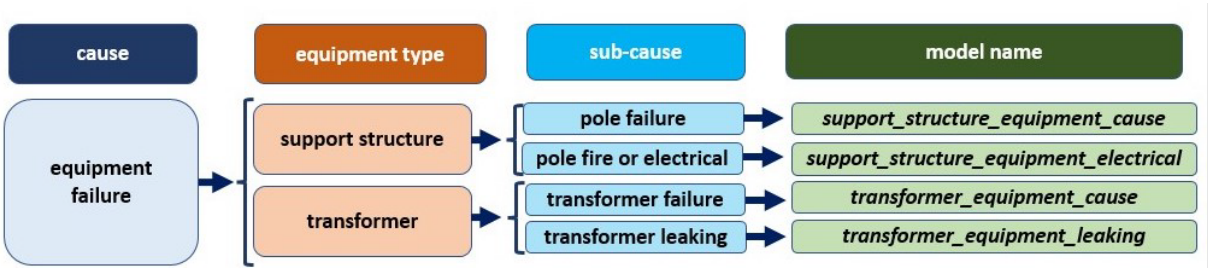


Figure 24 – Support Structures & Transformers P(o) Model Subsets

A support structure includes the components of a pole that support the overhead system:

- Pole
- Crossarm
- Guy or anchor wire
- Insulators
- Supporting hardware and connections

A support structure equipment failure can include the malfunction of one or more of these components. The model predicts the annual probability of an equipment failure per pole or support structure. The support structure model considers only wood distribution poles, limiting predictions to those poles constructed with wood, throughbore, or centerbore material. Each structure is tracked in the PG&E data model by its primary key, the SAP equipment id.

Transformers are nodes on the distribution grid, identified by a California Grid Coordinate (cgc12), and contain one to three transformer units that are identified by an equipment id. When any unit on a transformer fails, typically all individual transformer units are replaced. Thus, while planning focuses on the transformer level, failures occur at the transformer unit. The transformer model subsets therefore predict transformer unit failures.

The support structure and transformer asset attribute models were modeled using a Random Forest Classifier. Random Forest Classifier is a tree-based algorithm that implements a machine learning method called bagging to combine the results of many different decision trees fit to sub-samples of training data into one model result, minimizing the error. The Random Forest Classifier algorithm was identified as the top performing algorithm for the annual asset-level support structure and transformer datasets. Other algorithms explored included: logistic regression, maximum entropy, and survival models.

A support structure or transformer asset is often removed from service after a major equipment failure event. Therefore, it is critical to develop a training data set that can connect the historical conditions of the grid to historical equipment failure events to identify patterns in the types of assets that fail. PG&E's historical asset data was used to develop a training dataset that includes the asset attributes for 2015 through 2021 where each entry, or row, in the training dataset is considered an "asset-year". The historical equipment failure events were joined to the asset-year data set using the asset's primary key and the year in which the equipment failure occurred. If the asset primary key for the equipment-failure event couldn't be identified, then the event was dropped during join process. Ultimately, as detailed in in the Training Event Counts table, the four model subset training datasets included:

- 3,019 support structure equipment failure events
- 11,078 transformer equipment failure events.

WDRM v3 highlights for the modeling the P(o) support structures & transformers subsets:

- The model predicts the probability the asset will experience an equipment failure outage event. This tabular asset-level output can be aggregated to provide pixel-level probabilities so that it can be combined with other model subsets to form a composite.
- The models are asset-based, so asset-specific attributes can influence the likelihood of equipment failure. As summarized in the covariate selection section below, asset attributes were the most influential features in the model algorithm. The influence of asset attributes aligned with the equipment failure causal pathway, which helped promote trust and understanding by end users of the models.
- Model covariates include asset attributes aligned with the mitigation programs using the model. For example, asset age, a top performing covariate in all models, is reset to zero once the asset is replaced. The subsequent model probability output will decrease for the asset as the model covariates are aligned with the current asset state.
- The training data is structured into asset-years, which track whether there were one or more equipment failure events for each asset each fire season. The asset attributes during the year in which the equipment failure occurs are used to train the model. Thus, attributes like asset age and open maintenance tags, can be time varying and the asset attributes at the time of failure are directly used for training.
- The model predictions are asset-based, so work plans can be prioritized by simply joining the model results based on the asset primary key. However, the asset-based results can be summed to the pixel level and used for pixel-based work planning if necessary.

The support structures & transformers model subsets generally demonstrated good performance. Future work will focus on improving performance for the transformer equipment cause subset through additional covariate engineering and algorithm selection experimentation. The ROC performance plots for the four support structures & transformers subsets generated against predicted test event outages are shown to the right.

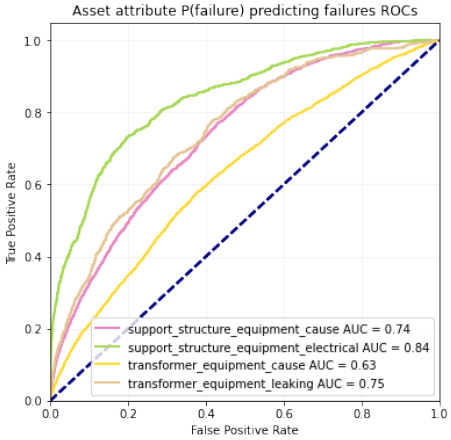
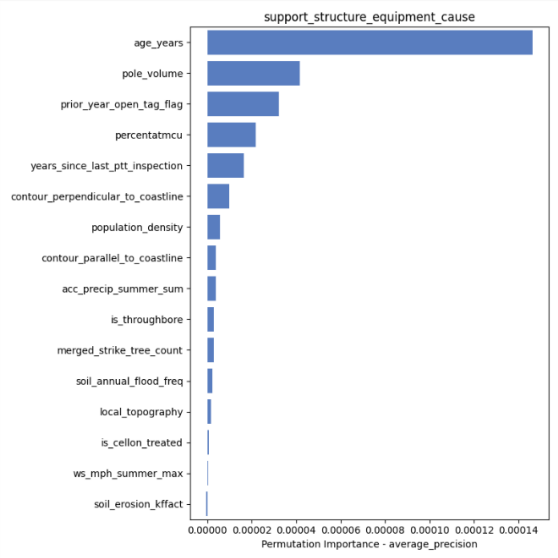


Figure 25 – Support Structures & Transformers Subsets P(o) ROC Curves

Model subset covariates were initially chosen based on discussions with Subject Matter Experts regarding common equipment failure causal pathways. The covariate selections were then refined through an iterative experimentation process where model performance was used to define new covariate sets. Two methods of measuring covariate importance were used: permutation importance and Gini importance score. Feature permutation importance is defined as the average precision that is lost in the model performance when that feature's values are randomly shuffled around. The Gini importance score is specific to tree-based models, like Random Forest, and captures the helpfulness of the feature to the model by quantifying the feature's influence on the tree splits. The resulting feature permutation scores for each model are included in the figures below.



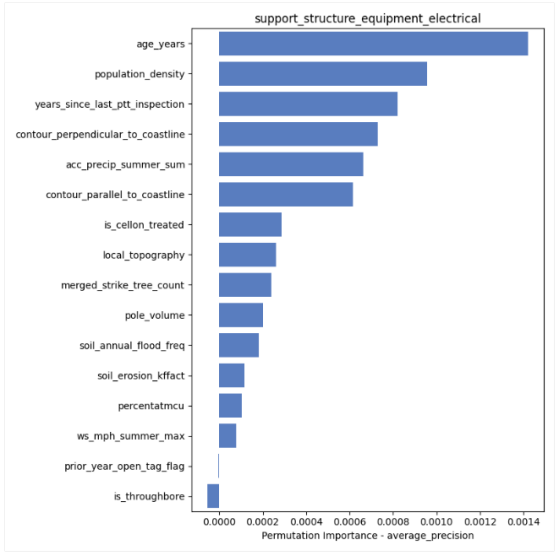
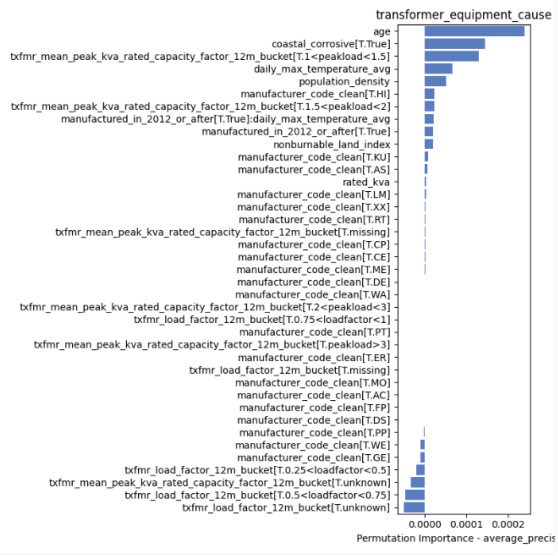


Figure 26 – Support Structure Subsets Covariates Permutation Importance



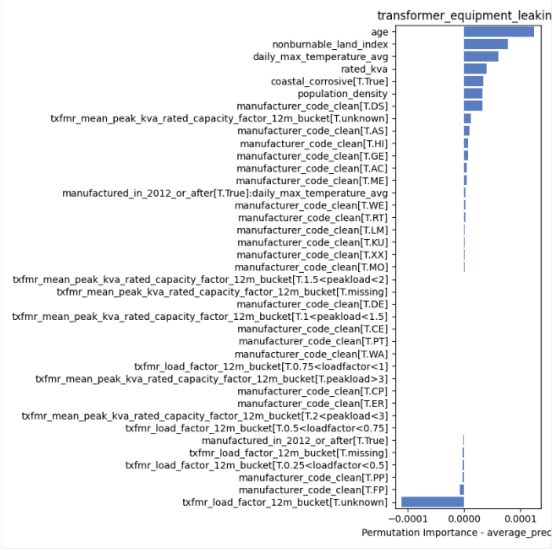


Figure 27 – Transformers Subsets Covariates Permutation Importance

Asset characteristics were particularly important when modeling equipment failure events. Asset age was the most important feature in each of the equipment-failure models. Asset loading information was also a top feature for the pole structural failure model and the transformer failure model.

For more detailed information on the support structures and transformers asset attribute modeling and covariate selection, see the appendix on Asset Time Series.

3.3.4 Predicting Probability of Ignition Given Outage – P(i|o)

The prior sections have described how the P(o) is predicted for the 17 WDRM model subsets. The probability of an ignition given an outage model, or P(i|o), is used to determine the probability that an outage will result in an ignition event. By design, the P(i|o) model transforms our P(o) predictions into P(i) probability of ignition predictions.

The WDRM predicts wildfire risk from all causes impacting all equipment types. To do so, it separates event data into subsets by cause, sub-cause, and equipment involved. Training predictive models directly on ignition event data for each subset is not feasible because several of the resulting subsets lack sufficient data to achieve statistical power. In other words, the models of the subsets lacking sufficient data would be incapable of distinguishing signal from noise and would generate unreliable predictions. Therefore, the WDRM first trains models on cause, sub-cause, and equipment involved subsets of outages, which are far more numerous that ignitions, to produce estimates of the probability of outages for each subset, P(o). A second model is used to estimate the probability of an ignition occurring from the circumstances that caused the outages. When multiplied together, these two probabilities produce the probability of an ignition:

$$P(ignition) = P(ignition|outage)P(outage)$$

The simplest formulation for the P(i|o) model would be to use the historical ratio of the count of ignitions to outages in each subset. Multiplying P(o) by that ratio would re-scale the P(o) data to the expected count of ignitions, but would preserve the spatial pattern, and thus rank order, found in the original P(o) result. The WDRM does not presume that the occurrence rate of ignitions, in both time and space, is a constant fraction of the outage occurrence rate. The heat maps below provide visual evidence that the locations most likely to experience ignitions are not identical to those most likely to experience outages.

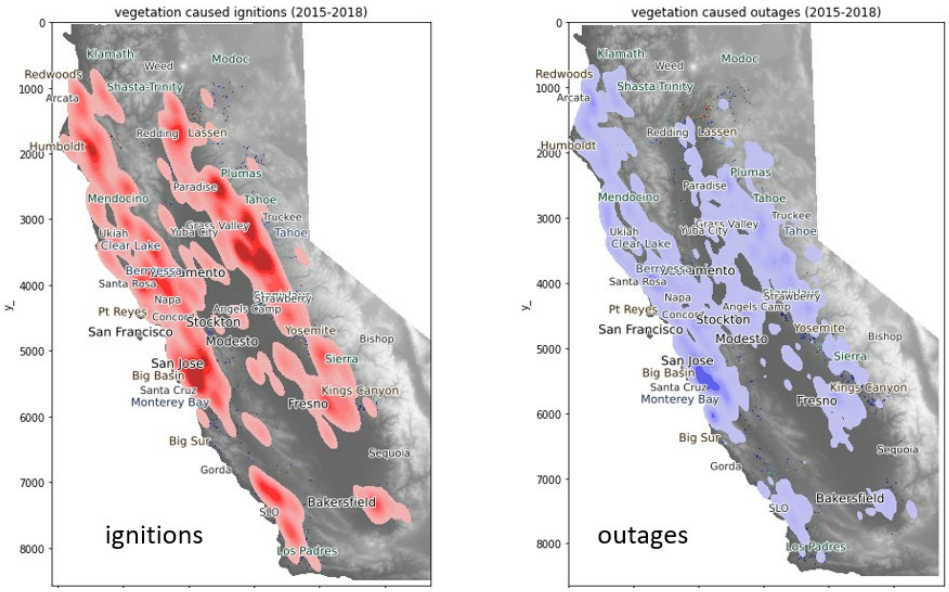


Figure 28 – Density Plots: Vegetation-Related Ignitions and Outages

Reasoning from the observed conditions under which ignitions tend to occur and exploratory data analysis, the P(i|o) is fit using a predictive model that can account for event attributes and local environmental conditions.

The P(i|o) model is a logistic regression trained on all the outage event records in the target set, joined with information on cause, sub-cause, equipment involved, local vegetation, fuel moisture, and wind conditions at the time and location of each outage, augmented with a column of data indicating whether each outage was associated with an ignition. A few special cases for the dataset:

- Ignitions without known outages are considered outages with ignitions
- PSPS hazards and damages are considered outages with ignitions, down weighted to the historical occurrence rate of ignitions under PSPS conditions
    - There are more PSPS damages and hazards than would have occurred with the power on, so PSPS is down weighted to the historical occurrence rate of ignitions under PSPS conditions. Therefore, they only contribute the estimated avoided ignition count to the overall fit.
    - This number was estimated by calculating the sustained-outage-to-ignition rate (614 outages & 47 reportable ignitions) during High Fire Danger Days (64 days) in HFTD between 2015-2018.

The P(i|o) model is built to predict that, when given an outage, with all its known characteristics, what is the probability that an ignition would occur.

The following table summarizes the spatial data sets available as covariates to the P(i|o) model.

Table 8 – WDRM v3 P(i|o) Model Datasets

Dataset	Vintage	Date Span	Description
HFTD	2018	n/a	High Fire Threat District (HFTD) - produced by California Public Utilities Commission (CPUC)
NED	n/a	n/a	National Elevation Database (NED) - produced by United States Geological Survey (USGS)
NLCD	2016	n/a	National Land Cover Database (NLCD) - produced by United States Geological Survey (USGS)
National Terrestrial Ecosystems	2011	n/a	The GAP/LANDFIRE National Terrestrial Ecosystems data represents a detailed vegetation and land cover classification for the Conterminous U.S., Alaska, Hawaii, and Puerto Rico - produced by the United States Geological Survey (USGS)
PG&E Meteorology	Jul-21	2010-2020	POMMS data used as inputs to, or the outputs of, the Fire Potential Index (FPI) model.
Vegetation Satellite	2019	2019	High resolution, territory-wide vegetation data - tree heights, density, etc - from Salo Sciences-processed satellite data.
WorldPop	2010 - 2020	n/a	WorldPop estimates number of people residing in each 100m grid cell

The Recursive Feature Elimination (RFE) implementation from scikit-learn was used to identify covariates that lack explanatory power for the P(i|o). The final model uses 27 covariates. 23 covariates are for differing ignition rates by model subset, and the remaining 4 covariates are related to ignition conditions:

- Presence in a high fire threat district
- Maximum wind speed at the outage location on the day it occurred
- The least 10 hour dry fuel moisture on the day of the outage
- 1000 hour dry fuel moisture on the day of the outage

The resulting model can be used to estimate the P(i|o) from a given subset, in a specific location, for the weather and fuel conditions on the day of the outage. However, the WDRM v3 predicts typical fire season risk values. For compatibility with the WDRM, the P(i|o) model is used to predict daily values across multiple years of historic fire season weather and fuel conditions, with the daily estimates aggregated into seasonal predictions. The aggregation process weights each day's contribution to the aggregate by its observed count of outages. In this way, the seasonal estimate automatically emphasizes the conditions associated with elevated outage rates, capturing their outsized importance in determining ignition probabilities. For example, the elevated rate of vegetation caused outages associated with late summer winds leads to those late summer conditions factoring more heavily into the seasonal P(i|o) for vegetation-caused outages.

Using O_d as the count of outages for a given day, we can express the logic of the marginalization across days as follows:

$$P(ignition|outage) = \sum_{days=d} O_d \cdot P(ignition|outage; conditions)_d$$

The figure below presents the feature importance for each covariate in the P(i|o) model fit. The prominence of subset terms underscores how important subset-specific ignition rates are to the prediction of probability of ignition. However, fuel moisture, location relative to the HFTDs, and wind speed are also in the top 10 most important features.

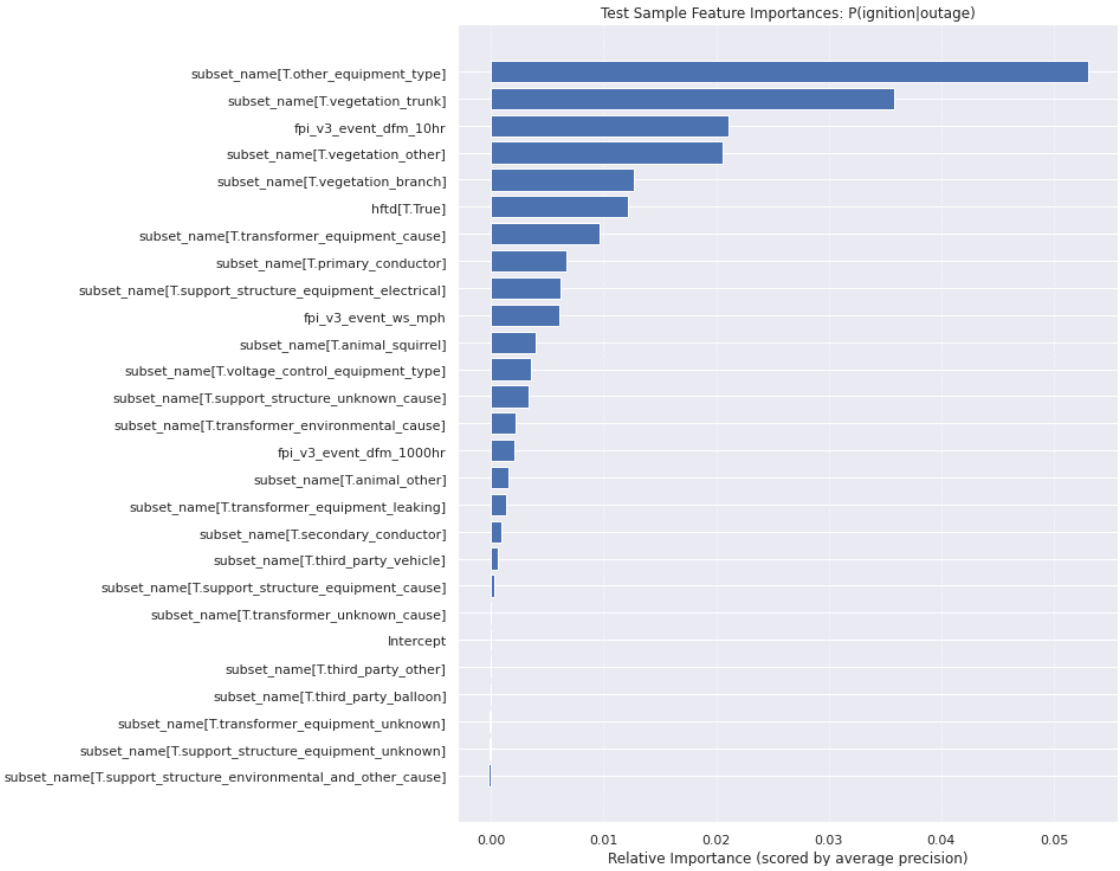


Figure 29 – P(i|o) Covariate Importance

The ROC curve for out of sample predictive performance of the P(i|o) model is shown to the right. The curve and associated AUC value capture the skill with which the model assigns elevated ignition probabilities to outages that were associated with real ignitions. It can be verified by consulting the curve that the model does quite well, capturing nearly 80% of true ignitions after covering just 20% of the target set in rank order.

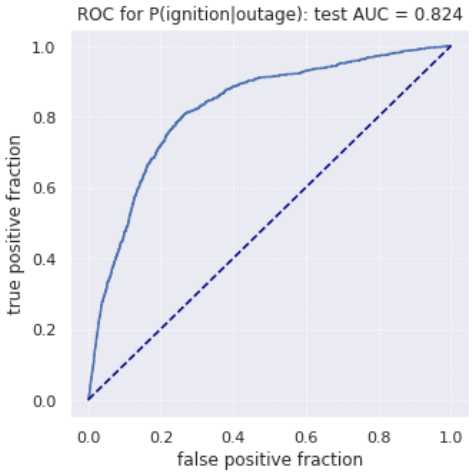


Figure 30 – P(i|o) ROC Curve

3.3.5 Calculate Probability of Ignition – P(i)

As described previously, the probability of ignition, P(i), is the product of the probability of outage, P(o), and the probability of ignition given outage, P(i|o).

$$P(ignition) = P(ignition|outage)P(outage)$$

P(i) is calculated for every distribution grid pixel and asset location for each of the 17 model subsets. The results are calibrated to match the predicted total number of ignition events that would be expected annually based on historical rates.

The plots below show the ROC curves for P(i) for all the model subsets, with the exception for transform equipment leaking. Leaking transformers have not been historically associated with any ignitions, so P(i) cannot be determined for this subset.

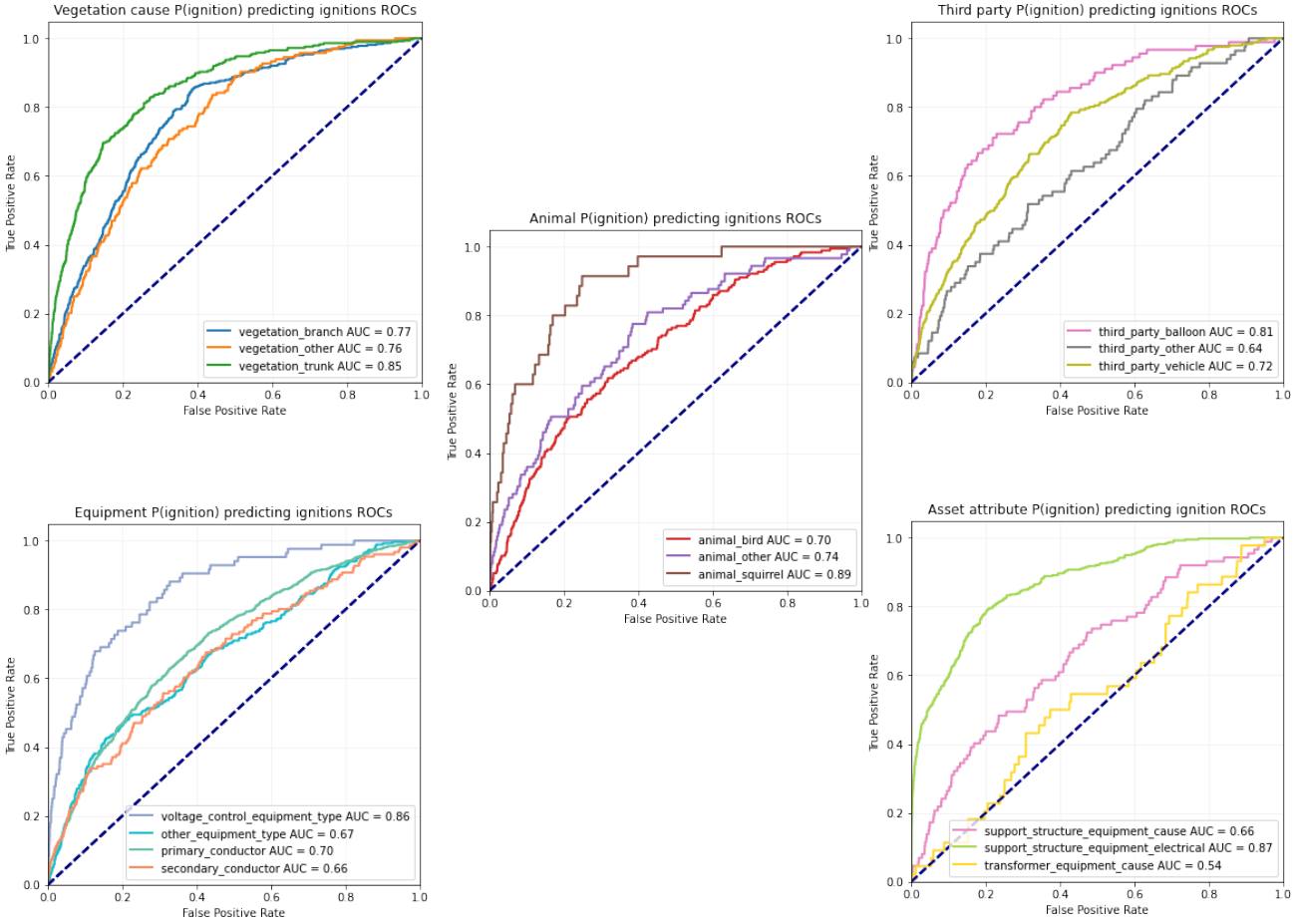


Figure 31 – P(i) ROC Curves

Several performance metrics are used to analyze the P(i) predictions. The table below presents the AUC, average precision, and top 20% concentration factors for the model subsets P(i) predictions. The average precision is the concentration of ignitions per ranked location up through all ignitions being accounted for. The concentration factor, used to understand model performance for highly ranked predictions, measures the concentration of true positives in the top 20% of predictions as a multiplicative factor compared to random chance. Improving model concentration factor can yield significant gains in the risk performance of resources budgeted for mitigation programs.

Table 9 – Model Subset Performance, Ranked by AUC

Subset	P(i) AUC	P(i) Precision	Top 20% Conc. Factor	Outage Count	Ignition Count
animal_squirrel	0.889	0.0011	4.000	3,694	40
support_structure_equipment_electrical	0.868	0.0047	3.896	2,096	582
voltage_control_equipment_type	0.857	0.0012	3.690	502	99
vegetation_trunk	0.851	0.0016	3.693	4,388	329
third_party_balloon	0.813	0.0009	3.389	2,127	103
vegetation_branch	0.775	0.0009	2.775	6,912	406
vegetation_other	0.755	0.0006	2.622	1,655	184
animal_other	0.742	0.0009	2.528	834	106
third_party_vehicle	0.725	0.0009	2.407	6,952	265
animal_bird	0.703	0.0003	2.360	4,831	219
primary_conductor	0.702	0.0023	2.334	12,343	974
other_equipment_type	0.670	0.0005	2.321	46,981	316
support_structure_equipment_cause	0.664	0.0000	2.126	4,631	194
secondary_conductor	0.663	0.0014	2.053	3,801	216
third_party_other	0.636	0.0010	1.867	2,202	102
transformer_equipment_cause	0.541	0.0000	1.023	8,809	62

It is often informative to view P(i) on color-coded maps of the distribution grid. Viewing the model subsets, we can easily see differences in risk patterns. Some interesting examples are shown on the following pages.

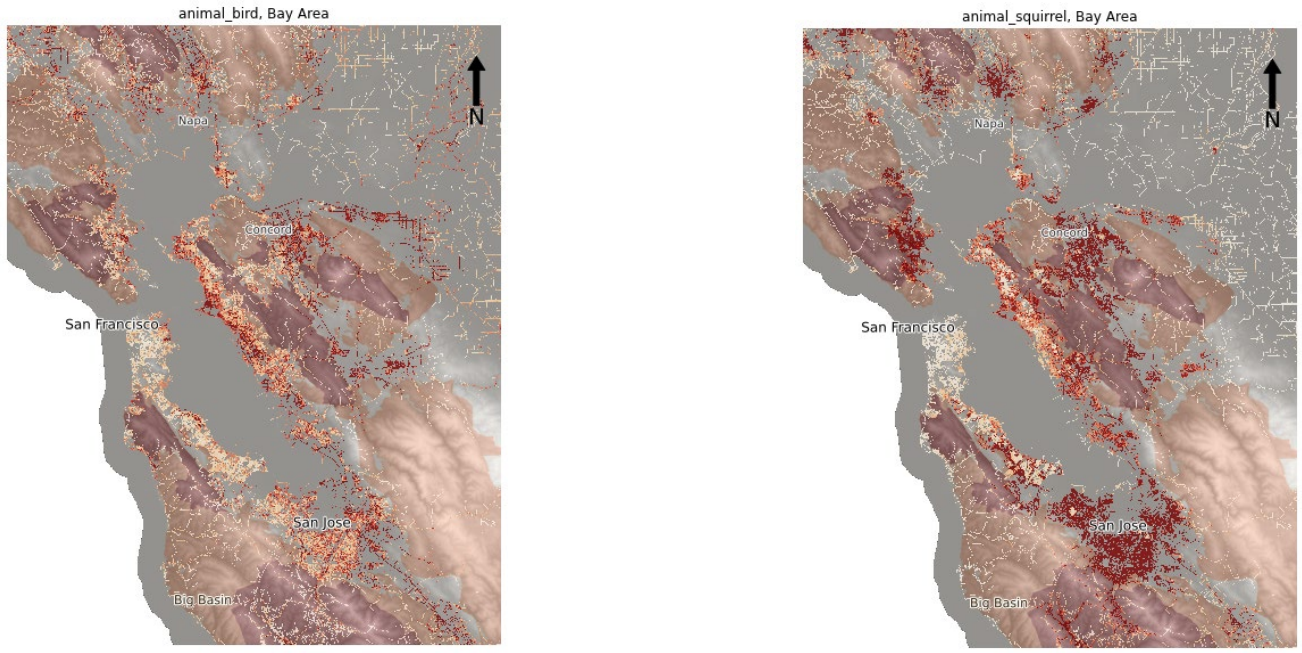


Figure 32 – Bay Area P(i), Bird vs. Squirrel

Bird events are more widely distributed but concentrated along the bay shore and other bird habitat and migratory routes.

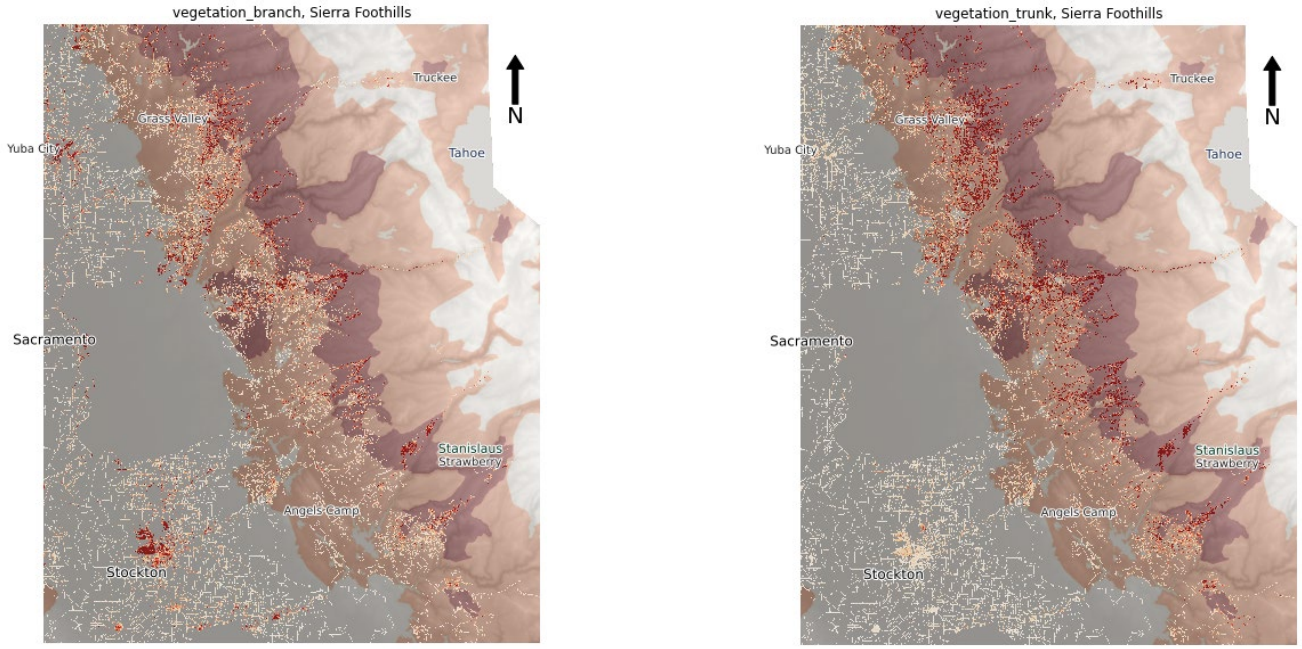


Figure 33 – Sierra Foothills P(i), Vegetation Branch vs. Trunk

Branch failures extend more consistently into suburban and urban areas while trunk failures are most likely in heavily forested areas.

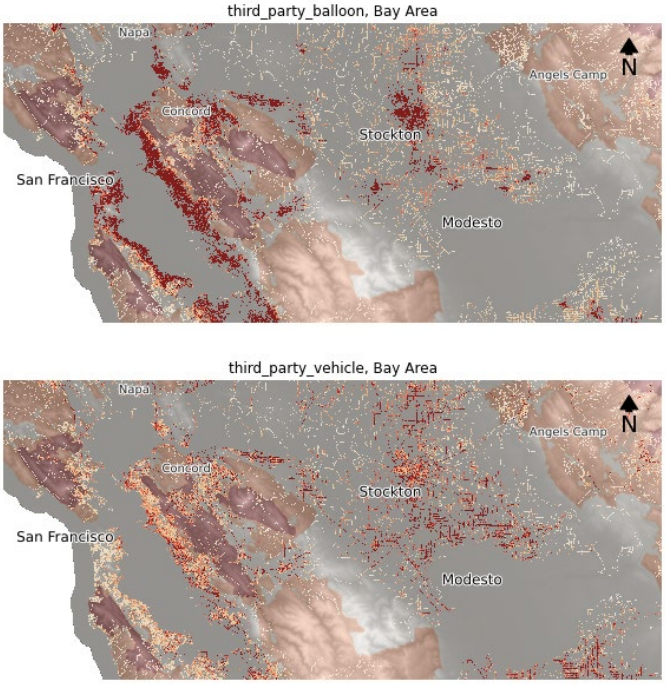


Figure 34 – Bay Area P(i), Balloon vs. Vehicle

The probability of balloon ignition events is, as expected, highly concentrated in densely settled areas. Vehicle ignition events are more likely transportation corridors rather than in urban core areas.

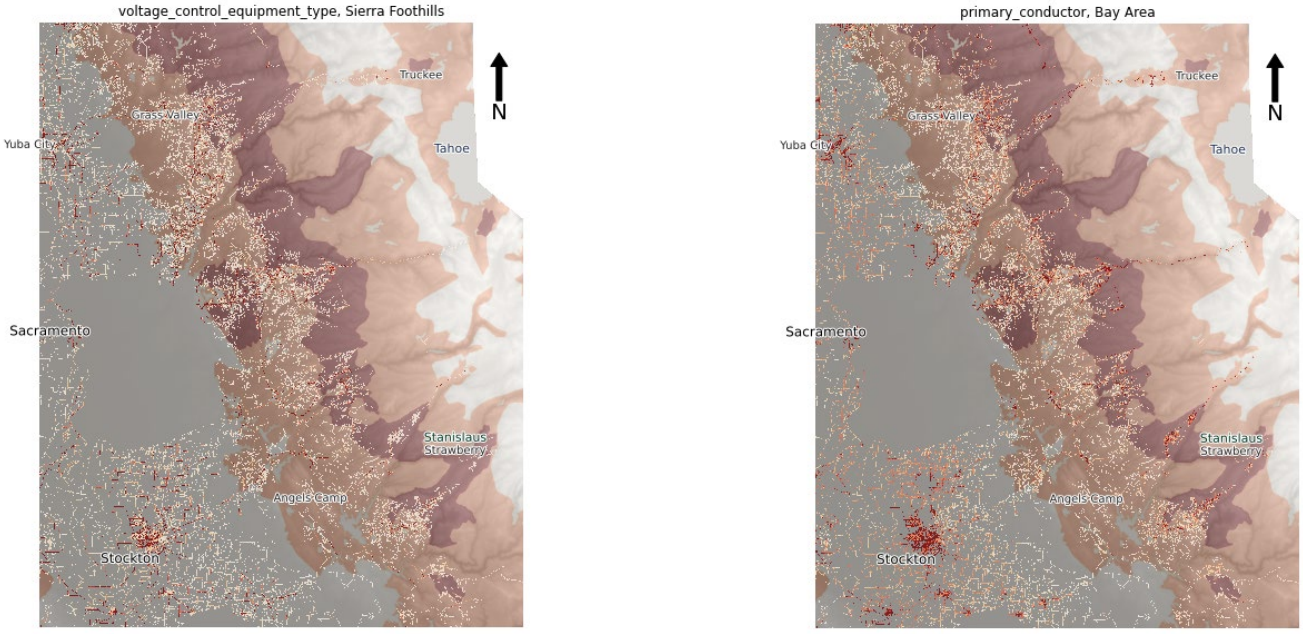


Figure 35 – Sierra Foothills P(i), Voltage Control Devices vs. Primary Conductors

Voltage control device probabilities of ignition driven by line characteristics and are spatially heterogeneous, with some pockets of high concentration. Primary conductor ignitions are more common in Sierra foothill and in urban centers.

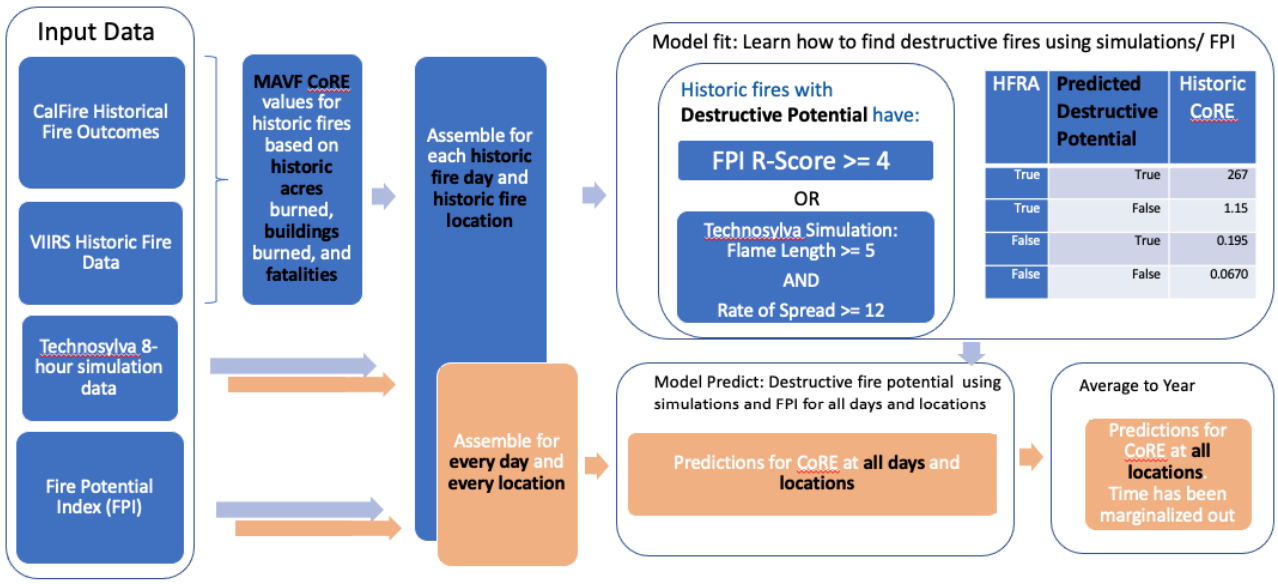
3.4 CoRE – Consequence Value

The v2 consequence model drew upon Tecnosylva simulations of acres burned, structures impacted, flame length and rate of spread returned (with flame length and rate of spread combined into a 1-5 Fire Behavior Index, called FBI), and then determining fire severity via thresholds of those values for every day of weather simulated. The daily severity for each simulation location was assigned a consequence consistent with MAVF CoRE. Final consequence values were determined as averages over all simulation days for each location.

The v3 consequence model draws upon 4 sources of data: the same physical outputs from 2021 updated simulations from Tecnosylva, satellite detected fires from VIIRS (infrared satellite), CalFire data on fire outcomes correlated to VIIRS fires (used to assign MAVF CoRE values), and daily estimates of the 1-5 scaled R-score produced for every 2x2km square in the territory on a daily basis by the models behind PSPS events. For v3, fire severity for a given day is assessed for destructive potential for all days in the June through November fire season. If either approach evaluates to destructive potential, the day/location is considered to have consequences consistent with the expectation value of MAVF CoRE assigned to destructive fires from the VIIRS data set. The use of FPI R-score in addition to the Tecnosylva simulations allows for the marginalization of consequence values across the entire fire season, not just the worst weather days approach used by Tecnosylva.

3.4.1 Overview Diagram

For CoRE the process steps are outlined as follows:



3.4.2 The relationship between asset location conditions and fire potential

Historic fires from the VIIRS data are combined with CalFire and other agency data on outcomes (buildings burned, acres burned, fatalities) to produce MAVf CoRE consequences for historic fires. The available data is joined to Tecnosylva WRRM 8hr simulations and to FPI R-score for all times and locations.

The relationship between the Tecnosylva WRRM simulations and the historic data on destructive fires is illustrated below.

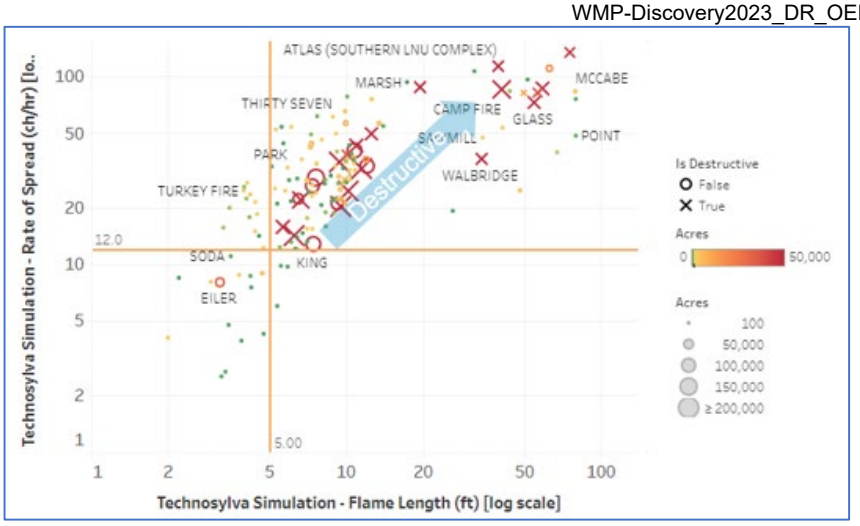


Figure 36 – Tecnosylva Simulation and Destructive Fire Relationship

The relationship between the FPI R-score and the historic fire data was examined as well.

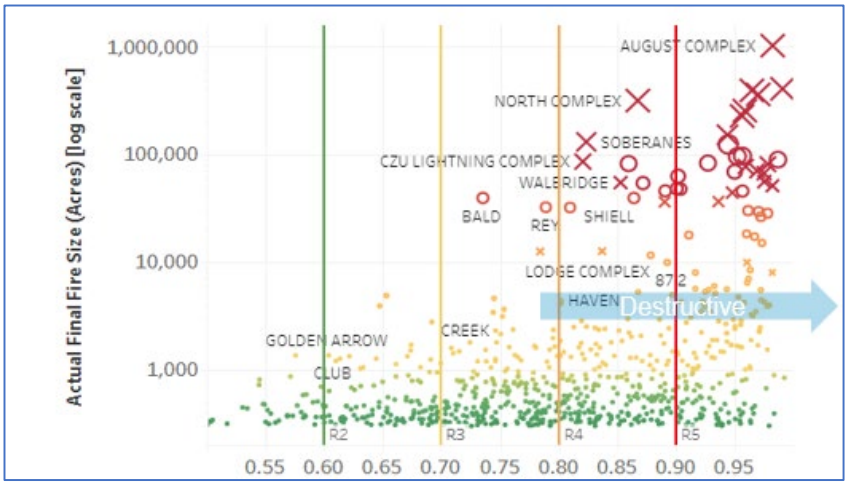


Figure 37 – FPI R-score and Destructive Fire Relationship

Analysis of the historic fire data led to the discovery of a partition function, shown to the right, that is used as a classifier. When the conditions for this classifier are met, then a potentially destructive fire is predicted. Conversely, non-destructive potential is predicted when the classifier conditions are not met. Each of the predicted destructive/non-destructive outcomes has an associated mean MAVf CoRE consequence from the observed, historic outcomes. Predicted destructive potential/non-destructive potential are computed both inside and outside the HFRA to complete the partition.

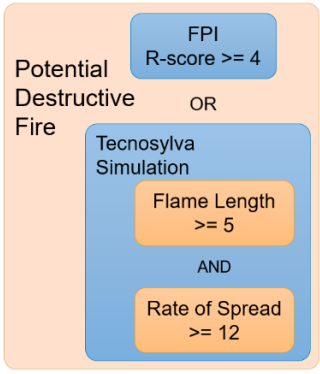


Figure 38 – Destructive Potential Classifier

Table 10 – WFC FPI Model

Using the classifier described above and the starting locations of historical fires, the mean MAVf was determined for a matrix of HFRA designation and the destructive potential prediction for each historical fire location.

HFRA	Predicted Destructive Potential	CoRE from mean MAVf of historic fires
True	True	267
True	False	1.15
False	True	0.195
False	False	0.0670

3.4.3 Assigning grid pixel CoRE values from destructive potential classification

To project current CoRE consequences the covariates are computed for as many pixels as possible. FPI R scores are computed for all times and most pixels. Tecnosylva fire simulations are computed for worst condition days and at roughly 200m intervals along the locations of grid assets. However, Tecnosylva does not produce simulations in locations where it is believed that that a wildfire cannot be sustained such as urban, industrial, and agricultural areas.

For each day in the fire season, the FPI R-score and Tecnosylva simulation results are classified for each pixel as shown in Figure 14 – Destructive Potential Classifier. From the pixel destructive potential classification, the appropriate CoRE value is assigned from Table 5 – WFC FPI Model. The final CoRE value for each pixel is the aggregate of the daily CoRE values.

3.4.4 Assigning backstop CoRE values for unclassified pixels

As noted, some locations have Not Available (NA) for their predicted destructive potential. Overwhelmingly, by a ratio of 100:1, Not Available pixel results are in non-HFRA locations. Tecnosylva simulations are absent for reasons such as the simulation point is too far from a specific pixel or a simulation point is considered 'unburnable' due to physical characteristics such as urban or agricultural settings.

Table 11 – WFC FPI Model Backstop Values

An analysis determined that historic fires outside the HFRA area where Tecnosylva declined to simulate are similar to non-destructive fires outside the HFRA where simulations were performed. Therefore, outside the HFRA, when destructive fire potential is Not Available, the WDRM uses CoRE from the mean MAVf of historic fires outside of HFRA.

HFRA	Predicted Destructive Potential	CoRE from mean MAVf of historic fires
True	NA	NA
False	NA	0.0670

Inside the HFRA, no value of CoRE is assigned to pixels with Not Available predictions because the reasons for missing a simulation point are divergent. This missing value could be the result of a location that is considered unburnable and Tecnosylva therefore declined to simulate an ignition at that location. It could be that the pixel

is too far from the nearest asset location simulation point, as the grid provided by Tecnosylva is irregular. Therefore, the WDRM always retains Not Available results for the CoRE for HFRA pixels. Less than 0.1% of all HFRA pixels have a Not Available value.

3.4.5 WDRM v3 Consequence pixel map

Figure 17 provides a color-coded map of the fire season average Consequence value. For WDRM v3, the high consequence values are typically found in the foothill regions of the distribution grid.

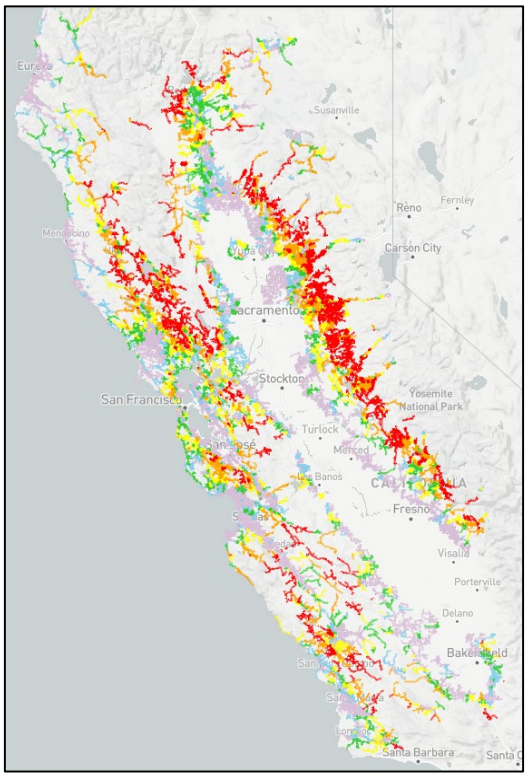


Figure 39 – WDRM v3 Consequence Map

3.5 Wildfire Risk Scores

For each grid pixel along the overhead distribution system, the WDRM assigns a wildfire risk score based upon the product of probability of ignition, P(i), or LoRE and consequence, or CoRE. The principal output of the WDRM is an assigned wildfire risk score for each grid pixel for each model subset. The subset-level grid pixel risk values can be summed across subsets to compute composite risk values for specific mitigation programs. Grid pixel risk values can also be aggregated with their associated circuit segments to deliver circuit segment risk values. These computations are described in the following sections.

3.5.1 Pixel Wildfire Risk Score

For each pixel in each model subset, the pixel probability of ignition, P(i), is multiplied by the its consequence value to determine the final wildfire risk value.

$$Risk = LoRE * CoRE\ (= P(igntion) * Consequence)$$

For any pixels that are missing either its P(i) or consequence value, the resulting wildfire risk value will be missing. Missing risk values are rare but can happen due to small gaps in covariate or fire simulation spatial coverage.

3.5.2 Compositing Wildfire Risk Scores

Ultimately, the purpose for the WDRM is to inform the prioritization of work planning for risk mitigation programs. The WDRM model subsets can be flexibly composited to provide risk values and priority rankings for specific mitigation programs. Risk can be composited for grid pixels or for specific assets. Mitigation planners and Subject Matter Experts can focus on the drivers of risk for which they are responsible with confidence that their composited view is relevant to their work planning needs. The ability to build custom composites of risk is a key improvement of WDRM v3 over prior versions. The following figure provides a visual explanation of compositing.

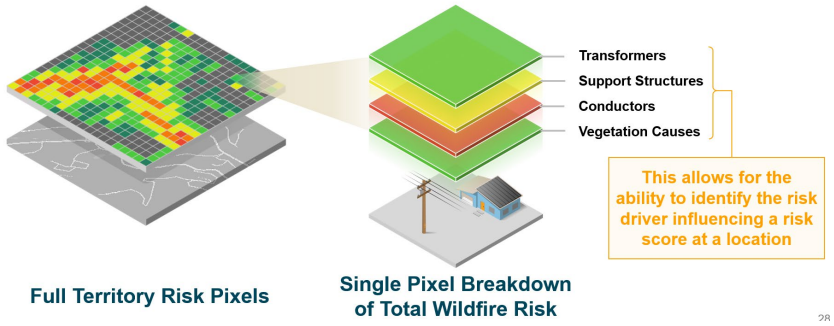


Figure 40 – Compositing Model Subset Risk

WDRM v3 composites can be specified to support specific mitigation strategies. Directly supported mitigation programs include:

- System Hardening
- Enhanced Vegetation Management (EVM)
- Support Structures
- Transformers

The following table provides examples of how compositing has been done for two mitigation programs, EVM and System Hardening.

Table 12 – EVM and System Hardening Compositing

Subset	Total Pixels	EVM	System Hardening
Vegetation subsets			
Vegetation (trunk) caused	1,418,832	Info	X
Vegetation (branch) caused	1,418,832	X	X
Vegetation (other) caused	1,418,832	Info	X
Animal subsets			
Animal bird	1,366,059		X
Animal squirrel	1,366,059		X
Animal other	1,366,059		X
Third party subsets			
Third party vehicle	1,366,059		X
Third party balloon	1,366,059		X
Third party other	1,366,059		X
Conductor and other equipment subsets			
Primary conductor	1,366,059		X
Secondary conductor	408,846		X
Other equipment type	1,366,059		X
Voltage control equipment	1,366,059		X
Support structure subsets			
Support structure – equipment caused	1,211,592		X
Transformer subset			
Transformer – equipment caused	540,457		X

Note that many model subsets have different pixel counts. This is for two reasons. First, many assets types are covered by a fraction of the total pixels for the distribution grid. Second, in some cases there are small gaps in covariate coverages for some model subsets.

EVM, while primarily interested in branch-related risk for planning purposes, composites other vegetation causes for informational reasons.

The figure to the right is a map of the wildfire risk values for the System Hardening composite. Interestingly, the composite risk is more strongly influenced by consequence values than probability of ignition. Intuitively, this is because the potential for a catastrophic wildfire at a location is of much higher magnitude that the likelihood of an ignition occurring at the location.

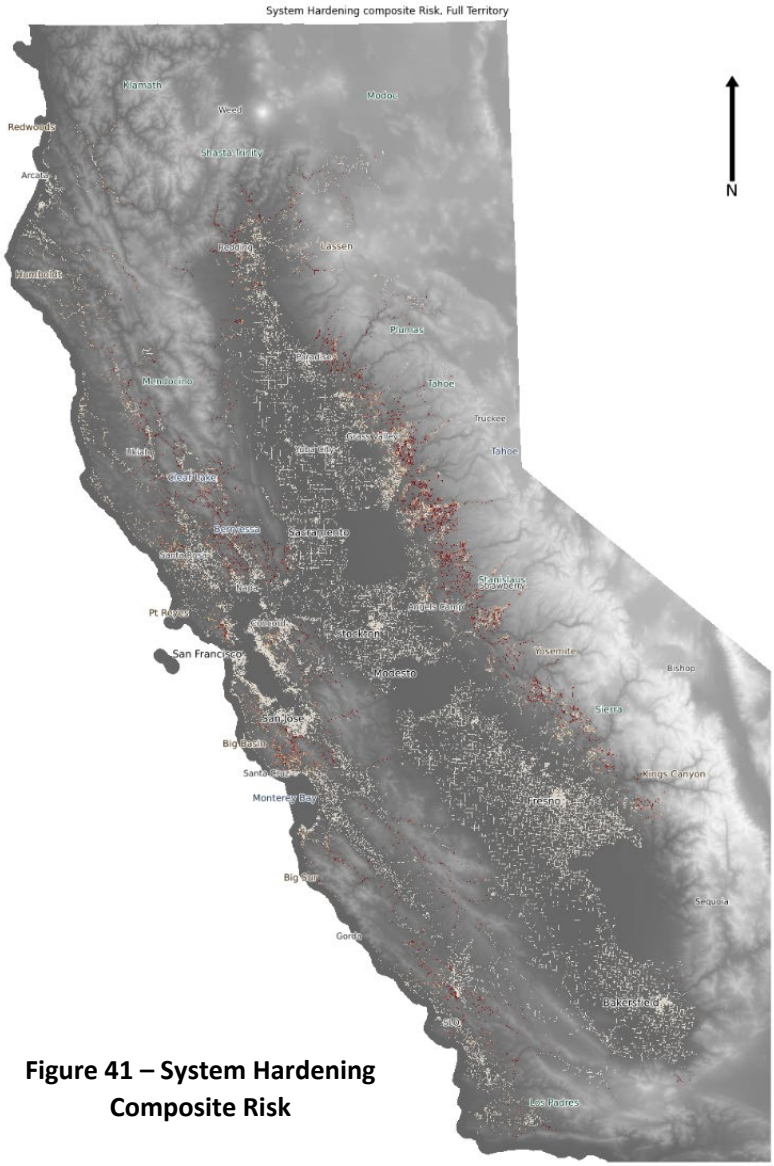


Figure 41 – System Hardening Composite Risk

3.5.3 Circuit Segment Risk Aggregation

Model subset and composite level grid pixel information can be too fine grained for effective planning. Some mitigation programs prefer to consider risk by circuit segments, where a segment is composed of all assets protected by a dynamic protective device on the grid. The pixels that correspond to each circuit segment can have their values rolled up into circuit segment aggregates that can then be ranked by probability of ignition, consequence value, or wildfire risk as inputs into a mitigation planning process.

3.6 Wildfire Risk Mitigation

Estimating the potential reduction in ignition events, and therefore future wildfire risk from various wildfire mitigation programs was a key objective of the WDRM v3. Mitigation programs such as System Hardening, and Enhanced Vegetation Management (EVM) are conducted to reduce wildfire risk. Mitigation programs address the LoRE, or P(i), component of wildfire risk. The program goals are to reduce the likelihood of future grid outages and failures which can lead to ignition events and potentially catastrophic wildfires.

A well-defined procedure to estimate the potential reduction in wildfire risk from PG&E mitigation program activities is vital for prioritization for work plans. Risk reduction estimates inform the comparison between alternate mitigation strategies and programs and help to optimize work planning and allocation of mitigation resources for systemic risk reduction.

WDRM v3 addresses four mitigation programs:

- System Hardening
- Enhanced Vegetation Management (EVM)
- Pole Replacement Prioritization
- Transformer Replacement Prioritization

Mitigation values are provided at a pixel level as a residual expectation of ignition after applying a specific mitigation action. The residual ignition value can be multiplied by the CoRE, or Consequence, value, to determine the expected future wildfire risk value.

The WDRM mitigation calculations have some limitations:

- v3 only considers one year of mitigation efforts
    - Mitigations vary in timeframe effectiveness
        - System Hardening effectiveness degrades slowly over decades
        - EVM effectiveness may only last a few years as vegetation grows back quickly
- Mitigation effectiveness does not currently consider costs in v3
    - Proposed improvement is to include some form of Net Present Value (NPV) for a future version

3.6.1 System Hardening Mitigation Effectiveness Factors

System Hardening Subject Matter Experts (SMEs) were consulted to understand how to implement mitigation calculations for WDRM v3. The SMEs estimated the effectiveness of 4700 historical failure modes. Each failure mode was defined by four characteristics:

- Outage cause
- Supplemental cause
- Equipment involved
- Equipment condition

Each failure mode corresponded with one or more failures as found in historical records from 2015 through 2019. The SMEs provided categorical mitigation effectiveness for each failure mode The SMEs subsequently assigned numeric effectiveness values for each category:

Table 13 – System Hardening Categorical Mitigation Effectiveness Estimates

Effectiveness Category	Effectiveness Estimate
None	0.00
Low	0.20
Medium	0.40
High	0.75
All	0.90

The System Hardening mitigation effectiveness for each subset is the weighted average over the failure modes,

$$\zeta_S = \frac{\sum_m N_{S,m} \times \zeta_m}{\sum_m N_{S,m}}$$

Where:

N_S^{outage} : total failure events in subset S
ζ_m : SH effectiveness in mitigating failure mode m
$N_{S,m}$: number of outages in S matching failure mode m

NOTE: Typically, $\sum_m N_{S,m} = N_S^{eff.match} \leq N_S^{outage}$ since not all failures have corresponding effectiveness estimates, either due to uncertain labels, or due to being notifications.

The above calculation produces the following effectiveness factors for the System Hardening subsets:

Table 14 – System Hardening Mitigation Effectiveness

Subset	Outage Count	Ignition Count	Mitigation Effectiveness
animal_bird	4831	219	0.663
animal_other	834	106	0.705
animal_squirrel	3694	40	0.662
other_equipment_type	46981	316	0.370
primary_conductor	12343	974	0.461
secondary_conductor	3801	216	0.524
support_structure_equipment_cause	4631	194	0.779
third_party_balloon	2127	103	0.794
third_party_other	2202	102	0.418
third_party_vehicle	6952	265	0.482
transformer_equipment_cause	8809	62	0.726
vegetation_branch	6912	406	0.552
vegetation_other	1655	184	0.541
vegetation_trunk	4388	329	0.333
voltage_control_equipment_type	502	99	0.150

3.6.2 EVM Mitigation Effectiveness Factors

The mitigation effectiveness factors for EVM subsets were directly provided by the relevant SMEs as fixed values. Whereas the effectiveness of EVM is often a complex consideration of the grid equipment, tree species mix, natural environment, and local climate, for the purposes of identifying the potential risk reduction due to EVM, the SME provided values are considered useful.

The EVM mitigation effectiveness factors used for WDRM v3:

Table 15 – EVM Mitigation Effectiveness

Subset	Outage Count	Ignition Count	Mitigation Effectiveness
vegetation_branch	6912	406	0.85
vegetation_other	1655	184	0.60
vegetation_trunk	4388	329	0.50

3.6.3 Transformer Replacement Mitigation Effectiveness Factors

The transformer replacement mitigation effectiveness factors were developed in collaboration with transformer replacement SMEs and were applied to the *transformer_equipment_cause* subset. The calculation procedure requires first estimating a base effectiveness value informed by the failure rate of low-risk, young transformers. This base effectiveness value is modified by SME-provided asset attribute specific multipliers for higher-risk, older transformers.

The base efficacy value ζ_b evaluates to 0.28 as the numerical solution of the following equation:

$$\zeta_b = 1 - \frac{r_{young}}{r_{all}} * \left(\sum_{p \in \{permutations\}} f_p * \frac{1 - \zeta_b}{1 - x_p * \zeta_b} \right)$$

The following multipliers were established by the transformer replacement SMEs:

$$f_{bad\_actor} = \begin{cases} 1 & not\ bad \\ 1.2 & bad \end{cases}$$

$$f_{average\_peak\_load\_percent} = \begin{cases} 1 & 0 - 100\% \\ 1.1 & 100 - 150\% \\ 1.2 & 150 - 200\% \\ 1.3 & 200 - 300\% \\ 1.4 & > 300\% \end{cases}$$

$$f_{load\_factor} = \begin{cases} 1 & 0.00 - 0.25 \\ 1.05 & 0.25 - 0.50 \\ 1.1 & 0.50 - 0.75 \\ 1.2 & 0.75 - 1.00 \end{cases}$$

$$f_{manufacture\_date} = \begin{cases} 1 & \geq 2012 \\ 1.1 & < 2012 \end{cases}$$

The following formula was used to calculate the effectiveness values 40 potential multiplier combinations:

$$\zeta = \zeta_b * f_{bad\_actor} * f_{average\_peak\_load\_percent} * f_{load\_factor} * f_{manufacture\_date}$$

3.6.4 Pole Replacement Mitigation Effectiveness Factors

The *support_structure_equipment_cause* model subset was used for prioritizing pole replacement work planning. The pole replacement mitigation program is aimed at replacing poles identified from an inspections process as structurally deficient. The pole replacement mitigation effectiveness factor was estimated from the annual failure rate of previously replaced poles and the pole failure model's weighting of pole replacement-related features. This mitigation effectiveness estimation approach is called the Similarity Method.

These pole replacement-related covariates include,

- pole age
- loading percent at Mcu
- prior year open tag count (for B, E, F, & H tags)
- number of years since last invasive PT&T inspection

Each of these covariates can be set to zero in the model after a pole has been replaced. Since these covariates are highly influential on probability of outage, the WDRM predicts a significantly reduced likelihood of failure upon replacement.

Initially, the Similarity Method buckets the population of poles into groups based on the above features. The categorical buckets used to group the pole population are:

- Age
  - 0-40 years
  - >40 years
- Percent at Mcu
  - <=31%
  - Unknown
  - >31%
- Open tag in prior year
  - True
  - False
- Years since last PT&T inspection
  - 0-4
  - >4

Each of the deficient pole categorical groups are compared to the ideal young & healthy pole group. The ideal young & health pole is less than 40-years-old, with zero maintenance tags in the prior year, with a PT&T inspection within the last 4 years, and estimated at less than 31% of the Mcu. A recently replaced pole always falls into this young & healthy group. The effectiveness factor, ζ, of pole replacement is calculated as:

$$\zeta = 1 - \frac{expected\ annual\ failure\ rate\ for\ mitigated\ pole\ group}{expected\ failure\ rate\ for\ deficient\ pole\ group}$$

The following table provides the mitigation effectiveness factors for replacing poles with different deficiencies as characterized in the list of categorical buckets. Note that the first that the first row in the table represents the ideal and healthy young pole.

Table 16 – Pole Replacement Effectiveness

Age (years)	Percent at Mcu	Open Tag, Prior Year	Last PT&T Inspection (years)	Replacement Effectiveness
0-40	<= 31%	False	0-4	0
0-40	<= 31%	False	4+	0.054041
0-40	<= 31%	True	0-4	0.292436
0-40	<= 31%	True	4+	0.350737
0-40	unknown	False	0-4	0.279867
0-40	unknown	False	4+	0.376295
0-40	unknown	True	0-4	0.513729
0-40	unknown	True	4+	0.583575
0-40	> 31%	False	0-4	0.413562
0-40	> 31%	False	4+	0.449270
0-40	> 31%	True	0-4	0.554744
0-40	> 31%	True	4+	0.587601
40+	<= 31%	False	0-4	0.600210
40+	<= 31%	False	4+	0.627793
40+	<= 31%	True	0-4	0.719409
40+	<= 31%	True	4+	0.750524
40+	unknown	False	0-4	0.723916
40+	unknown	False	4+	0.749747
40+	unknown	True	0-4	0.804098
40+	unknown	True	4+	0.830226
40+	> 31%	False	0-4	0.787688
40+	> 31%	False	4+	0.805365
40+	> 31%	True	0-4	0.825908
40+	> 31%	True	4+	0.847246

4 Results and Model Performance

4.1 WDRM v3 Performance

Subset-level model predictions produced AUCs ranging up to 0.88 (animal squirrel), with 5 subsets higher than 0.80, 6 subsets between 0.70 and 0.80, 4 subsets between 0.63 and 0.70, and transformer_equipment_cause and transformer_equipment_leaking at 0.54 and one without any ignitions for establishing truth respectively. The full dataset that trained the overall model (i.e., "all modeled subsets") achieved an AUC score of 0.70. With full territory coverage and events associated with all causes and equipment types, this result is for a much more extensive model than presented in v2.

Model Performance: subset-level P(outage) and P(ignition)

Table 9 provides ROC-AUC values for P(outage) predicting outages (the immediate job of the subset models), and ROC-AUC, average precision, and the top 20% ranked predictions ignition concentration factors for P(ignition) predicting ignitions (the goal of the WDRM) for all modeled subsets. It is sorted from highest to lowest for P(ignition) predicting ignitions AUC, with color scales shared by both AUC metrics for reference. Note that the leaking transformer subset has no historical ignitions and therefore lacks ground truth for computing these performance metrics.

Table 17 – Color Coded Subset Prediction Performance

Subset	P(o) AUC	P(i) AUC	P(i) Precision	P(i) Top 20% Conc. Factor
animal_squirrel	0.882	0.889	0.00112	4.000
support_structure_equipment_electrical	0.838	0.868	0.00472	3.896
voltage_control_equipment_type	0.882	0.857	0.00117	3.690
vegetation_trunk	0.867	0.851	0.00160	3.693
third_party_balloon	0.892	0.813	0.00091	3.389
vegetation_branch	0.831	0.775	0.00090	2.775
vegetation_other	0.839	0.755	0.00055	2.622
animal_other	0.770	0.742	0.00088	2.528
third_party_vehicle	0.765	0.725	0.00093	2.407
animal_bird	0.744	0.703	0.00030	2.360
primary_conductor	0.735	0.702	0.00229	2.334
other_equipment_type	0.755	0.670	0.00045	2.321
support_structure_equipment_cause	0.738	0.664	0.00002	2.126
secondary_conductor	0.706	0.663	0.00139	2.053
third_party_other	0.713	0.636	0.00101	1.867
transformer_equipment_cause	0.632	0.541	0.00002	1.023
transformer_equipment_leaking	0.749	NA	NA	NA

Model Performance: P(ignition|outage)

When working with P(outage) results directly, without the correction provided by P(ignition|outage) model, the implicit assumption is that P(ignition) will follow the same temporal and spatial patterns as P(outage), i.e.

rescaled to account for the lower count of ignitions annually. The net impact of the use of the P(ignition|outage) model can therefore be assessed via a comparison of predictions between P(outage) and P(ignition) for each subset. As illustrated below, the p(i) estimate's out of sample AUC score is higher or neutral for all but 3 subsets. Just a single subset, support_structure_equipment_cause, is significantly worse at predicting ignition locations after the p(i|o) correction. This is presumably because the vast majority of support structure ignitions stay on the poles, rendering the vegetation and fuel conditions on the ground (data consumed by and found generally predictive for the p(ignition|outage) model) moot.

Table 18 – Predictive ROC-AUC Performance Gain from P(i|o)

Subset	*P(o) AUC*	*P(i) AUC*	*P(o) Top 20% Conc. Factor*	*P(i) Top 20% Conc. Factor*	*Differential*
animal_bird	0.666	0.703	2.02	2.36	0.037
animal_other	0.718	0.742	2.53	2.53	0.024
secondary_conductor	0.643	0.663	1.92	2.05	0.020
third_party_vehicle	0.707	0.725	2.24	2.41	0.017
primary_conductor	0.689	0.702	2.13	2.33	0.013
vegetation_branch	0.762	0.775	2.58	2.77	0.013
third_party_balloon	0.800	0.813	3.33	3.39	0.013
animal_squirrel	0.876	0.889	3.71	4.00	0.012
transformer_equipment_cause	0.529	0.541	1.02	1.02	0.012
vegetation_trunk	0.843	0.851	3.66	3.69	0.008
third_party_other	0.630	0.636	1.81	1.87	0.006
vegetation_other	0.751	0.755	2.59	2.62	0.005
voltage_control_equipment_type	0.854	0.857	3.75	3.69	0.003
transformer_equipment_leaking	0.000	0.000	NaN	NaN	0.000
support_structure_equipment_electrical	0.872	0.868	3.98	3.90	-0.004
other_equipment_type	0.675	0.670	2.21	2.32	-0.005
support_structure_equipment_cause	0.687	0.664	2.41	2.13	-0.023

4.1.1 WDRM v2 vs. v3 comparisons

Due to changes in the categories of events modeled, the geographic territory covered and circuit segment name and topology changes, there is no direct comparison between v2 and v3, the closest comparison is a comparison of the vegetation models. In v2 of the WDRM, the vegetation model had an overall AUC score of 0.73, while this year's model (which models trunks, branches, and other separately) improves over the v2 performance, with an AUC score for the vegetation composite of 0.78. The v2 conductor-involved events (AUC 0.73) were largely vegetation caused (v3 composite AUC score of 0.78), but also drew upon conductor equipment events (v3 AUC of 0.70).

Nevertheless, if the resulting HFTD circuit segment prioritization for the System Hardening composite from the v2 and v3 model are compared, the movement of circuit segments can be observed. The Sankey chart below displays movement between the top two quartiles and the lower 50% from the v2 model on the left to the v3 model on the right. The following observations are made which correspond to flows on the Sankey chart:

1. 25% of v3 1st and 2nd quartile circuit segments are new and were not present in the v2 model
2. The majority of new v3 circuit segments are ranked in the lower half in the v3 model
3. Approximately 1/3 of v2 1st and 2nd quartile circuit segments move to the lower 50%

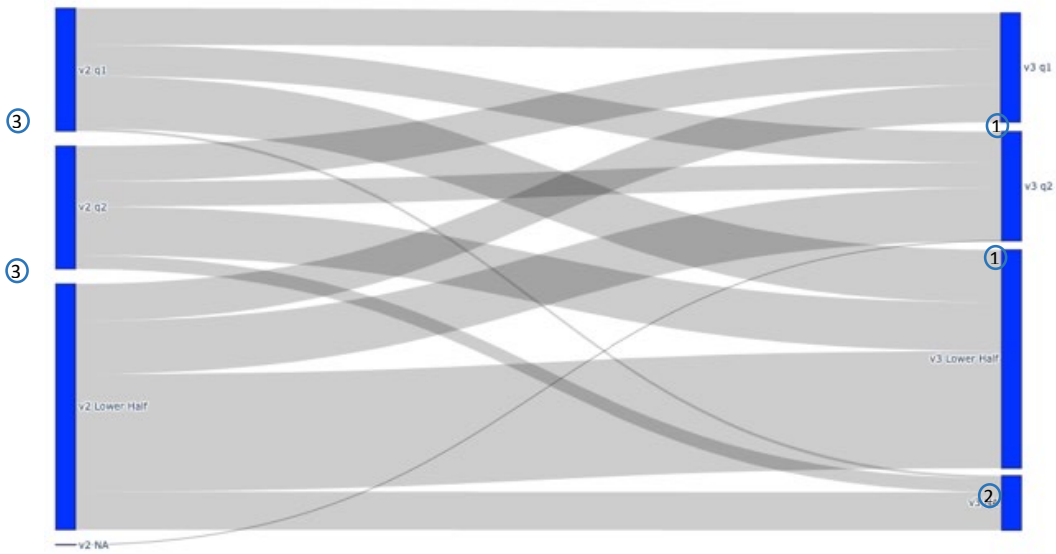


Figure 42 – Sankey View System Hardening Circuit Segment Risk, v2 vs. v3

In most cases, the central cause of decreased risk was updated consequence values, however, the following causes were also found to play a role:

1) Geometry changes including addition, subtraction, and the splitting of the CPZ into two or more distinct circuit segments in v3
2) Name changes including the absorption of CPZs into others resulting in the original CPZ no longer existing
3) The handling of CPZs that were partially included in high fire threat districts. In v2, the total risk was found by only summing those risk pixels belonging to the CPZ whose centroid was within the HFTD while in v3, all pixel risks belonging to the CPZ were summed regardless of HFTD inclusion.

As most work plans prioritize the top 20% of circuit segments, risk changes were studied for those circuit segments which were in the top 20% of risk in v2 but had dropped out of the top 20% in v3. As shown in Figure 38 below, of the 727 circuit segments in the top 20% of v2 risk, 88 were absorbed into another circuit segment name and another 68 were less than 1km long. Of the remaining 571 circuit segments, 208 remained in the v3 top 20% (or top 1,000). Of the top 727 circuit segments in the v2 top 20%, 363 dropped to the lower 80%. From among these, most of these moves (343) were dominated by a large shift in the wildfire consequence value and rank. A small portion of these moves (12) were influenced by both a large shift in the circuit segment mileage and wildfire consequence. Finally, eight circuit segments moved due to a shift in the ignition probability and were minimally influenced by wildfire consequence or a change in length.

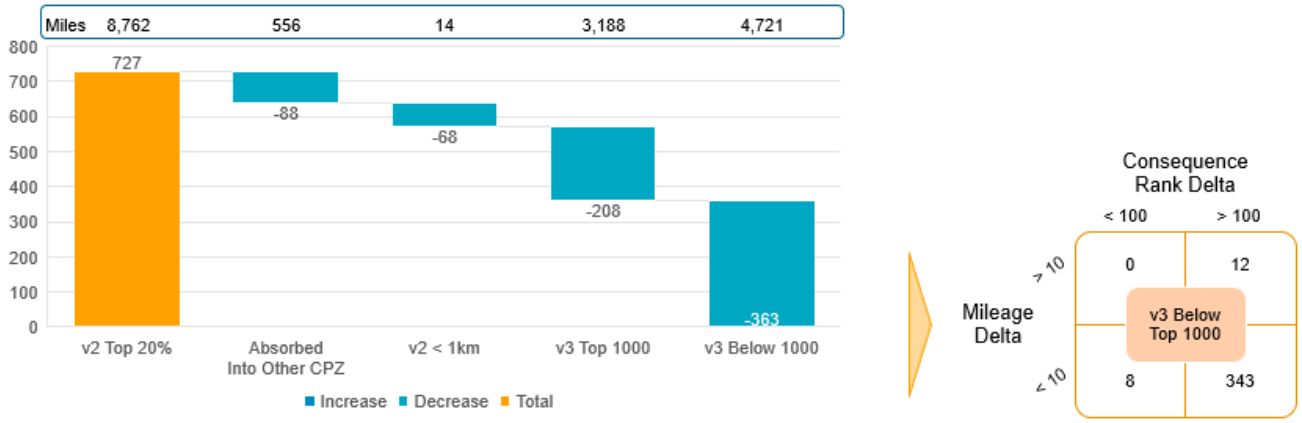


Figure 43 – v2 Top 20% Breakdown by Segment

A few examples of these movements are illustrative:

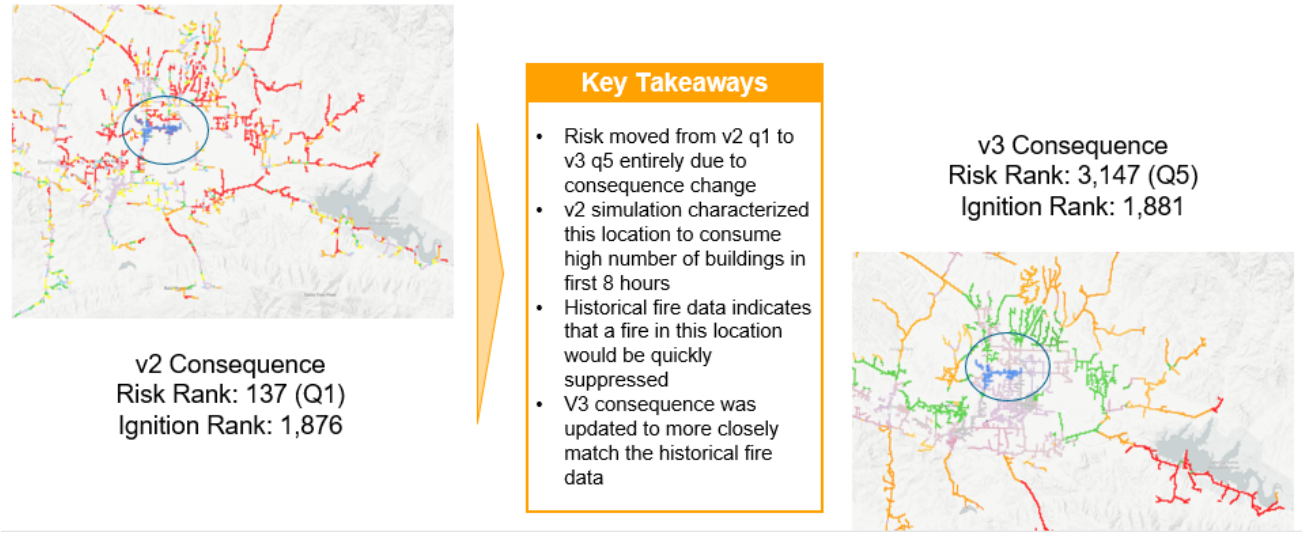
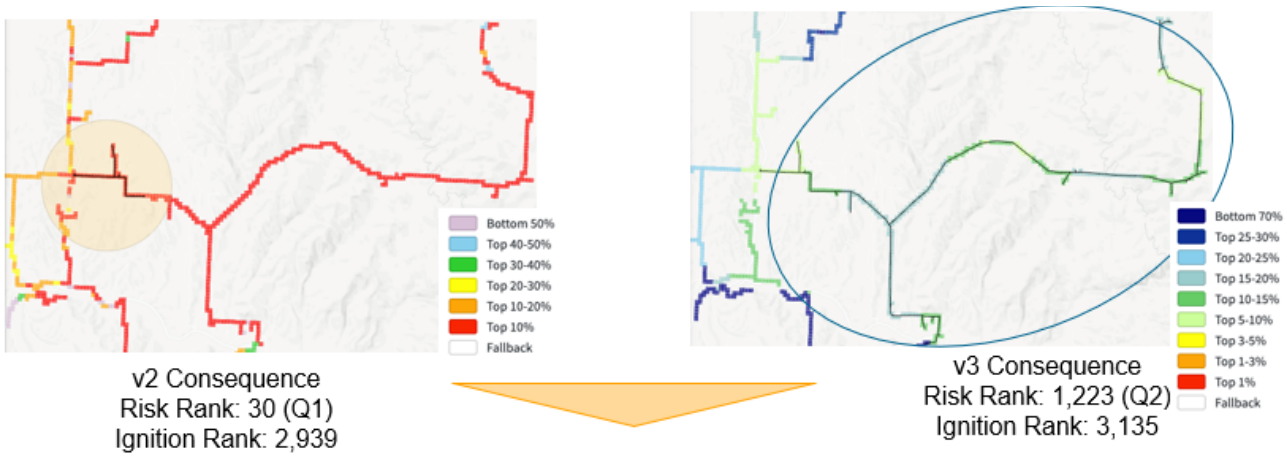


Figure 44 – Santa Ynex 1104CB: Large Consequence Change, Small Mileage Change



Key Takeaways

- While consequence values for the area are somewhat lower in V3 than V2, the circuit segment is much larger in V3 providing a more diverse set of pixels for the median consequence value.

Figure 45 – Poso Mountain 2103CB: Small Consequence Change, Large Mileage Change

V3 Risk Drivers

V2 Risk Drivers

Subset name	v2 System Hardening	v3 System Hardening	v3 veg branch	v3 veg trunk	v3 veg other	v3 primary conductor	v3 secondary conductor	v3 other equipment	v3 support structure	v3 txformer	v3 voltage control	v3 animal bird	v3 animal squirrel	v3 animal other	v3 third party balloon	v3 third party vehicle	v3 third party other
mean annual ignition count	60	517	58	47	26	139	31	45	28	9	14	31	5.7	15	15	38	15
APPLE HILL 2102CB	770	205	767	348	207	1315	4952	353	1261	6565	342	1218	68	3157	5604	1275	440
GARBERVILLE 11011770	51	5,480	1,331	216	2,778	9,374	10,219	8,420	9,964	6,086	8,974	10,861	8,122	9,887	10,655	8,223	6,324
LAYTONVILLE 1102141962	372	6,673	2,977	489	3,294	8,858	10,398	8,973	8,195	3,041	8,867	10,474	6,727	9,649	10,791	10,406	7,195
LAYTONVILLE 1102572	372	7,326	3,169	721	3,722	9,958	10,084	9,486	9,125	5,228	8,794	9,710	6,436	9,610	10,667	9,811	9,439
DIAMOND SPRINGS 1104705100	915	8,389	5,857	1,349	5,385	9,474	9,597	9,602	10,886	8,020	6,192	8,711	3,611	8,716	9,807	8,767	10,175
BANGOR 11017446	949	9,545	6,481	1,493	5,965	10,581	10,335	10,194	9,256	6,196	8,612	10,195	4,874	7,126	10,415	10,639	10,465
APPLE HILL 210277546	1,762	9,657	7,267	1,486	4,727	10,300	9,588	10,598	10,250	8,134	10,257	9,777	6,093	10,611	10,492	10,603	5,976
WHITMORE 11011594	1,089	9,751	6,548	1,458	4,806	10,826	9,692	10,298	7,357	7,329	9,569	10,986	4,437	10,394	11,053	10,581	8,824
FORT SEWARD 1121CB	400	9,850	6,563	1,172	4,154	10,522	8,540	10,951	9,798	9,541	10,231	10,969	7,180	10,539	10,745	10,554	10,951
ATASCADERO 1103A16	1,665	10,698	7,955	4,147	9,201	10,635	8,552	9,385	10,889	9,837	9,746	8,268	8,991	9,289	10,526	10,601	10,833

Key Takeaways

- V2 considered the four veg and conductor risk drivers only, while V3 includes 11 additional risk drivers.
- The mean annual ignition counts included in the top row, demonstrate the influence (purple) of each risk driver in the System Hardening composite. The primary conductor risk driver is dominant in the composite.
- The Apple Hill 2102CB is a reference case where ignition probability is generally high across all risks.
- The rest of the table displays cases where ignition probability is high in V2 but the inclusion of the additional risk drivers in V3 drop the ranking.

Figure 46 – WDRM Risk Drivers, v2 vs. v3

4.1.2 Power System Specialists (PSS) Review

PG&E Power System Specialists (PSS) developed a circuit-based risk assessment review. This qualitative assessment is based on their collective 300+ years of fire experience as most of the PSS had a previous career in fire with CalFire or other Fire agencies. The PSS assessment assigned a qualitative score of 0, 5, 15, or 30 for five categories focused not on ignition probabilities but factors that will contribute to the ability to manage a fire given their unique history in fighting fires in these locations. These categories were: Fire History, Ingress/Egress

Impacts, Resistance to Control, Community Risk Factors, and Other Unique Local Factors. These five values are then combined to achieve a total value for all circuits in the HFTD and HFRA.

These qualitative values were compared with the System Hardening Risk Composite scores by assembling the PSS assessment along the WDRM v3 risk buy-down curve. As circuits with higher PSS values tend toward the upper end of the risk buy-down curve the PSS assessment and WDRM v3 correlate particularly in the first half of the curve.

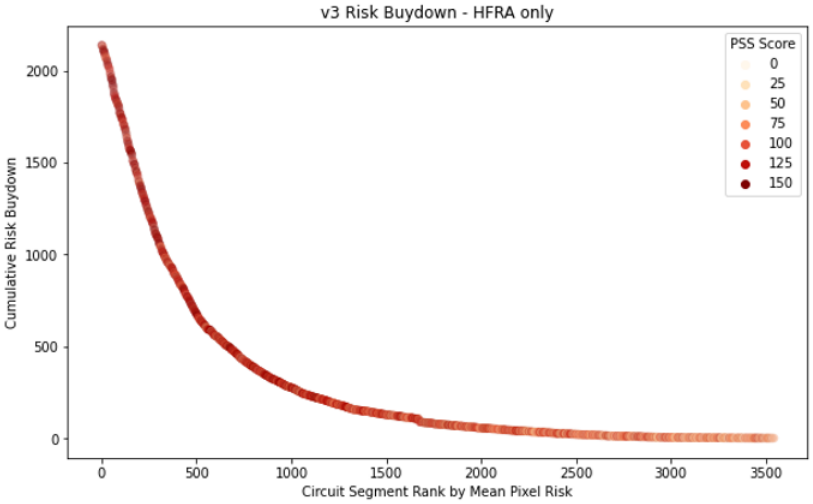


Figure 47 – HFRA WDRM v3 System Hardening Risk Buydown

4.1.3 E3 Review

An independent, third-party review of the WDRM v3 was conducted by E3 – Energy, Environment and Economics. The objective of the review was three-fold:

- Review the suitability and applications of consequence data in the modelling framework,
- Review the specific use of the Risk Model Information in each of its operations areas; and,
- Describe potential future uses of v3 and longer-term multi-year wildfire planning models.

This was a deliberate expansion of the objective for the E3 review of the WDRM v2 model which was to determine whether the model was 'fit for purpose'.

As result of the review E3 concluded the following:

"PG&E has made substantial progress in transforming its model from one that was primarily used to validate mitigation measures chosen by its subject matter experts (SME) within high fire zone areas to a model that can be used to supplement and prioritize the targeting of mitigation measures across its entire service territory." The "construct of v3 appears to be consistent with their commitment in their WMP to refocus mitigation work to achieve a target where 80 percent of their work is focused on mitigating the risk of the highest 20 percent of identified line segments."

"PG&E has made a substantial effort to incorporate feedback from the CPUC, stakeholders, E3 and its internal users to update the WDRM between versions 2 and 3. The updates made represent real improvements in several critical areas. From E3's review, the modeling team includes a group of highly skilled professionals from inside and outside of PG&E. The model is leveraging the best available data and methods to prioritize risk levels by geographic area and ignition type allowing for evidence-based decision-making. This model represents an improvement from v2…Most of modeling limitations are driven by limitations in data and resources which are difficult for the modeling team to directly solve."

In line with the third objective to 'Describe potential future uses of v3 and longer-term multi-year wildfire planning models', E3 identified a number of items for future improvement of the WDRM in future iterations.

"*While PG&E should be commended for its rapid development of a model that shows substantial promise to increase the effectiveness of their mitigation work, our recommendations focus on a few existing gaps:*

- *Standardizing and documenting the relationship between the model and subject matter experts*
- *The transparency and validity of the consequence portion of the model*
- *Establishing a data quality control process*
- *Establishing a roadmap for model direction*
- *Exploring potential further use cases of the model*
- *Coordination of PG&E's process with broader State-wide wildfire planning*"

E3 also assessed PG&E's progress on the 11 recommendations from the WDRM v2 validation report. E3 found that the WDRM v3 completely addressed 8 of these recommendations and made progress on the remaining 3. These three, a more detailed modeling roadmap, tighter coordination with SME input, and transparency of the Wildfire Consequence data, are highlighted in the WDRM v3 validation report for continued progress. The detailed table of

For more information, please see the complete E3 validation report in Appendix X

4.1.4 Wildfire Governance Steering Committee Review

The WDRM v3 was presented for review and approval to the Wildfire Risk Governance Steering Committee (WRGSC) on two occasions. The first occasion was to document the development of model features that were 2021 Wildfire Mitigation Plan (WMP) commitments and approve a new annual model development schedule and the second was to review and approve the WDRM v3 for use in developing wildfire mitigation work plans.

On December 15, 2021, the first presentation was made to the WRGSC. The objective of this presentation was to 1) Inform the committee of the improvements made to the WDRM v3 and 2) To approve the future model development schedule. As part of the reported improvements 6 features were detailed that met WMP commitments. These were:

1. Model extends beyond HFTD to entire distribution system
2. Added models for support structures and transformers
3. Updated training data sets with 2020 outages, ignitions and PSPS damages
4. Developed models to composite or add probabilities and risks
5. Automation of composite model framework (CPUC Recommendation
6. Risk reduction for mitigation options at a granular level

In addition, the development of four features recommended in the E3 Validation report were presented:

1. Develop and evaluate wider range of model algorithms
2. Improve coordination between PSPS Operational and WDRM Planning Models
3. Include model parameters that relate more closely to risk mitigation measures
4. Strengthen link between experts and models

Supporting these 10 added features, improvements on data sets such as historic weather, LiDAR tree survey data and pole loading calculations were presented.

The annual model development schedule was all presented and approved. In order to provide improved continuity for work plan execution, the new schedule proposed a phased approach for model application. As shown in Figure 3, after a model is approved at the end of the first quarter, work plans with shorter implementation cycles could begin to use the model while longer term work would look to apply the new model to plan work for the next year and beyond.

As a result of the meeting the new features were documented, and the model development schedule was approved.

The second presentation of WDRM v3 to the WRGSC was held on April 13, 2022 for the purpose of approving the updated WDRM v3 for use in work planning. Originally, the meeting was scheduled for March 30, 2022, in compliance with the commitment to develop the final WDRM v3 model by the end of Q1 2022. Due to a lack of quorum from the voting members of the committee, the original meeting was postponed, and the WDRM v3 was approved by the WRGSC on April 13, 2022.

As a result of the meeting, the WDRM v3 was approved with a number of follow-up items.

5 WDRM Future Plans

A model that attempts to predict the future, in an inherently chaotic system like that of PG&E's energy system in the California environment, can always be improved. New data becomes available and new data science tools and techniques are developed. Even simply the availability of staff time to invest in data quality, model design, and tuning of hyperparameters – all while utilizing existing data and tools - can produce improvements in prediction quality and applicability to PG&E risk mitigation programs.

PG&E plans to continue to improve the WDRM over the coming years. The next iteration of the model, v4, which is planned to be developed during 2022, will include:

- Modeling related to additional assets, including capacitor banks, fuses and voltage regulators
- Modeling of additional causes of outages and ignitions, such as those caused by animals, third-party interactions, and lightning
- Modeling how long it takes for fire suppression to get into a wildfire area (ingress) and how long it takes residents and workers to evacuate an area (egress)
- Standardizing and documenting the relationship between the model and subject matter experts
- Establishing a data quality control process
- Expanding the current roadmap for model direction
- The ability to account for the effects on EPSS (protective device settings that make them more likely to trigger protective outages under fire conditions).
- Improve the ability of model algorithms to specify the circumstances under which specified mitigations are applicable. This will enable a more precise application of mitigation effectiveness factors.

PG&E plans to incorporate the following into WDRM v5, which may be developed in 2023-2024:

- Modeling of additional assets such as distribution protection devices
- Modeling of outages and ignitions with unknown causes

PG&E is planning to incorporate the following in the WDRM v6:

- Understanding other challenges in the area, such as from seismic risk so that work plans can balance between different risk.

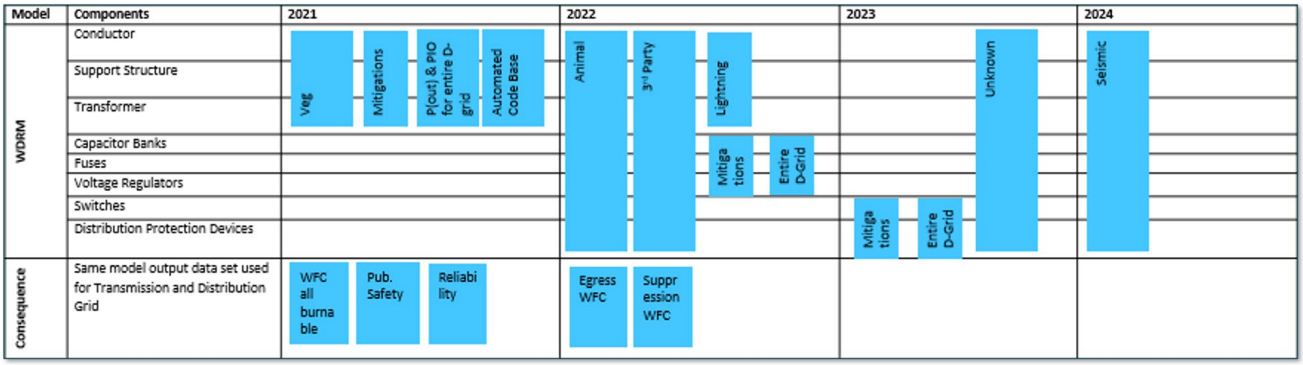


Figure 48 – WDRM Improvement Schedule (OEIS Presentation - October 5th, 2021)

In addition, note that the WDRM is currently modeled separately from PG&E's Operational Public Safety Power Shutoff (PSPS) model, and other wildfire risk-related modeling efforts. While the RaDA team communicated with the PG&E staff and contractors involved in these other efforts, there is an overlap since investments in long-

term risk reduction decrease the likelihood of PSPS shutoffs and vice versa. Future modeling efforts will seek to integrate the WDRM with other PG&E modeling efforts, and with similar efforts at other California utilities, as makes sense, and as is required by regulatory guidance.

This roadmap for improvements to the WDRM will be revisited on a continual basis as PG&E seeks to quantify wildfire risk and risk reduction and target work plans to the end that catastrophic wildfires will stop.

6 Appendix: Covariate Pools

In the WDRM v3, the set of all covariates available for modeling is called the covariate pool. In this section we document the covariate pools available to the P(o) and P(i|o) models and which covariates used by the model subsets.

6.1 P(o) Model Subsets Covariates

The following tables summarize the pool of covariates developed for use in P(o) model development.

Covariate	Dataset	Vintage	Source Type	Units	Description
local_topography	National elevation database	n/a	raster	Unitless positive or negative score	Topographic Position Index (TPI) helps distinguish topographic features such as a hilltop, valley bottom, exposed ridge, flat plain, upper or lower slope. It is calculated by comparing the elevation of each pixel to its surrounding neighbours.
vapor_pressure_deficit_avg	gridmet	2019	raster	kPa	Mean vapor pressure deficit. The daily max value is averaged over the period of fire season (June 1 to Nov 30) during the years 2015-2019
specific_humidity_avg	gridmet	2019	raster	kg/kg	Mean specific humidity, averaged over the period of fire season (June 1 to Nov 30) during the years 2015-2019.
burn_index_avg	gridmet	2019	raster	NFDRS fire danger index - Burning index	Mean NFDRS fire danger index (Burning index), averaged over the period of fire season (June 1 to Nov 30) during the years 2015-2019.
energy_release_avg	gridmet	2019	raster	NFDRS fire danger index - Energy release component	Mean NFDRS fire danger index (Energy release component), averaged over the period of fire season (June 1 to Nov 30) during the years 2015-2019.
wind_avg	Pge meteorology	2021	raster	Miles per hour	Sustained hourly windspeed at 10m above ground level, averaged over the period of fire season (June 1 to Nov 30) during the years 2015-2020.
wind_max	Pge meteorology	2021	raster	Miles per hour	

daily_max_temperature_avg	Pge meteorology	2021	raster	°F	Max of the hourly max temperature each day, averaged over the period of fire season (June 1 to Nov 30) during the years 2015-2020.
100_hour_fuels_avg	Pge meteorology	2021	raster	%	100 hour fuel moisture content - measure of the amount of water in "100 hour" fuel (downed logs and branches 1-3" in diameter), expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across fire season (June through Nov) for the years 2015-2020
1000_hour_fuels_avg	Pge meteorology	2021	raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter), expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across fire season (June through Nov) for the years 2015-2020
precipitation_avg	Pge meteorology	2021	raster		
windy_summer_day_pct	Pge meteorology	2021	raster	%	Fraction of fire season (June through Nov) days with max hourly wind speed over 15 mph.
impervious	National land cover database	2016	raster	%	Percent of the pixel covered by developed impervious surface.
nonburnable_land_index	Landfire surface fuels	2016	raster	%	Percentage of raster pixel considered unburnable, because it is of land types not considered to have combustible fuel (agricultural land, snow and ice, water, and barren/rocky areas)
gusty_summer_day_pct	rtma	Late 2019	raster	%	Fraction of fire season (June through Nov) days with max wind gusts over 20 mph
merged_strike_tree_count	Vegetation merged	2019	raster	Count of trees in each 100m pixel that could strike conductors.	
merged_strike_tree_ count_perlinemile	Vegetation merged	2019	raster	Count of trees in each 100m pixel that could strike conductors, divided by	

				the line miles present in each pixel.	
merged_strike_tree _height_avg	Vegetation merged	2019	raster	meters	
merged_strike_tree _height_min	Vegetation merged	2019	raster	meters	
merged_strike_tree _height_max	Vegetation merged	2019	raster	meters	
tree_species_treemap	treemap	2014	raster	Numerical species descriptor	
cwhr_aspen	cwhr	2014	raster	1/0	
cwhr_blue_oak _foothill_pine	cwhr	2014	raster	1/0	
cwhr_blue_oak _woodland	cwhr	2014	raster	1/0	
cwhr_coastal_oak _woodland	cwhr	2014	raster	1/0	
cwhr_closed_cone _pine_cypress	cwhr	2014	raster	1/0	
cwhr_douglas_fir	cwhr	2014	raster	1/0	
cwhr_desert_riparian	cwhr	2014	raster	1/0	
cwhr_eastside_pine	cwhr	2014	raster	1/0	
cwhr_eucalyptus	cwhr	2014	raster	1/0	
cwhr_jeffrey_pine	cwhr	2014	raster	1/0	
cwhr_joshua_tree	cwhr	2014	raster	1/0	
cwhr_juniper	cwhr	2014	raster	1/0	
cwhr_klamath _mixed_conifer	cwhr	2014	raster	1/0	
cwhr_lodgepole_pine	cwhr	2014	raster	1/0	
cwhr_montane _hardwood_conifer	cwhr	2014	raster	1/0	
cwhr_montan e_hardwood	cwhr	2014	raster	1/0	

cwhr_montane_riparian	cwhr	2014	raster	1/0	
cwhr_pinyon_juniper	cwhr	2014	raster	1/0	
cwhr_palm_oasis	cwhr	2014	raster	1/0	
cwhr_ponderosa_pine	cwhr	2014	raster	1/0	
cwhr_redwood	cwhr	2014	raster	1/0	
cwhr_red_fir	cwhr	2014	raster	1/0	
cwhr_subalpine_conifer	cwhr	2014	raster	1/0	
cwhr_sierran_mixed_conifer	cwhr	2014	raster	1/0	
cwhr_valley_oak_woodland	cwhr	2014	raster	1/0	
cwhr_valley_foothill_riparian	cwhr	2014	raster	1/0	
cwhr_white_fir	cwhr	2014	raster	1/0	
soil_AWC	gnatsgo	2019	raster	inches/inch	Available water capacity. The amount of water that an increment of soil depth, inclusive of fragments, can store that is available to plants. AWC is expressed as a volume fraction, and is commonly estimated as the difference between the water contents at 1/10 or 1/3 bar (field capacity) and 15 bars (permanent wilting point) tension and adjusted for salinity, and fragments.
soil_CLAY	gnatsgo	2019	raster	%	Clay content of soil. Mineral particles less than 0.002mm in equivalent diameter as a weight percentage of the less than 2.0mm fraction.
soil_erosion_kffact	gnatsgo	2019	raster	n/a	An erodibility factor which quantifies the susceptibility of soil particles to detachment by water.
soil_OM	gnatsgo	2019	raster	%	Organic matter content. The amount by weight of decomposed plant and animal residue expressed as a weight percentage of the less than 2 mm soil material.
soil_PERM	gnatsgo	2019	raster	inches/hour	Permeability rates.
soil_THICK	gnatsgo	2019	raster	inches	Total thickness of all sampled soil layers.
soil_HYGRP	gnatsgo	2019	raster	Index	Soil index variable - 1 = well drained to 4 = poorly drained.
soil_DRAIN	gnatsgo	2019	raster	Index	Soil index variable - 1 = well drained to 7 = poorly drained.
soil_SLOPE	gnatsgo	2019	raster	%	Average slope. The difference in elevation between two points, expressed as a percentage of the distance between those points.

soil_LL	gnatsgo	2019	raster	% Moisture by weight	Liquid limit. The water content of the soil at the change between the liquid and plastic states.
soil_IFHYDRIC	gnatsgo	2019	raster	Index	Hydric soil indicator - 1 if hydric.
soil_annual_flood_freq	gnatsgo	2019	raster	Index	Annual flood frequency - 1 = frequent (>50% chance), 2 = occasional (5-50% chance), 3 = rare (<5% chance).
population_density	world_pop	2010 - 2020	raster	Count of people residing in each grid cell	
temp2m_f_summer _max_high	Pge meteorology	2021	location time series raster	°F	
temp2m_f_summer _avg_high	Pge meteorology	2021	location time series raster	°F	
acc_precip_summer_sum	Pge meteorology	2021	location time series raster	%	
ws_mph_summer _gt15mph_pct	Pge meteorology	2021	location time series raster	Miles per hour	

6.2 P(o) MaxEnt Subset Models Covariates

The MaxEnt model fitting process includes a feature generation step, where all the covariates offered to the model are interacted and recast with hinges and thresholds. The resulting set of features are then regularized as a part of the model fitting procedure, resulting in a subset of all features remaining in the final model fits. The following table summarizes all covariates offered to each MaxEnt subset model.

Subset	Animal bird	Animal other	Animal squirrel	Other equipment type	Primary conductor	Secondary conductor	Third party balloon	Third party other	Third party vehicle	Vegetation branch	Vegetation other	Vegetation trunk	Voltage control equipment type
local_topography	X	X	X	X	X	X	X	X	X	X	X	X	X
merged_strike_tree_count	X	X	X	X	X	X	X	X	X	X	X	X	X
merged_strike_tree count_perlinemile	X	X	X	X	X	X	X	X	X	X	X	X	X
merged_strike_tree height_avg	X	X	X	X	X	X	X	X	X	X	X	X	X
merged_strike_tree height_min	X	X	X	X	X	X	X	X	X	X	X	X	X
merged_strike_tree height_max	X	X	X	X	X	X	X	X	X	X	X	X	X
100_hour_fuels_avg	X	X	X	X	X	X	X	X	X	X	X	X	X
1000_hour_fuels_avg	X	X	X	X	X	X	X	X	X	X	X	X	X
burn_index_avg	X	X	X	X	X	X	X	X	X	X	X	X	X
energy_release_avg	X	X	X	X	X	X	X	X	X	X	X	X	X
nonburnable_land_index	X	X	X	X	X	X	X	X	X	X	X	X	X
impervious	X	X	X	X	X	X	X	X	X	X	X	X	X
tier_2	X	X	X	X	X	X	X	X	X	X	X	X	X
tier_3	X	X	X	X	X	X	X	X	X	X	X	X	X
coastal	X	X	X	X	X	X	X	X	X	X	X	X	X
operatingvoltage	X	X	X	X	X	X	X	X	X	X	X	X	X
subtypecd	X	X	X	X	X	X	X	X	X	X	X	X	X
conductor_size and_material	X	X	X	X	X		X	X	X				X
psps_segment	X	X	X	X	X		X	X	X				X
precipitation_avg	X	X	X	X	X	X	X	X	X	X	X	X	X
specific_humidity_avg	X	X	X	X	X	X	X	X	X	X	X	X	X
daily_max temperature_avg	X	X	X	X	X	X	X	X	X	X	X	X	X
vapor_pressure deficit_avg	X	X	X	X	X	X	X	X	X	X	X	X	X
gusty_summer_day_pct	X	X	X	X	X	X	X	X	X	X	X	X	X
windy_summer_day_pct	X	X	X	X	X	X	X	X	X	X	X	X	X
wind_avg	X	X	X	X	X	X	X	X	X	X	X	X	X
wind_max	X	X	X	X	X	X	X	X	X	X	X	X	X
tree_species_treemap	X	X	X	X	X	X	X	X	X	X	X	X	X

cwhr_aspen	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_blue_oak foothill_pine	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_blue_oak woodland	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_coastal_oak woodland	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_closed_cone pine_cypress	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_douglas_fir	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_desert_riparian	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_eastside_pine	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_eucalyptus	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_jeffrey_pine	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_joshua_tree	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_juniper	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_klamath mixed_conifer	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_lodgepole_pine	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_montane hardwood_conifer	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_montane hardwood	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_montane riparian	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_pinyon_juniper	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_palm_oasis	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_ponderosa_pine	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_redwood	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_red_fir	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_subalpine_conifer	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_sierran mixed_conifer	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_valley_oak woodland	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_valley_foothill riparian	X	X	X	X	X	X	X	X	X	X	X	X	X
cwhr_white_fir	X	X	X	X	X	X	X	X	X	X	X	X	X
soil_AWC	X	X	X	X	X	X	X	X	X	X	X	X	X
soil_CLAY	X	X	X	X	X	X	X	X	X	X	X	X	X
soil_erosion_kffact	X	X	X	X	X	X	X	X	X	X	X	X	X
soil_OM	X	X	X	X	X	X	X	X	X	X	X	X	X
soil_PERM	X	X	X	X	X	X	X	X	X	X	X	X	X

soil_THICK	X	X	X	X	X	X	X	X	X	X	X	X	X
soil_HYGRP	X	X	X	X	X	X	X	X	X	X	X	X	X
soil_DRAIN	X	X	X	X	X	X	X	X	X	X	X	X	X
soil_SLOPE	X	X	X	X	X	X	X	X	X	X	X	X	X
soil_LL	X	X	X	X	X	X	X	X	X	X	X	X	X
soil_IFHYDRIC	X	X	X	X	X	X	X	X	X	X	X	X	X
soil_annual_flood_freq	X	X	X	X	X	X	X	X	X	X	X	X	X

6.3 Asset Attribute Models Covariates

[('support_structure_equipment_cause',
'failure_ind ~ age_years + prior_year_open_tag_flag + pole_volume + percentatmcu + years_since_last_ptt_inspection + merged_strike_tree_count + population_density + ws_mph_summer_max + acc_precip_summer_sum + soil_erosion_kffact + soil_annual_flood_freq + local_topography + contour_parallel_to_coastline + contour_perpendicular_to_coastline + is_throughbore + is_cellon_treated + 1'),
('support_structure_equipment_electrical',
'failure_ind ~ age_years + prior_year_open_tag_flag + pole_volume + percentatmcu + years_since_last_ptt_inspection + merged_strike_tree_count + population_density + ws_mph_summer_max + acc_precip_summer_sum + soil_erosion_kffact + soil_annual_flood_freq + local_topography + contour_parallel_to_coastline + contour_perpendicular_to_coastline + is_throughbore + is_cellon_treated + 1'),
('transformer_equipment_cause',
'event_flag ~ age + rated_kva + manufacturer_code_clean + manufactured_in_2012_or_after * daily_max_temperature_avg + txfmr_load_factor_12m_bucket + txfmr_mean_peak_kva_rated_capacity_factor_12m_bucket + coastal_corrosive + population_density + nonburnable_land_index + 1'),
('transformer_equipment_leaking',
'event_flag ~ age + rated_kva + manufacturer_code_clean + manufactured_in_2012_or_after * daily_max_temperature_avg + txfmr_load_factor_12m_bucket + txfmr_mean_peak_kva_rated_capacity_factor_12m_bucket + coastal_corrosive + population_density + nonburnable_land_index + 1')]

6.4 P(i|o) Model Available Covariates

The following table presents the covariates available to the P(i|o) model

Covariate	Dataset	Vintage	Source Type	Units	Description
landcover_mean	national_land cover_database	2016	raster	index	Indicator of land cover type - see **https://www.mrlc.gov/data/legends/national-land-cover-database-class-legend-and-description**

impervious_mean	national_land cover_database	2016	raster	%	Percent of the pixel covered by developed impervious surface.
impervious_descriptor_mean	national_land cover_database	2016	raster	index	Defines which impervious layer pixels are roads and provides the best fit description for impervious pixels that are not roads.
shrubland_annual_herbaceous_mean	national_land cover_database	2016	raster	%	The annual only grass and forb proportion in each in 30m pixel.
shrubland_bare_ground_mean	national_land cover_database	2016	raster	%	The bare ground proportion in each pixel.
shrubland_big_sagebrush_mean	national_land cover_database	2016	raster	%	The proportion of big sagebrush canopy in each pixel.
shrubland_herbaceous_mean	national_land cover_database	2016	raster	%	The annual and perennial grass and forb proportion in each pixel.
shrubland_litter_mean	national_land cover_database	2016	raster	%	The dead plant material proportion in each pixel.
shrubland_sagebrush_mean	national_land cover_database	2016	raster	%	The proportion of sagebrush canopy in each pixel.
shrubland_sagebrush_height_mean	national_land cover_database	2016	raster	centimeters	Average height of sagebrush.
shrubland_shrub_mean	national_land cover_database	2016	raster	%	The proportion of shrub canopy in each pixel.
shrubland_shrub_height_mean	national_land cover_database	2016	raster	centimeters	Average height of shrubs.
percent_tree_cover_mean	national_land cover_database	2016	raster	%	Percent of the pixel that is covered by tree canopy.
usgs_gap100m_cat	national_terrestrial ecosystems	2011	raster	584 numerical values, indicating different landcover descriptions	
TPI-LOCAL	national_elevation database	Not applicable	raster	Unitless positive or negative score	Topographic Position Index (TPI) helps distinguish topographic features such as a hilltop, valley bottom, exposed ridge, flat plain, upper or lower slope. It is calculated by comparing the elevation of each pixel to its surrounding neighbours.
hftd100m_zone	high_fire threat_district	Jan-18	raster	Numerical indicators of the HFTD Tier	
elevation100m_m	national_elevation database	Not applicable	raster	meters (above sea level)	

pop_density100m_cnt_100m2	world_pop	2010 - 2020	raster	Count of people residing in each grid cell	
satellite_tree_height_max	vegetation_satellite	2019	raster	meters	
satellite_tree_count	vegetation_satellite	2019	raster	Count of trees in each 100m pixel that could strike conductors.	
satellite_tree_height_perlinemile	vegetation_satellite	2019	raster	Count of trees in each 100m pixel that could strike conductors, divided by the line miles present in each pixel.	
satellite_tree_height_avg	vegetation_satellite	2019	raster	meters	
fpi_fire_season_dfm_1000hr_avg	pge_meteorology	Jul-21	raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter), expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across fire season (June through Nov) for the years 2015-2020
dfm_100hr_month_1	pge_meteorology	Jul-21	raster	%	100 hour fuel moisture content - measure of the amount of water in "100 hour" fuel (downed logs and branches 1-3" in diameter), expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 1 (Jan) for the years 2015-2019
dfm_100hr_month_2	pge_meteorology	Jul-21	raster	%	100 hour fuel moisture content - measure of the amount of water in "100 hour" fuel (downed logs and branches 1-3" in diameter), and is expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 2 (Feb) for the years 2015-2019
dfm_100hr_month_3	pge_meteorology	Jul-21	raster	%	100 hour fuel moisture content - measure of the amount of water in "100 hour" fuel (downed logs and branches 1-3" in diameter), and is expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 3 (March) for the years 2015-2019
dfm_100hr_month_4	pge_meteorology	Jul-21	raster	%	100 hour fuel moisture content - measure of the amount of water in "100 hour" fuel (downed logs and branches 1-3" in diameter), and is expressed as a percentage of the dry

					weight of that fuel. Hourly metrics averaged across month 4 (April) for the years 2015-2019
dfm_100hr_month_5	pge_meteorology	Jul-21	raster	%	100 hour fuel moisture content - measure of the amount of water in "100 hour" fuel (downed logs and branches 1-3" in diameter), and is expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 5 (May) for the years 2015-2019
dfm_100hr_month_6	pge_meteorology	Jul-21	raster	%	100 hour fuel moisture content - measure of the amount of water in "100 hour" fuel (downed logs and branches 1-3" in diameter), and is expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 6 (June) for the years 2015-2019
dfm_100hr_month_7	pge_meteorology	Jul-21	raster	%	100 hour fuel moisture content - measure of the amount of water in "100 hour" fuel (downed logs and branches 1-3" in diameter), and is expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 7 (July) for the years 2015-2019
dfm_100hr_month_8	pge_meteorology	Jul-21	raster	%	100 hour fuel moisture content - measure of the amount of water in "100 hour" fuel (downed logs and branches 1-3" in diameter), and is expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 8 (Aug) for the years 2015-2019
dfm_100hr_month_9	pge_meteorology	Jul-21	raster	%	100 hour fuel moisture content - measure of the amount of water in "100 hour" fuel (downed logs and branches 1-3" in diameter), and is expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 9 (Sept) for the years 2015-2019
dfm_100hr_month_10	pge_meteorology	Jul-21	raster	%	100 hour fuel moisture content - measure of the amount of water in "100 hour" fuel (downed logs and branches 1-3" in diameter), and is expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged

					across month 10 (Oct`) for the years 2015-2019
dfm_100hr_month_11	pge_meteorology	Jul-21	raster	%	100 hour fuel moisture content - measure of the amount of water in "100 hour" fuel (downed logs and branches 1-3" in diameter), and is expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 11 (Nov) for the years 2015-2019
dfm_100hr_month_12	pge_meteorology	Jul-21	raster	%	100 hour fuel moisture content - measure of the amount of water in "100 hour" fuel (downed logs and branches 1-3" in diameter), and is expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 12 (Dec) for the years 2015-2019
dfm_1000hr_month_1	pge_meteorology	Jul-21	raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter), expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 1 (Jan) for the years 2015-2019
dfm_1000hr_month_2	pge_meteorology	Jul-21	raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter), expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 2 (Feb) for the years 2015-2019
dfm_1000hr_month_3	pge_meteorology	Jul-21	raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter), expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 3 (March) for the years 2015-2019
dfm_1000hr_month_4	pge_meteorology	Jul-21	raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter),

					expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 4 (April) for the years 2015-2019
dfm_1000hr_month_5	pge_meteorology	Jul-21	raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter), expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 5 (May) for the years 2015-2019
dfm_1000hr_month_6	pge_meteorology	Jul-21	raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter), expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 6 (June) for the years 2015-2019
dfm_1000hr_month_7	pge_meteorology	Jul-21	raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter), expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 7 (July) for the years 2015-2019
dfm_1000hr_month_8	pge_meteorology	Jul-21	raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter), expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 8 (Aug) for the years 2015-2019
dfm_1000hr_month_9	pge_meteorology	Jul-21	raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter), expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 9 (Sept) for the years 2015-2019
dfm_1000hr_month_10	pge_meteorology	Jul-21	raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter), expressed as a percentage of the dry weight of

					that fuel. Hourly metrics averaged across month 10 (Oct) for the years 2015-2019
dfm_1000hr_month_11	pge_meteorology	Jul-21	raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter), expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 11 (Nov) for the years 2015-2019
dfm_1000hr_month_12	pge_meteorology	Jul-21	raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter), expressed as a percentage of the dry weight of that fuel. Hourly metrics averaged across month 12 (Dec) for the years 2015-2019
fpi_v3_grass	pge_meteorology	Jul-21	raster	%	Fraction of fuel category in POMMS grid cell attributed to grass
fpi_v3_grass-shrub	pge_meteorology	Jul-21	raster	%	Fraction of fuel category in POMMS grid cell attributed to grass-shrub
fpi_v3_shrub	pge_meteorology	Jul-21	raster	%	Fraction of fuel category in POMMS grid cell attributed to shrub
fpi_v3_timber_litter	pge_meteorology	Jul-21	raster	%	Fraction of fuel category in POMMS grid cell attributed to timber litter
fpi_v3_timber_understory	pge_meteorology	Jul-21	raster	%	Fraction of fuel category in POMMS grid cell attributed to timber understory
fpi_v3_urban	pge_meteorology	Jul-21	raster	%	Fraction of fuel category in POMMS grid cell attributed to urban
fpi_v3_slope_degree_mean	pge_meteorology	Jul-21	raster		Slope of terrain averaged over POMMS grid cell.
fpi_v3_terrainrugged_mean	pge_meteorology	Jul-21	raster		Terrain ruggedness average in POMMS grid cell.
fpi_v3_aspect_most_common_angle	pge_meteorology	Jul-21	raster		
fpi_v3_R_score	pge_meteorology	Jul-21	Location time raster	Score of 1 to 5	
fpi_v3_prob_detect	pge_meteorology	Jul-21	Location time raster	%	
fpi_v3_prob_large	pge_meteorology	Jul-21	Location time raster	%	

fpi_v3_prob_catastrophic	pge_meteorology	Jul-21	Location time raster	%	
fpi_v3_prob_large_or_catastrophic	pge_meteorology	Jul-21	Location time raster	%	
fpi_v3_event_dfm_10hr	pge_meteorology	Jul-21	Location time raster	%	10 hour fuel moisture content - measure of the amount of water in "10 hour" fuel (downed logs and branches 0.25"-1" in diameter), expressed as a percentage of the dry weight of that fuel.
fpi_v3_event_dfm_100hr	pge_meteorology	Jul-21	Location time raster	%	100 hour fuel moisture content - measure of the amount of water in "100 hour" fuel (downed logs and branches 1-3" in diameter), expressed as a percentage of the dry weight of that fuel.
fpi_v3_event_dfm_1000hr	pge_meteorology	Jul-21	Location time raster	%	1000 hour fuel moisture content - measure of the amount of water in "1000 hour" fuel (downed logs and branches 3-8" in diameter), expressed as a percentage of the dry weight of that fuel.
fpi_v3_event_acc_precip	pge_meteorology	Jul-21	Location time raster		
fpi_v3_event_sfcdownshortwaveflux	pge_meteorology	Jul-21	Location time raster		
fpi_v3_event_temp2m_f	pge_meteorology	Jul-21	Location time raster	°F	
fpi_v3_event_rh2m	pge_meteorology	Jul-21	Location time raster		
fpi_v3_event_ws_mph	pge_meteorology	Jul-21	Location time raster	miles per hour	Wind speed at the surface.
fpi_v3_event_ws_mph_300m	pge_meteorology	Jul-21	Location time raster	miles per hour	Wind speed at 300 meters above the surface.
fpi_v3_event_wind_dir_f	pge_meteorology	Jul-21	Location time raster		Wind direction.
fpi_v3_event_lfm_chamise_new	pge_meteorology	Jul-21	Location time raster	%	Live fuel moisture content of Chamise (new growth) species.
fpi_v3_event_lfm_chamise_old	pge_meteorology	Jul-21	Location time raster	%	Live fuel moisture content of Chamise (old growth) species.
fpi_v3_event_land_use	pge_meteorology	Jul-21	Location time raster		
fpi_v3_event_ffwi	pge_meteorology	Jul-21	Location time raster		Fosberg Fire Weather Index
fpi_event_u10_ms	pge_meteorology	Jul-21	Location time raster		

fpi_event_v10_ms	pge_meteorology	Jul-21	Location time raster		
fpi_event_dfm_1hr	pge_meteorology	Jul-21	Location time raster	%	1 hour fuel moisture content - measure of the amount of water in "1 hour" fuel (fine dead fuels up to 0.25" in diameter, such as grass), expressed as a percentage of the dry weight of that fuel.
fpi_event_fpi	pge_meteorology	Jul-21	Location time raster	Score of 1 to 5	

6.5 Model Covariates Offered to P(i|o)

The following table lists the covariates used by the WDRM v3 P(i|o) model

'ign_ind ~ fpi_v3_urban + subset_name + fpi_v3_grass + impervious_mean + fpi_v3_shrub + fpi_v3_timber_understory + fpi_v3_timber_litter + hftd + fpi_v3_event_ws_mph + fpi_v3_event_dfm_10hr + fpi_v3_event_dfm_1000hr

7 Appendix: Special Topics

7.1 Wind

7.1.1 Overview

The ingredients in dangerous wildfire are, in order from least to most ephemeral, favorable topography (fire burns better uphill), the presence of fuels, low fuel moisture, and winds. For the most part, topography doesn't change; outside of wildfire consuming them, fuel levels are fairly stable over whole fire seasons at least; and fuel moisture is a lagged effect of long term exposure to hot/dry air – 100 hour fuels are lagged over 4 days and 1000 hour fuels are lagged over 40 days – and is primarily a seasonal metric, with lower fuel moisture in more locations later in the dry season (end of summer and beginning of fall all the way until the first soaking rains). Wind, however, operates on much finer timescales. For this reason, wind can be considered the final, "right now" condition that makes bad wildfires worse and catastrophic wildfires possible.

The role of wind in wildfire is complex. Hot dry winds at high speeds can be the primary driver of breathtaking runs of fire spread – not just by driving fire to spread directly from a given location to neighboring locations downwind, but also by driving embers that can transport fire much faster than contiguous spread. However, as any camper can tell you, elevated wind speeds can also snuff out small fires before they are large enough to spread, and as recent experience with megafires has shown, wind need not be a pre-existing condition: sufficiently large fires create their own wind though superheated updrafts that support the formation of massive pyro-cumulus clouds while drawing in ground level replacement air – causing winds to flow into the fire, sometimes from all directions. On top of all this, the lightning strike fires of 2020 and other examples like the Dixie Fire of 2021 and many before it have illustrated the degree to which established fire can persist for days and weeks with little to no wind *only to spread explosively when days of extreme wind do occur.* On calm days, fires are far more manageable and, in a wildland setting, can even be beneficial to landscapes that co-evolved with fire. But on extreme wind event days, fires can be unpredictable, impossible to control, and devastating to everything in their path. The longer a fire burns, the more likely it is to experience such winds and therefore the more likely it is to escape control, with potentially catastrophic effects.

For all of the above reasons, the most dangerous fire conditions occur where there has been significant buildup of low moisture fuels dried by preceding hot, dry weather during periods of high sustained winds. Any grid events that occur under these conditions are high risk because they satisfy all of the pre-conditions for dangerous fire and are also guaranteed to experience the most ephemeral one: wind. In recognition of such risk, these conditions have become the criteria for PG&E's PSPS program, which turns off grid assets, and EPSS program, which tunes the response of protective devices to aggressively trigger outages without the "recloser" behaviors that can automatically recover from transient faults.

7.1.2 Considerations when modeling with wind

When modeling ignition events and their consequences, there are several wind-related factors to consider, and judgement is required:

- When modeling on annual or longer timeframes, it is inevitable that dangerous winds will be present some of the time and such winds make new and existing fires more dangerous. So the ignition consequence can be thought of as tied to the likelihood that a fire will be exposed to dangerous winds during its lifetime. In the WDRM v3 this is captured through the FPI R score thresholds in the consequence model (R Scores factor in weather), wind data from the day of events used to train the P(ignition|outage) model, and in the percentage of days with elevated winds in the P(outage) models.
- Some grid events are more likely to be associated with dangerous conditions. For example, vegetation-caused events are quite likely to be in the presence of fuels and some are caused by wind, which also factors into dangerous conditions. In the WDRM v3, this is captured in the independent fit of cause and sub-cause-categorized subsets and in the subset-specific terms in the P(ignition|outage) model, where the strength of wind importance varies substantially by subset.
- In locations that are always windy, winds strong enough to cause branches or grid equipment to break could happen on any given day – they are statistically more likely to occur under fuel and moisture conditions not conducive to extreme wildfire. Further, trees are smaller and stouter and grid assets built to a more stringent specification in locations with known and consistent stresses. We might say that the wind "harvests" the faults more steadily, and the resulting repairs and corrections make the grid (and vegetation) more resilient to wind. In that sense, unusually strong winds can be more dangerous than consistently strong winds. In the WDRM v3, this is accounted for by the derivation of wind covariates that capture the percentage of days that witness winds above specific threshold values and by providing min and max wind speeds as well as their average over time.
- PSPS and EPSS are not perfect, but they are greatly reducing the possibility of ignitions under the most extreme fire conditions. In the WDRM v3, the lack of new data from PSPS and EPSS events is accounted for by treating PSPS Hazards and Damages as expected outages, and as ignitions weighted to reproduce historical annual totals consistent with pre-PSPS counts under PSPS conditions.
- Also resulting from the effects of PSPS and EPSS events, fewer of the remaining ignitions originate during extreme fire weather conditions. This circumstance has underscored the potential for destructive fire behavior from fires that evade suppression long enough to be exposed to high winds that drive runs days or even weeks after their initial ignition. Examples including the lightning-caused LNU, CZU, SCU, and August Complex fires from 2020, the Dixie Fire, and many others. The potential for dangerous fire weather in the days that follow ignitions is accounted for by identifying destructive potential using simulations trained on worst fire weather days and by calibrating consequence against VIRRS final wildfire size and fire agency data on consequences.
- Because extreme events are comparatively rare in time and space, ignitions under such conditions tend to lack statistical power. This means that models fit to only such events produce estimates with very large uncertainties and risk over-fitting the data. However, the failure process behind most ignitions is often the same on extreme fire risk days or ordinary days. It is therefore possible to study those processes under the assumption that they could have been dangerous but for the lack of extreme weather and with the knowledge that mitigations effective against one will be effective against the other. By recognizing that many ignitions share root causes with those under the most dangerous conditions models trained on a broader sample of events and ignitions can establish greater statistical power.

7.1.3 The role of wind in the WDRM v3

The WDRM v3 is not a monolithic model. Rather it estimates risk based on three separate estimation steps:

- The probability of outages given asset and environmental conditions
- The probability of ignitions given an outage and its cause, location, and other characteristics
- The consequences of an ignition, given its location of origin.

7.1.3.1 Tabulation of wind conditions for modeled events

Wind is highly turbulent and interacts with the ground and surface features. No matter the spatial scale of measurements, there will always be unobserved structure to wind speeds. Wind data consumed by the WDRM v3 is thus a combination of ground truth measurements and weather model simulation. In this case, the PG&E meteorology department produces 2km gridded wind data on an hourly timescale. Gusts often occur when turbulent flow brings unobstructed wind all the way to the ground, and wind events are often associated with phenomena that drive higher altitude flows downward for sustained periods. For these reasons, meteorology data on winds above the ground is a useful proxy for the gusts that may be experienced on the ground. To break down windy vs. not windy conditions (just for the purposes of this discussion) we define “high wind” as hourly wind speed values exceeding 35 mph at 300m above the ground. Target set events occurring on days where their (2km grid) location experienced high winds for at least one hour are labeled “high wind” and all others are labeled “low wind” although “not high wind” is probably more accurate a name.

The count of high and low wind events with different groupings are tabulated below.

	low_wind	high_wind	low_wind_pct	high_wind_pct
No ignition	95,919	17,788	84.36	15.64
Ignition	2,757	345	88.88	11.12

Table 19: High and low wind event counts by ignition status

By failure type:

	low_wind	high_wind	low_wind_pct	high_wind_pct
Damage	472	549	46.23	53.77
Hazard	205	246	45.45	54.55
forced_outage	17,163	5,847	74.59	25.41
Ignition_no_outage	1,223	130	90.39	9.61
outage	79,613	11,361	87.51	12.49

Table 20: High and low wind event counts by event type

The WDRM v3 is focused on disaggregating all events into subsets with shared causes and sub-causes. The two figures below illustrate the fraction of outages (first one) and ignitions (second one) in each subset experiencing high winds.

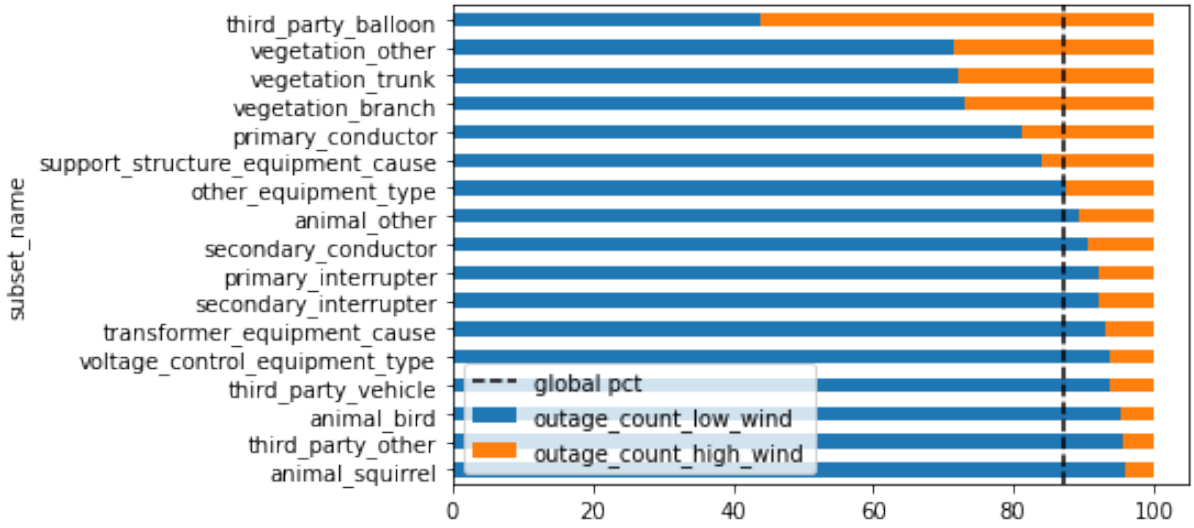


Figure 49: Fraction of outages on high or low wind days, by WDRM v3 subset. Note that third_party_balloon has very few outages and its prominence may be a statistical fluke

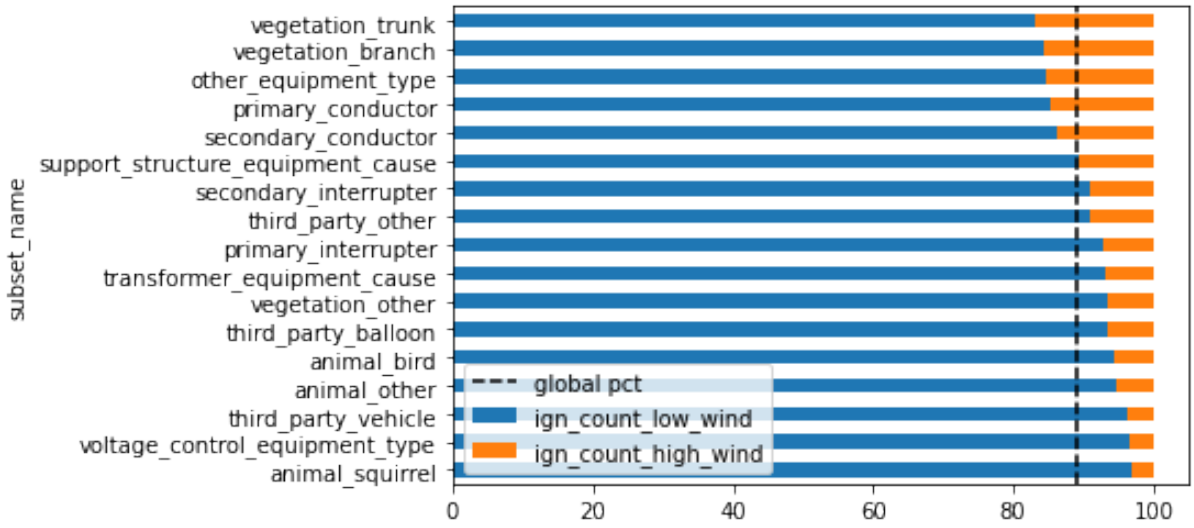


Figure 50: Fraction of ignitions on high or low wind days, by WDRM v3 subset

7.1.4 Role of wind in P(outage)

The importance of wind covariates to each of the MaxEnt P(outage) models can be compared using covariate jackknife results. Jackknifing is the process of isolating one covariate at a time, re-running a given model without the covariate, and running a model fit on only that covariate to measure how much worse the model is without the covariate (loss) and how good the model with only the covariate is (gain). The figures below illustrate the loss and gain from wind covariates for all MaxEnt subsets. The variability in loss/gain the figures illustrate confirms that wind plays a different role in different causes of outages and that the methods employed to predict P(outage) are capturing this differentiation.

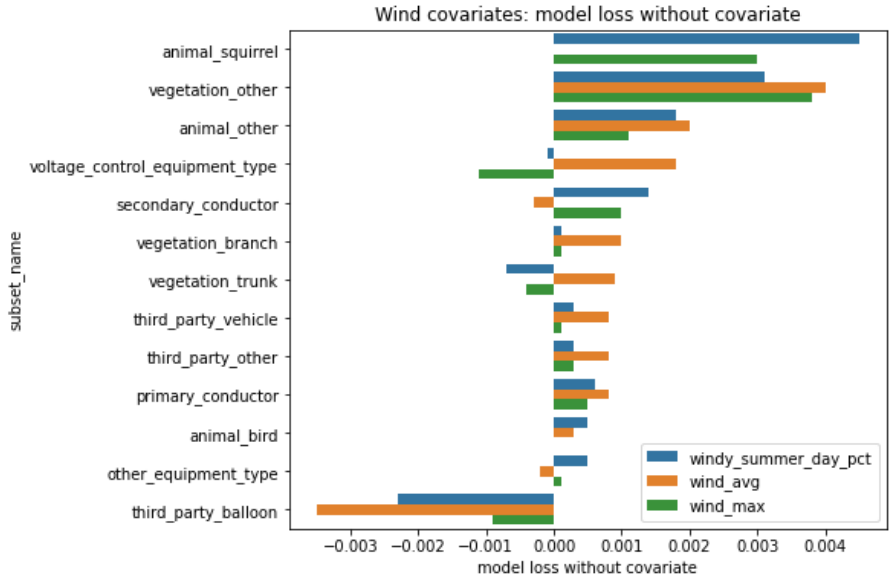


Figure 51: Loss of predictive performance for each subset missing each wind covariate

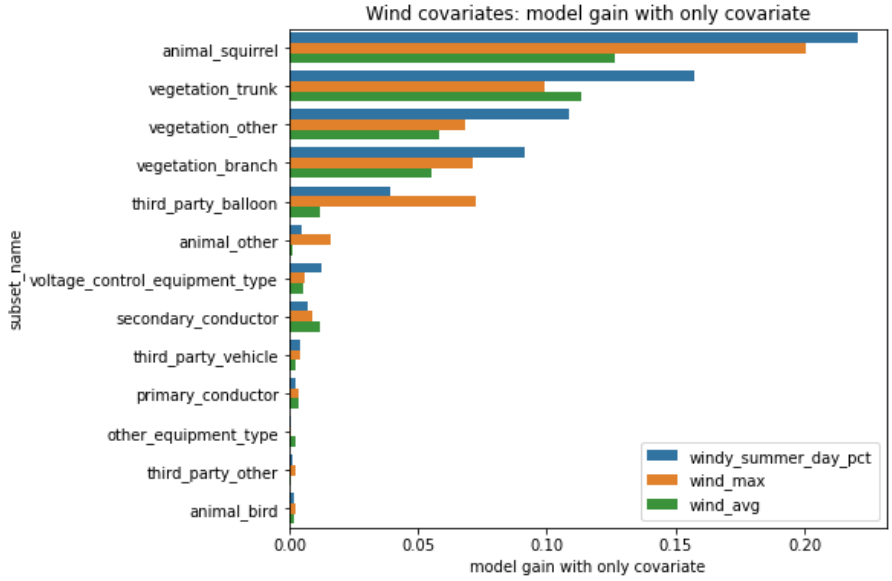


Figure 52: Predictive performance of each wind covariate alone for events drawn from each subset

7.1.5 Role of wind in P(ignition|outage)

The P(ignition|outage) model is trained on the event target set of data, which includes all outage, ignition without outage, and PSPS Hazards and Damages events, with cause, sub-cause, location-specific environmental conditions, and weather data covariates explaining/predicting 0/1 ignition indicator labels. In the case of PSPS Hazards and Damages, actual ignition outcomes are unknowable, but all are labeled with a 1 ignition indicator, but weighted to produce expected ignition counts consistent with ignition rates observed under PSPS conditions before such events were called.

Since it trains on event data, the P(ignition|outage) model consumes day-of event wind speed data – daily maximum wind speeds at 10m altitude. The model exhibits strong subset-level fixed effects (the faults associated with different causes and sub-causes have inherently different energetics and pathways to ignitions). Among the spatially varying covariates, wind speed are the third most important (as measured by influence over average precision) behind only 10hr fuel moisture and HFTD tier indicators, each related to the pre-conditions of fire.

The output of the P(ignition|outage) model are predictions for all grid locations with daily temporal resolution due to fuel moisture and wind covariates. However, the WDRM v3 predicts typical fire season risk values. For compatibility with the WDRM, the P(ignition|outage) model is used to predict daily values across multiple years of historic fire season weather and fuel conditions, with the daily estimates aggregated into seasonal predictions. The aggregation process weights each day's contribution to the aggregate by *its observed count of outages*. In this way, *the seasonal estimate automatically emphasizes the conditions associated with elevated outage rates*, capturing their outsized importance in determining ignition probabilities. For example***, the elevated rate of vegetation caused outages associated with late summer winds leads to those late summer conditions factoring more heavily into the seasonal P(i|o) for vegetation-caused outages***.

7.1.5.1 Role of wind in Consequence

The most well-known role that wind plays in consequence is that the Tecnosylva fire simulations at 200m locations along the grid are performed for hundreds of historically dangerous fire weather conditions. The resulting fire size, rate of spread, and flame length from these simulation results are used to estimate the potential for ignitions in each location to become destructive. The more consistently severe the simulation results are across simulated days, the greater the estimated consequence.

However, that is only part of the story for the WDRM v3. The WDRM v3 also considers what is known as the v3 FPI R score when determining destructive potential. The R score, which ranges from 1-5 is a measure of how likely a fire detectable from space (via the VIRRS satellite) is to spread into large or destructive size categories. The v3 FPI model, developed to support the process of forecasting and calling PSPS events and therefore well-tuned to the "right now" conditions that govern the severity of wildfire, is trained on fuels and hourly wind and weather data and delivers R scores on an hourly basis for all grid locations. The WDRM v3 consequence model is supported by historical R scores for all fire season days 2015-2020, where ignitions on any day/location experiencing an R score of 4 or higher is also classified as having destructive potential.

8 Appendix: Methodological Details

8.1 MaxEnt

8.1.1 Overview

To answer the question of where outage events for each subset are likely to occur, we have estimated fire season outage probabilities using maximum entropy models (as implemented in software called MaxEnt[4]) developed by computer scientists at AT&T research and Princeton University for the modeling of ecological ranges of species[5]. These models are trained on outage locations and gridded spatial (raster) environmental and asset attribute data for all "grid pixels", which are defined as raster pixels that host distribution grid assets on a 100mx100m scale. Event locations are considered "presence" observations and the set of all grid pixels are considered the "background". To accrue enough data to identify spatial patterns in the events, event training data typically spans several years of data collection. Although it is possible to filter training data by coarse temporal criteria, the model is dedicated to spatial, not temporal, patterns. The so-called "raw output" of MaxEnt models provides relative scores for fire-season outages per "grid pixel" of input data. Given know outage rates, the raw outputs can be calibrated through a logistic mapping with free parameter, τ, into annual fire season probabilities of outages for all grid pixels.

8.1.2 How MaxEnt frames estimation problems

We are interested in which environmental conditions and asset attributes (collectively called the model covariates) are more common (or less common) among outage locations (presence samples) than they are among all distribution grid locations (background samples). For example, tall trees are more common among vegetation-caused outage locations than they are among typical distribution grid locations. Metrics of fuel dryness, HFTD tier assignments, prevailing weather conditions, prevailing conductor materials and size, land use types, and others, can all be examined for such patterns. The ratio of the distribution of covariate value prevalence at outage locations to their prevalence across all grid locations is called the relative occurrence rate.

MaxEnt provides a way of estimating the relative occurrence rate given a fairly modest number of outage locations. The way it does this is to fit a statistical distribution of covariate values for outage locations that is consistent with the values at known outage locations, but otherwise as similar as possible to the distribution of values found everywhere else along the distribution grid. The similarity criteria is enforced using a metric called the *relative information entropy* between the distributions of covariate values associated with outage locations and the distribution grid locations, where the larger that metric is, the more similar the two distributions are. For this reason, the overall approach is referred to as a *maximum entropy* estimation of the relative occurrence rate.

Because historic outage counts are available, the relative occurrence rate can be scaled via a logistic transform with τ scale parameter ensuring expected counts match known outcomes. This results in an estimate of the annual probability an outage for all combinations of covariate values – the calibrated model fit. The fit can then predict outage probabilities based on the covariate values found at each distribution grid location.

[4] https://biodiversityinformatics.amnh.org/open_source/maxent/Maxent_tutorial2017.pdf

[5] Note that "maximum entropy" can be the optimization target for any number of models. The specific approach used in the v3 WDRM, which applies a maximum entropy criteria to the relationship between presence observation locations and all locations, is implemented in software called MaxEnt.

For more technical details of the math and ideas behind MaxEnt, readers are directed to:

- The original papers explaining the implementation of MaxEnt
  - http://rob.schapire.net/papers/maxent_icml.pdf
  - http://www.bio-nica.info/biblioteca/phillips2006maximumentropy.pdf
- An explanation that places MaxEnt more in the language, notation, and context of statistical learning methods
  - https://hastie.su.domains/Papers/maxent_explained.pdf
- Because the "maximum entropy" constraint has been broadly applied to many unrelated problem domains, technical readers should not assume that other search results for "maximum entropy models" and similar are directly related to these methods.

8.1.3 Reasons for choosing MaxEnt

One prominent distinction between MaxEnt and more traditional Machine Learning classification methods is that MaxEnt is what is called an "presence only" model. This means that MaxEnt can be trained on locations with known events without requiring ground truth absence data. Instead, event locations are compared to all locations. At first glance, this might seem like an odd choice – whether there is an event or not on the grid is a clear criterion – however, there are several aspects of the wildfire modeling problem that benefit from the choice.

The following aspects of MaxEnt make it a logical choice for the P(outage) modeling of the v3 WDRM.

1. Location data in imprecise: The most concrete and reliable location data available for each outage is the location of the protective device whose activation triggered the outage. However, the root cause of outage producing conditions is by definition "downstream" from protective devices, often without any way to know the exact location. P(outage) models must be robust to these locational uncertainties.
2. The absence of an outage does not guarantee an absence of an ignition. We are ultimately using outages as proxies for ignitions and there are ignitions that occur without outages.
3. Mitigation efforts, including enhanced vegetation management, grid hardening, PSPS events, and EPSS settings are implemented in the places (and times) where the conditions that produce outages and ignitions are deemed most risky. As a result, mitigation causes "negatives", i.e. avoided outages, where untreated conditions would tend to produce them. The modeling goal is to capture untreated risks so the benefits of mitigation can be assessed, so we do not want the models to mistake the correlation of suppressed outages with treatments for a lack of risk in treated locations.
4. Many degradation processes and proximate outage causes are environmentally driven. Environmentally driven failure processes benefit from support for spatial/raster data sets.

5. Asset attributes are not always fully known. A logical alternative to modeling spatial locations is to train models that attribute failures to the state of individual assets at their time of failure. By nature, such models require high quality asset attribute data spanning the same timeframe as event data and are best suited to equipment-caused outages, where internal degradation and failures of specific assets were the drivers. The v3 WDRM does fit asset-specific models to support structure and transformer failures, but each required labor intensive reconstructions of subset-specific probable asset attribute data sets spanning years, and the corresponding data for other asset types is either unavailable, less mature, or only incidental to the failure mode in question. Where detailed asset attribute data is either unavailable or incidental to the failure mode, a spatial model like MaxEnt is necessary.
6. MaxEnt (the software) is not just an implementation of the estimation algorithm. It includes support for spatial data sets and outputs, feature generation from covariates – i.e. interactions, hinges, thresholds, etc., regularization that adaptively down weights features that lack or degrade predictive power, and support for the production of train and test performance metrics. It is implemented as open-source software with a robust community of users and peer review of its many applications. All of these combined mean that there is a built-in audience that understands exactly how the models work and how to interpret their results.

8.1.4 Implementation

1. Data preparation
    a. Rasterize asset attributes – given assets with both geometries and attributes characterizing them, the attribute values are “burned to raster” using the geometries to assign them to each grid pixel. Note that the attribute values of several assets must be aggregated to assign a prevailing value at the pixel level and that the raster pixels with values are, by definition, the grid pixels for the corresponding assets. The v3 WDRM rasterizes asset data for primary conductors, secondary conductors, primary and secondary conductors, poles, and transformers.
    b. Prepare event data – a data set of outages, PSPS hazards and damages, and ignitions without outages, with ignition indicators and locations is the training event data for a MaxEnt model, where the locations are mapped to grid pixel locations as “presence” observations. The event data can be split into train and test samples and/or into leave-one-out folds for use in cross validation.
    c. Derive weather and other time varying data set extracts – covariate data should represent the conditions under which the events occurred. For multi-year event data sets, that means statistical extracts of multi-year sequences of weather data. Weather data, including wind speed and fuel moisture is aggregated into average values, but also percentile values (1% considered as a minimum and 99% considered the max) and counts of days crossing thresholds in wind speed and fuel moisture related to vegetation stress and fire risk.
    d. Prepare all covariates as raster data – all raster data, including asset data, aggregated weather data, and spatial attributes like tree height, vegetation types, land use categories, soil types, etc. are masked so only values for grid pixel locations are made available to the MaxEnt models and thus the probabilities are only estimated for grid pixels. Note that the grid pixels are specific to a given asset type, and thus asset types are associated with the data preparation for each MaxEnt model specification.

2. MaxEnt invocation (Java process invoked from python through the ccb open source library - https://github.com/stanford-ccb/ccb)
    a. A MaxEnt model fit is invoked through configuration options that
        i. Provide data files containing training and test event data
        ii. Provide raster files for all covariates – masked to include only grid pixel locations
        iii. Specify options for generating derived interaction, threshold, and hinge features from covariates
        iv. Specify options for the strength of regularization of feature selection
        v. Control whether to additionally perform jackknife model runs that systematically isolate or eliminate each covariate, recording the resulting (degraded) model performance compared to the full set of covariates
        vi. Provide a set of covariate raster files for use in out of sample (e.g. covariates drawn from other time periods) prediction
    b. Logistic form outputs (as raster data) are scaled to expected annual fire season counts of outages using the τ parameter
    c. Outputs include
        i. A tif of raster data for P(outage) predictions for all grid locations
        ii. Jackknife results
        iii. Model fit parameters
        iv. Model performance metrics, for both train and test event data
        v. Event data used to train the model with labels on events that were dropped due to insufficient covariate data
3. Multiple model fits
    a. If necessary, several MaxEnt models are executed for each subset. These can include a train/test run that sets aside data for testing predictive performance and flagging indications of over-fitting, a “standard” run that trains on all available data, jackknife runs with one run that isolated and one run that drops each covariate (2x N covariates are run), K-fold cross validation runs using random folds or leave-one-out folds based on year or other categorical data.
    b. The standard run fits, trained on as much event data as possible, are used for official predictions, with performance reported for test data from trial/test runs.

8.2 Asset Time Series

8.2.1 Introduction

The purpose of the asset equipment failure modeling approach is to reproduce the historical layout of the assets on the distribution system to reveal the asset attributes at the time equipment failure occurs. An asset is often removed from service after a major equipment failure event, so it is critical to develop a dataset that can connect the historical conditions of the grid to historical equipment failure events to identify patterns in the types of assets that fail. Thus, the asset equipment failure models were modeled and then forecasted utilizing an annual approach and then employing a Random Forest algorithm to predict the annual probability of a support structure or transformer unit equipment failure for an individual support structure or transformer unit currently installed on the distribution system within PG&E's service territory.

The support structure includes the components of a pole that support the overhead system: the pole, crossarm, guy or anchor wire, insulators, and supporting hardware and connections. A support structure equipment failure can include the malfunction of one or more of these components. The model predicts the annual probability of an equipment failure per pole or support structure (i.e. a prediction is provided for each support structure based on that support structure's attributes). The support structure model explores wood distribution poles, limiting predictions to those poles constructed with wood, throughbore, or centerbore material. Each structure is tracked in the PG&E data model by a primary key, the SAP equipment id.

Transformers are nodes on the grid (identified by a California Grid Coordinate - cgc12) containing 1-3 transformer units (identified by an SAP or GIS equipment id). When any unit on a transformer fails, they are typically all replaced. Thus, planning focuses on the transformer level, but failures occur at the transformer unit level. The RaDA team provide predictions at the transformer unit level.

8.2.2 Model dataset

8.2.2.1 Equipment failure events

The categories used to filter outage events into the appropriate asset equipment failure model are shown in **Figure 1**. Outage events were first filtered to the events caused by equipment failure. Then those events were split out according to the type of equipment involved (support structure or transformer). Those were each further split up according to sub-cause categories.

- Support structure equipment failure events with sub-cause categories related to pole failure or other structural failures were allocated to the *support_structure_equipment_cause* model.
- Support structure equipment failure events with sub-cause categories related to pole tracking fires or electrical-related outage events were allocated to the *support_structure_equipment_electrical* model.
- Transformer equipment failure events with sub-cause categories related to transformer overload or transformer failure were allocated to the *transformer_equipment_cause* model.
- Transformer equipment failure events with sub-cause categories related to leaking transformer were allocated to the *transformer_equipment_leaking* model.

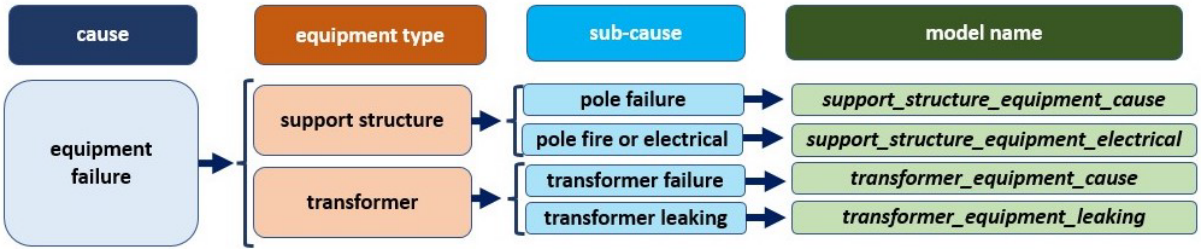


Figure 53. Categories used to filter events into the appropriate asset equipment failure model

8.2.2.2 Asset dataset structure

The asset equipment failure models utilize PG&E's available historical data to develop annual current and historical representations of the asset on the distribution system. The following current and historical data is used:

1. Current EDGIS asset dataset
2. Annual snapshots of historical EDGIS asset dataset
3. Asset replacement mapping.

The method to trace the history of each asset type varies in PG&E's current data model. Generally, the asset replacement mapping, which links an existing asset primary key to the prior asset primary key that was removed from service, can be used to identify the primary key for a removed asset. The historical EDGIS snapshots, which were extracted from backed-up versions of EDGIS for the following years: 2015, 2017, 2018, 2019, 2020, and 2021, can be used to identify the attributes of an asset removed from service. The historical view of an asset type is transformed to a dataset where each row in the historical asset dataset represents an "asset-year", or the asset's primary key and associated attributes for that year. The asset-years are traced back from 2022 to 2015.

8.2.2.3 Joining equipment failure events to asset-years

Historical equipment failure events are cleaned, filtered, and joined to the asset-year dataset by the asset's primary key and the year in which the equipment failure event occurred. There were several notable limitations in joining the event data:

- For support structures, many events didn't include an asset primary key for the equipment failure event. If the primary key couldn't be identified, then the event was dropped during this process resulting in an undercounting of events.
- For transformers, joining the event set with the covariate set is a challenge because the event set is based on transformers whereas the covariate set is based on the transformer unit. A given equipment failure may appear in multiple rows when there are multiple units installed at the same transformer resulting in an undercounting of events. Transformers may have up to three units installed, and records are not kept on which of the units caused the outage.
- For transformers, the asset primary key and CGC12 were only located for outage events, resulting in an undercounting of the other event types (i.e. non-outage ignitions and PSPS damages and hazards).

Ultimately, the four training data set count of events is included in **Table 3** and is summarized as:

- 4,631 *support_structure_equipment_cause* equipment failure events
- 2,096 *support_structure_equipment_electrical* equipment failure events
- 8,809 *transformer_equipment_cause* equipment failure events
- 1,126 *transformer_equipment_leaking* equipment failure events

An indicator column is created representing if there were one or more asset equipment failures during the wildfire season that asset-year (1) or if no equipment failure events were recorded during wildfire season that year (0).

An example view of the resulting dataset for a transformer unit is shown in **Table 1** and an example view of the resulting dataset for a support structure is shown in **Table 2**. **Figure 2** displays the general process flow for an asset equipment failure model data pipeline. More detailed information about developing the support structure model time series dataset can be found in the support structure technical documentation. More detailed information about developing the transformer unit model time series dataset can be found in the transformer technical documentation.

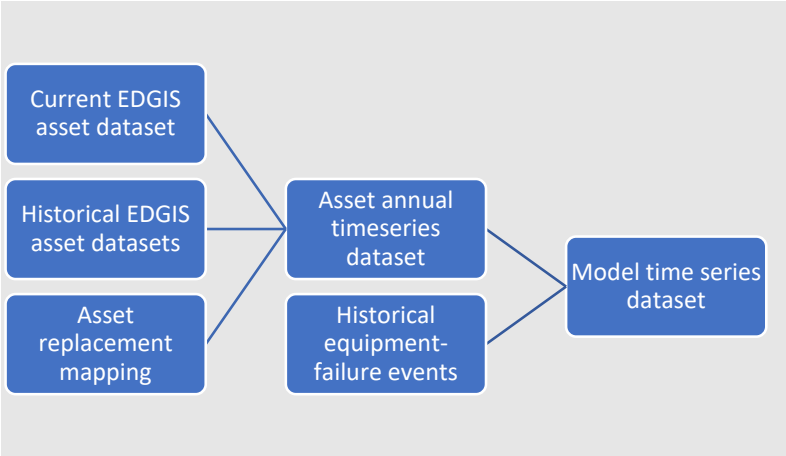


Figure 54. General data pipeline flow chart to develop the model time series dataset

Table 21. Example of the transformer unit asset-year dataset. Note that the old transformer failed in 2019 and was removed from service. Then a new SAP equipment id with the same CGC12 is used for the years following the failure.

CGC12	SAP equipment id	year	Age (years)	Equipment failure indicator
114303444276	103876345	2015	41	0
114303444276	103876345	2016	42	0
114303444276	103876345	2017	43	0
114303444276	103876345	2018	44	0
114303444276	103876345	2019	45	1
114303444276	226937054	2020	1	0
114303444276	226937054	2021	2	0
114303444276	226937054	2022	3	N/A

Table 22. Example of support structure asset-year representation. Note that the old pole failed in 2019, and a new SAP equipment id is used for the years following the failure. The global id in the asset-year dataframe always represents the current globalid.

global_id	SAP equipment id	year	age (years)	Equipment failure indicator
{D18CBF1C-64EB-4DA0-8513-0E874FF63332}	103876345	2015	63	0
{D18CBF1C-64EB-4DA0-8513-0E874FF63332}	103876345	2016	64	0
{D18CBF1C-64EB-4DA0-8513-0E874FF63332}	103876345	2017	65	0
{D18CBF1C-64EB-4DA0-8513-0E874FF63332}	103876345	2018	66	0
{D18CBF1C-64EB-4DA0-8513-0E874FF63332}	103876345	2019	67	1
{D18CBF1C-64EB-4DA0-8513-0E874FF63332}	226937054	2020	1	0
{D18CBF1C-64EB-4DA0-8513-0E874FF63332}	226937054	2021	2	0
{D18CBF1C-64EB-4DA0-8513-0E874FF63332}	226937054	2022	3	N/A

8.2.2.4 Model train, test, predict dataset splits

Once a model dataset is prepared, the dataset is split into a training dataset, a test dataset, and a dataset used to forecast probabilities for 2022. The asset-year dataset rows for 2015 to 2021 are used for the train and test datasets. The asset-year dataset rows for 2022 are extracted into a separate dataset used make predictions for the likelihood of equipment failure during 2022 wildfire season.

The support structure asset type models use a train and out-of-sample test datasets that are randomly split across years into 25% for the test dataset and 75% for the train dataset. A stratification of the split is used to ensure historical outage events are included in the test dataset despite the imbalance in the target data. The test dataset is set aside to evaluate the model's performance on an out-of-sample test dataset. The remaining 75% training dataset is used to fit the model algorithm to the equipment failure events.

The transformer asset type models use a train and out-of-sample test dataset that are split by month in order to preserve the temporal isolation among the transformer equipment failure events. The test dataset includes the events that occur during the months of July and October in addition to 1/3 of the transformer units that didn't experience an equipment failure event during the wildfire season. The train dataset includes the events occurring during the remaining wildfires season months (i.e. June, August, September, and November) as well as the remaining 2/3 of transformer units that didn't experience an equipment failure event.

Table 23. Model training dataset size summary

	Model training data event counts		
Model name	**Training events**	**train**	**Test**
support_structure_equipment_cause	4,631	1,363	518
support_structure_equipment_electrical	2,096	854	284
transformer_equipment_cause	8,809	5,102	4,751
transformer_equipment_leaking	1,126	626	599

8.2.3 Feature summary

8.2.3.1 Feature selection

Covariates were initially chosen based on discussions with SMEs about common equipment failure causal pathways. The covariates were then fed into the Random Forest algorithm and were manually selected through an experimentation process based on the feature importance scores described in the next section: feature permutation importance and the feature Gini importance score.

- For support structures, the model features were not highly correlated and included:
  - *pole attributes:* age, volume (i.e. height & circumference), material, treatment type, structural loading, last invasive inspection, and prior year maintenance tags.
  - *environmental:* count of strike trees, population density, soil erosion, soil annual flood frequency, surrounding topography, distance from the coastline, and contours running perpendicular to the coastline.
  - *weather*: cumulative precipitation and maximum wind speed during wildfire season.
- For transformers, the model features included:
  - *transformer unit attributes:* age, rated kVA, manufacturer, whether the transformer was manufactured after 2012, load factor, mean peak kVA divided by rated kVA, and whether the transformer was in the coastal corrosive zone.
  - *environmental:* population density and non-burnable land index.
  - *weather:* mean daily maximum temperature during wildfire season from 2015-2020.

The features used for each model asset type are listed in **Table 4**. More detailed feature descriptions are provided in the respective model technical documentation.

Table 24. Support structure feature definitions

Support structure Feature name	Feature description
Asset attributes	
age_years	Pole estimated age
pole_volume	Pole volume calculated as a cylinder using original circumference and height – essentially the pole class.
prior_year_open_tag_flag	Whether there was at least one open maintenance tag in the prior year, only considering B, E, F, and H tags.
percentatmcu	Structural loading descriptor indicating the estimated percent at the design loading capacity.
years_since_last_ptt_inspection	Number of years since the last recorded invasive pole test & treat (PT&T) inspection
is_throughbore	Whether the wood pole is constructed as through bore
is_cellon_treated	Whether the pole's original treatment type was Cellon
Environmental	
population_density	Estimate of population density within a 100m x 100m pixel
local_topography	
merged_strike_tree_count	Number of trees estimated to be within striking distance based on both LiDAR and satellite tree data.

soil_annual_flood_freq	USGS soil annual flood frequency
soil_erosion_kffact	USGS soil erosion K factor
Meteorological	
ws_mph_summer_max	Maximum windspeed during wildfire season (Jun-Nov)
acc_precip_summer_sum	Cumulative precipitation during wildfire season (Jun-Nov)
Spatial	
contour_perpendicular_to_coastline	Contour perpendicular to coastline
contour_parallel_to_coastline	Distance from coastline

Table 25. transformer unit feature definitions

Transformer feature name	Feature description
Asset attributes	
age	Transformer unit estimated age
rated_kva	Rated kVA
manufacturer_code_clean	Manufacturer
manufactured_in_2012_or_after	Manufactured year >= 2012
txfmr_load_factor_12m_bucket	Load factor
txfmr_mean_peak_kva_rated_capacity_factor_12m_bucket	Mean peak kVA / rated kVA
coastal_corrosive	Coastal corrosive
Environmental	
population_density	Population density
nonburnable_land_index	non-burnable land index
Meteorological	
daily_max_temperature_avg	mean daily max temperature during wildfire season from 2015-2020
manufactured_in_2012_or_after* daily_max_temperature_avg	Cross term using manufactured year >= 2012 and mean daily max temperature during wildfire season

8.2.3.2 Feature importance

Feature importance for the asset equipment failure models was measured in two ways:

1. feature permutation importance measured with average precision and
2. feature Gini importance score.

Feature permutation importance was implemented using the scikit-learn Python package's *permutation_importance* method and is defined as the average precision that is lost in the model performance when that feature's values are randomly shuffled around. More technical details can be found in the Python package documentation. The permutation importance scores for each support structure model are included in **Figure 3**, and the permutation importance scores for each transformer model are included in **Figure 5**.

Feature Gini importance was implemented using the scorer included in the Random Forest Classifier scikit-learn Python package. This scorer is specific to tree-based models and captures the helpfulness of the feature to the model by quantifying the feature's influence on the tree splits. The Gini importance is calculated for each node

and then aggregated to each feature for a decision tree. Then the feature importance for each decision tree in the Random Forest is averaged across all trees in the Forest. This is a standard method for calculating feature importance for Random Forest models, and additional technical details and mathematical implementation are widely available online. The feature Gini importance scores for each support structure model are included in **Figure 4,** and the feature Gini importance scores for each transformer model are included in **Figure 6**.

Asset attributes were particularly important when modeling equipment failure events. The asset's age was the most important feature in each of the equipment failure models. Information about the asset's loading was also a top feature.

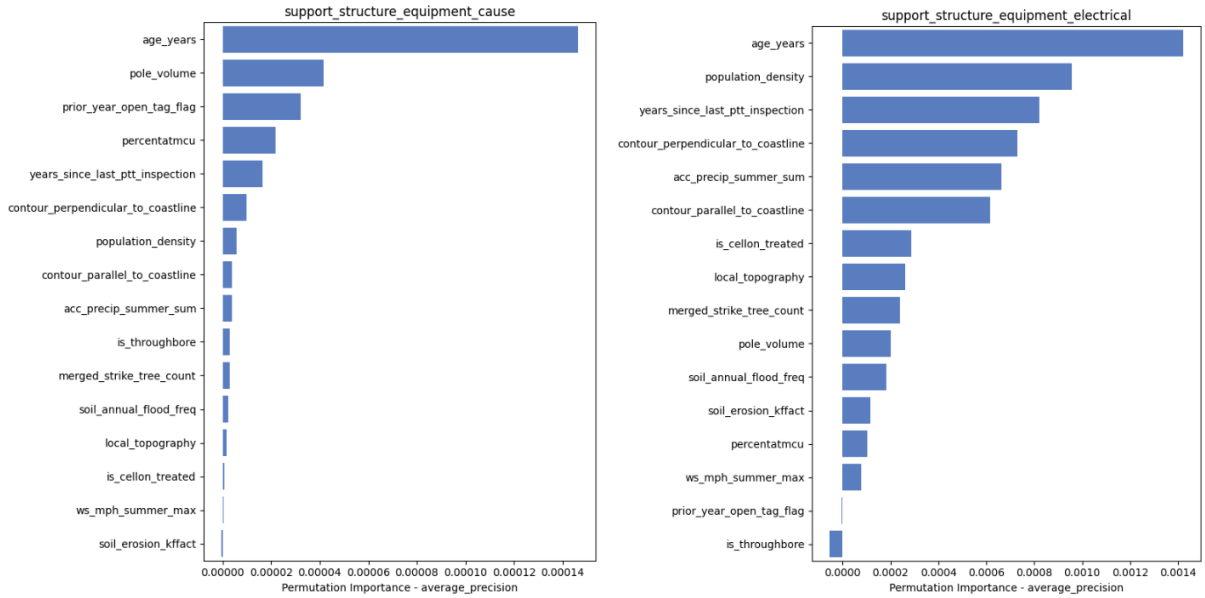


Figure 55. Support Structure models permutation importance scores (left: support_structure_equipment_cause model and right: support_structure_equipment_electrical model).

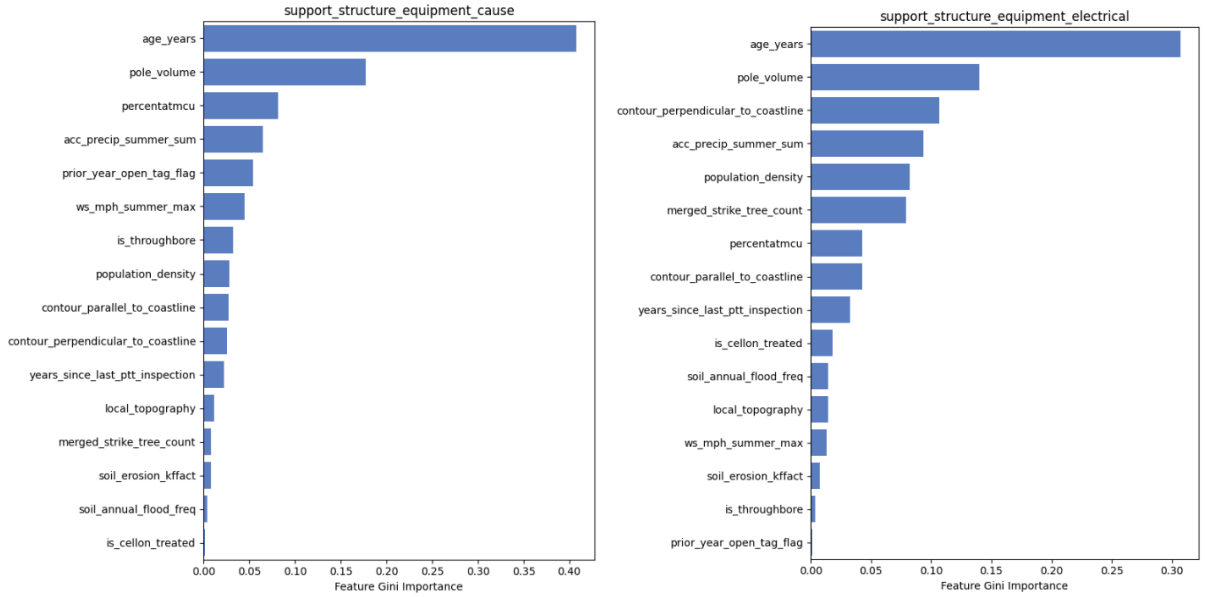
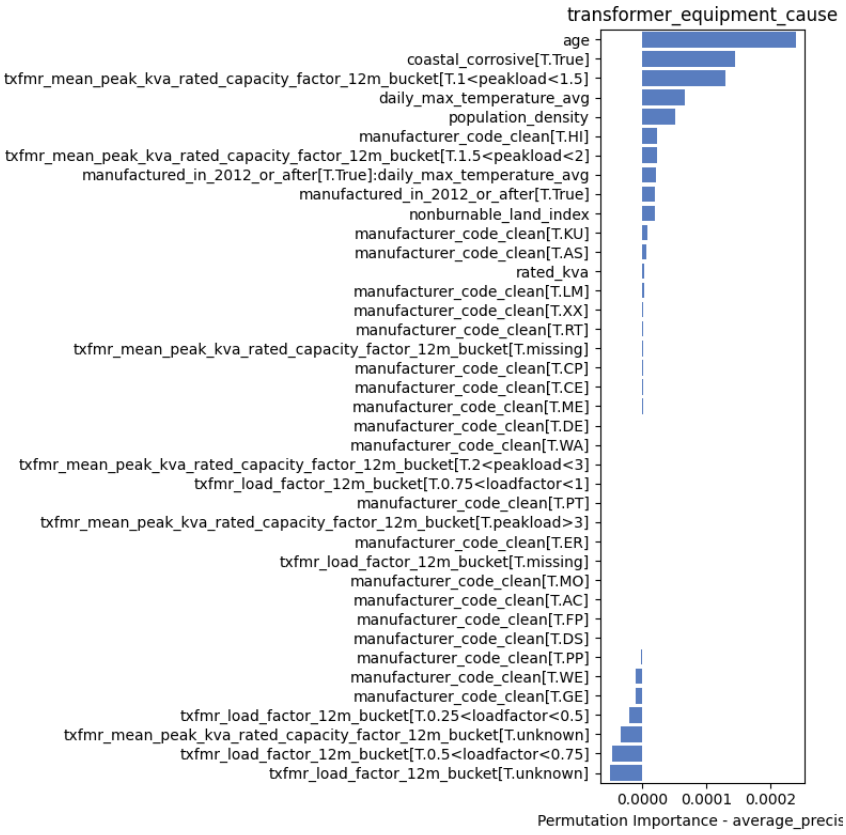


Figure 56. Support structure models feature Gini importance scores (left: support_structure_equipment_cause model and right: support_structure_equipment_electrical model)


transformer_equipment_cause
age
coastal_corrosive[T.True]
txfmr_mean_peak_kva_rated_capacity_factor_12m_bucket[T.1<peakload<1.5]
daily_max_temperature_avg
population_density
manufacturer_code_clean[T.HI]
txfmr_mean_peak_kva_rated_capacity_factor_12m_bucket[T.1.5<peakload<2]
manufactured_in_2012_or_after[T.True]:daily_max_temperature_avg
manufactured_in_2012_or_after[T.True]
nonburnable_land_index
manufacturer_code_clean[T.KU]
manufacturer_code_clean[T.AS]
rated_kva
manufacturer_code_clean[T.LM]
manufacturer_code_clean[T.XX]
manufacturer_code_clean[T.RT]
txfmr_mean_peak_kva_rated_capacity_factor_12m_bucket[T.missing]
manufacturer_code_clean[T.CP]
manufacturer_code_clean[T.CE]
manufacturer_code_clean[T.ME]
manufacturer_code_clean[T.DE]
manufacturer_code_clean[T.WA]
txfmr_mean_peak_kva_rated_capacity_factor_12m_bucket[T.2<peakload<3]
txfmr_load_factor_12m_bucket[T.0.75<loadfactor<1]
manufacturer_code_clean[T.PT]
txfmr_mean_peak_kva_rated_capacity_factor_12m_bucket[T.peakload>3]
manufacturer_code_clean[T.ER]
txfmr_load_factor_12m_bucket[T.missing]
manufacturer_code_clean[T.MO]
manufacturer_code_clean[T.AC]
manufacturer_code_clean[T.FP]
manufacturer_code_clean[T.DS]
manufacturer_code_clean[T.PP]
manufacturer_code_clean[T.WE]
manufacturer_code_clean[T.GE]
txfmr_load_factor_12m_bucket[T.0.25<loadfactor<0.5]
txfmr_mean_peak_kva_rated_capacity_factor_12m_bucket[T.unknown]
txfmr_load_factor_12m_bucket[T.0.5<loadfactor<0.75]
txfmr_load_factor_12m_bucket[T.unknown]
0.0000
0.0001
0.0002
Permutation Importance - average_precis

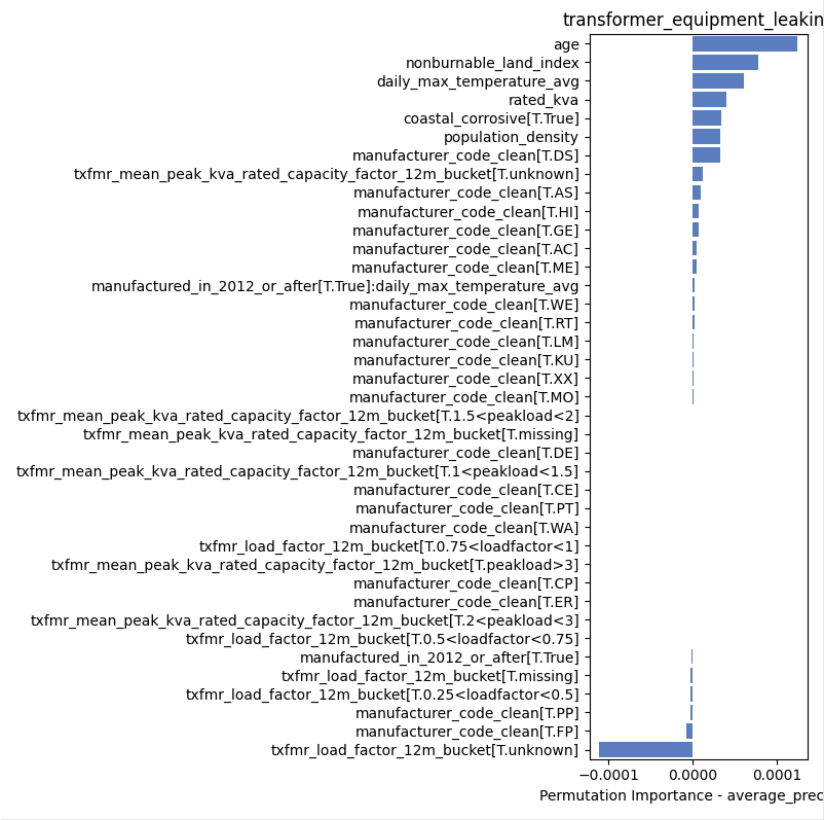
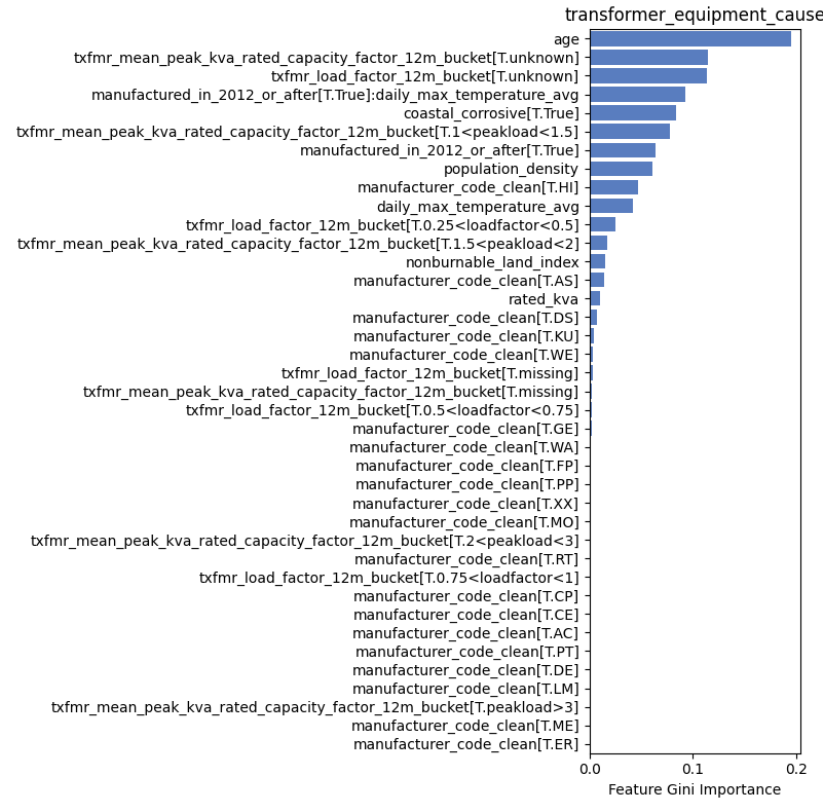


Figure 57. Transformer models permutation importance scores



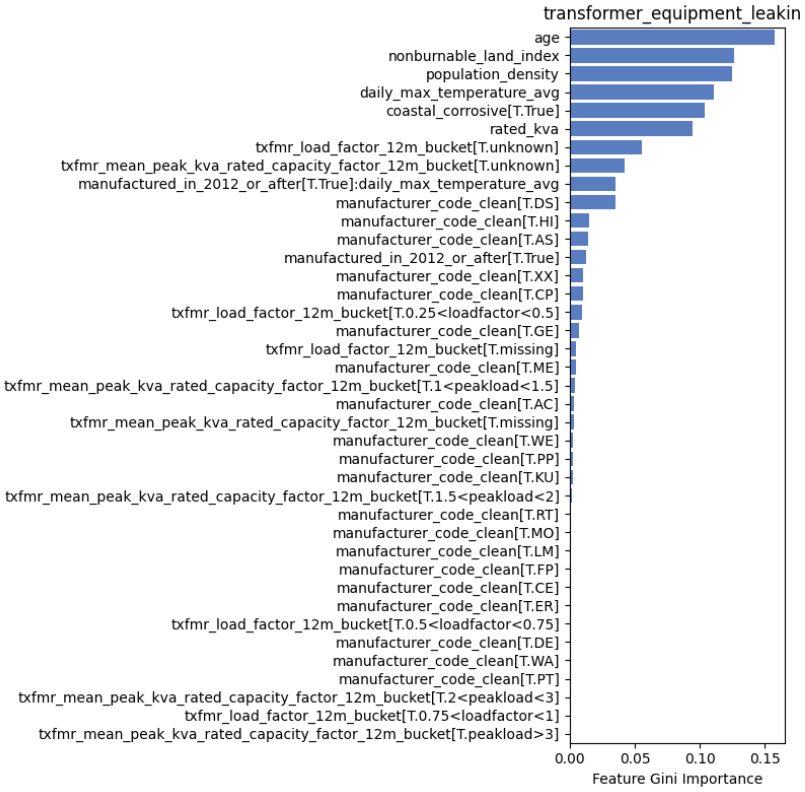


Figure 58. transformer models feature Gini importance scores

8.2.4 Algorithm

8.2.4.1 Algorithm selection

The following algorithms were tested on an early prototype of each asset equipment failure model prior to selecting the Random Forest Classifier as the top performing algorithm for both asset types. Ultimately, each equipment failure model was developed using Scikit-learn's Random Forest Classifier. Random Forest Classifier is a tree-based algorithm that implements an ML method called bagging to combine the results of many different decision trees fit to sub-samples of training data into one model result, minimizing the error. Although Random Forest can be a black-box algorithm, model explainablity tools, such as the Shapley Python library, were used to better understand the resulting model tree-based fit.

Table 26. support structure prototype model performance during the algorithm selection process

Support structure **Algorithm type**	**Prototype model performance** **Test AUC**
Logistic regression	0.50
• with up-sampling minority class	0.60
Maximum entropy	0.66
Random Forest	0.67

Table 27. transformer unit prototype model performance during the algorithm selection process

Transformer Unit **Algorithm type**	**Prototype model performance** **Test AUC**
Survival Model: Kaplan Meier Estimator	0.52
Random Forest	0.60

8.2.4.2 Hyperparameter selection

Random Forest can overfit to the training data set, and the model hyperparameters were adjusted to force the algorithm to build a less complex model to prevent overfitting. The hyperparameters were identical across all models except for the *max_depth* hyperparameter, which was reduced from 4 to 3 for the *transformer_equipment_cause* model to prevent overfitting. The hyperparameters that were adjusted from the default algorithm hyperparameters are listed in **Table 8**. The default parameters are listed on the Scikit-learn library's RandomForestClassifier: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html.

Table 28. Model-specific hyperparameters that deviated from the scikit-learn Python package's default hyperparameter values.

	hyperparameters		
Model name	**class_weight**	**max_depth**	**n_estimators**

support_structure_equipment_cause	balanced_subsample	4	120
support_structure_equipment_electrical	balanced_subsample	4	120
transformer_equipment_cause	balanced_subsample	3	120
transformer_equipment_leaking	balanced_subsample	4	120

8.2.4.3 Calibration

Calibration of the Random Forest model results was required to produce annual probability values due to the Random Forest algorithm's tendency to result in a bi-modal probability distribution gathered around 0.5. The calibration was performed using the Python Scikit-learn library's CalibratedClassifierCV method, which uses cross-validation to calibrate the Random Forest model fit. Utilizing an early prototype model, **Figure 7** demonstrates the histogram for Random Forest probabilities prior to calibration, and **Figure 8** shows the resulting histogram of probability values after CalibratedClassifierCV calibration is applied.

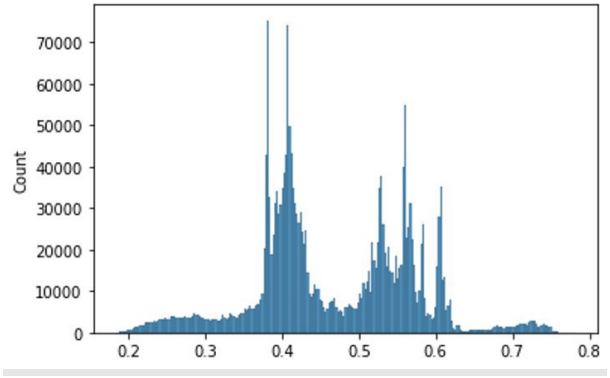


Figure 59. histogram of Random Forest probabilities

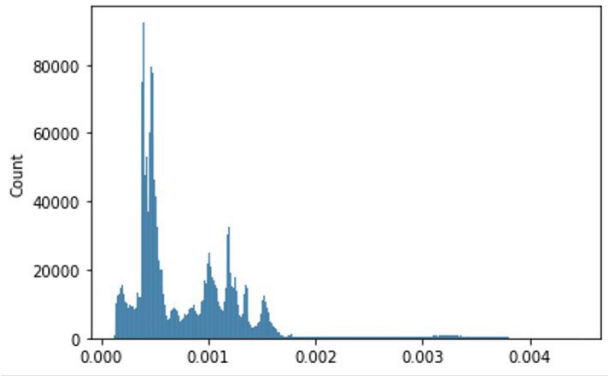


Figure 60. histogram of calibrated Random Forest probabilities

8.2.5 Summary of model results

8.2.5.1 Model performance

The resulting predictions of each model provide an asset-level probability of equipment failure outage event for *each* asset, taking into consideration asset-specific attributes. The final performance of each model on an out-of-sample random test dataset is recorded in Error! Reference source not found..

- The *support_structure_equipment_electrical* model performed very well with an out-of-sample test AUC score of 0.81 (**Figure 10**).
- The *support_structure_equipment_cause* performed adequately with an out-of-sample-test AUC score of 0.71 (**Figure 9**).
- The *transformer_equipment_leaking* model performed adequately with an out-of-sample AUC score of 0.72 (**Figure 12**).
- Although there was significant improvement in the *transformer_equipment_cause* model performance over the model development process, the final performance was below expectations with an AUC score of 0.61 (**Figure 12**). Future work will focus on expanding the feature engineering and algorithm selection experimentation process with the goal of improving performance for the *transformer_equipment_cause* model.

Table 29. Summary of model performance on an out-of-sample test dataset

Equipment failure model name	Model Performance Test AUC	Model Performance Test Top 20% Concentration Factor
support_structure_equipment_cause	0.71	2.4
support_structure_equipment_electrical	0.81	4.0
transformer_equipment_cause	0.61	0.8
transformer_equipment_leaking	0.72	2.2

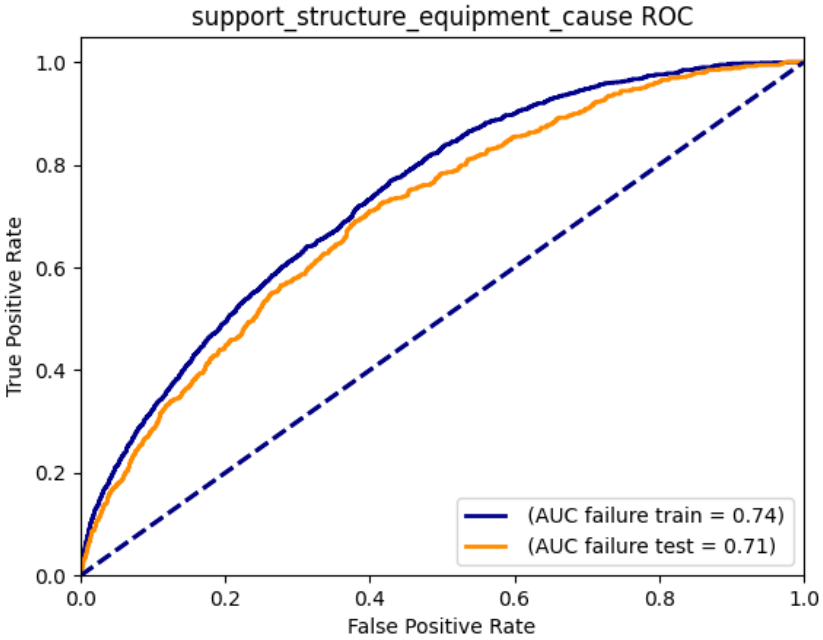


Figure 61. support_structure_equipment_cause model AUC performance on train and test data set

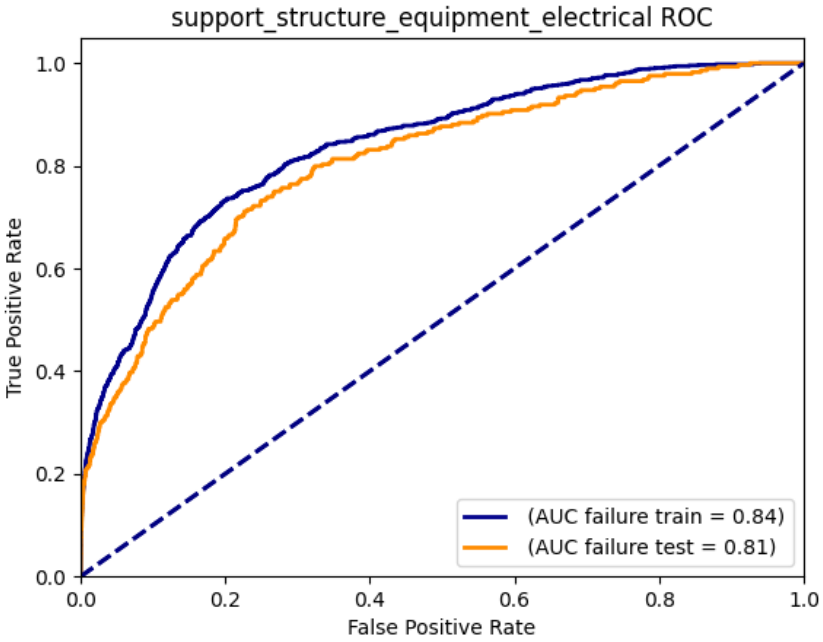


Figure 62. support_structure_equipment_electrical model AUC performance on train and test data set

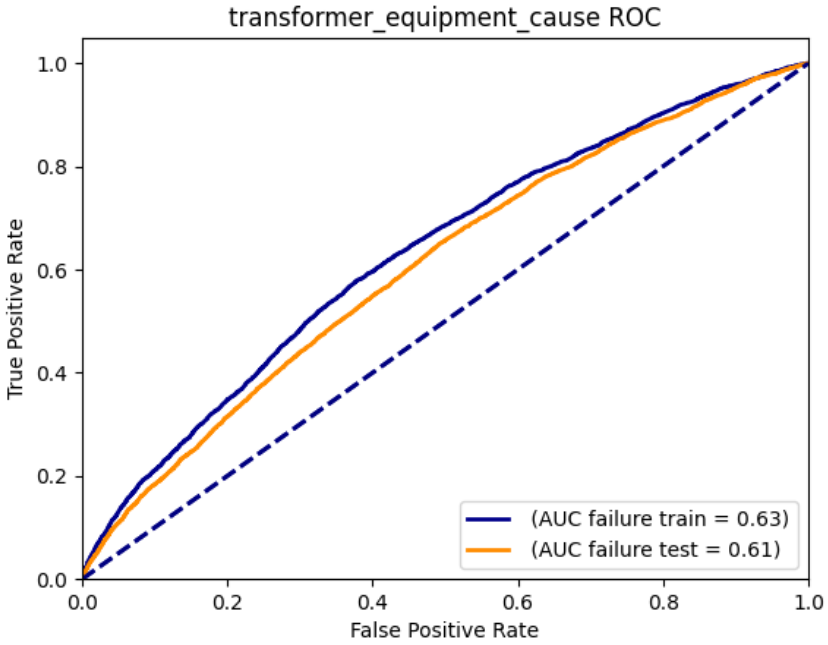


Figure 63. transformer_equipment_cause model AUC performance on train and test data set

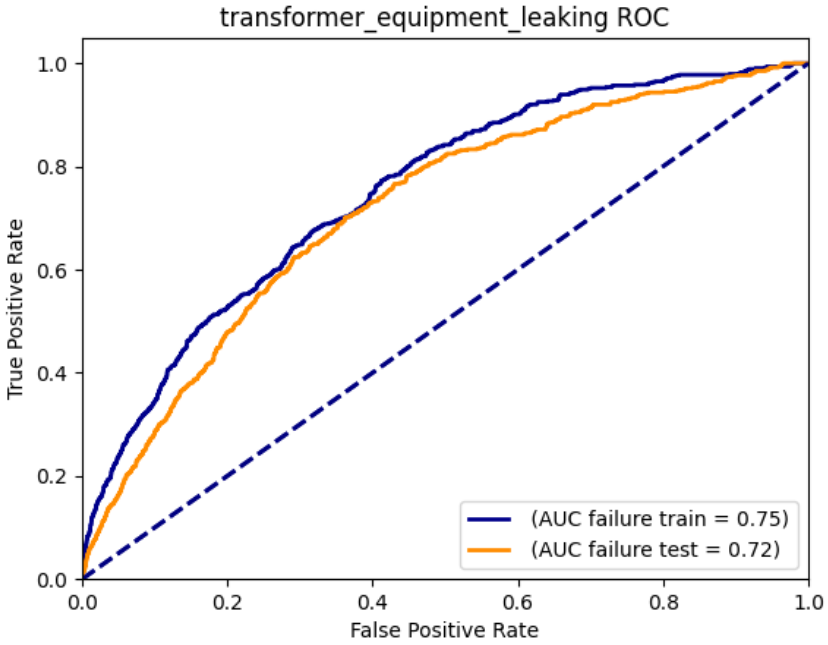


Figure 64. transformer_equipment_leaking model AUC performance on train and test data set

8.2.5.2 Model benefits

The equipment failure model results proved to provide adequate model performance and flexible model output for the first generation of asset-based equipment failure models. Notable benefits included:

- *Asset based.* The model predicts the probability the asset will experience an equipment failure outage event. This tabular asset-level output can be aggregated to provide pixel-level probabilities so that it can be combined with other subsets to form a composite.
- *Asset attribute focused.* Since the models are asset-based, asset-specific attributes can influence the likelihood of equipment failure. As summarized in the feature selection section below, asset attributes were the most influential features in the model algorithm. The influence of asset attributes aligned with the equipment failure causal pathway, which helped promote trust and understanding by the model's end users.
- *Mitigation aligned.* Model covariates include asset attributes related to the mitigation programs utilizing the model. For example, the asset age is the top performing feature in all models, which is reset to zero once the asset is replaced as part of the asset replacement work planning. As assets are replaced, the resulting model probability output will decrease since the model features are aligned with the mitigation work.
- *Temporally aware.* The training data is structured into asset-years, which track whether there were one or more equipment failure events for each asset each fire season. The asset attributes during the year in which the equipment failure occurs are used to train the model. Thus, attributes can be time varying (e.g. asset age, open maintenance tags, details from the last inspection, and changes in electrical loading), and the asset attributes at the time of failure are considered.
- *Flexible for work planning.* Because model predictions are asset-based, work plans can be prioritized by simply joining the model results based on the asset primary key – no allocation of pixel-based risk to assets or spatial analyses are required. Alternatively, the asset-based results can be summed to the pixel level and used for pixel-based work planning.

8.2.6 Summary of Mitigations Methodology

Incorporating mitigation efficacy estimates into the model results allows the associated mitigation program to separate baseline risk into mitigatable risk and the residual risk remaining after mitigation. The mitigatable risk can be used to prioritize assets that the mitigation program has a higher efficacy for to promote risk spend efficiency (RSE). It also allows for identification of where the efficacy is low and the residual risk may remain high even after mitigation and a different mitigation may be more effective. The mitigation efficacy for pole replacement and transformer replacement were estimated utilizing the asset-specific attributes associated with each probability result. The approach specific to each model asset type is summarized in the following subsections.

8.2.6.1 Support structure mitigations

The *support_structure_equipment_cause* model was used for prioritizing pole replacement work planning, which is aimed at replacing poles that have been identified during the inspections process as requiring replacement due to a structural deficiency. The efficacy of pole replacement in the following year was estimated based on a review of historically replaced poles' annual failure rate in the year following replacement and the model algorithm's utilization of the pole replacement-related features. This mitigation efficacy estimation approach was labeled the "Similarity Method".

The pole replacement-related features in the *support_structure_equipment_cause* model included the pole's age, loading percent at Mcu, prior year open tag count (for B, E, F, & H tags), and the number of years since last invasive PT&T inspection. Each of these mitigation-related features in the model can be mitigated and set to zero once the pole is replaced. Since these features are some of the most influential features in the model based on the feature permutation importance scores, the model algorithm will significantly decrease the likelihood of failure once the pole is replaced.

Initially, the Similarity Method buckets the population of poles into groups based on the pole replacement-related features. The categorical buckets used to group the pole population are:

- age
  - 0-40 years
  - >40 years
- Percent at Mcu
  - <=31%
  - Unknown
  - >31%
- Open tag in prior year
  - True
  - False
- Years since last PT&T inspection
  - 0-4
  - >4

Then each of the deficient pole groups are compared to the ideal young & healthy pole group. The ideal young & health pole is listed in the first row of **Table 10** and is less than 40-years-old, had zero maintenance tags in the prior year, had a PT&T inspection within the last 4 years, and is estimated to be at less than 31% of the Mcu. A

recently replaced pole (i.e. a mitigated pole) always falls into this young & healthy group. The formula used to compare a deficient pole group to the mitigated pole group to calculate the efficacy value is:

$$\varepsilon = 1 - \frac{expected\ annual\ failure\ rate\ for\ mitigated\ pole\ group}{expected\ failure\ rate\ for\ deficient\ pole\ group}$$

The resulting efficacy values for pole replacement are included in the last column of **Table 10**. Observe that the effectiveness values for replacing a pole range from 0% to 83%. More detailed documentation of the mitigation calculation and implementation process, can be found in the support structure model's mitigations technical documentation.

Table 30. Pole replacement effectiveness values for poles grouped by mitigation-related asset attribute

Age (years)	Percent at Mcu	Open tag in prior year	Number of years since last PT&T inspection	Pole replacement estimated effectiveness
0-40	<= 31%	False	0-4	0
0-40	<= 31%	False	4+	0.054041
0-40	<= 31%	True	0-4	0.292436
0-40	<= 31%	True	4+	0.350737
0-40	unknown	False	0-4	0.279867
0-40	unknown	False	4+	0.376295
0-40	unknown	True	0-4	0.513729
0-40	unknown	True	4+	0.583575
0-40	> 31%	False	0-4	0.413562
0-40	> 31%	False	4+	0.449270
0-40	> 31%	True	0-4	0.554744
0-40	> 31%	True	4+	0.587601
40+	<= 31%	False	0-4	0.600210
40+	<= 31%	False	4+	0.627793
40+	<= 31%	True	0-4	0.719409
40+	<= 31%	True	4+	0.750524
40+	unknown	False	0-4	0.723916
40+	unknown	False	4+	0.749747
40+	unknown	True	0-4	0.804098
40+	unknown	True	4+	0.830226
40+	> 31%	False	0-4	0.787688
40+	> 31%	False	4+	0.805365
40+	> 31%	True	0-4	0.825908
40+	> 31%	True	4+	0.847246

8.2.6.2 Transformer unit mitigations

Mitigation efficacy values for transformer replacement were developed in collaboration with transformer replacement SMEs and were applied to the *transformer_equipment_cause* model predictions to estimate and prioritize by the risk mitigated by replacing a transformer unit. The mitigation effectiveness values for transformer unit replacement were computed by first identifying a base efficacy value for a low-risk, young transformer. Then the base efficacy is increased for more deficient transformers by applying multipliers assigned by SMEs based on bucketed mitigation-related asset attributes.

The base efficacy value was numerically solved to be 28% using the following formula:

$$\epsilon_b = 1 - \frac{r_{young}}{r_{all}} * \left(\sum_{p \in \{permutations\}} f_p * \frac{1 - \epsilon_b}{1 - x_p * \epsilon_b} \right)$$

The following multipliers were established in collaboration with transformer replacement SMEs:

$$f_{bad\_actor} = \begin{cases} 1 & not\ bad \\ 1.2 & bad \end{cases}$$

$$f_{average\_peak\_load\_percent} = \begin{cases} 1 & 0 - 100\% \\ 1.1 & 100 - 150\% \\ 1.2 & 150 - 200\% \\ 1.3 & 200 - 300\% \\ 1.4 & > 300\% \end{cases}$$

$$f_{load\_factor} = \begin{cases} 1 & 0.00 - 0.25 \\ 1.05 & 0.25 - 0.50 \\ 1.1 & 0.50 - 0.75 \\ 1.2 & 0.75 - 1.00 \end{cases}$$

$$f_{manufacture\_date} = \begin{cases} 1 & \geq 2012 \\ 1.1 & < 2012 \end{cases}$$

And the following formula was used to calculate the efficacy for each combination of the groups listed above (40 in total):

$$\epsilon = \epsilon_b * f_{bad\_actor} * f_{average\_peak\_load\_percent} * f_{load\_factor} * f_{manufacture\_date}$$

8.2.7 Utilization for work planning

Asset-level probability results have proven to be sufficiently flexible for work planning. Several approaches to implement model results to prioritize work plans are to calculate:

1. *Wildfire Risk*: The model results can be combined with the p(ignition|outage) model result to produce p(ignition) and then further combined with the Wildfire Consequence model to produce Wildfire Risk values.
2. *Public Safety Risk*: Alternatively, the likelihood of equipment failure can be combined with Public Safety Consequence to produce Public Safety Risk values.
3. *Mitigated Risk*: Asset-specific mitigation efficacy can be utilized to estimate the magnitude of risk mitigated for each asset due to a specific mitigation program.

The predictions include asset-specific primary keys, which can be joined directly to lists of assets in Foundry to quickly prioritize mitigation work plans. Additional documentation on how the *support_structure_equipment_cause* model were used for work planning are included in *Prioritization of support structure model results for work planning*.

Another advantage of providing a result for the asset itself as opposed to a pixel-level probability include the ability to provide variation in probability values even within a 100m x 100m pixel (see **Figure 13** and **Figure 14**).

Figure 65. Example 100m x 100m pixel contains likelihood of equipment failure values for five poles with significantly different rankings (from the bottom 50% to the top 10-20%)

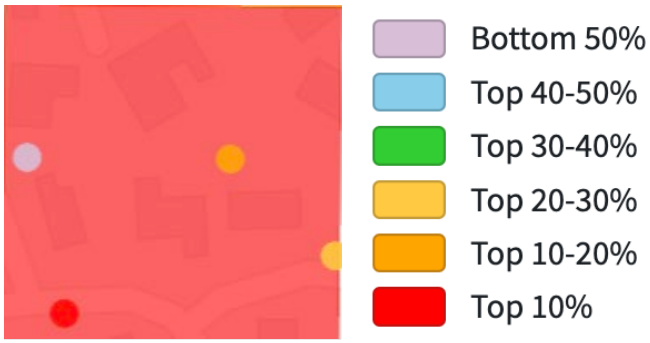


Figure 66. Example pixel for transformer units demonstrating the variability in likelihood of equipment failure within a pixel (from the bottom 50% to the top 10%)

8.2.8 References

1. Transformer unit technical documentation: Transformer Model Documentation.docx
2. Transformer unit mitigations methodology: Transformer Effectiveness Factors.docx
3. Support structure technical documentation: Technical Documentation - Support structure equipment failure models (structural & electrical).docx
4. Methodology for support structure mitigation efficacy: Methodology for quantifying the effectiveness of pole mitigation programs at mitigating failures with potential for ignition.docx
5. Prioritization of support structure model results for work planning: Methodology for using model results to prioritize support structure work plans.docx
6. Scikit-learn library's RandomForestClassifier: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html
7. Scikit-learn library's CalibratedClassifierCV: https://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html

8.3 Probability of Ignition Given Outage

Model Development and Validation

The RaDA team explored various Random Forest, XGBoosted trees, and Logistic Regression model specifications using 80/20% random splits of training vs. testing data and in- and out-of-sample ROC-AUC to quantify predictive performance and monitor for overfitting to training data. Using these metrics, it was determined that the tree-based methods tended to overfit the training data (significant gap in AUC between train and test – with test performance lagging behind training performance). With similar performance between train and test data and good predictive power, the logistic regression models were found to predict well out of sample.

The RaDA team explored three logistic regression model structures of cumulative/increasing complexity. In each case, the RaDA team first applied Recursive Feature Elimination (RFE) as implemented in scikit-learn to eliminate poorly performing covariates and reduce the risk of overfitting to training data. The model formulations tested were:

(1) **The subset-only model** consisting of subset fixed effects only – allowing each subset to have its own rate of ignitions per outage. This is equivalent of calculating the rate of ignitions per outage within each subset's target set data.
(2) **The spatially varying model**: 1, plus spatially varying covariates, like vegetation and prevailing weather and fuel moisture assigned by location to every grid pixel.
(3) **The time varying model**: 2, with the addition of daily time-varying elements such as weather and weather-dependent covariates.

(1) The subset-only model uses only subset_name indicators to predict the arrival rate of ignitions per outage for each subset. This approach surfaces significantly different ignition rates for each of the subsets. Recall that subsets correspond to different voltages, root causes, and equipment types, so this result affirms that voltage, cause, and equipment type play significant roles in determining how dangerous a given outage is.

Note that when using only subset indicators, the arrival rates are constant for each subset, which means the predictions are independent of location and constant over time. Because the likelihood of an outage leading to an ignition should vary spatially, the RaDA team moved beyond the basic model to look at a model that varies spatially across PG&E's landscape.

(2) the spatially varying model:

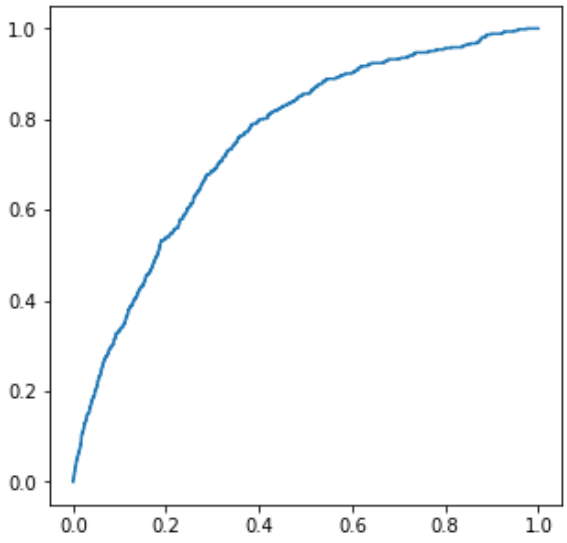
This model employed the covariase found in 1 plus 'hftd', 'impervious_mean', 'fpi_fire_season_dfm_1000hr_avg', 'satellite_tree_count', and 'fpi_v3_veg_types' which included percentages that fall into the following categories: 'fpi_v3_urban', 'fpi_v3_grass', 'fpi_v3_shrub', 'fpi_v3_timber_litter', 'fpi_v3_timber_understory'. The fpi_v3 prefix refers to the data source being the same PG&E meteorology data using in the PSPS-facing FPI model.

This formulation introduces vegetation and land use data as well as average long term windspeed, weather, and fuel moisture values derived from PG&E meteorology data. These values are pulled in for each grid location using spatial raster data. This differentiates each location from the others and captures some of the pre-conditions for the viability of ignitions, like the presence of fuels.

One drawback of this model, however, is that while it differentiates across space, it cannot account for daily or monthly fluctuations in such ignition-relevant factors as wind and fuel moisture. Although outages happen at a

specific time and location where weather conditions, including wind and fuel moisture at the time and location of the event vary, our time invariant model is only able to include average, max, min or other aggregated data for the time period spanning entire fire seasons.

Figure 67. Spatial p(i|o) test ROC




```
Train ROC AUC=0.761
Test ROC AUC=0.753
```

(3) The time varying model:

```
'ign_ind ~ subset_name + hftd + impervious_mean + fpi_fire_season_dfm_1000hr_avg +
satellite_tree_count + fpi_v3_grass + fpi_v3_timber_litter + fpi_v3_shrub + fpi_v3_timber_understory
+ fpi_v3_urban + fpi_v3_event_dfm_10hr + fpi_v3_event_dfm_1000hr + fpi_v3_event_ws_mph'
```

This model adds fpi_v3_event_dfm_10hr, fpi_v3_event_dfm_10hr, fpi_v3_event_dfm_1000hr, fpi_v3_event_ws_mph (dry fuel moisture over different time periods and wind speeds) for all days/locations to fit the model. This model considers the conditions of every historic weather day with at least one outage separately and then “marginalizes” the days into a fire-season probability for every location. Thus, the minority of days with elevated rates of outages and elevated fire “susceptibility” via dry fuels and wind contribute more to the marginalized fire season probabilities than do the “regular” days.

Note that it is straight forward to look up wind speeds or fuel moisture for the day and location of a given outage. That weather data is sufficient to train the p(ignition|outage) model using time-varying conditions, however, the p(outage) estimates used for prediction are subset-specific *annual fire season estimates without specific daily timing*. To adopt time varying conditions as elements of the p(ignition|outage) model, it is necessary to predict p(ignition|outage) for all days of the fire season, multiply the value for each day by the probability (in time) of outages for each day, and sum across all days. This process, known as marginalization, provides a framework for summing daily probabilities into seasonal probabilities, where the daily time dimension of the data is said to have been marginalized out.

$$P(i|o;L) = \sum_{t} P(i|o;L,T=t) \cdot P(o;T=t)$$

We define L as location specific covariates and T as the covariates that vary in both location and time, with a daily timestep, where t stands for each day in the historical data set of outages. We note that using L and T in our logistic regression, gives us the ability to estimate P(i|o;L,T). We want separate estimates for all locations, but only a single seasonal estimate at each location, thus P(i|o;L). Here P(o;T=t) is the historical probability of an outage from the given subset of outages on day t, which can be derived from historical data. Note that by multiplying by the probability of outage for a given day, marginalization places greater weight on weather conditions that occur on days that see elevated rates of outages – like windstorms.

8.4 Consequence

8.5 Applying Mitigation Effectiveness to Ignition Probability and Wildfire Risk

Section 3.6, Wildfire Risk Mitigation, presents the details of estimating the effectiveness of mitigation, for each of four PG&E mitigation programs. Within each composite program, this results in an effectiveness factor ζ_S for each of the subsets, that comprise it.

The prediction models operate at the unit of a pixel. The subset effectiveness factor ζ_S, is used to derive the mitigated probability and risk estimates, in each pixel, for each composite.

Denoting,

$$P_{S,i}^{ignition}: \text{ subset } S \text{ prediction of ignition probability in pixel } i$$

With C_i the wildfire consequence in pixel i, the wildfire risk in pixel i is given by,

$$R_{S,i} = C_i * P_{S,i}^{ignition}$$

Using the weighted mitigation effectiveness for each subset, the **potential reduction** in the probability and risk, is,

$$\Delta P_{S,i}^{ignition} = \zeta_S \times P_{S,i}^{ignition}$$

$$\Delta R_{S,i} = \zeta_S \times R_{S,i}$$

From obvious arithmetic, the corresponding residual probability and risk values are

$$\mathbb{P}_{S,i}^{ignition} = P_{S,i}^{ignition} - \Delta P_{S,i}^{ignition} = (1 - \zeta_S) \times P_{S,i}^{ignition}$$

$$\mathbb{R}_{S,i} = R_{S,i} - \Delta R_{S,i} = (1 - \zeta_S) \times R_{S,i}$$

The derivation of the probability (or risk) from all the subsets, in a given pixel i is,

$$P_i = 1 - \prod_S (1 - P_{S,i}) \approx \sum_S P_{S,i}$$

since, $P_{S,i} \ll 1.0$

Therefore, the expectation and risk values, and mitigations in pixel i are,

$$P_i^{ignition} = \sum_S P_{S,i}^{ignition} \quad ; \quad \Delta P_i^{ignition} = \sum_S \Delta P_{S,i}^{ignition}$$

$$R_i = \sum_S R_{S,i} \qquad ; \qquad \Delta R_i = \sum_S \Delta R_{S,i}$$

We calibrate the model predictions of ignitions to the annual expectation of the number of events, viz. the annual average over the years in the training set. Let's use the lowercase n to denote the annual expectation values. The model outputs are calibrated so that,

$$\sum_i P_{S,i}^{ignition} \approx n_S^{ignition} = \frac{N_S^{ignition}}{num.\,years}$$

Finally, summing over all the pixels and subsets, provides system-wide estimates for the distribution grid. The predictions of the bare probability/risk values, potential reductions due to mitigations, and residuals are,

$$P^{ignition} = \sum_{S,i} P_{S,i}^{ignition}$$

$$\Delta P^{ignition} = \sum_{S,i} \Delta P_{S,i}^{ignition} = \sum_{S,i} \zeta_S \times P_{S,i}^{ignition}$$

$$\mathbb{P}^{ignition} = \sum_{S,i} \mathbb{P}_{S,i}^{ignition} = \sum_{S,i} (1 - \zeta_S) \times P_{S,i}^{ignition}$$

$$R = \sum_{S,i} R_{S,i}$$

$$\Delta R = \sum_{S,i} \Delta R_{S,i} = \sum_{S,i} \zeta_S \times R_{S,i}$$

$$\mathbb{R} = \sum_{S,i} \mathbb{R}_{S,i} = \sum_{S,i} (1 - \zeta_S) \times R_{S,i}$$

Therefore, due to the aforementioned calibration, the model predicts the number of ignition events system-wide, in a typical year, to decrease from $P^{ignition}$ to $\mathbb{P}^{ignition}$