Table of Contents

List of Figures

List of Tables

1 Technical Model Documentation

1.1 Purpose

The Office of Energy Infrastructure (OEIS) requires transparency in risk calculation methodologies supporting Wildfire Mitigation. Per the guidelines, OEIS has specific requirements for technical documentation, substantiation, and data governance of the models used in risk calculations for the WMP. This template outlines the required technical documentation and substantiation for the models.

1.2 Applicability

The applicability of the model documentation and governance applies to models included in the Wildfire Mitigation Plan (WMP) filed with the OEIS for San Diego Gas & Electric (SDG&E).

2 Technical Documentation

2.1 Problem or Function

2.1.1 Problem Modeled

Define the problem modeled for function performed by the program, for example, calculation of fire growth, smoke spread, people movement, etc.

Wildfire Next Generation System for Operations (WiNGS-Ops) is a real-time risk assessment model built to evaluate and compare Wildfire and Public Safety Power Shutoff (PSPS) risks at the asset level (pole/span) and the sub-circuit/segment level at hourly intervals. The primary purpose of the model is to help inform decision makers in real-time about wildfire and PSPS risks, which will guide risk-based de-energization decisions during risk events. The model outputs used to help guide decision makers are understood to represent a range of potential risk of wildfire versus PSPS comparisons.

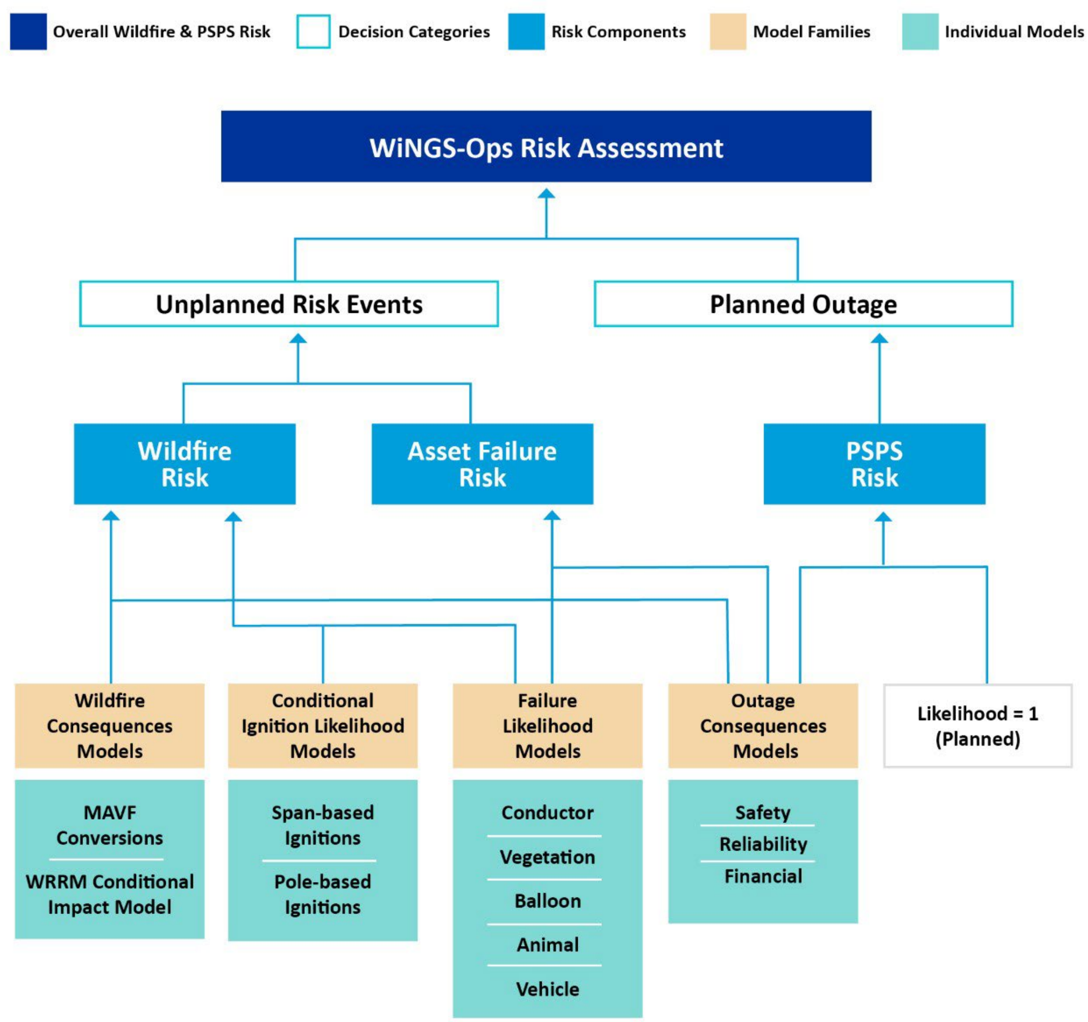
2.1.2 Problem Environment

Describe the total fire problem environment. General block or flow diagrams may be included here.

Changing weather, events that may cause ignitions, and planned work all need to be evaluated to determine the risk of ignition and/or the need for a PSPS de-energization. To make the best operational decisions regarding work and to make decisions to de-energize, real-time information on what is happening in the service territory, especially in the High Fire Threat District (HFTD), is needed.

The primary function of WiNGS-Ops is to provide the ability to weigh the quantified risks of a binary choice of actions: de-energization or not. To this end, the overall WiNGS-Ops model considers several risk components that are comprised of more specific individual sub-models. These components and their relations are shown in Figure 1 and are defined below.

Figure 1: WiNGS-Ops Process Flow Diagram



Tier 1: Decision Categories (white boxes in Figure 1):

- Unplanned Risk Events – The total expected risk of events not including planned de-energization.
- Planned Outage – The total expected risk of de-energization.

Tier 2: Risk components (turquoise boxes in Figure 1):

- Wildfire Risk – The expected risk of wildfire reduced by de-energization, which includes direct impacts of fire and secondary impacts of outages that occur due to wildfire.
- Asset Failure Risk – The expected risk of an unplanned asset failure that does not lead to a wildfire. This equates to the PSPS risk if the failure is at the sectionalizing device. This includes only the impacts of a single asset-based outage that does not cause an ignition.
- PSPS Risk – The expected risk of de-energization at 100 percent likelihood.

Tier 3: Model Families (orange boxes in Figure 1):

- Wildfire Consequence Models – The collection of sub-models required to quantify the consequence of a risk event for wildfire. This includes direct fire impacts.
- Conditional Ignition Likelihood Models – The collection of statistical sub-models used to estimate the conditional probability of an ignition given that an asset failure had occurred.
- Failure Likelihood Models – The collection of statistical sub-models used to estimate the probability of an asset failure.
- Outage Consequence Models – The collection of sub-models required to quantify the consequence of a risk event for outages. This includes fire-caused outages.
- Likelihood – the PSPS risk for WiNGS-Ops comprises only the consequence; the risk event is certain for planned outages. Therefore, the likelihood always equals 1.

Tier 4: Individual Sub-Models (aqua boxes in Figure 1)

- Multi Attribute Value Framework (MAVF) Conversions – Engineering equations used to determine the safety, financial, and reliability consequences of wildfire given outputs from the Wildfire Risk Reduction Model (WRRM).
- WRRM Conditional Impact Model – The computational fire spread model for determining the direct impacts of a wildfire (e.g., buildings destroyed, acres burned) originating from a specified ignition point location.
- Span-based ignitions – A statistical model for estimating the likelihood of an ignition given that an asset failure at a span has occurred.
- Pole-based ignitions – A statistical model for estimating the likelihood of an ignition given that an asset failure at a pole has occurred.
- Conductor – A statistical model for estimating the likelihood of conductor failure.
- Vegetation – A statistical model for estimating the likelihood of asset failure caused by contact with vegetation.
- Balloon – A statistical model for determining the likelihood of asset failure caused by contact with balloons.

- Animal – A statistical model for determining the likelihood of asset failure caused by contact with animals.
- Vehicle – A statistical model for determining the likelihood of asset failure caused by contact with vehicles.
- Safety (Outage) – Quantification of the safety consequences of an outage given the nature of the outage.
- Reliability (Outage) –Calculation of reliability consequences given an outage duration.
- Financial (Outage) – Quantification of financial consequences of an outage given the nature of the outage and the known characteristics of customers subject to the outage.

2.1.3 Background Environment

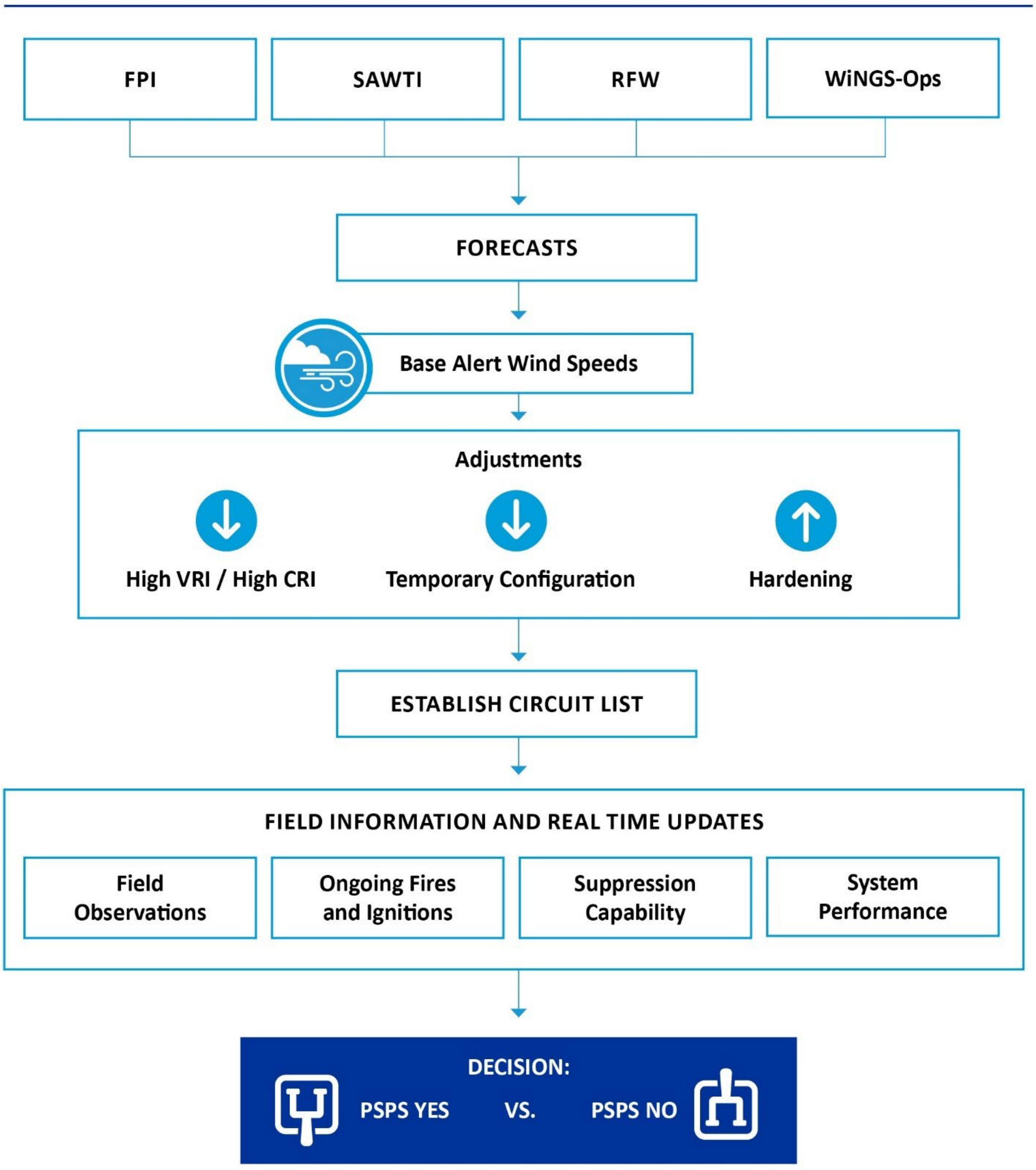
Include any desirable background information, such as feasibility studies or justification statements.

To determine whether the benefit of de-energization (PSPS) outweighs the potential public safety risks, opposing scenarios are quantified in accordance with the risk assessment framework reported in the 2021 Risk Assessment Mitigation Phase (RAMP), which uses a MAVF to quantify risk. Risk components include the likelihood and consequence of a risk event. The subsequent sections of this document primarily report on the methods used to determine the likelihood component through machine learning.

The consequence component, primarily comprised of WRRM, developed by Technosylva, Inc., is a computational dynamics model. The approach of combining machine learning for likelihood modeling and computational dynamics for consequence modeling was discussed at length during energy-led risk modeling workshops and is common among California Investor-Owned Utilities (IOUs).

While the machine learning models discussed in this document produce hourly quantities at the asset-level given weather conditions, they are most useful to emergency operators when aggregated or summarized. For example, the PSPS decision flow diagram (Figure 2) summarizes the results of the conductor risk portion of WiNGS-Ops into a high-medium-low circuit risk index (CRI) for alignment with other risk factors and user digestibility. Likewise, the decision to de-energize often includes a consideration of “what-if” scenarios, such as in understanding the risks if wind gusts were to exceed forecasts. These operational applications inform the nature and format of WiNGS-Ops outputs.

Figure 2: PSPS Decision Framework



2.2 Technical Description

2.2.1 Theoretical and Mathematical Foundations

Convey a thorough understanding of the theoretical and mathematical foundations, referencing the open literature where appropriate.

The WiNGS-Ops model is comprised of individual sub-models that were developed using multiple theoretical and mathematical foundations. This document focuses on the probabilistic models developed using data science methods, namely machine learning. This approach applies statistical learning to the historic record of outages and ignitions to determine the inputs ("features") that are predictive of future outages and ignitions (Hastie, 2009). For models that use machine learning, (Conditional Ignition Likelihood Models and Failure Likelihood Models in WiNGS-Ops) historical data serves to create a past record of binary outcomes: when an outage or ignition occurred and when it did not. Algorithms correlate features within each outcome in a process called supervised learning, and more specifically binary classification. A model is said to be "trained" on the historical data, and then can "infer" probabilities given future data, such as with weather forecasts. The model training process is discussed in Section 2.4.1 and the inference process is discussed in Section 2.4.2. The limitation of this approach is discussed in Section 2.4.3.

Within the realm of machine learning are a variety of algorithms and approaches. The sub-models in WiNGS-Ops leverage both regression and decision-tree-based algorithms. The specific algorithms are detailed in Table 1. Additionally, some approaches require discretionary judgement by the modeler to select parameters that balance statistical rigor with intuition, while other approaches rely almost entirely on the algorithm to detect patterns (Breiman, 2001). When automated algorithms are used, a process must be employed to prevent the algorithm from overfitting the model to the historic data, i.e., the model detects patterns specific only to the historic data, but the patterns are not generalizable to new data unseen by the algorithm. The following section details the theory behind the use of these algorithms.

2.3 Theoretical Foundation

2.3.1 Phenomenon and Physical Laws (Model Basis)

Describe the theoretical basis of the phenomenon and the physical laws on which the model is based.

2.3.1.1 Ensemble Learning, Bootstrap Aggregation, and Gradient Boosting

A foundational concept in machine learning is ensemble learning, the combining of many weak predictors to form an overall better predictor. When paired with predictors that are based on decision trees, it enables the relatively simple creation of powerful models that can avoid overfitting. In standard decision trees, a tradeoff between bias and variance (overfitting and generalizability) can be achieved by modifying the tree structure. The larger and more complex the tree, the more overfitted and less generalizable. However, the smaller the tree, the weaker its predictive power. Finally, an ensemble of smaller trees can yield strong predictive power while preventing overfitting.

Perhaps the most common approach for creating an ensemble of trees is bootstrap aggregation, more commonly referred to by the portmanteau "bagging" (Breiman, 1996). This process uses bootstrapping, a random statistical resampling method, for building many small decision trees from a subset of data and features and then aggregates the predictions from each individual tree. The structure of each tree is optimized (in machine learning, this is called hyperparameter tuning) to ensure that the individual trees are sufficiently small to avoid overfitting while sufficiently large to have predictive power. This overall approach, which is used in WiNGS-Ops, is commonly referred to as a "random forest".

Another approach for creating an ensemble of trees is gradient boosting. Like with random forests, it is an ensemble learning algorithm which combines multiple learning decision tree models to obtain better performance. However, rather than utilizing the random process behind bootstrapping, gradient boosting builds decision trees in sequence and uses the error of one iteration to build the next. Each iteration uses a regularized learning objective function such that when the trees are combined, overfitting can be avoided (Chen, 2016). The implementation of gradient-boosted trees used in WiNGS-Ops is "XGBoost", which has been shown to consistently perform better and more efficiently than random forests.

2.3.1.2 Measuring Model Performance with Cross-Validation

To avoid overfitting in machine learning, the performance of the model cannot be measured on the data with which the model was trained. The common method for measuring model performance is cross-validation, a statistical method of evaluating and comparing learning algorithms by dividing data into two segments: one used to learn or train a model and the other used to validate the model (James, 2013).

When data is limited, resampling techniques can be used to train and test a set of model parameters. For example, a process akin to the statistical method of "jackknifing", leave-one-out cross validation (LOOCV), selects one observation from the dataset for testing and the remaining observations for training, and this process is repeated until each sample serves as the test sample once. The final model performance is estimated as the average of all test scores. Since LOOCV can be very compute-intensive, an alternate approach is K-fold cross validation, which divides the data into k number of subsets. Each set is used once as a test set, with the remaining unselected sets collectively used for training for each iteration.

There are several metrics for measuring the performance of machine learning models. For binary classification, one of the most used metrics is the area under the receiver operating characteristic curve (ROC AUC) (Hanley, 1982). The ROC curve is created by plotting a model's True-Positive Rate (TPR) versus its False-Positive Rate (FPR) across all possible classification thresholds where TPR and FPR are represented on y-axis and x-axis. TPR and FPR are values derived from the confusion matrix generated from cross-validation. ROC AUC values vary from 0 and 1; a value of 0.5 resembles a model with random outputs.

2.3.2 Governing Equations

Present the governing equations and the mathematical model employed.

For the purposes of WiNGS-Ops, the terms "risk" and "expected risk" are used interchangeably, and their values are calculated as the product of the likelihood of a risk event (LoRE) and consequence of the same risk event (CoRE):

$$Expected\ Risk = LoRE \times CoRE$$

In the case of ignition likelihood, the probability of ignition (PoI) is calculated by Bayes' Theorem,[1] where PoI is the product of the probability of asset failure (PoF) and the conditional probability of ignition given a failure (PoI_F) (i.e., the probability that an ignition occurs following an asset failure), divided by the conditional probability of a failure given an ignition had resulted (PoF_I which is approximately equal to one):

$$PoI = \frac{PoF \times PoI_F}{PoF_I}$$

2.3.3 Independent Review Results

Provide the results of any independent review of the theoretical basis of the model.

In 2022 WiNGS-Ops underwent an internal review to determine areas of improvement. The model was updated to align with software development best practices by integrating source control, code optimization, and a multi-stage production environment. In 2023 a review of the machine-learning process will be explored. An independent reviewer has been engaged to begin the process.

2.4 Mathematical Foundation

2.4.1 Techniques, Procedures, Algorithms

Describe the mathematical techniques, procedures, and computational algorithms employed to obtain numerical solutions.

The following subsections explain the techniques, procedures, and algorithms employed to develop each of the tiers in Figure 1.

2.4.1.1 Likelihood Sub-Model Development Process

The likelihood sub-models are developed by a machine learning process. The general steps to this process are shown in Figure 3. Details for each step specific to WiNGS-Ops are described below:

[1] The ignition likelihood (PoI) is approximated by the Bayesian relation between the probability of failure (PoF) and conditional probability of ignition (PoI_F).

Figure 3: Machine Learning Process

Data collection

Machine learning training requires historic data. The likelihood models use historic records of outages and ignitions collected by SDG&E for reasons independent of WiNGS-Ops.

Create labeled set

From the historic record, a dataset is constructed for each model to be used for model training. The training set must match the desired level of granularity of the model output and, in the case of binary supervised classification, must also be labeled with a binary field indicating whether the event (outage or ignition) had occurred or not. An ideal training set has an equal number of event occurrences and non-occurrences.

For the failure likelihood models, each historic outage since 2015 was mapped to a specific asset and hour of the year and subsequently labeled as “event occurred”. The alternate set of samples labeled as “event did not occur”, i.e., no outage, could in theory consist of all asset-hours since 2015 that did not see an outage. However, this would create considerable imbalance of label classes, which would lead to computational and modeling challenges. Therefore, a sample of the non-occurrence set was taken and deemed as representative of asset-hours for which outages did not occur. The two sets are appended to form the fully labeled training set.

Feature engineering

After a labeled dataset is created, features that describe the observable characteristics of the asset-hour are joined to the observations. The key features used in the WiNGS-Ops are asset characteristics (e.g., wire length, pole material) and weather conditions (e.g., wind gust, temperature). Features can be model-specific, such as “road proximity” for the vehicle contact probability model, or derived from external sources, such as the public land-use designation around the asset. The engineering of features is considered an “art”, requiring creativity, input and feedback from subject matter experts and trial and error. The process of feature engineering also includes creating variations of a single feature, normalizing or scaling numerical features, imputing missing features, or numerical encoding of categorical features.

After the creation of features, some modeling algorithms require careful selection of features to avoid overfitting. For algorithms that are relatively insensitive to a large number of features, selection should still be considered to optimize computing, operational, and latency requirements. In both cases, some level of discretion from the modeler is required.

Algorithmic modeling

The labeled dataset and features, collectively referred to as the "training set", is processed through statistical algorithms. This process is known as "fitting" the data, or "training" the model. The algorithms are typically available for public use in open-source code repositories, such as GitHub. Table 1 shows the software versions for the algorithms used to train the failure likelihood sub-models in WiNGS-Ops.

The algorithms are not entirely a "black box". As discussed in Section 2.3.1.1, the modeler must tune "hyperparameters" to achieve the optimized fit for the select algorithm. This process requires several iterations of model training, or the modeler may opt for an automated tuning process, which is a more recent advancement in the field of machine learning, and therefore not currently employed in WiNGS-Ops.

Model validation

Each trained model is evaluated on data that is set aside from model training by cross-validation, the statistical approach detailed in Section 2.3.1.2. This approach measures the model performance by how it would generalize to new data, rather than how it fits to training data, ensuring the models are predictive in nature. The candidate model is then accepted or rejected.

Operational deployment

After a model is created, evaluated, and accepted, it is deployed in a software environment for use in an automated inference pipeline or other software applications. This process has its own cost and technical considerations, which creates limitations to model complexity, and sometimes performance, in exchange for reliability and robustness. For operational models such as WiNGS-Ops, deployment considerations should be considered during early stages of model development. In other words, model features that create deployment challenges should be precluded.

2.4.1.2 Outage Consequence Network Analysis

To quantify the impacts of a specific outage (planned or unplanned), the customers that would be impacted are first identified by analysis of the electrical system. For PSPS, this is simply the collection of customers downstream of the sectionalizing supervisory control and data acquisition (SCADA) device. For unplanned outages, such as those caused by wind conditions, the customers impacted from a specific asset failure are determined by first identifying the upstream sectionalizing device (SCADA or non-SCADA), and then identifying all customers downstream of that device. This calculation is performed daily for every span considered in WiNGS-Ops (the cadence that the electric network is logged to the analytical warehouse).

For the consequence of an outage caused by a wildfire, the rate-of-spread of the worst-case fire, as determined by WRRM, is used to estimate a geometry within which assets are damaged. The network analysis is performed for all damaged assets, and all customers that would lose power are identified, which could span multiple circuits even though the fire would originate from the failure of one specific asset.

2.4.1.3 References to Techniques and Algorithms

Provide references to the algorithms and numerical techniques.

The machine learning sub-models developed using the procedures detailed in Section 2.4.1.1 utilize open-source software and frameworks (called "libraries" in Python). Software versions and methods are summarized in Table 1.

Table 1: Machine Learning Sub-Models Development Summary

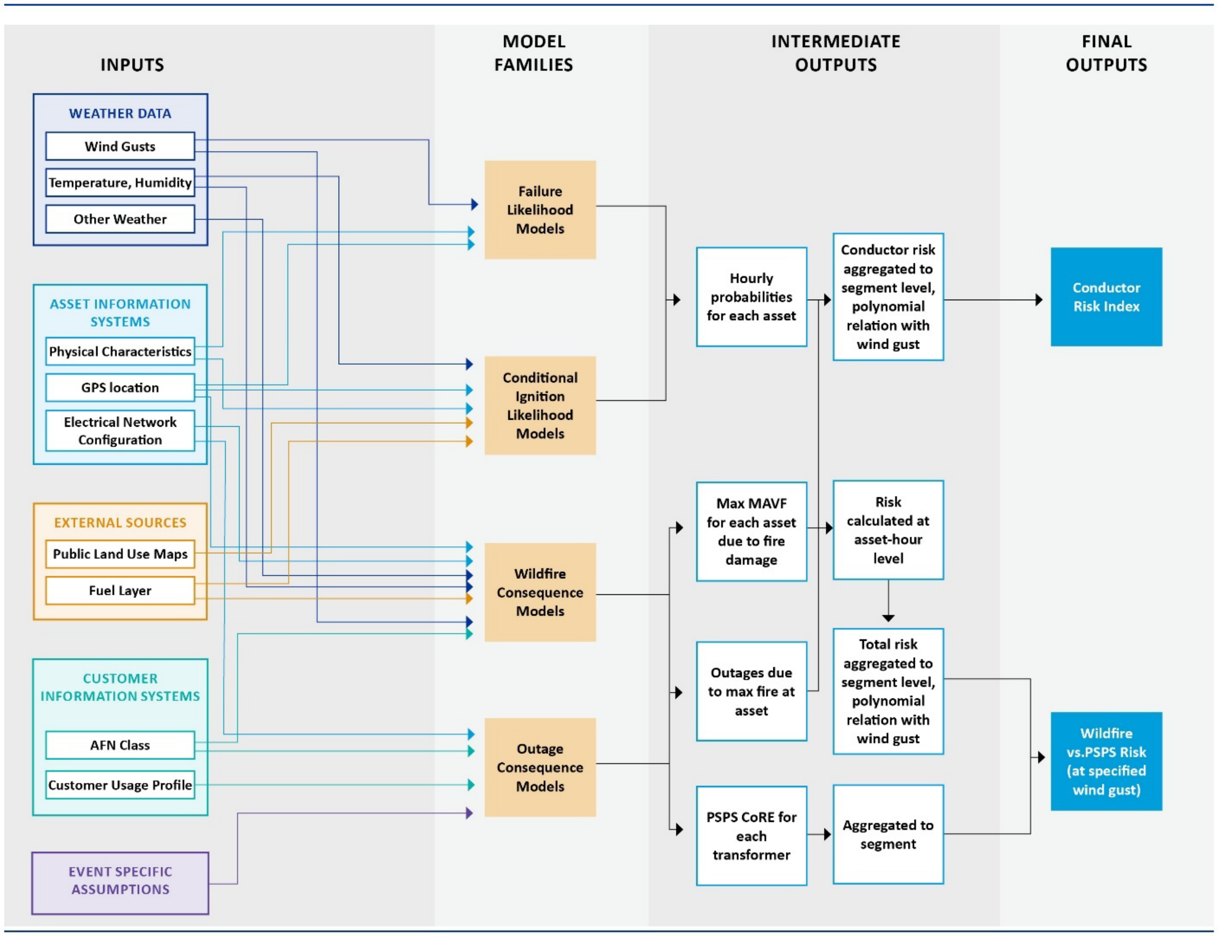
Likelihood Sub-Models	Algorithm	Data Manipulations	Feature Selection Methodology	Model Validation	Python Library (version)
Conductor Failure	Linear regression (log-log)	Misc	Bottom-up p-value	n/a	Statsmodels (0.13.0)
Balloon Contact	Logistic regression	Imputation of missing land-use feature using mean value	Bottom-up p-value	n/a	Statsmodels (0.13.2)
Animal contact	Logistic regression	n/a	Best judgement	n/a	Statsmodels (0.13.2)
Vegetation	Logistic regression	Normalization, scaling	Best judgement	n/a	Statsmodels (0.13.2)
Vehicle	Extreme gradient boosted trees	One-hot encoding	Review of the Feature Importance list	k-fold cross-validation	Xgboost (1.5.1)
Span	Ensemble decision trees (random forest)		Gini importance, discretionary	k-fold cross-validation	Pycaret (2.3.2)
Pole	Engineering equation	n/a	n/a	n/a	n/a

2.4.2 Equations and Implementation

Present the mathematical equations in conventional terminology and show how they are implemented in the code.

The WiNGS-Ops model takes input data from a variety of internal and external data sources. For machine learning models, inputs are determined by the feature selection methodology, reported in Table 1. Figure 4 details the inputs, outputs, and interdependencies of the data flowing through the model. The sub-sections below elaborate on the specific implementations and calculations of this flow diagram.

Figure 4: WiNGS-Ops Calculation Schematic



2.4.2.1 Individual Sub-Model Implementation

After a machine learning model is trained and saved as described in Section 2.4.1, a separate piece of software code (inference pipeline) generates the model output probabilities. This inference pipeline connects to all input data sources to generate the features required by each sub-model. Additionally, the software should use the same versions of the open-source Python libraries from which the models were trained to ensure the model performs as expected.

The consequence models used in WiNGS-Ops (Wildfire Consequence models and Outage Consequence models) that are not developed with machine learning are also implemented as software code in the inference pipeline. The equations are "hard-coded" into the software, which is version-controlled for auditability. Table 2 summarizes the equations used as of the time of this filing.

Table 2: WiNGS-Ops Equations

	Outage Consequence	MAVF Conversion of WRRM Conditional Impact Model
Health and Safety	Customer impact scaling factor × number of affected customers × outage duration × Serious Injuries and Fatalities (SIF) per customer-minutes	structures destroyed × SIF per structure impacted + Total acres burned × MAVF Conversion Factor
Reliability	SAIDI + SAIFI (based on outage duration + assumed restoration duration)	n/a (calculated in Outage Consequence column)
Financial	number of affected customers × dollars per affected customer	structures impacted × dollars per structure + acres impacted × dollars per acre + acres impacted × suppression dollars per acre

2.4.2.2 Model Families Aggregation

After each individual sub-model generates outputs from the data inputs, those values are aggregated to the model family tier. For failure likelihood models, since the modeled events are considered independent, this is a straight summation of the probable outputs from each model for a given asset.

For the consequence models, each component (safety, reliability, financial) is combined in accordance with the weighting scheme established in RAMP.

2.4.2.3 Risk Component Calculations

For wildfire models, risk components are calculated at the asset level, as shown in the Intermediate Outputs in Figure 4. Conversely, the PSPS risk components are calculated at the transformer device level. These components are ultimately aggregated to the segment level for comparison.

To calculate the likelihood of a wildfire event for an asset-hour, the likelihood model families are combined using Baye's Theorem, described in Section 2.3.2. For the likelihood of an asset failure that does not lead to a wildfire, the following equation based on the same statistical principles is used:

$$\binom{PoF}{without\ ignition} = PoF \times (1 - PoI_F)$$

This likelihood (and consequence) is considered to ensure that all risks that would be mitigated by PSPS are accounted for. In other words, the risk of a planned versus unplanned outage at the same time and location are considered equal in WiNGS-Ops. Although a planned outage should amount to less risk in practice (e.g., due to customer notification of potential PSPS), these factors are currently not considered in WiNGS-Ops.

While the consequence of an outage is a straightforward mapping to the model family outputs (for both PSPS and asset failure), the consequence of wildfire is determined by combining both fire and outage consequences. However, in this case, the outage consequence also considers additional customer outages that would occur as a result of the wildfire. The process for this determination is explained in Section 2.4.1.2.

The following equation provides a generalization of the aspect risk calculation of the LoRE and CoRE components:

$$Expected\ Risk_t = \sum_{i=0}^{m} (\sum_{j=0}^{n} (PoF_i \times PoI_i) \times CoRE_i\ )$$

Where:

PoF_i = Probability of Failure at pole/span location
PoI_i = Conditional Probability of Ignition (given a Failure has occurred) at each pole/span location
$CoRE_i$ = Consequence of risk event at each pole/span location
i = variable iterating through each pole/span location
j = variable iterating through each failure sub-model
m = number of poles/spans per segment
n = number of failure sub-models
t = time interval

2.4.2.4 Overall Risk Calculus

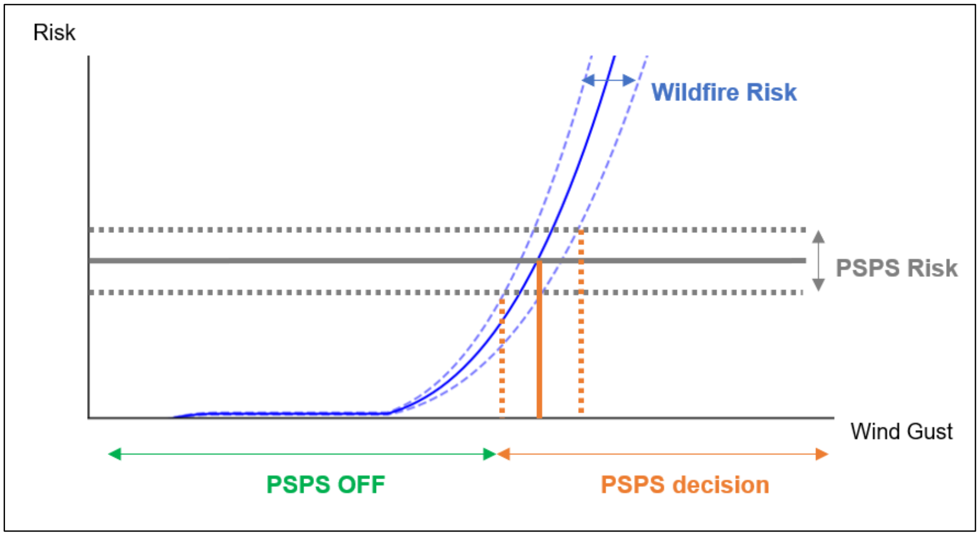
For PSPS events, PSPS risk is primarily a function of shut-off duration, which may vary based on weather conditions. However, for purposes of comparison, a constant duration that is independent of wind conditions is assumed and, unlike wildfire risk, the risk score is assumed to remain constant over wind gust variations. LoRE is not considered since this operational model weighs the expected wildfire risk against the full consequence of a PSPS de-energization. CoRE is calculated utilizing the MAVF, applying the same assumptions used in the WiNGS-Planning model.

After the Expected Outcome values are calculated for each span and pole downstream of a sectionalizing device, a simple regression is performed to acquire an average value for the segments as a function of wind speed. With the calculation of PSPS risk and wildfire risk curves as a function of wind gust, the intersection of risk scores is computed and evaluated as the wind gust at which the wildfire risk surpasses the PSPS risk (see Figure 5).

Wildfire and PSPS risk curves are evaluated as a range of possible values using bands that account for variation of risks within individual spans/poles of a given segment (specifically for wildfire risk), uncertainties, and variations in other assumptions, such as those made around PSPS de-energization consequence. The range of values for each metric allows decision makers to balance flexible decision-making with risk-informed situational awareness, thereby adding to a more holistic approach to PSPS de-energization decision making capabilities.

Figure 5 shows wind gust ranges where the expected wildfire risk impact will start to be greater than the risk impact of performing a PSPS de-energization on that segment. This helps decide at which ranges of wind gusts to consider de-energization of a particular segment. It is important to note that WiNGS-Ops outputs are not the sole decision-making points. Other variables and dynamic input from the field are also considered.

Figure 5: Impact of Wind Gust on Wildfire Risk Impact



2.4.3 Limitations (see Guide ASTM E 1895)

Identified the limitations of the model based on the algorithms and numerical techniques.

2.4.3.1 Limitations

Machine learning models are limited by the characteristics of the training data, such as the number of risk event observations and the temporal-spatial granularity and range of the data collected. For example, if data is collected from only the past few years, then the model results will be biased towards patterns observed during those years. Additional data that is used to generate machine learning features should ideally match the temporal-spatial range and granularity of the training data, although this is not a strict limitation if appropriately managed. This limitation commonly occurs when integrating external data. For example, public land use maps used for training some models cover only San Diego County. Therefore, the values used for assets located outside of San Diego County must be imputed or approximated for some algorithms, typically by using a mean value.

Another general limitation for machine learning is the computational limitations during both model training and inference. While the algorithms described in this document are conceptually simple, they often have heavy computational requirements that limit model complexity. For this reason, considerable effort has been given to migrate WiNGS-Ops to a cloud-based data science platform, such that distributed cloud resources can be leveraged. However, this process also comes with added challenges in software and runtime environments, data access and security, and human capital required to navigate cloud tools.

2.4.3.2 Key Modeling Assumptions

Assumptions made specific to each model to represent the physical world and to simplify calculations

For machine learning models:

- Outages (SAIDIDAT) assumed to be failures
- Reportable ignitions as totality of outages that have caused ignitions
- Representativeness of assets and time and sampling of training set
- Relatively coarse mapping of asset to anemometers is sufficient
- The mapping of outage or ignition to a specific asset in some cases when this was not recorded or could not be determined by field workers.

For wildfire consequence, WRRM estimates tangible impacts (e.g., structures destroyed, acres burned) from which the MAVF core attributes are derived. Therefore, an additional layer of modeling determines the MAVF attribute values from WRRM outputs. This step requires several discretionary assumptions that are assessed prior to PSPS de-energizations and are typically reported in post-PSPS reports.

2.4.3.3 Stability of Assumptions

Stability of assumptions in the program, including historical and projected changes

As with all machine learning models, as more data is collected, the model outputs will improve. The data collection efforts leveraged by WiNGS-Ops are expected to continue and expand, and therefore it can be expected that the model performance will improve over time due to both increase amount and quality of data.

2.5 Data Libraries

Provide background information on the source, contents, and use of data libraries.

Data used for WiNGS Ops is collected from enterprise resources and centralized in an Amazon Web Services (AWS)-based cloud environment. Data sources that are external to the enterprise are brought into AWS in raw format and transformed into structured data where necessary. All data in the cloud data repository must be structured at this time. Refresh schedule varies based on the source system refresh rate. All additional data transformations are done in accordance with machine learning modeling best practices to prepare the data for ingestion and prediction.

Table 3: Data Sources

Source	Collection Frequency	Spatial Granularity	Temporal Granularity	Comment
Asset information systems (or EAMP)	Daily refresh to analytics databases	Asset-level	n/a	

Source	Collection Frequency	Spatial Granularity	Temporal Granularity	Comment
Weather station data	Daily refresh to analytics databases	222 weather stations	10 min	Historical data used for training
Weather forecast (SDSC)	Twice a day	2 km	Hourly	Made publicly available by SDG&E/SDSC
Public maps (SANGIS)*	Approx. annually	Vectorized	n/a	Public resource; does not include Orange County
SDG&E Fuel layer	Approx. annually	9 m raster	n/a	Generated by Technosylva, Inc.
Customer information systems	On demand	Transformer-level	n/a	Not currently in cloud (BPP)

*External data dependencies

2.5.1 External Dependencies

The WiNGS-Ops model is dependent on both internal Enterprise data processes with robust maintenance protocols and external data sources with varying maintenance procedures.

Public maps data is provided by the San Diego Geographic Information Source (SanGIS)[2].

2.6 Substantiation (SFPE Guide 2010)]

For machine learning models, substantiation requires review of both data and code, which occurs during several stages of the development and deployment process. During early stages of development, verification mainly occurs through a series of ad-hoc analyses, commonly referred to as “exploratory data analysis”, where the data scientist develops an understanding of the data and how code interacts with the data. The model building stage that follows is where a series of experimental models are validated using statistical methods (see Section 2.3.1.2). High-performing models may then be deployed to a “production environment”, and the process may include code revisions and/or new data pipelines to meet software production requirements. The deployment process therefore includes several additional steps of data and code verification. Finally, the resulting model outputs undergo user acceptance testing to assure that the deployed model produces the intended results.

2.6.1 Verification

Describe efforts to verify the model is working as designed and that the equations are properly being solved (e.g., independent review of source code, testing, user training, and certification).

[2] https://www.sangis.org/

Production Code Verification

In the production environment, each model run is saved with a timestamp and version number. The version schema allows for tracking software patches, minor enhancements, and major enhancements. Code revisions must undergo a quality control review where the full model is run and tested for accuracy and errors before changes are accepted. This method of code verification is adherent to software development best practices and ensures that new functionality is reviewed by multiple developers before being accepted.

Data Quality Verification

Data quality is verified in the model production environment using heuristic checks, such as identifying when certain values are beyond an expected range. Additionally, data quality is verified at the source:

- GIS Electric System data - Data obtained from GIS is digitized internally from As-built drawings and undergoes a rigorous series of quality assurance tests prior to being released as official As-built GIS features. Field quality validation is accepted on an as-needed basis.
- Outage data - Outage data undergoes an internal audit process by qualified reliability staff to verify the details surrounding the outage. The reliability staff obtains outage information from the OUA application and verifies the relevant details of the outage (such as root causes, time stamps, and customer counts) and its effects using NMS.
- Ignition data - Ignition data is collected and investigated by qualified fire coordinators. Data includes information on fires started by SDG&E electric assets.
- Weather data - Weather data is collected by real time location system (RTLS) units [anemometers and Remote Automated Weather Stations (RAWS)] and coalesced into the OSI Pi database. Meteorology maintains relationships between the weather stations and electric assets.
- Vegetation data - Vegetation data is collected and maintained by Vegetation Management, who has ongoing maintenance to ensure inspection information is current and correct.

2.6.2 Validation

Identify existing data that can be used to validate model performance. Describe how model predictions are compared to observations from historical events or experiments.

Sub models are validated each year, see Table 4 for the sub-model performance validation summary.

Table 4: Machine Learning Sub-Models Validation Summary

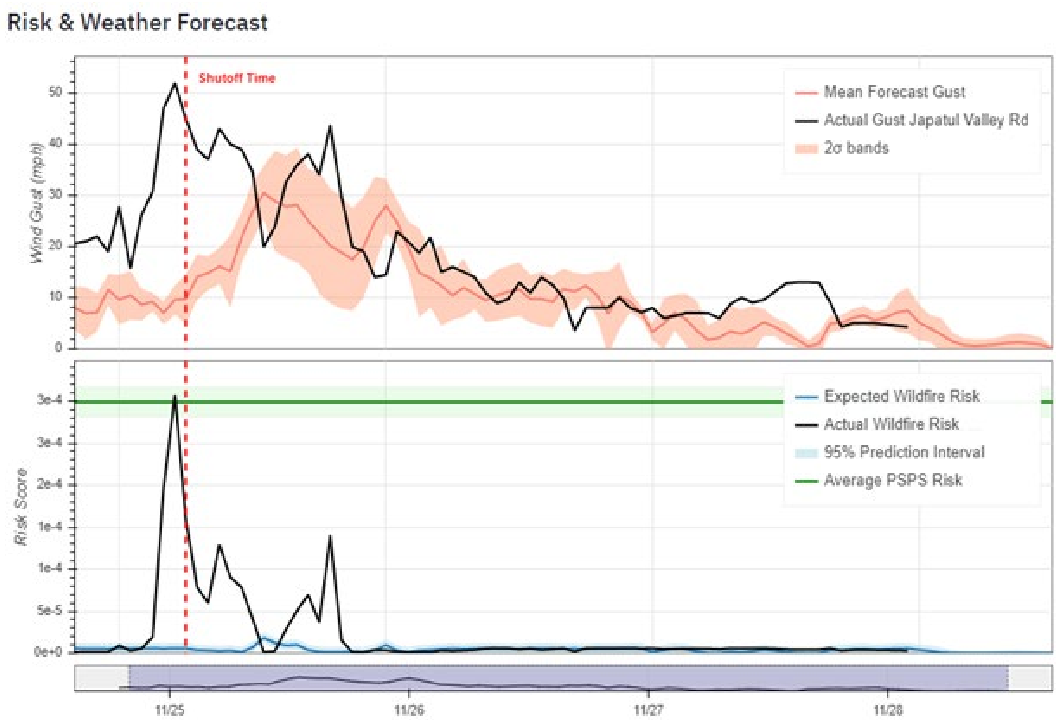
Model	Function	Validation Method	SME	Performance Metric
Conductor Failure	Wind gust, wind direction, conductor type, elevation	Goodness of fit	Electric District Operations	R-squared

Model	Function	Validation Method	SME	Performance Metric
			feedback on feature selection	
Balloon Contact	Time of day, day of week, month, land use (population) density	n/a	n/a	MSE
Animal contact	Outages, conductor attributes, species habitat models, vegetation, spatial	n/a	n/a	n/a
Vegetation	Number of trees, tree species	Cross validation	Vegetation management	n/a
Vehicle	Pole location, attributes, road attributes, landmarks	Cross validation	Electric District Operations feedback on feature selection	n/a
Span	Forecasted wind gust, forecasted temperature, fuel source prevalence, wire type, wire length	Cross validation	Fire science review, Technosylva feedback and fuel layer	ROC AUC
Pole	Pole age, pole material, pole class, number of wires, upstream sectionalizer type, WRRM value	n/a	n/a	n/a

WiNGS-Ops model outputs undergo significant review, observation, and scrutiny. Due to its novelty, there are no clear methods for validation. However, sensitivity analyses are being conducted and results are benchmarked with past decisions to determine areas of improvement and whether the quantifications are adequate.

Figure 6 demonstrates the ongoing analysis of the WiNGS-Ops results. The graphs show data from an actual PSPS de-energization where PSPS protocols were activated. The top graph shows weather forecast and actual weather data over time, and the bottom graph shows the corresponding wildfire and PSPS risk calculations during the same timeframe. Figure 6 suggests that model outputs are responding to weather patterns as intended, with a correlation observed between high wind gust and high calculated wildfire risk.

Figure 6: WiNGS-Ops Analysis



* Hourly wind gust forecast is from SDSC GFS003 Model averaged all spans downstream of the device. 2σ bands are from the standard deviations of this distribution.

** Risk scores are calculated by CPUC-approved S-MAP MAVF Framework.

2.6.3 Calibration

Describe how model inputs and parameters are modified to achieve better agreement for a specific scenario. Calibration limits the propagation of error by correcting new data but they have limited effectiveness in improving the quality of the forecast.

A calibration step is required to ensure that the probability of failure model outputs are congruent with annual outage rates since only a sampling of “non-event” observations were used in the machine learning training set, as discussed in Section 2.4.1.1. The same sampling ratio used to generate the training set is used to scale the predicted probabilities (i.e., probabilities are lowered) so that the total number of hours in which an outage did not occur within the observation period is considered. When the probabilities are summed across all hours in a specified year, the modeled annual outage rate is estimated. Since the sampling of “non-event” observations is random and presumed to be representative, this process is not expected to introduce any bias; it is performed for computational simplicity.

3 References

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York, Springer.

Gareth James; Daniela Witten; Trevor Hastie; Robert Tibshirani (2013). An Introduction to Statistical Learning. Springer. P. vii.

Hanley, James A.; McNeil, Barbara J. (1982). "The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve". Radiology. 143 (1): 29–36. doi:10.1148/radiology.143.1.7063747. PMID 7063747. S2CID 10511727.

Statistical Modeling: The Two Cultures Author(s): Leo Breiman Source: Statistical Science, Vol. 16, No. 3 (Aug., 2001), pp. 199-215

Breiman, L. Bagging predictors. Mach Learn 24, 123–140 (1996). https://doi.org/10.1007/BF00058655