

Contents

Supplemental Appendix B: Supporting Documentation for Risk Methodology and Assessment	1
Detailed Model Documentation	2
POI	5
Wildfire Consequence	8
Model Substantiation	32
POI	34
Wildfire Consequence	37
Additional Models Supporting Risk Calculation	39
Weather Analysis	41
Fuel Conditions	42
Calculation of Risk and Risk Components	44
Likelihood	44
Consequence	48
<i>IRC3: Wildfire Consequence</i>	49
<i>FRC5: Wildfire Hazard Intensity</i>	50
<i>FRC6: Wildfire Exposure Potential</i>	50
<i>FRC7: Wildfire Vulnerability</i>	50
PSPS Consequence	51
<i>FRC8: PSPS Exposure Potential</i>	51
<i>FRC9: Vulnerability of a Community to PSPS</i>	51
Risk	51
<i>R2: Ignition Risk</i>	52
<i>R3: PSPS Risk</i>	52
<i>R1: Overall Utility Risk</i>	52
Attachment A OH-Capacitor Sub-Model	56
Attachment B OH-Conductor Sub-Model	80
Attachment C OH-Switch Sub-Model	115
Attachment D OH-Transformer Sub-Model	143

Supplemental Appendix B: Supporting Documentation for Risk Methodology and Assessment

Note: As part of its 2023-2025 WMP, the electrical corporation is required to provide the “Summary Documentation” as defined by this appendix. For all other requirements in this appendix, the electrical corporation must be readily able to provide the defined documentation in response to a data request by Energy Safety or designated stakeholders.

The risk modeling and assessment in the main body of these Guidelines and electrical corporation’s WMP are focused on providing a streamlined overview of the electrical corporation risk framework and key findings from the assessment necessary to understand the wildfire mitigation strategy presented in Section 7.

The focus of this appendix is to provide additional information pertaining to the risk modeling approach used by the electrical corporation. This includes the following:

- *Additional detail on model calculations supporting the calculation of risk and risk components*
- *Additional detail on the calculation of risk and risk components*
- *More detailed presentation of the risk findings*

The following sections establish the reporting requirements for the approaches used by the electrical corporation to calculate each risk and risk component. These have been synthesized and adapted from guidance documents on model quality assurance developed by many agencies, with a focus on guidance related to machine learning, artificial intelligence, and fire science and engineering. These guidance documents include those from the Institute of Electrical and Electronics Engineers (IEEE),¹ the Society of Fire Protection Engineers (SFPE),² the American Society for Testing and Materials (ASTM International),³ the U.S. Nuclear Regulatory Commission (NRC),⁴ the Electric Power Research Institute (EPRI),⁵ the National Institute of Standards and Technology (NIST),⁵ and the International Organization for Standardization (ISO).⁶

¹ IEEE, 2022, “P2841/D2: Draft Framework and Process for Deep Learning Evaluation.”

² SFPE, 2010, “Substantiating a Fire Model for a Given Application,” Engineering Guides.

³ ASTM, 2005, “ASTM E1472: Standard Guide for Documenting Computer Software for Fire Models,” ASTM International.

ASTM, 2005, “ASTM E1895: Standard Guide for Determining Uses and Limitations of Deterministic Fire Models,” ASTM International.

ASTM, 2005, “ASTM E1355: Standard Guide for Evaluating the Predictive Capability of Deterministic Fire Models,” ASTM International.

⁴ U.S. NRC, EPRI, Jensen Hughes, NIST, 2016, “NUREG-1824: Verification and Validation of Selected Fire Models for Nuclear Power Plant Applications. Supplement 1.”

U.S. NRC, EPRI, Hughes Associates, Inc., NIST, California Polytechnic State University, Westinghouse Electric Company, University of Maryland, Science Applications International Corporation, ERIN Engineering, 2012, “NUREG-1934: Nuclear Power Plant Fire Modeling Application Guide.”

⁵ NIST, 1981, “NBS SP 500-73: Computer Model Documentation Guide.”

⁶ ISO, 2013, “ISO/TR 16730:2013: Fire Safety Engineering: Assessment, Verification and Validation of Calculation Methods.”

Detailed Model Documentation

The electrical corporation must be readily able to provide, if requested by Energy Safety or designated stakeholders, detailed documentation for each model and sub-model discussed in the summary documentation. The electrical corporation should not provide this information as part of its WMP submission. At a minimum, this documentation to be made available on request must include each of the following:

- **Purpose of the model / problem identification:**
 - Define the objectives/goals of the model.
 - Summarize and define the relevant outcomes to be predicted by the model.
 - Define the circumstances in which the model is to be used.
 - Time horizon (i.e., real time, annual planning, or both)
 - Spatial scales (i.e., service territory, region, local)
 - Deterministic (specific forecasts) or probabilistic (statistical)
- **Model version:**
 - Provide the name and version number of the software, including major and minor release number. Provide version control (git) commit level if available.
 - Document any utility-specific changes to the model and provide the reason for the change(s).
- **Theoretical foundation:**
 - Describe the theoretical basis of the model and the governing equations or physical laws on which the model is based.
 - Identify assumptions made in the model, their impact in the governing equations, and resulting limitations.
- **Mathematical foundation:**
 - Describe numerical techniques and computational algorithms used to solve/approximate the governing equations.
 - Describe the precision of the results and any reliance on specific computing hardware or facilities.

ISO, 2021, "ISO/IEC TR 24027:2021: Information Technology: Artificial Intelligence (AI) – Bias in AI Systems and AI Aided Decision Making."

ISO, 2021, "ISO/IEC TR 24029:2021: Artificial Intelligence (AI): Assessment of the Robustness of Neural Networks."

- *Discuss model convergence criteria, studies, and resulting grid resolution required to meet the criteria.*
- *Identify any additional limitations in the model based on the numerical techniques and implementation.*
- **External dependencies:**
 - *Describe external programs or software libraries used by the software.*
 - *Describe data used by the software, including utility-collected and external sources. This should include the following:*
 - *Characteristics of the data (field definitions/schema, uncertainties, acquisition frequency).*
 - *Scope and granularity (or resolution) of data in time and location (i.e., date range, spatial granularity for each data element).*
 - *Sources of data, frequency of data updates, and verification of data quality. Explain in detail measurement approaches and procedures.*
 - *Any processes used to modify the data (such as adjusting vegetative fuel models for wildfire spread based on prior history and vegetation growth).*
- **Model substantiation:**
 - *Identify existing data that can be used to validate model performance.*
 - *All models need to be verified and validated for the specific application in which they are to be used in accordance with the guidance provided in Section “Model Substantiation,” below.*
- **Sensitivity**
 - *Describe the efforts to evaluate the impact of model and input parameter uncertainty on the model predicted outcomes.*
 - *Describe the efforts to evaluate the propagation of uncertainty into downstream models.*

One approach to fulfill these requirements is to provide the following documents to demonstrate the substantiation of each model:

- **Technical documentation** according to ASTM E 1472 – Standard Guide for Documenting Computer Software for Fire Models. Include a listing of assumptions and known limitations of the model according to ASTM E 1895 – Standard Guide for Determining Uses and Limitations of Deterministic Fire Models.
- **Verification and validation documentation** according to the SFPE’s Guidelines for Substantiating a Fire Model for a Given Application or ASTM E 1355 – Standard Guide for

Evaluating the Predicting Capability of Deterministic Fire Models.

In lieu of providing customized documentation, the electrical corporation may provide a copy of documentation generated by a commercial provider of a model or an open-source project if all the following conditions are met:

- The specific version documentation of the model and any underlying data in use by the electrical corporation are the same.*
- Any custom modifications to the model by the electrical corporation have been integrated into the model documentation and are available in the same format as the model (i.e., custom modules to an open-source project must be open source and integrated into the project).*
- The electrical corporation lists and justifies the options used within the model for its application, including all non-default features or assumptions.*

POI

SCE's Probability of Ignition (POI) model is composed of the following sub-models:

- OH Conductor Model (includes both CFO & EFF)
- OH Transformer
- OH Switch
- OH Capacitor

SCE has prepared sub-model documentation that is provided in Attachments A through D of this document.

Below SCE has responded to each Energy Safety requirement or indicated where it is discussed in the sub-model documentation. In most cases SCE's responses are provided at the level of the sub-model, with the intention that this provides a more granular level of information. The overall POI model is discussed in the main body of the WMP (see Section 6).

Purpose of the model / problem identification:

Define the objectives/goals of the model.

The objectives and goals of SCE's POI sub-models is discussed in Section 1.1 of the model documentation.

Summarize and define the relevant outcomes to be predicted by the model.

The relevant outcomes to be predicted by the model of SCE's POI sub-models is discussed in Section 1.1 of the model documentation.

Define the circumstances in which the model is to be used.

The circumstance when the SCE's POI sub-models is used is discussed in Section 1.1 of the model documentation.

Time horizon (i.e., real time, annual planning, or both)

The time horizon of the data used in SCE's POI sub-models is discussed in Section 1.1 of the model documentation.

Spatial scales (i.e., service territory, region, local)

SCE models at the individual asset or segment level in SCE's territory depending on the type of asset (e.g. linear assets such as conductor or cable are modelled as segments, pieces of equipment like transformers and switches are modelled at the asset location).

Deterministic (specific forecasts) or probabilistic (statistical)

The type of modeling that SCE's sub-models uses is discussed in Section 1.2 of the model documentation.

Model version:

Provide the name and version number of the software, including major and minor release number. Provide version control (git) commit level if available.

The name and version of SCE's sub-models can be found on Title page of the model documentation and under the Document Version History section.

Document any utility-specific changes to the model and provide the reason for the change(s).

The specific changes to SCE's sub-models are tracked in the Document Version History section of the model documentation.

Theoretical foundation:

Describe the theoretical basis of the model and the governing equations or physical laws on which the model is based.

The theoretical basis of the model and the governing equations or physical laws on which the model is based of SCE's sub-models is discussed in Section 2 of the model documentation.

Identify assumptions made in the model, their impact in the governing equations, and resulting limitations.

The assumptions made in SCE's sub-models can be found in Section 1.5 of the model documentation.

Mathematical foundation:

Describe numerical techniques and computational algorithms used to solve/approximate the governing equations.

The numerical techniques and computational algorithms used in SCE's sub-models is discussed in Section 2.2 of the model documentation.

Describe the precision of the results and any reliance on specific computing hardware or facilities.

The precision of the results and any reliance on specific computing hardware or facilities used in SCE's sub-models is discussed in Section 3 of the model documentation.

Discuss model convergence criteria, studies, and resulting grid resolution required to meet the criteria.

The model convergence criteria, studies, and resulting grid resolution used in SCE's sub-models is discussed in Section 2.2 of the model documentation.

Identify any additional limitations in the model based on the numerical techniques and implementation.

Limitations in the model based on the numerical techniques and implementation in SCE's sub-models is discussed in Section 1.6 and Section 2.5 of the model documentation.

External dependencies:

Describe external programs or software libraries used by the software.

External programs or software libraries used by the software in SCE's sub-models is discussed in Section 1.2, Section 3.1 and Section 3.4 of the model documentation.

Describe data used by the software, including utility-collected and external sources. This should include the following:

- *Characteristics of the data (field definitions/schema, uncertainties, acquisition frequency).*
- *Scope and granularity (or resolution) of data in time and location (i.e., date range, spatial granularity for each data element).*
- *Sources of data, frequency of data updates, and verification of data quality. Explain in detail measurement approaches and procedures.*
- *Any processes used to modify the data (such as adjusting vegetative fuel models for wildfire spread based on prior history and vegetation growth).*

External programs or software libraries used by the software in SCE's sub-models is discussed in Section 1.2, Section 3.1 and Section 3.4 of the model documentation.

Model substantiation:

Identify existing data that can be used to validate model performance.

The process used to validate model performance used in SCE's sub-models is discussed in Section 3 of the model documentation.

All models need to be verified and validated for the specific application in which they are to be used in accordance with the guidance provided in Section "Model Substantiation," below.

The process used to verify and validate the model used in SCE's sub-models is discussed in Section 3 of the model documentation.

Sensitivity

Describe the efforts to evaluate the impact of model and input parameter uncertainty on the model predicted outcomes.

The sensitivity analysis used in SCE's sub-models is discussed Section 3.2 of the model documentation.

Describe the efforts to evaluate the propagation of uncertainty into downstream models.

The efforts to evaluate the propagation of uncertainty into downstream models used in SCE's sub-models is discussed in Section 3.2 of the model documentation.

Wildfire Consequence

SCE provides the detailed documentation of its Wildfire Consequence model below. Please note that the information provided was adapted from materials provided to SCE by Technosylva.

Purpose of the model / problem identification:

Define the objectives/goals of the model.

The purpose of the model is to simulate the consequence associated with ignitions emanating from overhead utility assets.

Summarize and define the relevant outcomes to be predicted by the model.

The results of these ignition simulations are used to assess the relative consequences associated with ignitions across a wide range of fuel and wind weather scenarios. These consequences include acres burned, population impacted, and buildings impacted. Consequences are determined based on an 8-hour duration representative of a typical first burning period (unsuppressed).

Define the circumstances in which the model is to be used.

The intent of the model is to provide a relative ranking of wildfire consequence based on a representation of the maximum consequences simulated for all 29 million ignition points. Simulations were run using the worst observed fire weather condition across 444 weather days. These days represent critical fire weather conditions that occurred throughout the SCE service area.

Time horizon (i.e., real time, annual planning, or both)

Planning is based on historical climatology (design scenario) along with associated fuel regrowth for major fire scars projected out to the year 2030.

Spatial scales (i.e., service territory, region, local)

Ignition points simulations in proximity to all overhead utility assets in HFTD, plus an additional 20-mile buffer (Wildfire Risk Reduction Model 6). Ignition simulations are conducted within 250 meters of each overhead SCE asset. Additionally, independent ignition simulations are in HFTD locations in which SCE assets are not present. These simulations are spaced roughly 1,000 meters apart.

Deterministic (specific forecasts) or probabilistic (statistical)

Deterministic simulations are performed for each ignition location across 444 weather scenarios representative of fire weather conditions (e.g., windy and/or dry fuel conditions).

Model version:

Provide the name and version number of the software, including major and minor release number.

Provide version control (git) commit level if available.

The software is Technosylva Wildfire Analyst Enterprise (WFA-E).

Document any utility-specific changes to the model and provide the reason for the change(s).

The specific version of inputs and outputs used by SCE is known as the SCE Wildfire Risk Reduction Model (WRRM) version 6.0. This version of the model employs SCE specific asset, historical weather, and fuel information.

Theoretical foundation:

Describe the theoretical basis of the model and the governing equations or physical laws on which the model is based.

Fire is a self-sustained and uncontrolled chemical processes as a result of the combination of fuel, oxygen, and an ignition source.

Wildfires are a type of fire in which the fuel source is comprised of dead and living foliage; oxygen is abundantly present in the atmosphere; and the ignition source is usually an external source (e.g., electrical equipment, lightning, etc.). Additional combustion could also occur due to radiation and convection of heat through and within surrounding vegetation.

Wildland fire behavior is highly dependent on vegetation composition and type (e.g., spatial patterns, vertical arrangement, state of dryness), terrain (e.g., slope, aspect, geology), and weather (e.g., temperature, wind velocity, duration, direction, relative humidity).

The behavior of wildfire can be described through fire intensity metrics based on a combination of the conditions present through the fire event. These observable factors include Rate of Spread (ROS), flame length, flame intensity, heat per unit area, flame depth, and residence time.

Based on this behavior, a fire may be classified as a surface or a crown fire. Surface fires burn loose needles, moss, lichen, herbaceous vegetation, shrubs, small trees and saplings that are at or near the surface of the ground.

Crown fires burn forest canopy fuels, which include live and dead foliage/branches, lichens in trees, and tall shrubs that lie well above the surface fuels. They are usually ignited by a surface fire. Crown fires can be passive or active. Passive crown fires involve the burning of individual trees or small groups of trees (often called torching). Active crown fires, or also referred to as running crown fires, present a solid wall of flame from the surface through the canopy fuel layers.

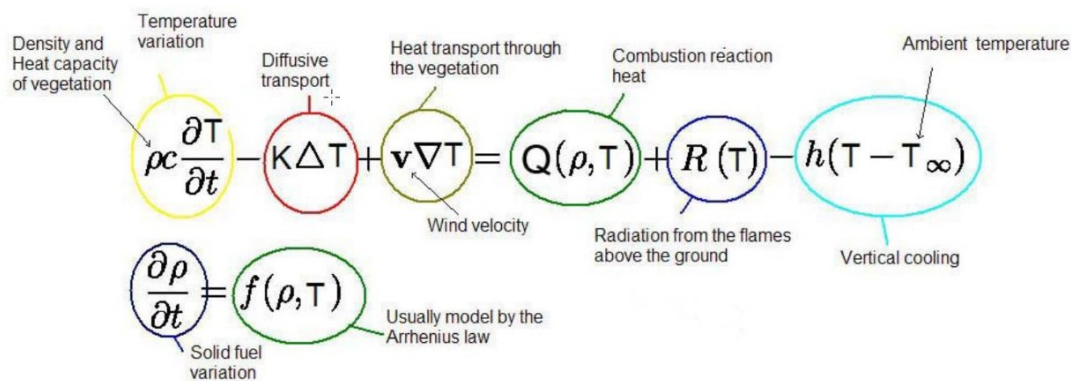
Fire growth from an ignition point can be split into four distinct phases (Fire Science 2017). In the

first phase, the fire starts to burn slowly as the influx of air caused by the buoyancy flow of hot gasses causes the flames to tilt inwards. Once the fire has spread enough from the ignition point, wind is able to enter the already burn vegetation and pushes the flames away from the center and tilts them towards the unburned fuels, increasing the heat transfer, and therefore accelerating the fire. As the fire moves further away from the center, the acceleration of the fire depends more on the local characteristics of the curvilinear front. Finally, the fire may reach a steady state when the fire line is uniform enough so that it can be considered of infinite length.

Wildfire modelling is a highly complex challenging problem based on both the physical parameters involved (e.g., wildfire attributes), as well as the amount of data and computational intensity involved. Wildfire modelers need to balance the accuracy of the model with practical limitations in data availability, computing power, as well as the intended use of the end result. Approaches to wildfire modeling can be primary broad approaches are physical models, quasi-empirical models, and empirical ones.

Physical models are the most complex and are considered to be valid across different fuels and weather conditions (Cruz 2017). These models are usually described as a set of coupled differential equations derived from conservation laws and generally involve representing the vegetation layer as a porous medium in which wildfires progress. The degree of approximation of the initial semi-physical description of the problem, as well as the rest of physical effects considered in the modelling may vary from one model to another. Despite these different approaches, a conventional 2D multiphase model involves sketching vegetation temperature through a convection reaction diffusion equation, and a solid combustible material evolution over temporal and spatial domains. These models can be represented by the following equation:

Figure Appendix B-01 - Illustrative physical model describing wildfire progression across vegetation



Even though physical models are a fair approximation of wildfire spread, these models are difficult to operationalize due to limitations in the amount of data available, as well as the computing intensity required.

Alternatively, empirical and semi-empirical models are based on experimental data (e.g., laboratory

runs, controlled outdoor fires, or well documented wildland fires). The difference between the empirical and semi-empirical approach is that the former contains no physical basis and are generally solely statistical experiments, while the later uses a form of physical framework, in which the statistical model is based (Andrews 2018, Sullivan 2009). These models are largely developed to support decision making and are the main operational models SCE's vendor, Technosylva employs in its wildfire modeling.

Identify assumptions made in the model, their impact in the governing equations, and resulting limitations.

The following are the assumptions made in the model.

- Ignition simulations represent an idealized situation in steady state spread which may not fit some extreme behavior of fires
- Fuels are assumed to be continuous, and uniform based on the spatial scale of the input data (typically between 10-to-30-meter (m) resolution)
- Wildfire characteristics for a given ignition point are dependent on the conditions at that location. There may be some non-local phenomena which may not be fully captured in the ignition simulation, such as:
 - Increased ROS due to a concave front.
 - Fire interaction between different parts of the same fire or a different one.
 - Fire spread is assumed to be elliptical although there are several variations of wildfire spread in observed fires, such as double ellipse, oval, egg-shaped, etc.
 - Weather is inputted into the model at hourly increments and is assumed to remain constant between those increments. There is no interpolation of those hourly time components to compute a more granular evolution of weather between hours.
 - The reliability of weather inputs in the mid-range forecast (2 to 5 days).
 - Wildfire is not coupled with the atmosphere in any way. This may seem like a major limitation in the model as wind is a main contribution to fire spread and at present many models (especially physical ones) try to couple wind and fire. The main reasons for us not to consider the coupling is:
 - It would make it unfeasible to run millions of simulations considering the coupling effect.
 - Empirical and semi-empirical models have been developed using an average wind speed as an input, so it is not clear that considering more granular wind at the front is advisable.
 - Fire is always assumed to be fully developed. Fire acceleration, flashover, or decay is not considered.
 - Atmospheric instability which may have a deep impact on ROS (Beer 1991) is not considered in the model.
 - Gusts are not considered in the model.
 - There is no interaction between slope and wind other than creating an effective

or equivalent wind. This means that fire is assumed to have an elliptical shape irrespective of the alignment of wind and slope.

- Models have been developed with extensive amounts of empirical data.
- Fuel array description of the vegetation may not perfectly describe fuel characteristics.
- Spotting is only considered in surface fires.

Mathematical foundation:

Describe numerical techniques and computational algorithms used to solve/approximate the governing equations.

The mathematical model used to simulate surface fire spread is the model developed by Rothermel (1972) with some modifications from Albini (1976) and some minor adjustments from Technosylva. It accepts the initial 13 fuel models (Anderson 1982) as well as Scott and Burgan's (2005) dynamical fuels where there is a transfer load between the herbaceous and dead classes.

Among other outputs this model provides the surface fire rate of spread, flame length and flame intensity in the direction of maximum spread (head front). Crown fire is implemented using the model developed by Van Wagner (1977, 1993) which computes the transition viability to crown fire, as well as the expected ROS and intensity in active crown fires. Spotting is modeled as a pseudo random event. The maximum expected spotting distance from the fire is obtained using the wind-driven model developed by (Albini 1983a; Albini 1983b; Chase 1984) and then embers are generated randomly on the front of the fire and the actual traveled distance is computed also randomly based on the maximum distance available.

In this modeling there is no tracking of individual embers in the air. Wind speed profiles at different heights (2m, 10m, 20ft) are obtained through a logarithm wind profile found in Andrews (2012). Fire is assumed to spread following an elliptical shape dependent only on the effective wind speed (Andrews 2012). The time evolution is done using a Fast-Marching method on a regularly spaced landscape grid of a Cellular Automata.

Surface Fire

The default propagation engine implemented in WFA-E is Rothermel's (1972) surface model with the modifications proposed by Albini (1976) and the requirements to accept Scott and Burgan (2005) fuel models. The basic equation in the model predicts the heads fire rate of spread without wind or slope:

$$R_0 = I_R \xi / \rho_b \epsilon Q_{ig}$$

Here I_R is the reaction intensity (energy released rate per unit area of the fire front), ξ the propagating flux ratio, ρ_b the bulk density, ϵ the effective heating number, and Q_{ig} the heat of ignition. The equation is derived by applying the energy conservation to a unit volume of fuel ahead of a steadily advancing fire in a homogeneous fuel bed. In this model, the ROS may be viewed as the ratio between the heat flux received by the unburned fuel ahead of the fire (numerator) and the heat required to ignite it (denominator).

The input parameters to compute the ROS in the case of no wind or slope are the moisture content and the characteristics of the vegetation. Moisture content is given by the 1h, 10h and 100h dead moisture content, and the woody and herbaceous live moisture content. Fuels are assumed to be a mixture of different vegetation types depending on their class (dead or live) and size (less than 0.25 inch, 0.25-1 inch, 1-3 inch), with each class having different surface to volume ratio and loads. The inputs required to define a fuel type is given in the following table:

Table Appendix B-01 - Input variables for each fuel type

			LOAD				SAV					
Fuel	1h	10h	100h	herb	woody	1h	herb	woody	Dyn	Depth	MoistExt	heat

Here Dyn (dynamic) is a boolean variable to define if there should be a transfer between the herbaceous load and the dead load based on the herbaceous content. In general, SAV values (the fineness of the fuel) strongly affects the ROS and flame length of the fire, while the fuel load does not affect the rate of spread but can have a strong effect on the flame length.

The effect of wind and slope can be incorporated in the model through a couple of dimensionless parameters depending on the midflame wind speed U and the terrain angle θ :

$$ROS = R_0 (1 + \Phi_w + \Phi_s)$$

with

$$\Phi_s = 5.275 \beta - 0.3 (\tan \theta)$$

$$\Phi_w = C * U^B (\beta / \beta_{op})^{-E}$$

Where β_{op} and β are the optimum and standard packing ratios respectively, and C , B , and E are parameters depending on the surface to volume ratio σ :

$$C = 7.47 * \exp(-0.133 \sigma^{0.55});$$

$$B = 0.02526 \sigma^{0.54}$$

$$E = 0.715 * \exp(-0.000359 * \sigma)$$

The slope and wind factors are summed together to obtain the final ROS. If they are not aligned, the resultant vector defines the direction of maximum spread (which will be between the direction of wind and the direction of slope). This final slope-wind factor can also be used to compute an equivalent or effective wind speed causing the same effect as the combined effect of wind and slope. To do that we simply inverse the equation of the wind factor to obtain:

$$U_e = [\Phi_e (\beta / \beta_{op})^E / C]^{-B}$$

The Rothermel model predicts fire characteristics (ROS, flame length, etc) only in the direction of maximum spread (head front) obtained from the combined effect of wind and slope. To compute the ROS in a direction different from the direction of maximum spread, and to be able to use the model in a 2D landscape, it is assumed that a free burning fire perimeter from a single ignition point has an elliptical shape. There are several different approaches to compute the ellipse (or ellipses) eccentricity based on wind and slope (Albini [2], Anderson 1983 [6], Alexander, etc). The present implementation follows the equations in Andrews (2008) depending on the effective wind speed U_e in mi/h in the direction of maximum spread. The length to width ratio is given by:

$$L/W = 0.1 + 0.25 U_e$$

Or equivalently the eccentricity e is given by

$$e = (Z^2 - 1)^{0.5} / Z$$

so that the ROS in any direction ϕ is given by

$$ROS(\phi) = ROS(1-e) / (1+e)$$

One of the most important variables of fire is the amount of heat it generates as this is the main contributor to fire spread and fire severity. The amount of heat can be measured using different variables

like the reaction intensity (IR), the Heat per Unit Area (HPA) or the fireline intensity. The Reaction intensity is the rate of energy release per unit area within the flaming front (with units of energy/area/time), heat per unit area is the amount of heat energy released per unit area within the flaming front (units of energy/area), fire line intensity is the rate of heat energy released per unit time per unit length of the fire front (units of energy/distance/time). Fireline intensity is independent of the depth zone and it is calculated as the product of the available fuel energy and the ROS of the fire (Byram 1959):

$$I_B = H_A \cdot ROS$$

Where the heat per unit area depends on the reaction intensity of the fire (IR) and the time that the area is in the flaming front (residence time t_r)

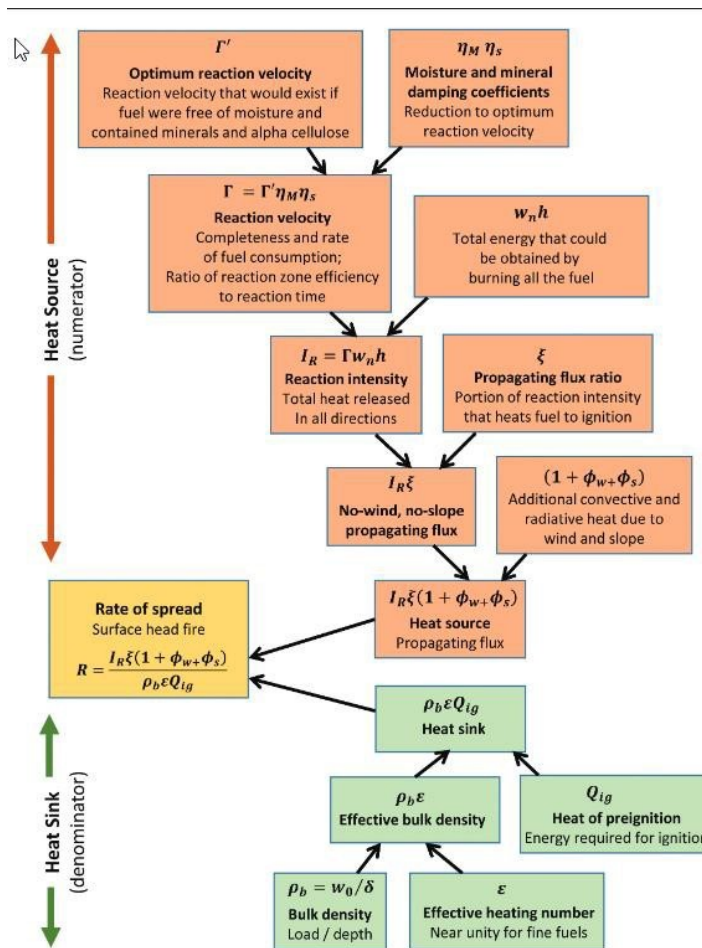
$$H_A = I_R \cdot t_r = 384 \cdot I_R / \sigma$$

In this model the flame length and Byram's intensity are closely related by:

$$FL = 0.45 I^{0.46}$$

Where the flame length is in feet and the intensity in Btu/ft/sc.

Figure Appendix B-02 - Flow of Calculation provided in Andrews (2018)



For a much more in-depth discussion of the Rothermel surface model please read Andrews (2018) and Rothermel (1972).

Crown fire

Crown fires burn forest canopy fuels. They are usually generated by surface fires and represent a major change in fire behavior due to an increased rate of spread and heat released. Crown fires can be passive, active or conditional based on the capacity of the surface fire to move into areal fuels, and to the capacity of the burning canopy to move between individual trees.

Crown fire initiation occurs when the surface fire provides enough heat to raise the temperature of the canopy fuel to ignition temperature. In Van Wagner (1977) model, this minimum intensity is given by:

$$I_{ini} = (0.01 \cdot CBH (460 + 25.9 FMC))^{1.5}$$

Where CBH is the canopy base height (m) and FMC is the foliage moisture content of the canopy cover. Foliar moisture content (FMC) is usually not known, but it is assumed that for most species old foliage should be around 100 percent and this value has been used as a default value when no other information is available (Scott 2001). This approach however does not consider any known humidity conditions of the site and in WFA the FMC is computed based on the 100h moisture content as follows:

$$FMC = 75 + 2 \cdot m100h$$

Once the fire has transitioned to the canopy it is necessary to have a critical mass-flow rate for the fire to be self-sustained. Vang Wagner found this critical mass to be $0.05 \text{ kg m}^{-2} \text{ sec}^{-1}$ (Scott 2001) which can be used to determine a minimum crown fire rate of spread only dependent on the Canopy Bulk Density (CBD) and given by:

$$R_{active} = 3 / CBD$$

Other existing models not used in WFA-E are Alexander (1998) which is very similar to Van Wagner (1977) but includes additional inputs like flaming residence time, plume angle and fuel bed characteristics, Cruz et al. (1999) fire transition model, and Cruz et al. (2002) crown fire spread model given by:

$$ROS = c1 U^{c2} CBD \cdot C3 \cdot e^{c4 \cdot EFM}$$

Where U is the wind at 10m, CBD the canopy bulk density, EFM is the fine dead moisture content, and C1, C2, C3, C4 are a set of regression coefficients.

The model for the ROS of crown fires was computed by Rothermel (1991) through a linear regression between observed crown ROS and the surface fire model. It states that the crown fire of an active ROS is 3.34 times the rate of spread of the surface model 10 assuming a 0.4 wind reduction factor.

$$R = 3.34(R_{10})_{40\%}$$

Based on these conditions, crown fire may be classified as:

- Surface fire if neither the intensity nor the minimum crown ROS is met
- Passive Crown fire (torching): Fire spreads through the surface fuels, occasionally torching overstory trees. Overall ROS is that of the surface fire.
- Conditional Crown: Fire cannot transition to crown, but active crown fire is possible if there was a fire transition to crown by other means
- Active Crown: Fire spreads through the overstory tree canopy if both conditions are met

Figure Appendix B-03 - Crown fire classification as shown in BehavePlus

Fire Type		Active crown fire?	
		No	Yes
Transition to crown fire?	No	Surface	Conditional Crown
	Yes	Torching	Crowning

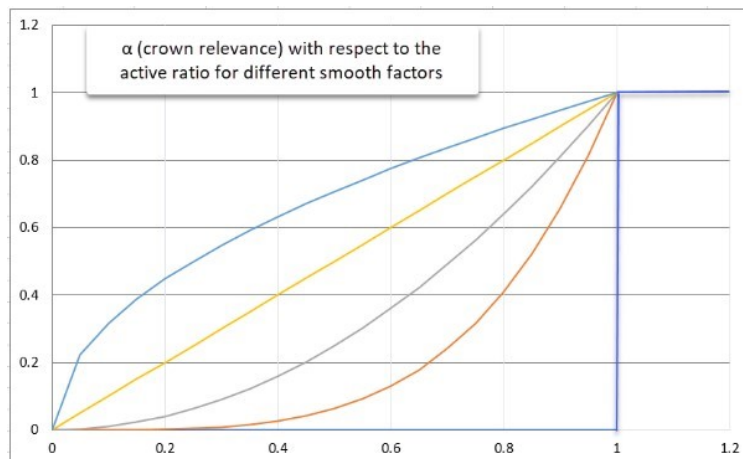
Van Wagner's crown fire transition and propagation models are well known and used by fire suppression agencies but have shown to have a significant underprediction bias when used in assessing potential crown fire behavior in conifer forests of western North America (Cruz et al. 2010). To try to correct this bias, Technosylva has introduced two new parameters in the model that has been adjusted based on the analysis carried out by the scientific team using data from the last two fire seasons in California. The model introduces two new parameters 1) a crown factor multiplier for the Canopy Bulk Density (CBD) which decreases the minimum crown ROS required to have an active crown fire, and a factor that forces a smooth transition between the surface and the crown fire behavior. The final ROS of the overall fire when crown fire type is conditional, or crowning is a weighted average of surface and crown ROS

$$ROS = surfROS * (1 - \alpha) + \alpha * crownRos$$

Where the value α ranges from 0 to 1 and depends on the **active ratio** in the following way:

$$\alpha = activeRatio^{1/smoothFactor}$$

Figure Appendix B-04 - Example effect of the smooth factor (0 blue, 0.25 red, 0.5 gray, 1 yellow) in the crown contribution for active ratios lower than 1



At present, with WFA-E the crown CBD factor is set to 1.2 and the smooth factor to 0.4. This approach to provide a gradual transition in the fire's rate of spread (and flame length) from the initial onset of crowning similar to the crown fraction burned (CFB) (Alexander 1998) used in other modelling systems like FlamMap, FARSITE or Nexus, with the main difference being the smoothing function itself. Cruz et al. observed that there is no evidence of such a smooth transition between surface and crown fire regimes in the experimental data but rather an abrupt transition is observed far more commonly. In this context, however, where the main aim is to produce a forecast risk and not to simulate an individual fire, is important to reflect the fact that the fire conditions are close to generating an active crown fire. For a more in-depth discussion of the crown fire models please read Cruz et al (2010) Scott et al. (2006)

Wind adjustment factor

Fire simulations require wind speed at midflame to compute surface fire spread and at 20ft to compute crown fire characteristics. To convert the wind between the two heights, WFA-E uses the wind adjustment factor (WAF) found in Andrews (2012) and implemented in the software BehavePlus and Farsite. The model is based on the work of Albini and Baughman (1979) and Baughman and Albini (1980), using some assumptions made by Finney (1998). This implementation considers two different models for sheltered and unsheltered conditions from the overstory. As described in Andrews (2012), the unsheltered WAF is based on an average wind speed from the top of the fuel bed to a height of twice the fuel bed depth. The sheltered WAF is based on the assumption that the wind speed is approximately constant with height below the top of a uniform forest canopy. Sheltered WAF is based on the fraction of crown space occupied by tree crowns. The unsheltered WAF model is used if crown fill portion is less than 5 percent. Midflame wind speed is the 20-ft wind multiplied by the WAF.

Unsheltered WAF depends on the surface fuel bed depth (in feet):

$$WAF = \frac{1.83}{20 + 0.36H} \ln(0.13H)$$

Sheltered WAF:

$$WAF = \frac{0.555}{\sqrt{fH} * \ln(0.13H)}$$

With H, the canopy height, and f, the crown fill portion, depending on the canopy cover (CC) and the crown ratio (CR):

$$f = CC * CR / 3$$

$$CR = (CH - CBH) / CH$$

CR is the ratio of the crown length to the total height of a tree.

Time evolution

The fire models can predict the potential ROS of the front at any point and direction but is not able to compute the evolution of the fire perimeter in time. The main models to do that are:

- 1) Using Huygens principle of wave propagation like in Farsite software and discretizing in time
- 2) Using a Minimum Travel Time Algorithm or Fast Marching method, and discretizing in space
- 3) Using the more general but usually slower Level Set Method.

In the context of wildfires, Huygens principle states that each point on a fire front is in itself the source of an elliptical wavelet (fire) which spreads out in an independent way in the forward direction. This approach is numerically solved by splitting the perimeter into a set of nodes, computing the evolution of those nodes in the direction normal to the perimeter based on the ROS given by the propagation model and a given time step, and then reconstructing the front based on the position of the transported nodes.

The main weakness of vector-based approaches is the need for a computationally costly algorithm for generating the convex hull fire-spread perimeter at each time step, especially in the presence of fire crossovers and unburned islands (Ghisu et al. 2014). Raster based implementations are computationally more efficient (Glaser et al. 2008) but can suffer from significant distortion of the produced fire shape if the number of neighboring cells considered (number of possible spread directions) is low.

Encroachment

Encroachment is a critical component in the WFA-E fire modeling simulations as it affects the number of buildings, assets, facilities and population impacted. It does not have a relevant effect on other impact metrics. To take advantage of enhanced algorithms for spread encroachment using adjacent fuels and fire behavior data, the non-burnable (and especially urban) fuel classification needed to be updated to provide better granularity and characterization of the type of urban/WUI. Accordingly, to test these methods an enrichment of the current fuels data was developed by Technosylva to delineate urban fuels into different types of urban coupled with a density level of buildings. This enhancement of the basic Scott and Burgan fuel models is used in combination with enhanced encroachment algorithms to more accurately calculate potential impacts to buildings and population.

Urban areas have been classified into classes depending on their structure (roads, urban core, isolated, sparse) and their surrounding fuels, characterized as high versus low fire behavior fuels). Specific encroachment factors can then be applied to each grouping.

Spark Modeling

Electrical failures can cause sparks and produce an ignition meters away from the asset location. To take this into account, the WFA-E allows the ignition point location to be displaced if the underlying vegetation type is either non-combustible or WUI. This displacement is in the direction of the wind and is proportional to the wind speed. The displacement distance and wind speed algorithm has been developed using expert opinion from electric utility engineers familiar with asset failure and ignition probability.

Weather

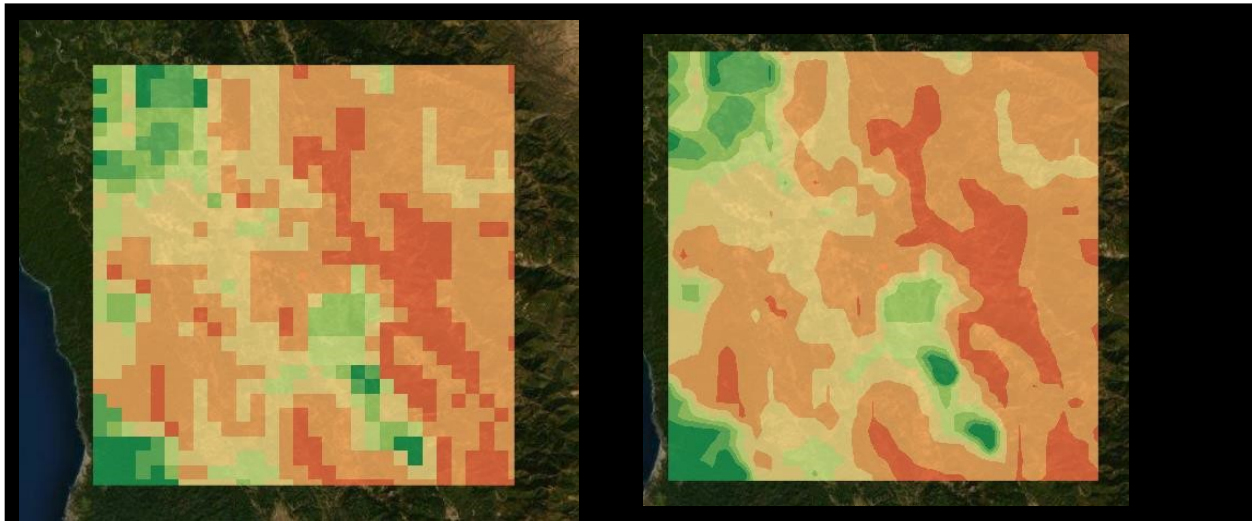
WFA-E requires historical daily weather data to run the fire simulations. The minimum required variables are the wind speed at 10m, the dead moisture content, and the live moisture content. More explicitly:

- Northward 10m wind speed
- Eastward 10m wind speed
- Dead moisture content 1hr
- Dead moisture content 10hr
- Dead moisture content 100hr
- Herbaceous moisture content
- Woody moisture content

The dead moisture is provided by SCE, but the herbaceous moisture is computed using Technosylva's Machine Learning algorithm based on historical NDVI output. The input wind speed required by the propagation model is 20ft; to convert the initial 10m wind speeds to 20ft, we use a logarithmic profile from Andrews (2012) leading to a 13% wind speed reduction.

SCE's weather and fuels forecasts are generated using the Weather Research and Forecasting (WRF) Model, in which large scale weather models from the National Center for Environmental Prediction (NCEP) are downscaled locally to 2 km. This can sometimes lead to sharp changes in weather conditions between neighboring cells. In order to increase accuracy and meet the underlying 30m cell size resolution of the fuels data, weather data is interpolated spatially using a bilinear interpolation scheme. The smoothing of the source weather data ensures that integration with the wildfire behavior models results in outputs that do not have hard edges in the data.

Figure Appendix B-05 - Left: Initial weather definition; Right: interpolated weather definition



Impact and consequence value calculation

Wildfire spread modeling is undertaken with asset ignition locations to derive potential impacts. The output impact values (risk metrics) are assigned back to the asset ignition point location. Using this approach allows SCE to differentiate between the risk output associated with different assets (and their ignition locations) using the same weather data although weather values may vary based on spatial location and time of day (hourly). For both operational and mitigation applications, the wildfire spread modeling is conducted using High Performance Computers (HPC) and typically involves hundreds of millions of spread simulations. The amount of simulation will vary depending operational use with daily forecasts versus mitigation planning use with hundreds of weather scenarios.

The main goal for the WFA-E simulations is to create a forecast risk associated to each ignition point and surrounding area. This is done by running individual simulations and associating the following main risk metrics back to each ignition point. The following baseline risk metrics are calculated from the spread simulations

- Acres Burned (referred to as Fire Size Potential)
- Number of Buildings Threatened
- Estimated Number of Buildings destroyed
- Population impacted

Numerous conventional fire behavior outputs are also calculated, the most important being:

- Rate Of Spread (ROS)
- Flame Length (FL)
- Fire Behavior Index (FBI) – combination of ROS and FL

Describe the precision of the results and any reliance on specific computing hardware or facilities.

Technosylva WFA-E is run in a cloud computing environment and does not require any additional specialized computing hardware or facilities.

Discuss model convergence criteria, studies, and resulting grid resolution required to meet the criteria.

The fire propagation model in WFA-E is a point-punctual model where the fire characteristics at a given point (cell) only depends on the conditions at that cell (weather, terrain, vegetation). This fits well in fire simulation as most of wildfire characteristics mainly depend on local characteristics (Di Gregorio et al 2003) but excludes the effects of non-local phenomena.

The overall resolution is done using a Cellular Automata (CA) where space is discretized into cells (from 10 m to 30 m resolution), and physical quantities take on a finite set of values at each cell. The potential rate of spread (ROS) at each cell at any time is given by the propagation models (surface and crown fire). CA models directly incorporate spatial heterogeneity in topography, fuel characteristics, and meteorological conditions, and they can easily accommodate any empirical or theoretical fire propagation mechanism, even complex ones (Collin et al. 2011).

Spotting is introduced as a random event where firebrands can be lifted and generate secondary ignition points ahead of the fire (in the direction of the wind). The time evolution is done using a Minimum Travel Time (Fast-Marching) algorithm. This algorithm is similar to the well-known Dijkstra's (1959) algorithm but more adapted to grids instead of the original model that uses graphs. This approach has been used with success in many forest fires propagation models like FlamMap (Finney 2002) and many others (CITES).

The algorithm provides a solution of the Eikonal equation of a spreading curve subject to a given speed function $ROS(x)$. This is done by searching for the fastest fire travel time along straight line transects of neighboring cells in the lattice. The number of neighboring cells considered determines the angle discretization of the spreading fire. The neighborhood or degrees of freedom, u , in WFA-E ranges from 8 cells (Moore neighborhood) to 32 cells.

The Technosylva WFA-E platform utilizes numerous models to address specific operational requirements. These models are integrated into an extendible platform that facilitates continued improvement as R&D advancements are made. The following table lists the primary models employed on WFA-E.

Table Appendix B-02 - Primary Models Employed

Model	Model Reference	Notes
Surface fire	Rothermel 1972, Albini 1976 Kitral IntecChile	WFA-E uses the core Rothermel model for fire propagation, however it can be configured for custom versions to support any empirical or semi empirical fire model. This has been done for different models employed in other countries, i.e., Chile, Canada, etc. In this regard, WFA-E platform is easily extended for use in unique geographies.
Crown Fire	Van Wagner (1977,1989,1993); Finney (1998); Scott and Reinhardt (2001)	Critical surface intensity and critical ROS for crown fire initialization. Expected ROS and flame intensity.
Time Evolution	Technosylva (Monedero, Ramirez 2011)	Fast-Marching method adapted to fire simulations. Minimum Travel Time algorithm with 32 degrees of freedom.
High-Definition Wind	Forthoffer et al (2009)	High resolution wind model obtained through the integration of the USFS WindNinja software. Note Technosylva is also the contractor for the USFS Missoula Fire Sciences Lab. for the on-going enhancement and customization of the WindNinja software. This provides Technosylva a unique understanding of the model science

Model	Model Reference	Notes
Surface fire	Rothermel 1972, Albini 1976 Kitral IntecChile	WFA-E uses the core Rothermel model for fire propagation, however it can be configured for custom versions to support any empirical or semi empirical fire model. This has been done for different models employed in other countries, i.e., Chile, Canada, etc. In this regard, WFA-E platform is easily extended for use in unique geographies.
		foundation and implementation approaches.
Wind Adjustment Factor	Andrews 2012	Wind speed conversion with height. Based on Albini and Baughman (1979); Baughman and Albini (1980); Rothermel (1983); Andrews (2012)
Fire Shape	Andrews 2018,	Unique ellipse based solely on the effective wind speed.
Live Moisture Content	Cardil et al.	Machine learning Algorithm based on historical NDVI weather reading
Dead Moisture Content	Nelson (2002)	SCE supplied based on fuel sampling
Spark Modelling	Technosylva	Ignition point displacement based on wind speed
Urban Encroachment	Technosylva 2016	Includes several variations of urban encroachment algorithms developed internally to facilitate spread of fires into non-burnable urban fuels. This incorporates a distance-based friction model. Based on research publications by NIST.
Spotting	Technosylva 2019	Surface spotting model for wind driven fires. Albini (1983a, 1983b); Chase (1984); Morris (1987)
Building Loss Factor⁷	Technosylva (Cardil xxx)	Machine Learning algorithm taking into account building conditions. Based on historical damage inspection data on buildings affected by fires over the past 13 years

Many of these models were originally published from research by the USFS Missoula Fire Sciences Laboratory. Technosylva has implemented, and enhanced these models, in addition to developing new models. Most Technosylva custom developed models are supported by journal publications as part of its corporate R&D program. Some of these models are referenced on the Technosylva web site at <https://technosylva.com/scientific-research/>.

- Beer, T. The interaction of wind and fire. Boundary-Layer Meteorology 54, 287–308 (1991). <https://doi.org/10.1007/BF00183958>
- Cruz Miguel G., Alexander Martin E. (2010) Assessing crown fire potential in coniferous forests of western North America: a critique of current approaches and recent simulation studies. International Journal of Wildland Fire 19, 377-398.
- Cruz, Miguel G.; Alexander, Martin E. (2013). Uncertainty associated with model predictions of surface and crown fire rates of spread. Environmental Modelling & Software. 47: 16-28.
- Scott, J.H. 2006. Comparison of crown fire modeling systems used in three fire management applications. USDA For. Serv. Res. Pap. RMRS-RP-58.
- Scott, J.H., and Reinhardt, E.D. 2001. Assessing crown fire potential by linking models of surface and crown fire behavior. USDA For. Serv. Res. Pap. RMRS-RP-29.
- Bennett, M., S.A. Fitzgerald, B. Parker, M. Main, A. Perleberg, C.C. Schnepf, and R.
- Mahoney. 2010. Reducing Fire Risk on Your Forest Property. PNW 618: 40 p.

⁷ SCE has not incorporated the Building Loss Factor methodology in the current version of its risk modeling.

- Fire Science Core Curriculum. 2017. OSU Extension Service, EM 9172: 197p.
- Gould, James. (1991). Validation of the Rothermel fire spread model and related fuel parameters in grassland fuels. *Proceedings of the Conference on Bushfire Modelling and Fire Danger Rating Systems*. 51-64.
- Di Gregorio, Salvatore & Bendicenti, E. (2003). Simulations of Forest Fires by the Cellular Automata Model.
- J. Glasa and L. Halada. On elliptical model for forest fire spread modeling and simulation. *Mathematics and Computers in Simulation*, 78(1):76–88, 2008.
- T. Ghisu, B. Arca, G. Pellizzaro, and P. Duce. A level-set algorithm for simulating wildfire spread. *CMES Computer Modeling in Engineering & Sciences*, 102(1):83–102, 2014.
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1), 269–271.
- Finney, M A, (2002). Fire growth using minimum travel time methods. *Canadian Journal of Forest Research*, 1420-1421, 32(8).
- Sanders, Kristen A., "Validation and calibration of the FARSITE fire area simulator for Yellowstone National Park" (2001). *Graduate Student Theses, Dissertations, & Professional Papers*. 3990. <https://scholarworks.umt.edu/etd/3990>
- A. Collin, D. Bernardin & O. Séro-Guillaume (2011) A Physical-Based Cellular Automaton Model for Forest-Fire Propagation, *Combustion Science and Technology*, 183:4, 347-369.
- Zigner, K.; Carvalho, L.M.V.; Peterson, S.; Fujioka, F.; Duine, G.-J.; Jones, C.; Roberts, D.; Moritz, M. Evaluating the Ability of FARSITE to Simulate Wildfires Influenced by Extreme, Downslope Winds in Santa Barbara, California. *Fire* **2020**, 3, 29. <https://doi.org/10.3390/fire3030029>
- Phillips, Ross J.; Waldrop, Thomas A.; Simon, Dean M. 2006. Assessment of the FARSITE model for predicting fire behavior in the Southern Appalachian Mountains. *Proceedings of the 13th biennial Southern Silvicultural Research Conference*. Gen. Tech. Rep. SRS-92. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station: 521-525.

Identify any additional limitations in the model based on the numerical techniques and implementation.

The Technosylva WFA-E platform is an integration of numerous specialty models designed to address specific scientific requirements and methods.

The following assumptions applied to the models used in WFA-E:

- The physical framework development is based on an idealized situation in steady state spread.
- Rate Of Spread at a point only depends on the conditions at that point (point-functional models). This means that there is no increase in speed due to non-local contributions of the fire front.
- Fire model is not directly coupled with the atmosphere. Fire will not modify local atmosphere. However, this is being addressed with seamless integration with the WRF-SFIRE model in development at San Jose State University, Wildfire Interdisciplinary Research Center. WRF-SFIRE is an option available to WFA-E customers to address specific convection-based fire scenarios.
- Fire is always assumed to be fully developed with fire acceleration, flashover, or decay not being considered.
- Atmospheric instability, which may have a deep impact on ROS (Beer 1991), is not considered in the model in any way.
- Gusts are not considered in the model.
- No interaction between slope and wind other than creating an effective or equivalent wind. This means that fire is assumed to have an elliptical shape no matter the alignment of wind and slope.
- Experimental data is scarce, and the empirical adjustment of models have been based on wind tunnel experiments and a few well documented fires.
- Fuel array description of the vegetation may not perfectly describe fuel characteristics.
- Spotting is only considered in surface fires.

External dependencies:

Describe external programs or software libraries used by the software.

This section provides a brief summary of the key input datasets (i.e., software libraries) required for wildfire behavior analysis and risk analysis. The following categories of input data are:

1. Landscape characteristics
2. Weather and atmospheric data
3. Fuel moisture
4. Values at risk (highly valued resources and assets)
5. Possible ignition sources
6. Fire activity

Describe data used by the software, including utility-collected and external sources. This should include the following:

- *Characteristics of the data (field definitions/schema, uncertainties, acquisition frequency).*
- *Scope and granularity (or resolution) of data in time and location (i.e., date range, spatial granularity for each data element).*
- *Sources of data, frequency of data updates, and verification of data quality. Explain in detail measurement approaches and procedures.*
- *Any processes used to modify the data (such as adjusting vegetative fuel models for wildfire spread based on prior history and vegetation growth).*

See expanded descriptions of the key input data (i.e., software libraries).

Landscape Characteristics

This includes a range of possible data that describe the characteristics of the landscape. The most important data are related to surface and canopy fuels, and vegetation. There are many publications available that describe these datasets, many from the USFS Missoula Fire Lab. Most use the Scott & Burgan 2005 Fuels Model Set standard for classification of fuels data.

Standard fire behavior analysis input layers are:

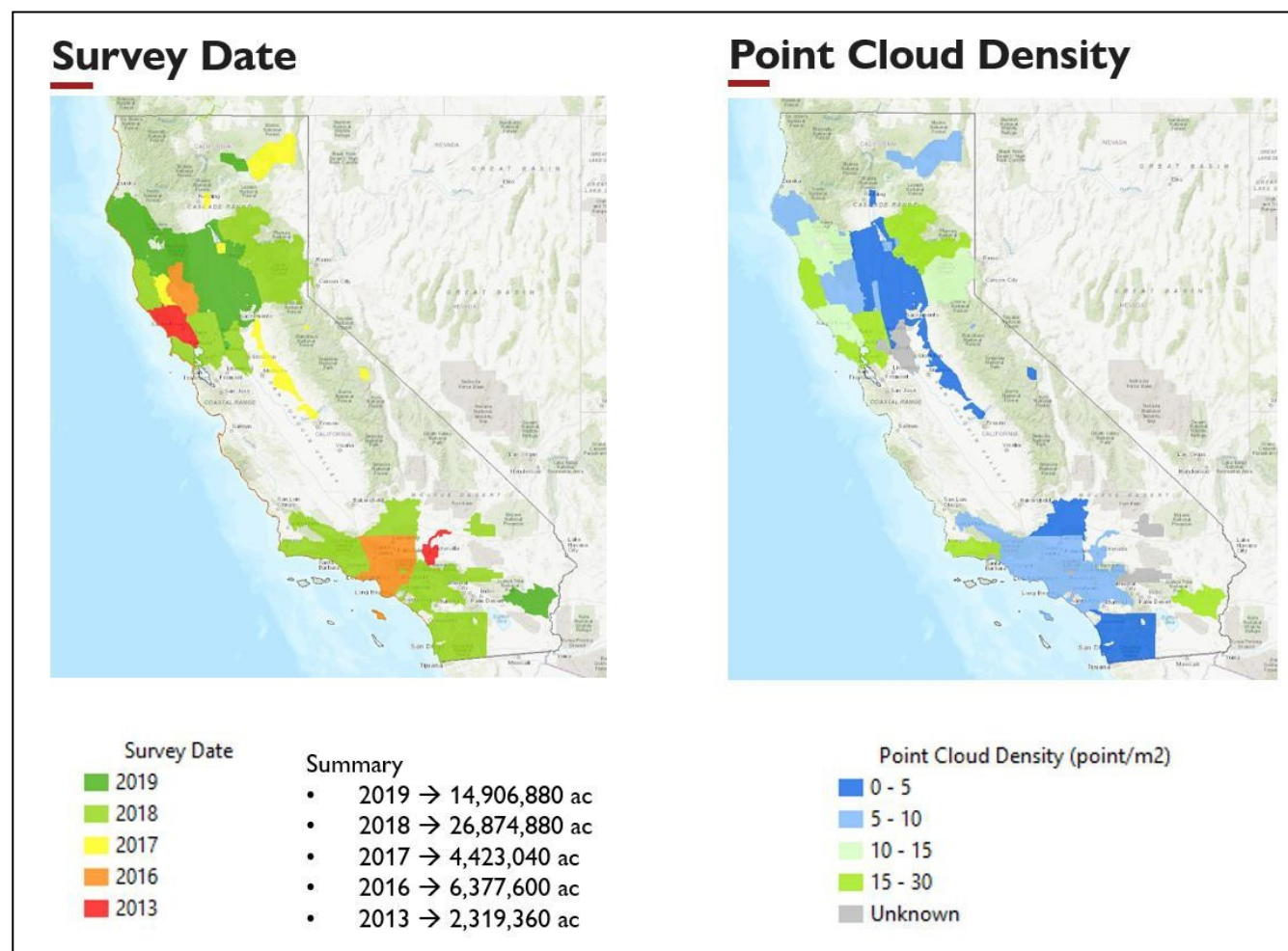
1. Terrain – elevation, slope, aspect
2. Surface fuels (Scott & Burgan 2005)
3. Canopy fuels
 - a. Canopy height
 - b. Canopy base height
 - c. Canopy bulk density
 - d. Canopy closure
4. WUI and Non-Forest Land Use classes (Technosylva, 2020)

Surface and Canopy Fuels

For these layers, data developed by Technosylva is used. Technosylva provides an annual fuel updating subscription where initial fuels is developed using advanced remote sensing object segmentation methods using high resolution imagery, available LiDAR & GEDI, and other standard imagery sources, as NAIP, Sentinel 2 and Landsat. This is supplemented with in-the-field surveys to verify the fuels for possible areas of concern and to validate the fuels classification. Surface and canopy fuels data is critical for accurate fire

behavior modeling, so it is paramount that this data is up-to-date, and when used, results in the observed and expected fire behavior.

Figure Appendix B-06 - LIDAR Data used for Technosylva Fuels 2021, with capture date and pointsdensity



Surface and canopy fuels are updated throughout the year, to accommodate changes to the fuels, typically monthly during fire season. This ensures that all major disturbances, such as fires, urban growth, landslides, etc. are updated in the fuels data. A variety of methods, including burn severity analysis, are used to update the fuels. Up to date fuels data is critical to ensuring the fire behavior outputs from our modeling are accurate, as it is a key input into risk analysis.

Technosylva continually tests new fuels datasets that become available from other sources, such as LANDFIRE, federal risk assessment regional projects, and independent sources, such as the California Forest Observatory data. Unfortunately, the publicly available data does not perform at the level required when confronted with operational testing. In general, these publicly available data do not result in fire behavior outputs that facilitated accurate predictions. Ultimately with any fuels dataset, the quality and accuracy of the fuels is measured on whether it produces 'observed and expected fire behavior'.

Technosylva is able to test this data, and other fuels data including their custom data, operationally on a daily basis with CAL FIRE and the IOUs against active wildfires to see how it performs.

Updates to the fuels, and algorithms that use the fuels data for fire behavior modeling is on-going, as Technosylva continues to enhance the data and algorithms to match observed fire behavior across the state. These methods and algorithms are proprietary.

WUI and Non-Forest Fuels Land Use classes are based on a Technosylva proprietary method that characterizes WUI and other land uses classes that have been a typical limitation of the Scott and Burgan classification, as they are defined in general non-burnable classes. In combination with the Surface Fuels, this provides a solid foundation for fire behavior and impact analysis.

The following two figures present an example of publicly available LANDFIRE data commonly used for fire modeling, and the custom Technosylva fuels used.

Figure Appendix B-07 - LandFire Fuels - Non-Burnable Classes

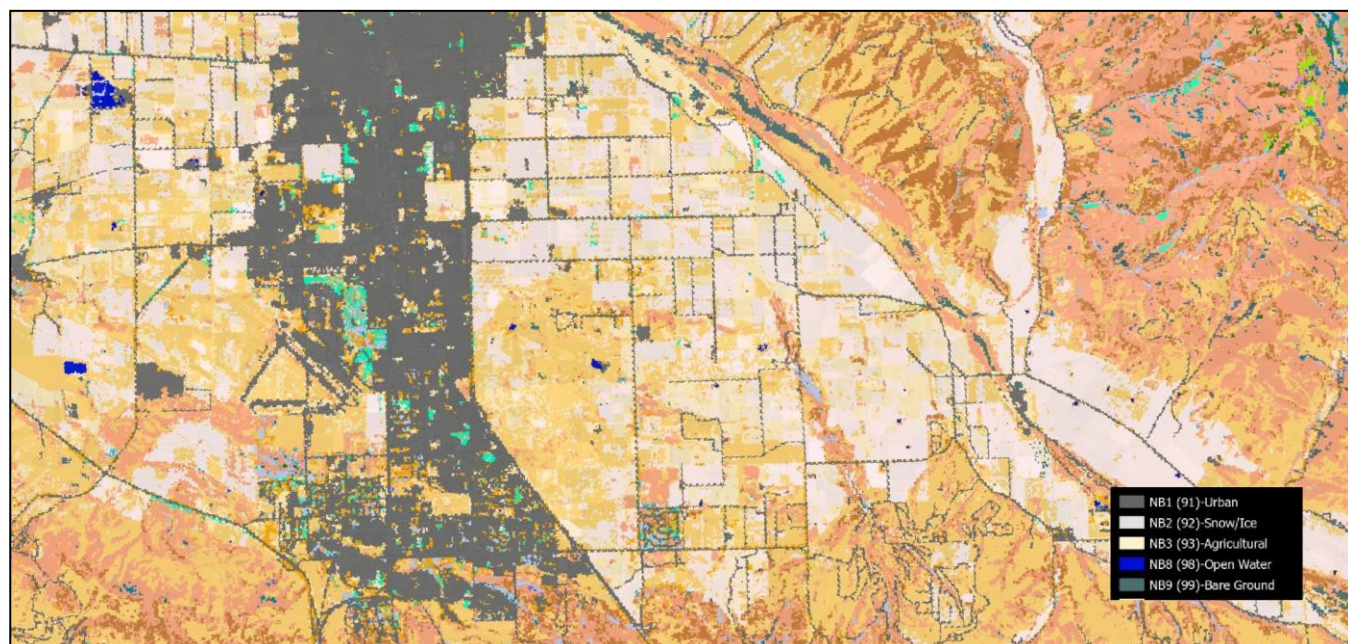
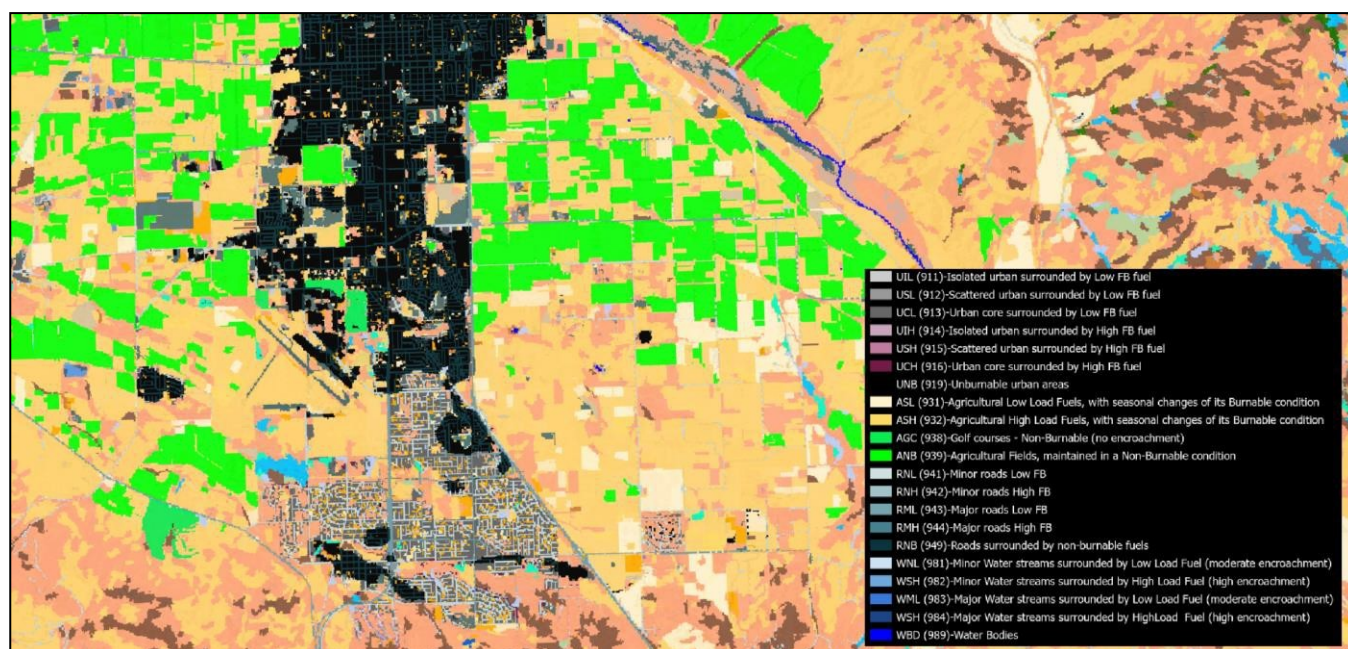


Figure Appendix B-08 - Technosylva Fuels Dec 2021 – WUI and Non-Forest Fuels Classes



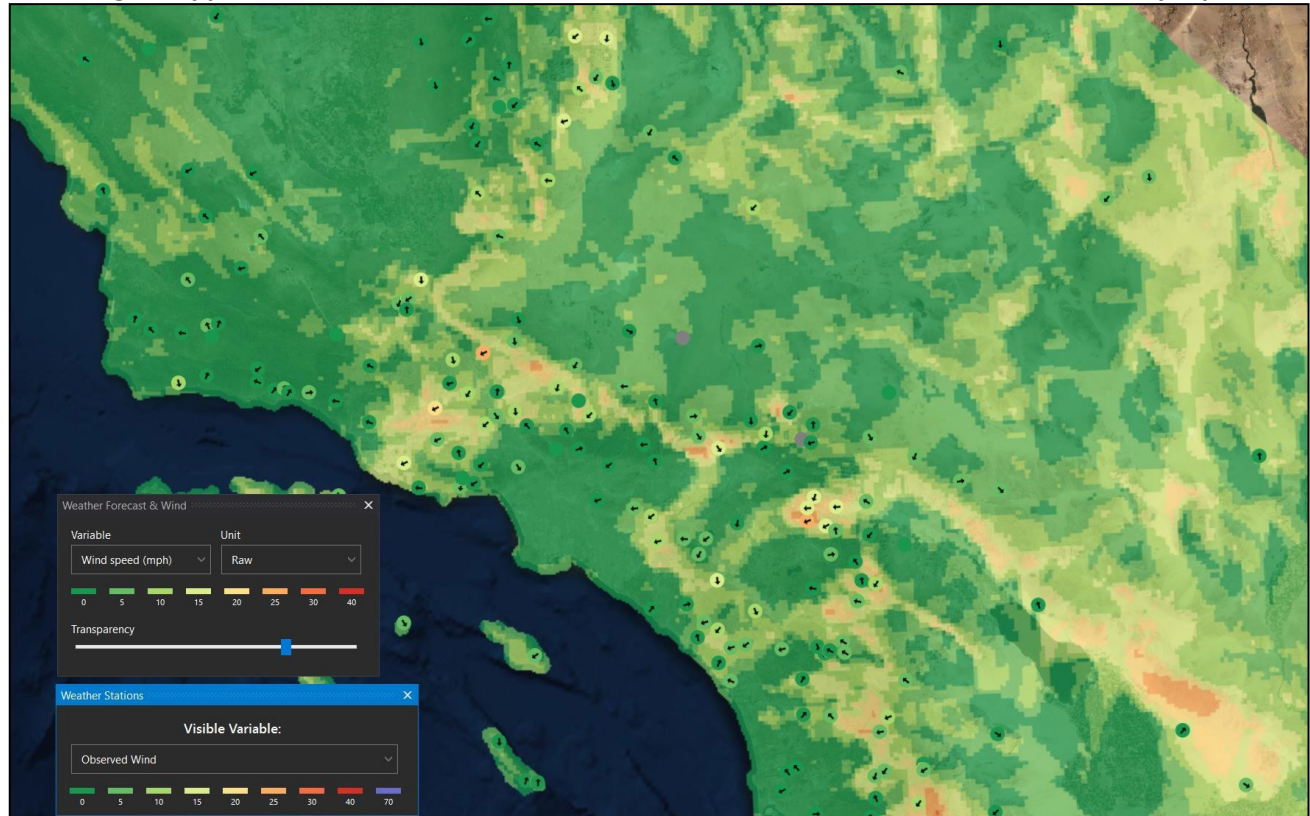
Weather and Atmospheric Data

WRF data is developed using third party weather and predictive services experts available through commercial providers. Data is at a 2 km spatial resolution and hourly (temporal) for a multi-day period, up to seven days. Multiple forecasts are generated daily. Weather observation data can also be used along

with, or independently, to support fire behavior analysis. This data is typically available through published weather stations on MesoWest, or through commercial providers, such as Synoptic. The methods of how this data can be integrated within the Technosylva software and processes is proprietary.

The following figure shows a typical 2km WRF model of wind speed overlaid with weather stations data (WFA-E software example).

Figure Appendix B-09 - Predicted (WRF model) and Observed Wind (Weather Stations, Synoptic)



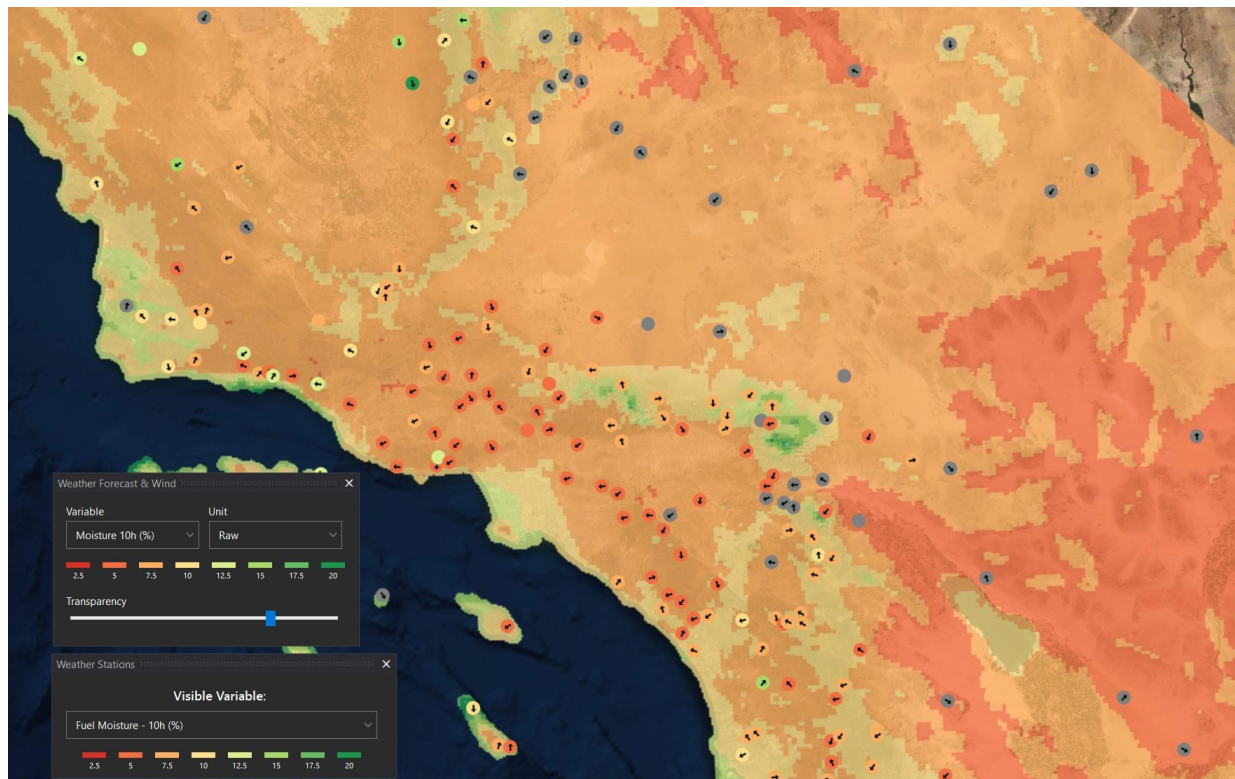
Fuel Moisture

Fuel moisture data is also a key input into fire behavior modeling. Fuel moisture can be characterized as either Dead or Live fuel moisture. Standard methods for measuring and quantifying fuel moistures are well documented in publications by the USFS Missoula Fire Lab and other research agencies.

However, to date the ability to accurately predict live and dead fuel moistures at high resolution has been limited. Only a few IOUs and commercial vendors are producing daily estimates that can be integrated into fire modeling. Technosylva produces both a dead and live fuel moisture data product that combines historical and current sample data with remotely sensing imagery in a machine learning model to estimate daily data products. These methods are proprietary although they are substantiated with several publications and on-going collaboration between the IOUs, Technosylva and fire weather and behavior research agencies. This fuel moisture data product is used by CAL FIRE and several IOUs across seven western US states.

The following figure shows the Technosylva Dead Fuel Moisture overlaid with weather stations data (WFA-E software example).

Figure Appendix B-10 - Technosylva Dead Fuel Moisture overlaid with weather stations



Predicted (WRF model) and Observed 10-hr Fuel Moisture (Weather Stations, Synoptic)

Possible Ignition Sources

Wildfire ignition data varies greatly depending on the organization and purpose of the wildfire risk analysis. Traditionally, agency driven risk assessments will use historical fire location data to create Historical Fire Occurrence datasets, reflecting ignition density over a specific time period. This data is obtained from federal and state fire reporting systems.

Risk can be assessed related to the probability of ignition for electric utility assets, or more commonly with the potential spread and impacts of a wildfire ignited by an asset. Technosylva provides integration of both ignition and spread analysis to derive risk metrics using VAR data. This focuses on assigning possible consequence back to the electric utility assets to identify those assets more prone to having significant impacts should a wildfire ignite. Different proprietary methods exist to integrate and model probability of ignition data for electric utility assets with consequence modeling. Referred to as “asset wildfire risk” this information can be used to support operational decisions, such as PSPS, resource allocation and placement, and stakeholder communication, in addition to short- and long-term mitigation planning efforts, reflected in IOU WMPs. The weather and fuels inputs will vary depending on the purpose of these risk analyses.

Model substantiation:

Identify existing data that can be used to validate model performance.

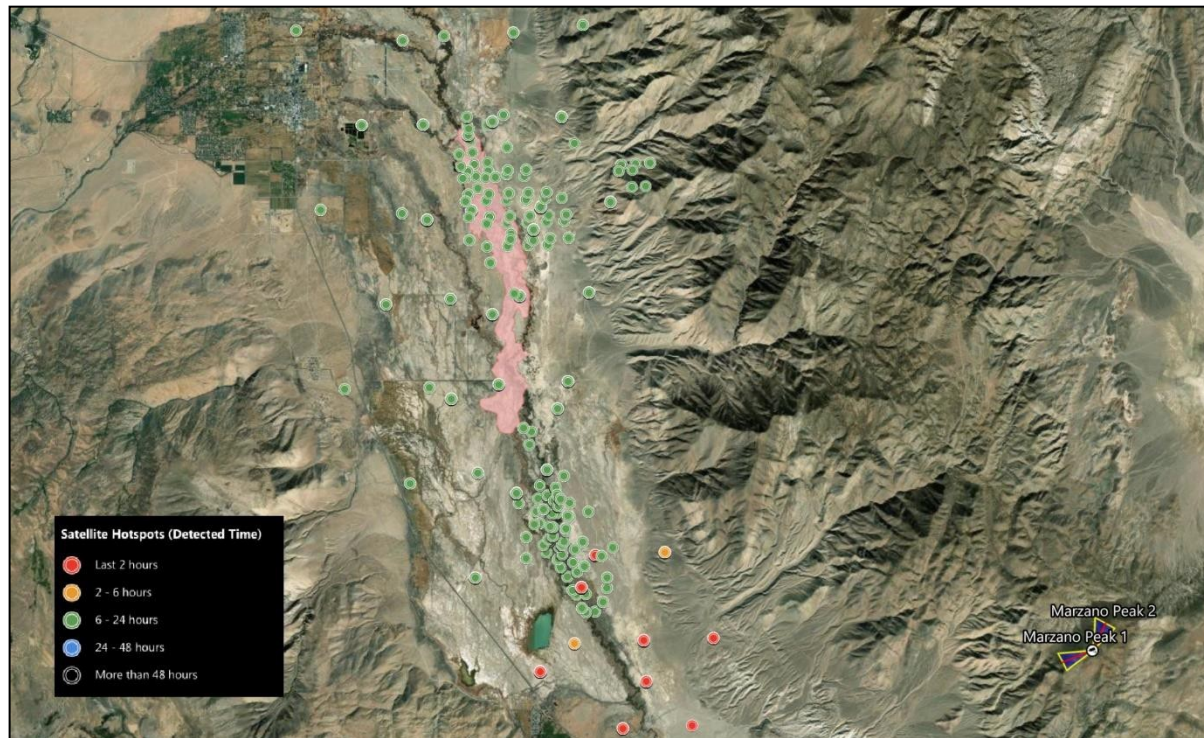
The fire activity data used validate model performance is captured from different sources:

- VIIRS and MODIS Satellite hotspots, from public sources (FIRMS)
- GOES 16 and 17 data based on agreement with providers to the IOUs
- Lighting data also from IOU’s providers

- Fire Perimeters from Open Wildfire data from NIFC
- Fire activity from National Guard data from Fire Guard program
- Alert Wildfire Cameras integration

The following figure shows an example of Fire Activity data integrated into the Technosylva WFA-E system. All data is temporal and displayed color coded based on a selected time from the software timeline.

Figure Appendix B-11 - Hotspots, Fire Perimeters and Alert Wildfire Cameras



All models need to be verified and validated for the specific application in which they are to be used in accordance with the guidance provided in Section “Model Substantiation,” below.

See verification and validation of individual models as referenced on the Technosylva web site at <https://technosylva.com/scientific-research/>.

Sensitivity

Describe the efforts to evaluate the impact of model and input parameter uncertainty on the model predicted outcomes.

Many of these models were originally published from research by the USFS Missoula Fire Sciences Laboratory. Technosylva has implemented, and enhanced these models, in addition to developing new models. Most Technosylva custom developed models are supported by journal publications as part of its corporate R&D program. See sensitivity of individual models as referenced on the Technosylva web site at <https://technosylva.com/scientific-research/>.

- Beer, T. The interaction of wind and fire. Boundary-Layer Meteorology 54, 287–308 (1991). <https://doi.org/10.1007/BF00183958>

- Cruz Miguel G., Alexander Martin E. (2010) Assessing crown fire potential in coniferous forests of western North America: a critique of current approaches and recent simulation studies. *International Journal of Wildland Fire* 19, 377-398.
- Cruz, Miguel G.; Alexander, Martin E. (2013). Uncertainty associated with model predictions of surface and crown fire rates of spread. *Environmental Modelling & Software*. 47: 16-28.
- Scott, J.H. 2006. Comparison of crown fire modeling systems used in three fire management applications. USDA For. Serv. Res. Pap. RMRS-RP-58.
- Scott, J.H., and Reinhardt, E.D. 2001. Assessing crown fire potential by linking models of surface and crown fire behavior. USDA For. Serv. Res. Pap. RMRS-RP-29.
- Bennett, M., S.A. Fitzgerald, B. Parker, M. Main, A. Perleberg, C.C. Schnepf, and R.
- Mahoney. 2010. Reducing Fire Risk on Your Forest Property. PNW 618: 40 p.
- Fire Science Core Curriculum. 2017. OSU Extension Service, EM 9172: 197p.
- Gould, James. (1991). Validation of the Rothermel fire spread model and related fuel parameters in grassland fuels. *Proceedings of the Conference on Bushfire Modelling and Fire Danger Rating Systems*. 51-64.
- Di Gregorio, Salvatore & Bendicenti, E. (2003). Simulations of Forest Fires by the Cellular Automata Model.
- J. Glasa and L. Halada. On elliptical model for forest fire spread modeling and simulation. *Mathematics and Computers in Simulation*, 78(1):76–88, 2008.
- T. Ghisu, B. Arca, G. Pellizzaro, and P. Duce. A level-set algorithm for simulating wildfire spread. *CMES Computer Modeling in Engineering & Sciences*, 102(1):83–102, 2014.
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1), 269–271.
- Finney, M A, (2002). Fire growth using minimum travel time methods. *Canadian Journal of Forest Research*, 1420-1421, 32(8).
- Sanders, Kristen A., "Validation and calibration of the FARSITE fire area simulator for Yellowstone National Park" (2001). *Graduate Student Theses, Dissertations, & Professional Papers*. 3990. <https://scholarworks.umt.edu/etd/3990>
- A. Collin, D. Bernardin & O. Séro-Guillaume (2011) A Physical-Based Cellular Automaton Model for Forest-Fire Propagation, *Combustion Science and Technology*, 183:4, 347-369,
- Zigner, K.; Carvalho, L.M.V.; Peterson, S.; Fujioka, F.; Duine, G.-J.; Jones, C.; Roberts, D.; Moritz, M. Evaluating the Ability of FARSITE to Simulate Wildfires Influenced by Extreme, Downslope Winds in Santa Barbara, California. *Fire* 2020, 3, 29. <https://doi.org/10.3390/fire3030029>
- Phillips, Ross J.; Waldrop, Thomas A.; Simon, Dean M. 2006. Assessment of the FARSITE model for predicting fire behavior in the Southern Appalachian Mountains. *Proceedings of the 13th biennial Southern Silvicultural Research Conference*. Gen. Tech. Rep. SRS-92. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station: 521-525.

Describe the efforts to evaluate the propagation of uncertainty into downstream models.

SCE utilizes a single deterministic forecast.

Many of these models were originally published from research by the USFS Missoula Fire Sciences Laboratory. Technosylva has implemented, and enhanced these models, in addition to developing new models. Most Technosylva custom developed models are supported by journal publications as part of its corporate R&D program. See sensitivity of individual models as referenced on the Technosylva web site at <https://technosylva.com/scientific-research/>.

- Beer, T. The interaction of wind and fire. *Boundary-Layer Meteorology* 54, 287–308 (1991). <https://doi.org/10.1007/BF00183958>
- Cruz Miguel G., Alexander Martin E. (2010) Assessing crown fire potential in coniferous forests of western North America: a critique of current approaches and recent simulation studies.

International Journal of Wildland Fire 19, 377-398.

- Cruz, Miguel G.; Alexander, Martin E. (2013). Uncertainty associated with model predictions of surface and crown fire rates of spread. *Environmental Modelling & Software*. 47: 16-28.
- Scott, J.H. 2006. Comparison of crown fire modeling systems used in three fire management applications. USDA For. Serv. Res. Pap. RMRS-RP-58.
- Scott, J.H., and Reinhardt, E.D. 2001. Assessing crown fire potential by linking models of surface and crown fire behavior. USDA For. Serv. Res. Pap. RMRS-RP-29.
- Bennett, M., S.A. Fitzgerald, B. Parker, M. Main, A. Perleberg, C.C. Schnepf, and R. Mahoney. 2010. Reducing Fire Risk on Your Forest Property. PNW 618: 40 p.
- Fire Science Core Curriculum. 2017. OSU Extension Service, EM 9172: 197p
- Gould, James. (1991). Validation of the Rothermel fire spread model and related fuel parameters in grassland fuels. *Proceedings of the Conference on Bushfire Modelling and Fire Danger Rating Systems*. 51-64.
- Di Gregorio, Salvatore & Bendicenti, E. (2003). Simulations of Forest Fires by the Cellular Automata Model.
- J. Glasa and L. Halada. On elliptical model for forest fire spread modeling and simulation. *Mathematics and Computers in Simulation*, 78(1):76–88, 2008.
- T. Ghisu, B. Arca, G. Pellizzaro, and P. Duce. A level-set algorithm for simulating wildfire spread. *CMES Computer Modeling in Engineering & Sciences*, 102(1):83–102, 2014
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1), 269–271.
- Finney, M A, (2002). Fire growth using minimum travel time methods. *Canadian Journal of Forest Research*, 1420-1421, 32(8)
- Sanders, Kristen A., "Validation and calibration of the FARSITE fire area simulator for Yellowstone National Park" (2001). *Graduate Student Theses, Dissertations, & Professional Papers*. 3990. <https://scholarworks.umt.edu/etd/3990>
- A. Collin, D. Bernardin & O. Séro-Guillaume (2011) A Physical-Based Cellular Automaton Model for Forest-Fire Propagation, *Combustion Science and Technology*, 183:4, 347-369,
- Zigner, K.; Carvalho, L.M.V.; Peterson, S.; Fujioka, F.; Duine, G.-J.; Jones, C.; Roberts, D.; Moritz, M. Evaluating the Ability of FARSITE to Simulate Wildfires Influenced by Extreme, Downslope Winds in Santa Barbara, California. *Fire* **2020**, 3, 29. <https://doi.org/10.3390/fire3030029>
- Phillips, Ross J.; Waldrop, Thomas A.; Simon, Dean M. 2006. Assessment of the FARSITE model for predicting fire behavior in the Southern Appalachian Mountains. *Proceedings of the 13th biennial Southern Silvicultural Research Conference*. Gen. Tech. Rep. SRS-92. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station: 521-525

Model Substantiation

Model substantiation is the process used to ensure that a model is correct and suitable to an application. The following relevant terms are defined in Appendix A “Definitions:”

- **Calibration**
- **Model uncertainty**
- **Parameter uncertainty**
- **Sensitivity**
- **Uncertainty**
- **Validation**
- **Verification**

For each model, the electrical corporation must be readily able to provide, if requested by Energy Safety or designated stakeholders, documentation of the following model substantiation studies:

- **Validation data** – Identify existing data that can be used to validate model performance.
- **Model verification** – Describe efforts to verify that the model is working as designed and that the equations are being properly solved. Verification is often conducted through independent review of source code and use of unit and integration test suites by the software developer. If the end user of a model is not the same as the model developers, the SFPE guidance includes an additional step on user training and certification to the verification process. The verification study of each model must include each of the following:
 - Verification of the basic functionality of the model through simple test cases.
 - Verification of consistency of input parameters. For example, wind speed varies substantially as a function of height and space. Individual wildfire models may assume wind speed is specified at a fixed height (such as 20 feet, 32 feet, or mid-flame height). Specifying the wind speed at the wrong height may result in incorrect model predictions.
 - Independent review, which may consist of one of the following:
 - Independent third-party review of software implementation and data integration where the third-party is neither an employee nor a subcontractor of the electrical corporation or software supplier.
 - Software verification suite, including software source code and automated verification code, provided by the electrical corporation to Energy Safety. See the Fire Dynamics Suite (FDS) developed by NIST for an example.⁸
- **Model validation** – Models are validated by comparing model predictions to observations

⁸ Fire Dynamics Simulator, FDS Verification Process - <https://github.com/firemodels/fds/wiki/FDS-Verification-Process>.

from historic events or experiments. It is important to note that validation does not mean that a model's predictions are perfect. Rather, the predictions are good enough for the intended use case. The validation study and uncertainty assessment of each model must do each of the following:

- Document the efforts undertaken by the electrical corporation to quantify the uncertainty in the model when input parameters are known (i.e., open calculation). This should include a discussion of relevant experiments/datasets used to benchmark performance as well as a statistical summary of performance. See the FDS validation suite developed by NIST.⁹*
- Document the efforts undertaken by the electrical corporation to quantify the variability in input parameters in practice. This should include a discussion of the input data currently used in the model, the process used to update these data, the sensitivity of model predictions to this variability, and the degree to which this variability is within the validation range presented for the software model.*
- Document the type of model validation based on the characterizations defined in ASTM E 1355 (i.e., blind calculation, specified calculation, open calculation).*
- Open calculations consist of modeling efforts where the expected model output and input parameters are based on post-event knowledge. This is a reasonable approach for risk assessment where there is time to gather and process these data. However, the accuracy of a model in open calculation may not directly translate to accuracy in other calculation classes.*
- The predictive power of the model to generate forecasts of ongoing events is best captured through blind validation due to the impact of uncertainties in model inputs. For example, in forecasting the spread of a wildfire, there is high uncertainty in vegetation and weather conditions. The focus of blind validation is to understand how accurate the forecasts are when the inputs include uncertainty.*
- Model calibration** – *Calibration in the context of wildfire risk assessment is focused on modifying model inputs and model parameters to achieve better agreement for a specific scenario. Calibration is an important process to develop validation scenarios as well as to support real-time decision making. In general, calibration approaches limit the propagation of error by correcting to new data but have limited effectiveness in improving the quality of the forecast. However, calibrating the model to each individual scenario does not provide confidence in the predictive capability of the model for new scenarios. For each model that uses real-time calibration, the following must be documented:*
 - Data sources used in calibrating the model*
 - Model parameters that are modified during calibration and the process used to modify parameters*
 - Uncertainty as a function of lead time (i.e., forecast time) with and without*

⁹ Fire Dynamics Simulator, FDS Validation Process - <https://github.com/firemodels/fds/wiki/FDS-Validation-Process>.

calibration

- *The degree to which a model predicted value might differ from the true value, including systematic bias and statistical variance (i.e., model uncertainty assessment). This should be presented in an open calculation.*
- *Uncertainty as a function of lead time (i.e., forecast time) with and without calibration*
- *The degree to which a model predicted value might differ from the true value, including systematic bias and statistical variance (i.e., model uncertainty assessment). This should be presented in an open calculation.*

POI

SCE has prepared sub-model documentation that is provided in Attachments A through D of this document.

Below SCE has responded to each Energy Safety requirement or indicated where it is discussed in the sub-model documentation. In most cases SCE's responses are provided at the level of the sub-model, with the intention that this provides a more granular level of information. The overall POI model is discussed in the main body of the WMP (see Section 6).

Validation data

Identify existing data that can be used to validate model performance.

Existing data that is used to validate model performance SCE's sub-models is discussed in Section 3.2 of the model documentation.

Model verification

Describe efforts to verify that the model is working as designed and that the equations are being properly solved. Verification is often conducted through independent review of source code and use of unit and integration test suites by the software developer. If the end user of a model is not the same as the model developers, the SFPE guidance includes an additional step on user training and certification to the verification process. The verification study of each model must include each of the following:

- *Verification of the basic functionality of the model through simple test cases.*
- *Verification of consistency of input parameters. For example, wind speed varies substantially as a function of height and space. Individual wildfire models may assume wind speed is specified at a fixed height (such as 20 feet, 32 feet, or mid-flame height). Specifying the wind speed at the wrong height may result in incorrect model predictions.*
- *Independent review, which may consist of one of the following:*
 - *Independent third-party review of software implementation and data integration where the third-party is neither an employee nor a subcontractor of the electrical corporation or software supplier.*
 - *Software verification suite, including software source code and automated verification code, provided by the electrical corporation to Energy Safety. See the Fire Dynamics*

SCE discusses each of the following requested model verification information in the following:

- Verification of the basic functionality of the model through simple test cases is discussed in Section 3.1 and Section 3.2.
- Verification of consistency of input parameters is discussed in Section 2.1.
- Please see Section 6.7.1 of SCE's 2023 WMP for a discussion of SCE's plans for increased independent review.

Model validation

Models are validated by comparing model predictions to observations from historic events or experiments. It is important to note that validation does not mean that a model's predictions are perfect. Rather, the predictions are good enough for the intended use case. The validation study and uncertainty assessment of each model must do each of the following:

- *Document the efforts undertaken by the electrical corporation to quantify the uncertainty in the model when input parameters are known (i.e., open calculation). This should include a discussion of relevant experiments/datasets used to benchmark performance as well as a statistical summary of performance. See the FDS validation suite developed by NIST.¹¹*
- *Document the efforts undertaken by the electrical corporation to quantify the variability in input parameters in practice. This should include a discussion of the input data currently used in the model, the process used to update these data, the sensitivity of model predictions to this variability, and the degree to which this variability is within the validation range presented for the software model.*
- *Document the type of model validation based on the characterizations defined in ASTM E 1355 (i.e., blind calculation, specified calculation, open calculation).*
- *Open calculations consist of modeling efforts where the expected model output and input parameters are based on post-event knowledge. This is a reasonable approach for risk assessment where there is time to gather and process these data. However, the accuracy of a model in open calculation may not directly translate to accuracy in other calculation classes.*
- *The predictive power of the model to generate forecasts of ongoing events is best captured through blind validation due to the impact of uncertainties in model inputs. For example, in forecasting the spread of a wildfire, there is high uncertainty in vegetation and weather conditions. The focus of blind validation is to understand how accurate the forecasts are when the inputs include uncertainty.*

SCE discusses each of the following requested model validation information in the following:

- SCE discusses how it quantifies the uncertainty of its model in Section 1.7 of the model documentation.
- SCE discusses the input parameters of its model in Section 2.1 of the model documentation.
- SCE discusses the model validation methodology in Section 2.2 of the model documentation.

¹⁰ Fire Dynamics Simulator, FDS Verification Process - <https://github.com/firemodels/fds/wiki/FDS-Verification-Process>.

¹¹ Fire Dynamics Simulator, FDS Validation Process - <https://github.com/firemodels/fds/wiki/FDS-Validation-Process>.

- SCE discusses modeling efforts where the expected model output and input parameters are based on post-event knowledge is discussed in Section 3.3 of the model documentation.
- SCE discusses the predictive power of the model in Section 3.2 of the model documentation.

Model calibration

Calibration in the context of wildfire risk assessment is focused on modifying model inputs and model parameters to achieve better agreement for a specific scenario. Calibration is an important process to develop validation scenarios as well as to support real-time decision making. In general, calibration approaches limit the propagation of error by correcting to new data but have limited effectiveness in improving the quality of the forecast. However, calibrating the model to each individual scenario does not provide confidence in the predictive capability of the model for new scenarios. For each model that uses real-time calibration, the following must be documented:

- *Data sources used in calibrating the model*
- *Model parameters that are modified during calibration and the process used to modify parameters*
- *Uncertainty as a function of lead time (i.e., forecast time) with and without calibration*
- *The degree to which a model predicted value might differ from the true value, including systematic bias and statistical variance (i.e., model uncertainty assessment). This should be presented in an open calculation.*

SCE discusses each of the following requested model validation information in the following:

- SCE discusses the data sources used in calibrating its POI model in Section 2.6 of the model documentation.
- SCE discusses model parameters that are modified during calibration and the process used to modify parameters in Section 2.6.
- SCE does not account for uncertainty as a function of lead time as part of the model calibration.
- SCE discusses the degree to which a model predicted value might differ is discussed in Section 1.7.

Wildfire Consequence

Validation data

Identify existing data that can be used to validate model performance.

Please see verification and validation of individual models as referenced on the Technosylva web site at <https://technosylva.com/scientific-research/>.

Model verification

Describe efforts to verify that the model is working as designed and that the equations are being properly solved. Verification is often conducted through independent review of source code and use of unit and integration test suites by the software developer. If the end user of a model is not the same as the model developers, the SFPE guidance includes an additional step on user training and certification to the verification process. The verification study of each model must include each of the following:

- *Verification of the basic functionality of the model through simple test cases.*
- *Verification of consistency of input parameters. For example, wind speed varies substantially as a function of height and space. Individual wildfire models may assume wind speed is specified at a fixed height (such as 20 feet, 32 feet, or mid-flame height). Specifying the wind speed at the wrong height may result in incorrect model predictions.*
- *Independent review, which may consist of one of the following:*
 - *Independent third-party review of software implementation and data integration where the third-party is neither an employee nor a subcontractor of the electrical corporation or software supplier.*
 - *Software verification suite, including software source code and automated verification code, provided by the electrical corporation to Energy Safety. See the Fire Dynamics Suite (FDS) developed by NIST for an example.¹²*

See verification of individual models as referenced on the Technosylva web site at <https://technosylva.com/scientific-research/>.

Please see Section 6.7.1 of SCE's 2023 WMP for a discussion of SCE's plans for increased independent review.

Model validation

Models are validated by comparing model predictions to observations from historic events or experiments. It is important to note that validation does not mean that a model's predictions are perfect. Rather, the predictions are good enough for the intended use case. The validation study and uncertainty assessment of each model must do each of the following:

- *Document the efforts undertaken by the electrical corporation to quantify the uncertainty in the model when input parameters are known (i.e., open calculation). This should include a discussion of relevant experiments/datasets used to benchmark performance as well as a statistical summary of performance. See the FDS validation suite developed by NIST.¹³*
- *Document the efforts undertaken by the electrical corporation to quantify the variability in input parameters in practice. This should include a discussion of the input data currently used*

¹² Fire Dynamics Simulator, FDS Verification Process - <https://github.com/firemodels/fds/wiki/FDS-Verification-Process>.

¹³ Fire Dynamics Simulator, FDS Validation Process - <https://github.com/firemodels/fds/wiki/FDS-Validation-Process>.

in the model, the process used to update these data, the sensitivity of model predictions to this variability, and the degree to which this variability is within the validation range presented for the software model.

- *Document the type of model validation based on the characterizations defined in ASTM E 1355 (i.e., blind calculation, specified calculation, open calculation).*
- *Open calculations consist of modeling efforts where the expected model output and input parameters are based on post-event knowledge. This is a reasonable approach for risk assessment where there is time to gather and process these data. However, the accuracy of a model in open calculation may not directly translate to accuracy in other calculation classes.*
- *The predictive power of the model to generate forecasts of ongoing events is best captured through blind validation due to the impact of uncertainties in model inputs. For example, in forecasting the spread of a wildfire, there is high uncertainty in vegetation and weather conditions. The focus of blind validation is to understand how accurate the forecasts are when the inputs include uncertainty.*

See validation of individual models as referenced on the Technosylva web site at

<https://technosylva.com/scientific-research/>.

Model calibration

Calibration in the context of wildfire risk assessment is focused on modifying model inputs and model parameters to achieve better agreement for a specific scenario. Calibration is an important process to develop validation scenarios as well as to support real-time decision making. In general, calibration approaches limit the propagation of error by correcting to new data but have limited effectiveness in improving the quality of the forecast. However, calibrating the model to each individual scenario does not provide confidence in the predictive capability of the model for new scenarios. For each model that uses real-time calibration, the following must be documented:

- *Data sources used in calibrating the model*
- *Model parameters that are modified during calibration and the process used to modify parameters*

- *Uncertainty as a function of lead time (i.e., forecast time) with and without calibration*
- *The degree to which a model predicted value might differ from the true value, including systematic bias and statistical variance (i.e., model uncertainty assessment). This should be presented in an open calculation.*

See calibration of individual models as referenced on the Technosylva web site at <https://technosylva.com/scientific-research/>.

Additional Models Supporting Risk Calculation

The electrical corporation must be able to provide, if requested by Energy Safety or designated stakeholders, the following information regarding additional models that support risk calculation. The electrical corporation does not need to provide this information as part of its WMP submission.

Weather Analysis

The electrical corporation must evaluate weather history within its service territory to determine realistic design scenarios. Energy Safety considers the following to be key elements in the calculation of the weather history:

- *Inclusion of at least the following **model outputs**:*
 - *Air temperature*
 - *Barometric pressure*
 - *Fuel moisture*
 - *Relative humidity*
 - *Wind velocity (speed and direction)*
- *Evaluation of the **sensitivity** of downstream models to uncertainty in weather modeling.*
- *Use of **separate modules** for local weather analysis and local vegetation analysis.*
- *Use of **spatial granularity** of forecasts that at a minimum include:*
 - *Horizontal resolution ≤ 4 km*
 - *Vertical resolution sufficient to evaluate average conditions at environmental monitoring system locations*
- *Use of at least a 30-year **time horizon** of the weather analysis throughout the service territory.*
- *Calculation of the **uncertainty** of the input parameters and model assumptions, limitations, and parameterizations on the model results.*

Fuel Conditions

The electrical corporation must describe how it monitors and accounts for the contribution of fuel conditions to ignition risk in its decision-making processes. The electrical corporation must track, calculate, and report the following:

- *Measurement and calculation methods used for assessing fuel conditions (e.g., live and dead fuel moisture, fuel density)*
- *Methodology used for projecting future fuel conditions*
- *Calculation of any proprietary fuel condition indices (or other measures tracked)*
- *Thresholds used to identify extreme fuel conditions, including any factors used to modify thresholds (e.g., fuel type, topography)*
- *Geospatial polygons of extreme fuel conditions within the service territory as defined in the geospatial schema (GIS Data Reporting Standard, current version)*
- *Geospatial statistical frequency of extreme fuel conditions over the last five years throughout the service territory*

Weather Analysis

Model Outputs

*Inclusion of at least the following **model outputs**: Air temperature, Barometric pressure, Fuel moisture, Relative humidity, and Wind velocity (speed and direction)*

SCE developed a list of extreme weather days which were used to calculate wildfire risk from a historic perspective. These weather events considered air temperature, barometric pressure, relative humidity, and wind speed. Fuel moisture was assessed separately but served as an additional layer when assessing wildfire risk.

Sensitivity

Evaluation of the sensitivity of downstream models to uncertainty in weather modeling.

Wildfire spread modeling is contingent on a number of upstream model inputs that can exacerbate errors during the initialization phase which affects output projections, however, specific sensitivities of these model inputs, while unknown, are not believed to be significant at this time.

Separate Modules

Use of separate modules for local weather analysis and local vegetation analysis.

SCE's fuel models run in a post processed environment and are separate from the meteorological modeling that takes place. Estimated conditions of meteorological and fuel moisture outputs are not based upon local conditions but are based on model forecasts.

Spatial Granularity

Use of spatial granularity of forecasts that at a minimum include:

- *Horizontal resolution ≤ 4 km*

Models are run at a 2-kilometer horizontal resolution.

- *Vertical resolution sufficient to evaluate average conditions at environmental monitoring system locations*

SCE's in-house modeling has multiple vertical levels such as 2m, 10m, etc. that account for conditions at environmental monitoring system locations.

Time Horizon

Use of at least a 30-year time horizon of the weather analysis throughout the service territory.

SCE has a 40+ years of historical weather and fuels data at a 2-kilometer horizontal resolution covering a domain that includes the entire SCE service territory.

Uncertainty

Calculation of the uncertainty of the input parameters and model assumptions, limitations, and parameterizations on the model results.

SCE has worked with its weather modeling vendor to assess the uncertainty in key model input

parameters and assumptions on its weather model results. These include evaluating the impacts weather model parameterization choice and initial condition source on weather forecast accuracy. Based on these efforts, SCE's weather modeling vendor has configured the 2-km deterministic forecast to use a configuration that minimized forecast error and maximized correlation with observations for wind. Additionally, SCE's weather analysis includes use of an ensemble approach designed to sample these key uncertainties provide a range of outcomes for subject matter experts to assess.

Fuel Conditions

Measurement and calculation methods used for assessing fuel conditions (e.g., live and dead fuel moisture, fuel density)

SCE measures the moisture content within the living vegetation bi-monthly by physically going to specific locations and collecting samples of the vegetation. These samples are then sent to a lab in which they undergo a gravimetric process to determine the exact amount of moisture in the vegetation sample. This information is also used to develop and train machine learning models to estimate live fuel moisture daily on a 1-kilometer by 1-kilometer grid. In addition, SCE estimates the moisture content within the dead vegetation through a series of mathematical algorithms that have been established by the US Forest Service decades ago and are used extensively by fire agencies countrywide. All these estimates of fuel moisture allow SCE to more accurately assess vegetation conditions related to ignition risk.

Methodology used for projecting future fuel conditions

Through its vendor, Atmospheric Data Solutions (ADS), SCE has created a machine learning model which leverages soil moisture information as well as various atmospheric parameters to estimate live fuel moisture on a daily basis out to seven days. An in-depth explanation of this approach has been published in the International Journal of Wildland Fire ([CSIRO PUBLISHING | International Journal of Wildland Fire](#)).

Mathematical algorithms ([National Fire-Danger Rating System - Google Books](#)) are used to estimate dead fuel moisture. These algorithms have been widely accepted and extensively used by the fire community for decades and represent a best practice for how dead fuel moisture is assessed. SCE calculates these algorithms using meteorological output from its internal modeling to produce daily forecast out to seven days across a 1-kilometer by 1-kilometer grid.

Calculation of any proprietary fuel condition indices (or other measures tracked)

Aside from future improvements including FPI 2.0, SCE does not leverage additional proprietary fuel condition indices.

Thresholds used to identify extreme fuel conditions, including any factors used to modify thresholds (e.g., fuel type, topography)

No thresholds exist within SCE that identify such measurements.

Geospatial polygons of extreme fuel conditions within the service territory as defined in the geospatial schema (GIS Data Reporting Standard, current version)

SCE does not obtain polygons of extreme fuel conditions within the service territory as defined in the geospatial schema.

Geospatial statistical frequency of extreme fuel conditions over the last five years throughout the service territory

SCE does not obtain geospatial statistical frequency of extreme fuel conditions over the last five years throughout the service territory.

Calculation of Risk and Risk Components

This section identifies the key components of a wildfire risk analysis that the electrical corporation must quantify. The electrical corporation must be readily able to provide, if requested by Energy Safety or designated stakeholders, the information described in the following subsections: Likelihood, Consequence, PSPS Consequence, and Risk.

Likelihood

The following subsections describe likelihood risk components. Each subsection includes elements which Energy Safety considers key to the calculation of the relevant risk component; these elements are intended to establish baseline evaluation and reporting for all electrical corporations. If the electrical corporation defines other key factors as important, it should report them in a similar format.

These risk components may be combinations of other fundamental risk components. The process the electrical corporation uses to combine these risk components must be documented in section 6.2.2 of its WMP. If the electrical corporation approach uses a MAVF, the electrical corporation must be able to provide justification of each parameter (e.g., limits, scaling functions, and weights) used.

IRC1: Ignition Likelihood

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine the likelihood of an ignition throughout its service territory. Energy Safety considers the following elements key to the calculation:

- *Equipment likelihood of ignition*
- *Contact by vegetation likelihood of ignition*
- *Contact by object likelihood of ignition*

SCE considers Ignition Likelihood to be synonymous with Probability of Ignition (POI). POI is the sum of the ignition component probabilities at that location (i.e., Equipment Ignition Likelihood (FRC1), Contact from Vegetation Ignition (FRC2), and Contact by Object Ignition Likelihood (FRC3). POI is used to assess overall utility wildfire risk at a given location.

Please see Section 6.2.2 and Appendix B of SCE's 2023 WMP on SCE's methodology used to determine the likelihood of an ignition in its service territory.

FRC1: Equipment Likelihood of Ignition

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine the equipment likelihood of ignition throughout its service territory by equipment type. The types of equipment it may include:

- *Arrestors*
- *Capacitors / Capacitor banks*
- *Circuit breakers*

- *Conductors*
- *Connection points (conductors, insulators, splices, hotline clamps, and other connectors)*
- *Crossarms*
- *Fuses*
- *Poles*
- *Splices*
- *Switches*
- *Transformers*
- *Tie wires*

Energy Safety considers the following elements key to the calculation:

- *Typical operating conditions*
- *Equipment-specific failure rates*
- *Spark generation rates from normal operation*
- *Age of equipment*
- *Presence of mitigation (i.e., covered conductors, vibration dampers)*
- *Protective equipment and device settings*
- *Time since most recent asset inspection*
- *Open work requests*
- *Local weather conditions*
- *Local surface vegetation conditions*

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine ignition likelihood from events and include basis data used, such as past ignition events, number of risk events, description of events, and the statistical tools used as part of the analysis.

Equipment Ignition Likelihood, also referred to as Equipment/Facility Failure Probability of Ignition (EFF POI). EFF POI is the sum of the ignition component sub models (which are conductor POI, transformer POI, switch POI, and capacitor POI) probabilities at a given location. SCE does not specifically include arrestors, circuit breakers, connection points, crossarms, fuses, poles, splices, or tie wires in the equipment likelihood calculations.

Please see Section 6.2.2 and Appendix B of SCE's 2023 WMP submission on SCE's methodology used to determine the equipment likelihood of an ignition in its service territory.

SCE does not consider all the elements listed as input data for the sub models used to determine equipment likelihood of an ignition. SCE discusses its basis data used to determine equipment likelihood of an ignition in Section 2.1 of its model documentation provided in Attachments X through X of this document.

FRC2: Contact from Vegetation Likelihood of Ignition

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine the contact from vegetation likelihood of ignition throughout its service territory. This may include:

- *Contact from vegetation grow-in*
- *Contact from vegetation fall-in*
- *Contact from vegetation blow-in*

Energy Safety considers the following elements key to the calculation:

- *Type of contact (i.e., grow-in, fall-in, blow-in)*
- *Vegetation species evaluated*
- *Protective equipment and device settings*
- *Time since most recent vegetation inspection*
- *Local weather conditions*
- *Local surface vegetation conditions*

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine ignition likelihood from events and include basis data used, such as past ignition events, number of risk events, and description of events, and the statistical tools used as part of the analysis.

Contact from Vegetation Ignition Likelihood, also referred to as Contact from Foreign Object -Vegetation Probability of Ignition (CFO-Veg POI). In determining its Contact from Vegetation Ignition Likelihood, SCE does not differentiate contact from vegetation grow-in, vegetation fall-in and vegetation blow-in.

Please see Section 6.2.2 and Appendix B of SCE's 2023 WMP submission on SCE's methodology used to determine the contact from vegetation likelihood of an ignition in its service territory.

SCE considers vegetation species, time since most recent vegetation inspection, local weather conditions and local surface vegetation conditions within our sub model calculations. SCE discusses its basis data used to determine equipment likelihood of an ignition in Section 2.1 of its model documentation provided in Attachments X through X of this document.

FRC3: Contact from Object Likelihood of Ignition

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine the contact from object likelihood of ignition throughout its service territory. This may include:

- *Vehicle contact (pole strike)*
- *Balloon contact*
- *Animal contact*

- *Unknown contact*

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine ignition likelihood from events, including data used, such as past ignition events, number of risk events, and description of events, and the statistical tools used as part of the analysis.

Contact from Object Ignition Likelihood, also referred to as Contact from Foreign Object Probability of Ignition (CFO POI), includes contact with objects other than vegetation such as vehicles, balloon, animals, other, and unknown.

Please see Section 6.2.2 and Appendix B of SCE's 2023 WMP submission on SCE's methodology used to determine the contact from object likelihood of an ignition in its service territory.

SCE discusses its basis data used to determine equipment likelihood of an ignition in Section 2.1 of its model documentation provided in Attachments X through X of this document.

FRC4: Burn Probability

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine the likelihood wildfire will burn individual locations within its service territory. Energy Safety considers the following elements key to the calculation:

- *Local topography (i.e., elevation, slope, aspect)*
- *Local weather (i.e., statistical extreme conditions based on a 30-year average and seasonal weather)*
- *Local vegetation (i.e., type/class/species/fuel model, canopy height/base height/cover, growth rates, and moisture content)*
- *Climate change impact on fuel aridity (i.e., impact in seasonal extreme moisture content)*

Please see Section 6.2.1 of SCE's 2023 WMP for SCE's approach to this risk component.

IRC4: PSPS Likelihood

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to evaluate the annual likelihood of its issuing a PSPS for a circuit segment within its service territory. Energy Safety considers the following elements key to the calculation:

- *Weather (i.e., statistical extreme conditions based on a 30-year average and seasonal weather)*
- *Ignition risk*

To estimate PSPS Likelihood (also referred to by SCE as POD), SCE derived a 10-year historical climatology of PSPS weather conditions along distribution circuits. This historical climatology was used to determine the extent by which recent years experienced de-energization conditions at above- or below-average frequency, and to what degree mitigations reduce de-energization frequency.

Please see Section 6.2.2 and Appendix B of SCE's 2023 WMP submission on SCE's methodology used to determine the evaluate its PSPS likelihood.

Consequence

The following subsections describes consequence risk components. Each subsection includes elements which Energy Safety considers key to the calculation of the relevant risk component; these elements are intended to establish baseline evaluation and reporting for all electrical corporations. If the electrical corporation identifies other key factors as important, it should report them in the WMP in a similar format.

These risk components may be the combination of other fundamental risk components. The process the electrical corporation uses to combine these risk components must be documented in section 6.2.2 of its WMP. If the electrical corporation approach uses a MAVF, the electrical corporation must provide a table in this section along with discussion and justification of each parameter (e.g., limits, scaling functions, and weights) used.

In the table below, SCE summarizes the associated attributes, units, weights, ranges and scaling functions to convert natural units of consequences (e.g., CMI, dollars, safety) into a unit-less risk score. These components were based on the principles set forth in the S-MAP Settlement and presented in SCE's 2022 RAMP filing.

Table Appendix B-02 - Summary of MAVF Attributes

Attribute	Units	Weight	Range	Scaling Factor
<i>Safety</i>	<i>Index</i>	<i>50%</i>	<i>0 - 100</i>	<i>Linear</i>
<i>Reliability</i>	<i>Customer Minutes of Interruption (CMI)</i>	<i>25%</i>	<i>0 - 2 Billion</i>	<i>Linear</i>
<i>Financial</i>	<i>Dollars</i>	<i>25%</i>	<i>0 - 5 Billion</i>	<i>Linear</i>

SCE developed its MAVF based on the principles as set forth in the S-MAP settlement. Below is a discussion and description of each of the components shown above.

Attributes and Units

Attribute Hierarchy: Attributes are combined in a hierarchy, such that the top-level Attributes are typically labels or categories and the lower-level attributes are observable and measurable.

SCE identified 3 top-level attributes: 1) Safety, 2) Reliability, and 3) Financial. These three attributes comport with the S-MAP Settlement requirements that Safety, Reliability and Financial consequences are included. Pursuant to the referenced S-MAP Settlement Principle, the lower level attribute for Safety are a combination of observable and measurable attributes, namely serious injuries and fatalities. For purposes of risk modelling, SCE used a safety index to combine these two lower-level safety attributes in the following manner:

$$\text{Safety Index} = (\# \text{ of fatalities}) + \frac{1}{4} * (\# \text{ of serious injuries})$$

Ranges

Measured Observations: Each lower-level Attribute has its own range (minimum and maximum) expressed in natural units that are observable during ordinary operations and as a consequence of the occurrence of a risk event.

SCE selected the safety range to be between 0 and 100. The maximum safety range was chosen based on the 2018 Camp Fire, which caused over 80 fatalities. For the reliability attribute, SCE used the 2011

Southwest blackout event on September 8, 2011 as the basis for the maximum range of 2 billion CMI. Finally, to set a range for the financial attribute, SCE used the Woolsey Fire (2018) as the basis for the maximum range of 5 billion dollars. Although there have been other more destructive and catastrophic financial losses observed in California's history, SCE chose the financial range based on its recency and in light of the paramount importance of mitigating wildfire risk.

Scaling Functions

Scaled Units: Construct a scale that converts the range of natural units (Principle 2) to scaled units to specify the relative value of changes within the range, including capturing aversion to extreme outcomes or indifference over a range of outcomes.

SCE selected a linear scaling function, which converts an attribute's natural units to a scaled unitless score between 0 to 100, for all three attributes. A key difference between our 2018 RAMP report and 2022 RAMP Report is the shift of the safety consequence from a non-linear to a linear scaling function to reflect that each incremental safety event is valued the same as the previous one. The scaled score of 100 was taken directly from the S-MAP lexicon of a "Scaled Unit of an Attribute," which prescribes the value to be in terms of 0-100.

Weights

Relative Importance: Each Attribute in the MAVF should be assigned a weight reflecting its relative importance to other Attributes identified in the MAVF. Weights are assigned based on the relative value of moving each Attribute from its least desirable to its most desirable level, considering the entire range of the Attribute.

SCE selected the following weights for each attribute: Safety – 50%, Reliability – 25% and Financial – 25%. The 50% Safety weight complies with the S-MAP Settlement minimum Safety weight of 40% and is consistent with what SCE used in its 2018 RAMP. Having allocated 50% to Safety, the remainder of 50% is left to allocate between the Reliability and Financial Attributes. Based on the relative value of moving the Reliability range from 2 billion CMI to 0 and the Financial range from \$5 billion to \$0, SCE believes that equal weighting is appropriate. Thus, for purposes of RAMP analysis SCE assigned 25% to the Reliability Attribute, and 25% to the Financial Attribute.

IRC3: Wildfire Consequence

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine the consequence of a wildfire at each location throughout its service territory. Energy Safety considers the following elements key to the calculation:

- *Wildfire hazard intensity*
- *Wildfire exposure potential*
- *Wildfire vulnerability*

SCE does not consider Wildfire Hazard Intensity (FRC5) and Wildfire Exposure Potential (FRC6) to determine Wildfire Consequence. Please see Section 6.2.1 of SCE's 2023 WMP for SCE's approach to these risk components.

Please see Section 6.2.2 and Appendix B of SCE's 2023 WMP submission on SCE's methodology used to determine the consequence of a wildfire.

FRC5: Wildfire Hazard Intensity

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine the intensity of a wildfire at a location it reaches within the community. Energy Safety considers the following elements key to the calculation:

- *Local topography (i.e., elevation, slope, aspect)*
- *Local weather (i.e., statistical extreme conditions based on a 30-year average and seasonal weather)*
- *Local vegetation (i.e., type/class/species/fuel model, canopy height/base height/cover, growth rates, and moisture content)*
- *Local fire behavior (e.g., heat release rate, flame length)*

Please see Section 6.2.1 of SCE's 2023 WMP for SCE's approach to this risk component.

FRC6: Wildfire Exposure Potential

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine the exposure potential of a wildfire that reaches a community. Energy Safety considers the following elements key to the calculation:

- *Population density*
- *Residential, community, and critical infrastructure*
- *Environmental resources*
- *Social or cultural assets*
- *Economic factors (businesses and individual livelihoods)*

Please see Section 6.2.1 of SCE's 2023 WMP for SCE's approach to this risk component.

FRC7: Wildfire Vulnerability

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine the vulnerability/resilience of a community to a wildfire that reaches the community. Energy Safety considers the following elements key to the calculation:

- *Vulnerable populations (AFN, LEP, elderly)*
- *Legacy building codes*
- *Community collaborative wildfire preparedness initiatives (e.g., Firewise USA)*
- *Availability of ingress and egress*

In determining its Wildfire Vulnerability, SCE considers Access and Functional Needs (AFN), and Non-

Residential Critical Infrastructure (NRCI), and egress. Please see Section 6.2.2 a and Appendix B of SCE's 2023 WMP submission on SCE's methodology used to determine the vulnerability of a community to a wildfire

PSPS Consequence

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine the consequence of a PSPS at each location throughout its service territory. The calculation must include a combination of at least the following:

- *PSPS exposure potential*
- *Vulnerability of community to PSPS*

FRC8: PSPS Exposure Potential

The electrical corporation must be able to outline the methodology used to determine the exposure potential of a PSPS at an affected location within the community. Energy Safety considers the following elements key to the calculation:

- *Population density*
- *Residential, community, and critical infrastructure*
- *Social or cultural assets*
- *Economic factors (businesses and individual livelihoods)*

Please see Section 6.2.1 of SCE's 2023 WMP for SCE's approach to this risk component.

FRC9: Vulnerability of a Community to PSPS

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine the vulnerability/resilience of a community to a PSPS that affects the community. Energy Safety considers the following elements key to the calculation:

- *Vulnerable populations (e.g., AFN, LEP, elderly)*
- *Presence of critical infrastructure*
- *Presence of redundant systems (e.g., secondary power systems)*

In determining its Vulnerability of a Community to PSPS in MARS, SCE considers Access and Functional Needs (AFN) and Non-Residential Critical Infrastructure (NRCI). Please see Section 6.2.2 and Appendix B of SCE's 2023 WMP submission on SCE's methodology used to determine the vulnerability of a community to PSPS.

Risk

The following subsections describe ignition risk, PSPS risk, and overall utility risk. Each subsection includes elements which Energy Safety considers key to the calculation of these risk; these elements are intended to establish baseline evaluation and reporting for all electrical corporations. If the electrical corporation identifies other key factors as important, it should report them in the WMP in a

similar format.

These risks are combinations of other risk components. The process the electrical corporation uses to combine these risk components must be documented in section 6.2.2 of its WMP. If the electrical corporation approach uses a MAVF, the electrical corporation must provide a table in this section along with discussion and justification of each parameter (e.g., limits, scaling functions, and weights) used.

R2: Ignition Risk

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine the ignition risk throughout its service territory. Energy Safety considers the following elements key to the calculation:

- *Ignition likelihood (ignition LoRE)*
- *Ignition consequence (ignition CoRE)*

The calculation of ignition risk should be in alignment with the most recent CPUC decision governing RAMP filings. In the 2018 S-MAP process, this is the direct multiplication of the ignition LoRE and ignition CoRE (see S-MAP, step 3, row 13).

Ignition Risk (synonymous with Wildfire Risk) is calculated as the product of the sum of all Ignition Likelihood components and Wildfire Consequence for each asset in SCE's HTFD. Please see Section 6.2.2 and Appendix B of SCE's 2023 WMP submission on SCE's methodology used to determine its Ignition Risk.

R3: PSPS Risk

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine the PSPS risk throughout its service territory. Energy Safety considers the following elements key to the calculation:

- *PSPS likelihood (PSPS LoRE)*
- *PSPS consequence (PSPS CoRE)*

The calculation of PSPS risk should be in alignment with the most recent CPUC decision governing RAMP filings. In the 2018 S-MAP process, this is the direct multiplication of the PSPS LoRE and PSPS CoRE (see S-MAP, step 3, row 13).

PSPS risk is calculated as the product of PSPS Likelihood (synonymous with Probability of Deenergization (POD)) and PSPS Consequence for each asset in SCE's HTFD. Please see Section 6.2.2 a and Appendix B of SCE's 2023 WMP submission on SCE's methodology used to determine its PSPS Risk.

R1: Overall Utility Risk

The electrical corporation must be readily able to outline, if requested by Energy Safety or designated stakeholders, the methodology used to determine the overall utility risk throughout its service territory. Energy Safety considers the following elements key to the calculation:

- Ignition risk
- PSPS risk

The calculation of overall risk should be in alignment with the most recent CPUC decision governing RAMP filings. The 2018 S-MAP process does not explicitly cover the combination of ignition risk and PSPS risk to determine overall utility risk. However, combination through MAVFs (see step 1A) is a logical extension of the concepts presented in the settlement agreement.¹⁴ The electrical corporation may choose an alternative approach to combine these risks; however, it must describe the process in its WMP submission.

Overall Utility Risk is calculated as the sum of Ignition Risk and PSPS Risk for each asset in SCE's HTFD. Please see Section 6.2.2 a and Appendix B of SCE's 2023 WMP submission on SCE's methodology used to determine its Overall Utility Risk.

¹⁴ (D.) 16-08-018 Interim Decision Adopting the Multi-Attribute Approach (or Utility Equivalent Features) and Directing Electrical corporations to Take Steps Toward a More Uniform Risk Management Framework. CPUC, 2016.

Attachment A

OH-Capacitor Sub-Model

**Southern California Edison (SCE)
Model Documentation
Prepared for 2023 WMP Appendix B**

OH Capacitor Sub-Model

3/27/23

Table of Contents

1. Executive Summary.....	4
1.1 Model Purpose and Intended Use	4
1.2 Model Description Summary	5
1.3 Model Risk Rating	5
1.4 Model Dependency and Interconnectivity	5
1.5 Model Assumptions	6
1.6 Model Limitations	6
1.7 Overall Model Performance Assessment.....	6
1.8 Contingency Plan for Vendor Model.....	7
2. Model Framework and Theory	8
2.1 Model Inputs and Data Quality.....	8
2.2 Methodology.....	10
2.3 Suitability	13
2.4 Assumptions.....	13
2.5 Limitations and Compensating Controls	13
2.6 Model Outputs	14
3. Model Performance and Testing.....	16
3.1 Model Specification Testing.....	16
3.2 Sensitivity Analysis	20
3.3 Outcome Analysis / Backtesting	21
3.4 Benchmarking Analysis	22
4. Model Management and Governance	23
4.1 Ongoing Monitoring Plan	23
4.2 Security and Control	24
5. References	25

1. EXECUTIVE SUMMARY

1.1 Model Purpose and Intended Use

The OH (Over Head) Capacitor Model is a Probability of Ignition (POI) Sub-Model developed by SCE (Southern California Edison). At SCE, models are developed using ML (Machine Learning) algorithms for each asset, i.e., OH Capacitor, OH Switches, etc., and at each contact type level like animal, balloon, etc., as the drivers vary by asset and contact type. The OH capacitor model is refreshed annually and used to predict the probability of failure (POF) for distribution overhead capacitors.

The calibrated outputs of the OH Capacitor model—i.e., failure events—are broadly used by three categories of programs described below:

1. The Inspections and Remediations programs, which considers POI as an element in prioritization and scoping.
2. Asset Class Strategies are developed using the capacitor model to prioritize high risk capacitors for replacement strategies.
3. Risk analyses via SCE's MARS Framework.

1.2 Model Description Summary

The OH Capacitor model is a binary classification model using Random Forest—a Machine Learning technique. It predicts the probability of a capacitor igniting a spark due to equipment failure by considering available capacitor attributes and condition data (i.e., age, etc.) and other environmental and operational attributes (i.e., historical wind, number of switches, notifications, etc.).

The model is programmed in python using the libraries scikit-learn and pandas and is connected to databases such as SAP, ADS Weather, etc. The model is run once a year manually by the Data Science and Asset Analytics team. The model is calibrated every year with the full historical outage data.

Cross-references: Please refer to Section 2.1 for more information about the inputs used by the OH capacitor model along with data processing details.

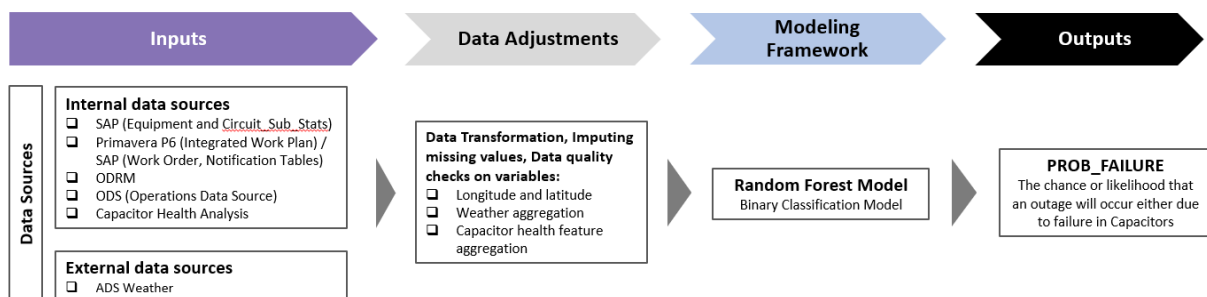


Figure 1: OH Capacitor model framework

The OH capacitor model uses the Random Forest methodology to perform classification tasks; it is considered a good choice for the OH capacitor model as the prediction is a classified event, i.e., failure. This methodology predicts output with high accuracy, runs efficiently on large datasets, and maintains accuracy with minimal adjustments for missing values and data treatments.

1.3 Model Risk Rating

There is no defined mechanism of identifying model risk rating at SCE, however certain factors—like frequency of risk events and use case—are considered while flagging model risk. Based on the Wildfire Mitigation Plan quarterly report, the frequency of outages in a year average around 338 which is moderately low compared to other sub-drivers. Figure 2 provides a snapshot of the count of outages over the years by the causes captured in the OH Capacitor model. In addition, the output of this model is considered important as it is considered in the strategy of a few programs which are discussed in section 1.1. Hence, the OH Capacitor model is deemed to be a medium risk model.

					Number of risk events														Projected risk events											
					Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4						
Table 7.1: Key recent and projected drivers of risk events																														
Risk Event category	Cause category	#	Sub-cause category	Are risk events tracked for ignition driver? (yes / no)	2015	2016	2017	2018	2019	2020	2020	2020	2020	2021	2021	2021	2021	2022	2022	2022	2022	2023	2023	2023	2023					
Outage - Distribution	18. Equipment / facility failure - Distribution	18.a.	Capacitor bank damage or failure- Distribution	Yes	280	275	372	337	426	126	159	72	46	110	98	124	80	100	96	102	90	96	96	102	90					

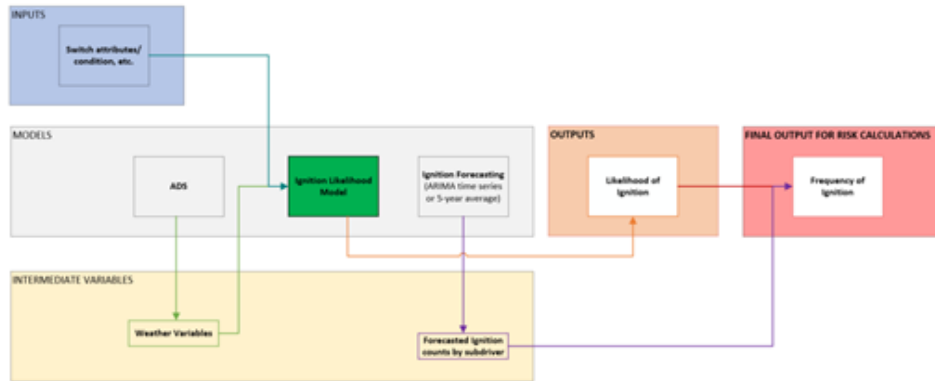
Figure 2: Key recent and projected risk events due to capacitor damage or failure from SCE Q1 2022 Quarterly Data Report, Table 7.1

Cross-references: Refer to link [RF 1: SCE’s WMP 2022 Q1 Quarterly Data Report submission

] in Section 5 for SCE’s Wildfire Mitigation Plan Q1 2022 Quarterly Data Report submission.

1.4 Model Dependency and Interconnectivity

The OH Capacitor model is an “Ignition Likelihood” model which uses the inputs from the ADS (Atmospheric Data Solutions) modeling output along with other data sources to calculate the probability of ignition.



***Ignition Likelihood model (highlighted in green) comprises of OH switch model along with other POI models

Figure 3: Model Interconnectivity Schema

OH Capacitor model uses the ADS model as one of its data sources to get the inputs for the weather variables. ADS's Next Generation Weather Modeling System (NGWMS) provides an extensive upgrade to SCE's in-house weather modeling capabilities and enhances SCE's ability to make more targeted PSPS decisions. The ADS model generates 10 years of hourly weather data between 2010 and 2019. That information is then processed and aggregated to calculate statistical measures, like mean and standard deviation, of wind, humidity, rain, snow, etc. The output data from the OH Capacitor model, i.e., POI, is used by three categories of programs, further discussed in Section 1.1, to inform their strategic decisions.

1.5 Model Assumptions

The business assumptions and model assumptions for the OH Capacitor model are summarized below:

1. There is no change in the OH Capacitor technical specification over time.
2. The calibration methodology assumes that fires are a subset of failures.
3. The model is designed to work in both base weather and extreme weather conditions.
4. The feature variables in the dataset should have some actual values so that the classifier model can predict accurate results.
5. The predictions from each tree must have very low correlations.

A detailed explanation of these assumptions is available in Section 2.4.

1.6 Model Limitations

The model limitations for the OH Capacitor model are summarized below:

1. Unavailability of linear/non-linear representation in the form of intuitive equation or correlation statistic.
2. Time consumption for model execution is high.
3. Resource utilization in terms of system capacity and higher configuration for model execution is high.
4. Model accuracy may reduce if the dataset experiences covariate shift.

A detailed explanation of these data limitations is available in Section 2.5.

1.7 Overall Model Performance Assessment

The machine learning model used to build the OH Capacitor model is the Random Forest algorithm. The model's overall performance is determined by the Area Under the ROC Curve (AUC) value and Confusion Matrix results. The performance of the OH Capacitor model was evaluated on test data with full historical outage information until 2022.

- The AUC value is 0.91.
- Confusion matrix results capture the accuracy rate as 85%.

The above metrics were derived by re-running the model as of Dec 2022 to capture an exhaustive set of statistical
OH-Capacitor Sub-Model A-6

results for documentation purposes.

1.8 Contingency Plan for Vendor Model

A contingency plan is not applicable for this model as it is an inhouse SCE model. This is not a vendor model.

2. MODEL FRAMEWORK AND THEORY

The OH Capacitor model is a binary classification model pertaining to equipment failures and employs a random forest algorithm to predict the likelihood of a capacitor experiencing an ignition event. The random forest approach was chosen for the classification task over other modeling approaches—such as logistic regression, gradient boosting, etc.—because it predicts output with high accuracy, runs efficiently on large datasets, and maintains accuracy with minimal adjustments for missing values and data treatments.

2.1 Model Inputs and Data Quality

Data Sources

This model refers to multiple internal and external data sources. The internal data sources used by the model are:

- **SAP** houses circuit¹, structure, and equipment characteristics. It contains latitude and longitude information of the assets.
- **OMS** refers to Outage Management System which contains information about switching operations. This data source contains the information about Underground switches, Overhead switches, Fuse, RAR and circuit data.
- **ODRM** refers to Outage Database and Reliability Metrics. It contains the detail for all historical outages. This information is used in conjunction with data from ODS (Operational Data Store) containing information about devices like active underground and overhead switches which is used to identify locations impacted by an outage. It is also used to record historical reliability indices which are used at a circuit level to inform the models of historic stressors from outages or transients.

The external data sources used by the model are:

- **ADS** (Atmospheric Data Solutions) model provides 10 years of hourly gridded weather data from 2010-2019. These are aggregated to individual locational measures and matched to the capacitors through spatial join to the nearest grid by the latitude and longitude as a part of the data engineering step.

Quality Checks

SCE has internal data management teams for ensuring data quality, including EAD (Enterprise Asset Data) and Master Data. They work on processing asset data corrections (E2 notifications) in SAP and fixing largely known data issues like missing or erroneous latitude and longitude information for assets in their territory. Some of the data quality checks that are performed in the OH Capacitor model to ensure the accuracy, validity, integrity, and consistency are provided below. Quality checks (QC) are incorporated coded in Python.

The QC steps performed by python code are as follows:

- Duplicate values that are identified in CKT_NAME and Circuit_Sub_Information variables are removed to maintain consistency in data by considering only the distinct values.
- The FLOC data file that is used to fetch the latitude and longitude values of the FLOCs from SAP has data quality issues. Coordinates are converted to floating point numbers and entries that are in degrees are converted to decimals. Missing or zero valued coordinates are imputed ordinally by sorting FLOCs.
- ODRM provides the information about all the outages encountered by SCE. Only the relevant information like failures specific to capacitors are loaded into this model. All the other non-relevant information i.e., for equipment other than capacitor are removed before loading the data into the model.

These data quality checks are performed across different Python programs with help of user defined functions. The data is deemed adequate as the pre- and post-performance tests during data adjustments are not conducted. The manual QC steps are as follows:

¹ Circuit comprises a collection of segments that altogether form a path for electrical current floating from the power source (including but not limited to a substation) to another power source or circuit endpoint.

- ADS weather data is validated against actual weather observations.
- Asset data obtained from SAP is validated and updated through inspections and other programs.

Data Sampling

Since this is a classification model to predict the outages, there are no sampling strategies used in the model other than the random split strategy to bifurcate the train and test data. The dataset used for the model are randomly divided to have 80% in train data and remaining 20% for test data.

Data Cleansing and Transformation

The data cleansing and transformation activities that are incorporated in the python scripts as a part of automation to ensure the completeness of data used for model training and estimation are provided below. Python codes used in the OH Capacitors model are added in the cross-references section below.

- Missing data for the below specified numeric variables are handled by imputing the mean value across the associated circuit.
 - FLOC_Latitude
 - FLOC_Longitude
 - EQ_StartUpDate
 - ICA_GEN_UDF
 - ICA_LOAD_UDF
 - INSTALLED_YEAR_UDF
- Data consistency is ensured by correcting formatting issues in date variables e.g., EQ_StartUpDate variable can have different formats of data, format is corrected in python program code for data formats to be consistent.
- Categorical variables are one hot encoded
 - EQ_SubType
 - EQ_SwitchType
- Reliability indices SAIDI, SAIFI, and MAIFI are summed to the circuit, for a year, then averaged over 5 years to get an average yearly circuit reliability which is applied to appropriate capacitors based on what circuits they are attached to.

Data Assumptions

The accuracy of the predicted results is dependent on the accuracy of the data used to build the predictive models. Following are the data assumptions:

1. The assumptions used for the data imputation utilized SCE's Distribution Design Standard (DDS), engineering judgement, manufacturer data and acceptable engineering practices.
2. The target labeling process used to label the failures and non-failures as '1' and '0' is considered accurate. This is performed by comparing the outage count in a segment against the mean value of total outages. If outage count in a segment is greater than the mean value of total outages, the '1' is assigned which represents a failure. Else '0' is assigned for non-failures.
3. For performing the mean-by-circuit imputation for locations, it is assumed that distribution circuits do not cover more than a few miles of territory and since the locations are used to assign weather values, missing locations provide a reasonable estimate within the resolution of the weather data. Dates are used to calculate in service age, as such it is assumed that missing dates are the same as the median startup date making capacitors with missing dates the same age as most capacitors on the circuit.
4. Input data with respect to asset information, weather information and engineering information are assumed to be stable and will not change over time until the subsequent data refresh. Example: If there is an update in the structure information specific to an asset, that updated information will be reflected only

in the subsequent data refresh. So, it is assumed that the updated structure information is not drastically different from the previous information which might alter the model outcomes.

Data Limitations

Following are data limitations across internal and external data sources:

1. Some of the data used by the model faces accuracy issues in terms of consistency in data labelling, missing data for a specific feature (predictive variable) which might impact model prediction power.
 - Data labelling issues might be caused due to manual errors during data entry task. While updating the type, different label might be used in different data entries which affects the consistency of the data. Hence these consistency issues in data needs to be addressed before using them in the model.
2. With respect to Failure targets, the starting location of the outage is recorded at the FLOC and associated with a circuit. This combination of FLOC and circuit is used to identify which capacitor experienced the failure.

Independent variables

The OH Capacitor model uses multiple variables/features. Some of these features are created based on engineering knowledge and some are selected based on expert advice. A subset of the independent variables used in the OH Capacitor model along with its data source and description, is provided below.

Feature	Data Source	Description
InServiceYear	SAP	Equipment age, calculated by subtracting EQ_StartUpDate from year of subset.
EQ_NumberOfSwitches	SAP	Total number of switches on the capacitor bank
Structure_StartUpYear	SAP	Parsed from SAP EQ_StartUpDate, engineers identified that there were certain design standards that were changed in certain years.
Avg_air_temperature_2m	ADS	Average hourly air temperature
Stddev_pop_air_temperature_2m	ADS	Standard deviation of the hourly air temperature

10 years of hourly data fetched from ADS Weather model is processed and aggregated to calculate the numerical measures like mean, max and standard deviation for wind, temperature, water vapor, turbulence kinetic energy, humidity, rain, and storm. Asset information in a segment is fetched from SAP.

Dependent Variable

In a typical classification risk model, defining the dependent variable is key for both model development and model performance assessment. The dependent variable in OH Capacitor model is PROB_FAILURE.

PROB_FAILURE represents the chance or likelihood that an outage will occur due to failure in capacitors due to equipment failure. The probability value ranges from 0 to 1 where '0' represents the least likelihood for an outage and '1' represent the high chance for an outage. The target variable represents the chance or likelihood that an outage will occur due to failure in capacitors.

2.2 Methodology

SCE utilizes machine learning to identify patterns that may lead to failures causing sparks from capacitors and uses the trained model to predict POIs at the asset level. The OH Capacitor model employs a random forest algorithm to predict failure events. The Random Forest approach can predict outputs with high accuracy, run efficiently for large

datasets, and maintain accuracy with minimal adjustments for missing values and data treatments.

A random forest is a supervised machine learning algorithm that is constructed from many decision trees. It can be used to solve both classification and regression problems. This approach utilizes ensemble learning, which is a technique that combines many classifiers to achieve greater predictive accuracy than that of a single classifier. A decision tree is a decision support technique that forms a tree-like structure. It consists of three components: decision nodes, leaf nodes, and a root node. The following diagram shows the three types of nodes in a decision tree.

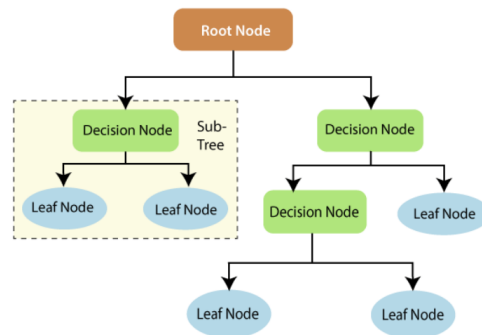


Figure 4: Decision Tree Structure

A decision tree algorithm divides observations of a dataset into branches, which further segregate into other branches. This sequence continues until a leaf node is attained. A leaf node cannot be segregated further. In more detail, the root node is the base of a decision tree, where the first of a chain of decisions is made. A branch is the connection path between nodes. A node is a potential splitting point on a tree. Decision nodes provide a link to the leaves. On the other hand, leaves, also known as terminal nodes, are the ends of a tree, representing the resulting classification or value for the sample.

The ‘forest’ generated by the random forest algorithm is trained through bagging, also known as bootstrap aggregating. Bagging is an ensemble meta-algorithm that fits multiple models on different subsets of a training dataset and then combines the predictions from all models. The diagram below shows a simple random forest classifier.

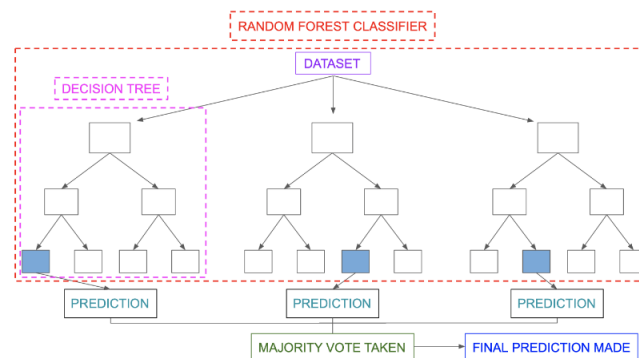


Figure 5: Structure of Random Forest Classifier model

The selection of the final output follows a majority-voting system. In this classification model case, the output chosen by a majority of the decision trees becomes the final output of the random forest system. The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.

Train test split is a model validation procedure that simulates how a model would perform on new/unseen data. Figure 6 shows the logic of dividing the dataset into train data and test data. First the data is consolidated and prepared for train test split. Then the historical input datasets are split into a training dataset (80%) and testing dataset (20%) based on simple random sampling strategy with a split ratio of 4:1 without replacement. Simple random sampling is a technique that ensures each observation has an equal likelihood of being selected for a set.

It is a fair strategy as it helps in avoiding any bias involved compared to other modeling techniques and it has no restrictions on the sample size which makes it suitable to handle vastly sized input data. The predictive algorithm is developed using the training dataset and is built by looking at the interactions between all the features to find patterns and predict the likelihood of equipment failure.

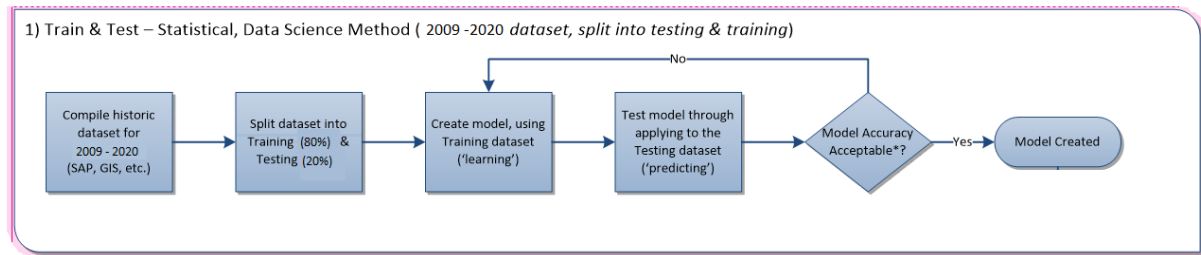


Figure 6: Train and Test data split logic

In the next step, the algorithm is tested on the ‘testing’ dataset. The model is run on the test dataset to make a prediction of a failure or success. Then an internal validation of the model is conducted by comparing the predicted results to the actual results which indicates the predictive capabilities of the features as well as the model. AUC is the metric used to assess the performance of the model on test data.

Area Under the Curve (AUC) – Area Under the Receiver Operating Characteristic (ROC) Curve is a measure to estimate the model discriminatory power (degree of separability) for the binary classification problem. The ROC curve plots True Positive Rate against different thresholds with False Positive Rate (FPR) or True Negative Rate (TNR). The higher the AUC, the better the model is at predicting True Negatives (non-events) and True Positives (events).

Hyperparameter Tuning:

Hyperparameters are parameters that are explicitly defined by the user to control the learning process. The process of selecting the optimal hyperparameters to use is known as hyperparameter tuning, and the tuning process to achieve the best-defined performance statistic is known as hyperparameter optimization. Cartesian Grid search and Random Grid search are the most widely used strategies for hyperparameter optimization.

- In the Cartesian grid search approach, the machine learning model is evaluated for a range of hyperparameter values, and it searches for the best set of hyperparameters from a grid of hyperparameters values. The disadvantage of grid search model is that it will go through all the intermediate combinations of hyperparameters which increases the time consumed by grid search computations. An example of how a cartesian grid search can affect different model performance values is studied in the code fragments and resulting plot below as the class weight hyperparameter is varied.
- In the random grid search approach, the machine learning model is evaluated for a range of hyperparameter values like that in Cartesian Grid Search approach. However, search criteria parameters are added to control the type and extent of the search, and it moves randomly within the grid to find the best set of hyperparameters to achieve maximum performance in terms of the metric defined by the user. As search criteria, the user can set a maximum runtime for the grid, a maximum number of models to create, or metric-based automatic early stopping. If many of these requirements are supplied, the algorithm will end when the first of the criteria is met. This approach reduces the time taken for computation thereby solves the drawbacks of the cartesian grid search approach.

The OH Capacitor model uses the Halving Random Grid Search method for Hyperparameter tuning. The reference literature link to understand the efficiency between Cartesian Grid search and Random Grid search is provided below. The criterion used for the hyperparameter tuning in OH Capacitor Model are:

- **N_estimators**: Total number of trees used in the random forest.
- **max_depth**: Specifies the maximum size of the sample data drawn for training each tree. A higher value for this feature will make the model more complex and can lead to overfitting issue.

- **min_samples_split:** This parameter defines the minimum number of observations required in order to allow for another decision node to split.
- **Class_weight:** This parameter gives heavier preference to specific classifications and is useful for training models with large class imbalance.

Halving Random Grid Search method uses random grid search methodology along with recursive halving and k-fold cross validation to find optimal hyperparameters.

Once the grid search completes, the grid object containing the list of models is queried, and models are sorted by a performance metric defined by the user. The model with better performance is chosen as the best model and it is validated on the test data.

[Cross-references:](#) Refer to [RF RF 1: SCE's WMP 2022 Q1 Quarterly Data Report submission

] in Section 5 to understand the efficiency between Cartesian Grid search and Random Grid search.

2.3 Suitability

Theoretically, the Random Forest methodology exhibits higher level of accuracy, stability and handles non-linear parameters efficiently than other approaches. Additionally, the random search approach used in hyperparameter tuning controls the maximum depth of the sample data drawn for training each tree and involves stopping criterion which reduces the computation time and avoids the overfitting issue.

Hence, the usage of Random Forest for the OH Capacitor model is deemed to be fit.

2.4 Assumptions

The key business assumptions that were considered during model development are specified below:

BA 01: There is no change in OH Capacitor technical specification over time. The model assumes the type of OH Capacitors used in the model building process have same characteristics in terms of build and quality.

BA 02: The Calibration model assumes that fires are a subset of failures. Capacitor failures do not always result in service loss to customers. As such, outages only represent a subset of capacitor failures. All failures are captured in the inventory of capacitors, the SAP equipment tables. These removal codes detail various reasons for removal. The removal codes used for the capacitor predictive model are related to failure, deterioration, and damage. Codes pertaining to removal due to circuit redesigns or idling are not used to quantify failures. Hence capacitor removals due to failure, damage, or deterioration are considered as events in model development. These removal conditions indicate capacitor conditions that can potentially spark an ignition as the failure target which can turn into a fire, but all failures will not result in a fire. Hence, fire can be treated as a subset of failure.

BA 03: The model is designed to work in both base weather and extreme weather conditions. The weather variables considered by the model are represented as various statistical aggregations like max, mean and standard deviation on wind, wind speed, humidity, rain, and snow. Hence the model results can be used under both base weather and extreme weather conditions.

The functional/model methodology assumptions that were considered during model development are discussed in detail below:

MA 01: The feature variables in the dataset should have some actual values so that the classifier model can predict accurate results. In an ideal scenario, all the variables would not have estimated values and they would instead use actual values. The current model is able to provide accurate results even after using estimates as they are derived through imputation using actual values from other variables.

MA 02: The predictions from each tree must have very low correlations. It is difficult to differentiate between a real interaction effect, marginal effects, and just random variations in random forests. Hence, the presence of highly correlated variables in Random Forest approach will have an impact on its ability to identify strong predictors.

2.5 Limitations and Compensating Controls

The key model limitations that would impact the accuracy and performance of the model are discussed in detail below:

Limitation ID: L01

Limitation Title: Unavailability of linear/non-linear representation in the form of intuitive equation or correlation statistic.

Description: The Random Forest algorithm does not explain any linear or non-linear relationship in the form of an intuitive equation or correlation statistic to enable measurement of the scalability of impact of independent variables on the dependent variable.

Compensating Controls: The Random Forest model is considered a black box as it is difficult to understand the relationship between independent and dependent variables and how the independent variables influence the predictions. Since black box is a common limitation with most ML algorithms, usage of the model is considered appropriate as it provides better AUC results than other models.

Limitation ID: L02

Limitation Title: Time consumption for model execution is high

Description: Since Random Forest models use a bagging algorithm, they can provide more accurate predictions but slow down the process as they compute data for each decision tree.

Compensating Controls: To overcome the time consumption issues from grid search computations, random grid search is used in the hyperparameter tuning process. Random grid search is a proven technique to reduce the time consumption when testing multiple models with different combinations of hyperparameters by using stopping criterion like tolerance, maximum rounds, maximum run time, and performance improvement thresholds. It moves within the grid in a random fashion to find the best set of hyperparameters to achieve maximum performance in terms of the metric specified, here AUC. Since the model is not executed through computer program automatically at a defined frequency and is instead run only once a year manually, usage of the model is considered appropriate.

Limitation ID: L03

Limitation Title: Resource utilization for model execution is high

Description: Since Random Forest models process many decision trees, they need more resources with respect to system configuration and system capacity to store that data.

Compensating Controls: The resource utilization factor will have a major impact for real time models as they would run more frequently. Since the OH Capacitor model is run only once a year with reasonable use cases, the impact of resource utilization is low. Additionally, the usage of random grid search and stopping criterion like tolerance, maximum rounds, maximum run time, and performance improvement thresholds provide more control on the number of recurring instances run to identify the best fit hyperparameters to achieve optimal AUC. Since the model is not executed through computer program automatically at a defined frequency and is instead run only once a year manually, usage of the model is considered appropriate.

Limitation ID: L04

Limitation Title: Model accuracy might reduce if the dataset experiences covariate shift.

Description: Covariate shift is a type of model drift which occurs when the distribution of independent variables changes between the training environment and live/test environment. Since the Random Forest cannot extrapolate (i.e., predict outside the training space), the model performance might decrease if there is covariate shift in the dataset.

Compensating Controls: The covariate shift affects most machine learning models to some degree, as test data is never going to be the same as training data. Detecting and addressing covariate shift is therefore a key step to the machine learning process. The current model is run only once a year along with data refresh. It uses a random sampling mechanism to split the dataset into train (80%) and test (20%) data whenever it is run. The usage of random sampling mechanism is considered to resolve the issue of covariate drift and maintains the accuracy of the model results. Hence the usage of the Random Forest methodology along with the random sampling mechanism to split train/test data is considered appropriate.

2.6 Model Outputs

The OH Capacitor model predicts the probability of ignition (POI) arising from equipment (capacitor) failure. The model has a single output characterized by a continuous number between 0 and 1 for each OH Capacitor asset. The probabilities across different asset failure predictive models cannot be aggregated or compared and hence are calibrated to derive frequencies of ignition. The sum of the resulting frequencies of ignition for a sub-driver equals the total expected ignitions for the specified year.

$$\text{Frequency of Ignition} = \text{Probability of Ignition} \times \frac{\text{Calibrated Targets}}{\sum \text{Probability of Ignition}}$$

where Calibrated Targets = Forecasted Ignitions for that sub-driver

The output from this calibration exhibits the following features:

- Frequency: Each value can be specified as the frequency of fires per year.
- Comparability: The frequencies are comparable against sub-drivers and models.
- Additivity: The frequencies can be added across models to derive the aggregated fire forecast in a year.

This is achieved by forecasting fires by sub-driver and using these forecasts to weight model probabilities. The sum of probabilities from each calibrated model equals the forecast by sub-driver.

Figure 7 provides the calibration steps that are performed using the failure probability results from the OH Capacitor model. This methodology followed in the calibration model is provided below:

- Aggregate the probability output from each sub-driver model.
- Based on the forecast logic specified above, find the forecast results (expected fires) for each sub-driver.
- Generate the calibration factor for each sub-driver based on the values calculated in the above steps (B/A).
- Multiply each model probability by its calibration factor to arrive at the estimated frequency of fires from each sub-driver.

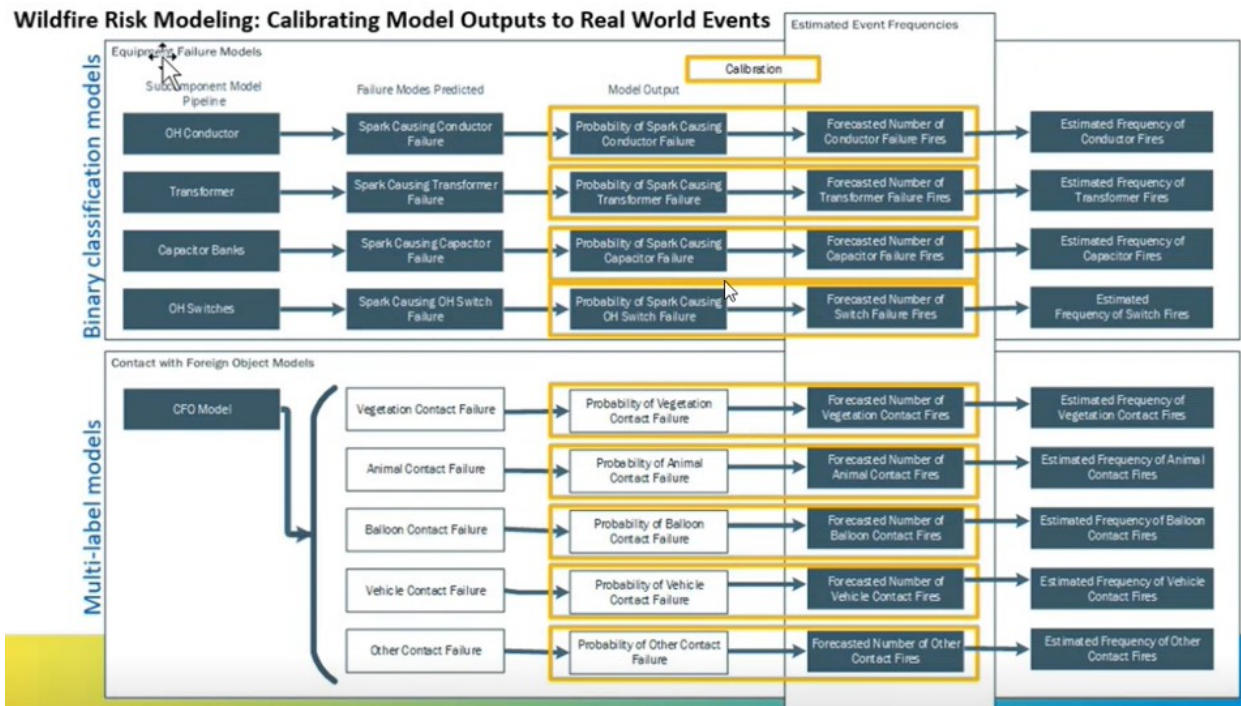


Figure 7: Calibration model schema

This estimated frequency of fires from each sub-driver can be added across the models to derive the expected frequency of ignition for each location.

The calibrated probabilities, frequencies of events, based on the output from OH Capacitor model is the data ingested to inform the programs mentioned in Section 1.1.

Model Changes:

Till September 2021, the OH Capacitor model only used SAP REMOVAL CODE to signify a failure. As part of the greater asset class strategy effort, two other ways of identifying failures were added: Outages from ODRM and Notifications from SAP. Removal code is a record in the capacitor inventory that records when specific capacitors are removed from the system due to deterioration, damage, or failure. Outages are records of failures that result in customer loss of service. Notifications are records from inspections teams to identify capacitors of concern. Although the three options exist, for wildfire risk assessment, SAP removal code is used because of its completeness and quantity compared to notifications and outages as well as to keep aligned with previous models.

3. MODEL PERFORMANCE AND TESTING

For each machine learning model developed, SCE tries to select the best algorithm based on the model train/test performance, which can be measured by Area Under the Curve (AUC) and other metrics from the Confusion Matrix.

3.1 Model Specification Testing

The model is developed and tested in python using the library scikit-learn. The model is run once a year manually by Data Science and Asset Analytics team.

The verification of the model implementation is performed by checking the variable importance results which provides the list of features implemented. The performance of the model is validated through the AUC, defined in Section 2.2 and provided in Section 3.3.

The validity and impact of the Model Assumptions are discussed below:

- The features used in the model are expected to have some actual values so that the model results can be accurate. In an ideal scenario, all the variables would not have estimated values and they would instead use actual values. Missing values are estimated by imputation using average values shared with like equipment on the same circuit. After using these estimates, the data quality is enhanced to support reliability of the current model in terms of improved predictive accuracy.
- Random Forest is considered as a strong approach for variable selection in high-dimensional data only when the variables have low correlation. The recursive structure of trees generally enables them to take dependencies into account in a hierarchical manner. Specifically, a different behavior in the two branches after a split indicates possible interactions between the predictor variables. However, some variable combinations without clear marginal effects might make the tree algorithm ineffective. To conclude, it is difficult to differentiate between a real interaction effect, marginal effects, and just random variations in random forests. Hence, the presence of highly correlated variables in Random Forest approach will have an impact on its ability to identify strong predictors. Adequate measures are taken to filter out the highly correlated features to overcome their impact in predicting the results.

Model Estimation:

The OH Capacitor model employs a number of independent variables. Section 2.1 contains a list of the independent variables utilized in this model.

The variable importance test results for the OH Capacitor model, Figure 8, shows the order of which features provided the most information gain in informing the correct prediction of failure or non-failure. The variable importance features test estimates the relative influence of each variable by calculating whether that variable was chosen to split during the tree building process and how much the squared error (over all trees) improved (or decreased) as a result.

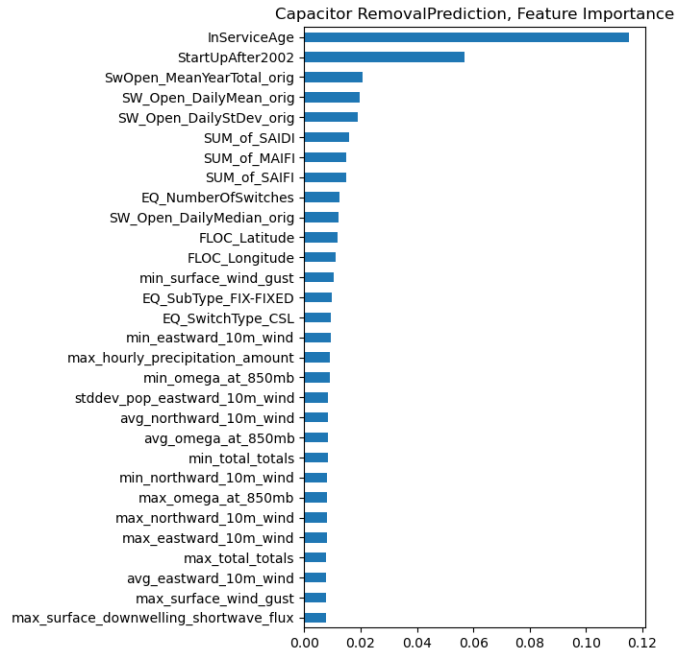


Figure 8: Variable Importance test results for OH Capacitor model

The results confirm that the design changes after 2002 affect failure rate, as well as switch actuation variables having to do with counts of openings, number of switches, and reliability indices exhibits high importance on model output.

Cross-references: Refer to link [RF 3RF 1: SCE's WMP 2022 Q1 Quarterly Data Report submission

] in Section 0 for description on the methodology used to perform the Variable Importance for tree-based methods. The OH Capacitor model uses the halving random grid search approach for hyperparameter optimization to select the best set of hyperparameters to achieve maximum performance in terms of AUC and recall. Once the grid search is completed, a list of models with their associated hyperparameter values is obtained. The acquired models are then sorted based on the AUC values for the model while also considering that maximizing capturing failures and minimizing misclassifications provide more operational benefit than classifying nonfailures. To this end, precision, recall, and f1 score are also considered as tradeoffs to AUC to find the optimal values for all. The best model is run on the respective test data, and the AUC metric is used to evaluate the performance of the models.

The AUC is used to estimate the model discriminatory power to predict the results in a binary classification problem. A higher AUC means the model can predict the results accurately. Figure 9 shows the AUC ROC for OH Capacitor based out of test dataset ran with holdout SAP removal data through 2021. The AUC value for the optimal model is 0.91. Optimal model hyperparameters are:

Hyperparameter	Optimal value
N_estimators	1000
Max_depth	34
Min_samples_split	49
Class_weight	7

In terms of model convergence, the random grid search for hyperparameter tuning uses a stopping criterion based on a specified tolerance in AUC. This means that the additional efforts involved in hyperparameters tuning and training is not likely to improve the model performance beyond the specified threshold.

	precision	recall	f1-score
False	0.94	0.86	0.90
True	0.62	0.81	0.70
accuracy			0.85
macro avg	0.78	0.83	0.80
weighted avg	0.87	0.85	0.86

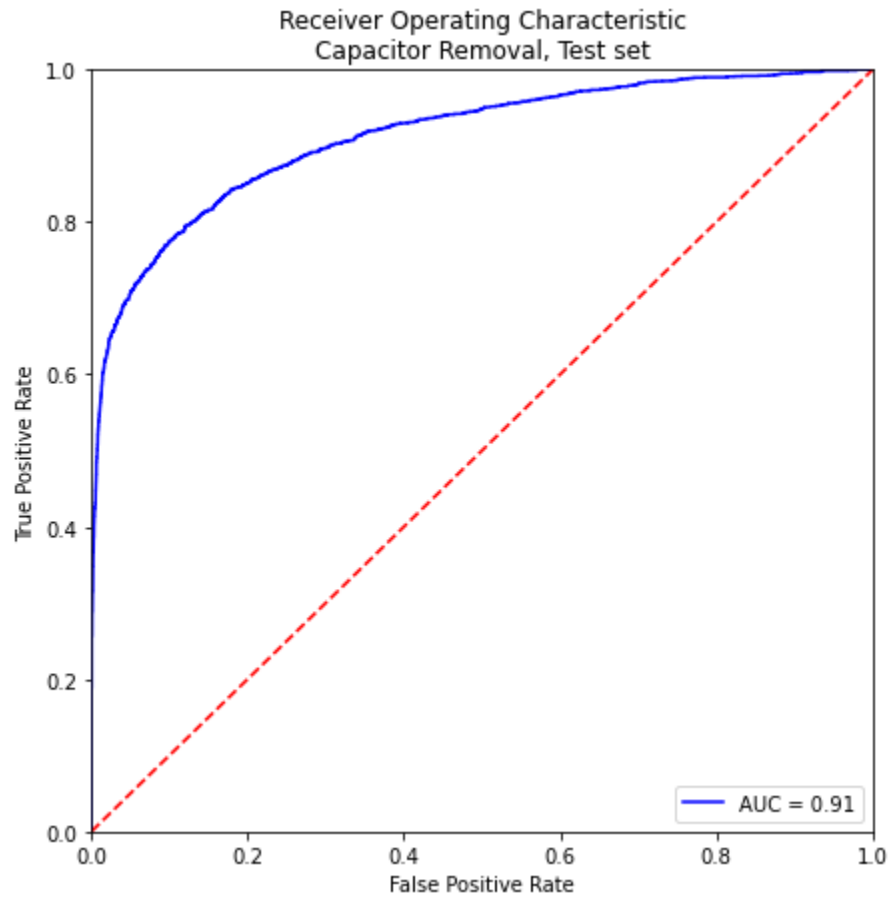
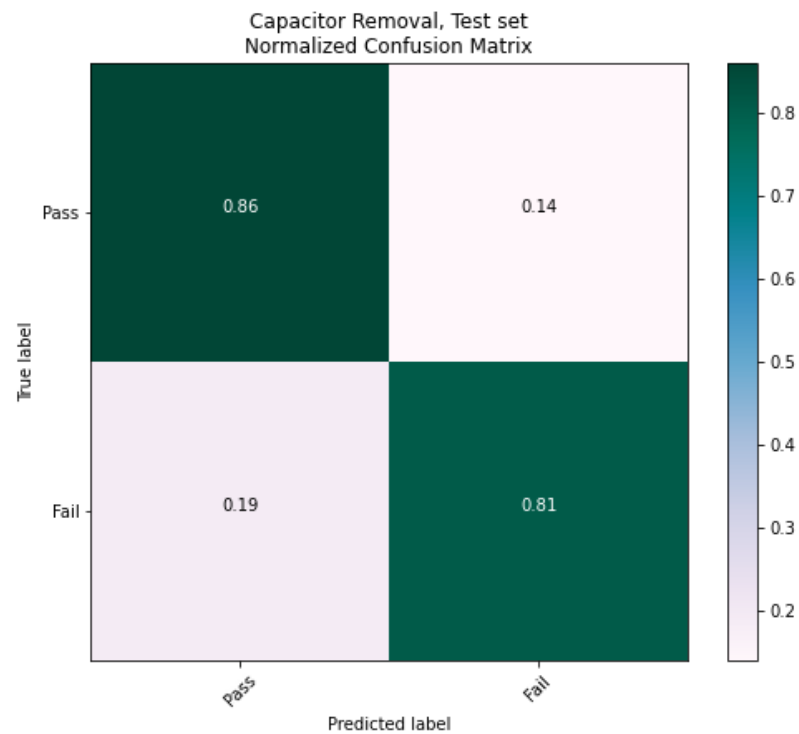
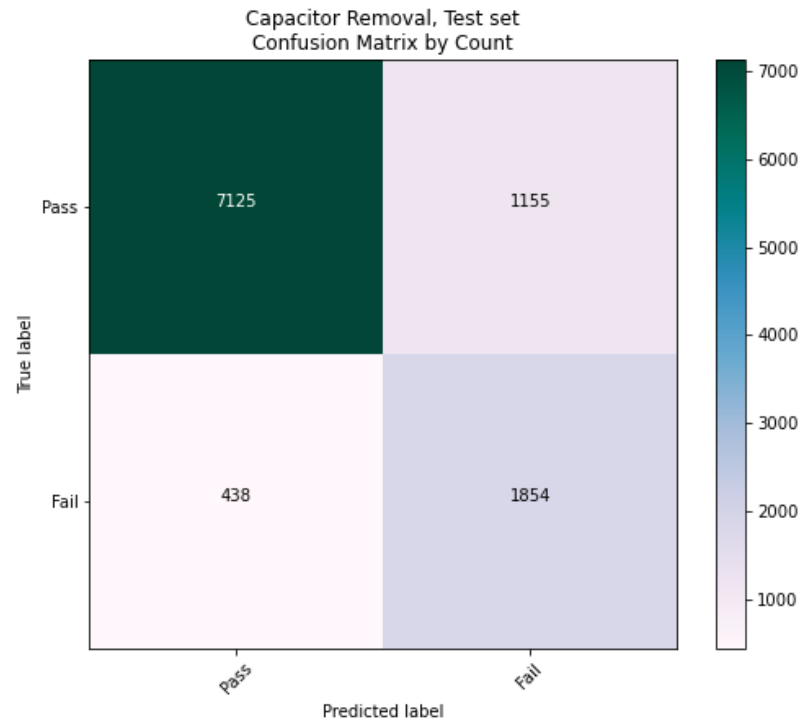


Figure 9: OH Capacitor Receiver Operating Characteristic curve and AUC

The accuracy of the model prediction, in addition to AUC, can be determined using the Confusion Matrix and Classification Error Rate results.

- A confusion matrix presents a tabular layout of the different outcomes of the prediction results of a classification problem and helps visualize its outcomes. It generates a table of all the predicted and actual values of a classifier model.



Figures 10 and 11: Confusion matrix results by count and normalized

- Figures 10 and 11 provide the confusion matrix results for the OH Capacitor model. It captures the accuracy rate as 85% on the test data.

- Classification error rate is used to estimate the proportion of instances misclassified over the whole set of instances. It is estimated using the below formula.

$$\text{Error Rate} = \frac{\text{False Positives} + \text{False Negatives}}{\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives}} * 100$$

The error rate turns out to be 15%. This means that the failure rate of the model prediction is moderately low and under control.

All these test results are performed on test dataset with holdout SAP removal data through 2021.

A detailed assessment of the model limitations and associated compensating controls are available in Section 2.5.

3.2 Sensitivity Analysis

Sensitivity analysis examines the impact of each feature on the model's prediction. It is a simple yet powerful technique to analyze a machine learning model. To determine the sensitivity of a feature, its value is changed while the values of all other features are held constant. The model's output is then examined. If the outcome of the model significantly changes when the feature value is changed, this indicates that the feature has a significant influence on the prediction. Based on the variable importance feature list shown in Figure 8, continuous variables from top five was chosen to perform the sensitivity analysis.

The feature variables used for sensitivity analysis of the model:

- InServiceAge
- StartUpAfter2002
- Sw_OpenMeanYearTotal
- SW_Open_DailyMean
- SW_Open_DailyStDev

The sensitivity of the model is examined using the test dataset, which contains 20% (10573 observations) of the entire processed data. For this analysis, 10% of the test dataset was selected and modified using extreme values. Stratified sampling was used to select the 10% of the test data to add randomization to eliminate sampling bias.

The test data is bound to a column of random numbers produced using a standard normal distribution, and the rank of these random numbers is used to sort the entire set of test data. The top 10% from each stratum was selected as the target observations to modify the input data. To test the sensitivity of a feature, the values of the selected observations were altered with extreme values (minimum and maximum) of the feature. As a result, for each feature, two sets of test data were generated for sensitivity analysis. The table below provides the extreme values (determined by historical data) used for each variable for the sensitivity analysis.

Extreme values used for Sensitivity testing		
Variables	Maximum Value	Minimum Value
InServiceAge (days)	36891	1
StartUpAfter2002	True	False
Sw_OpenMeanYearTotal	2164.75	0
SW_Open_DailyMean	6.956113	0
SW_Open_DailyStDev	18.542763	0

Extreme values used for Sensitivity testing

The table below provides the AUC and classification rate results of the unaltered test data i.e., test data without changing the variables values, and the various sensitivity tests that were performed by altering the inputs to minimum or maximum values for a single variable.

Variable	Variable Value	tnr	fpr	fnr	tpr	auc	change in AUC
Full Unaltered Test Set	Unaltered	0.89	0.11	0.18	0.82	0.91	0
InServiceAge	maximum	0.75	0.25	0.14	0.86	0.81	-0.10
InServiceAge	minimum	0.98	0.02	0.52	0.48	0.73	-0.18
StartUpAfter2002	maximum	0.91	0.09	0.30	0.70	0.81	-0.10
StartUpAfter2002	minimum	0.76	0.24	0.15	0.85	0.80	-0.11
Sw_OpenMeanYearTotal	maximum	0.90	0.10	0.14	0.86	0.88	-0.03
Sw_OpenMeanYearTotal	minimum	0.90	0.10	0.19	0.81	0.86	-0.05
SW_Open_DailyMean	maximum	0.89	0.11	0.21	0.79	0.84	-0.07
SW_Open_DailyMean	minimum	0.90	0.10	0.19	0.81	0.86	-0.05
SW_Open_DailyStDev	maximum	0.90	0.10	0.20	0.80	0.85	-0.06
SW_Open_DailyStDev	minimum	0.88	0.12	0.19	0.81	0.84	-0.07

Table 1: The sensitivity results based on AUC

3.3 Outcome Analysis / Backtesting

The subset of historical data on which a model is trained and optimized is referred to as the in-sample data. On the other hand, the subset of the dataset that has been reserved to test the model is known as the out-of-sample data. The OH Capacitor model uses a random sampling approach to split the dataset into Train (80%) and Test (20%) data. The results arrived from train data are considered as in-sample backtesting and the results arrived from test data are considered as out-of-sample backtesting.

Once the machine learning model is built with the training data, it is evaluated using a separate test dataset that has not yet been studied. The performance of the model is determined by the Area Under the ROC Curve (AUC) value. Figure 12 shows the AUC value and ROC for the capacitor model based on the test dataset ran with random forest. The AUC value of 0.91 implies that the model possesses high accuracy in terms of predicting the results.

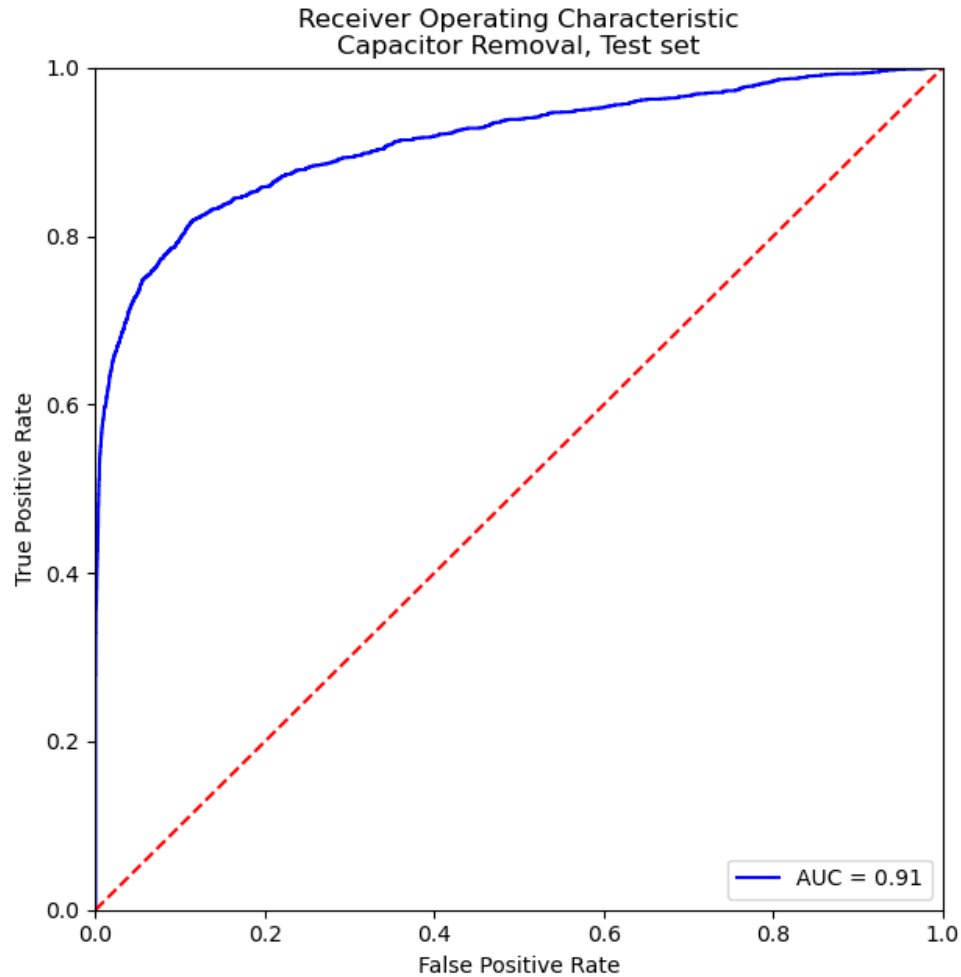


Figure 12: Out-sample backtesting result for OH Capacitor model based on test dataset

The impact of uncertainty in model inputs and parameters on model outputs are tested as a part of the sensitivity analysis and the results are captured in Section 3.2. In addition, the data imputations that are incorporated to address missing values before running the model are defined in Section 2.1.

3.4 Benchmarking Analysis

For the OH Capacitor model, different approaches like Gradient Boosting Machine (GBM) learning, Support Vector Machines (SVM), and Random Forest were considered during the model development phase in 2019. The analysis on these supervised machine learning approaches and the results are provided below.

- Gradient Boosting Machine (GBM) is one of the most popular forward learning ensemble methods in machine learning. It is a powerful technique for building predictive models for classification and regression tasks. GBM sequentially combines the predictions from various weak learner decision trees and builds a final predictive model with more accurate predictions by minimizing a defined loss function.
- Random Forest is a popular machine learning algorithm that can be used for both classification and regression problems. Random Forest is another ensemble method that combines the predictions of several decision trees to improve the predictive accuracy of the model. The individual decision trees are created based on a randomly selected subset of features at each node prior to determining the optimal split so each tree differs. The final output is determined by taking the majority vote of the predictions from the individual decision trees. The greater number of trees in the forest generally leads to higher accuracy and prevents

the problem of overfitting. Support Vector Machines attempt to calculate mappings in multivariate space that make for differentiation between failure classes

The benchmarking results of GBM and SVM shared in this section were developed using scikit learn library on the Test data with targets from the last 5 years of historical failure data (2017-2021). Since benchmark results were not saved during the model development phase, the benchmark models were executed in March 2023 for documentation purposes. Figure 13 provides the AUC values for the OH Capacitor model using the GBM, SVM, and Random Forest methodologies.

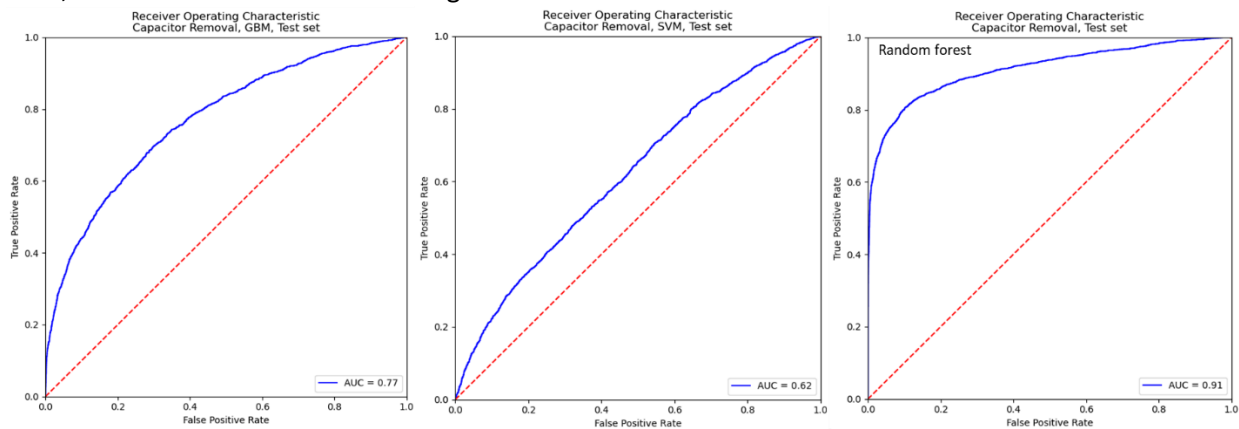


Figure 13: Gradient boosting, Support vector machines, and random forest machine learning algorithm performance on the capacitor predictive model data set compared

For the OH Capacitor model, the AUC results for GBM, SVM, and Random Forest were 0.77, 0.62, and 0.91 respectively.

Random Forest was chosen as the modeling algorithm for the OH Capacitor model because it achieved the highest AUC among the three approaches. Some additional advantages of using Random Forest over GBM and SVM are:

- Random Forest is less sensitive to overfitting issues than GBM.
- With fewer hyperparameters, hyperparameter tuning is relatively easy in Random Forest when compared with GBM.
- SVM attempts to create a functional map for the entire data set which is computationally expensive and does not handle intricate covariances very well.
- Random Forest offers much easier explainability and intuitive understanding than both GBM and SVM

4. MODEL MANAGEMENT AND GOVERNANCE

4.1 Ongoing Monitoring Plan

Ongoing monitoring is important for Machine Learning models especially when they are used to make predictions or when they are run on datasets with high volatility in variable values. The OH Capacitor model is run manually once a year, incorporating updated input datasets to reflect the latest available data and implementing any specific model enhancements—e.g., inclusion / replacement / removal of a feature, optimization of the hyperparameters, evaluation of a new performance metric, etc. During the model refresh, the limitations and assumptions of the model are also revisited by the model developers and necessary action items are conducted to address them. Performance monitoring is required only after running the model. Recalibration of the model has not been performed for the last two years, and it is performed only if the behavior of the model differs from that of the previous model or if there is a significant drop in model performance. The AUC and accuracy rate from confusion matrix results obtained after model refresh are compared against a threshold of 80%; if the value drops below this threshold, the reason behind the performance dip is investigated. Post-investigation, the steps required to improve

the performance of the model will be carried out. To monitor the model performance more thoroughly, developers of the model plan to additionally evaluate metrics like Precision and Recall. Precision is the positive predictive value which represents the proportion of predicted failures that were predicted correctly. Recall is the true positive rate which represents the proportion of actual failures that were predicted correctly.

The model documentation and the performance results are updated once a year immediately after the model refresh.

4.2 Security and Control

The Data Science and Asset Analytics team has access to the data inputs, code, and implementation for the model. Other business units, like the Grid Hardening Strategy team, are provided access to the model outputs upon request but cannot update or modify the code.

The model is run using Python programming and it can be executed in Python 3. Current model versioning is labeled by date of refresh (e.g., CapBankModel\ACS_2022). There are plans to move the code to GitHub, a platform that facilitates version control by tracking changes to the source code. Users with write or admin privileges to the repository can review proposed changes and approve them.

A contingency plan is not applicable for this model as it is an inhouse model for SCE.

5. REFERENCES

RF 1: SCE's WMP 2022 Q1 Quarterly Data Report submission

<https://www.sce.com/sites/default/files/AEM/Data%20Requests/2022/SCE%20Q1%202022%20Tables%201-12.xlsx>

RF 2: Literature reference on grid search vs random search approach for hyperparameter tuning.

<https://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf>

RF 3: Variable Importance methodology for tree-based methods

[Variable Importance — H2O 3.38.0.3 documentation](#)

Attachment B

OH-Conductor Sub-Model

**Southern California Edison (SCE)
Model Documentation
Prepared for 2023 WMP Appendix B**

OH Conductor Sub-Models (CFO & EFF)

3/27/23

Table of Contents

1.	EXECUTIVE SUMMARY	4
1.1	Model Purpose and Intended Use	4
1.2	Model Description Summary	5
1.3	Model Risk Rating	6
1.4	Model Dependency and Interconnectivity	6
1.5	Model Assumptions	7
1.6	Model Limitations	7
1.7	Overall Model Performance Assessment	7
1.8	Contingency Plan for Vendor Model	7
2.	MODEL FRAMEWORK AND THEORY	8
2.1	Model Inputs and Data Quality	8
2.2	Methodology	13
2.3	Suitability	15
2.4	Assumptions	16
2.5	Limitations and Compensating Controls	16
2.6	Model Outputs	17
3.	MODEL PERFORMANCE AND TESTING	20
3.1	Model Specification Testing	20
3.2	Sensitivity Analysis	25
3.3	Outcome Analysis / Backtesting	28
3.4	Benchmarking Analysis	30
4.	MODEL MANAGEMENT AND GOVERNANCE	31
4.1	Ongoing Monitoring Plan	31
4.2	Security and Control	31
5.	REFERENCES	33

1. EXECUTIVE SUMMARY

1.1 Model Purpose and Intended Use

The OH (Over Head) Conductor Model is a Probability of Ignition (POI) Sub-Model developed by SCE (Southern California Edison). The wildfire risk associated with SCE's OH Conductor model is further measured in additional sub-models, i.e., EFF (Equipment / Facility Failure) and CFO (Contact with Foreign Object). At SCE, models are developed using ML (Machine Learning) algorithms for each asset, i.e., OH Conductor, OH Switches, etc., and at each contact type level like animal, balloon, etc., as the drivers vary by asset and contact type. The OH Conductor model is refreshed annually and used to predict the probability of failure (POF) for distribution primary overhead conductors.

The calibrated outputs of the OH Conductor model—i.e., failure events—are broadly used by four categories of programs described below:

4. The Inspections and Remediations programs, which considers POI as an element in prioritization and scoping.
5. The output POI from the OH Conductor CFO Vegetation sub-model is used as one of its inputs into the Tree Risk Index.
6. Overhead Conductor Program (OCP) also uses the OH Conductor model POF to identify which conductors are more likely to experience a wire down event and inform bare wire replacement decisions.
7. Risk analyses via SCE's MARS Framework.

1.2 Model Description Summary

The OH Conductor model predicts the probability of failure of distribution primary overhead conductors using Random Forest—a Machine Learning technique—with two sub-models, i.e., EFF Conductor and CFO.

- **EFF Conductor sub-model:**

The EFF Conductor sub-model is a binary classification model. The EFF Conductor component predicts the probability of a conductor igniting a spark due to equipment failure.

- **CFO sub-model:**

The CFO sub-model is a multi-classification model. The CFO component predicts the probability of a conductor producing a spark as a result of contact with a particular type of foreign object, i.e., animal, vegetation, balloon, vehicle, unknown, and others. This multi-classification model is one approach to determine different failure probabilities for each sub-driver.

Both EFF Conductor and CFO sub-models use a few different variables to produce their failure targets but most of the variables are shared. Some of the common features used by both models are available conductor attributes and condition data (i.e., age, voltage, etc.) and other conductor and environmental attributes (i.e., historical wind, number of customers, etc.). In addition, EFF uses the splice information on the segment² to predict the probability of failure by the conductor whereas CFO uses animal incidents, car fatalities, and vegetation information to predict the probability of failure due to contact with its sub-drivers.

The model is implemented in R programming using the library h2o and is connected to databases such as Net9, SAP, WRF, ADS Weather, etc. The model is run once a year manually by the Data Science and Asset Analytics team. The model is calibrated every year with the full historical outage data.

Cross-references: Please refer to Section 2.1 for more information about the inputs used by the OH Conductor model along with data processing details.

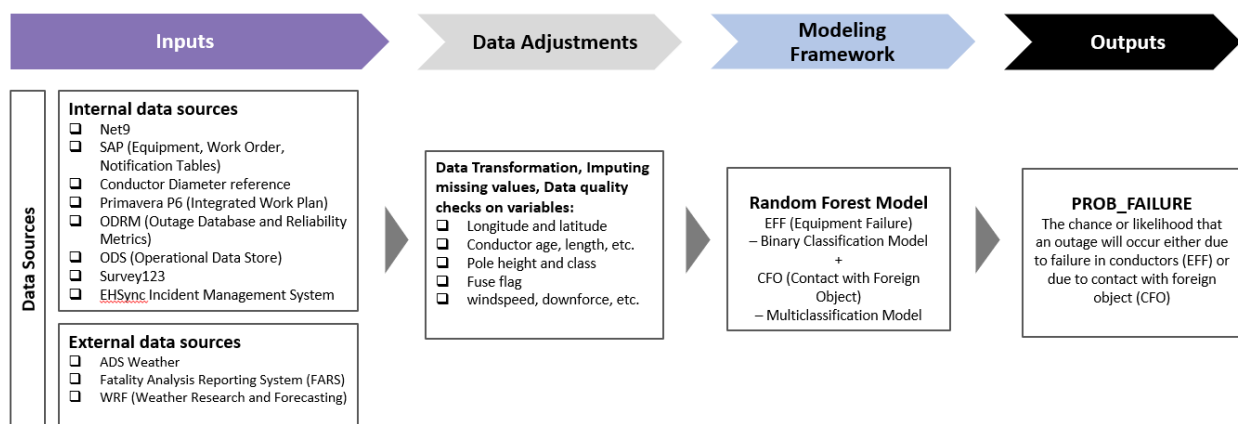


Figure 4: OH Conductor model framework

The OH Conductor model uses the Random Forest methodology for both EFF Conductor and CFO sub-models. Since the Random Forest methodology can perform both classification and regression tasks, it is considered a good choice

² Segment represents the span (conductor) between two structures with equipment installed on it. There could be structures in between that have no equipment installed but physically support the conductor. These structures without equipment are not considered while defining a segment.

for the OH Conductor model as the prediction is a classified event, i.e., failure. This methodology predicts output with high accuracy, runs efficiently on large datasets, and maintains accuracy with minimal adjustments for missing values and data treatments.

1.3 Model Risk Rating

There is no defined mechanism of identifying model risk rating at SCE, however certain factors—like frequency of risk events and use case—are considered while flagging model risk. Based on the Wildfire Mitigation Plan quarterly report, the frequency of outages in a year (from both EFF Conductor and CFO) averages around 3500 which is relatively high compared to other sub-drivers. Figure 2 provides a snapshot of the count of outages over the years by the causes captured in the OH conductor model. Also, the output of this model is considered important as it drives the strategy of several programs which are discussed in section 1.1. Hence, the OH Conductor model is deemed to be a high risk model.

Table 7.1: Key recent and projected drivers of risk events			Number of risk events																Projected risk events							
Risk Event category	Cause category	Sub-cause category	2015	2016	2017	2018	2019	2020	2020	Q3	Q4	2020	2021	2021	Q3	Q4	2021	2022	Q2	Q3	Q4	Q1	2022	2023	Q3	Q4
Outage - Distribution	17. Contact from object - Distribution	Veg. contact- Distribution	395	557	609	416	527	104	70	25	111	93	20	33	174	88	2	28	93	95	0	26	73			
		Animal contact- Distribution	655	598	622	648	686	122	202	169	163	78	169	143	103	56	159	126	97	58	156	115	93			
		Balloon contact- Distribution	758	785	911	975	776	178	348	272	191	245	436	246	166	201	375	232	172	232	360	225	168			
		Vehicle contact- Distribution	508	586	528	647	517	116	113	153	132	144	128	146	142	129	132	135	135	138	130	130	132			
		Other contact from object - Distribution	869	393	289	369	449	44	28	35	42	66	75	115	129	29	86	109	113	108	85	108	112			
	18. Equipment / facility failure - Distribution	Conductor damage or failure — Distribution	463	594	654	713	1116	206	144	211	252	276	109	133	319	228	235	209	296	294	229	204	288			

Figure 5: Key recent and projected risk events due to sub-drivers captured in the OH Conductor Model from SCE Q1 Quarterly Data Report, Table 7.1

Cross-references: Refer to link [RF 1: SCE’s WMP 2022 Q1 Quarterly Data Report submission

] in Section 5 for SCE’s Wildfire Mitigation Plan Q1 2022 Quarterly Data Report submission.

1.4 Model Dependency and Interconnectivity

The OH Conductor model is an “Ignition Likelihood” model which uses the inputs from the ADS (Atmospheric Data Solutions) and Weather Research Forecasting (WRF) modeling output along with other data sources to calculate the probability of ignition.

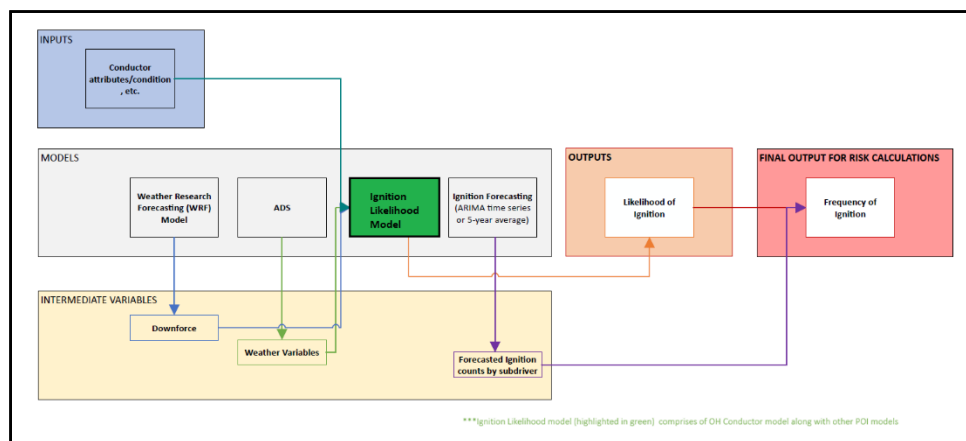


Figure 3: Model Interconnectivity Schema

OH Conductor model uses the ADS model as one of its data sources to get the inputs for the weather variables. ADS’s Next Generation Weather Modeling System (NGWMS) provides an extensive upgrade to SCE’s in-house weather modeling capabilities and enhances SCE’s ability to make more targeted PSPS decisions. The ADS model generates 10 years of hourly weather data between 2010 and 2019. That information is then processed and aggregated to calculate statistical measures, like mean and standard deviation, of wind, humidity, rain, snow, etc. These are used as locational measures and are matched to the conductors through the data engineering step of performing a spatial join to the nearest grid by their latitude and longitude coordinates.

The Weather Research and Forecasting (WRF) Model was run considering 2005 through 2013 data at a 2km-by-2km resolution grid across the entire SCE territory. The OH Conductor model uses WRF as a data source to develop the downforce information through a feature engineering process. Downforce is the perpendicular force applied on the wires due to wind which is termed as the wind factor. The hourly data is aggregated to calculate various statistical measures, like mean, standard deviation, skew and kurtosis, of downforce.

The output data from the OH Conductor model, i.e., POI, is used by four categories of programs, further discussed in Section 1.1, to drive their strategic decisions.

1.5 Model Assumptions

The business assumptions and model assumptions for both the EFF Conductor and CFO sub-models are summarized below:

6. There is no change in the OH Conductor technical specification over time.
7. The contact types that can cause a spark will remain the same throughout the prediction period.
8. The calibration methodology assumes that fires are a subset of failures.
9. The model is designed to work in both base weather and extreme weather conditions.
10. The feature variables in the dataset should have some actual values so that the classifier model can predict accurate results.
11. The predictions from each tree must have very low correlations.

A detailed explanation of these assumptions is available in Section 2.4.

1.6 Model Limitations

The model limitations for both the EFF Conductor and CFO sub-models are summarized below:

5. Unavailability of linear/non-linear representation in the form of intuitive equation or correlation statistic.
6. Time consumption for model execution is high.
7. Resource utilization in terms of system capacity and higher configuration for model execution is high.
8. Model accuracy may reduce if the dataset experiences covariate shift.

A detailed explanation of these data limitations is available in Section 2.5.

1.7 Overall Model Performance Assessment

The machine learning model used to build the OH Conductor model is the Random Forest algorithm. The model's overall performance is determined by the Area Under the ROC Curve (AUC) value and Confusion Matrix results.

The performance of the OH Conductor model was evaluated on test data with full historical outage information until 2022.

- The AUC values for the EFF Conductor and CFO sub-models are 0.9258 and 0.9453 respectively.
- Confusion matrix results capture the accuracy rate as 94.92% and 87.09% for EFF Conductor and CFO sub-models respectively.

The above metrics were derived by re-running the model as of Dec 2022 to capture an exhaustive set of statistical results for documentation purposes.

1.8 Contingency Plan for Vendor Model

A contingency plan is not applicable for this model as it is an inhouse SCE model. This is not a vendor model.

2. MODEL FRAMEWORK AND THEORY

The EFF Conductor sub-model is a binary classification model pertaining to equipment failures, whereas the CFO sub-model is a multiclassification model pertaining to various contacts with foreign object, i.e., animal, vegetation, balloon, vehicle, unknown, and others. Both sub-models of the OH conductor model employ a random forest algorithm to predict the likelihood of a segment experiencing an outage that can result in an ignition event. The random forest approach was chosen for the classification task over other modeling approaches—such as logistic regression, gradient boosting, etc.—because it predicts output with high accuracy, runs efficiently on large datasets, and maintains accuracy with minimal adjustments for missing values and data treatments.

2.1 Model Inputs and Data Quality

Data Sources

This model refers to multiple internal and external data sources. The internal data sources used by the model are:

- **Net9** is built on GESmallWorld data and MAP3D data. GESmallWorld contains all the asset attributes of conductors along with the connectivity of structures and segments. MAP3D is used for geospatial display, and it contains the geospatial attributes of the assets. Net9 is used to get conductor related features like conductor type, conductor size, conductor length, conductor material, etc., as inputs to the OH Conductor model.
- **SAP** houses circuit³, structure, and equipment characteristics. It contains latitude and longitude information of the assets which is used to determine the location of the segment by considering the midpoint between all the structures associated with the segment. SAP also provides features like base and height of the pole which are consumed by the model.
- **Conductor Diameter reference** is a flat file that is produced manually by engineering judgment. It contains the diameter of the conductor for each conductor size and material pairing.
- **Primavera P6 (Web Integrated Work Plan) / SAP (Work Order & Notifications)** track the planned and completed work for covered conductor installations. The statuses are tracked at the structure level which enables the calculation of the proportion of the segment that has covered conductor installed.
- **OMS** refers to Outage Management System which contains information about switching operations.
- **ODRM** refers to Outage Database and Reliability Metrics. It contains the detail for all historical outages. This information is used in conjunction with data from ODS (Operational Data Store) containing information about devices, like active underground and overhead switches, which is used to identify locations impacted by an outage.
- Outage code file (“Outage code Xref.xlsx”) contains the list of all outage cause codes and their mapping to different failure buckets. This is used to determine relevant outages to consider for both the EFF Conductor and CFO sub-models. It is refreshed manually when sub-driver mappings are updated by reviewing each outage cause code assignment.
- **Wire Down Database** contains detail for historical wire downs and their associated triggering event, or sub-driver. This is used to identify locations impacted by relevant wire down events.
- **Survey123** houses vegetation information, like the tree inventory across SCE’s territory, and splice information, like compression splices, automatic splices and preform splices. Features derived from the vegetation information like tree density, tree proximity, etc. are used only by the CFO sub-model and the Splice information is only used by the EFF Conductor sub-model.
- **EHSync** Incident Management System contains animal contact incident reports of deaths, injuries, etc. The animal incidents within a grid are derived from this source and mapped to the respective segment. This data is only used by the CFO sub-model.

³ Circuit comprises a collection of segments that altogether form a path for electrical current flowing from the power source (including but not limited to a substation) to another power source or circuit endpoint.

- **Atmospheric Corrosivity shape file** is used to fetch the atmospheric corrosivity intensity for each segment.

The external data sources used by the model are:

- **ADS** (Atmospheric Data Solutions) model provides 10 years of hourly gridded weather data from 2010-2019. These are aggregated to individual locational measures and matched to the conductor segments through spatial join to the nearest grid by the latitude and longitude as a part of the data engineering step.
- **Fatality Analysis Reporting System (FARS)** is a national database for car fatality information that tracks the location (latitude/longitude) and time of accidents. This data is only used by the CFO sub-model.
- **Weather Research and Forecasting Model (WRF)** is an open-source external model that provides meteorological data at a highly granular level. The downforce, or perpendicular force applied on the wires, is derived from an analysis using u-component of wind, v-component of wind, wind direction, and wind speed at 10m from the WRF data.

Cross-references: Please refer to link [RF] in Section 0 for NHTSA source.

Quality Checks

SCE has internal data management teams for ensuring data quality, including EAD (Enterprise Asset Data) and Master Data. They work on processing asset data corrections (E2 notifications) in SAP and fixing largely known data issues like missing or erroneous latitude and longitude information for assets in the territory. Some of the data quality checks that are performed in the OH Conductor model to ensure accuracy, validity, integrity, and consistency are provided below. Quality checks (QC) are coded in R or incorporated into the data gathering process.

The QC steps performed by automated R code are as follows:

- Duplicate values that are identified in CKT_NAME and Circuit_Sub_Information variables are removed to maintain consistency in data by considering only the distinct values.
- The FLOC data file that is used to fetch the latitude and longitude values of the FLOCs from Net9 would have duplicates entries in it. Those duplicates in latitude and longitude values are removed to improve data quality by considering the distinct values in data processing.
- Splice data file received from DOCI contains the splice information for each segment. This file will contain duplicate entries for the same segment based on the updates made in a different time period. The duplicate entries are removed by selecting the most recent inspection to improve data relevance. Also, splice entries without FLOC details are removed before loading the data into the model to improve data quality.
- Vegetation data that is fetched from the tree inventory file might contain duplicate entries. These duplicates are removed by Tree_ID to improve its quality. Additionally, the tree inventory data is filtered for relevant, recent data marked by record status fields.
- ODRM provides the information about all the outages encountered by SCE. Additionally, the Wire Down Database provides historical wire down events across the service territory. Only the relevant information like failures specific to conductor (EFF sub-model) and contact from foreign objects (CFO sub-model) are loaded into the respective model. All the other non-relevant information i.e., for equipment like switches, are excluded.

The manual QC steps are as follows:

- ADS weather data is validated against actual weather observations.
- Asset data obtained from SAP is validated and updated through inspections and other programs.
- Routine tree data and hazard tree data from Survey123 are validated by QC and field verifications.

Data Sampling

Since this is a classification model to predict conductor and contact failures, there are no sampling strategies used in the model other than the random split strategy to bifurcate the train and test data. The dataset used for the model is randomly divided to have 80% in train data and 20% in test data.

Data Cleansing and Transformation

The data cleansing and transformation activities that are incorporated in the R scripts as a part of automation to ensure the completeness of data used for model training and estimation are provided below.

- Missing data for the below specified numeric variables are handled by imputing the mean value on the shared circuit.
 - LAT_UDF
 - LONG_UDF
 - Conductor_Length
 - Conductor_AGE_UDF
 - Conductor_Diameter_UDF
 - ICA_GEN_UDF
 - ICA_LOAD_UDF
 - INSTALLED_YEAR_UDF
- The conductor's replacement information is not updated in the database which makes it difficult to track the age of conductor. Hence there are high possibilities to encounter missing data when calculating conductor age based on service date. To fill in missing values, mean imputation based on structure age from SAP is performed.
- Data consistency is ensured by correcting formatting issues in date variables. e.g., INSTALLED_DATE_UDF variable can have different formats of data; format is corrected in R program code for data formats to be consistent.
- The data issues that are identified with respect to conductor size (CONDUCTOR_SIZE_UDF) are updated using the user defined size features.

Data Assumptions

The accuracy of the predicted results is dependent on the accuracy of the data used to build the predictive models. The data assumptions follow:

5. The assumptions used for the data imputation utilized SCE's Distribution Design Standard (DDS), engineering judgment, and manufacturer data.
6. The target labeling process used to label the failures and non-failures as '1' and '0' is considered accurate. This is performed by comparing the outage and wire down count in a segment against the mean value of outages and wire downs across all segments. If outage and wire down count in a segment is greater than the mean value, the '1' is assigned to represent failure. Else '0' is assigned for non-failures.
7. Input data with respect to asset, weather, and engineering information are assumed to be stable and will not change over time until the subsequent data refresh. Example: If there is an update in the structure information specific to an asset, that updated information will be reflected only in the subsequent data refresh. So, it is assumed that the updated structure information is not drastically different from the previous information which might alter the model outcomes.

Data Limitations

The following are data limitations across internal and external data sources:

3. Some of the data used by the model faces accuracy issues in terms of consistency in data labelling or missing values which might impact model prediction power.
 - Data labelling issues may be due to manual errors during data entry. E.g., conductor type information is fed manually into the system. While updating the name, different labels for the same conductor type might be used in different data entries which affects the consistency of the data. Hence these consistency issues in data need to be addressed before using them in the model.
 - Missing data for a specific feature (predictive variable) might be due to unavailability of data. E.g., for the Conductor_AGE feature, conductor replacement information is not updated in the

database, which makes it difficult to track the age of the conductors. To overcome this issue, pole age information fetched from SAP data source is used as a proxy to estimate the conductor age which is further used in model processing. Other missing values for a conductor segment are filled using imputations by cross-referencing other fields or other data sources to mitigate the risk arising from missing predictors.

4. With respect to Failure targets, the starting location of the outage is not tracked every time. So, the outages associated to a segment needs to be mapped based on approximations specified in the Unknown Outage Mapping process below.

Mapping Unknown Outages with DOTS2.2 Method

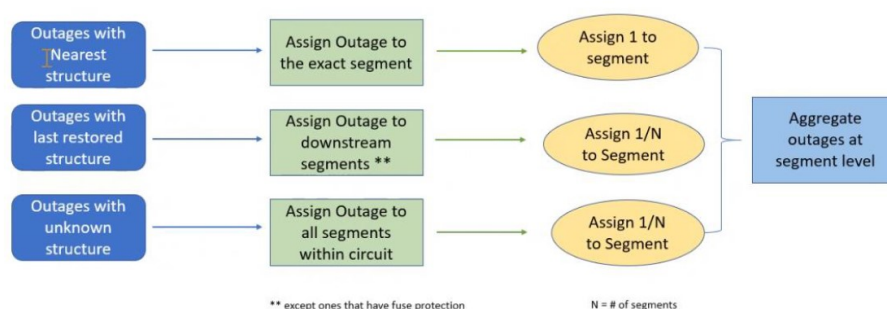


Figure 4: Unknown outage mapping process

- After assigning values to the segment, the mean of these non-zero assigned values is calculated and considered as the threshold to classify an event. The values assigned to a segment is compared with this threshold and the higher ones are identified as failures.

Independent variables

The OH Conductor model uses multiple variables/features. Most of the features are commonly used in both EFF Conductor and CFO sub-models. Some of these features are created based on engineering knowledge (like short circuit duty for each conductor) and some are selected based on expert advice like the logic to calculate tree density. A subset of the independent variables (inclusive of EFF Conductor and CFO sub-models) used in the OH Conductor model along with its data source and description, is provided below.

Feature	Data Source	Description
Conductor_AGE_UDF	Net9, SAP	Conductor age, calculated by referring to IN_SERVICE_DATE from GE Smallworld and imputing missing values using age of structures on the circuit
EQ_PoleHeight	SAP	Average pole height of all structures associated to the FLOC
EQ_PoleClass	SAP	Mode pole class (field in Equipment table: _BIC_ZCAE_P015; values = 1-6, H1-6) of all structures associated to the FLOC
SCD_Seg	Net9, SAP (Circuit_Sub_Stats)	$SCD_S1 = X1SubtoSeg / ((Circuit_Voltage)^2 / 100)$
Delta_SCD	Net9, SAP (Circuit_Sub_Stats)	Difference between SCD_Seg and SCD_Thresholds
LENGTH_SEG_CAL_TOTAL	Net9	Calculated Total Length of segment from the Substation, inclusive of current feature length

Feature	Data Source	Description
FUSE_FLAG	Net9	Indicator if conductor has fuse protection or not
Log_WindForce	SPIDA, Weather Research and Forecasting Model (WRF)	Log conversion of downforce
SECTION_FLAG	Net9	Indicator if conductor has RCS/RAR protection or not
skew_of_sum_of_seg_downforce	SPIDA, Weather Research and Forecasting Model (WRF)	Skewness of sum of segment downforce
CONDUCTOR_SIZE_UDF	Net9	Conductor size
DOWNSTREAM_CUST	Net9	Downstream Total Customer count found at the end of the conductor, exclusive of current conductor customer count. When no serial number match for a transformer, default to 1 customer
SCD_Ratio	Net9, SAP (Circuit_Sub_Stats)	SCD_Seg / SCD_Thresholds
max_of_sum_of_seg_downforce	SPIDA, Weather Research and Forecasting Model (WRF)	Maximum value of sum of segment downforce
DOWNSTREAM_KVA	Net9	Downstream Total KVA count found at the end of conductor, exclusive of the current conductor kva

In addition to the data utilized above, 10 years of hourly data fetched from ADS Weather model is processed and aggregated to calculate statistical measures like mean, max, and standard deviation for wind, temperature, water vapor, turbulence kinetic energy, humidity, rain, and snow. Asset information in a segment is fetched from Net9 with which the various inputs from other sources are combined using spatial join. Pole information like the pole base, pole height on a circuit is obtained from SAP to get a representative measure for height of attachment. The consolidated downforce information is processed from SPIDA/Weather WRF and aggregated to calculate statistical measures like mean, standard deviation, skew and kurtosis of segment downforce.

Splice information from DOCI is used only by the EFF sub-model to understand the joints or links connecting the various assets in a segment. The vehicle accident information fetched from FARS is not specific to grids in SCE territory. Based on the vehicle accident information with latitude and longitude, the KDE⁴ of there being a vehicular accident at the segment location is calculated using CrimeStat. Similarly, animal contact incidents are used to calculate the KDE of there being an avian accident at the segment. Vegetation features are further described under Model Changes in Section 2.6.

Cross-References: Refer to link [RF 3] in Section 0 for CrimeStat data.

Dependent Variable

In a typical classification risk model, defining the dependent variable is key for both model development and model performance assessment.

- EFF Conductor sub-model is a binary classification model, where the target variable represents whether the segment experienced a higher than average number of outages related to conductor failure.
- CFO sub-model is a multi-classification model, and it has 6 sub-categories which identifies the contact from several objects like animal, balloon, vehicle, vegetation, unknown, and other. The target variable represents the contact type if the segment experienced a higher than average number of outages related to contact.

The final output of the model is PROB_FAILURE, representing the chance or likelihood that an outage will occur either due to failure in conductors or due to contact from foreign objects. The CFO sub-model produces six failure probabilities (one for each sub-category). The h2o.predict (level = 0.05, type = 'response') function is used to specify the desired output (PROB_FAILURE) in probability values, rather than binary values. The probability value ranges

⁴ Kernel Density Estimation (KDE) is the application of kernel smoothing for probability density estimation.

from 0 to 1 where '0' represents the least likelihood of an outage and '1' represents high chance of an outage.

2.2 Methodology

SCE utilizes machine learning to identify patterns that may lead to failures causing sparks from conductors and uses the trained model to predict POIs at the asset level. The OH conductor model predicts the POI arising from two ignition drivers viz., asset and contact type separately. The POI model with asset as the driver is categorized as sub-model EFF whereas contact type is categorized as sub-model CFO. Both the EFF and CFO components of the OH Conductor model employ a random forest algorithm to predict failure events. The Random Forest approach can predict outputs with high accuracy, run efficiently for large datasets, and maintain accuracy with minimal adjustments for missing values and data treatments.

A random forest is a supervised machine learning algorithm that is constructed from many decision trees. It can be used to solve both classification and regression problems. This approach utilizes ensemble learning, which is a technique that combines many classifiers to achieve greater predictive accuracy than that of a single classifier. A decision tree is a decision support technique that forms a tree-like structure. It consists of three components: decision nodes, leaf nodes, and a root node. The following diagram shows the three types of nodes in a decision tree.

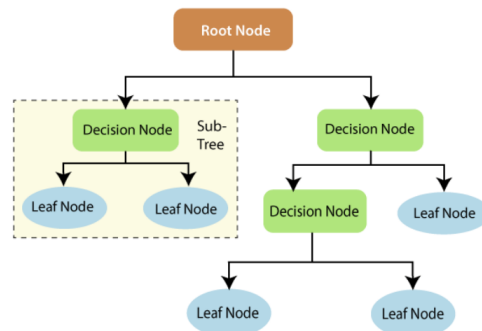


Figure 6: Decision Tree Structure

A decision tree algorithm divides observations of a dataset into branches, which further segregate into other branches. This sequence continues until a leaf node is attained. A leaf node cannot be segregated further. In more detail, the root node is the base of a decision tree, where the first of a chain of decisions is made. A branch is the connection path between nodes. A node is a potential splitting point on a tree. Decision nodes provide a link to the leaves. On the other hand, leaves, also known as terminal nodes, are the ends of a tree, representing the resulting classification or value for the sample.

The 'forest' generated by the random forest algorithm is trained through bagging, also known as bootstrap aggregating. Bagging is an ensemble meta-algorithm that fits multiple models on different subsets of a training dataset and then combines the predictions from all models. The diagram below shows a simple random forest classifier.

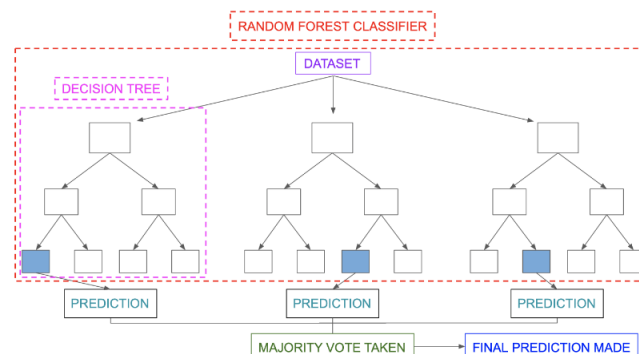


Figure 7: Structure of Random Forest Classifier model

The selection of the final output follows a majority-voting system. In this classification model case, the output chosen by a majority of the decision trees becomes the final output of the random forest system. The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.

Train test split is a model validation procedure that simulates how a model would perform on new/unseen data. Figure 7 shows the logic of dividing the dataset into train data and test data. First the data is consolidated and prepared for train test split. Then the historical input datasets are split into a training dataset (80%) and testing dataset (20%) based on simple random sampling strategy with a split ratio of 4:1 without replacement. Simple random sampling is a technique that ensures each observation has an equal likelihood of being selected for a set. It is a fair strategy as it helps in avoiding any bias involved compared to other modeling techniques and it has no restrictions on the sample size which makes it suitable to handle vastly sized input data. The predictive algorithm is developed using the training dataset and is built by looking at the interactions between all the features to find patterns and predict the likelihood of equipment failure.

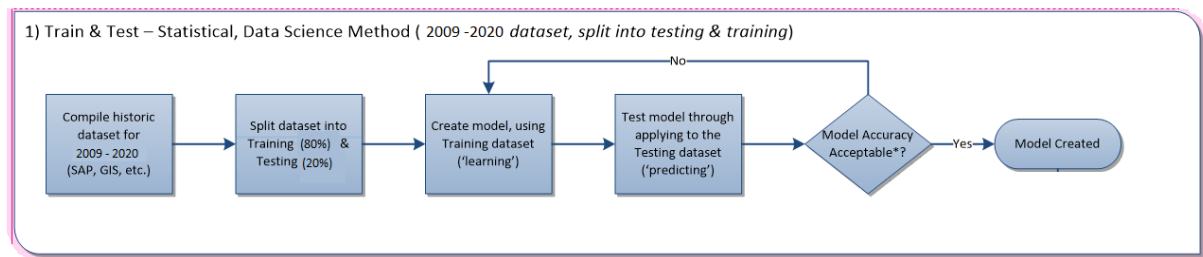


Figure 8: Train and Test data split logic

In the next step, the algorithm is tested on the ‘testing’ dataset. The model is run on the test dataset to make a prediction of a failure or success. Then an internal validation of the model is conducted by comparing the predicted results to the actual results which indicates the predictive capabilities of the features as well as the model. AUC is the metric used to assess the performance of the model on test data.

Area Under the Curve (AUC) – Area Under the Receiver Operating Characteristic (ROC) Curve is a measure to estimate the model discriminatory power (degree of separability) for the binary classification problem. The ROC curve plots True Positive Rate against different thresholds with False Positive Rate (FPR) or True Negative Rate (TNR). The higher the AUC, the better the model is at predicting True Negatives (non-events) and True Positives (events).

Hyperparameter Tuning:

Hyperparameters are parameters that are explicitly defined by the user to control the learning process. The process of selecting the optimal hyperparameters to use is known as hyperparameter tuning, and the tuning process to achieve the best-defined performance statistic is known as hyperparameter optimization. Cartesian Grid search and Random Grid search are the most widely used strategies for hyperparameter optimization.

- In the Cartesian grid search approach, the machine learning model is evaluated for a range of hyperparameter values, and it searches for the best set of hyperparameters from a grid of hyperparameters values. The disadvantage of grid search model is that it will go through all the intermediate combinations of hyperparameters which increases the time consumed by grid search computations.
- In the random grid search approach, the machine learning model is evaluated for a range of hyperparameter values like that in Cartesian Grid Search approach. However, search criteria parameters are added to control the type and extent of the search, and it moves randomly within the grid to find the best set of hyperparameters to achieve maximum performance in terms of the metric defined by the user. As search criteria, the user can set a maximum runtime for the grid, a maximum number of models to create, or metric-based automatic early stopping. If many of these requirements are supplied, the algorithm will end when the first of the criteria is met. This approach reduces the time taken for computation thereby solves the drawbacks of the cartesian grid search approach.

Both the EFF Conductor and CFO sub-models use Random Grid Search method for Hyperparameter tuning. The

reference literature link to understand the efficiency between Cartesian Grid search and Random Grid search is provided below. The criterion used for the hyperparameter tuning in OH Conductor Model are:

- **ntrees:** Total number of trees used in the random forest. For tuning this parameter, the EFF Conductor sub-model uses a range of values between 100 and 500 with an increment of 50, whereas the CFO sub-model uses values between 100 and 500 with an increment of 100, and then uses 1000 as the end value.
- **mtries:** Total number of predictors/variables that will be randomly selected in each node to search for the best split. This parameter is varied by using different percentages of the total number of independent variables in both models. For both models, the various percentages taken into account are 5%, 15%, 25%, 33.3%, and 40%.
- **max_depth:** Specifies the maximum size of the sample data drawn for training each tree. A higher value for this feature will make the model more complex and can lead to the issue of overfitting the training data. For the max_depth parameter tuning, the range of values used for the EFF Conductor sub-model are set between 25 and 50 with an increment of 5, and for the CFO sub-model, the range of values used are between 20 and 40 with an increase of 10.
- **min_rows:** This parameter defines the minimum number of observations required for a leaf to split. This parameter is tuned using the values 1, 2, 7, 10, and 15 for the CFO sub-model. For the EFF Conductor sub-model, this parameter is set to a default value of 1.
- **Sample_rate:** The size of the sample data drawn for training each tree. The scale goes from 0 to 1.0. For both models, this parameter is tweaked with values of 0.5, 0.632, 0.7, and 0.8.

Random Grid Search method uses the below specified stopping criterion in both the EFF Conductor and CFO sub-models to stop the random grid search. The conditions are provided below.

- **stopping_tolerance = 0.005 for EFF Conductor and 0.01 for CFO**
This will stop the random search if the tolerance level reaches 0.005 for EFF Conductor and 0.01 for CFO.
- **stopping_rounds = 15**
This will stop the random search if none of the last 15 models managed to have 0.5% improvement for EFF and 1% improvement for CFO compared to best model identified before that.
- **max_runtime_secs = 3600**
This is used to define the maximum number of seconds allowed for the search. The random search will stop if the search continues to find improvements after 30 min.
- **stopping_metric = AUC**
This defines the performance metric-based condition to stop the search. The random grid search will stop when the model's AUC value doesn't improve by 0.5% for the EFF sub-model and 1% for the CFO sub-model.

Once the random search completes, the grid object containing the list of models is queried, and models are sorted by a performance metric defined by the user. The model with better performance is chosen as the best model and it is validated on the test data.

Cross-references: Refer to [RF 4] in Section 5 to understand the efficiency between Cartesian Grid search and Random Grid search.

2.3 Suitability

During development of the model in 2019, Logistic Regression was used to construct the EFF Conductor and CFO sub-models. Then the other modelling approaches like GBM and Random Forest were tested. The test results proved that the Random Forest methodology fits well for the sub-models as it exhibited higher AUC than other approaches. AUC comparison of these three approaches is specified in Section 3.4.

Random Forest methodology can be used to solve both classification as well as regression problems and it can handle both categorical and continuous variables. One of the main advantages of the Random Forest methodology

is that it maintains accuracy with minimal adjustments for missing values and data treatments. It also runs efficiently on large datasets like the set of all OH Distribution Primary Conductors. Theoretically, the Random Forest methodology exhibits a higher level of accuracy and stability and handles non-linear parameters more efficiently than other approaches. Additionally, hyperparameter optimization prevents the issue with random forests overfitting. Random grid search is used for hyperparameter tuning; it controls the maximum depth of the sample data drawn for training each tree and involves stopping criterion which reduces the computation time. Hence, the usage of Random Forest for the OH Conductor model is deemed to be fit.

2.4 Assumptions

The key business assumptions that were considered during model development are specified below:

BA 01: There is no change in OH Conductor technical specification over time. The model assumes the type of OH conductors used in the model building process have the same characteristics in terms of build and quality. For example: If the conductor type is aluminum, conductor size is deemed to be '4'; and if the conductor type is copper, the conductor size is deemed to be '6'. These kinds of technical specifications are expected to remain the same over time.

BA 02: The contact types that can cause a spark will remain the same throughout the prediction period. The six sub-drivers included in the CFO component are Animal, Balloon, Vegetation, Vehicle, Unknown, and Other. The list of specific sub-drivers might not be exhaustive, but it is the best representation of the contact types that are main drivers of ignitions based on SME judgment. It is assumed that there will not be any requirement to add a new sub-driver to the existing list of six CFO sub-drivers.

BA 03: The Calibration model assumes that fires are a subset of failures. Outages are the representative failure targets used in place of few ignition events. Outages can potentially spark an ignition, but not all outages will result in a fire. Hence, fire can be treated as a subset of failure.

BA 04: The model is designed to work in both base weather and extreme weather conditions. The weather variables considered by the model are represented as various statistical aggregations like max, mean and standard deviation on wind, wind speed, humidity, rain, and snow. Hence the model results can be used under both base weather and extreme weather conditions.

The functional/model methodology assumptions that were considered during model development are discussed in detail below:

MA 01: The feature variables in the dataset should have some actual values so that the classifier model can predict accurate results. In an ideal scenario, all the variables would not have estimated values and they would instead use actual values. The current model is able to provide accurate results even after using estimates as they are derived through imputation using actual values from other variables. Example: Estimating conductor age based on pole age.

MA 02: The predictions from each tree must have very low correlations. It is difficult to differentiate between a real interaction effect, marginal effects, and just random variations in random forests. Hence, the presence of highly correlated variables in the Random Forest approach will have an impact on its ability to identify strong predictors.

2.5 Limitations and Compensating Controls

The key model limitations that would impact the accuracy and performance of the model are discussed in detail below:

Limitation ID: L01

Limitation Title: Unavailability of linear/non-linear representation in the form of intuitive equation or correlation statistic.

Description: The Random Forest algorithm does not explain any linear or non-linear relationship in the form of an intuitive equation or correlation statistic to enable measurement of the scalability of impact of independent variables on the dependent variable.

Compensating Controls: The Random Forest model is considered a black box as it is difficult to understand the

relationship between independent and dependent variables and how the independent variables influence the predictions. Since black box is a common limitation with most ML algorithms, usage of the model is considered appropriate as it provides better AUC results than other models.

Limitation ID: L02

Limitation Title: Time consumption for model execution is high

Description: Since Random Forest models use a bagging algorithm, they can provide more accurate predictions but slow down the process as they compute data for each decision tree.

Compensating Controls: To overcome the time consumption issues from grid search computations, random grid search is used in the hyperparameter tuning process. Random grid search is a proven technique to reduce the time consumption when testing multiple models with different combinations of hyperparameters by using stopping criterion like tolerance, maximum rounds, maximum run time, and performance improvement thresholds. It moves within the grid in a random fashion to find the best set of hyperparameters to achieve maximum performance in terms of the metric specified, here AUC. Since the model is not executed through computer program automatically at a defined frequency and is instead run only once a year manually, usage of the model is considered appropriate.

Limitation ID: L03

Limitation Title: Resource utilization for model execution is high

Description: Since Random Forest models process many decision trees, they need more resources with respect to system configuration and system capacity to store that data.

Compensating Controls: The resource utilization factor will have a major impact for real time models as they would run more frequently. Since the OH Conductor model is run only once a year with reasonable use cases, the impact of resource utilization is low. Additionally, the usage of random grid search and stopping criterion like tolerance, maximum rounds, maximum run time, and performance improvement thresholds provide more control on the number of recurring instances run to identify the best fit hyperparameters to achieve optimal AUC. Since the model is not executed through computer program automatically at a defined frequency and is instead run only once a year manually, usage of the model is considered appropriate.

Limitation ID: L04

Limitation Title: Model accuracy might reduce if the dataset experiences covariate shift.

Description: Covariate shift is a type of model drift which occurs when the distribution of independent variables changes between the training environment and live/test environment. Since the Random Forest cannot extrapolate (i.e., predict outside the training space), the model performance might decrease if there is covariate shift in the dataset.

Compensating Controls: The covariate shift affects most machine learning models to some degree, as test data is never going to be the same as training data. Detecting and addressing covariate shift is therefore a key step to the machine learning process. The current model is run only once a year along with data refresh. It uses a random sampling mechanism to split the dataset into train (80%) and test (20%) data whenever it is run. The usage of random sampling mechanism is considered to resolve the issue of covariate drift and maintains the accuracy of the model results. Hence the usage of the Random Forest methodology along with the random sampling mechanism to split train/test data is considered appropriate.

2.6 Model Outputs

The OH Conductor model predicts the probability of ignition (POI) arising from asset (conductor) and contact from foreign objects separately. The POI model with asset driver is categorized as EFF and has a single output characterized by a continuous number between 0 and 1. The POI model categorized as CFO contains 6 separate outputs (one for each sub-driver), each with a continuous number bounded by 0 and 1.

The probabilities across different asset failure predictive models cannot be aggregated or compared and hence are calibrated to derive frequencies of ignition. The sum of the resulting frequencies of ignition for a sub-driver equals the total expected ignitions for the specified year.

$$\text{Frequency of Ignition} = \text{Probability of Ignition} \times \frac{\text{Calibrated Targets}}{\sum \text{Probability of Ignition}}$$

where Calibrated Targets = Forecasted Ignitions for that sub-driver

The output from this calibration exhibits the following features:

- Frequency: Each value can be specified as the frequency of fires per year.
- Comparability: The frequencies are comparable against sub-drivers and models.
- Additivity: The frequencies can be added across models to derive the aggregated fire forecast in a year.

This is achieved by forecasting fires by sub-driver and using these forecasts to weight model probabilities. The sum of probabilities from each calibrated model equals the forecast by sub-driver.

Figure 8 provides the calibration steps that are performed using the failure probability results from the OH Conductor model. This methodology followed in the calibration model is provided below:

- E. Aggregate the probability output from each sub-driver model.
- F. Based on the forecast logic specified above, find the forecast results (expected fires) for each sub-driver.
- G. Generate the calibration factor for each sub-driver based on the values calculated in the above steps (B/A).
- H. Multiply each model probability by its calibration factor to arrive at the estimated frequency of fires from each sub-driver.

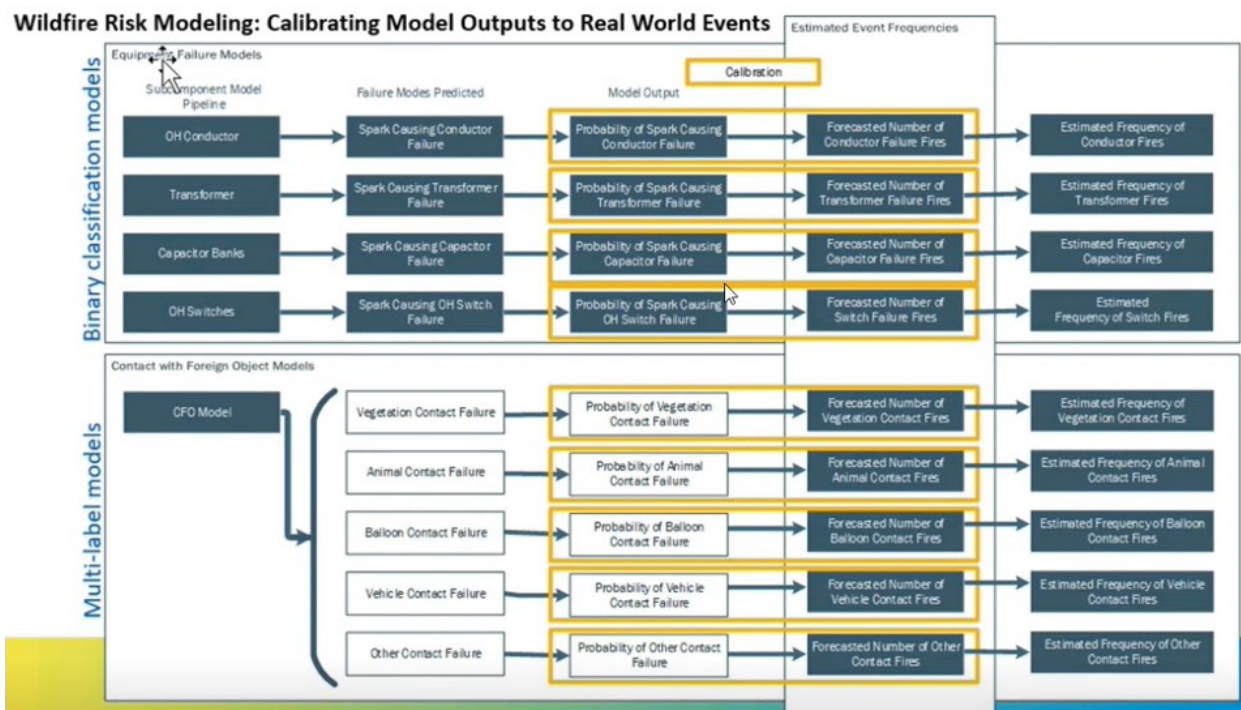


Figure 8: Calibration model schema

This estimated frequency of fires from each sub-driver can be added across the models to derive the aggregated fire forecast in a year.

The calibrated probabilities, frequencies of events, based on the output from OH Conductor model is the data ingested to inform the programs mentioned in Section 1.1.

Model Changes:

Till April 2022, the CFO sub-model was using hazard tree data to inform predictions of probability of contact due to vegetation. This logic was updated in May 2022 along with the annual model refresh to replace the hazard tree data with new features (tree density, tree proximity, tree growth rate, and tree risk rating).

- Tree density represents the total count of existing trees (including hazard and dead and dying trees) within 50 feet of a segment. These trees are mapped to a segment using ArcGIS spatial join.
- Tree proximity represents the minimum allowed clearance distance between the tree to the segment.

- Tree growth rate represents the fastest growth rate, which can be applied to grown-in potential, among all tree species mapped to a segment.
- Tree risk rating represents the most extreme risk rating, which can be applied for fall-in/blow-in potential, among all tree species mapped to a segment.

The impact on the model performance from the change in features can be determined by comparing the AUC of the model before (April 2021) and after (Dec 2022) refresh. Figure 9 and Figure 10 provide the AUC comparison for the EFF Conductor and CFO sub-models before (April 2021) and after (Dec 2022) refresh. The AUC value of the EFF Conductor sub-model was 0.87 before refresh and 0.9258 after refresh. The AUC value of the CFO sub-model was 0.945 before refresh and 0.9453 after refresh.

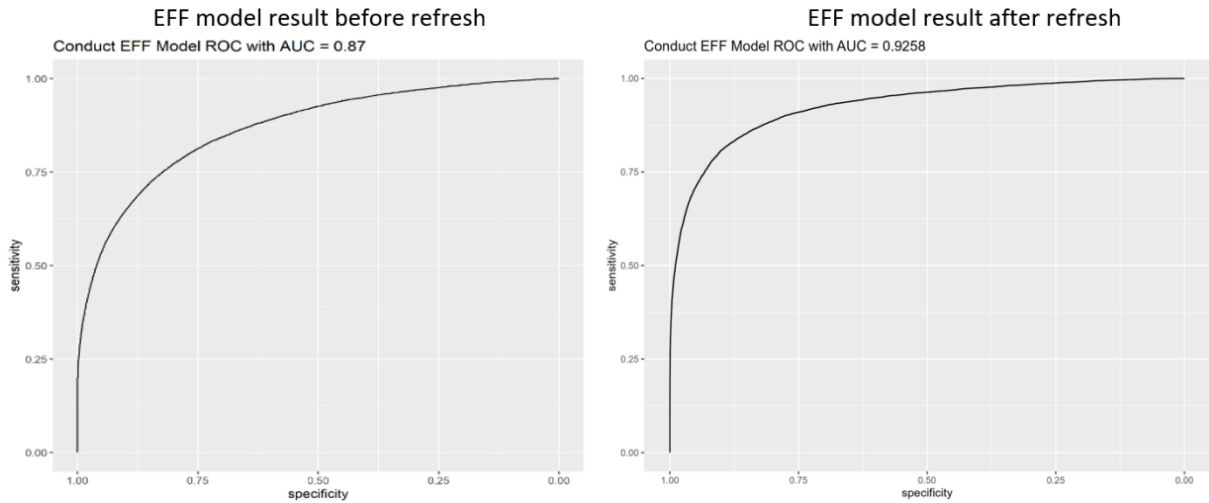


Figure 99: AUC results for EFF Conductor sub-model (before and after refresh)

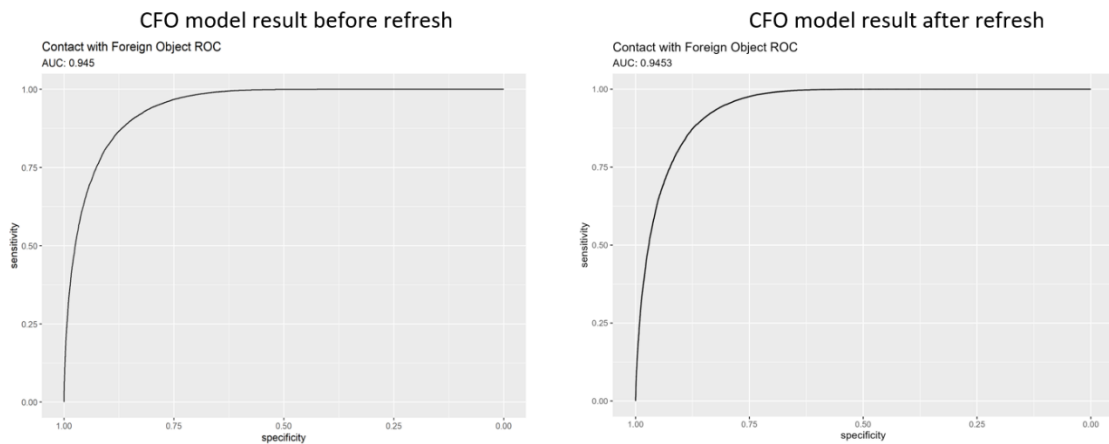


Figure 10: AUC results for CFO sub-model (before and after refresh)

3. MODEL PERFORMANCE AND TESTING

For each machine learning model developed, SCE tries to select the best algorithm based on the model train/test performance, which can be measured by Area Under the Curve (AUC) and other metrics from the Confusion Matrix.

3.1 Model Specification Testing

The model is developed and tested in R programming using library h2o. The model is run once a year manually by Data Science and Asset Analytics team. The model is calibrated every year with the full historical outage data.

The verification of the model implementation is performed by checking the variable importance results which provides the list of features implemented. The performance of the model is validated through the AUC, defined in Section 2.2 and provided in Section 3.3.

The validity and impact of the Model Assumptions, mentioned in Section 2.4, are discussed below:

- The features used in the model are expected to have some actual values so that the model results can be accurate. In an ideal scenario, all the variables would not have estimated values and they would instead use actual values. Some features like conductor age do not have the actual values in all scenarios so values are imputed for this feature with help of information in other variables like pole age. After using these estimates, the data quality is enhanced to support reliability of the current model in terms of improved predictive accuracy.
- Random Forest is considered a strong approach for variable selection in high-dimensional data only when the variables have low correlation. The recursive structure of trees generally enables them to take dependencies into account in a hierarchical manner. However, some variable combinations without clear marginal effects might make the tree algorithm ineffective. To conclude, it is difficult to differentiate between a real interaction effect, marginal effects, and just random variations in random forests. Hence, the presence of highly correlated variables in Random Forest approach will have an impact on its ability to identify strong predictors. Adequate measures are taken to filter out highly correlated features to overcome their impact in predicting the results.

Model Estimation:

The OH Conductor model employs a number of independent variables. Section 2.1 contains a list of the independent variables utilized in this model.

The variable importance test results for the OH Conductor model, Figures 11 and 12 for EFF Conductor and CFO respectively, shows the order of which features provided the most information gain in informing the correct prediction of failure or non-failure. The variable importance features test estimates the relative influence of each variable by calculating whether that variable was chosen to split during the tree building process and how much the squared error (over all trees) improved (or decreased) as a result.

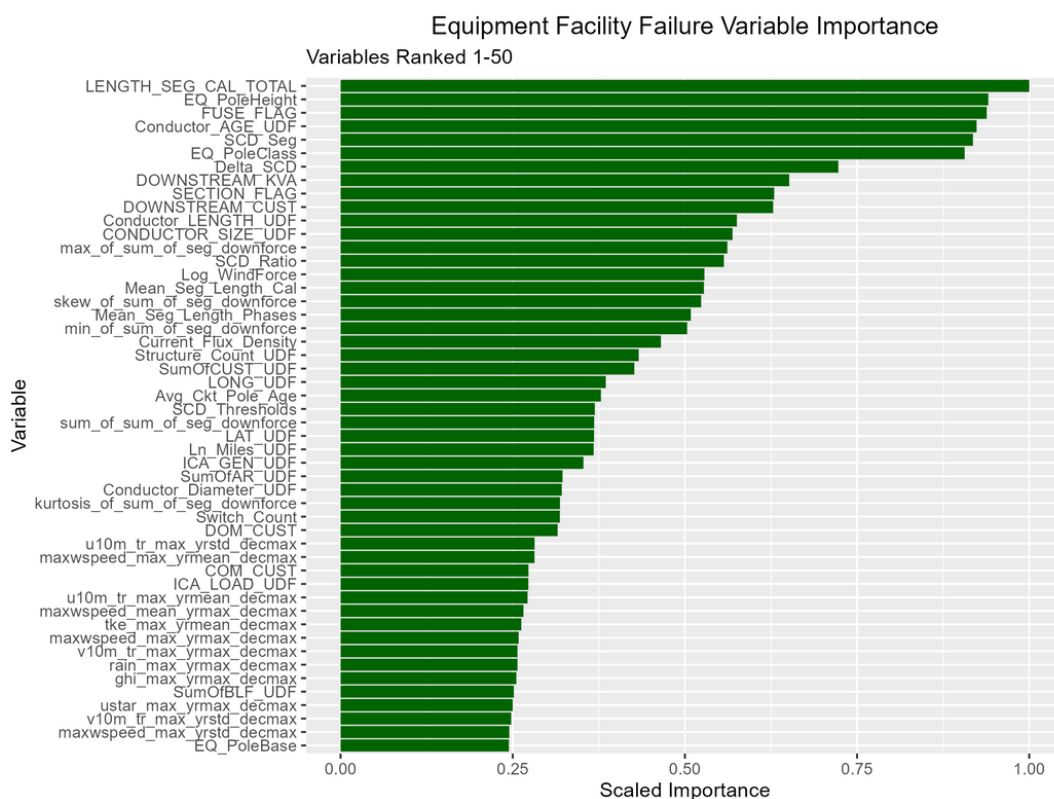


Figure 11: Variable Importance test results for EFF Conductor sub-model

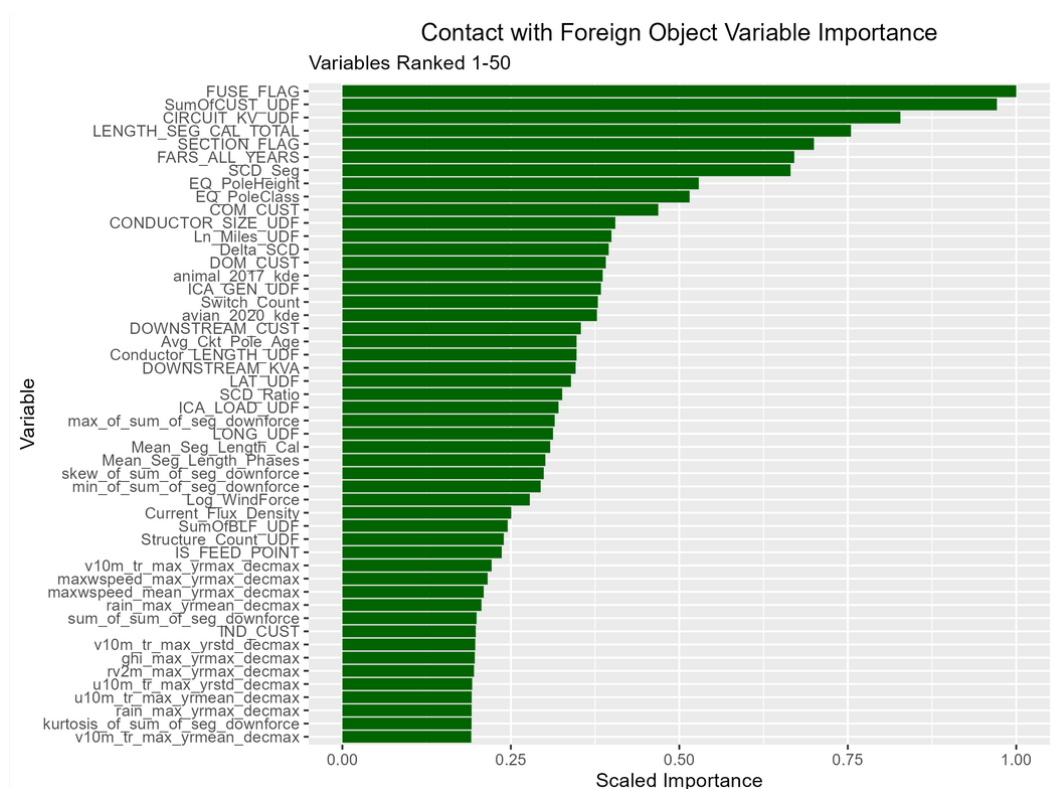


Figure 12: Variable Importance test results for CFO sub-model

The results confirm that the features LENGTH_SEG_CAL_TOTAL, EQ_PoleHeight, and FUSE_FLAG exhibit high importance on the EFF Conductor sub-model output and the features FUSE_FLAG, SumOfCUST_UDF, and

CIRCUIT_KV_UDF exhibit high importance on the CFO sub-model output.

Cross-references: Refer to link [RF 5] in Section 0 for description on the methodology used to perform the Variable Importance for tree-based methods.

Both EFF Conductor and CFO sub-models use the random grid search approach for hyperparameter optimization to select the best set of hyperparameters to achieve maximum performance in terms of AUC as described in Section 2.2. Once the grid search is completed, a list of models with their associated hyperparameter values is obtained for both EFF Conductor and CFO as shown in Figure and Figure . The acquired models are then sorted based on the AUC values for the OH Conductor model. The model with the highest AUC value is regarded the best fitted model. Figure and Figure show the best models obtained for the EFF Conductor and CFO sub-models, respectively. The best models are run on the respective test data, and the AUC metric is used to evaluate the model performance. The AUC is used to estimate the model discriminatory power to predict the results in a binary classification problem. A higher AUC means the model can predict the results accurately. Figure 17 and Figure 18 shows the ROC with AUC for EFF Conductor and CFO sub-models based out of test dataset ran with full historical outage data till 2022 (results derived in Dec 2022 R scripts re-rerun). The AUC values for the EFF Conductor and CFO sub-models are 0.9258 and 0.9453 respectively. The AUC values for sub-drivers Animal, Balloon, Other, Unknown, Vegetation and Vehicle are 0.8902, 0.9131, 0.8481, 0.9639, 0.8861, and 0.8771.

```
Grid ID: Grid_DRF_TRAIN_CONDUCTOR.hex_model_R_1671282333708_1
Used hyper parameters:
- max_depth
- mtries
- ntrees
- sample_rate
Number of models: 6
Number of failed models: 0

Hyper-Parameter Search Summary: ordered by increasing logloss
max_depth mtries ntrees sample_rate
1 35.00000 13.00000 400.00000 0.80000
2 45.00000 13.00000 250.00000 0.70000
3 40.00000 29.00000 300.00000 0.50000
4 30.00000 13.00000 200.00000 0.50000
5 50.00000 35.00000 200.00000 0.50000
6 30.00000 4.00000 250.00000 0.63200

model_ids logloss
1 Grid_DRF_TRAIN_CONDUCTOR.hex_model_R_1671282333708_1_model_5 0.14148
2 Grid_DRF_TRAIN_CONDUCTOR.hex_model_R_1671282333708_1_model_2 0.14576
3 Grid_DRF_TRAIN_CONDUCTOR.hex_model_R_1671282333708_1_model_6 0.15103
4 Grid_DRF_TRAIN_CONDUCTOR.hex_model_R_1671282333708_1_model_1 0.15495
5 Grid_DRF_TRAIN_CONDUCTOR.hex_model_R_1671282333708_1_model_4 0.15531
6 Grid_DRF_TRAIN_CONDUCTOR.hex_model_R_1671282333708_1_model_3 0.16137
```

Figure 13: List of models with their associated hyperparameter values produced after the grid search for the EFF Conductor sub-model

```
H2O Grid Details
=====

Grid ID: Grid_DRF_TRAIN_CONDUCTOR.hex_model_R_1671297828950_1
Used hyper parameters:
- max_depth
- min_rows
- mtries
- ntrees
- sample_rate
Number of models: 2
Number of failed models: 0

Hyper-Parameter Search Summary: ordered by increasing logloss
max_depth min_rows mtries ntrees sample_rate
1 30.00000 1.00000 36.00000 300.00000 0.70000
2 30.00000 1.00000 30.00000 100.00000 0.80000

model_ids logloss
1 Grid_DRF_TRAIN_CONDUCTOR.hex_model_R_1671297828950_1_model_2 0.44326
2 Grid_DRF_TRAIN_CONDUCTOR.hex_model_R_1671297828950_1_model_1 0.45628
```

Figure 14: List of models with their associated hyperparameter values produced after the grid search for the CFO sub-model.

Model Details:
=====

```
H2OBinomialModel: drf
Model ID: Grid_DRF_TRAIN_CONDUCTOR.hex_model_R_167128233708_1_model_5
Model Summary:
  number_of_trees number_of_internal_trees model_size_in_bytes min_depth max_depth mean_depth min_leaves
1             400                400          183445360          35          35    35.00000    31440
  max_leaves mean_leaves
1      39139 36427.63000
```

Figure 15: The best model for the EFF Conductor sub-model, along with the related hyperparameter values.

Model Details:
=====

```
H2OMultinomialModel: drf
Model ID: Grid_DRF_TRAIN_CONDUCTOR.hex_model_R_1671297828950_1_model_1
Model Summary:
  number_of_trees number_of_internal_trees model_size_in_bytes min_depth max_depth mean_depth min_leaves
1             100                800          153047668          17          30    28.87125    206
  max_leaves mean_leaves
1      39200 15223.45100
```

Figure 16: The best model for the CFO sub-model, along with the related hyperparameters values.

In terms of model convergence, the random grid search for hyperparameter tuning uses a stopping criterion based on a specified tolerance in AUC. This means that the additional efforts involved in hyperparameters tuning and training is not likely to improve the model performance beyond the specified threshold. The accuracy of the model prediction, in addition to AUC, can be determined using the Confusion Matrix and Classification Error Rate results.

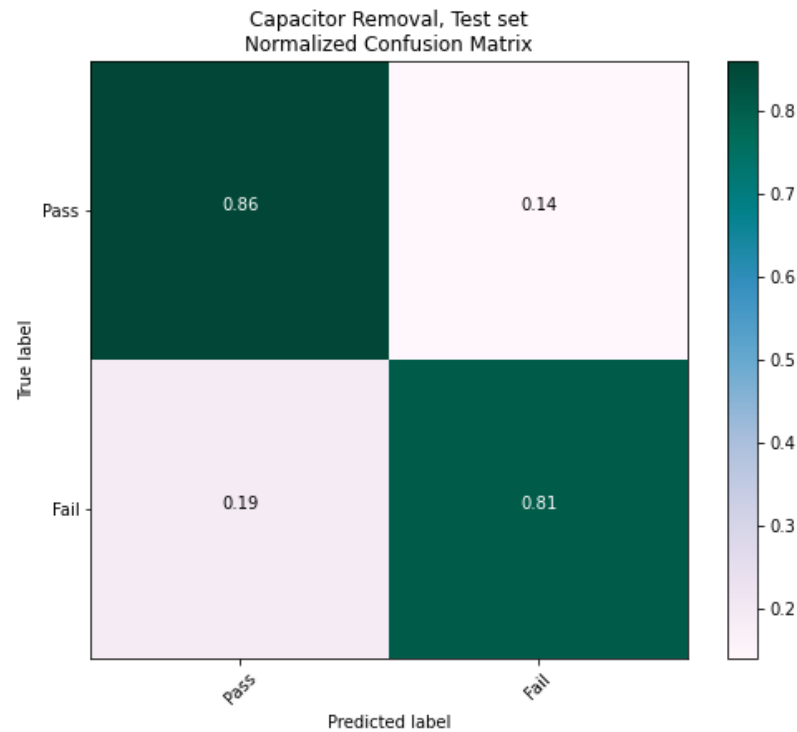
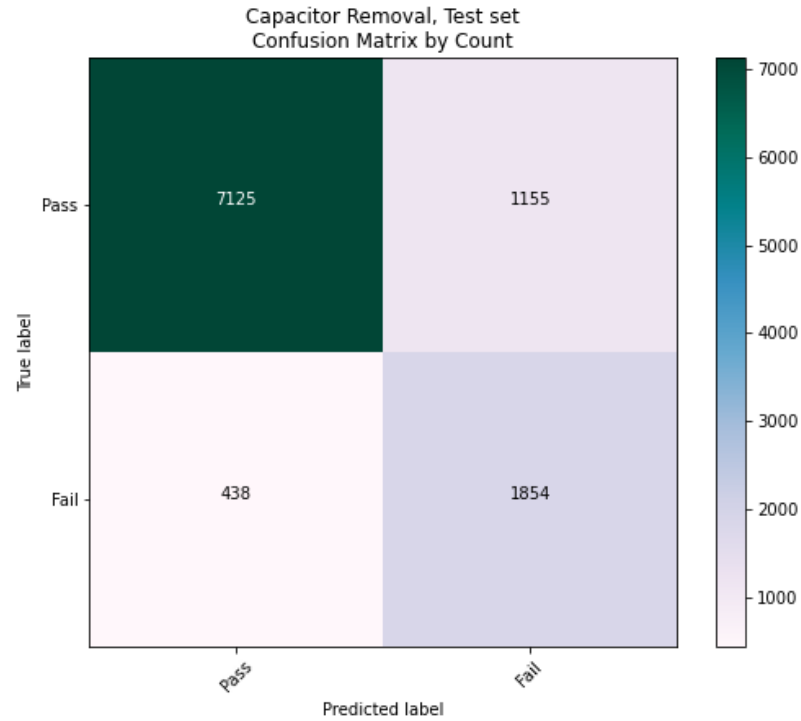
- A confusion matrix presents a tabular layout of the different outcomes of the prediction results of a classification problem and helps visualize its outcomes. It generates a table of all the predicted and actual values of a classifier model.

EFF - Confusion Matrix Results				
Actuals	Predicted			
		0	1	Error Rate
	0	127614	2813	0.021568
	1	4354	6332	0.407449
		131968	9145	0.050789

Table 1: Confusion matrix results for EFF Conductor sub-model

CFO - Confusion Matrix Results									
Actuals	Predicted								
		ANIMAL	BALLOON	EFF	NO	OTHER	UNK	VEGETATION	VEHICLE HIT
	ANIMAL	406	59	2	1654	118	633	61	50
	BALLOON	73	639	3	1853	75	949	66	30
	EFF	5	1	7	10	0	11	5	0
	NO	100	139	2	113012	76	312	95	86
	OTHER	125	94	2	1655	302	978	98	75
	UNK	620	907	10	1905	800	11979	473	287
	VEGETATION	84	66	9	1545	71	503	398	52
	VEHICLE HIT	42	47	0	1452	59	357	52	357
	Totals	1455	1952	35	123086	1501	15722	1248	937
		Error Rate							
		0.8639							
		0.8267							
		0.8205							
		0.0077							
		0.9093							
		0.2946							
		0.8541							
		0.8491							
		0.1291							

- Table 2: Confusion matrix results for CFO sub-modelA confusion matrix presents a tabular layout of the different outcomes of the prediction results of a classification problem and helps visualize its outcomes. It generates a table of all the predicted and actual values of a classifier model.



- and Table 2 provide the confusion matrix results for EFF Conductor and CFO sub-models respectively. It captures the accuracy rate as 94.92% and 87.09% for EFF and CFO sub-model respectively.
- Classification error rate is used to estimate the proportion of instances misclassified over the whole set of instances. It is estimated using the below formula.

$$Error\ Rate = \frac{False\ Positives + False\ Negatives}{True\ Positives + True\ Negatives + False\ Positives + False\ Negatives} * 100$$

The error rate for EFF Conductor and CFO sub-model turns out to be 5.07% and 12.91%. This means that the failure rate of the model prediction is low and under control.

All these test results are performed on test dataset with full outage data till 2022.

A detailed explanation of the compensating controls employed for these limitations are available in Section 2.5.

3.2 Sensitivity Analysis

Sensitivity analysis examines the impact of each feature on the model's prediction. It is a simple yet powerful technique to analyze a machine learning model. To determine the sensitivity of a feature, its value is changed while the values of all other features are held constant. The model's output is then examined. If the outcome of the model significantly changes when the feature value is changed, this indicates that the feature has a significant influence on the prediction. Based on the variable importance feature list shown in Figure 11 and Figure 12, the top five continuous variables of the EFF Conductor and CFO sub-models were chosen to perform the sensitivity analysis. Additionally, three categorical variables for EFF and CFO were also considered for the analysis based on the suggestion provided by the business function.

The feature variables used for sensitivity analysis of the EFF sub-model:

- LENGTH_SEG_CAL_TOTAL
- EQ_PoleHeight
- Conductor_AGE_UDF
- SCD_Seg
- DOWNSTREAM_KVA
- max_of_sum_of_seg_downforce

The feature variables used for sensitivity analysis of the CFO sub-model:

- SumOfCUST_UDF
- CIRCUIT_KV_UDF
- LENGTH_SEG_CAL_TOTAL
- CONDUCTOR_SIZE_UDF
- Log_WindForce

The sensitivity of the model is examined using the test dataset, which contains 20% (141113 observations for EFF and 141905 observations for CFO) of the entire processed data. For this analysis, 10% of the test dataset (14094 observations for EFF and 14577 observations for CFO) was selected and modified using extreme values. Stratified sampling was used to select the 10% of the test data to add randomization to eliminate sampling bias. Sensitivity tests are performed in Dec 2022 using outage data until 2022.

To set up the strata, three categorical features that are specified below were selected from the test dataset.

- Main-line (Yes/No)
- Conductor_type_udf (Aluminium/Copper)
- CC-installed (Yes/No)

Based on these variables, eight different strata were picked. The test data is bound to a column of random numbers produced using a standard normal distribution, and the rank of these random numbers is used to sort the entire set of test data. The top 10% from each stratum was as the target observations to modify the input data. To test the sensitivity of a feature, the values of the selected observations were altered with extreme values (minimum and maximum) of the feature. As a result, for each feature, two sets of test data were generated for sensitivity analysis. Table 3 and Table 4 provide the extreme values (determined by historical data) used for each variable during the sensitivity analysis.

Extreme values used for Sensitivity testing in EFF Conductor sub-model		
Variables	Maximum Value	Minimum Value
LENGTH_SEG_CAL_TOTAL	271774.592	6.5073

Extreme values used for Sensitivity testing in EFF Conductor sub-model		
EQ_PoleHeight	102.5	0
Conductor_AGE_UDF	122	0
SCD_Seg	32277.26	0
DOWNSTREAM_KVA	46619	0
max_of_sum_of_seg_downforce	689053102.9	465304.9234

Table 3: Extreme values used for Sensitivity testing in EFF Conductor sub-model

Extreme values used for Sensitivity testing in CFO sub-model		
Variables	Maximum Value	Minimum Value
SumOfCUST_UDF	7686	0
CIRCUIT_KV_UDF	33	2.4
LENGTH_SEG_CAL_TOTAL	271774.6	0
CONDUCTOR_SIZE_UDF	1000	1/0
Log_WindForce	25.54981911	13.7435954

Table 4: Extreme values used for Sensitivity testing in CFO sub-model

Table 5 and Table 6 provide the AUC results of the unaltered test data i.e., test data without changing the variables' values, and the various sensitivity tests that were performed. For both EFF Conductor and CFO sub-models, the difference in AUC values between the sensitivity tests and the unaltered test data results do not exceed 2% which means that the variations in the input values for these variables does not have a huge impact on the results in terms of AUC.

AUC result of Unaltered Test Data from EFF Conductor sub-model	0.9258
--	--------

EFF Conductor Sub-model Results				
Feature	Maximum value scenario		Minimum value scenario	
	AUC	% Decline in AUC Compared with Unaltered Test Data	AUC	% Decline in AUC Compared with Unaltered Test Data
LENGTH_SEG_CAL_TOTAL	0.9225	-0.33%	0.9201	-0.57%
EQ_PoleHeight	0.9246	-0.12%	0.9227	-0.31%
Conductor_AGE_UDF	0.9249	-0.09%	0.9232	-0.26%
SCD_Seg	0.923	-0.28%	0.9185	-0.73%
DOWNSTREAM_KVA	0.9256	-0.02%	0.9159	-0.99%
max_of_sum_of_seg_downforce	0.924	-0.18%	0.9172	-0.86%

Table 5: The sensitivity results based on AUC for EFF Conductor sub-model

AUC result of Unaltered Test Data from CFO sub-model	0.9453
--	--------

CFO Sub-model Results				
Feature	Maximum value scenario		Minimum value scenario	
	AUC	% Decline in AUC Compared with Unaltered Test Data	AUC	% Decline in AUC Compared with Unaltered Test Data
SumOfCUST_UDF	0.9305	-1.48%	0.9451	-0.02%
CIRCUIT_KV_UDF	0.9309	-1.44%	0.933	-1.23%
LENGTH_SEG_CAL_TOTAL	0.9429	-0.24%	0.936	-0.93%

AUC result of Unaltered Test Data from CFO sub-model				0.9453
CONDUCTOR_SIZE_UDF	0.9448	-0.05%	0.9437	-0.16%
Log_WindForce	0.9407	-0.46%	0.9397	-0.56%

Table 6: The sensitivity results based on AUC for CFO sub-model

Table 7 and Table 8 provide the True Positive Rate (TPR) for the unaltered test data and the various sensitivity tests determined using the prediction output provided by EFF Conductor and CFO sub-models. The increase and decrease in TPR among different tests can be observed from the results but the difference in values seems to be very low. Table 7 also provides the changes in True Positives, False Positives, True Negatives, and False Negatives. The change in the predicted outcome for EFF Conductor seems to be low when compared with the count of observations (14094) that were altered for performing this test.

True Positive rate and False Positive rate for EFF Conductor Sub-model						
	TP	FP	TN	FN	TPR	FPR
Unaltered Test Data	6315	4371	127640	2787	69.38%	3.31%

True Positive rate and False Positive rate for EFF Conductor Sub-model						
Sensitivity Test	TP Change	FP Change	TN Change	FN Change	TPR	FPR
Max Value for LENGTH_SEG_CAL_TOTAL	1	-1	-41	41	69.07%	3.31%
Min Value for LENGTH_SEG_CAL_TOTAL	-26	26	10	-10	69.37%	3.33%
Max Value for EQ_PoleHeight	7	-7	-71	71	68.87%	3.31%
Min Value for EQ_PoleHeight	5	-5	-32	32	69.15%	3.31%
Max Value for Conductor_AGE_UDF	11	-11	-56	56	68.99%	3.30%
Min Value for Conductor_AGE_UDF	-6	6	-30	30	69.13%	3.32%
Max Value for SCD_Seg	5	-5	-54	54	68.99%	3.31%
Min Value for SCD_Seg	-15	15	-2	2	69.31%	3.32%
Max Value for DOWNSTREAM_KVA	56	-56	-228	228	67.88%	3.28%
Min Value for DOWNSTREAM_KVA	-9	9	12	-12	69.44%	3.32%
Max Value for max_of_sum_of_seg_downforce	31	-31	-94	94	68.78%	3.29%
Min Value for max_of_sum_of_seg_downforce	-3	3	-29	29	69.15%	3.31%

Table 7: The sensitivity results based on predicted outcome for EFF Conductor sub-model

Since CFO sub-model produces six predictions in its outcome, the analysis on predictions is provided at the sub-driver level. Similar to the EFF Conductor sub-model results, the increase and decrease in TPR rate among different tests can be witnessed but the difference in values seems to be very low.

True Positive Rate (TPR) for CFO Sub-model							
	Animal	Balloon	No	Other	Unknown	Vegetation	Vehicle
Unaltered Test Data	13.61%	17.33%	99.29%	9.10%	70.54%	14.59%	15.09%

True Positive Rate (TPR) for CFO Sub-model							
Sensitivity Test	Animal	Balloon	No	Other	Unknown	Vegetation	Vehicle
Max Value for SumOfCUST_UDF	13.04%	17.11%	99.33%	8.74%	68.26%	14.37%	14.24%
Min Value for SumOfCUST_UDF	12.94%	17.19%	98.87%	8.26%	69.80%	14.33%	13.91%
Max Value for CIRCUIT_KV_UDF	12.50%	16.78%	99.30%	8.83%	69.62%	14.52%	15.09%
Min Value for CIRCUIT_KV_UDF	11.70%	17.03%	98.04%	8.86%	70.19%	14.41%	14.96%
Max Value for	12.91%	17.06%	99.27%	8.92%	70.58%	14.41%	15.00%

True Positive Rate (TPR) for CFO Sub-model							
LENGTH_SEG_CAL_TOTAL							
Min Value for LENGTH_SEG_CAL_TOTAL	13.01%	16.97%	99.30%	8.68%	70.42%	14.44%	14.96%
Max Value for CONDUCTOR_SIZE_UDF	12.77%	17.06%	99.31%	8.83%	70.53%	14.44%	14.79%
Min Value for CONDUCTOR_SIZE_UDF	13.07%	17.14%	99.31%	8.80%	70.53%	14.41%	14.84%
Max Value for Log_WindForce	13.38%	17.16%	99.29%	8.89%	70.53%	14.63%	15.05%
Min Value for Log_WindForce	13.17%	17.35%	99.29%	9.13%	70.56%	14.59%	15.09%

Table 8: The sensitivity results based on predicted outcome for CFO sub-model

Based on these test results, it can be determined that the variations in the input values for the high importance features alters some of the predictions from the model, but the magnitude of the impact seems to be low. Hence the model results from both EFF Conductor and CFO sub-models are robust and reliable post sensitivity testing for the variables defined in this section earlier with extremely high and low values tested for each of the defined variables.

3.3 Outcome Analysis / Backtesting

The subset of historical data on which a model is trained and optimized is referred to as the in-sample data. On the other hand, the subset of the dataset that has been reserved to test the model is known as the out-of-sample data. The OH Conductor model uses a random sampling approach to split the dataset into Train (80%) and Test (20%) data. The results arrived from train data are considered as in-sample backtesting and the results arrived from test data are considered as out-of-sample backtesting.

Once the machine learning model is built with the training data, it is evaluated using a separate test dataset that has not yet been studied. The performance of the model is determined by the Area Under the ROC Curve (AUC) value. Figure and Figure 18 shows the AUC ROC for EFF Conductor and CFO sub-models based on the test dataset ran with full historical outage data until 2022.

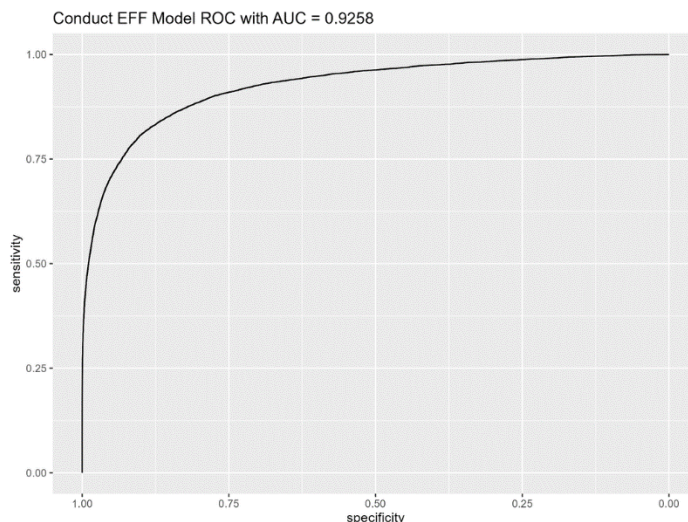


Figure 17: Out-sample backtesting result for EFF Conductor sub-model based on test dataset

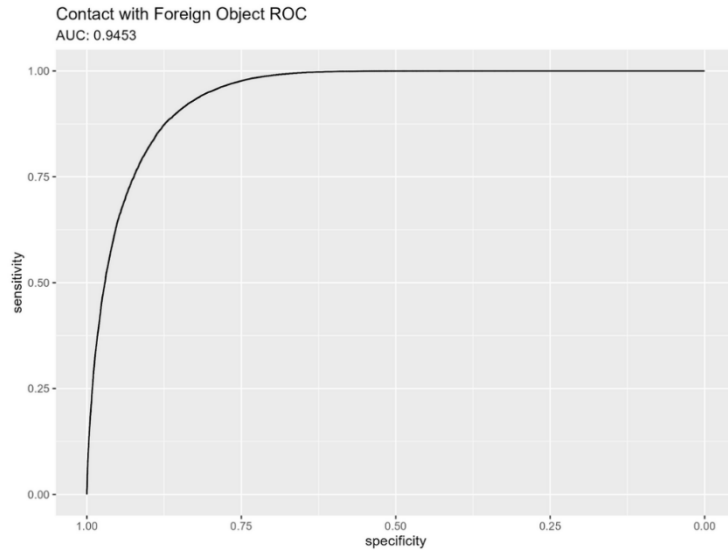


Figure 18: Out-sample backtesting result for overall CFO sub-model based on test dataset

The AUC values for the EFF Conductor and CFO sub-models are 0.9258 and 0.9453 respectively. The AUC values of both sub-models are higher than 0.91 which imply that the models possess high accuracy in terms of predicting the results.

Figure 19 provides the ROC curves and the associated AUC for the individual sub-drivers used in the CFO sub-model to derive the failure probabilities due to contact from different foreign objects.

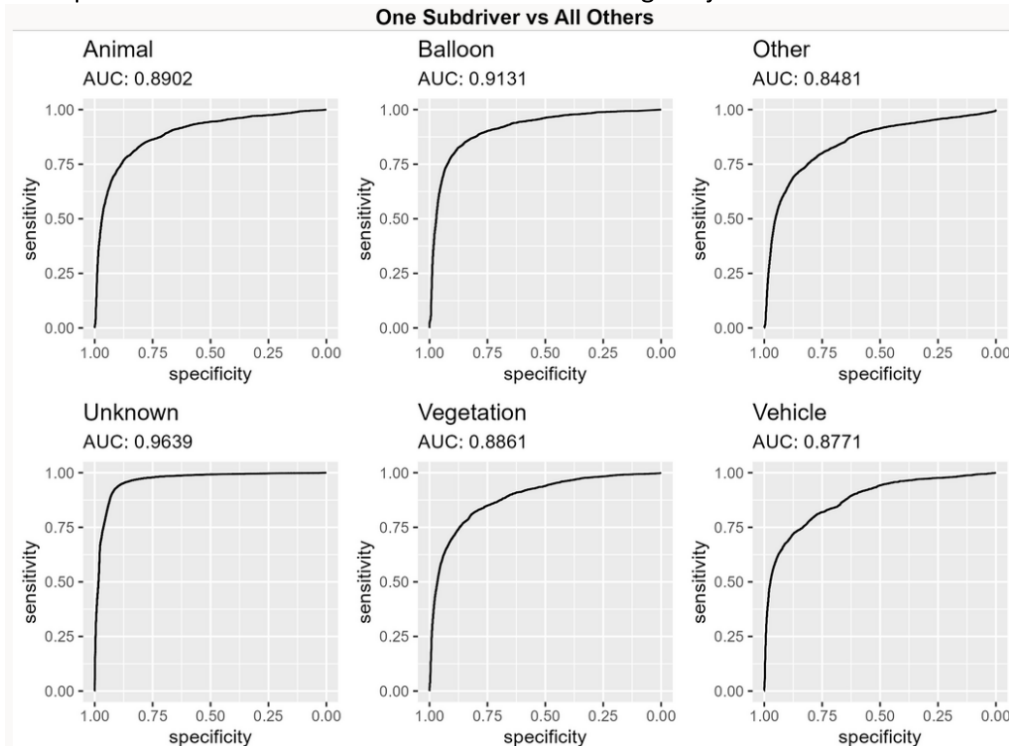


Figure 19: In-sample backtesting result for sub-drivers in CFO sub-model based on test dataset

The AUC values for sub-drivers Animal, Balloon, Other, Unknown, Vegetation and Vehicle are 0.8902, 0.9131, 0.8481, 0.9639, 0.8861, and 0.8771.

The impact of uncertainty in model inputs and parameters on model outputs are tested as a part of the sensitivity analysis and the results are captured in Section 3.2. In addition, the data imputations that are incorporated to address missing values before running the model are defined in Section 2.1.

3.4 Benchmarking Analysis

For the EFF Conductor and CFO sub-models, different approaches like Gradient Boosting Machine (GBM) learning, Logistic Regression, and Random Forest were considered during the model development phase in 2019. The analysis on these supervised machine learning approaches and the results are provided below.

- Gradient Boosting Machine (GBM) is one of the most popular forward learning ensemble methods in machine learning. It is a powerful technique for building predictive models for classification and regression tasks. GBM sequentially combines the predictions from various weak learner decision trees and builds a final predictive model with more accurate predictions by minimizing a defined loss function.
- Logistic regression is used to solve classification problems. The three types of logistic regression available are Binary logistic regression (handles binary outcomes), Multinomial logistic regression (handles multiple outcomes, i.e., multi-classification variable), and Ordinal logistic regression (handles ordered outcomes). In contrast, linear regression solves regression problems where the outcome is continuous and can be any possible numeric value.
- Random Forest is a popular machine learning algorithm that can be used for both classification and regression problems. Random Forest is another ensemble method that combines the predictions of several decision trees to improve the predictive accuracy of the model. The individual decision trees are created based on a randomly selected subset of features at each node prior to determining the optimal split so each tree differs. The final output is determined by taking the majority vote of the predictions from the individual decision trees. The greater number of trees in the forest generally leads to higher accuracy and prevents the problem of overfitting.

The benchmarking results of GBM and Logistic Regression shared in this section were developed using the h2o library in R on the Test data with targets based on the full historical outage data until 2022. Since benchmark results were not saved during the model development phase, the benchmark models were executed in Dec 2022 for documentation purposes. Figure 20 and Figure 21 provide the AUC values for the EFF Conductor and CFO sub-models using the GBM, Logistic Regression, and Random Forest methodologies.

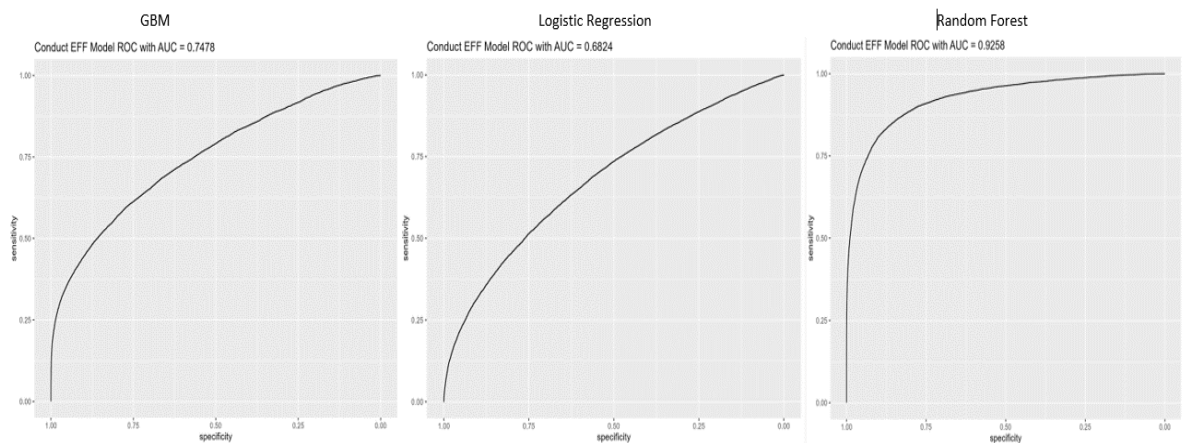


Figure 20: AUC Comparison for the EFF Conductor sub-model using GBM, Logistic Regression, and Random Forest methodologies

For the EFF Conductor sub-model, the AUC results for GBM, Logistic Regression, and Random Forest were 0.7478, 0.6824, and 0.9258 respectively.

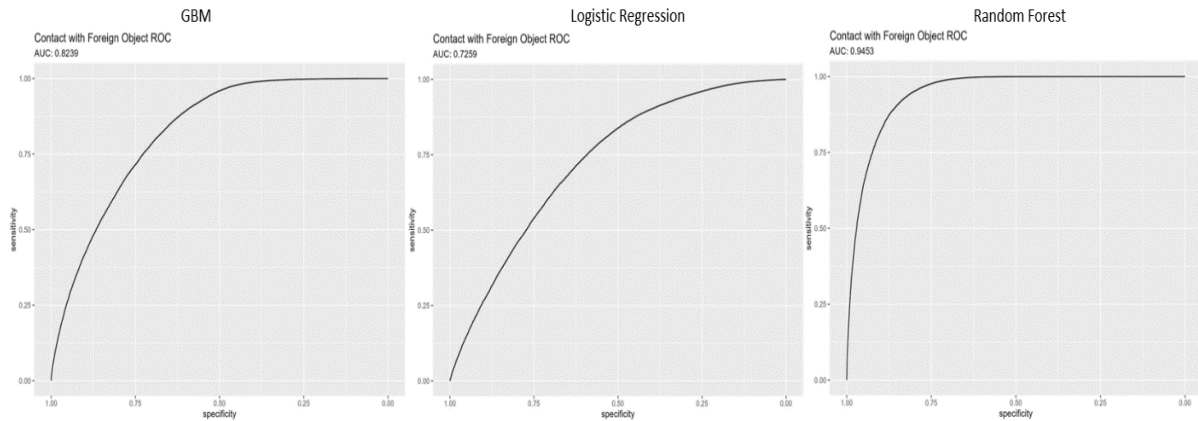


Figure 21: AUC Comparison for the CFO sub-model using GBM, Logistic Regression, and Random Forest methodologies

For the CFO sub-model, the AUC results for GBM, Logistic Regression, and Random Forest were 0.8239, 0.7259, and 0.9453 respectively.

Random Forest was chosen as the modeling algorithm for the EFF Conductor and CFO sub-models since it achieved the highest AUC among the three approaches. Some additional advantages of using Random Forest over GBM and Logistic Regression are provided below:

- Random Forest is less sensitive to overfitting issues than GBM.
- Hyperparameter tuning is relatively easy in Random Forest when compared with GBM.
- Random Forest is better at handling categorical variables while retaining the original encoding compared to weight-based algorithms like logistic regression which may treat categories of higher importance depending on the number assigned.

4. MODEL MANAGEMENT AND GOVERNANCE

4.1 Ongoing Monitoring Plan

Ongoing monitoring is important for Machine Learning models especially when they are used to make predictions or when they are run on datasets with high volatility in variable values. The EFF Conductor and CFO sub-models are run manually once a year, incorporating updated input datasets to reflect the latest available data and implementing any specific model enhancements—e.g., inclusion / replacement / removal of a feature, optimization of the code, evaluation of a new performance metric, etc. During the model refresh, the limitations and assumptions of the model are also revisited by the model developers and necessary action items are conducted to address them. Performance monitoring is required only after running the model. Recalibration of the model has not been performed for the last two years, and it is performed only if the behavior of the model differs from that of the previous model or if there is a significant drop in model performance. The AUC and accuracy rate from confusion matrix results obtained after model refresh are compared against a threshold of 80%; if the value drops below this threshold, the reason behind the performance dip is investigated. Post-investigation, the steps required to improve the performance of the model will be carried out. To monitor the model performance more thoroughly, developers of the model plan to additionally evaluate metrics like Precision and Recall. Precision is the positive predictive value which represents the proportion of predicted failures that were predicted correctly. Recall is the true positive rate which represents the proportion of actual failures that were predicted correctly.

The model documentation and the performance results are updated once a year immediately after the model refresh.

4.2 Security and Control

The Data Science and Asset Analytics team has access to the data inputs, code, and implementation for the model.

Other business units, like the Grid Hardening Strategy team, are provided access to the model outputs upon request but cannot update or modify the code.

The model is run using R programming and it can be executed in any recent versions of the R software. Current model versioning is labeled by date of refresh (e.g., CFO_EFF\update_20220527). There are plans to move the code to GitHub, a platform that facilitates version control by tracking changes to the source code. Users with write or admin privileges to the repository can review proposed changes and approve them.

A contingency plan is not applicable for this model as it is an inhouse model for SCE.

5. REFERENCES

RF 1: SCE's WMP 2022 Q1 Quarterly Data Report submission

<https://www.sce.com/sites/default/files/AEM/Data%20Requests/2022/SCE%20Q1%202022%20Tables%201-12.xlsx>

RF 2: Fatality Analysis Reporting System (FARS) from the National Highway Traffic Safety Administration (NHTSA)

<https://www.nhtsa.gov/node/97996/251>

RF 3: CRIMESTAT

<https://en.wikipedia.org/wiki/CrimeStat>

RF 4: Literature reference on grid search vs random search approach for hyperparameter tuning.

<https://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf>

RF 5: Variable Importance methodology for tree-based methods

[Variable Importance — H2O 3.38.0.3 documentation](#)

Attachment C

OH-Switch Sub-Model

**Southern California Edison (SCE)
Model Documentation
Prepared for 2023 WMP Appendix B**

OH Switch Sub-Model

3/27/23

Table of Contents

1.	EXECUTIVE SUMMARY	5
1.1	Model Purpose and Intended Use	5
1.2	Model Description Summary	6
1.3	Model Risk Rating	6
1.4	Model Dependency and Interconnectivity	6
1.5	Model Assumptions	7
1.6	Model Limitations	7
1.7	Overall Model Performance Assessment	7
1.8	Contingency Plan for Vendor Model	8
2.	MODEL FRAMEWORK AND THEORY	9
2.1	Model Inputs and Data Quality	9
2.2	Methodology	11
2.3	Suitability	14
2.4	Assumptions	14
2.5	Limitations and Compensating Controls	15
2.6	Model Outputs	16
3.	MODEL PERFORMANCE AND TESTING	19
3.1	Model Specification Testing	19
3.2	Sensitivity Analysis	23
3.3	Outcome Analysis / Backtesting	24
3.4	Benchmarking Analysis	25
4.	MODEL MANAGEMENT AND GOVERNANCE	26
4.1	Ongoing Monitoring Plan	26
4.2	Security and Control	27
5.	REFERENCES	28

1. EXECUTIVE SUMMARY

1.1 Model Purpose and Intended Use

The OH (Over Head) Switch Model is a Probability of Ignition (POI) Sub-Model developed by SCE (Southern California Edison). At SCE, models are developed using ML (Machine Learning) algorithms for each asset, i.e., OH Conductor, OH Switch, etc., and at each contact type level like animal, balloon, etc., as the drivers vary by asset and contact type. The OH Switch model is refreshed annually and used to predict the probability of failure (POF) for distribution overhead switches.

The calibrated outputs of the OH Switch model—i.e., failure events—are broadly used by two categories of programs described below:

8. The Inspections and Remediations programs, which considers POI as an element in prioritization and scoping.
9. Risk analyses via SCE's MARS Framework.

1.2 Model Description Summary

The OH Switch model is a binary classification model using Random Forest—a Machine Learning technique. It predicts the probability of a switch igniting a spark due to equipment failure by considering available switch attributes and condition data (i.e., age, voltage, etc.) and other environmental and operational attributes (i.e., historical wind, switching counts, etc.).

The model is implemented in R programming using the library h2o and is connected to databases such as SAP, ADS Weather, etc. The model is run once a year manually by the Data Science and Asset Analytics team. The model is calibrated every year with the last 5 years of historical failure data.

Cross-references: Please refer to Section 2.1 for more information about the inputs used by the OH Switch model along with data processing details.

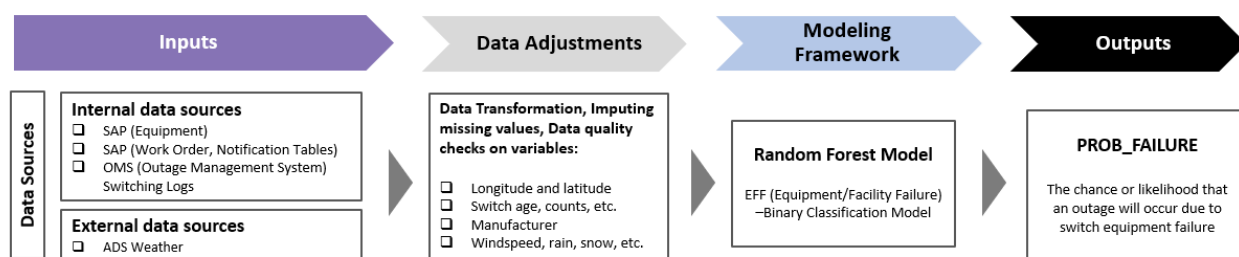


Figure 10: OH Switch model framework

The OH Switch model uses the Random Forest methodology. Since the Random Forest methodology can perform both classification and regression tasks, it is considered a good choice for the OH Switch model as the prediction is a classified event, i.e., failure. This methodology predicts output with high accuracy, runs efficiently on large datasets, and maintains accuracy with minimal adjustments for missing values and data treatments.

1.3 Model Risk Rating

There is no defined mechanism of identifying model risk rating at SCE, however certain factors—like frequency of risk events and use case—are considered while flagging model risk. Based on the Wildfire Mitigation Plan quarterly report, the frequency of outages in a year from switches averages around 60 which is low compared to other sub-drivers. Figure 2 provides a snapshot of the count of outages over the years by some Equipment/Facility Failure (EFF) sub-drivers, with switches in bold. In addition, the output of this model is considered important as it informs the strategy of a few programs which are discussed in section 1.1. Hence, the OH Switch model is deemed to be a medium risk model.

Table 7.1: Key recent and projected drivers of risk events			Number of risk events																Projected risk events							
Risk Event category	Cause category	Sub-cause category	2015	2016	2017	2018	2019	2020	2020	2020	2020	2021	2021	2021	2021	2022	2022	2022	2022	2023	2023	2023	2023	2023		
Outage - Distribution	18. Equipment / facility failure - Distribution	Capacitor bank damage or failure- Distribution	280	275	372	337	426	126	159	72	46	110	98	124	80	100	96	102	90	96	96	102	90			
		Conductor damage or failure - Distribution	463	594	654	713	1116	206	144	211	252	276	109	133	319	228	235	209	296	294	229	204	288			
		Fuse damage or failure - Distribution	232	195	245	508	1245	169	176	317	167	179	132	201	183	165	158	156	178	181	158	156	178			
		Lightning arrester damage or failure- Distribution	105	127	99	106	216	27	21	26	25	12	21	18	22	8	26	24	27	24	26	24	27			
		Switch damage or failure- Distribution	51	46	45	67	78	17	11	16	19	14	10	18	22	10	13	16	19	15	13	16	19			
		Pole damage or failure - Distribution	98	126	130	207	541	57	36	31	40	32	22	21	60	32	47	43	56	52	47	43	56			
		Insulator and brushing damage or failure - Distribution	42	75	79	123	121	28	14	11	43	30	13	22	45	17	15	21	37	27	15	21	37			
		Crossarm damage or failure - Distribution	127	143	138	354	834	98	45	29	45	39	17	17	61	34	46	36	63	56	46	36	62			

Figure 11: Key recent and projected risk events due to switch damage or failure from SCE Q1 2022 Quarterly Data Report, Table 7.1

Cross-references: Refer to link [\[RF 1\]](#) in Section 5 for SCE’s Wildfire Mitigation Plan Q1 2022 Quarterly Data Report submission.

1.4 Model Dependency and Interconnectivity

The OH Switch model is an “Ignition Likelihood” model which uses the inputs from the ADS (Atmospheric Data Solutions) modeling output along with other data sources to calculate the probability of ignition.

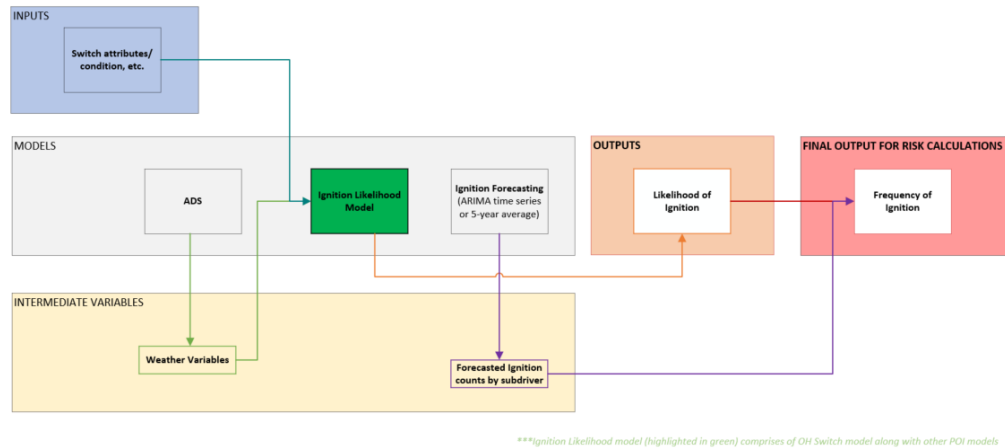


Figure 12: Model Interconnectivity Schema

OH Switch model uses the ADS model as one of its data sources to get the inputs for the weather variables. ADS's Next Generation Weather Modeling System (NGWMS) provides an extensive upgrade to SCE's in-house weather modeling capabilities and enhances SCE's ability to make more targeted PSPS decisions. The ADS model generates 10 years of hourly weather data between 2010 and 2019. That information is then processed and aggregated to calculate statistical measures, like mean and standard deviation, of wind, humidity, rain, snow, etc. These are used as locational measures and are matched to the switches through the data engineering step of performing a spatial join to the nearest grid by their latitude and longitude coordinates.

The output data from the OH Switch model, i.e., POI, is used by two categories of programs, further discussed in Section 1.1, to inform their strategic decisions.

1.5 Model Assumptions

The business assumptions and model assumptions for the OH Switch model are summarized below:

12. There is no change in the OH Switch technical specification over time.
13. The calibration methodology assumes that fires are a subset of failures.
14. The model is designed to work in both base weather and extreme weather conditions.
15. The feature variables in the dataset should have some actual values so that the classifier model can predict accurate results.
16. The predictions from each tree must have very low correlations.

A detailed explanation of these assumptions is available in Section 2.4.

1.6 Model Limitations

The model limitations for the OH Switch model are summarized below:

9. Unavailability of linear/non-linear representation in the form of intuitive equation or correlation statistic.
10. Time consumption for model execution is high.
11. Resource utilization in terms of system capacity and higher configuration for model execution is high.
12. Model accuracy may reduce if the dataset experiences covariate shift.

A detailed explanation of these data limitations is available in Section 2.5.

1.7 Overall Model Performance Assessment

The machine learning model used to build the OH Switch model is the Random Forest algorithm. The model's overall performance is determined by the Area Under the ROC Curve (AUC) value and Confusion Matrix results.

The performance of the OH Switch model was evaluated on test data with historical failure information between 2017-2021.

- The AUC value is 0.85.
- Confusion matrix results capture the accuracy rate as 93.8%.

The above metrics were derived at the time of the model refresh in June 2022 to capture an exhaustive set of statistical results for documentation purposes.

1.8 Contingency Plan for Vendor Model

A contingency plan is not applicable for this model as it is an inhouse SCE model. This is not a vendor model.

2. MODEL FRAMEWORK AND THEORY

The OH Switch model is a binary classification model pertaining to switch equipment failures. It employs a random forest algorithm to predict the likelihood of a switch experiencing a failure that can result in an ignition event. The random forest approach was chosen for the classification task over other modeling approaches—such as gradient boosting, etc.—because it predicts output with high accuracy, runs efficiently on large datasets, and maintains accuracy with minimal adjustments for missing values and data treatments.

2.1 Model Inputs and Data Quality

Data Sources

This model refers to multiple internal and external data sources. The internal data sources used by the model are:

- **SAP** houses circuit⁵, structure, and equipment characteristics. It contains latitude and longitude information of the assets. SAP also houses notification and work order records which track planned and completed work for issues like repairs, replacements, etc. The latter is used to develop failure targets for the model.
- **OMS** refers to Outage Management System which contains information about switching operations and about switch assets including latitude and longitude as well.

The external data sources used by the model are:

- **ADS** (Atmospheric Data Solutions) model provides 10 years of hourly gridded weather data from 2010-2019. These are aggregated to individual locational measures and matched to the switches through spatial join to the nearest grid by the latitude and longitude as a part of the data engineering step.

Quality Checks

SCE has internal data management teams for ensuring data quality, including EAD (Enterprise Asset Data) and Master Data. They work on processing asset data corrections (E2 notifications) in SAP and fixing largely known data issues like missing or erroneous latitude and longitude information for assets in the territory. Some of the data quality checks that are performed in the OH Switch model to ensure accuracy, validity, integrity, and consistency are provided below. Quality checks (QC) are coded in R or incorporated into the data gathering process.

The QC steps performed by automated R code are as follows:

- Duplicate records that are identified in the switching operations log are removed to maintain consistency in counting times operated by considering only the distinct records including date stamp.
- SAP includes all historical equipment records. The most recent record for each equipment is selected to filter out inactive records if a replacement occurred.
- SAP provides information about all the removals and remediations encountered by SCE. Only the relevant failures specific to switches are loaded into the respective model. All the other non-relevant information, i.e., for equipment other than switches, is excluded.

The manual QC steps are as follows:

- ADS weather data is validated against actual weather observations.
- Asset data obtained from SAP is validated and updated through inspections and other programs.

Data Sampling

Since this is a classification model to predict switch failures, there are no sampling strategies used in the model other than the random split strategy to bifurcate the train and test data. The dataset used for the model is randomly divided to have 80% in train data and 20% in test data.

Data Cleansing and Transformation

⁵ Circuit comprises a collection of segments that altogether form a path for electrical current floating from the power source (including but not limited to a substation) to another power source or circuit endpoint.

The data cleansing and transformation activities that are incorporated in the R scripts as a part of automation to ensure the completeness of data used for model training and estimation are provided below.

- Missing data in SAP for the below specified numeric variables are handled by referring to another database, OMS
 - latitude
 - longitude
- The asset's manufacturer is not always populated so values are imputed for this feature with help of information in other variables like switch subtype.
- Data consistency is ensured by correcting formatting issues in date variables; e.g., Start-Up Date variable can have different formats of data and is corrected in R program code by forcing the values to a consistent day, month, year format.

Data Assumptions

The accuracy of the predicted results is dependent on the accuracy of the data used to build the predictive models. The data assumptions follow:

8. The assumptions used for the data imputation utilized SCE's Distribution Design Standard (DDS), engineering judgment, and manufacturer data.
9. The target labeling process used to label the failures and non-failures as '1' and '0' is considered accurate. This is performed by observing historical failure records. If the switch experiences a removal, replacement, or repair within the study period (2017-2021) then '1' is assigned to represent failure. Else '0' is assigned for non-failures.
10. Input data with respect to asset, weather, and engineering information are assumed to be stable and will not change over time until the subsequent data refresh. For example: If there is an update in the structure information specific to an asset, that updated information will be reflected only in the subsequent data refresh. So, it is assumed that the updated structure information is not drastically different from the previous information which may alter the model outcomes.

Data Limitations

The following are data limitations across internal and external data sources:

Some of the data used by the model faces accuracy issues in terms of consistency in data labelling or missing values which might impact model prediction power.

- Data labelling issues may be due to manual errors during data entry. E.g., Manufacturer information is fed manually into the system. While updating the name, different labels for the same manufacturer might be used in different data entries which affects the consistency of the data. Hence these consistency issues in data need to be addressed before using them in the model.
- Missing data for a specific feature (predictive variable) might be due to unavailability of data. E.g., for the switching counts feature, some planned work and energized operations like fully operated load and load drops at lesser currents are not tracked. To overcome this issue, recorded switching counts pulled from OMS are used as a proxy to estimate times operated which is further used in model processing. Other missing values for a switch are filled using imputations by cross-referencing other fields or other data sources to mitigate the risk arising from missing predictors.

Independent variables

The OH Switch model uses multiple variables/features. A subset of the independent variables used in the OH Switch model, along with its data source and description, is provided below.

Feature	Data Source	Description
eq_sys_voltage_UDF	SAP	Voltage handled by the equipment
district_UDF	SAP	District of the FLOC in which the equipment is installed

AGE_UDF	SAP	Calculated age of switch equipment, from start-up date to the beginning of forecast period (2022)
oms_switching_cts	OMS	Counted # of times the switch was operated. An operation is either open or close.
manufacturer_UDF	SAP	Manufacturer of the switch
ASSET_SUBTYPE_UDF	SAP	Switch asset subtype (Accessory or Pole Disconnect, Dip, Horizontal, Tiered, Triangular, Vertical)
ASSET_TYPE_UDF	SAP	Switch asset type (Loadbreak, Non-loadbreak)
GROUP_UDF	SAP	Switch group (GOAB, Hookstick Operated Single Blade Disconnect Switch)

In addition to the data utilized above, 10 years of hourly data fetched from ADS Weather model is processed and aggregated to calculate statistical measures like mean, max, and standard deviation for wind, temperature, water vapor, turbulence kinetic energy, humidity, rain, and snow.

Dependent Variable

In a typical classification risk model, defining the dependent variable is key for both model development and model performance assessment. The dependent variable in the OH Switch model represents the observation of a switch equipment failure in terms of removals, replacements, or repairs. It is a binary status of failure or non-failure.

The final output of the model is `PROB_FAILURE`, representing the chance or likelihood that a switch failure will occur. The `h2o.predict (level = 0.05, type = 'response')` function is used to specify the desired output (`PROB_FAILURE`) in probability values, rather than binary values. The probability value ranges from 0 to 1 where '0' represents the least likelihood of failure and '1' represents high chance of failure.

2.2 Methodology

SCE utilizes machine learning to identify patterns that may lead to failures causing sparks from switches and uses the trained model to predict POIs at the asset level. The OH Switch model employs a random forest algorithm to predict failure events. The Random Forest approach can predict outputs with high accuracy, run efficiently for large datasets, and maintain accuracy with minimal adjustments for missing values and data treatments.

A random forest is a supervised machine learning algorithm that is constructed from many decision trees. It can be used to solve both classification and regression problems. This approach utilizes ensemble learning, which is a technique that combines many classifiers to achieve greater predictive accuracy than that of a single classifier. A decision tree is a decision support technique that forms a tree-like structure. It consists of three components: decision nodes, leaf nodes, and a root node. The following diagram shows the three types of nodes in a decision tree.

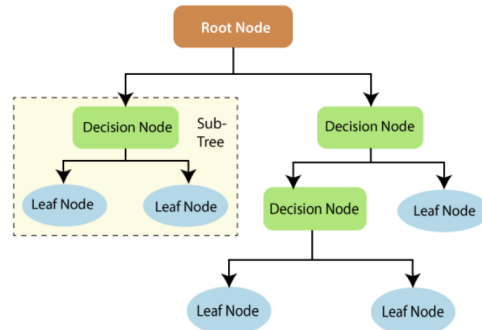


Figure 4: Decision Tree Structure

A decision tree algorithm divides observations of a dataset into branches, which further segregate into other branches. This sequence continues until a leaf node is attained. A leaf node cannot be segregated further. In more detail, the root node is the base of a decision tree, where the first of a chain of decisions is made. A branch is the connection path between nodes. A node is a potential splitting point on a tree. Decision nodes provide a link to the leaves. On the other hand, leaves, also known as terminal nodes, are the ends of a tree, representing the resulting classification or value for the sample.

The 'forest' generated by the random forest algorithm is trained through bagging, also known as bootstrap aggregating. Bagging is an ensemble meta-algorithm that fits multiple models on different subsets of a training dataset and then combines the predictions from all models. The diagram below shows a simple random forest classifier.

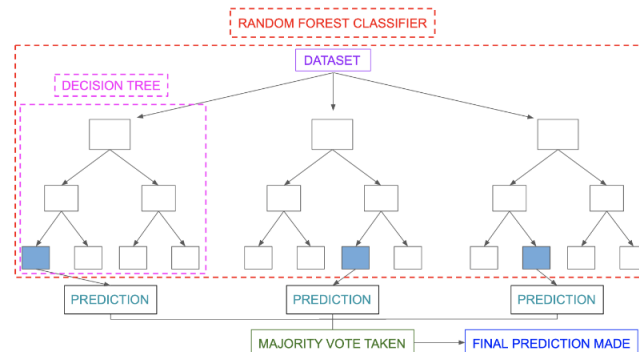


Figure 5: Structure of Random Forest Classifier model

The selection of the final output follows a majority-voting system. In this classification model case, the output chosen by a majority of the decision trees becomes the final output of the random forest system. The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.

Train test split is a model validation procedure that simulates how a model would perform on new/unseen data. Figure 6 shows the logic of dividing the dataset into train data and test data. First the data is consolidated and prepared for train test split. Then the historical input datasets are split into a training dataset (80%) and testing dataset (20%) based on simple random sampling strategy with a split ratio of 4:1 without replacement. Simple random sampling is a technique that ensures each observation has an equal likelihood of being selected for a set. It is a fair strategy as it helps in avoiding any bias involved compared to other modeling techniques and it has no restrictions on the sample size which makes it suitable to handle vastly sized input data. The predictive algorithm is developed using the training dataset and is built by looking at the interactions between all the features to find patterns and predict the likelihood of equipment failure.

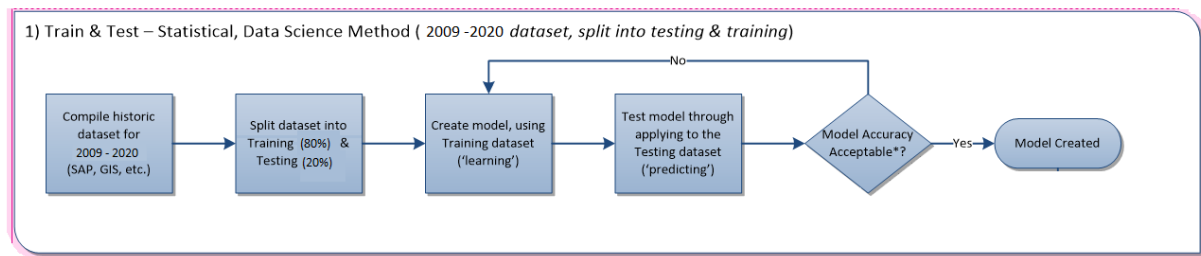


Figure 6: Train and Test data split logic

In the next step, the algorithm is tested on the ‘testing’ dataset. The model is run on the test dataset to make a prediction of a failure or success. Then an internal validation of the model is conducted by comparing the predicted results to the actual results which indicates the predictive capabilities of the features as well as the model. AUC is the metric used to assess the performance of the model on test data.

Area Under the Curve (AUC) – Area Under the Receiver Operating Characteristic (ROC) Curve is a measure to estimate the model discriminatory power (degree of separability) for the binary classification problem. The ROC curve plots True Positive Rate against different thresholds with False Positive Rate (FPR) or True Negative Rate (TNR). The higher the AUC, the better the model is at predicting True Negatives (non-events) and True Positives (events).

Hyperparameter Tuning:

Hyperparameters are parameters that are explicitly defined by the user to control the learning process. The process of selecting the optimal hyperparameters to use is known as hyperparameter tuning, and the tuning process to achieve the best-defined performance statistic is known as hyperparameter optimization. Cartesian Grid search and Random Grid search are the most widely used strategies for hyperparameter optimization.

- In the Cartesian grid search approach, the machine learning model is evaluated for a range of hyperparameter values, and it searches for the best set of hyperparameters from a grid of hyperparameters values. The disadvantage of grid search model is that it will go through all the intermediate combinations of hyperparameters which increases the time consumed by grid search computations.
- In the random grid search approach, the machine learning model is evaluated for a range of hyperparameter values like that in Cartesian Grid Search approach. However, search criteria parameters are added to control the type and extent of the search, and it moves randomly within the grid to find the best set of hyperparameters to achieve maximum performance in terms of the metric defined by the user. As search criteria, the user can set a maximum runtime for the grid, a maximum number of models to create, or metric-based automatic early stopping. If many of these requirements are supplied, the algorithm will end when the first of the criteria is met. This approach reduces the time taken for computation thereby solves the drawbacks of the cartesian grid search approach.

The OH Switch model uses Random Grid Search method for Hyperparameter tuning. The reference literature link to understand the efficiency between Cartesian Grid search and Random Grid search is provided below. The criterion used for the hyperparameter tuning in OH Switch Model are:

- **ntrees:** Total number of trees used in the random forest. For tuning this parameter, the OH Switch model uses a range of values between 100 and 500 with an increment of 50.
- **mtries:** Total number of predictors/variables that will be randomly selected in each node to search for the best split. This parameter is varied by using different percentages of the total number of independent variables in the model. The various percentages taken into account are 5%, 15%, 25%, 33.3%, and 40%.
- **max_depth:** Specifies the maximum size of the sample data drawn for training each tree. A higher value for this feature will make the model more complex and can lead to the issue of overfitting the training data. For the max_depth parameter tuning, the range of values used for the OH Switch model are set between 25 and 50 with an increment of 5.

- **min_rows:** This parameter defines the minimum number of observations required for a leaf to split. This parameter is tuned using the values 1, 2, 7, 10, and 15.
- **Sample_rate:** The percentage of the sample data drawn for training each tree. The scale goes from 0 to 1.0. For both models, this parameter is tweaked with values of 0.5, 0.632, 0.7, and 0.8.

Random Grid Search method uses the below specified stopping criterion in the OH Switch model to stop the random grid search. The conditions are provided below.

- **stopping_tolerance = 0.005**
This will stop the random search if the tolerance level reaches 0.005.
- **stopping_rounds = 15**
This will stop the random search if none of the last 15 models managed to have 0.5% improvement compared to best model identified before that.
- **max_runtime_secs = 3600**
This is used to define the maximum number of seconds allowed for the search. The random search will stop if the search continues to find improvements after 30 min.
- **stopping_metric = AUC**
This defines the performance metric-based condition to stop the search. The random grid search will stop when the model's AUC value doesn't improve by 0.5% for the OH Switch model.

Once the random search is complete, the grid object containing the list of models is queried, and models are sorted by a performance metric defined by the user. The model with better performance is chosen as the best model and it is validated on the test data.

Cross-references: Refer to [\[RF 2\]](#) in Section 5 to understand the efficiency between Cartesian Grid search and Random Grid search.

2.3 Suitability

During development of the model in 2017, Gradient Boosted Machine (GBM) was used to construct the OH Switch model. Then Random Forest, another modeling approach, was tested. The test results proved that the Random Forest methodology fits well for the OH Switch model as it exhibited similarly high AUC as other approaches. AUC comparison of these two approaches is specified in Section 3.4.

Random Forest methodology can be used to solve both classification as well as regression problems and it can handle both categorical and continuous variables. One of the main advantages of the Random Forest methodology is that it maintains accuracy with minimal adjustments for missing values and data treatments. Theoretically, the Random Forest methodology exhibits a higher level of accuracy and stability and handles non-linear parameters more efficiently than other approaches. Additionally, hyperparameter optimization prevents the issue with random forests overfitting. Random grid search is used for hyperparameter tuning; it controls the maximum depth of the sample data drawn for training each tree and involves stopping criterion which reduces the computation time. Hence, the usage of Random Forest for the OH Switch model is deemed to be fit.

2.4 Assumptions

The key business assumptions that were considered during model development are specified below:

BA 01: There is no change in OH Switch technical specification over time. The model assumes the type of OH switches used in the model building process have the same characteristics in terms of build and quality. For example, each switch asset type generally has its unique technical specifications, like physical attachments and mounting direction, that are expected to remain the same over time.

BA 02: The Calibration model assumes that fires are a subset of failures. Failures prompting the need for removals, replacements, and repairs are the representative failure targets used in place of few ignition events. Some of these issues left unaddressed can potentially spark an ignition, but not all failures will result in a fire. Hence, fire can be treated as a subset of failure.

BA 03: The model is designed to work in both base weather and extreme weather conditions. The weather variables incorporated in the model are represented as various statistical aggregations like max, mean, and

standard deviation on wind, wind speed, humidity, rain, and snow. Hence the model results can be used under both base weather and extreme weather conditions.

The functional/model methodology assumptions that were considered during model development are discussed in detail below:

MA 01: The feature variables in the dataset should have some actual values so that the classifier model can predict accurate results. In an ideal scenario, all the variables would not have estimated values and they would instead use actual values. The current model is able to provide accurate results even after using estimates as they are derived through imputation using actual values from other variables. Example: Inferring manufacturers based on switch subtype.

MA 02: The predictions from each tree must have very low correlations. It is difficult to differentiate between a real interaction effect, marginal effects, and random variations in random forests. Hence, the presence of highly correlated variables in the Random Forest approach will have an impact on its ability to identify strong predictors.

2.5 Limitations and Compensating Controls

The key model limitations that would impact the accuracy and performance of the model are discussed in detail below:

Limitation ID: L01

Limitation Title: Unavailability of linear/non-linear representation in the form of intuitive equation or correlation statistic.

Description: The Random Forest algorithm does not explain any linear or non-linear relationship in the form of an intuitive equation or correlation statistic to enable measurement of the scalability of impact of independent variables on the dependent variable.

Compensating Controls: The Random Forest model is considered a black box as it is difficult to understand the relationship between independent and dependent variables and how the independent variables influence the predictions. Since black box is a common limitation with most ML algorithms, usage of the model is considered appropriate as it provides better AUC results than other models.

Limitation ID: L02

Limitation Title: Time consumption for model execution is high

Description: Since Random Forest models use a bagging algorithm, they can provide more accurate predictions but slow down the process as they compute data for each decision tree.

Compensating Controls: To overcome the time consumption issues from grid search computations, random grid search is used in the hyperparameter tuning process. Random grid search is a proven technique to reduce the time consumption when testing multiple models with different combinations of hyperparameters by using stopping criterion like tolerance, maximum rounds, maximum run time, and performance improvement thresholds. It moves within the grid in a random fashion to find the best set of hyperparameters to achieve maximum performance in terms of the metric specified, here AUC. Since the model is not executed through computer program automatically at a defined frequency and is instead run only once a year manually, usage of the model is considered appropriate.

Limitation ID: L03

Limitation Title: Resource utilization for model execution is high

Description: Since Random Forest models process many decision trees, they need more resources with respect to system configuration and system capacity to store that data.

Compensating Controls: The resource utilization factor will have a major impact for real time models as they would run more frequently. Since the OH Switch model is run only once a year with reasonable use cases, the impact of resource utilization is low. Additionally, the usage of random grid search and stopping criterion like tolerance, maximum rounds, maximum run time, and performance improvement thresholds provide more control on the number of recurring instances run to identify the best fit hyperparameters to achieve optimal AUC. Since the model is not executed through computer program automatically at a defined frequency and is instead run only once a year manually, usage of the model is considered appropriate.

Limitation ID: L04

Limitation Title: Model accuracy might reduce if the dataset experiences covariate shift.

Description: Covariate shift is a type of model drift which occurs when the distribution of independent variables changes between the training environment and live/test environment. Since the Random Forest cannot extrapolate (i.e., predict outside the training space), the model performance might decrease if there is covariate shift in the dataset.

Compensating Controls: The covariate shift affects most machine learning models to some degree, as test data is never going to be the same as training data. Detecting and addressing covariate shift is therefore a key step to the machine learning process. The current model is run only once a year along with data refresh. It uses a random sampling mechanism to split the dataset into train (80%) and test (20%) data whenever it is run. The usage of random sampling mechanism is considered to resolve the issue of covariate drift and maintains the accuracy of the model results. Hence the usage of the Random Forest methodology along with the random sampling mechanism to split train/test data is considered appropriate.

2.6 Model Outputs

The OH Switch model predicts the probability of ignition (POI) arising from equipment (switch) failure. The model has a single output characterized by a continuous number between 0 and 1 for each OH Switch asset.

The probabilities across different asset failure predictive models cannot be aggregated or compared and hence are calibrated to derive frequencies of ignition. The sum of the resulting frequencies of ignition for a sub-driver equals the total expected ignitions for the specified year.

$$\text{Frequency of Ignition} = \text{Probability of Ignition} \times \frac{\text{Calibrated Targets}}{\sum \text{Probability of Ignition}}$$

where Calibrated Targets = Forecasted Ignitions for that sub-driver

The output from this calibration exhibits the following features:

- Frequency: Each value can be specified as the frequency of fires per year.
- Comparability: The frequencies are comparable against sub-drivers and models.
- Additivity: The frequencies can be added across models to derive the aggregated fire forecast in a year.

This is achieved by forecasting fires by sub-driver and using these forecasts to weight the model probabilities. The sum of probabilities from each calibrated model equals the forecast by sub-driver.

Figure 7 provides the calibration steps that are performed using the failure probability results from the OH Switch model. This methodology followed in the calibration model is provided below:

- I. Aggregate the probability output from each sub-driver model.
- J. Based on the forecast logic selected, find the forecast results (expected fires) for each sub-driver.
- K. Generate the calibration factor for each sub-driver based on the values calculated in the above steps (B/A).
- L. Multiply each model probability by its calibration factor to arrive at the estimated frequency of fires from each sub-driver.

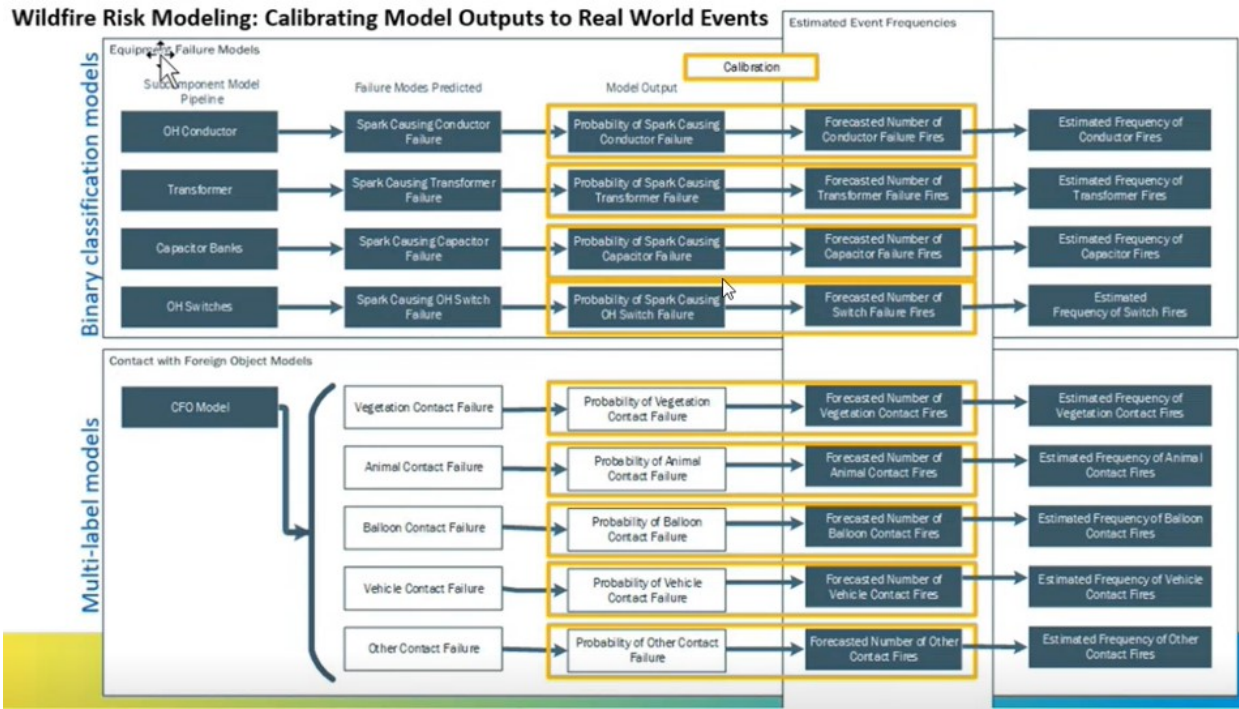


Figure 7: Calibration model schema

This estimated frequency of fires from each sub-driver can be added across the models to derive the expected frequency of ignition for each location.

The calibrated probabilities, frequencies of events, based on the output from OH Switch model is the data ingested to inform the programs mentioned in Section 1.1.

Model Changes:

The OH Switch model was enhanced with the following changes in June 2022 as part of the annual refresh:

- Replaced weather features to incorporate more granular (hourly, 2km x 2km gridded) ADS weather data; was previously referring to weather station observations that generalized larger territorial regions
- Filled in unknown manufacturer names by using vendor numbers and by referring to switch subtype
- Expanded failure dataset to include repairs and replacement notifications in addition to removals that were initially tracked

The impact on the model performance from the change in features can be determined by comparing the AUC of the model before (Sept 2020) and after (June 2022) refresh (Figure 8). The AUC value of the OH Switch model was 0.73 before refresh and 0.85 after refresh.

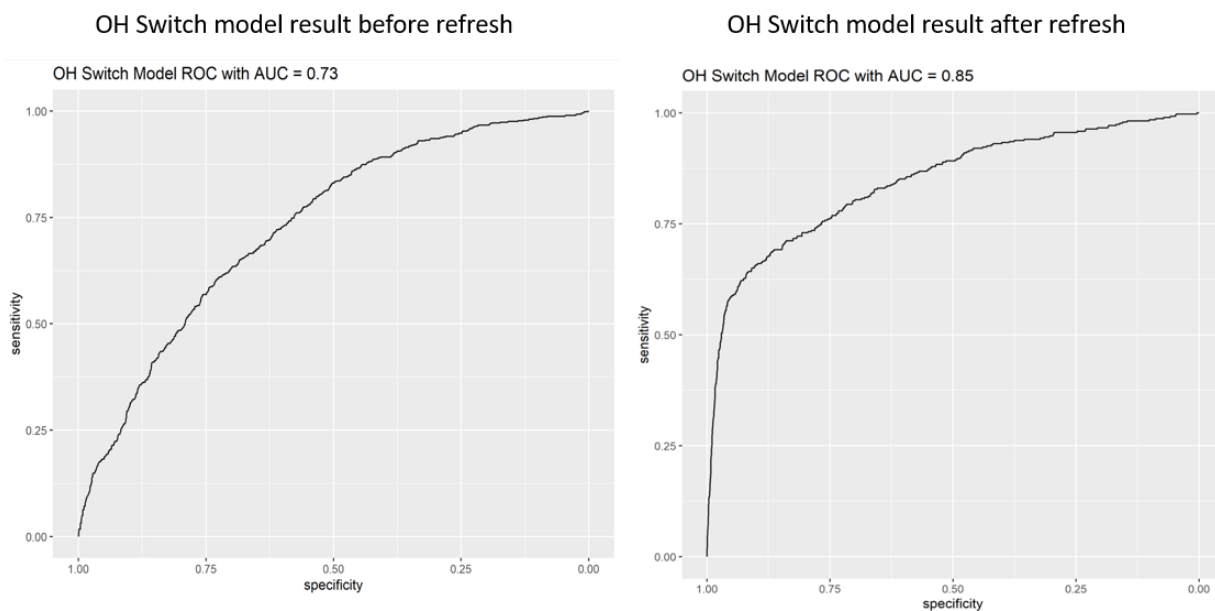


Figure 8: AUC results for OH Switch model (before and after refresh)

3. MODEL PERFORMANCE AND TESTING

For each machine learning model developed, SCE tries to select the best algorithm based on the model train/test performance, which can be measured by Area Under the Curve (AUC) and other metrics from the Confusion Matrix.

3.1 Model Specification Testing

The model is developed and tested in R programming using library h2o. The model is run once a year manually by Data Science and Asset Analytics team. The model is calibrated every year with the last 5 years of historical failure data.

The verification of the model implementation is performed by checking the variable importance results which provides the list of features implemented. The performance of the model is validated through the AUC, defined in Section 2.2 and provided in Section 3.3.

The validity and impact of the Model Assumptions, mentioned in Section 2.4, are discussed below:

- The features used in the model are expected to have some actual values so that the model results can be accurate. In an ideal scenario, all the variables would not have estimated values and they would instead use actual values. Some features like manufacturer do not have the actual values in all scenarios so values are imputed for this feature with help of information in other variables like switch subtype. After using these estimates, the data quality is enhanced to support reliability of the current model in terms of improved predictive accuracy.
- Random Forest is considered a strong approach for variable selection in high-dimensional data only when the variables have low correlation. The recursive structure of trees generally enables them to take dependencies into account in a hierarchical manner. However, some variable combinations without clear marginal effects might make the tree algorithm ineffective. To conclude, it is difficult to differentiate between a real interaction effect, marginal effects, and just random variations in random forests. Hence, the presence of highly correlated variables in Random Forest approach will have an impact on its ability to identify strong predictors. Adequate measures are taken to filter out highly correlated features to overcome their impact in predicting the results.

Model Estimation:

The OH Switch model employs a number of independent variables. Section 2.1 contains a list of the independent variables utilized in this model.

The variable importance test results for the OH Switch model, Figure 9, shows the order of which features provided the most information gain in informing the correct prediction of failure or non-failure. The variable importance features test estimates the relative influence of each variable by calculating whether that variable was chosen to split during the tree building process and how much the squared error (over all trees) improved (or decreased) as a result.

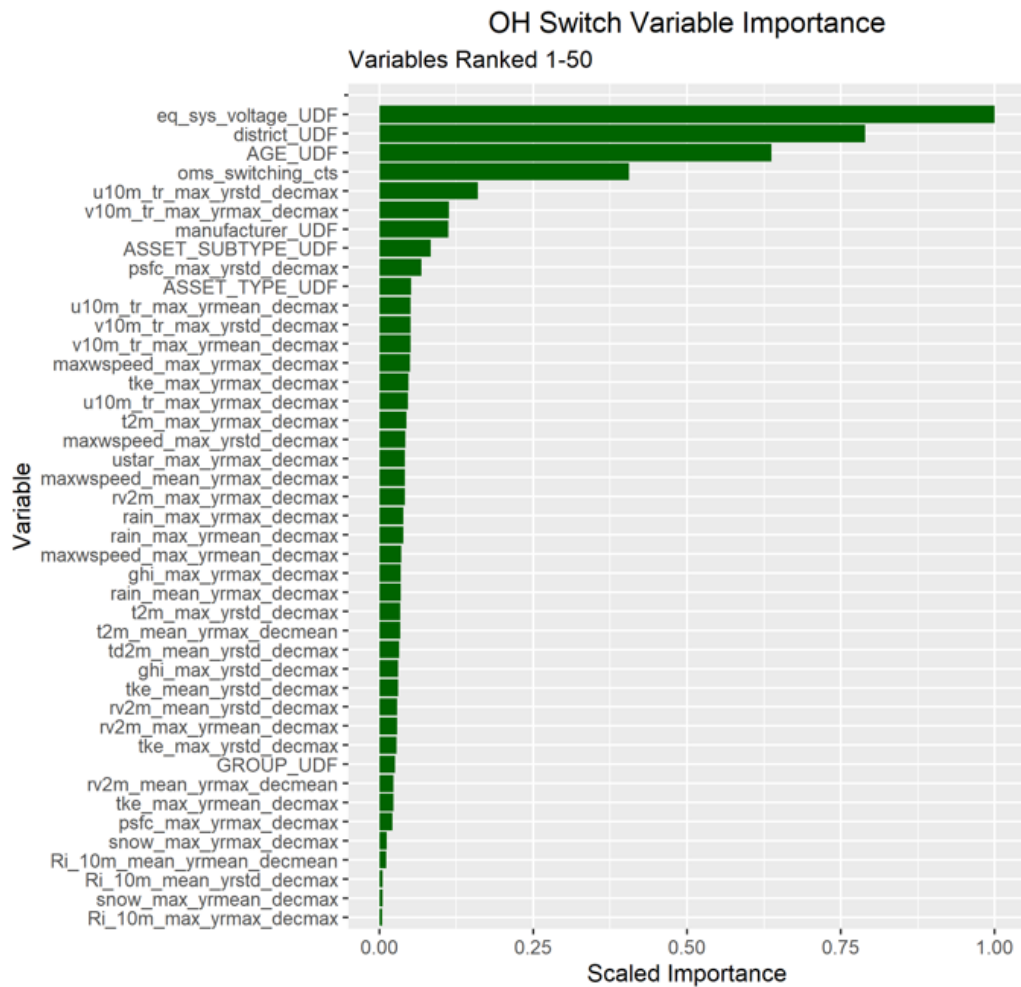


Figure 139: Variable Importance test results for the OH Switch model

The results confirm that the features eq_sys_voltage_UDF, district_UDF, and AGE_UDF exhibit high importance on the OH Switch model output.

Cross-references: Refer to link [\[RF 3\]](#) in Section 0 for description on the methodology used to perform the Variable Importance for tree-based methods.

The OH Switch model uses the random grid search approach for hyperparameter optimization to select the best set of hyperparameters to achieve maximum performance in terms of AUC as described in Section 2.2. Once the grid search is completed, a list of models with their associated hyperparameter values is obtained as shown in Figure 10. The acquired models are then sorted based on the AUC values for the OH Switch model. The model with the highest AUC value is regarded the best fitted model. Figure 11 shows the best model obtained for the OH Switch model. The best model is run on the respective test data, and the AUC metric is used to evaluate the model performance.

The AUC is used to estimate the model discriminatory power to predict the results in a binary classification problem. A higher AUC means the model can predict the results more accurately. Figure 12 shows the ROC with AUC for the OH Switch model based out of test dataset ran with the last 5 years of historical failure data (2017-2021) (results derived in June 2022 R scripts re-rerun). The AUC value for the OH Switch model is 0.85.

```

Grid ID: Grid_DRF_TRAIN_3Y.hex_model_R_1678475465188_1
Used hyper parameters:
- max_depth
- min_rows
- mtries
- ntries
- sample_rate
Number of models: 85
Number of failed models: 0

Hyper-Parameter Search Summary: ordered by increasing logloss
max_depth min_rows mtries ntries sample_rate model_ids logloss
1 30 15.0 11 200 0.7 Grid_DRF_TRAIN_3Y.hex_model_R_1678475465188_1_model_42 0.16447639758742755
2 35 15.0 17 150 0.5 Grid_DRF_TRAIN_3Y.hex_model_R_1678475465188_1_model_40 0.16451271762018357
3 50 15.0 11 400 0.5 Grid_DRF_TRAIN_3Y.hex_model_R_1678475465188_1_model_46 0.1646499215501404
4 50 15.0 11 150 0.5 Grid_DRF_TRAIN_3Y.hex_model_R_1678475465188_1_model_61 0.1648396021479463
5 25 15.0 14 350 0.632 Grid_DRF_TRAIN_3Y.hex_model_R_1678475465188_1_model_9 0.1650984652330101

---
max_depth min_rows mtries ntries sample_rate model_ids logloss
80 35 2.0 17 100 0.8 Grid_DRF_TRAIN_3Y.hex_model_R_1678475465188_1_model_3 0.24097288268967174
81 40 1.0 11 300 0.632 Grid_DRF_TRAIN_3Y.hex_model_R_1678475465188_1_model_39 0.24369984970761074
82 35 2.0 17 100 0.7 Grid_DRF_TRAIN_3Y.hex_model_R_1678475465188_1_model_10 0.250858114565438
83 50 1.0 11 250 0.8 Grid_DRF_TRAIN_3Y.hex_model_R_1678475465188_1_model_31 0.2741770196820364
84 50 1.0 14 200 0.8 Grid_DRF_TRAIN_3Y.hex_model_R_1678475465188_1_model_17 0.2799627250849247
85 40 1.0 17 100 0.7 Grid_DRF_TRAIN_3Y.hex_model_R_1678475465188_1_model_45 0.33986253752057427

```

Figure 10: List of models with their associated hyperparameter values produced after the grid search for the OH Switch model

```

Model Details:
=====
H2OBinomialModel: drf
Model ID: Grid_DRF_TRAIN_3Y.hex_model_R_1654078825750_7501_model_29
Model Summary:
number_of_trees number_of_internal_trees model_size_in_bytes min_depth max_depth mean_depth min_leaves max_leaves mean_leaves
1 250 250 892324 15 30 22.92000 92 300 269.40000

```

Figure 14: The best model selected for the OH Switch model, along with the related hyperparameter values.

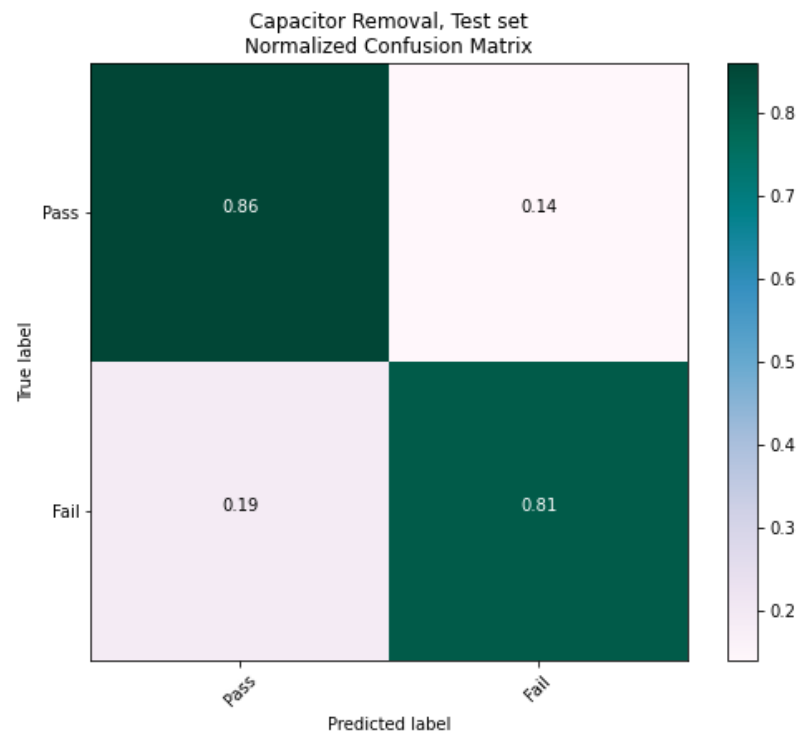
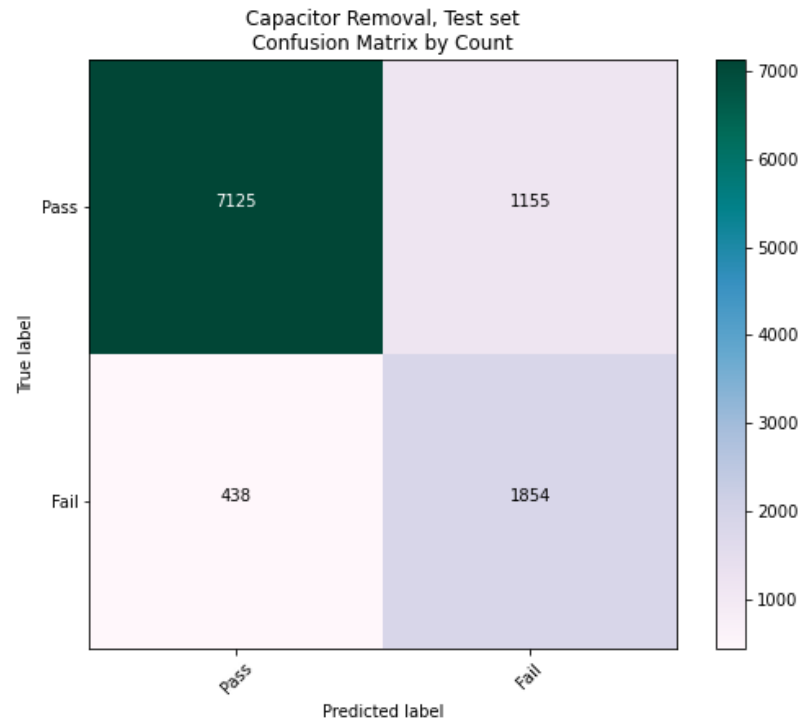
In terms of model convergence, the random grid search for hyperparameter tuning uses a stopping criterion based on a specified tolerance in AUC. This means that the additional efforts involved in hyperparameters tuning and training is not likely to improve the model performance beyond the specified threshold.

The accuracy of the model prediction, in addition to AUC, can be determined using the Confusion Matrix and Classification Error Rate results.

- A confusion matrix presents a tabular layout of the different outcomes of the prediction results of a classification problem and helps visualize its outcomes. It generates a table of all the predicted and actual values of a classifier model.

OH Switch - Confusion Matrix Results				
Actuals	Predicted			
		0	1	Error Rate
	0	5523	202	0.035284
	1	177	212	0.455013
		5700	414	0.061989

- Table 9: Confusion matrix results for OH Switch modelA confusion matrix presents a tabular layout of the different outcomes of the prediction results of a classification problem and helps visualize its outcomes. It generates a table of all the predicted and actual values of a classifier model.



- provides the confusion matrix results for OH Switch model. It captures the accuracy rate as 93.8%.
- Classification error rate is used to estimate the proportion of instances misclassified over the whole set of instances. It is estimated using the below formula.

$$Error\ Rate = \frac{False\ Positives + False\ Negatives}{True\ Positives + True\ Negatives + False\ Positives + False\ Negatives} * 100$$

The error rate for OH Switch model turns out to be 6.2%. This means that the failure rate of the model prediction is very low and under control.

All these test results are performed on test dataset with historical failure information between 2017-2021.

A detailed assessment of the model limitations and associated compensating controls are available in Section 2.5.

3.2 Sensitivity Analysis

Sensitivity analysis examines the impact of each feature on the model's predictions. It is a simple yet powerful technique to analyze a machine learning model. To determine the sensitivity of a feature, its value is changed while the values of all other features are held constant. The model's output is then examined. If the outcome of the model significantly changes when the feature value is changed, this indicates that the feature has a significant influence on the prediction. Based on the variable importance feature list shown in Figure 9, the top five continuous variables of the OH Switch model were chosen to perform the sensitivity analysis. Additionally, two categorical variables of the OH Switch model were also considered for the analysis based on the suggestion provided by the business function.

The feature variables used for sensitivity analysis of the OH Switch model:

- eq_sys_voltage_UDF
- AGE_UDF
- oms_switching_cts
- u10m_tr_max_yrstrd_decmax
- v10m_tr_max_yrmax_decmax

The sensitivity of the model is examined using the test dataset, which contains 20% (6114 observations) of the entire processed data. For this analysis, 10% of the test dataset (607 observations) was selected and modified using extreme values. Stratified sampling was used to select 10% of the test data to add randomization to eliminate sampling bias. Sensitivity tests were performed in June 2022 using the last 5 years of historical failure data.

To set up the strata, the categorical features ASSET_TYPE_UDF and ASSET_SUBTYPE_UDF were selected from the test dataset. Based on these variables, eight different strata were created. The test data is bound to a column of random numbers produced using a standard normal distribution, and the rank of these random numbers is used to sort the entire set of test data. The top 10% from each stratum was selected as the target observations to modify the input data. To test the sensitivity of a feature, the values of the selected observations were altered with extreme values (minimum and maximum) of the feature. As a result, for each feature, two sets of test data were generated for sensitivity analysis. Table 2 provides the extreme values (determined by historical data) used for each variable for the sensitivity analysis.

Extreme values used for Sensitivity testing in OH Switch model		
Variables	Maximum Value	Minimum Value
eq_sys_voltage_UDF	66	2.4
AGE_UDF	101	0
oms_switching_cts	782	0
u10m_tr_max_yrmax_decmax	25.85	4.55
v10m_tr_max_yrmax_decmax	25.93	3.95

Table 2: Extreme values used for Sensitivity testing in OH Switch model

Table 3 provides the AUC results of the unaltered test data, i.e., test data without changing the variables' values, and of the various sensitivity tests that were performed. For the OH Switch model, the difference in AUC values between the sensitivity tests and the unaltered test data results do not exceed 2% which means that the variations in the input values for these variables do not have a huge impact on the results in terms of AUC.

AUC result of Unaltered Test Data from OH Switch	0.8463
--	--------

OH Switch model Results				
Feature	Maximum value scenario		Minimum value scenario	
	AUC	% Decline in AUC Compared with Unaltered Test Data	AUC	% Decline in AUC Compared with Unaltered Test Data
eq_sys_voltage_UDF	0.8529	0.78%	0.8539	0.90%
AGE_UDF	0.8481	0.21%	0.8483	0.24%
oms_switching_cts	0.8449	-0.17%	0.8562	1.17%
u10m_tr_max_yrmax_decmax	0.8527	0.76%	0.8533	0.83%
v10m_tr_max_yrmax_decmax	0.8505	0.50%	0.8525	0.73%

Table 3: The sensitivity results based on AUC for OH Switch model

Table 4 provides the True Positive Rate (TPR) for the unaltered test data and the various sensitivity tests determined using the prediction output provided by the OH Switch model. The increase and decrease in TPR among different tests can be observed from the results but the difference in values seems to be very low. Table 4 also provides the changes in True Positives, False Positives, True Negatives, and False Negatives. The change in the predicted outcomes seems to be low when compared with the count of observations (607) that were altered for performing this test.

True Positive rate and False Positive rate for OH Switch model						
	TP	FP	TN	FN	TPR	FPR
Unaltered Test Data	212	202	5523	177	54.50%	3.53%

True Positive rate and False Positive rate for OH Switch model						
Sensitivity Test	TP Change	FP Change	TN Change	FN Change	TPR	FPR
Max Value for eq_sys_voltage_UDF	-11	-1	1	11	57.33%	3.55%
Min Value for eq_sys_voltage_UDF	-11	0	0	11	57.33%	3.53%
Max Value for AGE_UDF	-11	1	-1	11	57.33%	3.51%
Min Value for AGE_UDF	-11	-1	1	11	57.33%	3.55%
Max Value for oms_switching_cts	-14	-9	9	14	58.10%	3.69%
Min Value for oms_switching_cts	-11	-1	1	11	57.33%	3.55%
Max Value for u10m_tr_max_yrmax_decmax	-11	-1	1	11	57.33%	3.55%
Min Value for u10m_tr_max_yrmax_decmax	-11	-1	1	11	57.33%	3.55%
Max Value for v10m_tr_max_yrmax_decmax	-3	22	-22	3	57.27%	3.14%
Min Value for v10m_tr_max_yrmax_decmax	-11	0	0	11	57.33%	3.53%

Table 4: The sensitivity results based on predicted outcome for OH Switch model

Based on these test results, it can be determined that the variations in the input values for the high importance features alter some of the predictions from the model, but the magnitude of the impact seems to be low. Hence the model results from the OH Switch model are robust and reliable post sensitivity testing for the variables defined in this section earlier with extremely high and low values tested for each of the defined variables.

3.3 Outcome Analysis / Backtesting

The subset of historical data on which a model is trained and optimized is referred to as the in-sample data. On the other hand, the subset of the dataset that has been reserved to test the model is known as the out-of-sample data. The OH Switch model uses a random sampling approach to split the dataset into Train (80%) and Test (20%) data. The results arrived from train data are considered as in-sample backtesting and the results arrived from test data are considered as out-of-sample backtesting.

Once the machine learning model is built with the training data, it is evaluated using a separate test dataset that has not yet been studied. The performance of the model is determined by the Area Under the ROC Curve (AUC) value. Figure shows the AUC value and ROC for the OH Switch model based on the test dataset ran with historical failure information between 2017-2021. The AUC value of 0.85 implies that the model possesses moderately high accuracy in terms of predicting the results.

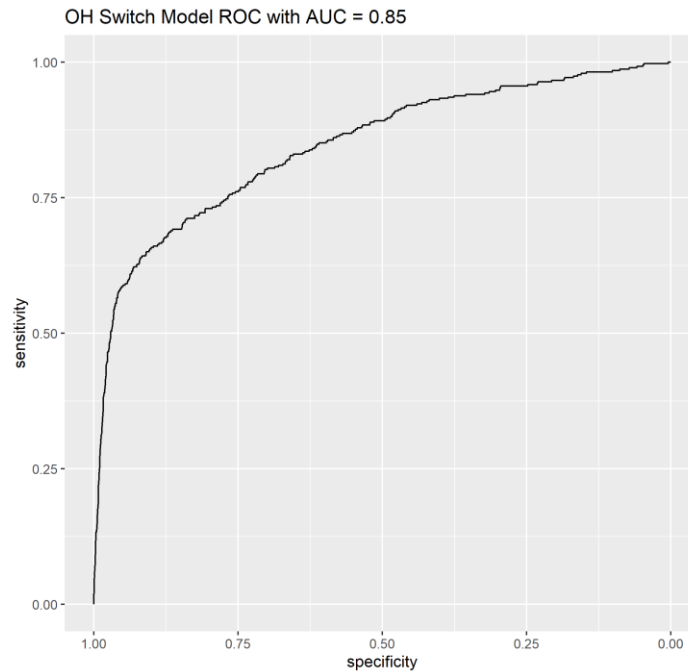


Figure 12: Out-sample backtesting result for OH Switch model based on test dataset

The impact of uncertainty in model inputs and parameters on model outputs are tested as a part of the sensitivity analysis and the results are captured in Section 3.2. In addition, the data imputations that are incorporated to address missing values before running the model are defined in Section 2.1.

3.4 Benchmarking Analysis

For the OH Switch model, different approaches like Gradient Boosting Machine (GBM) learning and Random Forest were considered during the model development phase in 2017. The analysis on these supervised machine learning approaches and the results are provided below.

- Gradient Boosting Machine (GBM) is one of the most popular forward learning ensemble methods in machine learning. It is a powerful technique for building predictive models for classification and regression tasks. GBM sequentially combines the predictions from various weak learner decision trees and builds a final predictive model with more accurate predictions by minimizing a defined loss function.
- Random Forest is a popular machine learning algorithm that can be used for both classification and regression problems. Random Forest is another ensemble method that combines the predictions of several decision trees to improve the predictive accuracy of the model. The individual decision trees are created based on a randomly selected subset of features at each node prior to determining the optimal split so each tree differs. The final output is determined by taking the majority vote of the predictions from the individual decision trees. The greater number of trees in the forest generally leads to higher accuracy and prevents the problem of overfitting.

The benchmarking results of GBM shared in this section were developed using the h2o library in R on the Test data with targets based on the last 5 years of historical failure data (2017-2021). Since benchmark results were not saved during the model development phase, the benchmark models were executed in June 2022 for documentation purposes. Figure 13 provides the AUC values for the OH Switch model using the GBM and Random Forest methodologies.

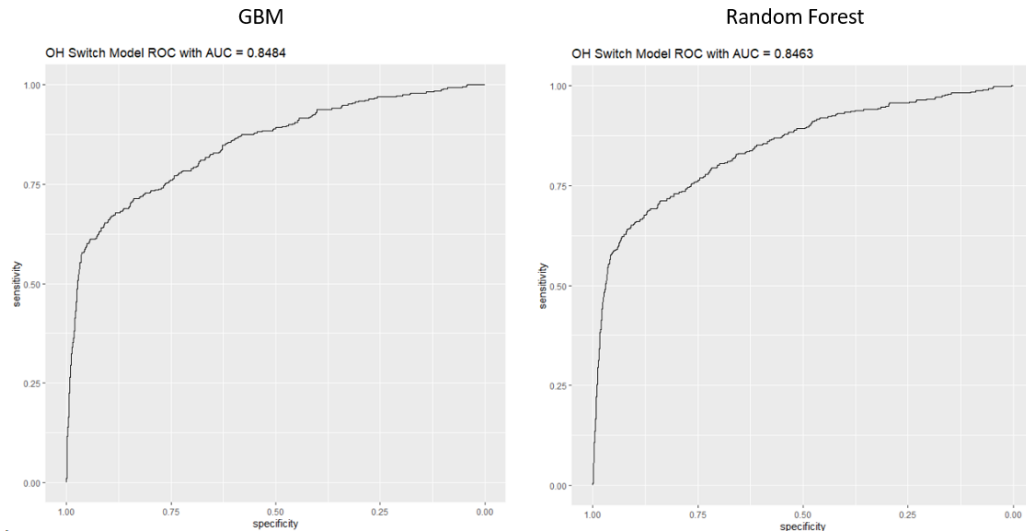


Figure 13: AUC Comparison for the OH Switch sub-model using GBM and Random Forest methodologies

For the OH Switch model, the AUC results for GBM and Random Forest were 0.8484 and 0.8463 respectively. Random Forest was chosen as the modeling algorithm for the OH Switch model as it aligns with the modeling approach for SCE’s other predictive asset failure models, and it achieved relatively the same AUC as GBM. GBM is a methodology that will continue to be considered for use as part of the annual refresh of the model. Some additional advantages of using Random Forest over GBM are provided below:

- Random Forest is less sensitive to overfitting issues than GBM.
- Hyperparameter tuning is relatively easy in Random Forest when compared with GBM.

4. MODEL MANAGEMENT AND GOVERNANCE

4.1 Ongoing Monitoring Plan

Ongoing monitoring is important for Machine Learning models especially when they are used to make predictions or when they are run on datasets with high volatility in variable values. The OH Switch model is run manually once a year, incorporating updated input datasets to reflect the latest available data and implementing any specific model enhancements—e.g., inclusion / replacement / removal of a feature, optimization of the code, evaluation of a new performance metric, etc. During the model refresh, the limitations and assumptions of the model are also revisited by the model developers and necessary action items are conducted to address them.

Performance monitoring is required only after running the model. Recalibration of the model has not been performed for the last two years, and it is performed only if the behavior of the model differs from that of the previous version or if there is a significant drop in model performance. The AUC and accuracy rate from confusion matrix results obtained after model refresh are compared against a threshold of 80%; if the value drops below this threshold, the reason behind the performance dip is investigated. Post-investigation, the steps required to improve the performance of the model will be carried out. To monitor the model performance more thoroughly, developers of the model plan to additionally evaluate metrics like Precision and Recall. Precision is the positive predictive value which represents the proportion of predicted failures that were predicted correctly. Recall is the true positive rate

which represents the proportion of actual failures that were predicted correctly.

The model documentation and the performance results are updated once a year immediately after the model refresh.

4.2 Security and Control

The Data Science and Asset Analytics team has access to the data inputs, code, and implementation for the model. Other business units, like the Grid Hardening Strategy team, are provided access to the model outputs upon request but cannot update or modify the code.

The model is run using R programming and it can be executed in any recent versions of the R software. Current model versioning is labeled by date of refresh (e.g., OH_Switches\update_20220610). There are plans to move the code to GitHub, a platform that facilitates version control by tracking changes to the source code. Users with write or admin privileges to the repository can review proposed changes and approve them.

A contingency plan is not applicable for this model as it is an inhouse model for SCE.

5. REFERENCES

RF 1: SCE's WMP 2022 Q1 Quarterly Data Report submission

<https://www.sce.com/sites/default/files/AEM/Data%20Requests/2022/SCE%20Q1%202022%20Tables%201-12.xlsx>

RF 2: Literature reference on grid search vs random search approach for hyperparameter tuning

<https://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf>

RF 3: Variable Importance methodology for tree-based methods

[*Variable Importance — H2O 3.38.0.3 documentation*](#)

Attachment D

OH-Transformer Sub-Model

**Southern California Edison (SCE)
Model Documentation
Prepared for 2023 WMP Appendix B**

OH Transformer Sub-Model

3/27/23

Table of Contents

1. EXECUTIVE SUMMARY	4
1.1 Model Purpose and Intended Use	4
1.2 Model Description Summary	5
1.3 Model Risk Rating.....	5
1.4 Model Dependency and Interconnectivity	6
1.5 Model Assumptions	6
1.6 Model Limitations	6
1.7 Overall Model Performance Assessment	6
1.8 Contingency Plan for Vendor Model	6
2. MODEL FRAMEWORK AND THEORY	8
2.1 Model Inputs and Data Quality.....	8
2.2 Methodology	10
2.3 Suitability.....	13
2.4 Assumptions	13
2.5 Limitations and Compensating Controls	13
2.6 Model Outputs	14
3. MODEL PERFORMANCE AND TESTING	17
3.1 Model Specification Testing.....	17
3.2 Sensitivity Analysis	20
3.3 Outcome Analysis / Backtesting	22
3.4 Benchmarking Analysis	22
4. MODEL MANAGEMENT AND GOVERNANCE	23
4.1 Ongoing Monitoring Plan	23
4.2 Security and Control.....	24
5. REFERENCES.....	25

1. EXECUTIVE SUMMARY

1.1 Model Purpose and Intended Use

The OH (Over Head) Transformer Model is a Probability of Ignition (POI) Sub-Model developed by SCE (Southern California Edison). At SCE, models are developed using ML (Machine Learning) algorithms for each asset—in this case OH transformer. The OH Transformer model is refreshed annually and used to predict the probability of failure (POF) at distribution overhead transformers.

The calibrated outputs of the OH Transformer model—i.e., failure events—are used by two programs described below:

10. The Inspections and Remediations programs, which considers POI as an element in prioritization and scoping.
11. Risk analyses via SCE's MARS Framework.

1.2 Model Description Summary

The OH Transformer model is a binary classification model using Random Forest—a Machine Learning technique. It predicts the probability of a transformer igniting a spark due to equipment failure by considering available transformer attributes and condition data (i.e., age, voltage, etc.) and other environmental and operational attributes (i.e., historical weather, loading, etc.).

The model is implemented in R programming using the h2o library and is connected to databases such as SAP, Weather, etc. The model is run once a year manually by the Data Science and Asset Analytics team. The model is calibrated every year with the full historical outage data.

Cross-references Please refer to Section 2.1 for more information about the inputs used by the OH Transformer model along with data processing details.

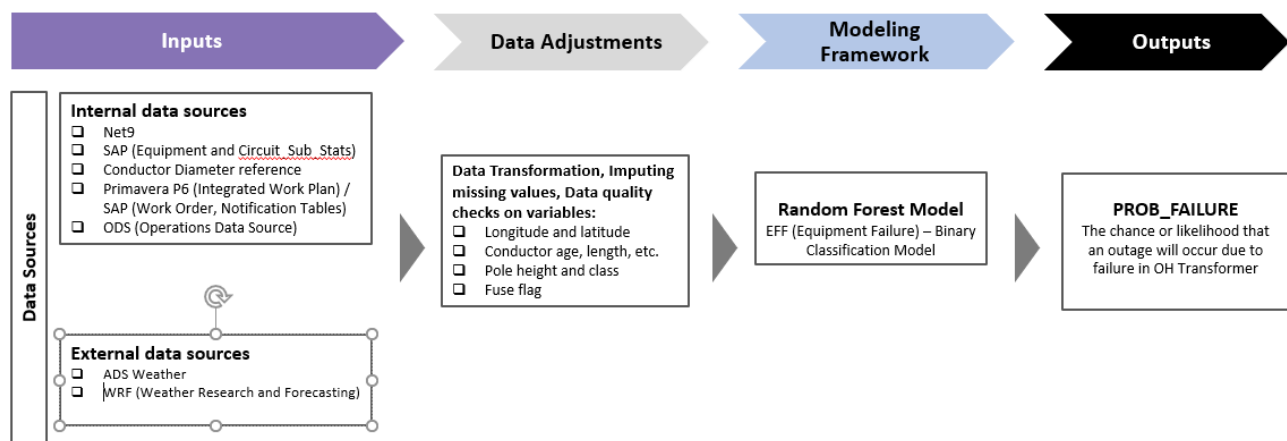


Figure 15: OH Transformer model framework

The OH Transformer model uses the Random Forest methodology. Since the Random Forest methodology can perform both classification and regression tasks, it is considered a viable choice for the OH Transformer model as the prediction is a classified event, i.e., failure. This methodology predicts output with high accuracy, runs efficiently on large datasets, and maintains accuracy with minimal adjustments for missing values and data treatments.

1.3 Model Risk Rating

There is no defined mechanism of identifying model risk rating at SCE, however certain factors—like frequency of risk events and use case—are considered while flagging model risk. Based on the Wildfire Mitigation Plan quarterly report, the frequency of transformer failures in a year from transformers averages 2311 which is medium compared to other sub-drivers. Figure 2 provides a snapshot of the count of transformer failures over the years by the causes captured in the OH transformer model. In addition, the output of this model is considered important as it informs the strategy of a few programs which are discussed in section 1.1. Hence, the OH Transformer model is deemed to be high risk model.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC
				Number of risk events																Projected risk events								
Risk Event category	Cause category	#	Sub-cause	Are risk events	2015	2016	2017	2018	2019	2020	2020	2020	2020	2021	2021	2021	2021	2022	2022	2022	2022	2023	2023	2023	2023	Unit(s)	Comments	
Outage - Distribution	18.n.	Transformer damage or failure - Distribution	Yes																			522	603	1029	537	# risk events (excluding ignitions)		
				1889	1649	1978	2594	2489	416	559	1890	536	403	547	724	501	288	613	1053	556								

Figure 16: Key recent and projected risk events due to transformer damage or failure from SCE Q1 2022 Quarterly Data Report, Table 7.1

Cross-references: Refer to link [RF 1: SCE's WMP 2022 Q1 Quarterly Data Report submission

] in Section 5 for SCE's Wildfire Mitigation Plan Q1 2022 Quarterly Data Report submission.

1.4 Model Dependency and Interconnectivity

The OH Transformer model is an "Ignition Likelihood" model which uses the inputs from weather stations along with other data sources to calculate the probability of ignition.

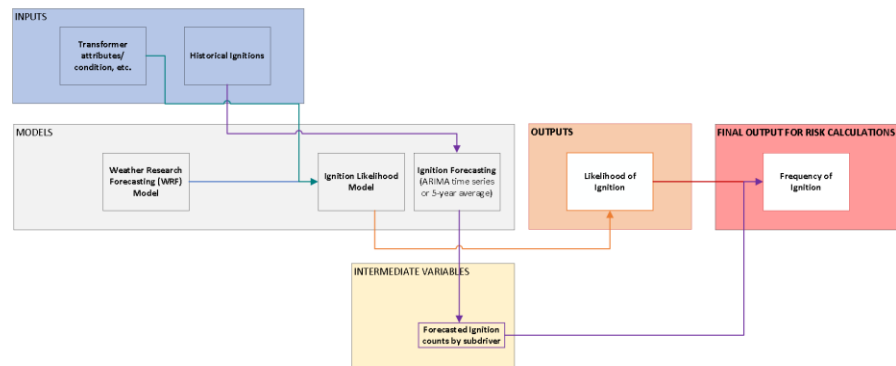


Figure 17: Model Interconnectivity Schema

The OH Transformer model uses weather data, and historical attributes of transformers as input. The output data from the OH Transformer model, i.e., POI, is used by two categories of programs, further discussed in Section 1.1, to inform their strategic decisions.

1.5 Model Assumptions

The business assumptions and model assumptions for the OH Transformer model is summarized below:

17. There is no change in the OH Transformer technical specification over time.
18. The calibration methodology assumes that fires are a subset of failures.
19. The model is designed to work in both base weather and extreme weather conditions.
20. The feature variables in the dataset should have some actual values so that the classifier model can predict accurate results.
21. The predictions from each tree must have very low correlations.

A detailed explanation of these assumptions is available in Section 2.4.

1.6 Model Limitations

The model limitations for the Transformer model are summarized below:

13. Unavailability of linear/non-linear representation in the form of intuitive equation or correlation statistic.
14. Resource utilization in terms of system capacity and higher configuration for model execution is high.
15. Model accuracy may reduce if the dataset experiences covariate shift.

A detailed explanation of these data limitations is available in Section 2.5.

1.7 Overall Model Performance Assessment

The machine learning model used to build the OH Transformer model is the Random Forest algorithm. The model's overall performance is determined by the Area Under the ROC Curve (AUC) value and Confusion Matrix results. The performance of the OH Transformer model was evaluated on test data with transformer replacement data.

- The AUC value is 0.8146.
- Confusion matrix results capture the accuracy rate as 81.3%.

The above metrics were derived at the time of the model refresh in July 2022 to capture an exhaustive set of statistical results for documentation purposes.

1.8 Contingency Plan for Vendor Model

A contingency plan is not applicable for this model as it is an inhouse SCE model. This is not a vendor model.

2. MODEL FRAMEWORK AND THEORY

The Transformer model is a binary classification model pertaining to equipment failures. The model employs a random forest algorithm to predict the likelihood of a transformer experiencing an ignition event due to transformer failure. The random forest approach was chosen for the classification task over other modeling approaches—such as logistic regression, gradient boosting, etc.—because it predicts output with high accuracy, runs efficiently on large datasets, and maintains accuracy with minimal adjustments for missing values and data treatments.

2.1 Model Inputs and Data Quality

Data Sources

This model refers to multiple internal and external data sources. The internal data sources used by the model are:

- **SAP** houses circuit⁶, structure, and equipment characteristics. It contains latitude and longitude information of the assets which is used to determine the location of the segment by considering the midpoint between all the structures associated to the segment. SAP also provides features like base and height of the pole which is consumed by the model.
- **Atmospheric Corrosivity shape file** is used to fetch the Atmospheric corrosivity intensity in the segment.
- **AMI** smart meter readings in HADOOP are used for loading information on the transformers.

The external data sources used by the model are:

- **Weather Research and Forecasting Model (WRF)** is an open-source external model used to fetch information about these wind factors. Downforce is the perpendicular force applied on the wires because of wind which is termed as the wind factor. The logic on how the downforce calculations were derived is attached below.

Quality Checks

SCE has internal data management teams for ensuring data quality, including EAD (Enterprise Asset Data) and Master Data. They work on processing asset data corrections (E2 notifications) in SAP and fixing largely known data issues like missing or erroneous latitude and longitude information for assets in their territory. Some of the data quality checks that are performed in the OH Transformer model to ensure the accuracy, validity, integrity, and consistency are provided below. Quality checks (QC) are incorporated coded in R.

The QC steps performed by automated R code are as follows:

- SAP provided data is checked for date issues, spelling issues and duplication.

These data quality checks are performed across different R programs with help of user defined functions. The data is deemed adequate as the pre- and post-performance tests during data adjustments are not conducted.

The manual QC steps are as follows:

- Asset data obtained from SAP is validated and updated through inspections and other programs.

Data Sampling

Since this is a classification model to predict the transformer failures, there are no sampling strategies used in the model other than the random split strategy to bifurcate the train and test data. The dataset used for the model are randomly divided to have 60% in train data and remaining 40% for test data.

Data Cleansing and Transformation

The data cleansing and transformation activities that are incorporated in the R scripts as a part of automation to ensure the completeness of data used for model training and estimation are provided below. R codes used in the OH Transformer model are added in the cross-references section below.

- Missing data for the below specified numeric variables are handled by imputing the mean value.

⁶ Circuit comprises a collection of segments that altogether form a path for electrical current floating from the power source (including but not limited to a substation) to another power source or circuit endpoint.

- AGE_UDF
- Region_UDF
- User_STATUS_UDF
- The transformer replacement information may be missing. hence there are high possibilities to encounter missing data variable. Mean imputation based on Structure Age is performed. If the information is still missing, then this information is proxied using the pole age data fetched from SAP data source.
- Region information is imputed based on location if it is missing.
- USER_STATUS imputed from other attributes if empty.

Data Assumptions

The accuracy of the predicted results is dependent on the accuracy of the data used to build the predictive models. Following are the data assumptions:

11. The assumptions used for the data imputation utilized SCE's Distribution Design Standard (DDS), engineering judgement, manufacturer data and acceptable engineering practices.
12. The failure data is identified as the transformer replacement due to a failure such as corrosion.
13. Input data with respect to asset information, weather information and engineering information are assumed to be stable and will not change over time until the subsequent data refresh. Example: If there is an update in the structure information specific to an asset, that updated information will be reflected only in the subsequent data refresh. So, it is assumed that the updated structure information is not drastically different from the previous information which might alter the model outcomes.

Data Limitations

Following are data limitations across internal and external data sources:

5. Some of the data used by the model faces accuracy issues in terms of consistency in data labelling, missing data for a specific feature (predictive variable) which might impact model prediction power.
 - Data labelling issues might be caused due to manual errors during data entry task. E.g., manufacturer/model is fed manually into the system. While updating the information, different label might be used in different data entries which affects the consistency of the data. These consistency issues in data needs to be addressed before using them in the model.
 - Missing data for a specific feature might be due to unavailability of data. The data is imputed from other asset or location features if possible.

Independent variables

The OH Transformer model uses multiple variables/features. Key features are provided below:

Feature	Data Source	Description
AGE_UDF	SAP	Transformer age, calculated by referring to IN_SERVICE_DATE
Model_Group	SAP	Identify model of transformer
Manufacturer_Group	SAP	Manufacturer of transformer
Latitude/Longitude	SAP (Circuit_Sub_Stats)	Data based on structure location
KVA_Group	SAP (Circuit_Sub_Stats)	KVA cleaned up and validated
Region_UDF	SAP	Region – geographic information from SAP, clean it up for consistency and

Feature	Data Source	Description
		impute missing values
Peak>Loading/Average>Loading Percent_Time_overloaded,Loading Ratio	AMI data	Indicates stress put on transformer
SubType	SAP	transformer sub type
Floodzone/Corrosion	GIS	GIS data
WRF Weathers	SCE Weather stations	Weather information such as temp, wind
User_UDF	SAP	new or used

Transformer age, manufacturer and models are provided through SAP. The consolidated downforce information is processed from SPIDA/Weather WRF and are aggregated to calculate the fiscal measures like mean, standard deviation, skew and kurtosis of segment downforce.

Dependent Variable

In a typical classification risk model, defining the dependent variable is key for both model development and model performance assessment. The dependent variable in OH Transformer model is PROB_FAILURE.

PROB_FAILURE represents the chance or likelihood that a transformer failure will occur either due to failure in transformer. The probability value ranges from 0 to 1 where '0' represents the least likelihood for a failure and '1' represent the high chance for a failure.

- a classification model, the target variable represents the chance or likelihood that a failure will occur due to failure in a transformer.

The failure targets are those identified as Transformer failure in SAP. Then the `h2o.predict (level = 0.05, type = 'response')` function is used to specify the desired output (PROB_FAILURE) in probability values, rather than binary values.

2.2 Methodology

SCE utilizes machine learning to identify patterns that may lead to failures causing sparks from transformers and uses the trained model to predict POIs at the asset level. The OH Transformer model employs a random forest algorithm to predict failure events. The Random Forest approach can predict outputs with high accuracy, run efficiently for large datasets, and maintain accuracy with minimal adjustments for missing values and data treatments.

A random forest is a supervised machine learning algorithm that is constructed from many decision trees. It can be used to solve both classification and regression problems. This approach utilizes ensemble learning, which is a technique that combines many classifiers to achieve greater predictive accuracy than that of a single classifier. A decision tree is a decision support technique that forms a tree-like structure. It consists of three components: decision nodes, leaf nodes, and a root node. The following diagram shows the three types of nodes in a decision tree.

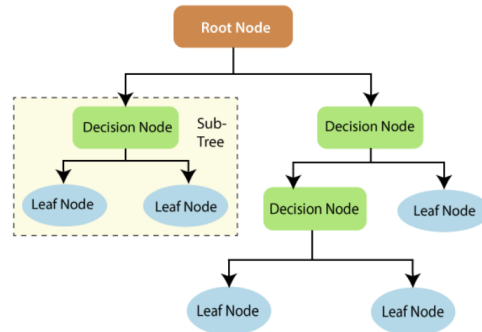


Figure 4: Decision Tree Structure

A decision tree algorithm divides observations of a dataset into branches, which further segregate into other branches. This sequence continues until a leaf node is attained. A leaf node cannot be segregated further. In more detail, the root node is the base of a decision tree, where the first of a chain of decisions is made. A branch is the connection path between nodes. A node is a potential splitting point on a tree. Decision nodes provide a link to the leaves. On the other hand, leaves, also known as terminal nodes, are the ends of a tree, representing the resulting classification or value for the sample.

The 'forest' generated by the random forest algorithm is trained through bagging, also known as bootstrap aggregating. Bagging is an ensemble meta-algorithm that fits multiple models on different subsets of a training dataset and then combines the predictions from all models. The diagram below shows a simple random forest classifier.

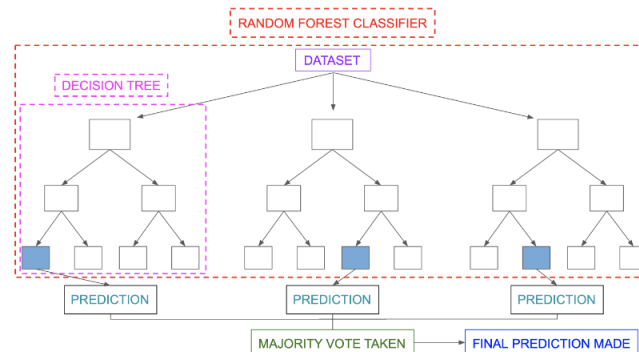


Figure 5: Structure of Random Forest Classifier model

The selection of the final output follows a majority-voting system. In this classification model case, the output chosen by a majority of the decision trees becomes the final output of the random forest system. The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.

Train test split is a model validation procedure that allows to simulate how a model would perform on new/unseen data. Figure provides the logic about dividing the dataset into train data and test data. First the data is consolidated and prepared for train test split. Then the historical input datasets are split into a training dataset (60%) and testing dataset (40%) based on simple random sampling strategy with a split ratio of 3:2 without replacement. Simple random sampling is a technique that ensures each observation has an equal likelihood of being selected for a set. It is a fair strategy as it helps in avoiding any bias involved compared to other modeling techniques and it has no restrictions on the sample size which makes it suitable to handle vastly sized input data. The predictive algorithm is developed using the training dataset and is built by looking at the interactions between all the features to find patterns and predict the likelihood of equipment failure.

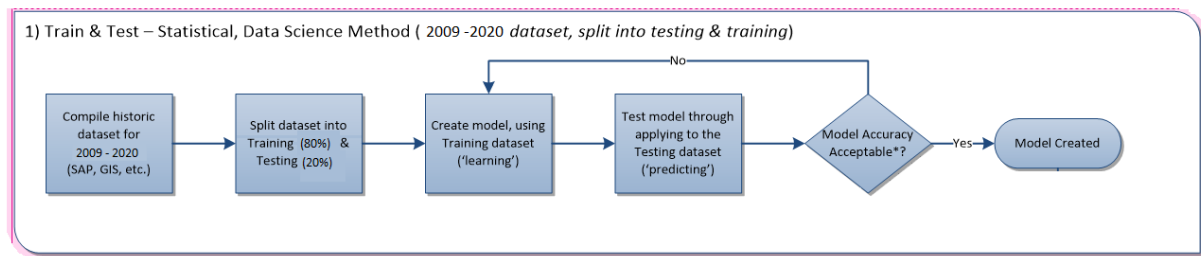


Figure 6: Train and Test data split logic

In the next step, the algorithm is tested on the ‘testing’ dataset. The model is run on the test dataset to make a prediction of a failure or success. Then an internal validation of the model is conducted by comparing the predicted results to the actual results which indicates the predictive capabilities of the features as well as the model. AUC is the metric used to assess the performance of the model on test data.

Area Under the Curve (AUC) – Area Under the Receiver Operating Characteristic (ROC) Curve is a measure to estimate the model discriminatory power (degree of separability) for the binary classification problem. The ROC curve plots True Positive Rate against different thresholds with False Positive Rate (FPR) or True Negative Rate (TNR). The higher the AUC, the better the model is at predicting True Negatives (non-events) and True Positives (events).

Hyperparameter Tuning:

Hyperparameters are parameters that are explicitly defined by the user to control the learning process. The process of selecting the optimal hyperparameters to use is known as hyperparameter tuning, and the tuning process to achieve the best-defined performance statistic is known as hyperparameter optimization. Cartesian Grid search and Random Grid search are the most widely used strategies for hyperparameter optimization.

- In the Cartesian grid search approach, the machine learning model is evaluated for a range of hyperparameter values, and it searches for the best set of hyperparameters from a grid of hyperparameters values. The disadvantage of grid search model is that it will go through all the intermediate combinations of hyperparameters which increases the time consumed by grid search computations.
- In the random grid search approach, the machine learning model is evaluated for a range of hyperparameter values like that in Cartesian Grid Search approach. However, search criteria parameters are added to control the type and extent of the search, and it moves randomly within the grid to find the best set of hyperparameters to achieve maximum performance in terms of the metric defined by the user. As search criteria, the user can set a maximum runtime for the grid, a maximum number of models to create, or metric-based automatic early stopping. If many of these requirements are supplied, the algorithm will end when the first of the criteria is met. This approach reduces the time taken for computation thereby solves the drawbacks of the cartesian grid search approach.

The OH Transformer model was optimized using the cartesian grid search in 2017 and then again 2020. Once the grid search completes, the grid object containing the list of models is queried, and models are sorted by a performance metric defined by the user. The model with better performance is chosen as the best model and it is validated on the test data.

Since then, the same parameters have been used.

- `ntrees`: Total number of trees used in the random forest. 500 trees have been used
- `mtries`: Total number of predictors/variables that will be randomly selected in each node to search for the best split. This parameter is varied by using different percentages of the total number of independent variables in both models. `mtries = -1` which means calculates `sqrt(p)` for classification and `p/3` for regression (where `p` is the # of predictor)
- `max_depth`: Specifies the maximum size of the sample data drawn for training each tree. A higher value for this feature will make the model more complex and can lead to overfitting issue. `max_depth = 20`

- **min_rows:** This parameter defines the minimum number of observations required for a leaf to split.
Min_rows = 10

Cross-references: Refer to [RF RF 1: SCE's WMP 2022 Q1 Quarterly Data Report submission

RF 1: SCE's WMP 2022 Q1 Quarterly Data Report submission

] in Section 5 to understand the efficiency between Cartesian Grid search and Random Grid search.

2.3 Suitability

During development of the model in 2019, GBM and Random Forest were tested with OH EFF Conductor Model that has very similar data. The test results showed that the Random Forest methodology fits well with the data and the results sought. The Transformer model uses the same methodology as the EFF OH Conductor Model using overlapping weather data but transformer appropriate asset information.

Random Forest methodology can be used to solve both classification as well as regression problems and it works well with both categorical and continuous variables. One of the main advantages of Random Forest methodology is to run efficiently for large dataset and maintains accuracy with minimal adjustments for missing values and data treatments. This characteristic to handle large dataset also makes the approach more suitable for its usage in many of the SCE Asset models.

Theoretically, the Random Forest methodology exhibits higher level of accuracy, stability and handles non-linear parameters efficiently than other approaches. Additionally, the random search approach used in hyperparameter tuning controls the maximum depth of the sample data drawn for training each tree and involves stopping criterion which reduces the computation time and also avoids the overfitting issue.

Hence, the usage of Random Forest for the OH Transformer model is deemed to be fit.

2.4 Assumptions

The key business assumptions that were considered during the model development are specified below:

BA 01: There is no change in OH Transformer technical specification over time. The model assumes the type of OH transformers used in the model building process have same characteristics in terms of build and quality. For example, **transformer voltage is constant.**

BA 02: The Calibration model assumes that fires are a subset of failures. Transformer replacements due to failure are the representative failure targets used in place of few ignition events. The failure can spark an ignition, but not all failures will result in a fire. Hence, fire can be treated as a subset of failure.

BA 03: The model is designed to work in both base weather and extreme weather conditions. The weather variables considered by the model are represented as various statistical aggregations like max, mean and standard deviation on wind, wind speed, humidity, rain, and snow. Hence the model results can be used under both base weather and extreme weather conditions.

The functional/model methodology assumptions that were considered during model development are discussed in detail below:

MA 01: The feature variables in the dataset should have some actual values so that the classifier model can predict accurate results. In an ideal scenario, all the variables would not have estimated values and they would instead use actual values. The current model is able to provide accurate results even after using estimates as they are derived through imputation using actual values from other variables.

MA 02: The predictions from each tree must have very low correlations. It is difficult to differentiate between a real interaction effect, marginal effects, and just random variations in random forests. Hence, the presence of highly correlated variables in Random Forest approach will have an impact on its ability to identify strong predictors.

2.5 Limitations and Compensating Controls

The key model limitations that would impact the accuracy and performance of the model are discussed in detail below:

Limitation ID: L01

Limitation Title: Unavailability of linear/non-linear representation in the form of intuitive equation or correlation statistic.

Description: The Random Forest algorithm does not explain any linear or non-linear relationship in the form of an intuitive equation or correlation statistic to enable measurement of the scalability of impact of independent variables on the dependent variable.

Compensating Controls: The Random Forest model is considered a black box as it is difficult to understand the relationship between independent and dependent variables and how the independent variables influence the predictions. Since black box is a common limitation with most ML algorithms, usage of the model is considered appropriate as it provides better AUC results than other models.

Limitation ID: L02

Limitation Title: Resource utilization for model execution is high

Description: Since Random Forest models process many decision trees, they need more resources with respect to system configuration and system capacity to store that data.

Compensating Controls: The resource utilization factor will have a major impact for real time models as they would run more frequently. Since the OH Transformer model is run only once a year with reasonable use cases, the impact of resource utilization is low. Additionally, the usage of random grid search and stopping criterion like tolerance, maximum rounds, maximum run time, and performance improvement thresholds provide more control on the number of recurring instances run to identify the best fit hyperparameters to achieve optimal AUC. Since the model is not executed through computer program automatically at a defined frequency and is instead run only once a year manually, usage of the model is considered appropriate.

Limitation ID: L03

Limitation Title: Model accuracy might reduce if the dataset experiences covariate shift.

Description: Covariate shift is a type of model drift which occurs when the distribution of independent variables changes between the training environment and live/test environment. Since the Random Forest cannot extrapolate (i.e., predict outside the training space), the model performance might decrease if there is covariate shift in the dataset.

Compensating Controls: The covariate shift affects most machine learning models to some degree, as test data is never going to be the same as training data. Detecting and addressing covariate shift is therefore a key step to the machine learning process. The current model is run only once a year along with data refresh. It uses a random sampling mechanism to split the dataset into train (60%) and test (40%) data whenever it is run. The usage of random sampling mechanism is considered to resolve the issue of covariate drift and maintains the accuracy of the model results. Hence the usage of the Random Forest methodology along with the random sampling mechanism to split train/test data is considered appropriate.

2.6 Model Outputs

The OH Transformer model predicts the probability of ignition (POI) arising from equipment (transformer) failure. The model has a single output characterized by a continuous number between 0 and 1 for each OH Transformer asset.

The probabilities across different asset failure predictive models cannot be aggregated or compared and hence are calibrated to derive frequencies of ignition. The sum of the resulting frequencies of ignition for a sub-driver equals the total expected ignitions for the specified year.

$$\text{Frequency of Ignition} = \text{Probability of Ignition} \times \frac{\text{Calibrated Targets}}{\sum \text{Probability of Ignition}}$$

where Calibrated Targets = Forecasted Ignitions for that sub-driver

The output from this calibration exhibits the following features:

- Frequency: Each value can be specified as the frequency of fires per year.
- Comparability: The frequencies are comparable against sub-drivers and models.
- Additivity: The frequencies can be added across models to derive the aggregated fire forecast in a year.

This is achieved by forecasting fires by sub-driver and using these forecasts to weight model probabilities. The sum of probabilities from each calibrated model equals the forecast by sub-driver.

Figure 7 provides the calibration steps that are performed using the failure probability results from the OH Transformer model. This methodology followed in the calibration model is provided below:

- M. Aggregate the probability output from each sub-driver model.
- N. Based on the forecast logic specified above, find the forecast results (expected fires) for each sub-driver.
- O. Generate the calibration factor for each sub-driver based on the values calculated in the above steps (B/A).
- P. Multiply each model probability by its calibration factor to arrive at the estimated frequency of fires from each sub-driver.

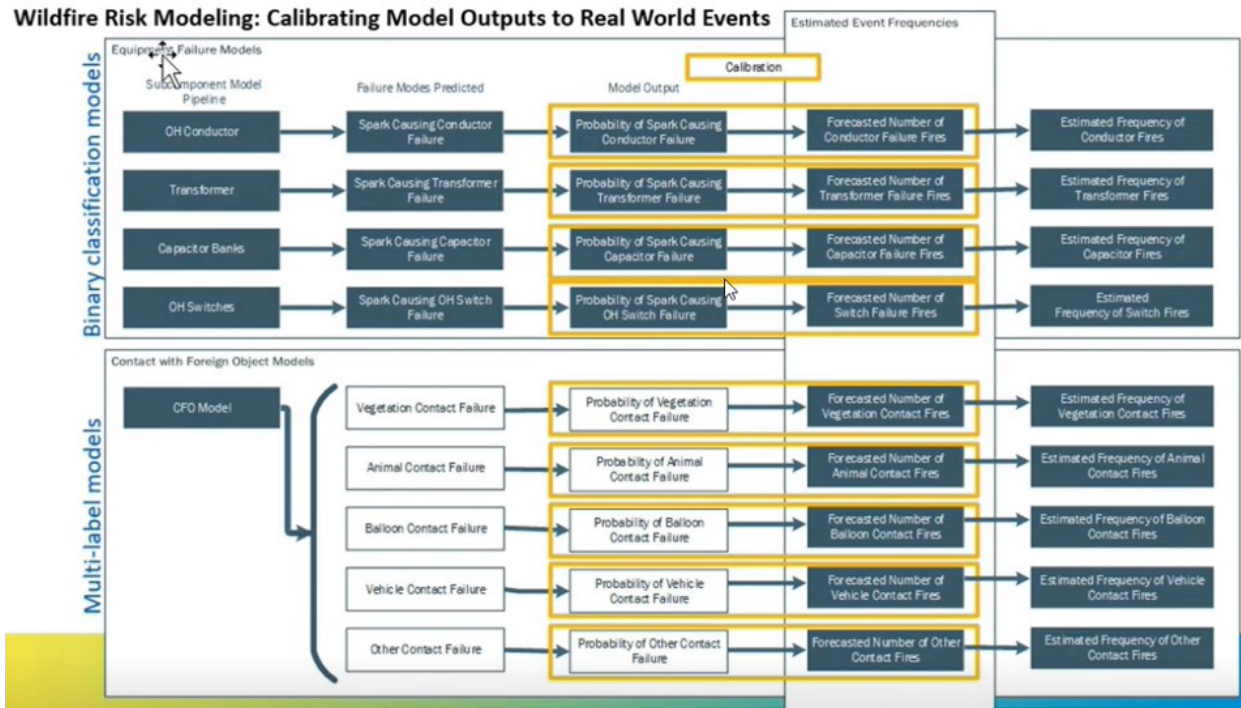


Figure 7: Calibration model schema

This estimated frequency of fires from each sub-driver can be added across the models to derive the expected frequency of ignition for each location.

The calibrated probabilities, frequencies of events, based on the output from OH Transformer model is the data ingested to inform the programs mentioned in Section 1.1.

Model Changes:

No major changes have been recorded for the Transformer model since inception. The data used has been refreshed yearly but the basic features remain the same.

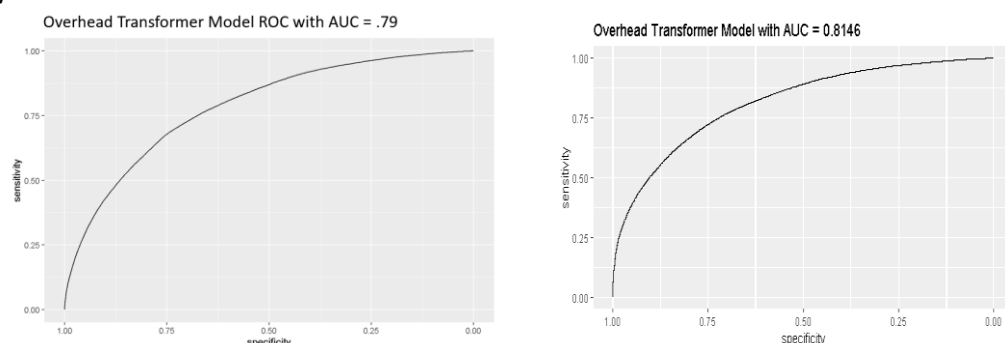


Figure 8: AUC results for 2021 and 2022 Transformer Model

The change in AUC determines the impact of model performance with respect to the changes made as a part of yearly refresh.

The impact on the model performance from the change in features can be determined by comparing the AUC of the model before (2021) and after (2022) refresh (Figure 8). The AUC value of the OH Transformer model was 0.79 before refresh and 0.8146 after refresh.

3. MODEL PERFORMANCE AND TESTING

For each machine learning model developed, SCE tries to select the best algorithm based on the model train/test performance, which can be measured by Area Under the Curve (AUC) and other metrics from the Confusion Matrix.

3.1 Model Specification Testing

The model is developed and tested in R programming using the h2o library. The model is run once a year manually by Data Science and Asset Analytics team with refreshed asset, outage and weather data.

The verification of the model implementation is performed by checking the variable importance results which provides the list of features implemented. The performance of the model is validated through the AUC, defined in Section 2.2 and provided in Section 3.3.

The validity and impact of the Model Assumptions, mentioned in Section 2.4, are discussed below:

- Random Forest is considered a strong approach for variable selection in high-dimensional data only when the variables have low correlation. The recursive structure of trees generally enables them to take dependencies into account in a hierarchical manner. However, some variable combinations without clear marginal effects might make the tree algorithm ineffective. To conclude, it is difficult to differentiate between a real interaction effect, marginal effects, and just random variations in random forests. Hence, the presence of highly correlated variables in Random Forest approach will have an impact on its ability to identify strong predictors. Adequate measures are taken to filter out highly correlated features to overcome their impact in predicting the results.

Model Estimation:

The OH Transformer model employs a number of independent variables. Section 2.1 contains a list of the independent variables utilized in this model.

The variable importance test results for the OH Transformer model, Figure 9, shows the order of which features provided the most information gain in informing the correct prediction of failure or non-failure. The variable importance features test estimates the relative influence of each variable by calculating whether that variable was chosen to split during the tree building process and how much the squared error (over all trees) improved (or decreased) as a result.

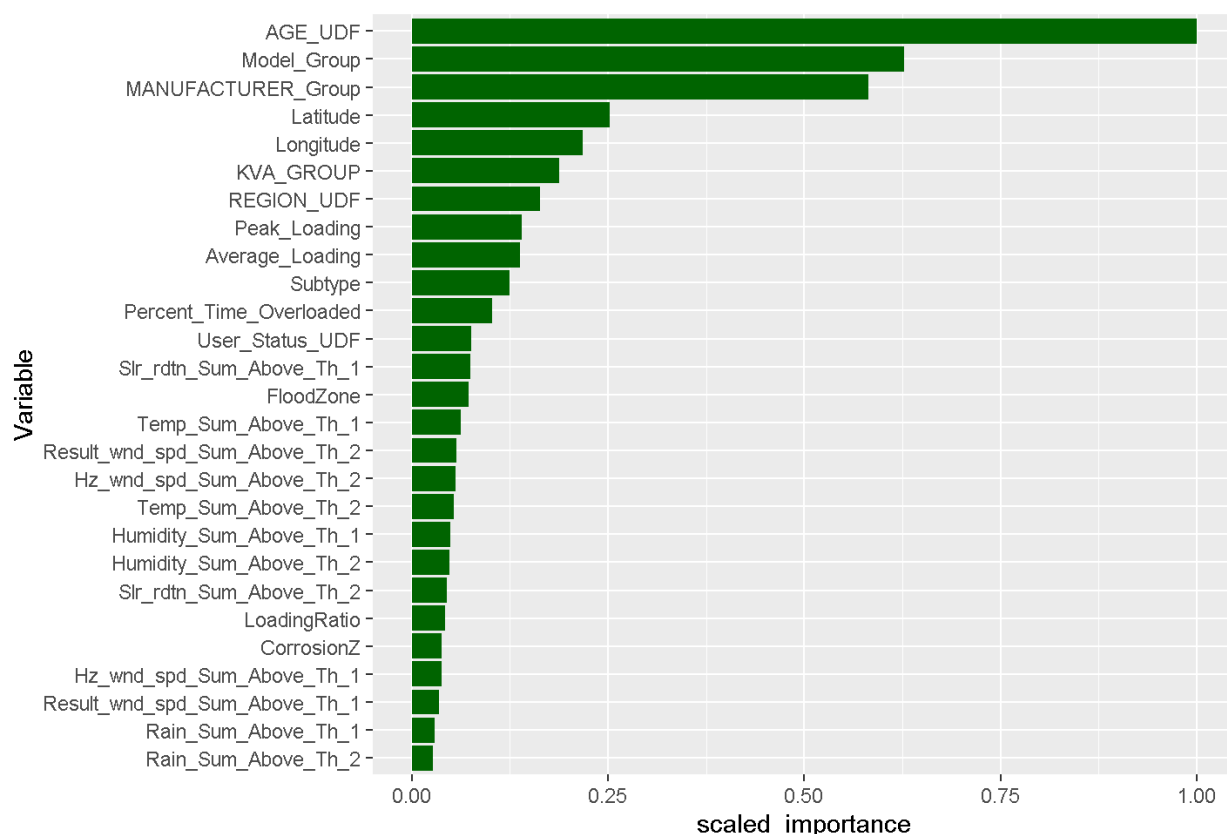


Figure 9: Variable Importance test results for OH Transformer model

The results confirm that Age, Model/Manufacturer, KVA and loading information exhibits high importance on the model output.

Cross-references: Refer to link [RF 3] in Section 0 for description on the methodology used to perform the Variable Importance for tree-based methods.

The OH Transformer model had some parameter tuning to select the best fit of the features in terms of AUC. The same parameters continued to be used in the refresh of the data.

In terms of model convergence, the model runs to max_depth = 20 deep, with min_row = 10 which is the minimum number of observations required for a leaf to split. The algorithm is stopped at no more than 20 deep. Making max_depth larger makes the model overfit the training data.

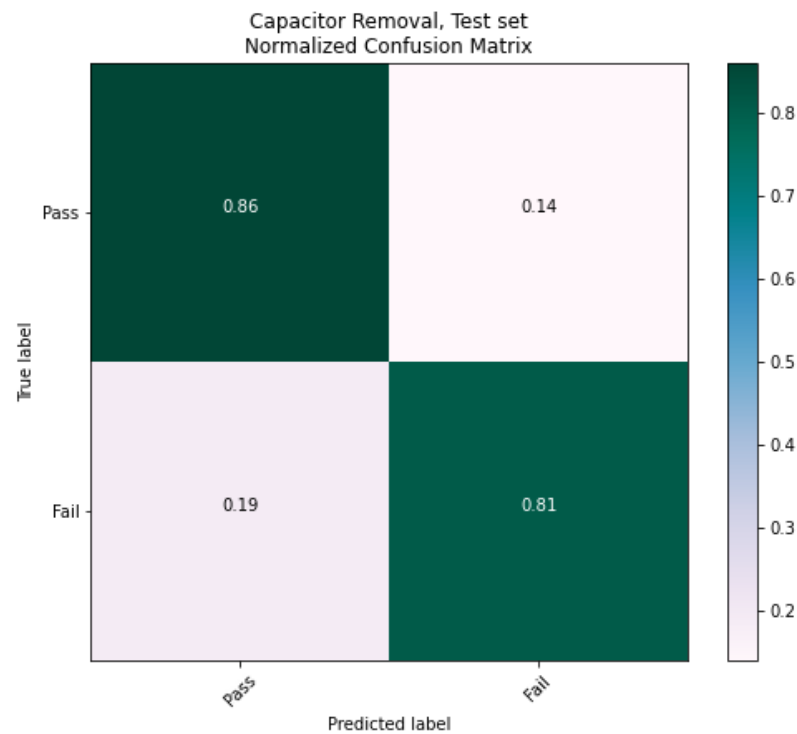
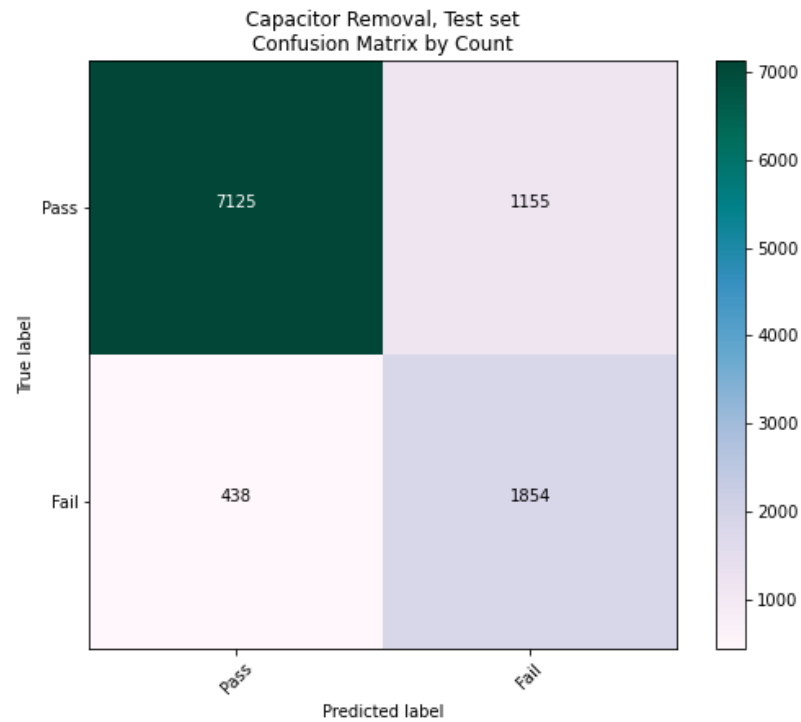
The accuracy of the model prediction, in addition to AUC, can be determined using the Confusion Matrix and Classification Error Rate results.

- A confusion matrix presents a tabular layout of the different outcomes of the prediction results of a classification problem and helps visualize its outcomes. It generates a table of all the predicted and actual values of a classifier model.

Confusion Matrix Results				
		Predicted		
		0	1	Error Rate
Actuals	0	305421	50421	0.141695
	1	29184	40913	0.416337
		334605	91334	0.1868929

Table 10: Confusion matrix results for Transformer Model

- A confusion matrix presents a tabular layout of the different outcomes of the prediction results of a classification problem and helps visualize its outcomes. It generates a table of all the predicted and actual values of a classifier model.



- provides the confusion matrix results for the OH Transformer model. It captures the accuracy rate as 81.3%.
- Classification error rate is used to estimate the proportion of instances misclassified over the whole set of instances. It is estimated using the below formula.

$$\text{Error Rate} = \frac{\text{False Positives} + \text{False Negatives}}{\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives}} * 100$$

The error rate for the transformer model is 18.7% .

All these test results are performed on test dataset with removal for failure data set from SAP.

A detailed assessment of the model limitations and associated compensating controls are available in Section 2.5.

3.2 Sensitivity Analysis

For sensitivity analysis SHAP values were calculated for the main features to see how each of the features impacted the model predictions.

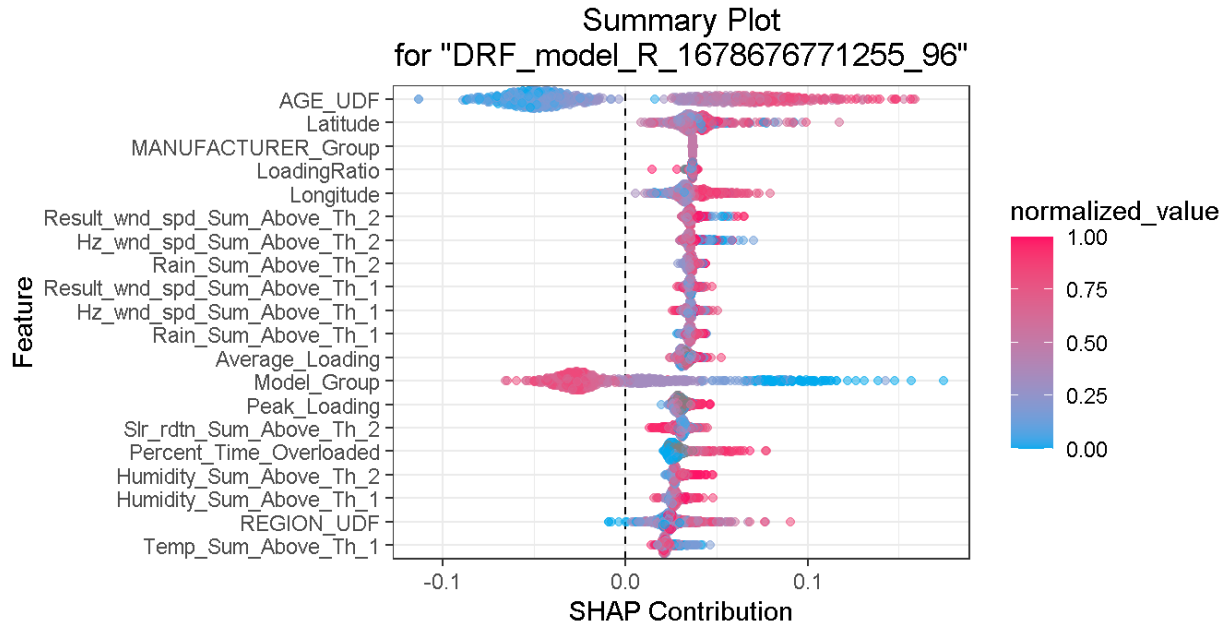


Figure 10: Variable Importance with SHAP values give further insight into how the variables interact with the probability of failure results.

From this Plot we extract features that were most important in prediction and some which had an interesting interaction with prediction for the Transformer model.

The

- AGE_UDF – this is the Age of the transformer and looking at the partial plot of it, yields expected results, newer transformers predict lower levels of failure.
- Hz_wnd_spd_Sum_Above_Th2 Horizontal wind speed exceeding a high threshold
- Result_wnd_spd_Sum_Above_Th2 The vectorial average of all wind directions and speeds exceeding a threshold

Both Hz_wnd_spd_Sum_Above_Th2 and Result_wnd_spd_Sum_Above_Th2 have increase in failure at low and high end of exceeding threshold – this result is surprising and we will take a deeper look at it. This may make more sense with a combination of information from other features -such as age or region.

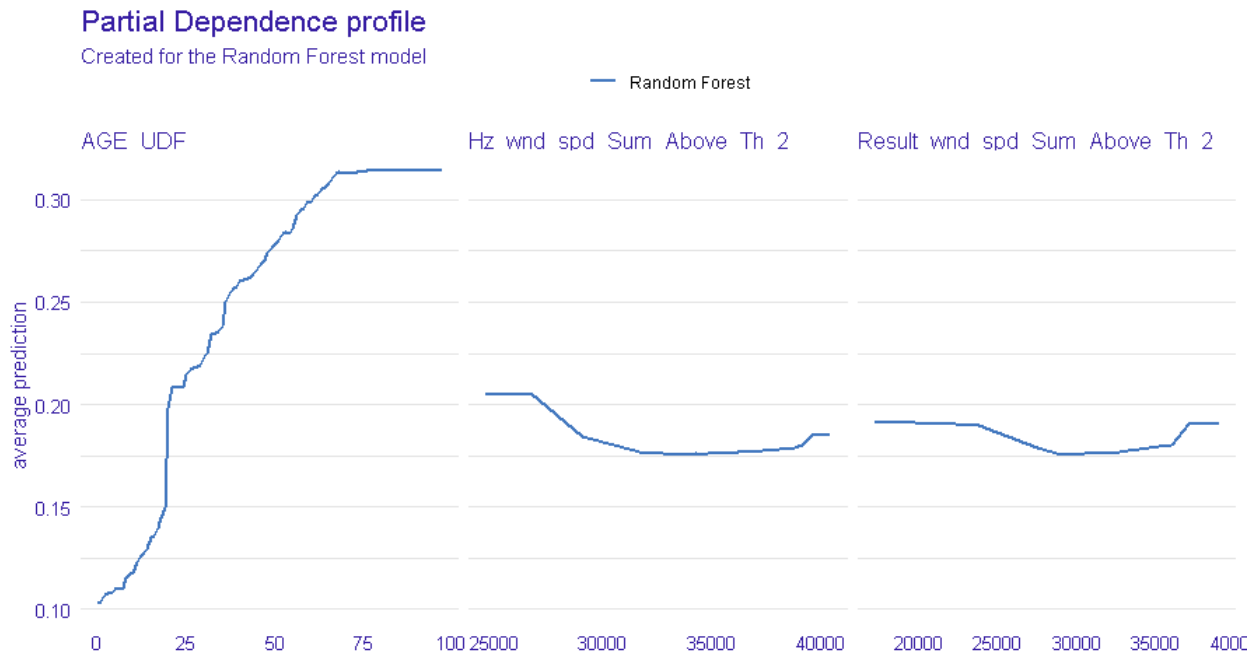


Figure 11: Partial Dependence Profile for Age, Horizontal wind speed exceeding a high threshold, and vectorial average of all wind directions and speeds exceeding a threshold

- Model Group – this was a calculation of bagging models based on their frequency of failure, it behaves consistently, Group 0 and Group 1 categories have highest predicted failure but the other groups have very similar predicted failures

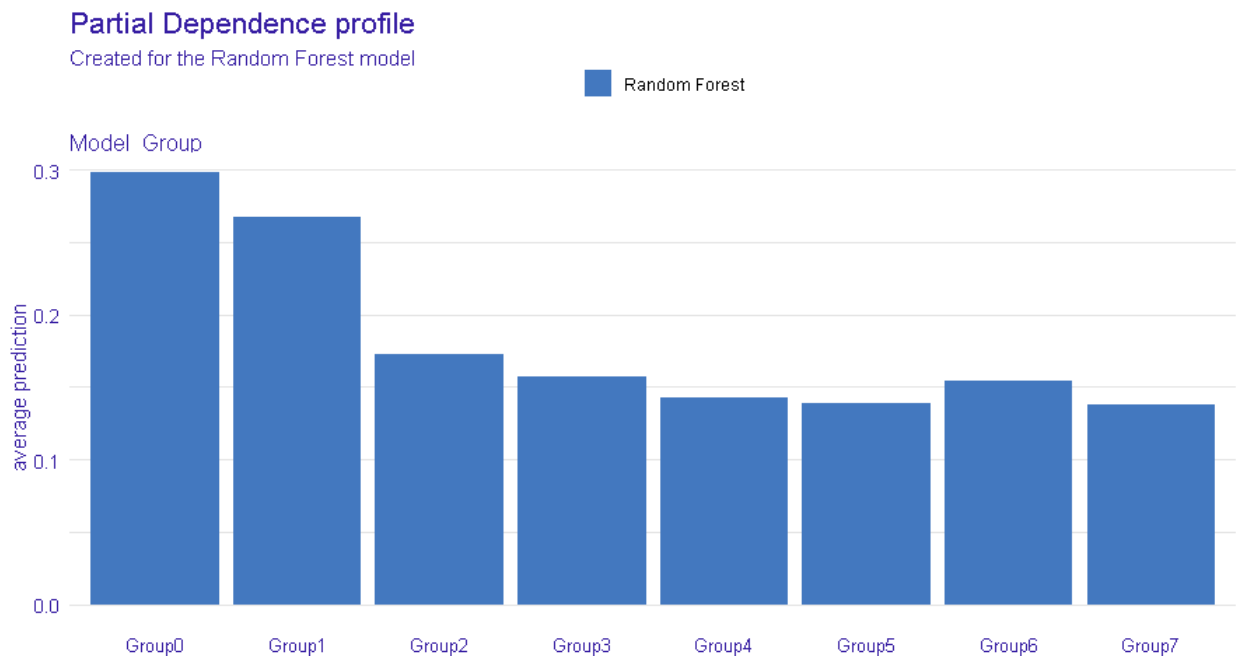


Figure 12: Categorical Partial Dependence Profile for Model Group. Model Group bags various models depending on frequency of failure 0 being highest and 7 the lowest. This plot confirms that the models with highest frequency failure do impact probability of further failure, but that is only true for the two highest tiers.

Further work is planned to look at feature inter-dependence to see how multiple feature information combined impacts the prediction for failure.

3.3 Outcome Analysis / Backtesting

The subset of historical data on which a model is trained and optimized is referred to as the in-sample data. On the other hand, the subset of the dataset that has been reserved to test the model is known as the out-of-sample data. The OH Transformer model uses a random sampling approach to split the dataset into Train (60%) and Test (40%) data. The results arrived from train data are considered as in-sample backtesting and the results arrived from test data is considered as out-of-sample backtesting.

Once the machine learning model is built with the training data, it is evaluated using a separate test dataset that has not yet been studied. The performance of the model is determined by the Area Under the ROC Curve (AUC) value. Figure 13 shows the AUC value and ROC for the OH Transformer Model based on the test dataset ran with transformer replacement due to failure data until 2022. The AUC value of 0.814 implies that the model possesses reasonable accuracy in terms of predicting the results, with room for improvement in future iterations.

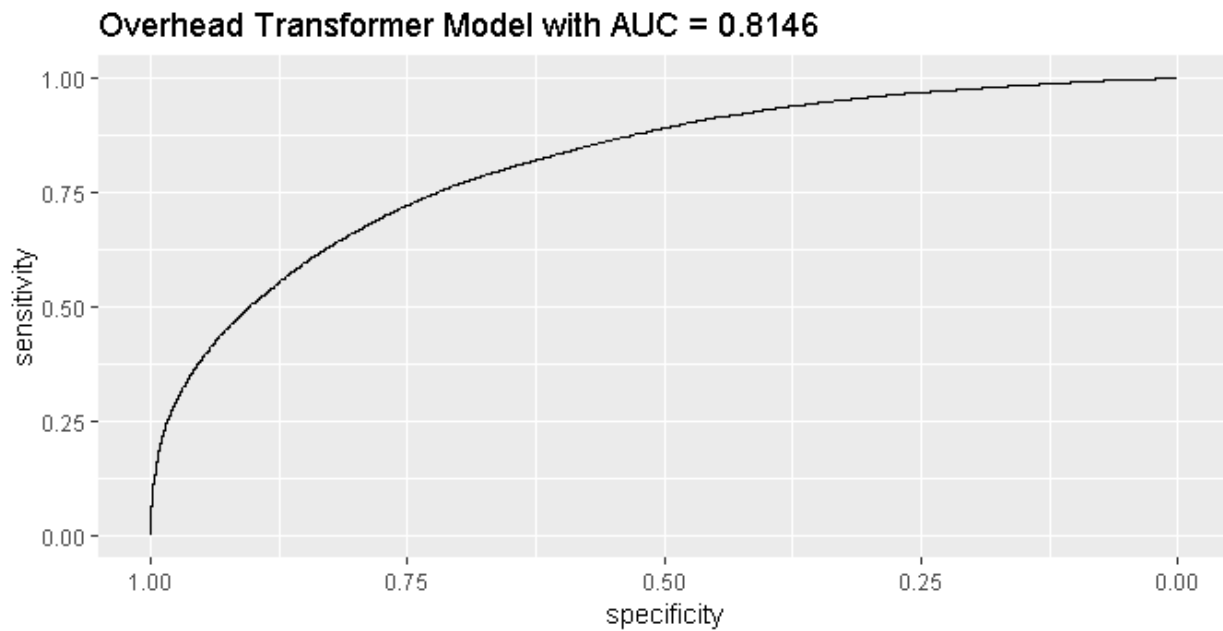


Figure 13: Out-sample backtesting result for the Transformer Model based on test dataset

The impact of uncertainty in model inputs and parameters on model outputs are tested as a part of the sensitivity analysis and the results are captured in Section 3.2. In addition, the data imputations that are incorporated to address missing values before running the model are defined in Section 2.1.

3.4 Benchmarking Analysis

For the OH Transformer Model, different approaches like Gradient Boosting Machine (GBM) learning, Logistic Regression, and Random Forest were considered during the model development phase in 2017. The analysis on these supervised machine learning approaches and the results are provided below.

- Gradient Boosting Machine (GBM) is one of the most popular forward learning ensemble methods in machine learning. It is a powerful technique for building predictive models for classification and regression tasks. GBM sequentially combines the predictions from various weak learner decision trees and builds a final predictive model with more accurate predictions by minimizing a defined loss function.

- Logistic regression is used to solve classification problems. The three types of logistic regression available are Binary logistic regression (handles binary outcomes), Multinomial logistic regression (handles multiple outcomes, i.e., multi-classification variable), and Ordinal logistic regression (handles ordered outcomes). In contrast, linear regression solves regression problems where the outcome is continuous and can be any possible numeric value.
- Random Forest is a popular machine learning algorithm that can be used for both classification and regression problems. Random Forest is another ensemble method that combines the predictions of several decision trees to improve the predictive accuracy of the model. The individual decision trees are created based on a randomly selected subset of features at each node prior to determining the optimal split so each tree differs. The final output is determined by taking the majority vote of the predictions from the individual decision trees. The greater number of trees in the forest generally leads to higher accuracy and prevents the problem of overfitting.

The benchmarking results of GBM and Logistic Regression shared in this section were developed using the h2o library in R on the Test data with targets based on the transformer replacement due to failure data until 2022. Since benchmark results were not saved during the model development phase, the benchmark models were executed in February 2023 for documentation purposes. Figure 14 provides the AUC values for the OH Transformer model using the GBM, Logistic Regression, and Random Forest methodologies.

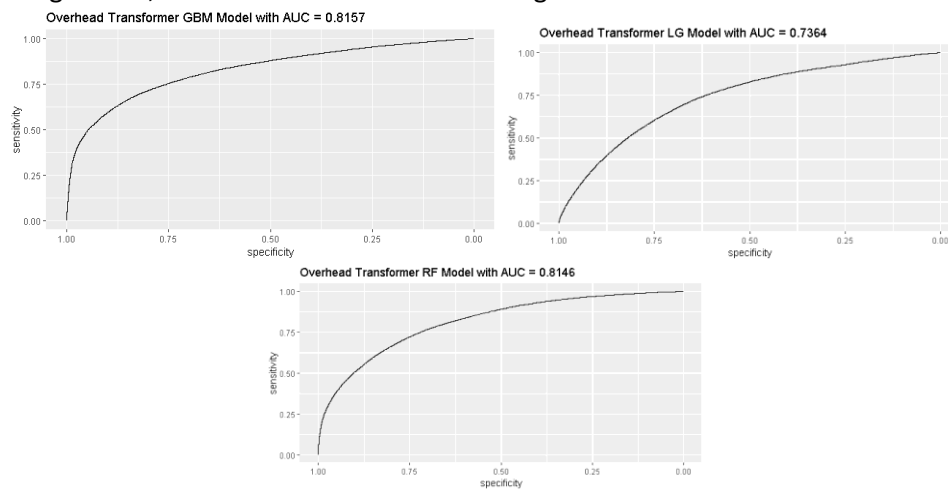


Figure 14: AUC Comparison for the Transformer Model using GBM, Logistic regression and Random Forest methodologies. GBM is the gradient boosting method. LG is the linear regression and RF is the random forest method.

Random Forest was chosen as the modeling algorithm for the OH Transformer model as it aligns with the modeling approach for SCE's other predictive asset failure models, and it achieved relatively the same AUC as GBM. GBM is a methodology that will continue to be considered for use as part of the annual refresh of the model. Some additional advantages of using Random Forest over GBM and Logistic Regression are provided below:

- Random Forest is less sensitive to overfitting issues than GBM.
- Hyperparameter tuning is relatively easy in Random Forest when compared with GBM.
- Random Forest is better at handling categorical variables while retaining the original encoding compared to weight-based algorithms like logistic regression which may treat categories of higher importance depending on the number assigned.

4. MODEL MANAGEMENT AND GOVERNANCE

4.1 Ongoing Monitoring Plan

Ongoing monitoring is important for Machine Learning models especially when they are used to make predictions

or when they are run on datasets with high volatility in variable values. The OH Transformer model is run manually once a year, incorporating updated input datasets to reflect the latest available data and implementing any specific model enhancements—e.g., inclusion / replacement / removal of a feature, optimization of the code, evaluation of a new performance metric, etc. During the model refresh, the limitations and assumptions of the model are also revisited by the model developers and necessary action items are conducted to address them.

Performance monitoring is required only after running the model. Recalibration of the model has not been performed for the last two years, and it is performed only if the behavior of the model differs from that of the previous version or if there is a significant drop in model performance. The AUC and accuracy rate from confusion matrix results obtained after model refresh are compared against a threshold of 80%; if the value drops below this threshold, the reason behind the performance dip is investigated. Post-investigation, the steps required to improve the performance of the model will be carried out. To monitor the model performance more thoroughly, developers of the model plan to additionally evaluate metrics like Precision and Recall. Precision is the positive predictive value which represents the proportion of predicted failures that were predicted correctly. Recall is the true positive rate which represents the proportion of actual failures that were predicted correctly.

The model documentation and the performance results are updated once a year immediately after the model is refreshed.

4.2 Security and Control

The Data Science and Asset Analytics team has access to the data inputs, code, and implementation for the model. Other business units, like the Grid Hardening Strategy team, are provided access to the model outputs upon request but cannot update or modify the code.

The model is run using R programming and it can be executed in any recent versions of the R software. The current model versioning is labeled by date of refresh (e.g., OverheadTransformer\updated2022). There are plans to move the code to GitHub, a platform that facilitates version control by tracking changes to the source code. Users with write or admin privileges to the repository can review proposed changes and approve them.

A contingency plan is not applicable for this model as it is an in-house model for SCE.

5. REFERENCES

RF 1: SCE's WMP 2022 Q1 Quarterly Data Report submission

<https://www.sce.com/sites/default/files/AEM/Data%20Requests/2022/SCE%20Q1%202022%20Tables%201-12.xlsx>

RF 2: Literature reference on grid search vs random search approach for hyperparameter tuning.

<https://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf>

RF 3: Variable Importance methodology for tree-based methods

[Variable Importance — H2O 3.38.0.3 documentation](#)