

Southern California Edison (SCE)
Model Documentation
Prepared for 2026-2028 WMP
Appendix B

OH Conductor Sub-Models (CFO & EFF)

5/16/25

Table of Contents

1. EXECUTIVE SUMMARY	3
1.1 Model Purpose and Intended Use	3
1.2 Model Description Summary	3
1.3 Model Risk Rating.....	4
1.4 Model Dependency and Interconnectivity	5
1.5 Model Assumptions	5
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	7
2.1 Model Inputs and Data Quality.....	7
2.2 Methodology	13
2.3 Suitability.....	16
2.4 Assumptions	17
2.5 Limitations and Compensating Controls	17
2.6 Model Outputs	19
3. MODEL PERFORMANCE AND TESTING	21
3.1 Model Specification Testing.....	21
3.2 Sensitivity Analysis	25
3.3 Outcome Analysis / Backtesting.....	29
3.4 Benchmarking Analysis	31
4. MODEL MANAGEMENT AND GOVERNANCE	34
4.1 Ongoing Monitoring Plan	34
4.2 Security and Control.....	34
5. REFERENCES.....	35

1. EXECUTIVE SUMMARY

1.1 Model Purpose and Intended Use

The Overhead (OH) Conductor Model is a Probability of Ignition (POI) Sub-Model developed by Southern California Edison (SCE). The wildfire risk associated with SCE's OH Conductor model is further measured in additional sub-models, i.e., Equipment / Facility Failure (EFF) and Contact with Foreign Object (CFO). At SCE, models are developed using Machine Learning (ML) 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 OH conductors.

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

1. Inspections and Remediations programs that consider POI as an element in prioritization and scoping.
2. Tree Risk Index that uses the output POI from the OH Conductor CFO Vegetation sub-model as one of its inputs.
3. Overhead Conductor Program (OCP) that 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.
4. Risk analyses via SCE's Multi Attribute Risk Scoring (MARS) Framework.

1.2 Model Description Summary

The OH Conductor model predicts the probability of failure of distribution primary OH conductors using Random Forest—a ML 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 (e.g., age, voltage) and other environmental and operational attributes (e.g., historical wind, number of customers). In addition, EFF uses the splice information on the segment¹ to predict the probability of failure by the conductor whereas CFO uses animal incidents, car

¹ 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.

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 Advanced Predictive Modeling team. The model is calibrated every year with the last 5 years of historical outage data.

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

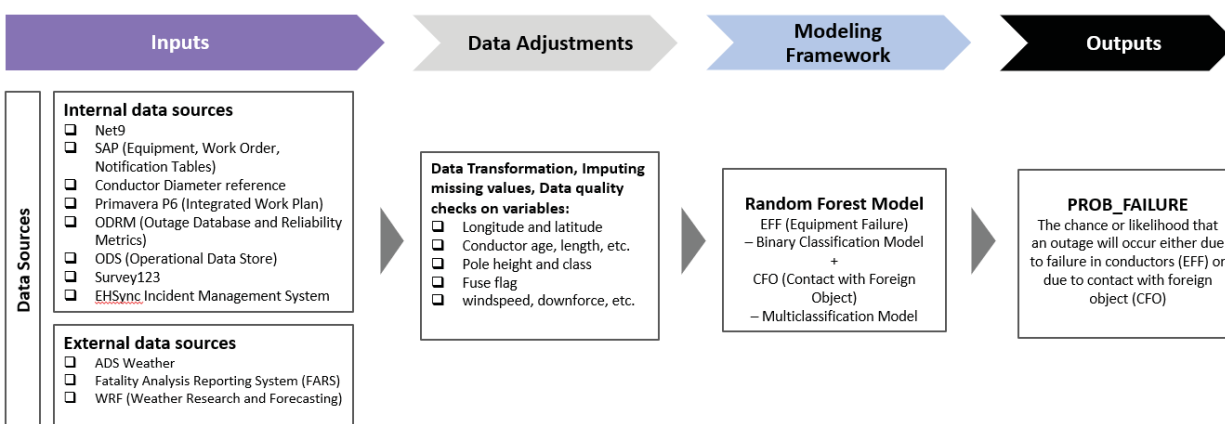


Figure 1: OH Conductor model framework

The OH Conductor model uses the Random Forest methodology for both EFF Conductor and CFO sub-models. Since the prediction is a classified event (i.e., failure) and the Random Forest methodology can perform both classification and regression tasks, the Random Forest methodology is considered a good choice for the OH Conductor model. 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 to identify model risk rating at SCE. However, certain factors—like frequency of risk events and use case—are considered when 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 3,500. This frequency 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 importantly informs the strategy of several programs, 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	2020	2020	2021	2021	2021	2021	2021	2022	2022	2022	2022	2023	2023	2023	2023	2023	
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 2: 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

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

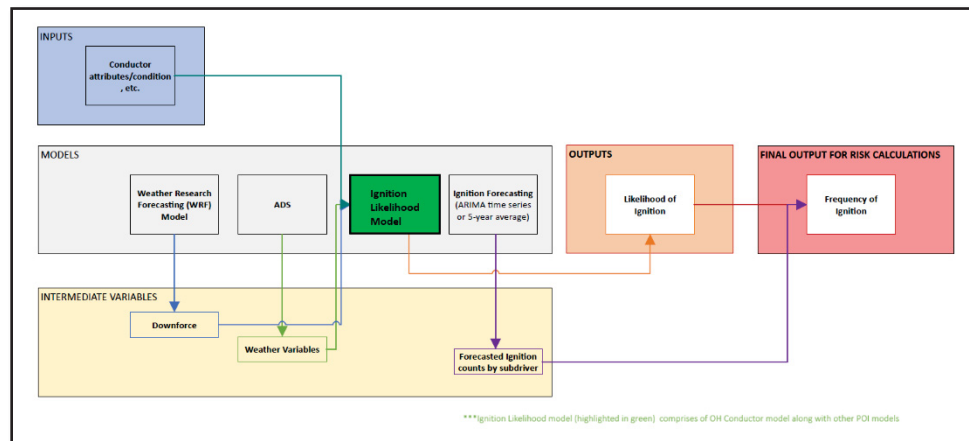


Figure 3: Model Interconnectivity Schema

ADS weather variables are used as one input in the OH Conductor model. ADS’ Next Generation Weather Modeling System (NGWMS) upgrades 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 2013 and 2022. That information is then processed and aggregated to calculate statistical measures such as mean and standard deviation of wind, humidity, rain, snow, etc. These are used as locational measures and are matched to the conductors by their latitude and longitude coordinates.

The Weather Research and Forecasting (WRF) Model was run considering 2013 through 2023 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 to inform the strategic decisions of the four categories of programs, discussed in Section 1.1.

1.5 Model Assumptions

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

1. There is no change in the OH Conductor technical specification over time.
2. The contact types that can cause a spark will remain the same throughout the prediction period.
3. The calibration methodology assumes that fires are a subset of failures.

4. The model is designed to work in both base weather and extreme weather conditions.
5. The feature variables in the dataset should have some actual values so that the classifier model can predict accurate results.
6. 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:

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 ML 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 using the last 5 years of historical outage information.

- The AUC values for the EFF Conductor and CFO sub-models are 0.8709 and 0.8069 respectively. Note that the CFO AUC is for the NO class; the average AUC of the other classes is 0.8707. The model was tuned to capture more True Positives, resulting in a higher False Positive rate.
- Confusion matrix results capture the accuracy rate as 93.29% and 95.57% for EFF Conductor and CFO sub-models respectively.

The above metrics were derived by re-running the model as of Sep 2024 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 in-house 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 circuit², 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 Operational Data Store (ODS) 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.

² Circuit comprises segments that collectively 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.

- **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. Copies of these files are re-housed on Sharepoint and downloaded for this 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.
- **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** model provides 10 years of hourly gridded weather data from 2013-2022. These are aggregated to individual locational measures and matched to the conductor segments through a k-nearest-neighbor distance-weighted spatial join 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.
- **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.

References: Please refer to link [RF 2] in Section 0 for NHTSA source.

Quality Checks

SCE has internal data management teams for ensuring data quality, including Enterprise Asset Data (EAD) 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 performed in the OH Conductor model to ensure accuracy, validity, integrity, and consistency are provided below. Quality checks (QC) are coded in R/Python and incorporated into the data gathering process.

The QC steps performed by automated R/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 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 and within 50ft of conductor, and details about the trees are included as features, e.g. growth rate.
- 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

This classification model has a heavy class imbalance for both EFF and CFO. Stratified under-sampling was used to improve model performance.

The dataset used for the CFO model is randomly divided by class to have 80% in train data and 20% in test data. This ensures that both the train and test data have proportional amounts of each class. The train data is then sampled class-wise to produce class selections that are more balanced (although still disproportionately weighted towards the majority class). The best-performing sampling method this year involved under-sampling the “NO” and “UNKNOWN” classes to 10% and 30% of their original training set size respectively.

The dataset used for the EFF model is randomly divided to have 70% in train data and 30% in test data. The “NO” classes in the train data are then under-sampled so that they represent 80% of the train 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.

- 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 Python 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:

1. The assumptions for the data imputation uses SCE's Distribution Design Standard (DDS), engineering judgment, and manufacturer data.
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 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.
3. 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 and would not 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 labeling or missing values that may impact model prediction power.

- Data labeling issues might be caused by manual errors during data entry. For example, when conductor type information is fed manually into the system, different labels for the same conductor type might be used in different data entries. This affects the consistency of the data and needs to be addressed before using the data 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.

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

Mapping Unknown Outages with DOTS2.2 Method

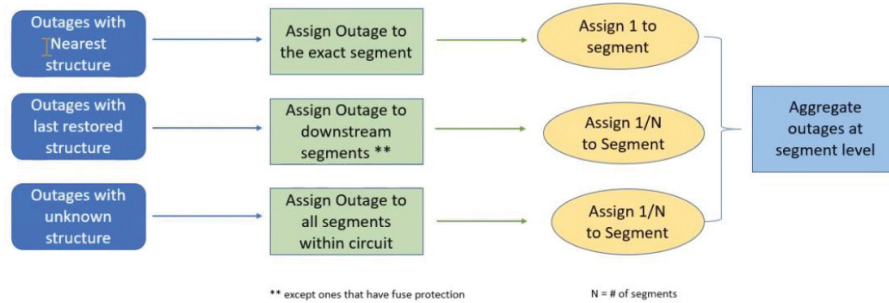


Figure 4: Unknown outage mapping process

- 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 for EFF, and the mean plus 0.5 standard deviation is considered the threshold for CFO. The values assigned to a segment are 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 segment
EQ_PoleClass	SAP	Mode pole class (field in Equipment table: _BIC_ZCAE_P015; values = 1-6, H1-6) of all structures associated to the segment
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
Log_WindForce	SPIDA, Weather Research and Forecasting Model (WRF)	Log conversion of downforce
SECTION_FLAG	Net9	Indicator if conductor has RCS/RAR/FUSE 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

Table 1: Key Features in the Model

In addition to the data 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 KDE³ of there being a vehicular accident at the segment location is calculated using a kernel density estimator from the Python sklearn package. Similarly, animal contact incidents are used to calculate the KDE of there being an avian accident at the segment.

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.

³ Kernel Density Estimation (KDE) is the application of kernel smoothing for probability density estimation.

- 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 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 uses ML to identify patterns that may lead to failures causing sparks from conductors and uses the trained model to predict Probability of Ignitions (POI)s 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.

A random forest is a supervised ML algorithm that is constructed from many decision trees. It can be used to solve both classification and regression problems. This approach uses 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.

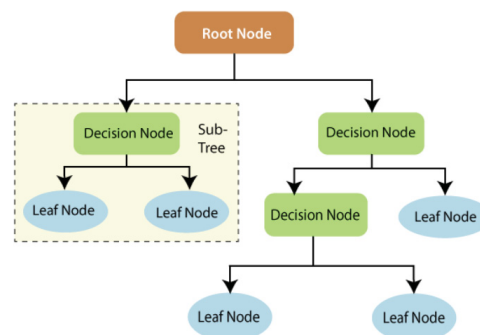


Figure 3: Decision Tree Structure

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.

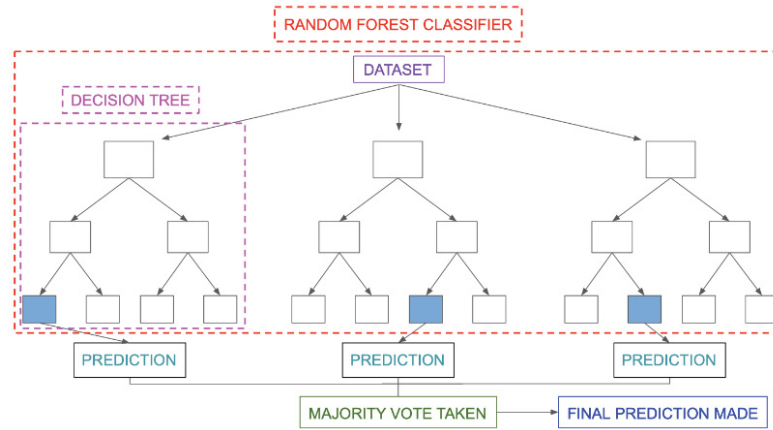


Figure 4: Structure of Random Forest Classifier model

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 57 shows the logic for 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 (CFO: 80%, EFF: 70%) and testing dataset (CFO: 20%, EFF: 30%) based on simple random sampling strategy with a split ratio of 4:1 or 7:3 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 vast amounts of input data. The predictive algorithm is developed using the training dataset and built by looking at the interactions between all the features to find patterns and predict the likelihood of equipment failure.

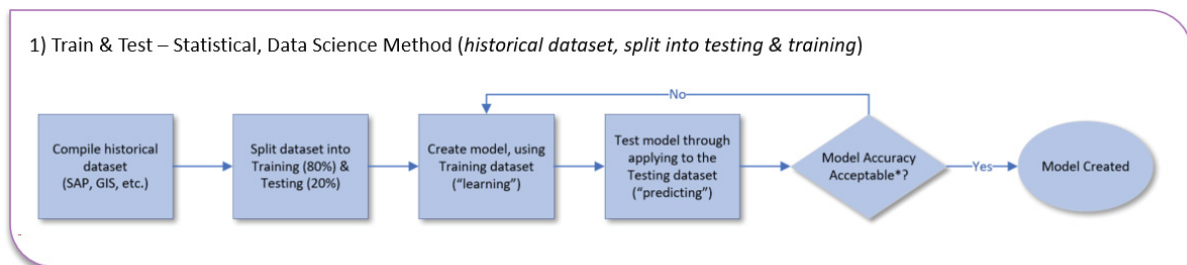


Figure 5: Train and Test data split logic

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. Area Under the Curve (AUC) is the metric used to assess the performance of the model on test data.

AUC – Area Under the Receiver Operating Characteristic (ROC) Curve estimates the model discriminatory power (i.e., 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 widely used strategies for hyperparameter optimization.

- In the Cartesian grid search approach, the ML 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 ML 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 solving 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 240 and 1080 with an increment of 120, whereas the CFO sub-model uses values between 120 and 1080 with an increment of 120, with 240 as the end value for EFF and 960 as the end value for CFO.
- **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 15%, 33.3%, and 40%.
- **max_depth:** The maximum number of decision splits allowed within a 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 both models are set between 20 and 60 with an increment of 10. The final models for EFF and CFO used max depths of 30 and 20 respectively.

- `min_rows`: This parameter defines the minimum number of observations required for a leaf to split. This parameter is tuned using the values 5, 10, 15, 20 for both models.
- `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 0.4, 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` = 14400
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 four hours.
- `stopping_metric` = AUC/Log-loss
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 when the model's log-loss value doesn't improve by 0.5% for the CFO sub-model. While AUC is the preferred metric, log-loss is used for CFO due limitations within H2O for producing multi-label AUC scores in a grid search.

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.

References: Refer to [RF] 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. See Section 3.4 for the AUC comparison of these three approaches.

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. Random Forest methodology provides a high level of accuracy and stability and handles non-linear parameters efficiently. 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 use of Random Forest for the OH Conductor model is deemed to be a suitable 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 variables would have actual, not estimated, values. The current model provides 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 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 that 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 ML 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 ML 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 and test data whenever it is run. The random sampling mechanism is used to resolve covariate drift and maintain the accuracy of 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 for each overhead distribution primary conductor. 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 for each overhead distribution primary conductor.

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

- A. Aggregate the probability output from each sub-driver model.
- B. Based on the forecast logic selected, find the forecast results (i.e., expected fires) for each sub-driver.
- C. Generate the calibration factor for each sub-driver based on the values calculated in the above steps (B/A).

- D. Multiply each model probability by its calibration factor to arrive at the estimated frequency of fires from each sub-driver.

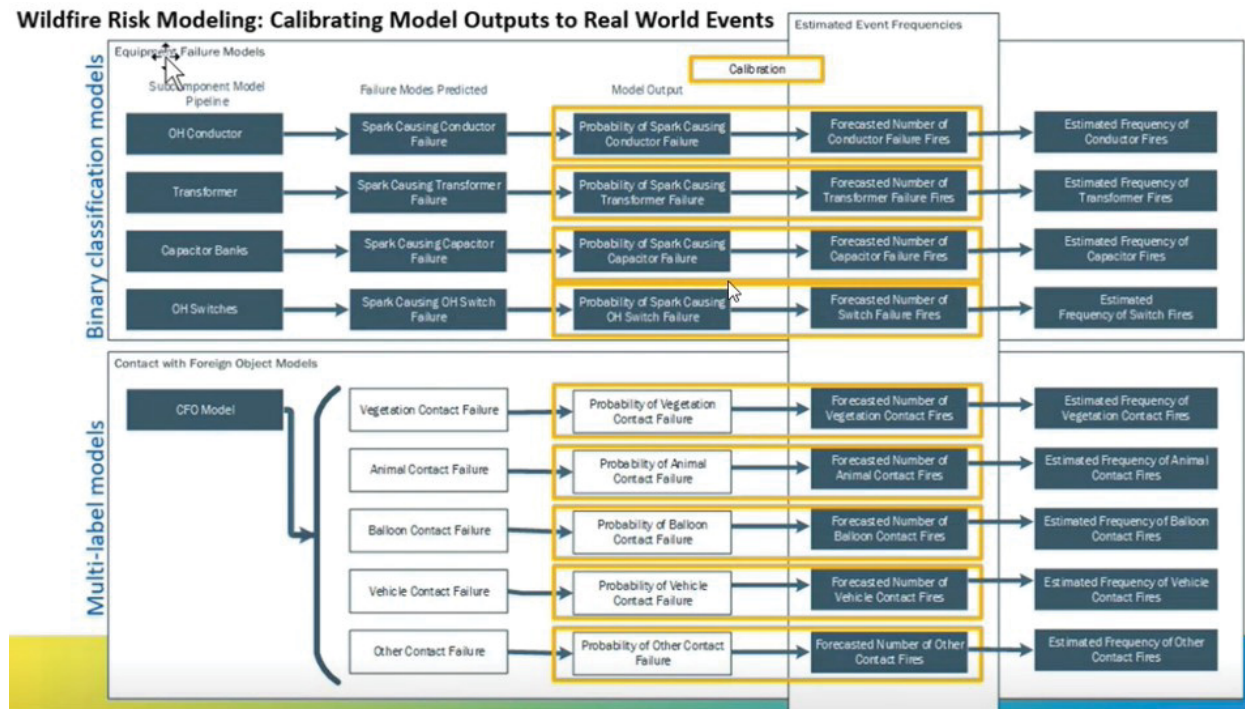


Figure 8: Calibration model schema

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

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

Model Changes:

In 2024 the majority of the feature engineering for the CFO and EFF sub-models was converted to Python. As part of this process, each feature was rigorously evaluated to ensure the new Python methods were producing results at parity or more accurately than the R processes. While the majority of the code maintains logical parity with the R code, improvements in data quality for features—such as Short Circuit Duty for each segment—required using updated data sources.

Much of this work was necessitated to maintain model consistency with changing data sources. For example, structure to segment mapping for all Data Science models now uses a standardized methodology and data source in Python and hosted on Snowflake. This necessitated updating the Overhead Conductor Model's methods for assigning structures to segments, which used custom R code. While the standardized segment to structure mapping was based on the model's methodology, the new data stream made improvements to the mapping, and thus was adopted for both standardization and data quality reasons.

Other features, such as the ADS weather data, used new imputation methods. ADS data is now assigned to segments using a distance-weighted k-nearest-neighbors method, with k=3.

3. MODEL PERFORMANCE AND TESTING

For each ML 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 the Advanced Predictive Modeling team with refreshed asset and weather data. The model is calibrated every year with the last five years of historical outage data.

SCE performs verification of the model implementation by checking the variable importance results (see Model Estimation section below for a detailed explanation of variable importance results). 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 several 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 9 and 10 for EFF Conductor and CFO respectively, shows the order of which features provide 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.

Southern California Edison (SCE) Model Documentation Prepared for 2026-2028 WMP Appendix B OH Conductor Sub-Models (CFO & EFF)

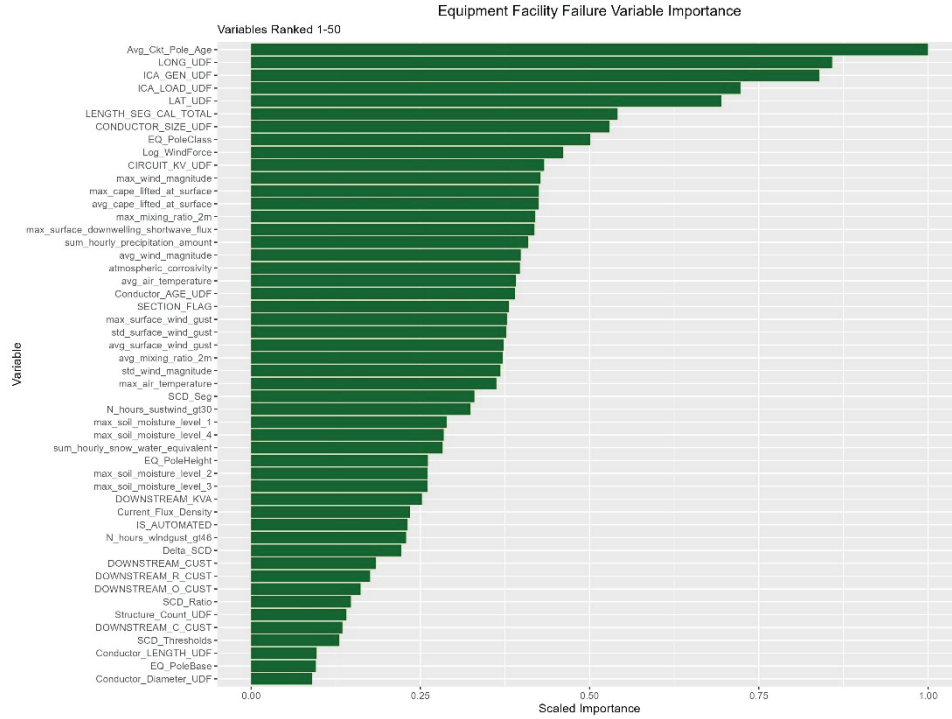


Figure 9: Variable Importance test results for EFF Conductor sub-model

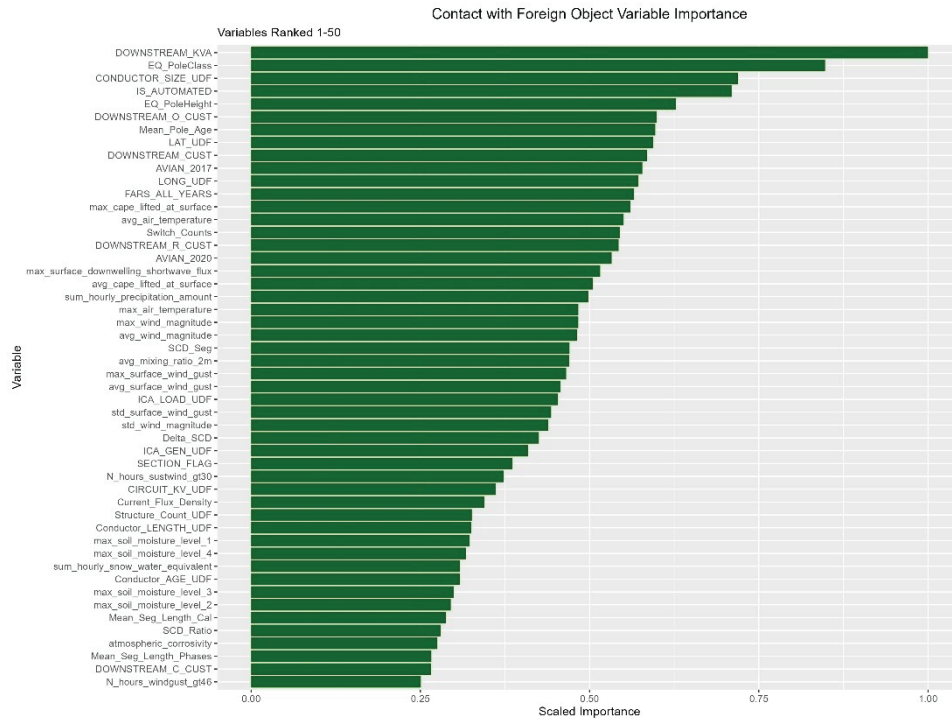


Figure 10: Variable Importance test results for CFO sub-model

The results confirm that the features Avg_Ckt_Pole_Age, LENGTH_SEG_CAL_TOTAL, and CONDUCTOR_SIZE_UDF exhibit high importance on the EFF Conductor sub-model output and the features

CONDUCTOR_SIZE_UDF, DOWNSTREAM_KVA, and SCD_Seg exhibit high importance on the CFO sub-model output.

References: Refer to link [RF 4] 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 11 and Figure 12. 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 13 and Figure 14 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 more accurately. Figure 15 and Figure 16 shows the ROC with AUC for EFF Conductor and CFO sub-models based out of test dataset ran with historical outage data from 2020-2024. The AUC values for the EFF Conductor and CFO sub-models are 0.8709 and 0.8069 respectively. The AUC values for sub-drivers Animal, Balloon, Other, Unknown, Vegetation and Vehicle are 0.9052, 0.8783, 0.8732, 0.7882, 0.9034, and 0.8759.

```
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 11: Example 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
- ntries
- sample_rate
Number of models: 2
Number of failed models: 0

Hyper-Parameter Search Summary: ordered by increasing logloss
max_depth min_rows mtries ntries 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 12: Example list of models with their associated hyperparameter values produced after the grid search for the CFO sub-model.

```
H2OBinomialModel: drf
Model ID: Grid_DRF_TRAIN_CONDUCTOR.hex_model_R_1726442445339_125_model_17
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
1 240 240 15220038 30 30 30.00000 4720 5233 5033.70000
```

Figure 13: The best model for the EFF Conductor sub-model, along with the related hyperparameter values.

```
H2OMultinomialModel: drf
Model ID: DRF_model_R_1726525477068_3
Model Summary:
number_of_trees number_of_internal_trees model_size_in_bytes min_depth max_depth mean_depth min_leaves
1 960 6720 157777625 20 20 20.00000 841
max_leaves mean_leaves
1 4863 1861.05250
```

Figure 14: 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 predicted and actual values of a classifier model.

EFF - Confusion Matrix Results				
Actuals	Predicted			
		0	1	Error Rate
	0	198248	10376	0.049735
	1	4237	4873	0.465093
		202485	15249	0.067114

Table 2: Confusion matrix results for EFF Conductor sub-model

CFO - Confusion Matrix Results									
Actual	Predicted								
		ANIMAL	BALLOON	NO	OTHER	UNK	VEGETATION	VEHICLE HIT	Error
	ANIMAL	117	19	597	14	18	8	9	0.850384
	BALLOON	9	136	757	14	24	8	19	0.859359
	NO	267	299	138697	134	208	226	268	0.010007
	OTHER	10	24	634	69	15	17	19	0.912437
	UNK	91	66	1278	77	122	54	68	0.930524
	VEGETATION	10	15	457	12	12	88	9	0.854063
	VEHICLE HIT	15	31	596	20	24	7	87	0.888462
	Totals	519	590	143016	340	423	408	479	0.044308

- Table 3: Confusion matrix results for CFO sub-model

EFF - Confusion Matrix Results				
Actuals	Predicted			
		0	1	Error Rate
	0	198248	10376	0.049735
	1	4237	4873	0.465093
		202485	15249	0.067114

- Table 2 and

CFO - Confusion Matrix Results									
Actual	Predicted								
		ANIMAL	BALLOON	NO	OTHER	UNK	VEGETATION	VEHICLE HIT	Error
	ANIMAL	117	19	597	14	18	8	9	0.850384
	BALLOON	9	136	757	14	24	8	19	0.859359
	NO	267	299	138697	134	208	226	268	0.010007
	OTHER	10	24	634	69	15	17	19	0.912437
	UNK	91	66	1278	77	122	54	68	0.930524
	VEGETATION	10	15	457	12	12	88	9	0.854063
	VEHICLE HIT	15	31	596	20	24	7	87	0.888462
	Totals	519	590	143016	340	423	408	479	0.044308

- 3 provide the Confusion Matrix results for EFF Conductor and CFO sub-models respectively. It captures the accuracy rate as 93.29% and 95.57% 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.

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

The error rate for EFF Conductor and CFO sub-model turns out to be 6.71% and 4.43%. This means that the failure rate of the model prediction is low and under control.

All these test results are performed on a test dataset with a sampling of outage data from 2020 to August 8th, 2024.

A detailed explanation of the compensating controls employed for these limitations is 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 ML 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 and Figure 10, 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:

- Avg_Ckt_Pole_Age
- LENGTH_SEG_CAL_TOTAL
- CONDUCTOR_SIZE_UDF
- Log_WindForce
- max_wind_magnitude
- Conductor_AGE_UDF

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

- DOWNSTREAM_KVA
- CONDUCTOR_SIZE_UDF
- DOWNSTREAM_O_CUST
- Mean_Pole_Age
- LAT_UDF

The sensitivity of the model is examined using the test dataset, which contains 30% (141113 observations) for EFF and 20% (145775 observations) for CFO of the entire processed data. For this analysis, 10% of the test dataset (14094 observations for EFF and 14576 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 were performed in September 2024.

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 the combinations of these variables, eight different strata were picked. The test data is split into 10% and 90% based on proportional sampling of these strata. The 10% group is then assigned altered values for a selection of features before being rebound to the 90%. 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 4 and Table 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
Avg_Ckt_Pole_Age	124	0
LENGTH_SEG_CAL_TOTAL	227184	0
CONDUCTOR_SIZE_UDF	19	0
Log_WindForce	24.96733	14.95205
max_wind_magnitude	70.18979	27.76869
Conductor_AGE_UDF	124	0

Table 4: 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
DOWNSTREAM_KVA	47546	0
CONDUCTOR_SIZE_UDF	19	1
DOWNSTREAM_O_CUST	555	0
Mean_Pole_Age	71	0
AVIAN_2017	15.94884	0

Table 5: Extreme values used for Sensitivity testing in CFO sub-model

Table 6 and Table 7 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 3% with the exception of Avg_Ckt_Pole_Age, which is the most important feature in EFF. While this is larger than expected and warrants investigation, it is not so large as to indicate a problem with the model, especially considering that Avg_Ckt_Pole_Age is a feature that is constant for each circuit, so the distribution is a step function rather than a continuum.

AUC result of Unaltered Test Data from EFF Conductor sub-model	0. 8709
--	---------

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
Avg_Ckt_Pole_Age	0.8679	-0.30%	0.7923	-7.86%
LENGTH_SEG_CAL_TOTAL	0.8685	-0.24%	0.8452	-2.57%
CONDUCTOR_SIZE_UDF	0.8705	-0.04%	0.8646	-0.63%
Log_WindForce	0.8698	-0.11%	0.8530	-1.79%
max_wind_magnitude	0.8701	-0.08%	0.8597	-1.12%
Conductor_AGE_UDF	0.8702	-0.07%	0.8558	-1.51%

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

Southern California Edison (SCE) Model Documentation Prepared for 2026-2028 WMP Appendix B
OH Conductor Sub-Models (CFO & EFF)

AUC result of Unaltered Test Data from CFO sub-model	0.8069
--	--------

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
DOWNSTREAM_KVA	0.7974	-0.95%	0.8053	-0.15%
CONDUCTOR_SIZE_UDF	0.8064	-0.05%	0.8061	-0.08%
DOWNSTREAM_O_CUST	0.7966	-1.03%	0.8061	-0.08%
Mean_Pole_Age	0.7993	-0.76%	0.8032	-0.37%
AVIAN_2017	0.8004	-0.65%	0.8035	-0.34%

Table 7: The sensitivity results based on AUC for CFO sub-model

Table 8 and Table 9 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 8 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	4383	8123	200501	4727	48.11%	3.89%

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 Avg_Ckt_Pole_Age	-229	-286	286	229	45.60%	3.76%
Min Value for Avg_Ckt_Pole_Age	-106	-119	119	106	46.95%	3.84%
Max Value for Conductor_AGE_UDF	8	110	-110	-8	48.20%	3.95%
Min Value for Conductor_AGE_UDF	11	78	-78	-11	48.23%	3.93%
Max Value for CONDUCTOR_SIZE_UDF	25	121	-121	-25	48.39%	3.95%
Min Value for CONDUCTOR_SIZE_UDF	-62	-217	217	62	47.43%	3.79%
Max Value for LENGTH_SEG_CAL_TOTAL	-17	39	-39	17	47.93%	3.91%
Min Value for LENGTH_SEG_CAL_TOTAL	-4	81	-81	4	48.07%	3.93%
Max Value for Log_WindForce	-153	-533	533	153	46.43%	3.64%
Min Value for Log_WindForce	-168	-591	591	168	46.27%	3.61%
Max Value for max_wind_magnitude	13	94	-94	-13	48.25%	3.94%
Min Value for max_wind_magnitude	32	170	-170	-32	48.46%	3.98%

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

Southern California Edison (SCE) Model Documentation Prepared for 2026-2028 WMP Appendix B OH Conductor Sub-Models (CFO & EFF)

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	15.0%	14.1%	99.0%	8.76%	6.95%	14.6%	11.2%

True Positive Rate (TPR) for CFO Sub-model							
Sensitivity Test	Animal	Balloon	No	Other	Unknown	Vegetation	Vehicle
Max Value for DOWNSTREAM_KVA	14.96%	13.55%	99.02%	8.63%	6.83%	14.26%	11.15%
Min Value for DOWNSTREAM_KVA	14.71%	12.93%	99.07%	8.50%	6.83%	13.43%	10.26%
Max Value for CONDUCTOR_SIZE_UDF	14.96%	13.75%	99.01%	8.63%	6.89%	14.43%	11.15%
Min Value for CONDUCTOR_SIZE_UDF	14.96%	13.75%	99.02%	8.63%	6.83%	14.26%	10.90%
Max Value for DOWNSTREAM_O_CUST	14.96%	13.75%	99.04%	8.63%	6.78%	14.10%	10.77%
Min Value for DOWNSTREAM_O_CUST	14.83%	13.44%	99.05%	8.63%	6.83%	14.26%	10.51%
Max Value for Mean_Pole_Age	14.96%	13.75%	99.01%	8.63%	7.00%	14.59%	11.28%
Min Value for Mean_Pole_Age	14.71%	12.93%	99.03%	8.63%	6.78%	14.43%	10.77%
Max Value for AVIAN_2017	14.71%	13.75%	99.04%	8.63%	6.78%	14.59%	10.64%
Min Value for AVIAN_2017	14.58%	13.65%	99.03%	8.63%	6.95%	14.26%	11.15%

Table 9: 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 alter some of the predictions from the model. However, the magnitude of the impact seems to be low relative to the initial rates, especially considering the features selected are the most important to the model. 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, while 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 and Test data. The results of the train data are considered in-sample backtesting and the results of the test data are considered out-of-sample backtesting.

Once the ML 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 AUC value. Figure 5 and Figure 16 shows the AUC ROC for EFF Conductor and CFO sub-models based on the test dataset using historical outage data from 2020 to August 8th, 2024.

Southern California Edison (SCE) Model Documentation Prepared for 2026-2028 WMP Appendix B
OH Conductor Sub-Models (CFO & EFF)

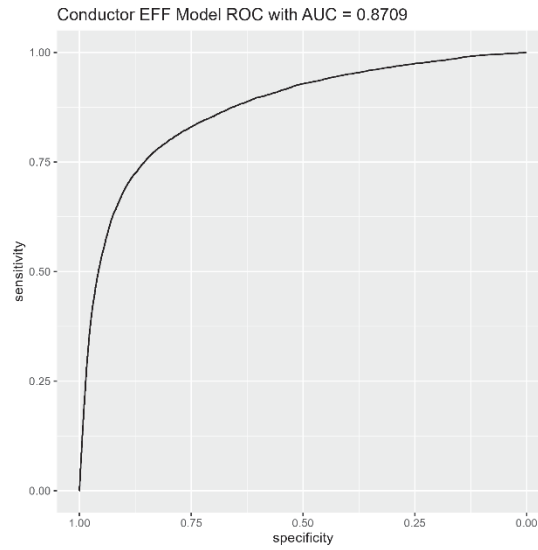


Figure 15: Out-sample backtesting result for EFF Conductor sub-model based on test dataset

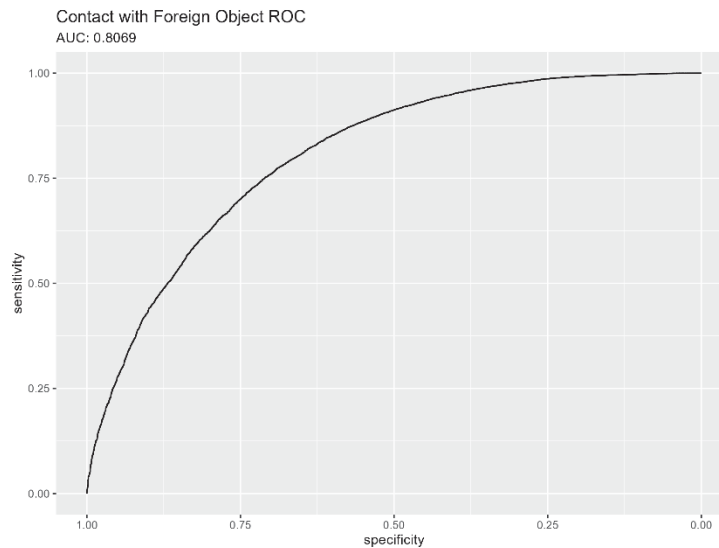


Figure 16: Out-sample backtesting result for overall CFO sub-model based on test dataset

The AUC values for the EFF Conductor and CFO sub-models are 0.8709 and 0.8069 respectively. The AUC values of both sub-models are higher than 0.8 which imply that the models possess high accuracy in terms of predicting the results.

Figure 17 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. Model tuning and sampling was biased toward better AUC scores for these subdrivers over the “NO” case.

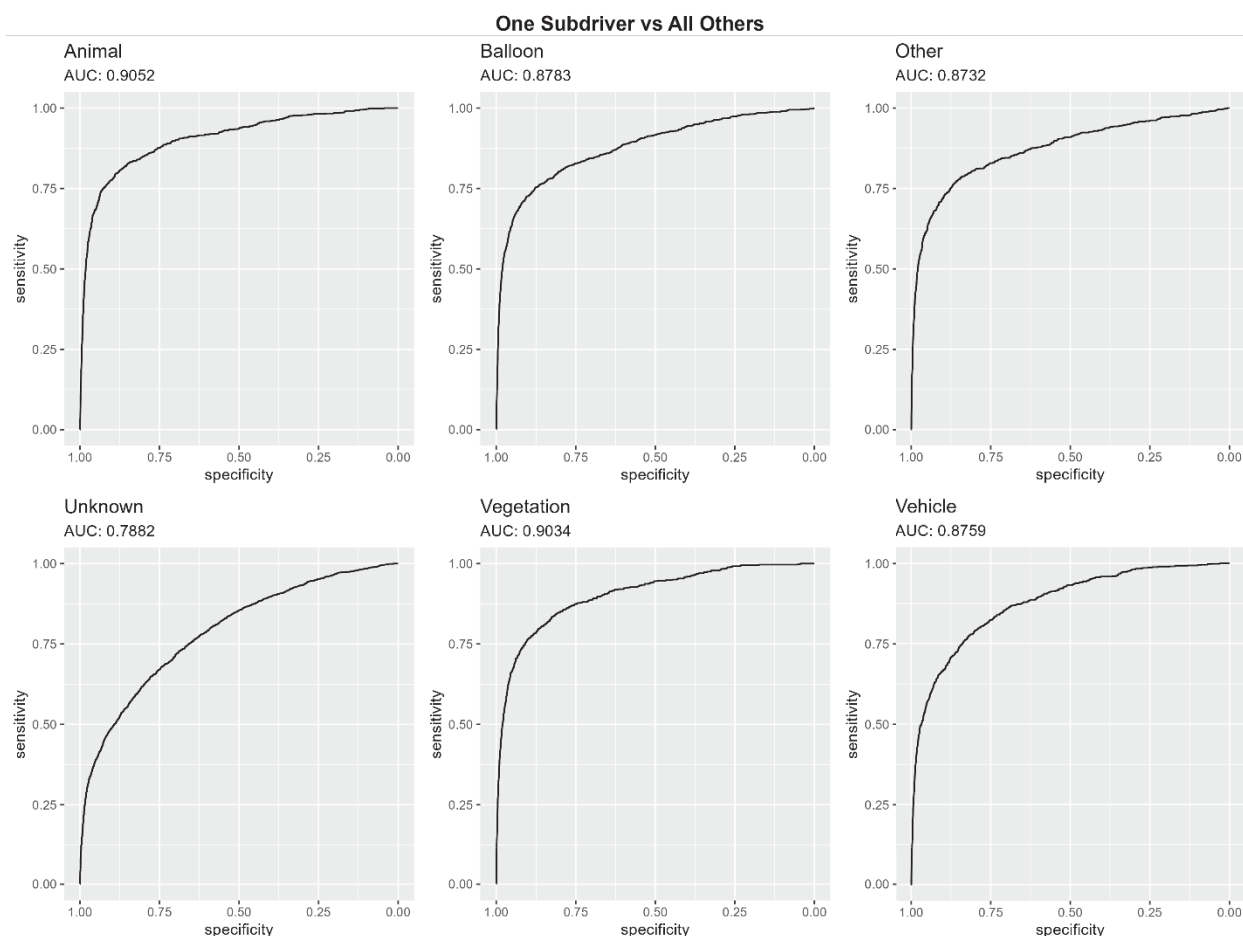


Figure 17: In-sample backtesting result for sub-drivers in CFO sub-model based on test dataset

The AUC values for sub-drivers Animal, Balloon, Other, Unknown, Vegetation and Vehicle are 0.9052, 0.8783, 0.8732, 0.7882, 0.9034, and 0.8759. The only sub-driver to experience significant AUC drift here is Unknown, which is being investigated.

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 (GLMM), and Directed Random Forest (DRF) were considered during the model development phase in 2019. The analysis on these supervised ML approaches and the results are provided below.

- **Gradient Boosting Machine (GBM)** is one of the most popular forward learning ensemble methods in ML. 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 (GLMM)** 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. GLMM is an extension of logistic regression known as a mixed-effects model that better incorporates binary predictors.
- **Random Forest (DRF)** is a popular ML 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 GLMM shared in this section were developed using the H2O library in R on the train and test data used in the Random Forest model. Since benchmark results were not saved during the model development phase, the benchmark models were executed in October 2024 for documentation purposes. Figure 18 and Figure 19 provide the AUC values for the EFF Conductor and CFO sub-models using the DRF, GBM, and GLMM methodologies.

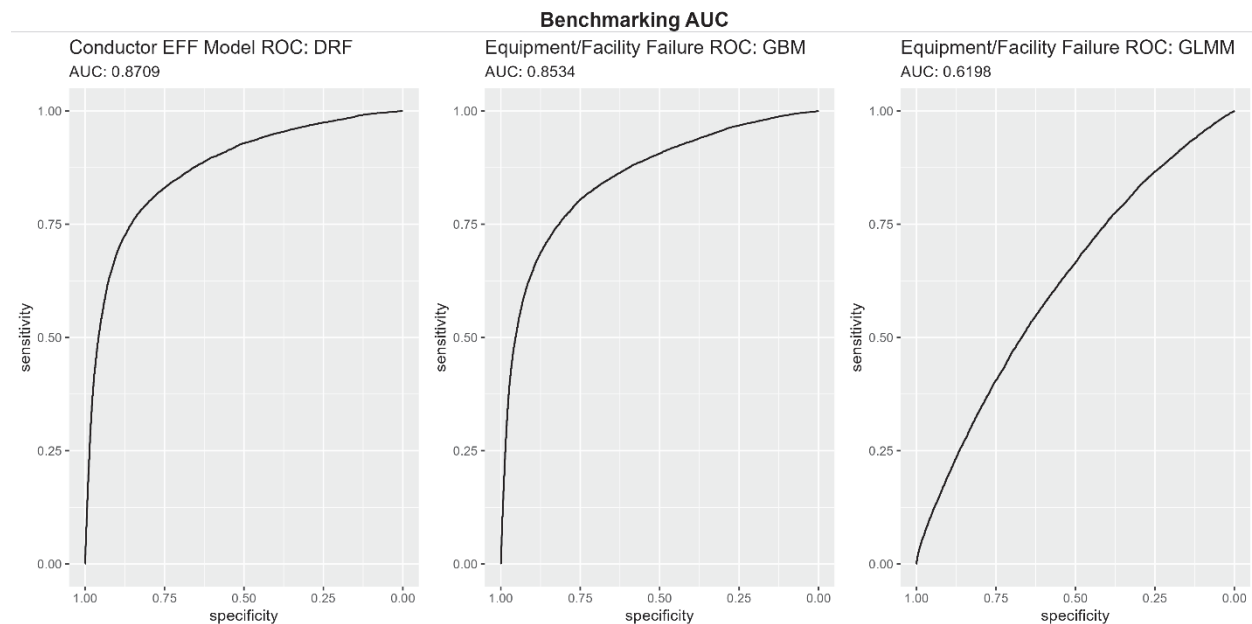


Figure 18: AUC Comparison for the EFF Conductor sub-model using DRF, GBM, and GLMM methodologies

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

Southern California Edison (SCE) Model Documentation Prepared for 2026-2028 WMP Appendix B
OH Conductor Sub-Models (CFO & EFF)

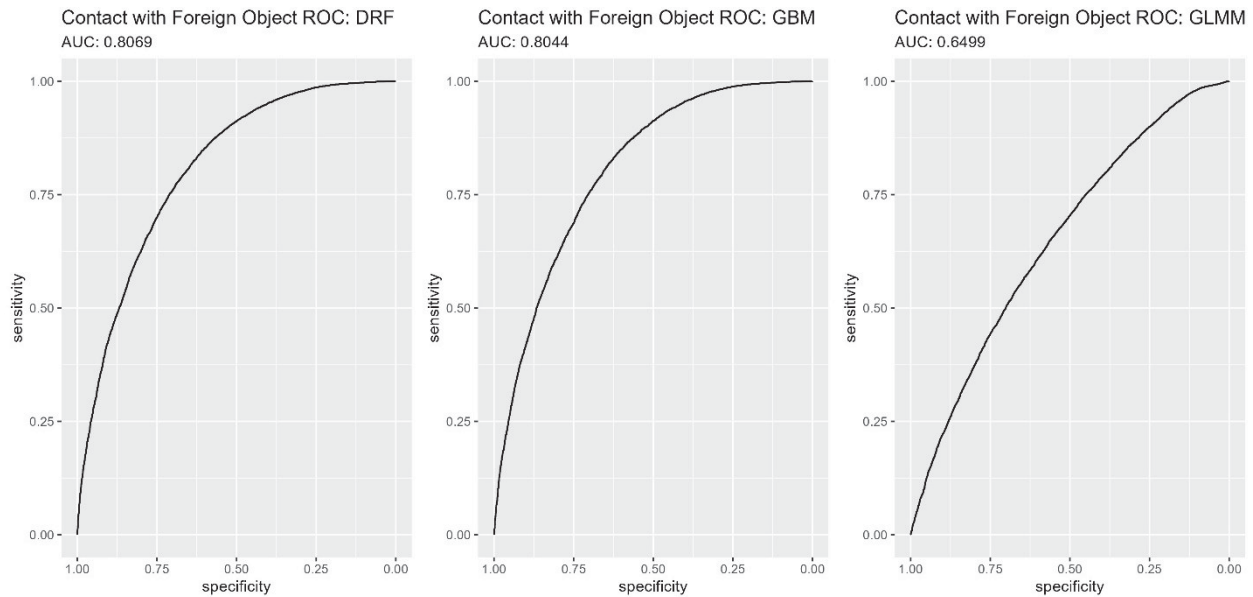


Figure 19: AUC Comparison for the CFO sub-model using DRF, GBM, and GLMM methodologies

For the CFO sub-model, the AUC results for Random Forest, Gradient Boosting Machine, and Logistic Regression were 0.8069, 0.8044, and 0.6499 respectively.

SCE chose Random Forest for the EFF Conductor and CFO sub-models since it achieved the highest AUC among the three approaches. SCE will continue to consider GBM as a methodology 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 ML models especially when 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 actions are taken to address them.

Performance monitoring is required only after running the model. The AUC and accuracy rate from Confusion Matrix results obtained after model refresh are compared against a threshold of 70%. 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, the 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 Advanced Predictive Modeling 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). The code is saved on 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: Fatality Analysis Reporting System (FARS) from the National Highway Traffic Safety Administration (NHTSA)

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

RF 3: Literature reference on grid search vs random search approach for hyperparameter tuning.

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(1), 281-305.

RF 4: Variable Importance methodology for tree-based methods

[Variable Importance — SHAP documentation](#)