PACIFIC GAS AND ELECTRIC COMPANY
PG&E Ref. DRU13643
Data Request
Office of Energy Infrastructure Safety
Requester DR No. OEIS-RMWG_2024-001

Requester: Biggs, Andie
Request Date: May 16, 2024
Response Date: June 11, 2024

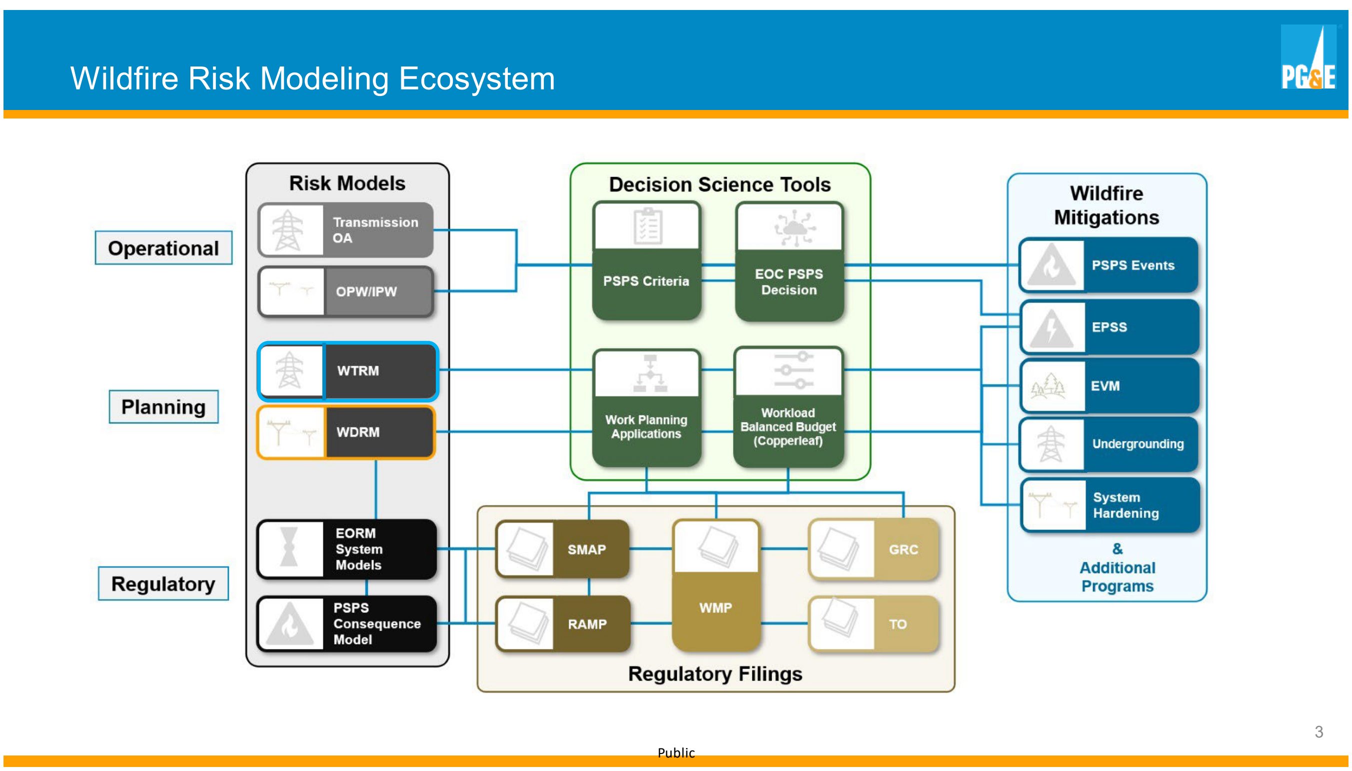
Question No. 001:

Regarding: visual depiction of risk models

Please provide a visual depiction of various models used by utilities and how such models are connected (e.g., swim lanes, flowchart).

Response to Question No. 001 Response No. 001:

The diagram below outlines the various risk models used by PG&E and their relationship with programs and regulatory processes.



Question No. 002:

Regarding: data usage by model

Please provide data usage broken down by model (e.g., vegetation model, conductor model, transformer model, etc.) using the example table provided below. Include the following data usages:

i. Scale and geographical context.
ii. Topography.
iii. Quality of historical outage, fault, and ignition data.
iv. Usage of outage and fault events to augment ignition data.
v. Integration of potential ignitions avoided due to PSPS events.
vi. Asset data (including asset age, health, inspection results, type, etc.).
vii. Impacts of system hardening and other initiative efforts.
viii. Climate conditions (include historical wind conditions, relative humidity, temperature, etc.).
ix. Vegetation (include type, density, height, etc.).
x. Fuel characteristics (include load, size, continuity, vertical arrangement, moisture, etc.).
xi. Impacts of routine and enhanced vegetation management activities (including tree trimming, tree removal, inspections, etc.).
xii. Frequency of updates to datasets and inputs, including any associated triggers to determine the need for updates.
xiii. Accuracy and quality checks for data and inputs.

Example of Table Illustrating Data Usage by Model

Data Usage	Model 1	Model 2	Model 3	Model 4
Scale and geographical context	EXAMPLE: 100m x 100m pixels across the HFTD Tier 2&3.			
Topography	EXMAPLE: USGS Topographic Position Index			
…				

Response to Question No. 002 Response No. 001:

The Wildfire Distribution Risk Model (WDRM) v4 provides individual sub-models for 22 discrete failure modes. These are generally categorized as asset or equipment models, vegetation models, and contact from object models. Information for this request are grouped in the table below according to these categories.

Data Usage	Equipment	Vegetation	Contact from object
Scale and geographical context	Each asset is represented	100m x 100m pixel across the entire PG&E service territory	100m x 100m pixel across the entire PG&E service territory
Topography	National Elevation Database	National Elevation Database	National Elevation Database
Quality of historical outage, fault, and ignition data	Relational according to failing equipment ID in company outage records	Lat and long and circuit segment according to company outage records	Lat and long and circuit segment according to company outage records
Usage of outage and fault events to augment ignition data	Test and training data sets	Test and training data sets	Test and training data sets
Integration of potential ignitions avoided due to PSPS events.	PSPS damages added to outage and ignition test and training data sets.	PSPS damages added to outage and ignition test and training data sets.	PSPS damages added to outage and ignition test and training data sets.
Asset data (including asset age, health, inspection results, type, etc.).	Age, type, location, inspection results Pole loading Remaining pole strength Pole volume Replacement open tag indicator Peak and avg. loading Splices	Age, type, location.	Age, type, location. Conductor size and material Region
Impacts of system hardening and other initiative efforts.	All system changes (hardened, replaced or removed assets, removed trees, etc) are recorded as new asset data. With time the model ascribes lower probabilities to these newer assets or conditions	All system and environmental changes (hardened, replaced or removed assets, removed trees, etc) are recorded as new asset data. With these improvements the model ascribes lower probabilities to these newer assets or conditions	All system and environmental changes (hardened, replaced or removed assets, removed trees, etc) are recorded as new asset data. With these improvements the model ascribes lower probabilities to these newer assets or conditions
Climate conditions (include historical wind conditions, relative humidity, temperature, etc.).	maximum wind speed windy summer day % soil slope available soil water capacity soil permeability annual soil flood frequency local topography is soil hydric soil clay content avg. specific humidity soil hydro-group avg. precipitation avg. wind speed avg. energy release avg. vapor pressure deficit	maximum wind speed windy summer day % soil slope available soil water capacity soil permeability annual soil flood frequency local topography is soil hydric soil clay content avg. specific humidity soil hydro-group avg. precipitation avg. wind speed avg. energy release avg. vapor pressure deficit	maximum wind speed windy summer day % soil slope available soil water capacity soil permeability annual soil flood frequency local topography is soil hydric soil clay content avg. specific humidity soil hydro-group avg. precipitation avg. wind speed avg. energy release avg. vapor pressure deficit

Data Usage	Equipment	Vegetation	Contact from object
	avg. daily max temperature soil organic matter soil horizon thickness percent gusty summer day	avg. daily max temperature soil organic matter soil horizon thickness percent gusty summer day	avg. daily max temperature soil organic matter soil horizon thickness percent gusty summer day
Vegetation (include type, density, height, etc.).	Strike tree count Avg. strike tree height maximum strike tree height minimum strike tree height canopy density tree species	Strike tree count Avg. strike tree height maximum strike tree height minimum strike tree height canopy density tree species	Strike tree count Avg. strike tree height maximum strike tree height minimum strike tree height canopy density tree species
Fuel characteristics (include load, size, continuity, vertical arrangement, moisture, etc.).	avg. 100-hr fuels avg. 1000-hr fuels avg. energy release available soil water capacity avg. burn index avg. vapor pressure deficit Dominant ground cover	avg. 100-hr fuels avg. 1000-hr fuels avg. energy release available soil water capacity avg. burn index avg. vapor pressure deficit Dominant ground cover	avg. 100-hr fuels avg. 1000-hr fuels avg. energy release available soil water capacity avg. burn index avg. vapor pressure deficit Dominant ground cover
Impacts of routine and enhanced vegetation management activities (including tree trimming, tree removal, inspections, etc.).	All system and environmental changes (hardened, replaced or removed assets, removed trees, etc) are recorded as new asset data. With these improvements the model ascribes lower probabilities to these newer assets or conditions	All system and environmental changes (hardened, replaced or removed assets, removed trees, etc) are recorded as new asset data. With these improvements the model ascribes lower probabilities to these newer assets or conditions	All system and environmental changes (hardened, replaced or removed assets, removed trees, etc) are recorded as new asset data. With these improvements the model ascribes lower probabilities to these newer assets or conditions
Frequency of updates to datasets and inputs, including any associated triggers to determine the need for updates	Risk models have historically been updated yearly. Now moving to a 3-year cycle coincident with WMP filings.	Risk models have historically been updated yearly. Now moving to a 3-year cycle coincident with WMP filings.	Risk models have historically been updated yearly. Now moving to a 3-year cycle coincident with WMP filings.
Accuracy and quality checks for data and inputs	AUC, top 20% concentration factor, and Precision-Recall curves. Cross fold validation runs on training data sets. Test code	AUC, top 20% concentration factor. Cross fold validation runs on training data sets. Test code	AUC, top 20% concentration factor. Cross fold validation runs on training data sets. Test code

<u>Question No. 003:</u>

Regarding: model descriptions

Please provide model descriptions for ignition, consequence, and PSPS models using the example table provided below. Include the following descriptions:

i. Algorithms used and machine learning capabilities.
ii. Inputs for the model.
iii. Outputs for the model.
iv. Description of any modules used, including but not limited to:
    (1) Climate change.
    (2) Ingress and egress.
    (3) Suppression.
    (4) Conflagration risks.
    (5) Smoke impacts.
    (6) Community vulnerability.
v. Modeling components, linkages, and interdependencies.
vi. Weight of each data component and input.
vii. Automatization implemented.
viii. Frequency of model updates, including the basis for each update.

Example of Table Illustrating Descriptions by Model

Descriptions	Ignition Model	Consequence Model	PSPS Model
Algorithms used and machine learning capabilities	EXAMPLE: Max Ent		
Inputs for the model			
. . .			

<u>Response to Question No. 003 Response No. 001:</u>

Much of the input and output data, along with modules and interdependencies requested are provided in graphical form per OEIS request in Appendix B.2 – High Level Bow Tie Schematics and Appendix B.3 High-Level Calculation Procedure Schematic in PG&E's 2023-2025 WMP.

Descriptions	Ignition	Consequence	PSPS Model
Algorithms used and machine learning capabilities.	Maximum Entropy XGBoost Logistic Regression	Binomial Regression MAVF	Balanced Random Forrest Classifier (FPI), CatBoost Decision Tree (OPW)
Inputs for the model	Asset, Meteorology and Vegetation data as described in response to Q002. Also, please see Figure B-5 Ignition Probability Calculation Procedure Schematic in	Historical ignitions Historical Meteorological data Also, please see Figure B-6 Wildfire Consequence Calculation Procedure	Please see Figure B-7 PSPS Calculation Schematic in PG&E's 2023-2025 WMP

Descriptions	Ignition	Consequence	PSPS Model
	PG&E's 2023-2025 WMP Historical outages, ignitions, and PSPS damages	Schematic in PG&E's 2023-2025 WMP	
Outputs for the model	Probability of outage, probability of ignition	MAVF risk units,	Probability of a small, large or catastrophic fire (FPI). Probability of outage per outage class (OPW). Probability of an Ignition (IPW)
Description of any modules used	Please see Figure B-5 Ignition Probability Calculation Procedure Schematic in PG&E's 2023-2025 WMP	Please see Figure B-6 Wildfire Consequence Calculation Procedure Schematic in PG&E's 2023-2025 WMP	Please see Figure B-7 PSPS Calculation Schematic in PG&E's 2023-2025 WMP
Modeling components, linkages, and interdependencies	Please see Figure B-5 Ignition Probability Calculation Procedure Schematic in PG&E's 2023-2025 WMP	Please see Figure B-6 Wildfire Consequence Calculation Procedure Schematic in PG&E's 2023-2025 WMP	Please see Figure B-7 PSPS Calculation Schematic in PG&E's 2023-2025 WMP
Weight of each data component and input	Machine learning models	Weighted according to historical fires.	Machine learning models
Automatization implemented	Semi-automated planning model	Semi-automated planning model	Automated operational model
Frequency of model updates, including the basis for each update	Annual in the past and every three years going forward.	Annual in the past and every three years going forward.	One to three years (FPI), Annual (OPW)

Question No. 004:

Regarding: model outputs

Please provide how model outputs are analyzed and utilized for each model using the example table provided below. Include:

i. Confidences for each modeling component, including how such confidences were determined.
ii. Range of uncertainty for model outputs, including how those ranges are determined and how uncertainty is minimized.
iii. Systems used to verify the model outputs, including verifier (subject matter experts, third-party)and mechanisms for implementing lessons learned.
iv. How uncertainty affects the interpretations of model outputs.
v. Determination of highest risk areas based on model outputs.
vi. Use of subject matter expertise for inputs and further verification.
vii. Scaling of outputs in final determinations.
viii. Risk tolerances used for decision-making.

Example of Table Illustrating Outputs by Model

Output	Model 1	Model 2	Model 3	Model 4
Confidences for each modeling component, including how such confidences were determined	EXAMPLE: Receiver Operating Characteristic (ROC) /Area Under the Curve (AUC/ROC)			
Range of uncertainty for model outputs, including how those ranges are determined and how uncertainty is minimized	EXAMPLE: Evaluation of ROC/AUC/ROC, Precision, and Recall values			
…				

<u>Response to Question No. 004 Response No. 001:</u>

Output	Equipment	Vegetation	Contact From Object
Confidences	AUC/ROC and Precision-Recall	AUC/ROC	AUC/ROC
Range of uncertainty	AUC/ROC and Precision-Recall	AUC/ROC	AUC/ROC
Systems used to verify the model outputs	Test set predictions, AUC/ROC and Precision-Recall curves, third party validation	Test set predictions, AUC/ROC curves, third party validation	Test set predictions, AUC/ROC curves, third party validation
How uncertainty affects the interpretations of model outputs	Based on AUC/ROC scores workplans will adjust application of the model results	Based on AUC/ROC scores workplans will adjust application of the model results	Based on AUC/ROC scores workplans will adjust application of the model results
Determination of highest risk areas based on model outputs	Depending on workplan, mean circuit segment risk value or individual assets values	Mean circuit segment risk value	Mean circuit segment risk value
Use of subject matter expertise for inputs and further verification	Models are developed with SME review from input data selection, initial model tuning and development to final validation and approval	Models are developed with SME review from input data selection, initial model tuning and development to final validation and approval	Models are developed with SME review from input data selection, initial model tuning and development to final validation and approval

Output	Equipment	Vegetation	Contact From Object
Scaling of outputs in final determinations	Ignition probabilities are calibrated to annual average counts. Risk values are calibrated to enterprise risk values in RAMP filing	Ignition probabilities are calibrated to annual average counts. Risk values are calibrated to enterprise risk values in RAMP filing	Ignition probabilities are calibrated to annual average counts. Risk values are calibrated to enterprise risk values in RAMP filing
Risk tolerances used for decision-making	Based on AUC/ROC scores workplans will adjust application of the model results	Based on AUC/ROC scores workplans will adjust application of the model results	Based on AUC/ROC scores workplans will adjust application of the model results

<u>Question No. 005:</u>

Regarding: description of any collaborations among the utilities

Please provide a description of all collaborations previously undertaken among the utilities, as well as details on any known consistency across utilities, including:

i. What modeling approaches are already consistent.
ii. Which modeling approaches have the potential for more consistency and how approaches would benefit from consistency.
iii. Where consistency is infeasible or not necessary.

<u>Response to Question No. 005 Response No. 001:</u>

i. What modeling approaches are already consistent.

PG&E's wildfire modeling approach follows the framework outlined in the Wildfire Mitigation Plan Guidance document. Specifically, risk is represented by PG&E's Multi Attribute Value Function (MAVF) as prescribed by the CPUC.

As outlined in section 6.2 Risk Analysis Framework of PG&E's 2023-2025 WMP, within this framework wildfire risk is a product of ignition likelihood and ignition consequence. SDG&E, SCE, and PG&E all employ a range of machine learning models and sub-models to produce a probability of ignition result. Please see pages 156 – 157 of PG&E's 2023-2025 WMP for a more detailed explanation of PG&E's ignition probability models.

For wildfire consequence SDG&E, SCE, and PG&E all employ variations of the Technosylva fire spread simulation results to represent potential consequence spatially along electric grid assets.

ii. Which modeling approaches have the potential for more consistency and how approaches would benefit from consistency.

PG&E does not see the need for an increased alignment on risk models because the service territories of SDG&E, SCE, and PG&E are varied and diverse. As all three utilities continue to experiment and explore methods for improving the predictive power and application of model results it is anticipated that methods will more closely align.

iii. Where consistency is infeasible or not necessary.

Note that model consistency does not necessitate optimal model effectiveness across all California utilities. In Australia, the wildfire consequence models are provided by a university team specialized in fire science and the electric utilities still use variations of this model in conjunction with their probability of ignition models to construct wildfire models. This benchmarked example might suggest that variation in models provides the ability to adapt to the specific terrain and wildland urban interface (WUI) characteristics of each California service territory.

Question No. 006:

Regarding: description of any additional collaborations

Please provide a description of all collaborations previously undertaken and/or ongoing with other entities.

Response to Question No. 006 Response No. 001:

In line with our company stand that catastrophic wildfire will stop, PG&E continues to collaborate with other entities in the utility and wildfire community. On May 16, 17, 2024 PG&E, SCE, and SDG&E hosted the Near-Term Wildfire Mitigations Conference which was attended by representatives of over 45 US and international utilities. In addition to the collaboration detailed in ACI PG&E-22-02 – Collaboration and Research in Best Practices in Integrating Climate Change Impacts and Wildfire Risk and Consequence Modeling and the table 8-44: State and Local Agency Collaboration(s) in PG&E's 2023 – 2025 WMP, below is a listing of recent collaborations.

CPUC

- Risk Informed Decision-Making Proceeding (RDF) F.20-07-0134 Working Groups
- Climate Change Adaptation Proceedings R.18-04-019 Working Groups

OEIS

- OEIS Risk Modeling Working Group
- OEIS Working Group on Social and Community Vulnerability
- OEIS Working Group on Climate Change

External Organizations:

- International Wildfire Risk Mitigation Consortium Risk Subcommittee
- California Department of Fire and Forestry (CalFire)
    - Wildfire Mitigation Advisory Committee
- Wildfire Interdisciplinary Research Center at San Jose State University

Question No. 007:

Regarding: attachments

Please provide attachments of:

i. All internal or third-party validations completed, and
ii. Description of any peer review of risk models utilized.

Response to Question No. 007 Response No. 001:

As detailed in section 6, pages 212 – 214 of PG&E's 2023-2025 Wildfire Mitigation Plan, Energy and Environmental Economics (E3) conducted a validation and review of the Wildfire Distribution Risk Model. In addition, in response to ACI PG&E-22-07- Applying Modeling Lessons-Learned from Third-Party Review. The E3 report was provided as supporting documentation in PG&E's 2023-2025 Wildfire Mitigation Plan as 2023-03-27_PGE_2023_WMP_R0_Section 6.6.1_Atch01 and is also attached to this response as *DRU13643_Q007_Atch01_E3 Review.pdf*.