PACIFIC GAS AND ELECTRIC COMPANY
Wildfire Mitigations Plans Discovery 2026-2028
Data Response

PG&E Data Request No.:	OEIS_022-Q001
PG&E File Name:	WMP-Discovery2026-2028_DR_OEIS_022-Q001
Request Date:	December 22, 2025
Requester DR No.:	OEIS-P-WMP_2025-PG&E-022
Requesting Party:	Office of Energy Infrastructure Safety
Requester:	Nathan Poon
Date Sent:	December 26, 2025

SUBJECT: REGARDING PG&E'S STRIKE TREE SPECIES MODEL PREDICTION

QUESTION 001

On page 85 of the draft Decision for the PG&E 2026-2028 Base WMP, PGE-26B-18 requires PG&E in its next WMP Update to include "[t]he number of outages caused by each [tree] species per 1,000 trees in PG&E's records of that species." In PG&E's Opening Comments on the draft Decision, it requested that Energy Safety modify PGE-26B-18 to require PG&E calculate "[t]he Number of outages caused by each species per 1,000 trees in its Strike Tree Species Model Prediction." PG&E indicated that this modification would improve the specificity of the requirement.

a. Regarding PG&E's Strike Tree Species Model:
    i. Describe the model's history (e.g., when it was created, who created it, and what dataset(s) informs its predictions) and purpose (i.e., how PG&E currently uses the model and plans to use the model).
    ii. List and describe the model's inputs (e.g., environmental factors, plant species, plant age, plant height, etc.)
    iii. List and describe the model's outputs (e.g., probability of outage by plant species or genus, probability of ignition by plant species or genus, etc.)
    iv. Is the model able to predict the species- or genus-specific probability of an outage due to both grow-in and fall-in events?
        1. If yes, is PG&E able to divide this prediction into separate grow-in and fall-in outage probabilities?
        2. If no, what types of vegetation failure modes does the model's output take into account to calculate the species- or genus-specific outage probability?
    v. What are the advantages and disadvantage of using PG&E's Strike Tree Species Model Prediction in place of "PG&E's records" to calculate the number of outages by each tree species or genus per 1,000 trees?

ANSWER 001

a. Responses Below:

i. The Tree Species Model, developed jointly by PG&E's Vegetation Management and Risk Management organizations, was initiated in 2023 and continued through 2024. The purpose was to create a machine learning-based tree species identification system to support wildfire mitigation strategies.

The model identifies how vegetation types across different ecoregions in High Fire-Threat Districts (HFTDs) impact mitigation planning. This per-tree model provides an updateable baseline of species variation across PG&E's service territory, providing input to future risk models (e.g., Wildfire Distribution Risk Model (WDRM), in addition to supporting future program development and planning.

ii. The initial model inputs include:

- LiDAR point cloud data — HFTD collection of ~5 million data points for tree geometry and proximity mapping.
- EVM Inspection data — species-level details from PG&E's vegetation management program, used as labels for training the model.
- Satellite Data — surface reflectance imagery for canopy height, cover, and mortality indicators. 8 bands, 4.7m resolution.
- Internal Wildfire Risk Models — various wildfire risk covariates.

iii. The initial raw model outputs include:

1. Predicted tree species label:
    - One of 20+ possible species classes.
2. Sensitivity/Confidence scores:
    - Probability or confidence of the tree belonging to each class.

Tree features like canopy height are also extracted through LiDAR point cloud data and each tree is represented as a cluster of points rather than a single pixel.

iv. This model is not intended to predict the species- or genus-specific probability of an outage due to grow-in and fall-in events.

1. N/A
2. This model could feed other models that account for vegetation failure to develop a species- or genus-specific probability of outage model. Currently, it includes outages caused by vegetation contact events, without the explicit separation of grow-in and fall-in categories.

   Instead, PG&E uses other models (like OPW/IPW and WDRM vegetation sub-models) to handle failure-mode differentiation, by 1) breaking down vegetation-caused outages into branch, trunk, and

other categories, 2) incorporating LiDAR and satellite-derived metrics to estimate fall-in risk, and 3) using weather covariates (i.e., wind speed, soil moisture) for dynamic failure probability modeling.

Additionally, this data is better suited for fall-in outage probabilities, as its prediction is associated with strike tree species determination. Any estimate of grow-in outage probabilities would be inferred from the strike tree species output.

v. The main advantage of using PG&E's Tree Species Model is that it provides comprehensive and consistent species identification at single-tree level across HFTDs. This level of granularity could be used as the "denominator" of a probability-of-outage estimate, as outage probabilities can be severely skewed by counting only trees recorded in PG&E records.

   The disadvantage of using only PG&E records is that they predominantly rely on trees identified as needing maintenance or mitigation. Other than the subset of miles inspected and documented in our 2023–2025 Focused Tree Inspection (FTI) program and our continuing 2026–2028 Elevated Inspections in areas of high risk, all other miles are largely based on recording trees that require work.