Data-Driven Analysis of Wildfire Post-Event Reports: Patterns, Causes, and Mitigation Strategies

Sarah Sami Khan*, Hamed Mohsenian-Rad†, Masood Parvania*
* Department of Electrical and Computer Engineering, The University of Utah, Salt Lake City, UT, USA
† Department of Electrical and Computer Engineering, The University of California, Riverside, CA, USA
Emails: {sarah.khan, masood.parvania}@utah.edu, hamed@ece.ucr.edu

***Abstract*—Public Safety Power Shutoffs (PSPS) are a critical yet disruptive wildfire mitigation strategy used by electric utilities to reduce ignition risk during periods of elevated fire danger. However, current PSPS decisions often lack transparency and consistency, prompting the need for data-driven tools to better understand utility behavior. This paper presents a Support Vector Machine (SVM) framework to model and interpret PSPS decision-making using post-event wildfire reports. Forecast-based weather and fire behavior features are used as model inputs to represent decision-relevant variables reported by utilities. The model is calibrated using Platt scaling for probabilistic interpretability and adapted across utilities using importance-weighted domain adaptation to address feature distribution shifts. A post-hoc clustering segments PSPS events into wildfire risk zones based on wildfire risk metrics excluded from model training. Results demonstrate that the proposed framework supports interpretable, transferable analysis of PSPS decisions, offering insight into utility practices and informing more transparent de-energization planning.**

***Index Terms*—Wildfire prevention, post-event analysis, data-driven modeling, risk-based clustering, Public Safety Power Shutoffs, importance-weighted support vector machine.**

I. Introduction

A. Background and Related Works

It is increasingly challenging to maintain a reliable operation of the power grid due to an increase in high-impact, low-probability (HILP) weather events such as wildfires, heatwaves and floods [1]. Wildfires especially pose unique risks due to their ability to spread rapidly and directly threaten power infrastructure [2]. Power line ignitions are particularly severe as they often occur during elevated fire danger conditions, exacerbating spread and damage [3], [4]. For instance, the 2021 Dixie Fire—ignited by distribution lines—became the largest single-source wildfire in California history, burning over 960,000 acres and destroying more than 1,300 structures [5]. Similarly, the 2025 Southern California wildfires, notably Eaton Fire, have been linked to alleged equipment failures, resulting in the destruction of over 9,000 buildings and 18 fatalities [6]. 55.1% of utility infrastructure assessed in the aftermath was destroyed, along with widespread damage to critical facilities. These statistics highlight the disproportionate vulnerability of grid infrastructure and the cascading societal impact of power line ignitions. These events underscore the need for better strategies to mitigate wildfire disasters, especially those aimed at minimizing grid-initiated ignitions.

This work was supported in part by the National Science Foundation under Grant Number 2330582, and by the University of California Riverside's Winston Chung Endowed Chair Professorship on Energy Innovation.

To reduce ignition risk from electrical infrastructure, power utility companies implement *Public Safety Power Shutoffs* (PSPS), proactively de-energizing power lines under extreme weather and dry fuel conditions. While effective in preventing fire starts, PSPS events often impact thousands of customers. Between 2013 and 2020, California utilities executed 51 PSPS events, affecting over 3.2 million customers and incurring billions in economic and social losses [7]. For instance, a recent study projects an annual average of 1.6 million person-days of de-energization under future climate conditions [8]. Rising wildfire mitigation costs have translated into rate increases, and general rate adjustments added an estimated $32 to the average monthly household bill in 2024 [9]. In order to promote transparency and public accountability of utility PSPS decisions, regulatory agencies such as California Public Utilities Commission (CPUC) are starting to require detailed post-event reports for all PSPS actions [5].

Wildfire modeling efforts span a broad spectrum, from physics-based fire spread simulations to probabilistic ignition forecasts. Prior work has developed machine learning models to assess wildfire risk and weather-induced outages using ensemble trees, regression, and deep learning [10]–[14]. However, these studies generally focus on forecasting ignition likelihood or outage occurrence rather than interpreting utility actions. Optimized shutoff strategies also exist, involving mixed-integer and stochastic optimization to balance ignition risk against load shed [15]–[17].

A small body of work has begun to directly address PSPS decisions. For instance, Multi-Attribute Value Function (MAVF)-based frameworks have been adopted by utilities to quantify wildfire-versus-shutoff risk tradeoffs from a planning perspective [18]. Recent reviews also propose guideline taxonomies for proactive shutoffs [19]. However, existing models lack validation against actual post-event outcomes. They also often rely on predefined thresholds or deterministic logic.

B. Contribution and Paper Structure

This paper presents a data-driven, behavioral modeling framework for analyzing how electric utilities mitigate wildfire risks through PSPS decisions. Rather than predicting wildfire ignitions, the objective of this framework is to diagnose and explain utility decision-making—i.e., how forecasted environmental and fire-behavior conditions influence the execution or cancellation of PSPS events. The proposed method differs from traditional predictive models by focusing on interpretability and diagnostic insight into utility decision-making processes, while also enabling cross-utility evaluation based on publicly available data.

The methodology is applied to two major California utilities—Pacific Gas and Electric (PG&E) and San Diego Gas and Electric (SDG&E)—as representative use cases. A Support Vector Machine (SVM) is trained on PG&E post-event reports, capturing decision boundaries based on forecast inputs. The SVM margins are scaled for probabilistic interpretation using Platt scaling. To account for class imbalance and domain shift—since SDG&E reports only executed events—an importance-weighted domain adaptation strategy aligns PG&E-trained models to SDG&E's feature distribution, supporting cross-utility generalization.

To move beyond binary classification, a post-hoc clustering segments the dataset into risk-informed decision clusters using wildfire consequence metrics and utility-defined risk and benefit scores-excluded from model training. This enables a diagnostic evaluation of how model confidence aligns with operational behavior across different wildfire risk zones.

The overall framework—illustrated in Fig. 1—includes feature preprocessing, SVM training and calibration, cross-utility domain adaptation, and post-hoc analysis. In summary, the contributions of this work are as follows:

- A behavioral decision-analysis framework for interpreting real-world PSPS actions, using case studies from two major California IOUs (Investor-Owned Utilities).
- Development of a calibrated and domain-adapted SVM classifier to uncover decision boundaries under feature shift and class imbalance.
- A post-hoc risk-based clustering approach to evaluate alignment between modeled confidence and real-world operational behavior.

The remainder of this paper is structured as follows: Section II presents the data construction, preprocessing, and SVM-based decision boundary formulation. Section III provides an extended ablation study, classifier benchmarking, and robustness testing under extreme wildfire scenarios. Section IV introduces a post-hoc risk space analysis to evaluate model confidence across wildfire risk zones Finally, Section V concludes the study and discusses directions for future research.

II. Data Framework and Model Development

This section outlines the construction of the dataset, processing steps, and the interpretable learning framework used to identify PSPS decision boundaries from post-event data.

A. Data Framework

1) Data Sources and labeling: The primary data sources include post-event wildfire reports from PG&E and SDG&E, covering 2021–2024. Each record corresponds to a unique circuit-day event, labeled as executed (+1) if a de-energization occurred, or canceled (–1) if a PSPS notification was issued

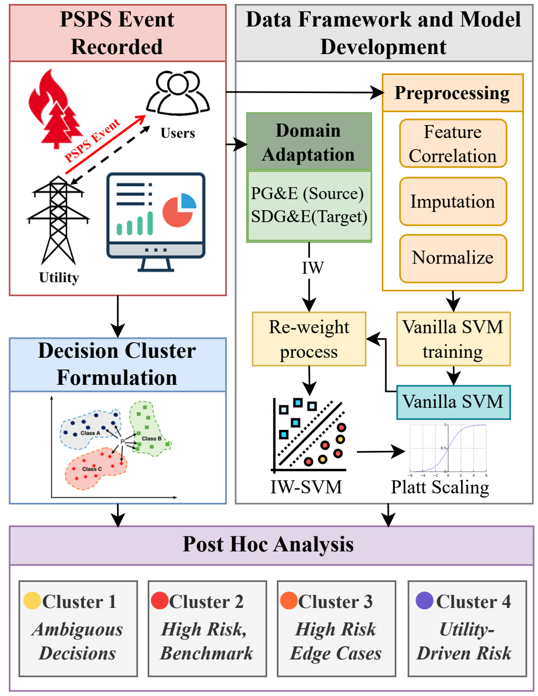

Fig. 1. Overview of the proposed PSPS decision-space analysis framework using SVM-based modeling and post-hoc clustering evaluation.

TABLE I
Comparison of PSPS Factors Reported by PG&E and SDG&E [5]

ID	Feature	Unit	PG&E	SDG&E
Weather Conditions				
X1	Sustained Wind Speed (10m)	mph	> 19	> 25
X2	Sustained Wind Speed (50m)	mph	Included	Not Reported
X3	Peak Wind Gust (Forecasted)	mph	> 30	> 35
X4	Temperature	°F	Included	Included
X5	Relative Humidity	%	< 30	< 20
X6	Vapor Pressure Deficit (2m)	mb	Included	Not Reported
Fuel Moisture				
X7	Dead Fuel Moisture (10-hr)	%	< 9	< 7
X8	Dead Fuel Moisture (100-hr)	%	< 12	< 10
X9	Dead Fuel Moisture (1000-hr)	%	< 11	Not Reported
X10	Live Fuel Moisture – Herbaceous	%	< 65	Not Reported
X11	Live Fuel Moisture – Woody	%	Included	Not Reported
X12	Live Fuel Moisture – Chamise	%	< 90	< 79
Fire Behavior				
X13	Flame Length (2hr forecast)	ft	> 10	Not Reported
X14	Rate of Spread (2hr forecast)	ch/hr	> 30	Not Reported
X15	Area Burned (8hr forecast)	acres	Included	Not Reported
X16	Fire Potential Index (FPI)	Probability ×10	< 0.22	Included
Vegetation				
X17	Tree Overstrike Potential	ft	Included	Not Reported

but later rescinded. The datasets include forecast-based environmental variables and fire behavior indicators (see Table I). The analysis focuses exclusively on distribution-level circuits, which face greater infrastructure exposure and pose elevated risks to surrounding communities [1].

2) Dataset Challenges: The dataset poses several modeling challenges: (i) limited sample size due to restricted public release of detailed post-event PSPS data, particularly from some utilities such as Southern California Edison (SCE), which limits generalization and increases the risk of overfitting; (ii) severe class imbalance, with canceled events representing only 11.4% of dataset; (iii) domain shift, as evident from discrepancies in feature distributions between utilities; and (iv) a high-dimensional input space, that increases complexity.

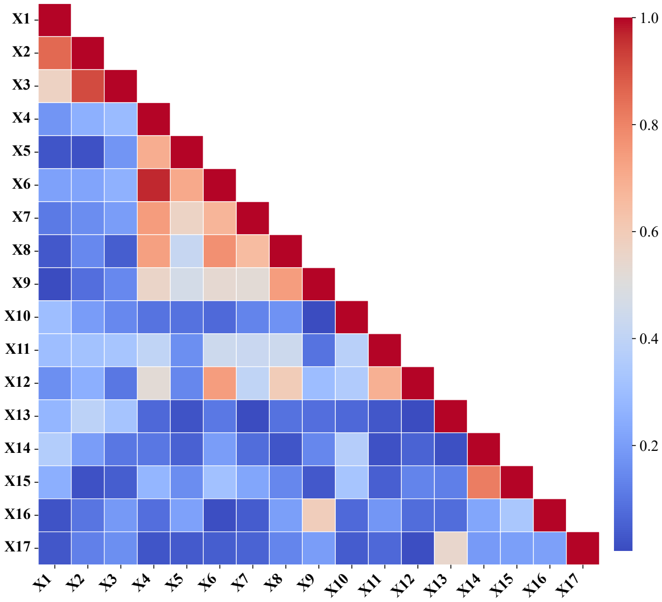

Fig. 2. Correlation heatmap of input features representing factors considered in PSPS decision-making. Highly collinear features ($\rho > 0.85$) were removed prior to model training. Specifically, *X2*, *X6*, and *X9* were removed due to their high redundancy with other retained features. See Table I for variable descriptions and activation thresholds defined by different utilities.

3) Feature Engineering and Preprocessing: To refine the input feature set, pairwise Pearson correlations were computed and visualized via a heatmap to highlight redundant and unique variable relationships, as illustrated in Fig. 2. The Pearson correlation coefficient is given by

$$\rho_{A,B} = \frac{\mathrm{cov}(A,B)}{\sigma_A \sigma_B}. \tag{1}$$

Missing data points are imputed using feature-wise means, and features exhibiting heavy-tailed distributions (e.g., X15 and X17) are log-transformed using $\log(1+x)$ to reduce skewness. All features are then normalized using robust scaling, defined as $z = (x - \text{median})/\text{IQR}$, where IQR is the interquartile range between the 75th and 25th percentiles. The PG&E dataset is split into 70% training and 30% testing subsets for model development and in-domain evaluation, while the positive-class SDG&E dataset is reserved entirely for cross-utility generalization.

B. SVM for Learning Utility Decision Boundaries

1) Baseline SVM Modeling: A binary classifier—referred to as the "vanilla" SVM in Fig. 1—is trained using PG&E's labeled data to capture operational decision boundaries. The model uses fixed training parameters without hyperparameter tuning, to preserve transparency and avoid overfitting given the limited dataset. The SVM formulation minimizes the following objective function:

$$\begin{aligned} \min_{\omega,b,\xi} \quad & \lambda\|\omega\|^2 + C\sum_{n=1}^{N}\xi_n, \\ \text{subject to} \quad & y_n\left[\langle \omega^T, x_n\rangle + b\right] \geq 1 - \xi_n, \\ & \xi_n \geq 0, \quad \forall n \end{aligned} \tag{2}$$

TABLE II
CLASSIFICATION RESULTS: VANILLA VS. WEIGHTED SVM

Class	Vanilla SVM			Weighted SVM		
	Precision	Recall	F1-Score	Precision	Recall	F1-Score
Canceled (-1)	0.75	0.50	0.60	0.35	0.84	0.50
Executed (+1)	0.94	0.98	0.96	0.97	0.80	0.88
Accuracy	0.925 (95% CI: [0.881, 0.962])			0.806 (95% CI: [0.771, 0.837])		

This formulation allows margin violations through slack variables ξ_n and balances margin maximization with classification accuracy using penalty term C [20]. The resulting model captures a behavioral boundary reflecting utility decision logic.

2) Margin-Based Decision Function Interpretation: A SVM classifier is trained using a custom `SVM_Classifier()` implementation. The model learns a weight vector $\boldsymbol{\omega}$ and bias b that define the separating hyperplane. The decision function, defined as the signed margin, quantifies how confidently a sample is classified :

$$f(x) = \boldsymbol{\omega}^\top x - b. \tag{3}$$

The final label is determined as $\hat{y} = \mathrm{sign}(f(x))$ [21]. Model performance is evaluated on PG&E test set using standard classification metrics—precision, recall, F1-score, and accuracy—as defined in [22]. Table II summarizes these metrics. The "vanilla SVM" serves as a baseline without domain adaptation or reweighting, enabling comparison with the IW-SVM model described in Section II-D.

C. Confidence Calibration

To convert SVM margins into interpretable confidence scores, we compare two techniques. Both map the margin $m_i = \mathbf{w}^\top x_i - b$ into a pseudo-probability, but differ in flexibility [23], formulated as following:

1) Sigmoid Transformation: The standard sigmoid function applies a fixed scaling to the margin:

$$P(y = 1 \mid x_i) = \frac{1}{1 + \exp(-a m_i)}. \tag{4}$$

with $a = 1$. This method is computationally efficient but does not guarantee that the output aligns with observed execution frequencies [23].

2) Logistic Calibration: Commonly referred to as *Platt scaling* —fits a sigmoid function to validation data as:

$$P(y = 1 \mid x_i) = \frac{1}{1 + \exp(-(\alpha m_i + \beta))}. \tag{5}$$

where α and β are learned parameters. Compared to (4), this data-driven mapping adjusts for over-confidence or under-confidence, yielding better-calibrated probabilities [23].

3) Final Selection: To transform SVM margin outputs into interpretable probabilities, we compare two scoring approaches on (4) and (5). Calibration performance is summarized in Table III. For PG&E, Platt scaling significantly outperforms the standard sigmoid in terms of Brier score and log loss, indicating better probability calibration. For SDG&E, only the Brier score is applicable due to the absence of a negative class; Platt scaling again yields the lowest score.

TABLE III
CALIBRATION METRICS COMPARISON (PG&E AND SDG&E)

Model	Brier Score	Log Loss	AUC
PG&E - Sigmoid	0.2488	0.6907	0.7898
PG&E - Platt	**0.1016**	**0.3564**	0.7898
SDG&E - Sigmoid	0.1676	N/A	N/A
SDG&E - Platt	**0.0105**	N/A	N/A

Given these results, Platt scaling is selected as the preferred method for producing probability confidence scores.

D. Importance-Weighted SVM (IW-SVM) model

The vanilla SVM—trained solely on PG&E data without adaptation—achieves superficially high accuracy on SDG&E due to the presence of only executed (positive class) events. However, this 100% classification accuracy is misleading, as the model faces no negative examples and thus defaults to classifying all instances as positive. The lack of canceled events prevents meaningful performance assessment using standard binary metrics and reveals the limitations of direct transfer in imbalanced domains. This discrepancy underscores the need for domain adaptation to correct for covariate shift and enable more reliable cross-utility generalization.

1) Domain Shift Analysis: To assess cross-utility mismatches, we compare the marginal feature distributions between PG&E and SDG&E datasets, as shown in Fig. 3. The overlaid histograms and kernel density estimates reveal distinct shifts in variables such as fuel moisture, relative humidity, and temperature. These distributional differences can significantly distort model behavior when applied across domains, further motivating the use of domain adaptation to calibrate decision logic for SDG&E-like operational contexts.

2) IW Domain Adaptation: The `SVM_Classifier` was modified to accept per-sample weights for domain adaptation using a class-conditional reweighting scheme. Specifically, the PG&E training samples are assigned weights that quantify their similarity to the SDG&E distribution, computed only for the positive class (executed) since the target domain lacks canceled events. The method proceeds as follows: For all executed events in PG&E, a logistic regression classifier is trained to distinguish these from SDG&E samples. The resulting class-conditional probabilities define importance weights:

$$w_i = \frac{P(\text{target} \mid x_i)}{P(\text{source} \mid x_i) + \epsilon}. \tag{6}$$

where $\epsilon = 10^{-8}$ prevents division by zero. Canceled samples retain unit weight, ensuring only the positive class distribution is adapted. This strategy emphasizes training samples that better reflect the operational conditions of SDG&E while maintaining PG&E-only training integrity.

The weighted SVM (IW-SVM) achieves a lower overall accuracy (80.6%) compared to the vanilla SVM (92.5%) but significantly improves recall on the underrepresented canceled class (from 0.50 to 0.84), indicating improved sensitivity to non-events. Conversely, the vanilla model demonstrates high precision and recall for executed events but performs poorly in identifying canceled cases.

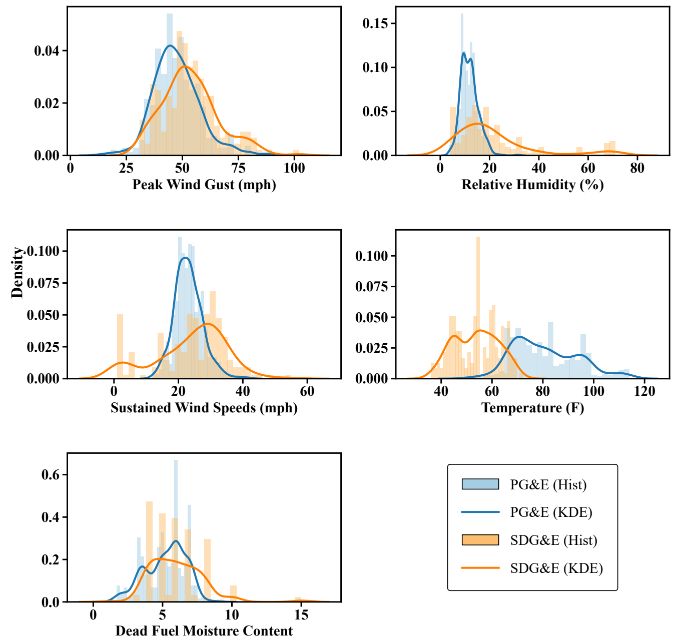

Fig. 3. Feature distribution comparison between PG&E and SDG&E post-event reports. Notable differences are observed in features such as temperature and relative humidity, where SDG&E values are heavily concentrated, suggesting lower variability and potential sensor thresholds or rounding.

Recommendation: Use the IW-SVM if the application prioritizes conservative de-energization (i.e., reducing false negatives for canceled events). Use the vanilla SVM when high precision or stability for executed events is preferred and the cost of false alarms is high. This tradeoff allows to select a model based on operational tolerance for risk versus reliability.

III. EXTENDED ABLATION STUDY

This section evaluates the robustness, generalizability, and interpretability of the proposed SVM framework through an extended ablation study. Together, these evaluations validate the design choices presented in earlier sections and highlight which variables and modeling strategies most effectively support PSPS decision-making.

A. Feature Space Ablation

Table IV presents the performance of the SVM classifier under different feature configurations. Accuracy is reported with 95% confidence intervals estimated via bootstrapping, while precision, recall, F1-score, and AUC are shown as point estimates. The results highlight the critical role of environmental predictors in PSPS decision modeling. Notably, removing fuel moisture features led to the most significant performance decline, underscoring their importance in capturing vegetation dryness and ignition potential—key factors for proactive shutoff decisions. Weather variables, while influential, had a slightly smaller impact, suggesting some redundancy among forecasted wind and humidity measures. The exclusion of fire behavior features resulted in minimal degradation, indicating that modeled fire spread characteristics may be less discriminative than raw environmental conditions near

TABLE IV
SVM ABLATION STUDY RESULTS

Configuration	Acc. [95% CI]	Prec.	Recall	F1	AUC
All Features	0.931 [0.888, 0.969]	0.940	0.986	0.962	0.950
Remove Weather	0.906 [0.856, 0.950]	0.915	0.986	0.949	0.928
Remove Fuel Moisture	0.887 [0.831, 0.931]	0.903	0.979	0.939	0.873
Remove Fire Behavior	0.938 [0.900, 0.969]	0.934	1.000	0.966	0.948
Remove Risk Features	0.931 [0.888, 0.969]	0.940	0.986	0.962	0.944
Forecast-Only	0.931 [0.888, 0.969]	0.940	0.986	0.962	0.944
Forecast + Risk	0.931 [0.888, 0.969]	0.940	0.986	0.962	0.950

the PSPS activation threshold. Furthermore, removing utility-defined risk scores (e.g., ignition probability, CFPD) had little effect on model performance, implying that forecast-based variables alone are sufficiently informative. The forecast-only model performed nearly identically to the full-feature model, and adding risk scores back yielded no further gains. These findings reinforce the value of using clean, forecast-derived indicators to support PSPS decisions and suggest that reliance on proprietary or post-processed risk indices may be unnecessary when environmental predictors are properly modeled.

B. Alternative Classifier Comparison

To benchmark performance, the SVM model was evaluated against alternative classifiers, including Logistic Regression, Random Forest, and XGBoost. As summarized in Table V, ensemble methods achieved slightly higher accuracy and F1-scores. However, the forecast-only SVM configuration (Table II) still reached strong performance with 0.925 accuracy and 0.98 recall. Unlike black-box ensemble models, SVM's margin-based formulation aligns with operational PSPS thresholds and supports calibrated confidence scoring. This interpretability makes SVM a practical and auditable choice for modeling utility decision behavior.

C. Generalization to Unseen Extreme Conditions

To evaluate the robustness of the SVM model under out-of-distribution scenarios, we perform an extreme condition test by masking and excluding high-risk cases—defined by the top quantile of wind speed and fire spread, and the bottom quantile of humidity and fuel moisture—from training. The SVM model is retrained exclusively on non-extreme data and evaluated on these withheld extremes.

The selected quantile-based thresholds, shown in Fig. 4, are not arbitrary; they reflect conditions commonly cited as operational triggers by utilities when deciding to initiate PSPS events. These include sustained winds, low relative humidity, and highly receptive fuels—all of which are recognized as wildfire risk indicators in utility PSPS guidelines [5], [19], [24]. When tested on the masked extreme cases, the SVM

TABLE V
COMPARISON OF CLASSIFIER METRICS ACROSS MODELS

Model	Accuracy [95% CI]	F1-Score	Recall
Logistic Regression	0.925 [0.881, 0.969]	0.958	0.972
Random Forest	0.963 [0.931, 0.988]	0.979	0.993
XGBoost	0.975 [0.950, 0.994]	0.986	0.993

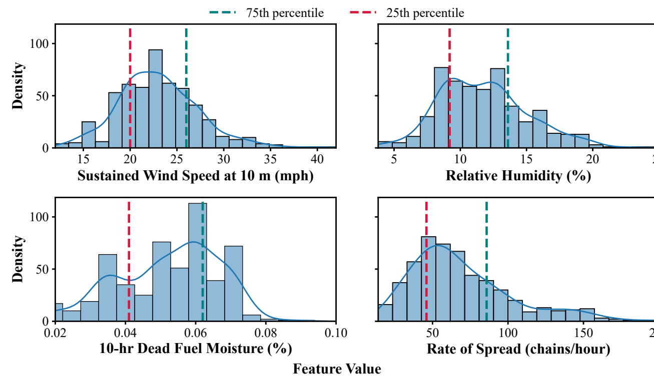

Fig. 4. Distributions of key wildfire predictors with 25% (black) and 75% (red) quantile thresholds. These thresholds define masked extreme scenarios used for generalization testing.

model achieves an overall accuracy of 74.7%, with a recall of 0.77 and precision of 0.23 for the minority class representing PSPS cancellations. This indicates that while the model can still detect many at-risk circuits under extreme conditions, it is prone to overestimating cancellations. Nonetheless, the high recall and weighted F1-score of 0.80 highlight its robustness in prioritizing fire-prone operational triggers even in scenarios not seen during training. These results underscore the relevance of weather and fuel features as core predictors in characterizing utility decision patterns.

IV. POST-HOC RISK SPACE ANALYSIS

In this section, a post-hoc diagnostic framework is constructed to evaluate how well the SVM-calibrated outputs align with wildfire risk conditions. This framework applies unsupervised clustering to define decision-relevant risk groupings and assess the PSPS model's behavior within them.

A. Formulation of the Post-Hoc Risk Space

The post-hoc feature space is constructed from wildfire consequence indicators available in utility post-event reports [24]. Specifically, the following metrics are included:

- Ignition Probability (IPW): The estimated likelihood of a wildfire ignition event, as computed by PG&E's operational model.
- Catastrophic Fire Probability Distribution (CFPD): A composite metric defined as the product of IPW and the *X16* (see Table I), capturing both the probability and potential severity of wildfire ignition.
- Risk and Benefit Indicators (MAVF): Binary MAVF scores reported by PG&E, based on a multi-attribute value function that evaluates PSPS justifications. MAVF combines attributes such as safety risk, reliability impact (e.g., customer minutes interrupted), and estimated financial costs, weighted by event probability to determine if a shutoff reduced wildfire harm (`Benefit = Yes/No`) or mitigated expected ignition risk (`Risk = Yes/No`).

KMeans clustering from the scikit-learn library [22] was used to segment the post-hoc risk-informed feature space into distinct decision clusters. It was selected due to its simplicity

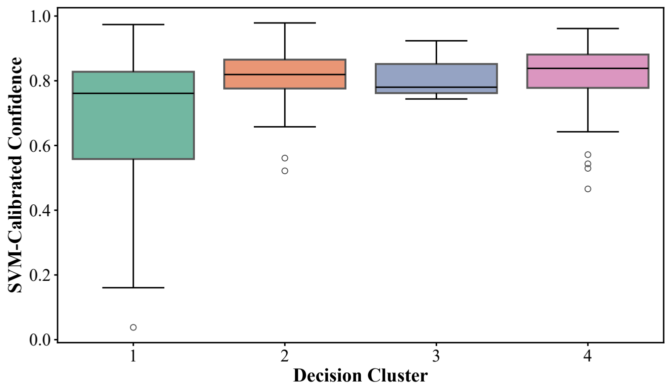

Fig. 5. SVM-calibrated confidence scores across post-hoc decision clusters. Cluster 0 shows wider variance and lower median, suggesting ambiguity. Clusters 1–3 exhibit higher, more consistent confidence aligned with high-risk designations.

and suitability for well-separated, moderately sized clusters. The steps include:

1) *Input Construction*: A matrix of samples $\times$ risk metrics (IPW, CFPD, MAVF indicators).
2) *Preprocessing*: All features are log-transformed where appropriate, followed by mean imputation and robust scaling. These post-hoc features are not included in model training but serve as independent validators of utility decisions and model behavior.
3) *Cluster Validation*: Silhouette scores and the elbow method are used to evaluate cluster compactness and stability across values of $k = 2$ to 7.
4) *Final Selection:* $k = 4$ selected to preserve interpretability across diverse decision behaviors.

These unsupervised clusters serve as independent validation zones for evaluating model outputs and real-world PSPS decision consistency under varied risk profiles. Once each event is assigned to a cluster, SVM margins and Platt-calibrated confidence scores—computed solely from forecast-based inputs—are analyzed across groups. This preserves a strict separation between training features and evaluation metrics. Figure 5 illustrates the distribution of confidence scores across clusters, and Table VI reports corresponding cluster-level averages.

B. Cluster-Level Interpretations and Recommendations

This analysis produced four meaningful decision clusters, each representing a distinct pattern of utility behavior under varying wildfire risk conditions. Note that these clusters are derived purely from risk-related attributes reported in post-event datasets and do not correspond to any geographic, spatial, or circuit-based grouping. Rather, they represent behavioral patterns across similar operational contexts.

Cluster 1 – Ambiguous Decisions: This group contains 152 events with low ignition probability (< 0.002), low SVM confidence ($\approx$ 0.68), and the lowest execution rate (19.7%). Utility PSPS decisions in this cluster appear precautionary and inconsistent, likely reflecting conservative activation under uncertain conditions, where the perceived risk may not be fully captured by forecasted inputs or model confidence.

- *Recommendation:* Flag for retrospective audit to refine PSPS thresholds under marginal risk conditions and reduce activation uncertainty.
- *Operational Suggestion:* Use sectionalizing devices or targeted automation in low-risk areas to minimize unnecessary customer disruptions.

This cluster underscores the importance of refining decision thresholds in low-risk conditions, where inconsistent PSPS actions suggest uncertainty or overly cautious activation logic.

TABLE VI
DECISION CLUSTER ID SUMMARY

ID	Avg Confidence	Exec. Rate	Cluster Size	Ignition Prob	CFPD
1	0.684	0.197	152	0.002	11.51
2	0.822	1.000	293	0.521	10.23
3	0.816	1.000	3	0.003	20.20
4	0.820	1.000	85	0.003	19.00

Cluster 2 – High-Risk Execution Benchmark: The largest cluster (293 events) exhibits the highest ignition probability (0.52), strong model confidence (0.82), and full execution. However, MAVF labels were not reported.

- *Recommendation:* Use this as a baseline for well-aligned PSPS decisions, and improve MAVF data reporting to better explain utility actions
- *Operational Suggestion:* Prioritize this cluster for microgrid deployment and CAVA (Climate Adaptation and Vulnerability Assessment)-informed hardening to support proactive shutdowns and recovery.

This cluster offers a reference for model-utility alignment in high-consequence scenarios and emphasizes the value of complete MAVF scoring to improve transparency in utility decision rationale.

Cluster 3 – High Risk Edge Cases: This cluster contains only 3 cases but exhibits the highest CFPD (> 20) and full MAVF flagging (`Risk = Yes`, `Benefit = Yes`).

- *Recommendation:* Leverage as reference cases for PSPS simulation and risk-informed planning.
- *Operational Suggestion:* Deploy advanced protection schemes and enhance vegetation management to mitigate ignition triggers in extreme-risk profiles.

This small but severe-risk cluster reinforces the value of edge-case preparedness and validates the need for proactive adaptation under extreme fire conditions.

Cluster 4 – Utility Driven Risk: This group includes events where the PSPS decision was executed with high model confidence, despite low ignition probability. Every case in this cluster was marked with `Risk = Yes`, indicating that utilities perceived a threat not captured by the current weather or fire-behavior inputs.

- *Recommendation:* Investigate alternative or unmeasured drivers—such as infrastructure vulnerability, limited access for emergency response, or operational constraints—that may have influenced these PSPS decisions beyond what the current model captures.

- *Model Insight:* Future model iterations should incorporate non-weather variables to better explain utility behavior in such cases,where decisions may hinge on asset vulnerabilities, operational policies, or situational constraints not captured by current inputs.

This cluster demonstrates how data-driven risk profiling can reveal blind spots in existing PSPS models and supports greater interpretability of utility actions under complex or unreported operational conditions.

V. Conclusions

This paper developed a data-driven framework for analyzing utility-initiated PSPS actions based on post-event wildfire reports. The primary objective is to characterize how utilities translate forecast-based wildfire risk into PSPS execution decisions. Using case studies conducted on data from PG&E and SDG&E, the proposed approach combines SVM-based classification with probabilistic calibration and domain adaptation to interpret decision-making patterns under class imbalance and utility-specific feature distributions. Platt scaling converts raw margins into interpretable confidence scores, and importance-weighted adaptation enables generalization to another power utility, SDG&E, which reports only executed events.

To move beyond binary outcomes, a post-hoc risk space is constructed using ignition probability, CFPD, and MAVF-based utility indicators. KMeans clustering segments this risk space into four decision clusters, each reflecting different PSPS activation behaviors. The results indicate that the model confidence aligns well with high-risk clusters, while low-risk clusters show greater ambiguity and inconsistent utility actions—providing useful zones for retrospective audit and operational threshold refinement.

The proposed framework offers a transferable and interpretable diagnostic tool for assessing PSPS actions using publicly available data. Further, it highlights the need for more structured utility reporting and sets the stage for future work incorporating geospatial, infrastructure, and temporal dynamics to inform advanced wildfire mitigation strategies.

References

[1] J. W. Muhs, M. Parvania, H. T. Nguyen, and J. A. Palmer, "Characterizing probability of wildfire ignition caused by power distribution lines," *IEEE Transactions on Power Delivery*, vol. 36, no. 6, pp. 3681–3688, 2021.

[2] S. Mohagheghi and S. Rebennack, "Optimal resilient power grid operation during the course of a progressing wildfire," *International Journal of Electrical Power & Energy Systems*, vol. 73, pp. 843–852, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061515002409

[3] J. E. Keeley and A. D. Syphard, "Historical patterns of wildfire ignition sources in california ecosystems," *International Journal of Wildland Fire*, vol. 27, pp. 781–799, Nov. 2018. [Online]. Available: https://www.publish.csiro.au/wf/WF18026

[4] J. T. Abatzoglou, C. M. Smith, D. L. Swain, T. Ptak, and C. A. Kolden, "Population exposure to pre-emptive de-energization aimed at averting wildfires in northern california," *Environmental Research Letters*, vol. 15, no. 9, p. 094046, aug 2020. [Online]. Available: https://dx.doi.org/10.1088/1748-9326/aba135

[5] "California Public Utilities Commission — cpuc.ca.gov," https://www.cpuc.ca.gov/, [Accessed 21-10-2024].

[6] Los Angeles County Economic Development Corporation, "The economic impact of the 2025 la wildfires," 2025, https://laedc.org/wpcms/wp-content/uploads/2025/02/LAEDC-2025-LA-Wildfires-Study.pdf.

[7] A. Arab, A. Khodaei, R. Eskandarpour, M. P. Thompson, and Y. Wei, "Three lines of defense for wildfire risk management in electric power grids: A review," *IEEE Access*, vol. 9, pp. 61 577–61 593, 2021.

[8] A. Kroepsch, S. St. George, and S. Mahnovski, "Public safety power shutoffs in california: Utility mitigation and societal adaptation," *Environmental Research Letters*, vol. 17, no. 9, p. 094035, 2022.

[9] Pacific Gas and Electric Company, "2023 general rate case," 2023, accessed: 2025-04-07. [Online]. Available: https://www.pge.com/en/regulation/general-rate-case.html

[10] A. D. Keety, T. J. Cova, and P. E. Dennison, "Modeling wildfire risk to infrastructure: Using machine learning to predict wildland fire ignition and spread near critical assets," *Fire*, vol. 4, no. 3, p. 42, 2021. [Online]. Available: https://www.mdpi.com/2571-6255/4/3/42

[11] A. Umunnakwe, M. Parvania, H. Nguyen, J. D. Horel, and K. R. Davis, "Data-driven spatio-temporal analysis of wildfire risk to power systems operation," *IET Generation, Transmission & Distribution*, vol. 16, no. 13, pp. 2531–2546, 2022. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/gtd2.12463

[12] M. Yao, M. Bharadwaj, Z. Zhang, B. Jin, and D. S. Callaway, "Predicting electricity infrastructure induced wildfire risk in california," *Environmental Research Letters*, vol. 17, no. 9, p. 094035, sep 2022. [Online]. Available: https://dx.doi.org/10.1088/1748-9326/ac8d18

[13] H. Mohsenian-Rad, A. Shahsavari, and M. Majidi, "Analysis of power quality events for wildfire monitoring: Lessons learned from a california wildfire," in *2023 IEEE PES Innovative Smart Grid Technologies Latin America (ISGT-LA)*, 2023, pp. 50–54.

[14] J. Garland, K. Baker, and B. Livneh, "Weather-induced power outage prediction: A comparison of machine learning models," in *2023 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2023, pp. 1–6.

[15] N. Rhodes, L. Ntaimo, and L. Roald, "Balancing wildfire risk and power outages through optimized power shut-offs," *IEEE Transactions on Power Systems*, vol. 36, no. 4, pp. 3118–3128, 2021.

[16] A. Lesage-Landry, F. Pellerin, J. A. Taylor, and D. S. Callaway, "Optimally scheduling public safety power shutoffs," 2022. [Online]. Available: https://arxiv.org/abs/2203.02861

[17] W. Hong, B. Wang, M. Yao, D. Callaway, L. Dale, and C. Huang, "Data driven power system optimal decision-making strategy under wildfire events," in *Proc. 55th Hawaii Int. Conf. Syst. Sci.*, 2022, pp. 1–10.

[18] C. Huang, Q. Hu, L. Sang, D. D. Lucas, R. Wong, B. Wang, W. Hong, M. Yao, and V. Donde, "A review of public safety power shutoffs (psps) for wildfire mitigation: Policies, practices, models and data sources," *IEEE Transactions on Energy Markets, Policy and Regulation*, vol. 1, no. 3, pp. 187–197, 2023.

[19] D. A. Z. Vazquez, F. Qiu, N. Fan, and K. Sharp, "Wildfire mitigation plans in power systems: A literature review," *IEEE Transactions on Power Systems*, vol. 37, no. 5, pp. 3540–3551, 2022.

[20] H. Mohsenian-Rad, *Smart Grid Sensors: Principles and Applications*. Cambridge University Press, 2022.

[21] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[22] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[23] J. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," *Adv. Large Margin Classif.*, vol. 10, 06 2000.

[24] Pacific Gas and Electric Company, "2023-2025 wildfire mitigation plan," https://www.pge.com/assets/pge/docs/outages-and-safety/outage-preparedness-and-support/120524-pge-2023-2025-wmp-r7-public.pdf, 2023, accessed: 2024-12-19.