

Earth's Future

RESEARCH ARTICLE

10.1029/2025EF006935

Key Points:

- Daily wildfire ignition prediction models were developed and validated over the western United States
- Robust sampling techniques to represent conditions associated with absence of wildfire ignitions were proposed
- Wildfire ignitions are governed by complex interactions among physical, biological, social, and administrative factors, not weather alone

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

M. Sadegh,
mojtabasadegh@gmail.com

Citation:

Pourmohamad, Y., Abatzoglou, J. T., Fleishman, E., Belval, E., Short, K. C., Williamson, M., et al. (2026). Predictive understanding of Wildfire ignitions across the western United States. *Earth's Future*, 14, e2025EF006935. <https://doi.org/10.1029/2025EF006935>

Received 10 JUL 2025

Accepted 5 DEC 2025

Author Contributions:

Conceptualization: Yavar Pourmohamad, John T. Abatzoglou, Erin Belval, Mojtaba Sadegh
Data curation: Yavar Pourmohamad
Formal analysis: Yavar Pourmohamad, Seyd Teymoor Seydi
Investigation: Yavar Pourmohamad, Seyd Teymoor Seydi, Mojtaba Sadegh
Methodology: Yavar Pourmohamad, John T. Abatzoglou, Mojtaba Sadegh
Project administration: Mojtaba Sadegh
Resources: Yavar Pourmohamad, John T. Abatzoglou, Mojtaba Sadegh
Software: Yavar Pourmohamad
Validation: Yavar Pourmohamad, Seyd Teymoor Seydi, Mojtaba Sadegh

© 2026 The Author(s).

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Predictive Understanding of Wildfire Ignitions Across the Western United States



Yavar Pourmohamad¹, John T. Abatzoglou² , Erica Fleishman³, Erin Belval⁴, Karen C. Short⁵ , Matthew Williamson⁶ , Michael Perlmutter⁷ , Seyd Teymoor Seydi¹ , and Mojtaba Sadegh^{1,8} 

¹Department of Civil Engineering, Boise State University, Boise, ID, USA, ²Management of Complex Systems Department, University of California, Merced, CA, USA, ³College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, USA, ⁴USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO, USA, ⁵USDA Forest Service, Rocky Mountain Research Station, Missoula, MT, USA, ⁶Human-Environment Systems, Boise State University, Boise, ID, USA, ⁷Department of Mathematics, Boise State University, Boise, ID, USA, ⁸United Nations University Institute for Water, Environment and Health, United Nations University, Richmond Hill, ON, Canada

Abstract Wildfires have increasingly affected human and natural systems across the western United States (WUS) in recent decades. Given that the majority of ignitions are human-caused and potentially preventable, improving the ability to predict fire occurrence is critical for effective wildfire prevention and risk mitigation. We used over 500,000 wildfire ignition records from 2000 to 2020 to develop machine learning models that predict daily ignition probability across the WUS and incorporate a wide range of physical, biological, social, and administrative variables. A key innovation of this work is development of novel sampling techniques for representing ignition absence. Unlike traditional purely random sampling or hyper-sampling, which does not account for temporally autocorrelated factors (such as droughts, insect outbreaks, and heatwaves) and spatially autocorrelated factors (such as proximity to human settlements, infrastructure presence, and fuel type), we introduce spatially and temporally stratified sampling of ignition absence. By drawing absence samples near the location and time of historical ignitions, we better captured the complex environmental and anthropogenic conditions associated with fire occurrence or lack thereof. Models trained without stratified sampling produced ignition probability maps that consistently overestimated fire risk during high fire danger periods, whereas models incorporating stratified fire absence samples more accurately captured the spatial and temporal variability of fire potential and achieved predictive accuracies exceeding 95%. In addition to operational utility for fire prevention and resource allocation, our approach offers insights into the drivers of wildfire ignitions and highlights the value of incorporating spatial and temporal structure in absence sampling for wildfire modeling.

Plain Language Summary Wildfire prevention is one of the most effective and economical risk mitigation strategies. Human-started wildfires account for over 60% of all recorded wildfires across the western United States and are responsible for the vast majority of wildfire-related societal impacts, underscoring the value of effective wildfire prevention strategies. To address this need, we developed machine learning models that not only effectively predict spatial and temporal patterns of wildfire ignitions but reveal the nuanced interactions among physical, biological, social, and administrative factors that govern wildfire ignition outcomes. Annual temperature (climate), discovery day-of-year (seasonal pattern), fire year (trend), and national preparedness level (management and fire danger) were the primary governing factors in models of all ignitions, natural ignitions, and human-caused ignitions. Secondary governing factors of natural and human-caused ignitions, respectively, were weather-related attributes and weather and social attributes. Our results indicated that although daily ignition probabilities generally track weather patterns, they can remain persistently high in areas where human factors dominate. Our results also show that models relying solely on weather do not accurately predict wildfire ignitions, reinforcing the fact that ignitions are caused by complex interactions among diverse factors.

1. Introduction

Wildfires, hereafter fires, are one class of climate-related extremes that have increasingly impacted several regions globally, including the western United States (WUS), with compound and cascading effects (Higuera et al., 2023; Modaresi Rad, Abatzoglou, Fleishman, et al., 2023; Modaresi Rad, Abatzoglou, Kreitler, et al., 2023). Fires impact a variety of systems and sectors, including the built environment (Seydi et al., 2024;

Visualization: Yavar Pourmohamad, Seyd
Teymoor Seydi

Writing – original draft:

Yavar Pourmohamad, John T. Abatzoglou,
Erica Fleishman, Erin Belval, Karen
C. Short, Seyd Teymoor Seydi,
Mojtaba Sadegh

Writing – review & editing: John
T. Abatzoglou, Erica Fleishman,
Matthew Williamson, Michael Perlmutter,
Mojtaba Sadegh

Wibbenmeyer & McDarris, 2021), public health (Buchholz et al., 2022), and land management and firefighting resources (Hosansky, 2023). Fires also affect the environment in numerous ways, such as altering vegetation structure and composition, increasing the risk of post-fire debris flows, modifying snowpack dynamics, and accelerating erosion in burned areas while contributing to sedimentation in downstream channels (Buchholz et al., 2022; Williams et al., 2022).

Increasing fire risks in the WUS have been attributed to intensification of fire weather (Abatzoglou & Williams, 2016; Zhuang et al., 2021), expansion of the wildland-urban interface (Radeloff et al., 2018), and increased fuel loads due to historical fire suppression (Boisramé et al., 2022). Strategies for mitigating increased fire risk include fuel treatments around values-at-risk and in critical ecological areas (Chung, 2015; Finney, 2001), home hardening to enhance the structure's ability to survive fire (Kodur et al., 2020), and fire prevention efforts (Edgeley et al., 2025). Fire prevention is among the least costly and most effective strategies in the WUS (Calkin et al., 2023) given that over 60% of all fire ignitions in the region are human-caused and potentially preventable. Because human-caused ignitions typically occur closer to human settlements and values-at-risk, they tend to be more destructive (Kumar, 2025) and more intense (Hantson et al., 2022).

Studies have focused on understanding the patterns and drivers of fire ignitions (Balch et al., 2017; Mann et al., 2016; Syphard et al., 2007). Weather conditions have been associated with fuel flammability and human outdoor activities, contributing to seasonal and interannual variability in the number and location of fire ignitions (Finney et al., 2011; Littell et al., 2016; Noonan-Wright et al., 2011). Widespread drying and increases in the number of critical fire weather days (Alizadeh et al., 2023; Khorshidi et al., 2020), in conjunction with human factors such as the expansion of the wildland-urban interface (Radeloff et al., 2018), increasing population size, and fire prevention efforts, have been linked with long-term trends in the number of ignitions (Nagy et al., 2018; Noonan-Wright et al., 2011). Although there is no widespread increase in the number of ignitions across the West (some studies show declines in the number of human-ignited fires; Jorge et al., 2025; Pourmohamad, Abatzoglou, et al., 2025; Pourmohamad, Sadegh, & Abatzoglou, 2025; Syphard et al., 2025), suggesting that various fire prevention efforts have been effective, the persistence of high-impact, human-ignited fires suggests that more fire prevention is needed. Essential to effective fire prevention is identifying not only where and when ignitions occur, but also what drives such variability.

In this paper, we develop machine learning models to predict daily fire ignition occurrence and investigate their drivers across the WUS. Such information is essential for understanding fire risk; prioritizing regional fuel management, fire suppression resources, and fire prevention efforts; and preparing communities for fire emergencies (Chen & Jin, 2022; Di Giuseppe et al., 2025). To train these models, we use extensive historical fire ignition records, and propose a novel method to develop robust samples that indicate absence of fire (Jiménez-Ruano et al., 2022). Our absence-of-fire samples not only are randomly distributed across space and time, but reflect conditions generally similar to those of ignition presence, with subtle yet important differences. We also evaluate the relative influence of various drivers on ignition occurrence. Gaining insight into the spatial and temporal patterns of fire ignitions is essential for informing fire and land management strategies aimed at reducing ignition risk (Chen et al., 2021; Faivre et al., 2014). Finally, we investigate the applicability of these models, which are trained on point samples, to develop gridded ignition probability maps for the WUS, and conduct a sensitivity analysis of these models to grid sizes.

2. Data and Methods

Our fire occurrence prediction framework (Figure 1) integrates a variety of factors identified in the literature as influences on fire ignition probability, including weather, fire danger, climate, land cover, topography, social factors, population density, and management. In the following sections, we first describe the samples representing fire presence and absence, and the attributes used to distinguish between them. We then describe our models and their evaluation, followed by a brief discussion of how our predictive maps of ignition probability are generated.

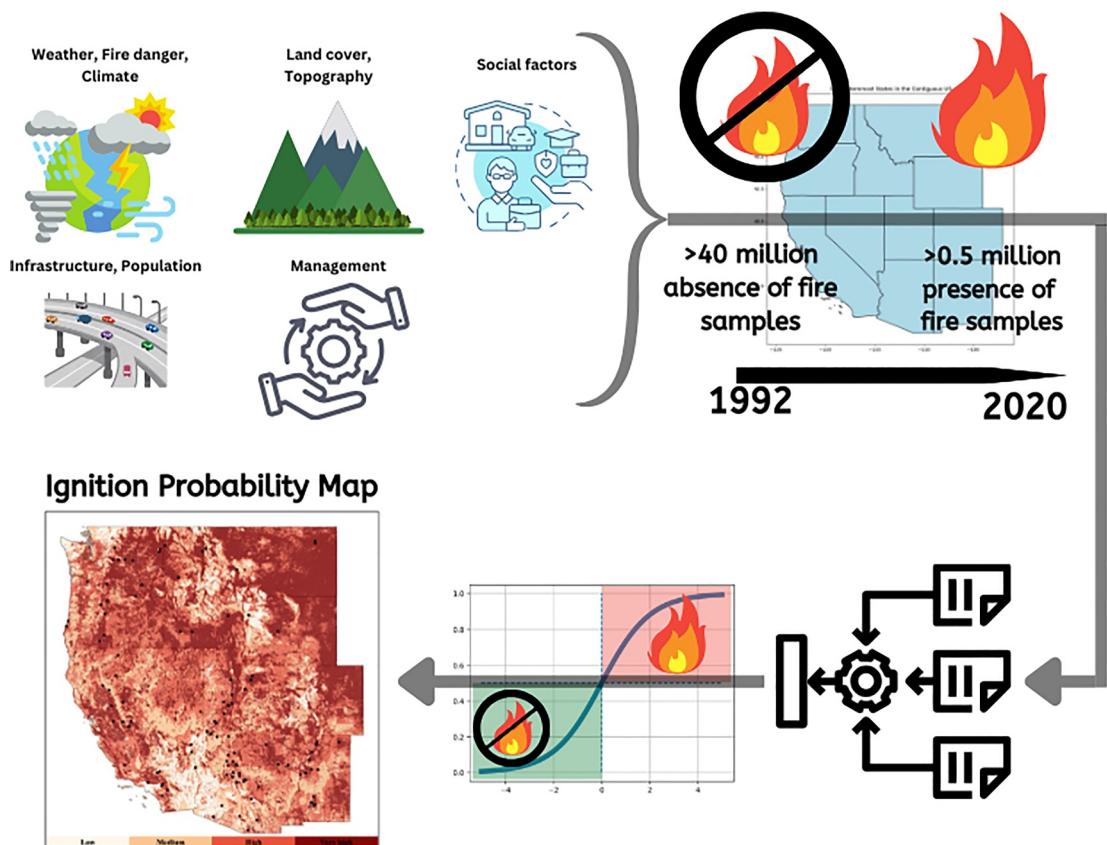


Figure 1. Fire ignition prediction framework.

2.1. Data

2.1.1. Data to Train, Validate, and Test Models

2.1.1.1. Presence of Ignition

We obtained fire records from the FPA FOD-Attributes data set (Pourmohamad et al., 2024), which augments the sixth version of the Fire Program Analysis-Fire Occurrence Database (FPA FOD v6; Short, 2014, 2022) with nearly 270 physical (e.g., weather, climate, topography, infrastructure), biological (e.g., land cover), social (e.g., population density, social vulnerability index (SVI)), and administrative (e.g., national preparedness level (NPL), jurisdiction) attributes that coincide with the date and location of each ignition. This data set contains information on the location, discovery time, cause, and final size of >2.3 million fires in the United States from 1992 to 2020. Of these fires, 752,461 occurred in the WUS (Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming), with final sizes from <0.1 acres to 1,068,802 acres. Here, we address fire occurrence without regard for fire size.

2.1.1.2. Absence of Ignition

We created three sets of samples that represent absence of ignition, and compiled physical, biological, social, and administrative attributes associated with each point-date.

■ *Hyper-sampling*: We created a 4 km grid mesh over the WUS (193,545 grid cells). Because fire occurrence peaks during summer and is minimal during winter in the WUS (Table S1 in Supporting Information S1) (Podschwit & Cullen, 2020), we randomly selected four dates in summer, one date in winter, and two dates each in spring and autumn (nine dates per year for each year from 2000 through 2020—timeline dictated by data availability, details later). For each date, we randomly selected one point in each grid cell to ensure a uniform spatial distribution of the samples. This resulted in 36,580,005 ($193,545 \times 9 \times 21$) samples of ignition absence (Figure 2).

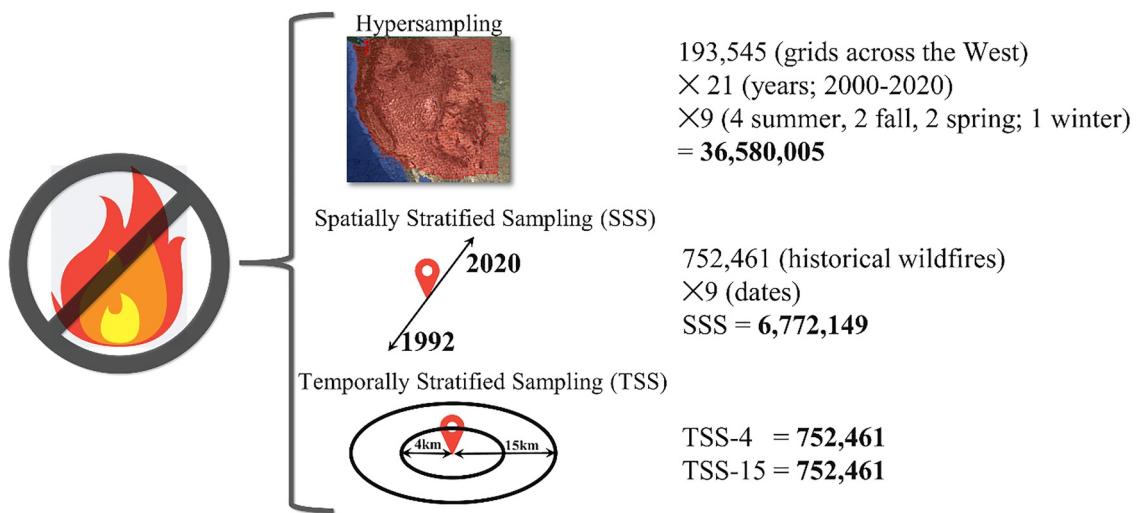


Figure 2. Sampling of point-dates that represent absence of fire ignition.

- *Spatially Stratified Sampling (SSS):* A majority of previous studies used either entirely random sampling (Zaidi, 2023) or a version of hyper-sampling (Moradi et al., 2024) to develop ignition absence samples. In contrast, we used spatially and temporally stratified sampling (TSS) to better capture the nuances in the attributes associated with the location and date of ignition presence and absence. By investigating these stratified samples, one can better understand the exact underlying conditions that lend themselves to a fire occurrence. We sampled 9 dates—4 in summer, 2 in fall, 2 in spring, 1 in winter—at the exact location of historical fire occurrences but on dates other than the ignition date. This resulted in 6,772,149 (752,461 × 9) SSS ignition absence samples (Figure 2).
- *Temporally Stratified Sampling (TSS):* We randomly sampled points that coincided with the date of an ignition and in close proximity to the ignition location, but not at the exact location. We randomly sampled points within a 4 km (TSS-4) radius of the ignition location, and then within a 15 km radius (TSS-15). This resulted in 752,461 ignition absence samples for each of the TSS-4 and TSS-15 sets (Figure 2).

Ignitions are more likely to occur in certain locations than others due to a variety of factors, including human activity and infrastructure, fuel type, and moisture regime. Similarly, specific periods—such as the Fourth of July for fireworks ignitions and growing season tails for ignitions caused by debris burning—are more prone to new ignitions. Hyper-sampling methods implicitly assume that ignition probability is uniformly distributed across space and to some extent time, although constraints on temporal distributions try to alleviate this shortcoming. In contrast, SSS and TSS techniques account for the spatial and temporal variability in ignition probability, thereby providing a more representative sample of ignition absence cases.

2.1.2. Data to Assess Predictive Models

We develop two sets of samples, beyond formal training, validation, and test data, to assess the predictive model. We first develop daily time series of attributes from 2000 to 2020 for a set of points to assess the dynamics of ignition probability over time. Next, we develop gridded maps of attributes at certain dates to assess the spatial distribution of ignition probabilities. The latter case also helps to evaluate whether a point model can develop predictive maps, and if so, to identify the grid resolution that can accurately represent attributes of the point samples used to train the models.

- *Time series data:* To assess how fire ignition probability changes over time, we selected the locations of 13 historical fires. We ensured that these fires were distributed over the entire WUS (Figure S1 in Supporting Information S1) and that each had a different ignition cause (arson or incendiary; debris and open burning; equipment and vehicle use; firearms and explosives use; fireworks; data missing, not specified, or undetermined; misuse of fire by a minor; natural; other causes; power generation, transmission, or distribution; railroad operations and maintenance; recreation and ceremony; and smoking). We extracted daily physical, biological, social, and administrative attributes at each location from 2000 to 2020.

- *Map data:* To assess whether point models can develop ignition probability maps, and if so, to determine the optimal grid size, we developed four sets of gridded data with different grid sizes (4 km, 1 km, 250 m, and 30 m). In each case, we acquired physical, biological, social, and administrative attributes associated with the centroids of the grids (Figure S2 in Supporting Information S1). This exercise aims to strike a balance in grid resolution: although finer resolutions more effectively capture local, bottom-up factors (e.g., land cover, population density), their information content is limited by the uncertainty in the locations of ignitions reported in the FPA FOD data set. We selected 5 September 2019, with 323 reported fires across the WUS, and 4 July 2020, with 250 reported fires across the WUS, to develop fire probability maps.

2.1.3. Attributes Used in the Predictive Model

Following Pourmohamad, Abatzoglou, et al. (2025), Pourmohamad, Sadegh, and Abatzoglou (2025), we selected 29 attributes with the highest predictive power for fire ignition modeling and the least collinearity. These attributes are enlisted in Table S2 in Supporting Information S1, and are summarized below:

- *General information:* Discovery day of year (DOY), fire year, and ignition cause. DOY captures intra-annual seasonality of ignitions, such as an increase in the number of fire ignitions around the Fourth of July, US Independence Day (Balch et al., 2017). Fire year acts as a proxy for long-term trends in fire ignitions, such as those due to fire prevention efforts. Each ignition cause is associated with distinct attributes (Pourmohamad, Abatzoglou, et al., 2025, Pourmohamad, Sadegh, & Abatzoglou, 2025). The National Wildfire Coordinating Group provides a list of 13 causes (National Wildfire Coordinating Group, 2025); here, we classify fires as all fires, naturally ignited fires, and human-ignited fires.
- *Weather and fire danger indices:* Daily precipitation, wind speed, minimum temperature, 100- and 1000-hr dead fuel moisture, burning index, vapor pressure deficit (VPD), energy release component (ERC), and ERC percentile. Fuels are more receptive to ignitions during dry-hot conditions, and wind enhances the likelihood of certain ignition causes, such as power-started fires (Balch et al., 2024). We note that there is a level of correlation between these attributes, but opted to include them in our model since each conveys nuanced information that can inform our predictive models. For example, ERC and VPD provide complementary information given that ERC evolves slowly whereas VPD evolves quickly. Daily precipitation enables differentiation between dry and wet lightning, which are associated with widely different probabilities of fire ignition (Kalashnikov, 2024).
- *Climate attributes:* Long-term average annual precipitation, temperature, and reference evapotranspiration. Climatic attributes largely shape fuel availability and type, and therefore the fire regimes in each region (Pausas & Paula, 2012).
- *Topography and land cover:* Average elevation, aspect, slope, and topographic position index within a 1 km radius of the ignition point, and existing vegetation cover and fire regime group (FRG) at the point of ignition. Topography is highly correlated with the type of ignition; for example, lightning-caused ignitions generally occur at higher elevations and slopes compared to human-caused ignitions (Narayananaraj & Wimberly, 2012). We selected average topographical indices within 1 km of the ignition point to address uncertainties in the reported ignition locations (Short, 2015). The 1-km radius also contributed to generalizing the point model to develop ignition probability maps (details later). Finally, existing vegetation cover helps capture different flammability levels among fuels and represents land-cover types such as roads (Fares et al., 2017).
- *Social factors:* Annual gross domestic product, global human modification (GHM), average human population density within a 1 km radius of the ignition point, and overall SVI. Various social attributes have been linked with human ignitions of fire (Flanagan et al., 2018). For example, moderate population densities were linked with a higher probability of ignition, whereas lower and higher population densities translate to absence of humans and limited vegetation cover, respectively, and hence were associated with lower ignition probabilities (Syphard et al., 2008).
- *Administrative factors:* Suppression difficulty index, FRG, management unit, protected area status (GAP Status Code), NPL, FIPS code—Land and fire management practices impact ignition loading (Keane, 2012). The number of human-caused ignitions is lower in protected areas (Nelson & Chomitz, 2011). NPL represents the national availability of fire suppression resources, with higher values indicating coincidence of multiple large fires across the country and strained suppression resources (Abatzoglou et al., 2021). FIPS code acts as a proxy for local factors such as demographics, suppression resources, and level of recreational activities, among

others, which are relevant to ignition modeling but are not readily available as independent data to be included in the model.

2.2. Modeling

We developed two ensemble machine learning models (Random Forest and XGBoost) and two deep learning models (one-dimensional convolutional neural networks, with convolution over attributes; CNN-1D with Keras and PyTorch; Figure S3 in Supporting Information S1) to estimate ignition probability in three ignition classes: all ignitions (regardless of reported cause), human-caused ignitions (12 classes of human-caused ignitions combined), and natural ignitions (almost entirely lightning-caused in the WUS). We used data from 2000 to 2018 to train, validate, and test our models, and used data from 2019 to 2020 as additional test cases (extra test). We did not use data before 2000 because certain data attributes (such as population density) were unavailable before 2000. We used dummy encoding to convert string-type attributes to digits. Importantly, we removed fire records that met three criteria: (a) reported by county or local agencies, (b) less than 0.1 acre in final size, and (c) associated with developed land cover with low, medium, or high development intensity. We deemed the uncertainty in reporting of these fires unacceptable due to temporal changes in governmental fire reporting procedures (Jorge et al., 2025). This process decreased the total number of ignition records between 1992 and 2020 from 752,461 to 574,522, of which 383,872 occurred between 2000 and 2020 and were used in our modeling effort. We divided the 2000–2018 machine learning-ready ignition records and ignition absence samples into training (~65%), validation (~15%), and test (~20%) data (Table S3 in Supporting Information S1). We used training data to learn model parameters, validation data to learn model hyperparameters, and test data to assess model accuracy. We considered ignition probabilities over 50% as an ignition incident. This threshold was selected through fine tuning of hyperparameters. We also performed a sensitivity analysis to examine how the model's performance changed with different cutoff thresholds used to distinguish between ignition presence and absence. We found a higher probability threshold to marginally reduce false positives at the expense of a large increase in false negatives (Figure S4 in Supporting Information S1).

We used a Bayesian optimizer and cross-validation to learn hyperparameters for each model. The training data had a significant class imbalance, with ignition absence incidents vastly outnumbering ignition presence incidents (98.9% vs. 1.1% of all samples, respectively). We developed separate models both with and without class weights to assess the potential impacts of class imbalance on the model performance. We then selected the best model on the basis of precision, recall, F1 score, and overall accuracy. We refrained from additional input data pre-processing, including temporal detrending, smoothing, or using a moving average, and spatial blocking. Specifically, temporal trends in the observed ignitions respond to interannual climatic fluctuations, the chronology of adoption of fire prevention strategies, and other factors that are directly or indirectly included in the modeling framework. Similarly, potential clustering of ignitions can be explained by underlying governing factors that are explicitly provided to the model.

We used the ignition absence samples to develop two categories of models. In the first category, *Ignition Model (IM)*, we included the samples representing absence of fire derived from hyper-sampling (Figure S5 in Supporting Information S1). In the second category, *Augmented Ignition Model (AIM)*, we included the absence of fire samples from hyper-sampling and from SSS and TSS-4 (Figure S6 in Supporting Information S1). In each category, we kept 20% of the fire absence samples for model testing. Furthermore, we withheld the entire group of TSS-15 data from the training process for additional testing. We used Shapley Additive explanations (SHAP) to calculate global feature importance, a measure of each feature's overall contribution to the model's predictive power. This metric is crucial for model interpretation and identifying the underlying relationships between features and the target variable (Nohara et al., 2022).

Interested audience are referred to Text S1, Table S4, and Figures S7 and S8 in Supporting Information S1 that describe machine learning versus weather-based ignition modeling.

3. Results and Discussion

3.1. Attributes Associated With Ignition Presence and Absence

Kernel density distributions of attributes associated with presence and absence of fire ignitions (Figure 3) have similarities and distinctions that can be used in machine learning models to predict an ignition or lack thereof.

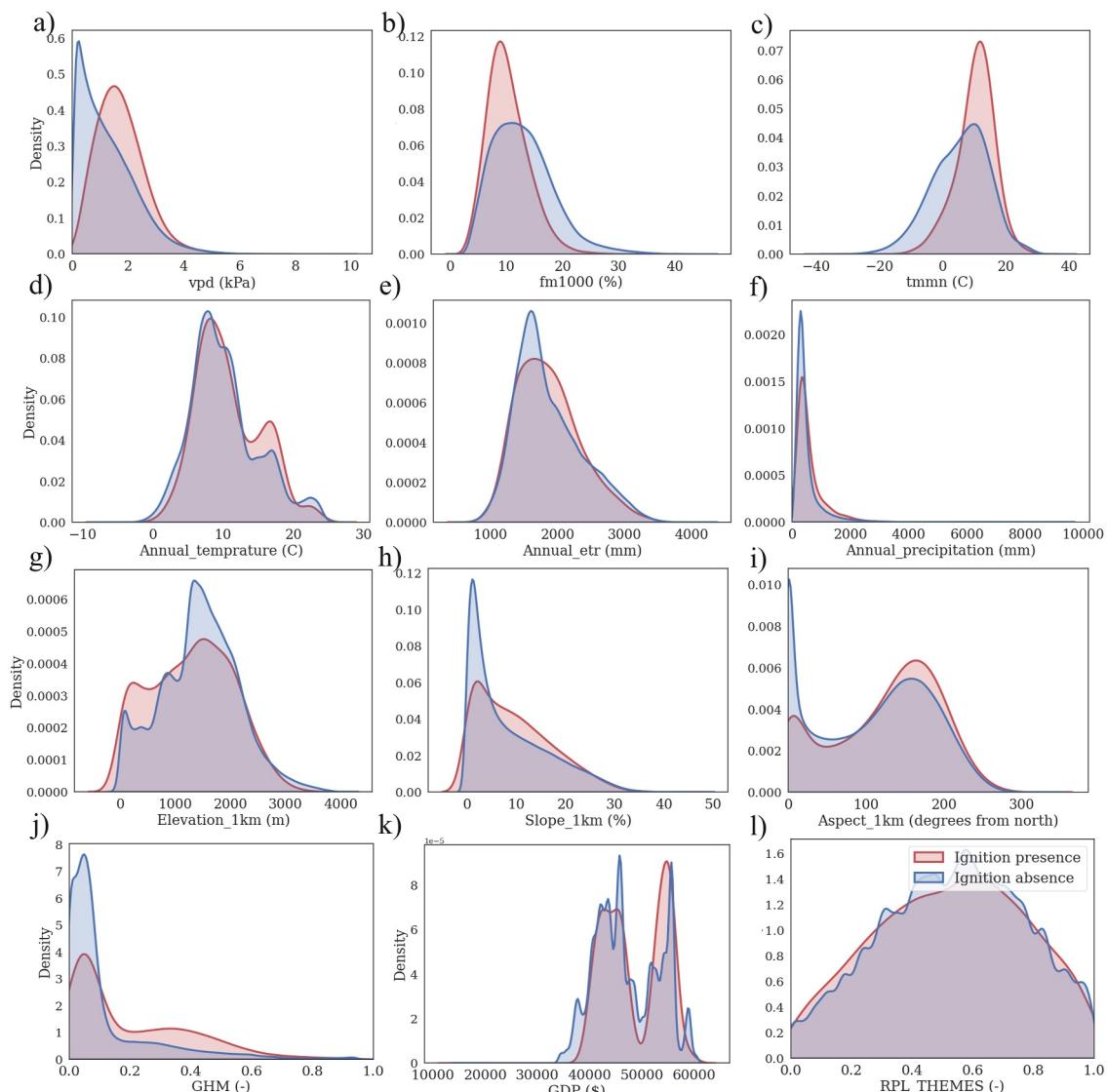


Figure 3. Kernel density distribution of attributes associated with ignition presence (all fire occurrences; red) and ignition absence (blue; combined hyper-sampled and stratified samples). The 12 attributes included in this figure are vpd: vapor pressure deficit (kPa), fm1000: 1,000-hr dead fuel moisture (%), tmmn: minimum daily temperature (Celsius), Annual_temperature (Celsius), Annual_etp: annual average reference evapotranspiration (mm), Annual_precipitation (mm), Elevation_1 km (m), Slope_1 km (%), Aspect_1 km (degrees from north), GHM: global human modification index (–), GDP: gross domestic product (US \$), RPL_THEMES: overall social vulnerability index (–).

Here, ignition absence includes both hyper-sampled and stratified (SSS, TSS-4) samples. Differences were more pronounced for weather-related attributes: ignition presence samples (all ignitions in Figure 3) were associated with drier and hotter conditions than ignition absence samples. For example, daily VPD (Figure 3a) and daily minimum temperature (Figure 3c) distributions shifted toward higher values, and 1,000-hr dead fuel moisture distribution shifted toward lower values (Figure 3b), for samples associated with presence of ignitions. This pattern is consistent with expectations given the higher likelihood of ignitions in drier and hotter conditions. Differences were less pronounced for other attributes, but still apparent for elevation (Figure 3g), aspect (Figure 3i) and global human modification index (GHM; Figure 3j), which are to some extent explained by the geography of fire occurrences (mainly human-caused ignitions, but also natural fires following mountain patterns; Pourmohamad et al., 2024) and our hyper-sampling method—generating a majority of ignition absence samples—drawing uniformly distributed samples for the absence of fire case. Figure S9 in Supporting Information S1 summarizes similarities and differences among the 27 numerical attributes associated with presence and absence of ignitions.

Nuanced differences emerged when separating natural from human-caused ignitions (Figures S10 and S11 in Supporting Information S1). Specifically, distributions of weather-related attributes associated with natural ignitions were more distinct from absence of fire samples (Figure S11 in Supporting Information S1) than those associated with human-caused ignitions (Figure S10 in Supporting Information S1). The latter is consistent with the concentration of natural ignitions in summer as compared to the expanded spatial and temporal presence of human-caused ignitions (Balch et al., 2017). These differences were also observed for topographical attributes with natural ignitions occurring at higher elevations and on steeper slopes than human-caused ignitions. The GHM distribution for natural ignitions, unlike that for human-caused ignitions, closely resembled the distribution of fire absence samples. Human-caused ignitions tend to occur in more developed areas and natural ignitions in less developed areas, and a majority of our fire absence samples were uniformly drawn from across the WUS, much of which is less developed. Finally, the distribution of climate attributes (annual temperature and reference evapotranspiration) for natural ignitions were more closely aligned with those of fire absence samples, as compared to human-caused ignitions (Figures S10 and S11 in Supporting Information S1), which we believe are due to differences in the spatial distributions of human-caused and natural ignitions.

3.2. Model Accuracy

We first used ignition presence samples (all causes) and the hyper-sampling-derived fire absence data to train two ensemble machine learning and two CNN-1D deep learning models. XGBoost provided the highest overall accuracy over the 20% out-of-sample test data among all models (Table S5 in Supporting Information S1). Our prior modeling efforts (Pourmohamad, Abatzoglou, et al., 2025, Pourmohamad, Sadegh, & Abatzoglou, 2025; Seydi et al., 2024) and those of others (Shwartz-Ziv & Armon, 2022) also showed superior performance of XGBoost for tabular data. We therefore employed XGBoost for the remainder of our modeling exercises in this paper. We also tested the XGBoost models with and without class weights, which resulted in very similar overall accuracy (with weight: 99.95%, without weight: 99.96%) with marginally superior accuracy for the case without weights; we hence used the XGBoost model without weights in the remainder of our analysis. We also note that within our conditional probability framework—where ignition probability is modeled based on 29 fire-related covariates and drivers—the decision boundaries between ignition presence and absence differ across modulating conditions. The XGBoost model can effectively capture these distinctions, so the imbalance in the data is not expected to fundamentally bias the predictions toward the more frequent label.

The IM framework demonstrated a robust performance on the test data, achieving high precision, recall, accuracy, and F1 score across all ignitions, natural ignitions and human-caused ignitions (84%–100%; Table 1). The model performance for the stratified fire absence data (SSS, TSS-4, and TSS-15), however, dropped to 66%–80% (Table 1). We attribute this performance drop to the divergent attributes of ignition absence samples used in training the model and those of the SSS and TSS data. The stratified ignition absence samples are spatially and/or temporally clustered around the ignition presences and hence are associated with attributes more similar to those of ignition presence. Performance of the AIM framework was marginally inferior to IM for the test data with precision, recall, accuracy, and F1 score ranging between 81% and 100%, but AIM performed markedly superior to IM by accurately capturing 96%–100% of fire absence samples for out-of-sample SSS, TSS-4, and TSS-15 data (Table 1). Refer to Table S6 for confusion matrices and to Figure S12 in Supporting Information S1 for Receiver Operating Characteristic curves and the Area Under the Curve for both IM and AIM frameworks and all modeling cases.

Model performance for both IM and AIM frameworks remained robust when applied to 2019 data, but performance metrics for both cases dropped markedly when applied to 2020 data (Table 1; extra test data). Societal disruptions and stay-at-home orders in 2020 associated with COVID-19 markedly changed human interactions with fire, causing a near record number of human-caused ignitions in 2020 (Jorge et al., 2025). Additionally, lower resource availability for fuel management and infrastructure maintenance likely contributed to the trends and patterns of fire ignitions in 2020. Nevertheless, our models were not trained with data that represent such a societal shock, and hence they were not able to accurately predict the 2020 ignitions.

3.3. Gridded Ignition Model

We used the trained models to develop predictive maps of ignition probability. Given that our model was trained on points, we executed the model for the grid centroids. In our analysis of sensitivity to grid sizes, we evaluated

Table 1

Performance Metrics for the Ignition Model (IM) and the Augmented Ignition Model Over Out-of-Sample and Extra Test Data

Data set	Metrics	Ignition model			Augmented ignition model		
		All fire causes (%)	Natural causes (%)	Human-caused (%)	All fire causes (%)	Natural causes (%)	Human-causes (%)
Test data	Precision	95.54	96.19	96.75	91.53	92.35	90.85
	Recall	82.96	84.87	78.30	80.71	82.78	74.63
	F1 score	88.81	90.18	86.55	85.78	87.30	81.94
	Accuracy	99.77	99.91	99.84	99.71	99.89	99.79
Accuracy for stratified data	SSS	76.04	76.06	79.87	96.18	99.95	99.75
	TSS-4	66.60	70.80	67.15	96.20	98.31	96.88
	TSS-15	67.71	70.90	68.09	96.18	98.26	96.94
2019	Precision	86.53	91.96	85.19	81.45	85.88	78.60
	Recall	83.34	92.56	76.24	73.88	79.97	64.94
	F1 score	84.91	92.26	80.47	77.48	82.82	71.12
	Accuracy	99.79	99.95	99.84	99.70	99.91	99.78
2020	Precision	69.70	64.50	72.69	68.68	56.47	71.24
	Recall	24.04	16.09	23.12	21.77	14.36	21.82
	F1 score	35.75	25.75	35.08	33.06	22.90	33.41
	Accuracy	99.22	99.80	99.40	99.22	99.79	99.41

Note. The green, light brown, light red show highest to lowest performance across scenarios, respectively.

the percentage of ignitions that were correctly classified by our point-based model, in essence calculating Recall for the model evaluated on grids. Model performance enhanced with increasing grid resolution (decreasing grid sizes), achieving 76%, 81%, 83%, and 86% accuracy for 4 km, 1 km, 250 m, and 30 m grid sizes (Figure S13 in Supporting Information S1). This outcome is expected, as finer resolutions place grid centroids closer to ignition points, resulting in their associated attributes being more similar—or even identical—to those of the ignition locations. Computational costs increase by a factor of 16–70 as grid sizes decrease. Considering trade-offs between accuracy and computational cost, we selected a 1 km grid to map the ignition probability over the WUS. We developed ignition probability maps for two dates in our extra test data, 5 September 2019 (Figure 4) and 4 July 2020 (Figure S14 in Supporting Information S1), both of which were associated with a large number of ignitions.

Although the IM framework almost perfectly identified ignitions in all classes (all ignitions, human-caused ignitions, and natural ignitions), it invariably assigned a high probability to the entire WUS on 5 September 2019 (Figure 4; left column) noting that small patches of low probability emerging in the human-caused ignitions model (Figure 4c). The IM yielded a low-information map due to the lack of spatial variability in the ignition probability distribution, rendering it not useful for resource allocation purposes. In contrast, the AIM framework generated a more informative ignition probability map (Figure 4; right column). The information content came at the expense of a marginal decline in correctly classification of observed ignition incidents but still captured 82% of human-caused ignitions and 93% of natural ignitions.

The AIM of natural ignitions exhibited a distinct spatial pattern, assigning elevated ignition probabilities to mountainous regions and lower probabilities to low elevations (Figure 4f), for example, in California's Central Valley, where irrigation of agricultural land minimizes the chances of lightning-started fires. Lower elevations are also generally associated with fewer ignitions than higher elevations, although our model does not include climatological lightning data and is therefore not affected by this factor. Our natural IM accurately captured the cluster of lightning storm-started fires on 5 September 2019 on the border of California and Oregon (Figure 4f), as well as other scattered ignitions across the WUS, but also assigned a high ignition probability to other mountainous regions that did not experience a fire. We attribute this behavior to the lack of representation of lightning strikes in our model—since our main focus was on modeling human-caused ignitions that potentially could be prevented. In other words, our model only captured the environmental conditions that are receptive to a lightning-ignited fire, not the presence of lightning strikes.

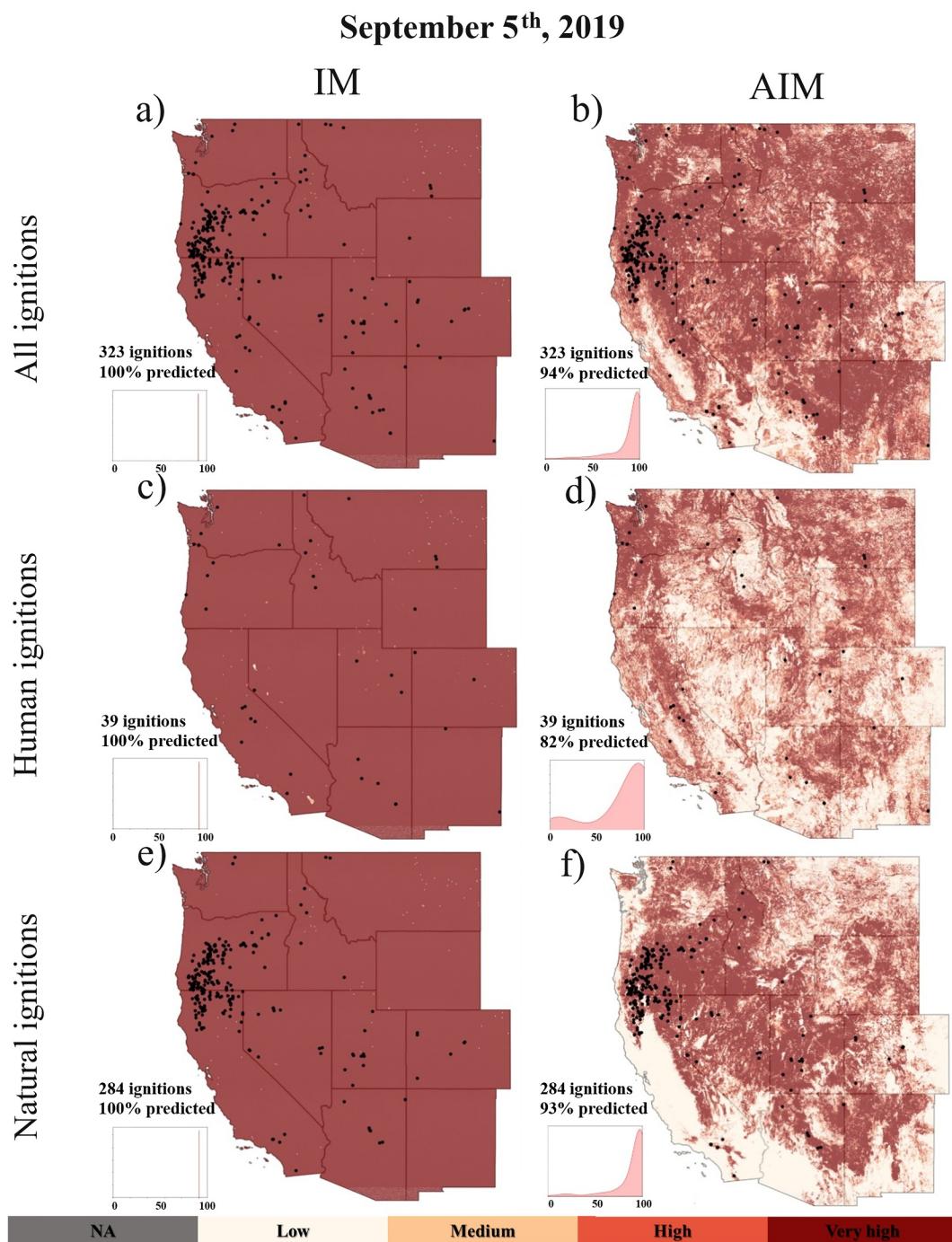


Figure 4. Probabilities of all ignitions, natural ignitions, and human-caused ignitions assigned by the Ignition Model (left column) and Augmented Ignition Model (right column) on 5 September 2019 at 1 km spatial resolution. Observed ignitions are displayed with black dots. Inset histograms show the probability that each model assigned to the grids encapsulated observed ignition points. An ignition probability of more than 50% is considered a presence of ignition. Ignition probability levels are shown in four categories: low (<25%), medium (25%–50%), high (50%–75%), and very high (>75%). Also see Table S7 in Supporting Information S1.

The human-caused ignitions also captured the geographical difference in high and low ignition likelihoods, assigning higher ignition probabilities to areas close to human settlements, roads, and infrastructure (Figure 4d). For example, foothills surrounding the Central Valley of California were assigned a high ignition probability, whereas the adjacent irrigated agricultural lands were assigned lower ignition likelihoods. High-elevation

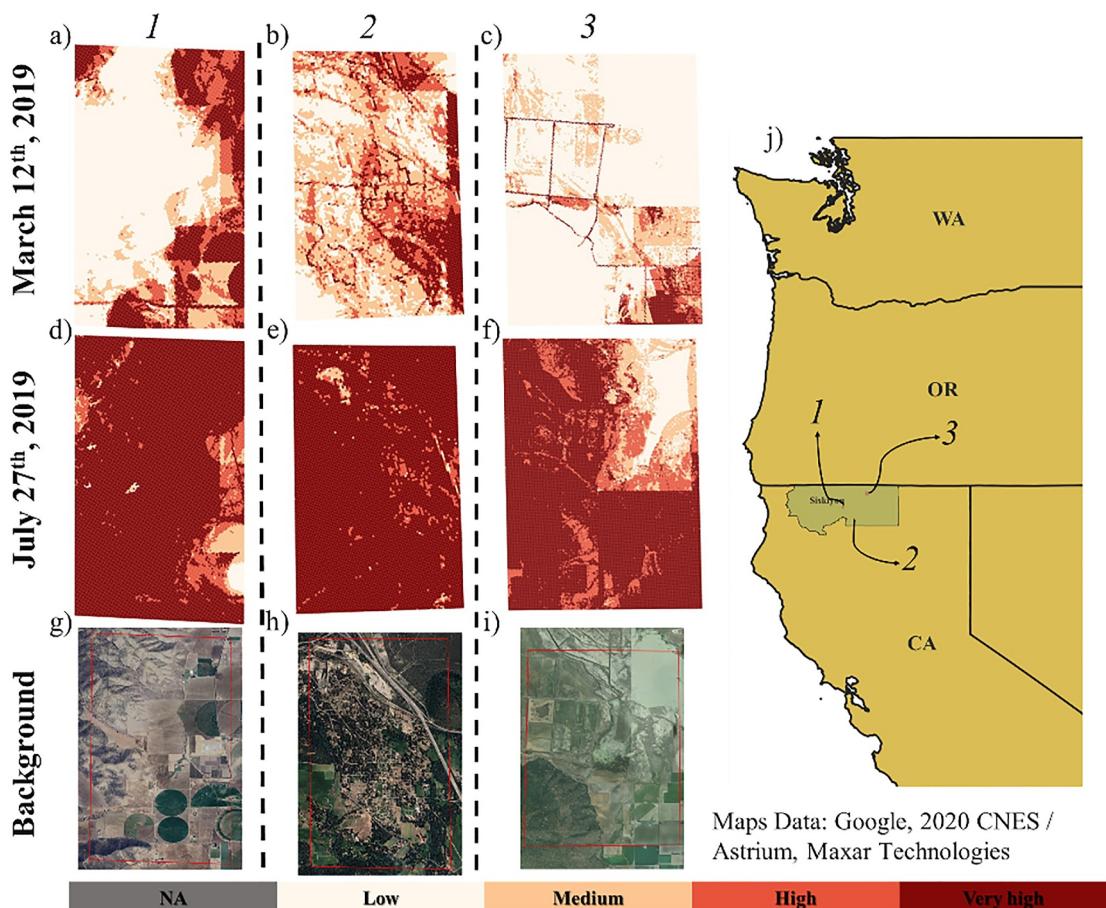


Figure 5. High-resolution (30 m) ignition probability maps generated by the Augmented Ignition Model for three locations in Siskiyou County, California. Panels (a)–(f) display ignition probabilities for 12 March 2019, and 27 July 2019. Ignition probability levels are shown in four categories: low (<25%), medium (25%–50%), high (50%–75%), very high (>75%). Panels (g)–(i) show satellite images of the case study locations.

mountains also received a low probability of human-caused ignition. The AIM model trained on all ignitions (Figure 4b) exhibited a more extensive spatial distribution of high ignition probabilities compared to the models trained on natural and human-caused ignition data separately. This is because the all-ignitions model integrates the predictive patterns of both natural and human-caused ignitions, capturing a more diverse range of factors that contribute to ignition risk. Probability maps for 4 July 2020 rendered similar general insights discussed here (Figure S14 in Supporting Information S1).

Given that the AIM framework was found to be more informative than the IM, we will use AIM in the remainder of this study. We applied the AIM to a more refined gridded map (30 m) in three case-study locations in Siskiyou County, California, on 12 March 2019 (during a period of relatively low ignition activity in the region) and 27 July 2019 (during a period of high ignition activity). The first location (Figure 5g) encompasses agricultural land and wildland vegetated areas. On 12 March, which represents wet conditions in the early spring, AIM assigned low ignition probabilities to the wildland areas, but assigned high ignition probabilities to agricultural lands (Figure 5a). We believe the latter is due to the probability of agricultural residue burns escaping and igniting a fire. Although our model does not include any information on soil moisture or vegetation health (e.g., Normalized Difference Vegetation Index (NDVI)), we believe the model captures agricultural residue burn probabilities through the interaction of seasonality metrics, such as DOY, combined with land cover. On 27 July, when conditions were hot and dry, our model assigned a high probability of ignition to the wildland areas, but a lower probability to the agricultural lands as by this time irrigation started and lowered the likelihood that vegetation would ignite, let alone the lower chance of residue burn—which we believe these conditions were captured through the seasonality metrics (Figure 5d).

The third case study location in northern California (Figure 5i) encompasses wildland vegetation, agricultural land, and wetlands. Similar to the first case study, AIM assigned high ignition probability to agricultural land and low ignition probability to wildlands on 12 March 2019. In this case, the model also predicted linear patterns of high ignition probability alongside roads. While our variables do not include distance to road, our model was able to capture higher ignition probabilities alongside roads using land cover data (Figure 5c). We emphasize that our model predicts an ignition of a fire of any final size (e.g., <0.1 acres), which is associated with loose thresholds of an ignition occurrence. On 27 July 2019, ignition probability across a majority of the case study location, except the wetland on the northeast side and in agricultural land, was high. Our model could have been further improved by incorporating vegetation productivity or greenness indices, such as the NDVI; however, we did not include these metrics among our variables due to incomplete coverage across the spatial and temporal extent of our study. The results of the second case study (Figures 5b, 5e, and 5h) were similar to those of the first and third cases.

3.4. Time Series of Ignition Probability

We used the AIM to develop daily time series of natural or human-caused ignition probabilities (Figure 6, Figure S15 in Supporting Information S1) from 2000 to 2020 for 13 locations across the WUS (Figure S1 in Supporting Information S1) that experienced a natural or human-ignited fire during that time period. The type of ignition we modeled corresponded to that of the reported fire. For a majority of the 13 time series, including that for the location that experienced a naturally ignited fire, the daily time series generally followed weather patterns, where dry and hot days were associated with a higher probability of ignition (Figure 6a). In some cases where the weather conditions were not hot and dry (low ERC), ignition probability was still high due to elevated wind speeds (Figure S15b in Supporting Information S1; fire ignited by debris burning). However, ignition probability dynamics widely diverged from weather conditions in the case of the location that observed an arson fire (Figure 6d, Figure S1 in Supporting Information S1). In that location, social factors dominated ignition probability, and the probability remained high except during some periods that were too wet and cold to be receptive to an ignition (Figure 6d).

3.5. Driving Factors of Predictions

Shapley value analysis revealed that the top four contributors to our model's classification outcomes were annual temperature, NPL, discovery DOY, and fire year, consistently for all ignitions, natural ignitions, and human ignitions (Figure 7). Annual temperature reflects a range of influences—including background climate conditions, prevailing fire regimes, and the geographic distribution of ignitions—all of which play important roles in shaping ignition patterns. Higher/lower values of annual temperature, however, were not consistently associated with positive/negative contribution to ignition incidence (Figure 7, right column), indicating that annual temperature works in conjunction with other factors to modulate fire occurrence. Higher values of NPL—which are associated with elevated fire activity across the country and strained firefighting resources—negatively contributed to the fire occurrence, lessening the probability of new ignitions. This is probably due to the management factors, such as burn bans, public land closures, and enhanced social awareness during high national preparedness levels. Discovery DOY mainly indicates certain celebrations such as 4 July, but also captures the seasonality of ignitions—for example natural ignitions occurring in summer months and agricultural residue burns occurring in the two tails of the growing season. Finally, the fire year represents the long-term trends in the number of ignitions, due to management, fire prevention, public awareness campaigns, and population dynamics, among others (Pourmohamad, Abatzoglou, et al., 2025, Pourmohamad, Sadegh, & Abatzoglou, 2025).

Divergence emerges in the importance of subsequent drivers, with VPD, FIPS Code, and 1000-hr dead fuel moisture (FM1000) being important for all ignition; FIPS Code, FM1000, and GHM index (GHM) being important for human ignitions; and daily minimum temperature, VPD and FM1000 being important for natural ignitions. Weather-related attributes are obviously important for both natural and human ignitions, but social factors represented through FIPS Code and GHM are more important for human ignitions. Dry-hot values of weather indices positively contribute to the fire occurrence, whereas the impacts of social factors are not as direct. High values of GHM generally lead to more fire ignitions, but also preventing ignitions in certain cases, and its lower values negatively contribute to fire occurrence. Increasing GHM is associated with greater human presence and a higher likelihood of ignition—up to a threshold where human density becomes high enough to limit fuel availability, thereby limiting the potential for fire ignition.

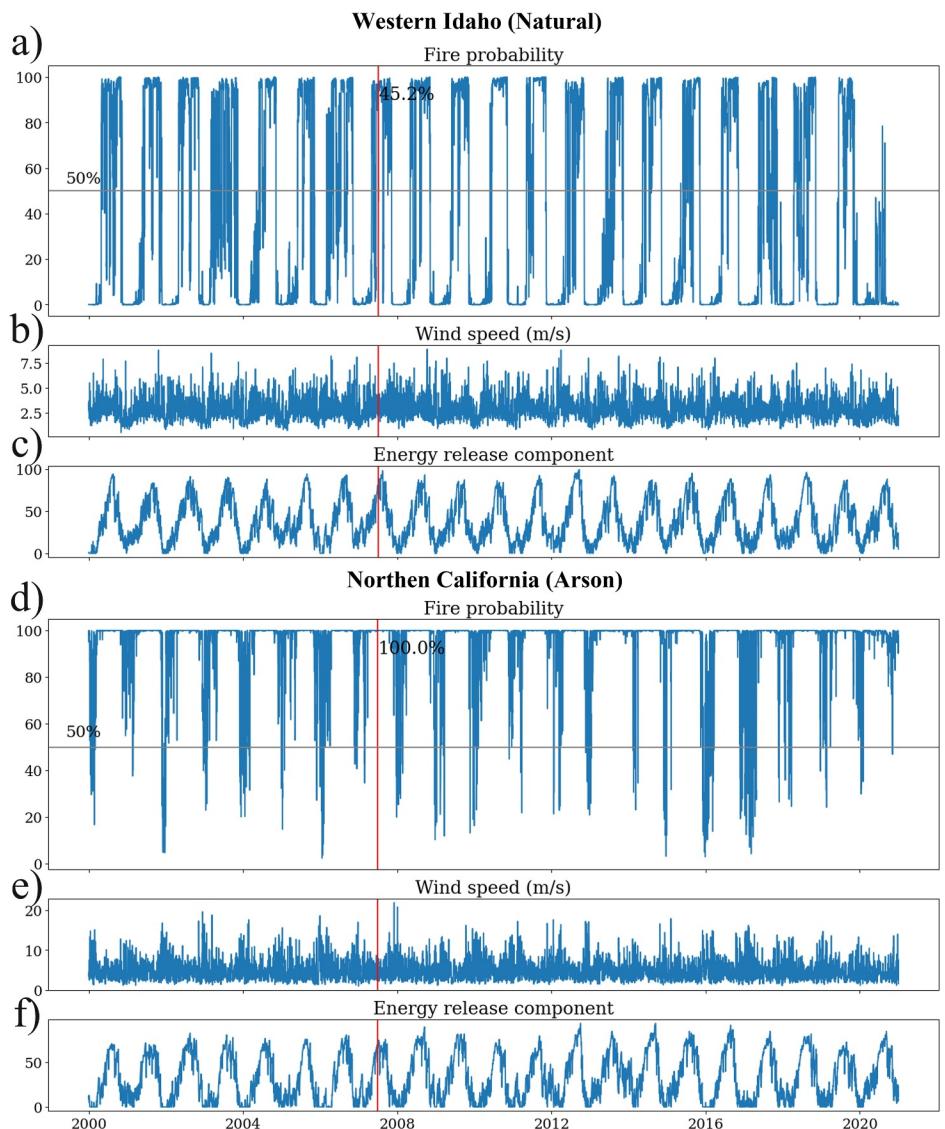


Figure 6. Daily dynamics of ignition probability, wind, and energy release component from 2000 through 2020 at two locations that experienced an ignition on the date marked by the vertical red line. The horizontal gray lines in (a) and (d) indicates a probability of 50%.

4. Conclusion

Accurate prediction of ignition likelihood is crucial for assessing fire risk, informing targeted fire prevention and fuel management strategies, and enhancing early response and community preparedness efforts (Chen & Jin, 2022). To address this need, we developed machine learning models trained on a comprehensive data set that integrated physical, biological, social, and administrative factors associated with ignition presence and absence samples. We used our models to develop ignition risk maps at daily temporal resolution and multiple spatial resolutions (4 km, 1 km, 250 m, and 30 m), providing a robust framework for ignition risk assessment and management.

A key contribution of our study is its use of spatially and TSS to represent fire absence conditions more accurately than conventional random or hyper-sampling approaches. Whereas random sampling implicitly assumes a uniform probability of ignition across space and time, our stratified approach acknowledges the inherent heterogeneity in ignition likelihood that is driven by factors such as human activity, infrastructure presence, fuel type, moisture regimes, and climate extremes such as prolonged droughts and heatwaves. By aligning the distribution

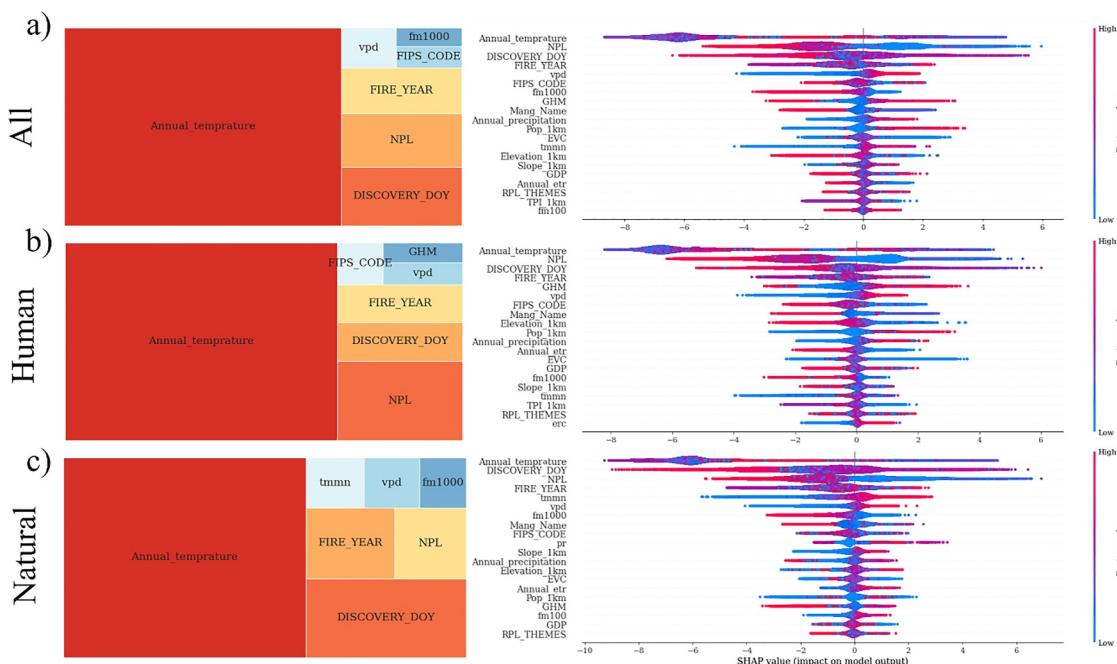


Figure 7. Driving factors of modeled probability of ignition (all ignitions, natural ignitions, and human-caused ignitions) in the western United States derived from Shapley Additive Explanations. Left column shows the global importance of each factor. Right column shows the contribution of each factor to each fire incident, with blue and red colors indicating low and high feature values. Positive SHAP values indicate positive contribution of the driver to ignition probability.

of absence samples with the spatial and temporal patterns of reported ignitions, our method provides a more realistic and informative foundation for modeling fire occurrence. We first used hyper-sampling to develop over 36 million fire absence samples, randomly selecting points in space and time to ensure uniform coverage across the WUS, while aligning the temporal distribution with seasonal ignition patterns (i.e., more samples in summer, fewer in winter). Trained on all historical ignition samples ($>500,000$, after filtering out uncertain records) and the hyper-sampled ignition absence data, our IM yielded near-perfect performance. Stratified ignition absence samples were then strategically drawn either at historical fire locations but on different dates (to capture spatial structure) or on the same dates as historical fires but at nearby locations (to capture temporal structure). When applied to these more realistic ignition absence samples, which were designed to closely resemble fire-prone conditions, the model's performance declined substantially. Further analysis revealed that the model tended to assign high ignition probabilities across much of the WUS on critical fire weather days, thereby limiting its effectiveness for operational use in resource allocation and fire prevention or response.

We then developed an AIM that incorporated not only hyper-sampled data but also spatially and temporally stratified samples representing ignition absence. Although this model slightly underperformed the original IM when applied to the test data, its performance substantially improved when applied to absence samples with background conditions that closely resembled those of ignition events. The AIM also produced more reliable ignition risk maps, effectively capturing spatial variability in fire occurrence likelihood across low- and high-risk areas. We used this model to generate time series of ignition probabilities for selected locations across the WUS. The model successfully tracked intra-annual cycles of dry-hot versus wet-cold weather to predict ignition likelihood. Additionally, in certain locations, ignition probabilities remained consistently high throughout the year except during brief cold and wet periods, indicating a dominant influence of consistent human factors. The model also effectively captured seasonal ignition patterns, such as those associated with agricultural debris burning, and was sensitive to the effect of infrastructure proximity, especially roads, on ignition risk.

Future research could build on our development of point-based models and their application to develop gridded maps of fire ignition likelihood by incorporating computer vision models to better capture spatial patterns in ignition drivers. Our models assume that temporal nonstationary in ignitions can be represented through proxy variables such as year. Although our models make accurate short-term predictions, they will need periodic retraining as new data become available to account for evolving trends. We also used the FIPS code as a proxy for

local factors such as resources, demographics, and community characteristics. Extending this framework to explicitly include these variables could further improve model performance. Lastly, projecting future ignition patterns in response to population dynamics, WUI expansion, and ongoing climate warming remains an active and important area for future research.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

All the data used and models developed in this study are publicly available from Pourmohamad, Abatzoglou, et al. (2025), Pourmohamad, Sadegh, and Abatzoglou (2025).

Acknowledgments

This research was supported by the Joint Fire Science Program (Grants L21AC10247 and L24AC00277) and the National Science Foundation (awards # 2429021 and 2521103).

References

Abatzoglou, J. T., Juang, C. S., Williams, A. P., Kolden, C. A., & Westerling, A. L. (2021). Increasing synchronous fire danger in forests of the western United States. *Geophysical Research Letters*, 48(2), e2020GL091377. <https://doi.org/10.1029/2020GL091377>

Abatzoglou, J. T., & Williams, A. P. (2016). Impact of anthropogenic climate change on wildfire across western US forests. *Proceedings of the National Academy of Sciences*, 113(42), 11770–11775. <https://doi.org/10.1073/pnas.1607171113>

Alizadeh, M. R., Abatzoglou, J. T., Adamowski, J., Modaresi Rad, A., AghaKouchak, A., Pausata, F. S. R., & Sadegh, M. (2023). Elevation-dependent intensification of fire danger in the western United States. *Nature Communications*, 14(1), 1773. <https://doi.org/10.1038/s41467-023-37311-4>

Balch, J. K., Bradley, B. A., Abatzoglou, J. T., Nagy, R. C., Fusco, E. J., & Mahood, A. L. (2017). Human-started wildfires expand the fire niche across the United States. *Proceedings of the National Academy of Sciences*, 114(11), 2946–2951. <https://doi.org/10.1073/pnas.1617394114>

Balch, J. K., Iglesias, V., Mahood, A. L., Cook, M. C., Amaral, C., DeCastro, A., et al. (2024). The fastest-growing and most destructive fires in the US (2001 to 2020). *Science*, 386(6720), 425–431. <https://doi.org/10.1126/science.adk5737>

Boisramé, G. F. S., Brown, T. J., & Bachelet, D. M. (2022). Trends in western USA fire fuels using historical data and modeling. *Fire Ecology*, 18(1), 8. <https://doi.org/10.1186/s42408-022-00129-4>

Buchholz, R. R., Park, M., Worden, H. M., Tang, W., Edwards, D. P., Gaubert, B., et al. (2022). New seasonal pattern of pollution emerges from changing North American wildfires. *Nature Communications*, 13(1), 2043. <https://doi.org/10.1038/s41467-022-29623-8>

Calkin, D. E., Barrett, K., Cohen, J. D., Finney, M. A., Pyne, S. J., & Quarles, S. L. (2023). Wildland-urban fire disasters aren't actually a wildfire problem. *Proceedings of the National Academy of Sciences*, 120(51), e2315797120. <https://doi.org/10.1073/pnas.2315797120>

Chen, B., & Jin, Y. (2022). Spatial patterns and drivers for wildfire ignitions in California. *Environmental Research Letters*, 17(5), 055004. <https://doi.org/10.1088/1748-9326/ac60da>

Chen, B., Jin, Y., Scaduto, E., Moritz, M. A., Goulden, M. L., & Randerson, J. T. (2021). Climate, fuel, and land use shaped the spatial pattern of wildfire in California's Sierra Nevada. *Journal of Geophysical Research: Biogeosciences*, 126(2), e2020JG005786. <https://doi.org/10.1029/2020JG005786>

Chung, W. (2015). Optimizing fuel treatments to reduce wildland fire risk. *Current Forestry Reports*, 1, 44–51. <https://doi.org/10.1007/s40725-015-0005-9>

Di Giuseppe, F., McNorton, J., Lombardi, A., & Wetterhall, F. (2025). Global data-driven prediction of fire activity. *Nature Communications*, 16(1), 2918. <https://doi.org/10.1038/s41467-025-58097-7>

Edgeley, C. M., Evans, A. M., Devenport, S. E., Kohler, G., Zamudio, Z. M., & DeGrandpre, W. D. (2025). Preventing human-caused wildfire ignitions on public lands: A review of best practices. *Forest Science*, 71(4), 1–29. <https://doi.org/10.1007/s44391-025-00025-9>

Faivre, N., Jin, Y., Goulden, M. L., & Randerson, J. T. (2014). Controls on the spatial pattern of wildfire ignitions in Southern California. *International Journal of Wildland Fire*, 23(6), 799–811. <https://doi.org/10.1071/wf13136>

Fares, S., Bajocco, S., Salvati, L., Camarretta, N., Dupuy, J.-L., Xanthopoulos, G., et al. (2017). Characterizing potential wildland fire fuel in live vegetation in the Mediterranean region. *Annals of Forest Science*, 74, 1–14. <https://doi.org/10.1007/s13595-016-0599-5>

Finney, M. A. (2001). Design of regular landscape scale fuel treatment patterns for modifying fire growth and behavior. *Forest Science*, 47(2), 219–228. <https://doi.org/10.1093/forestscience/47.2.219>

Finney, M. A., McHugh, C. W., Grenfell, I. C., Riley, K. L., & Short, K. C. (2011). A simulation of probabilistic wildfire risk components for the continental United States. *Stochastic Environmental Research and Risk Assessment*, 25(7), 973–1000. <https://doi.org/10.1007/s00477-011-0462-z>

Flanagan, B. E., Hallisey, E. J., Adams, E., & Lavery, A. (2018). Measuring community vulnerability to natural and anthropogenic hazards: The Centers for Disease Control and Prevention's social vulnerability index. *Journal of Environmental Health*, 80(10), 34–36.

Hantson, S., Andela, N., Goulden, M. L., & Randerson, J. T. (2022). Human-ignited fires result in more extreme fire behavior and ecosystem impacts. *Nature Communications*, 13(1), 2717. <https://doi.org/10.1038/s41467-022-30030-2>

Higuera, P. E., Cook, M. C., Balch, J. K., Stavros, E. N., Mahood, A. L., & St. Denis, L. A. (2023). Shifting social-ecological fire regimes explain increasing structure loss from Western wildfires. In *PNAS Nexus*, pgad005.

Hosansky, D. (2023). Simultaneous large wildfires will increase in Western U.S. Retrieved from <https://news.ucar.edu/132914/simultaneous-large-wildfires-will-increase-western-us>

Jiménez-Ruano, A., Jolly, W. M., Freeborn, P. H., Vega-Nieva, D. J., Monjáras-Vega, N. A., Briones-Herrera, C. I., & Rodrigues, M. (2022). Spatial predictions of human and natural-caused wildfire likelihood across Montana (USA). *Forests*, 13(8), 1200. <https://doi.org/10.3390/f13081200>

Jorge, A. L., Abatzoglou, J. T., Fleishman, E., Williams, E. L., Rupp, D. E., Jenkins, J. S., et al. (2025). COVID-19 fueled an elevated number of human-caused ignitions in the western United States during the 2020 wildfire season. *Earth's Future*, 13(4), 1–13. <https://doi.org/10.1029/2024EF005744>

Kalashnikov, D. A. (2024). Dry lightning in the Western United States: Meteorological conditions, wildfire ignition, air quality impacts, and future projections. In *ProQuest dissertations and theses*. Washington State University PP- United States. Retrieved from <https://libproxy.boisestate.edu/login?url=https://www.proquest.com/dissertations-theses/dry-lightning-western-united-states/docview/3071634457/se-2?accountid=9649>

Keane, R. E. (2012). Describing wildland surface fuel loading for fire management: A review of approaches, methods and systems. *International Journal of Wildland Fire*, 22(1), 51–62. <https://doi.org/10.1071/wf11139>

Khorshidi, M. S., Dennison, P. E., Nikoo, M. R., AghaKouchak, A., Luce, C. H., & Sadegh, M. (2020). Increasing concurrence of wildfire drivers tripled megafire critical danger days in Southern California between 1982 and 2018. *Environmental Research Letters*, 15(10), 104002. <https://doi.org/10.1088/1748-9326/abae9e>

Kodur, V., Kumar, P., & Rafi, M. M. (2020). Fire hazard in buildings: Review, assessment and strategies for improving fire safety. *PSU Research Review*, 4(1), 1–23. <https://doi.org/10.1108/prr-12-2018-0033>

Kumar, D. (2025). Analysing trends of multi-sensor data assimilations for wildfire-urban interface risk assessment and modeling: An observation networks and remote sensing Approach. In *105th Annual AMS Meeting 2025* (Vol. 105, p. 456990).

Littell, J. S., Peterson, D. L., Riley, K. L., Liu, Y., & Luce, C. H. (2016). A review of the relationships between drought and forest fire in the United States. *Global Change Biology*, 22(7), 2353–2369. <https://doi.org/10.1111/gcb.13275>

Mann, M. L., Battlori, E., Moritz, M. A., Waller, E. K., Berck, P., Flint, A. L., et al. (2016). Incorporating anthropogenic influences into fire probability models: Effects of human activity and climate change on fire activity in California. *PLoS One*, 11(4), e0153589. <https://doi.org/10.1371/journal.pone.0153589>

Modaresi Rad, A., Abatzoglou, J. T., Fleishman, E., Mockrin, M. H., Radeloff, V. C., Pourmohamad, Y., et al. (2023). Social vulnerability of the people exposed to wildfires in US West Coast states. *Science Advances*, 9(38), eadhf4615. <https://doi.org/10.1126/sciadv.adhf4615>

Modaresi Rad, A., Abatzoglou, J. T., Kreitler, J., Alizadeh, M. R., AghaKouchak, A., Hudyma, N., et al. (2023). Human and infrastructure exposure to large wildfires in the United States. *Nature Sustainability*, 6(11), 1–9. <https://doi.org/10.1038/s41893-023-01163-z>

Moradi, S., Hafezi, M., & Sheikhi, A. (2024). Early wildfire detection using different machine learning algorithms. *Remote Sensing Applications: Society and Environment*, 36, 101346. <https://doi.org/10.1016/j.rsase.2024.101346>

Nagy, R. C., Fusco, E., Bradley, B., Abatzoglou, J. T., & Balch, J. (2018). Human-related ignitions increase the number of large wildfires across US ecoregions. *Fire*, 1(1), 4. <https://doi.org/10.3390/fire1010004>

Narayananraj, G., & Wimberly, M. C. (2012). Influences of forest roads on the spatial patterns of human-and lightning-caused wildfire ignitions. *Applied Geography*, 32(2), 878–888. <https://doi.org/10.1016/j.apgeog.2011.09.004>

National Wildfire Coordinating Group. (2025). Proposed wildfire cause data standard. Retrieved from <https://www.nwcg.gov/page/proposed-wildfire-cause-data-standard>

Nelson, A., & Chomitz, K. M. (2011). Effectiveness of strict vs. multiple use protected areas in reducing tropical forest fires: A global analysis using matching methods. *PLoS One*, 6(8), e22722. <https://doi.org/10.1371/journal.pone.0022722>

Nohara, Y., Matsumoto, K., Soejima, H., & Nakashima, N. (2022). Explanation of machine learning models using shapley additive explanation and application for real data in hospital. *Computer Methods and Programs in Biomedicine*, 214, 106584. <https://doi.org/10.1016/j.cmpb.2021.106584>

Noonan-Wright, E. K., Opperman, T. S., Finney, M. A., Zimmerman, G. T., Seli, R. C., Elenz, L. M., et al. (2011). Developing the US wildland fire decision support system. *Journal of Combustion*, 2011(1), 168473. <https://doi.org/10.1155/2011/168473>

Pausas, J. G., & Paula, S. (2012). Fuel shapes the fire-climate relationship: Evidence from Mediterranean ecosystems. *Global Ecology and Biogeography*, 21(11), 1074–1082. <https://doi.org/10.1111/j.1466-8238.2012.00769.x>

Podschwit, H., & Cullen, A. (2020). Patterns and trends in simultaneous wildfire activity in the United States from 1984 to 2015. *International Journal of Wildland Fire*, 29(12), 1057–1071. <https://doi.org/10.1071/wf19150>

Pourmohamad, Y., Abatzoglou, J. T., Belval, E. J., Fleishman, E., Short, K. C., Reeves, M. C., et al. (2024). Physical, social, and biological attributes for improved understanding and prediction of wildfires: FPA FOD-Attributes dataset. *Earth System Science Data*, 16(6), 3045–3060. <https://doi.org/10.5194/essd-16-3045-2024>

Pourmohamad, Y., Abatzoglou, J. T., Fleishman, E., Short, K. C., Shuman, J., AghaKouchak, A., et al. (2025). Inference of wildfire causes from their physical, biological, social and management attributes. *Earth's Future*, 13(1), e2024EF005187. <https://doi.org/10.1029/2024EF005187>

Pourmohamad, Y., Sadegh, M., & Abatzoglou, J. T. (2025). Predictive understanding of wildfire ignitions across the western United States. *Zenodo*. <https://doi.org/10.5281/zenodo.15652532>

Radeloff, V. C., Helmers, D. P., Kramer, H. A., Mockrin, M. H., Alexandre, P. M., Bar-Massada, A., et al. (2018). Rapid growth of the US wildland-urban interface raises wildfire risk. *Proceedings of the National Academy of Sciences*, 115(13), 3314–3319. <https://doi.org/10.1073/pnas.1718850115>

Seydi, S. T., Abatzoglou, J. T., AghaKouchak, A., Pourmohamad, Y., Mishra, A., & Sadegh, M. (2024). Predictive understanding of links between vegetation and soil burn severities using physics-informed machine learning. *Earth's Future*, 12(8), e2024EF004873. <https://doi.org/10.1029/2024ef004873>

Short, K. C. (2014). A spatial database of wildfires in the United States, 1992–2011. *Earth System Science Data*, 6(1), 1–27. <https://doi.org/10.5194/essd-6-1-2014>

Short, K. C. (2015). Sources and implications of bias and uncertainty in a century of US wildfire activity data. *International Journal of Wildland Fire*, 24(7), 883–891. <https://doi.org/10.1071/wf14190>

Short, K. C. (2022). Spatial wildfire occurrence data for the United States, 1992–2020 [FPA_FOD_20221014].

Shwartz-Ziv, R., & Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion*, 81, 84–90. <https://doi.org/10.1016/j.inffus.2021.11.011>

Syphard, A. D., Keeley, J. E., Conlisk, E., & Gough, M. (2025). Regional patterns in US wildfire activity: The critical role of ignition sources. *Environmental Research Letters*, 20(5), 54046. <https://doi.org/10.1088/1748-9326/adc9c8>

Syphard, A. D., Radeloff, V. C., Keeley, J. E., Hawbaker, T. J., Clayton, M. K., Stewart, S. I., & Hammer, R. B. (2007). Human influence on California fire regimes. *Ecological Applications*, 17(5), 1388–1402. <https://doi.org/10.1890/06-1128.1>

Syphard, A. D., Radeloff, V. C., Keuler, N. S., Taylor, R. S., Hawbaker, T. J., Stewart, S. I., & Clayton, M. K. (2008). Predicting spatial patterns of fire on a southern California landscape. *International Journal of Wildland Fire*, 17(5), 602–613. <https://doi.org/10.1071/wf07087>

Wibbenmeyer, M., & McDarris, A. (2021). Wildfires in the United States 101: Context and consequences. *Resources for the Future*.

Williams, A. P., Livneh, B., McKinnon, K. A., Hansen, W. D., Mankin, J. S., Cook, B. I., et al. (2022). Growing impact of wildfire on western US water supply. *Proceedings of the National Academy of Sciences*, 119(10), e2114069119. <https://doi.org/10.1073/pnas.2114069119>

Zaidi, A. (2023). Predicting wildfires in Algerian forests using machine learning models. *Heliyon*, 9(7), e18064. <https://doi.org/10.1016/j.heliyon.2023.e18064>

Zhuang, Y., Fu, R., Santer, B. D., Dickinson, R. E., & Hall, A. (2021). Quantifying contributions of natural variability and anthropogenic forcings on increased fire weather risk over the western United States. *Proceedings of the National Academy of Sciences*, 118(45), e2111875118. <https://doi.org/10.1073/pnas.2111875118>

References From the Supporting Information

Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, 33(1), 121–131. <https://doi.org/10.1002/joc.3413>

Deeming, J. E., Burgan, R. E., & Cohen, J. D. (1977). *The national fire-danger rating system, 1978* (Vol. 39). Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station.

Kennedy, C. M., Oakleaf, J. R., Theobald, D. M., Baruch-Mordo, S., & Kiesecker, J. (2019). Managing the middle: A shift in conservation priorities based on the global human modification gradient. *Global Change Biology*, 25(3), 811–826. <https://doi.org/10.1111/gcb.14549>

Kummu, M., Taka, M., & Guillaume, J. H. A. (2018). Gridded global datasets for gross domestic product and Human Development Index over 1990–2015. *Scientific Data*, 5(1), 1–15. <https://doi.org/10.5061/dryad.dk1j0>

LANDFIRE. (2022a). LANDFIRE Existing Vegetation Type layer. <https://doi.org/10.5066/P974JF8W>

LANDFIRE. (2022b). LANDFIRE Topography type layer.

National Interagency Fire Center (NIFC). (2023). Wildland fire perimeters full history. Retrieved from <https://data-nifc.opendata.arcgis.com/data-sets/nifc::wfigs-wildland-fire-perimeters-full-history/explore>

Silva, F. R., O'Connor, C. D., Thompson, M. P., Martínez, J. R. M., & Calkin, D. E. (2020). Corrigendum to: Modelling suppression difficulty: Current and future applications. *International Journal of Wildland Fire*, 29(8), 752. https://doi.org/10.1071/wf19042_co

U.S. Geological Survey (USGS) Gap Analysis Project (GAP). (2018). *US geological survey–gap analysis project species range maps CONUS_2001*. US Geological Survey Data Release. <https://doi.org/10.5066/F7Q81B3R>

WorldPop. (2018). Global high resolution population denominators project. In *Funded by The Bill Melinda Gates Foundation (OPP1134076 School of Geography and Environmental Science University of Southampton; Department of Geographic Geoscience University of Louisville; Departement de Géographie Université de Namur Center for International Earth Science Information Network (CIESIN), Columbia University*.