**Wildfire Predictive Modelling in the Wildland–Urban Interface with Random Forest Analysis: Evaluating the Role of Social and Environmental Factors in the Santa Monica Mountains**

London School of Economics and Political Science
MSc Geographic Data Science
Ava Blanchette

**ABSTRACT**

Wildfires in California's wildland–urban interface (WUI) are growing in frequency and intensity, raising urgent questions about whether communities can meaningfully influence risk through preparedness measures. While most wildfire models focus on environmental drivers such as vegetation, weather, and drought indices, this study incorporates social and infrastructural factors to test their predictive relevance. Using data from 2015–2025 for the western Santa Monica Mountains and adjacent WUI neighbourhoods, a 300-metre spatial grid was developed and analysed with a Random Forest classification model. Model behaviour was interpreted using both impurity-based feature importance scores and SHapley Additive exPlanations (SHAP) to capture the relative weight and directional influence of each variable.

Environmental factors, including shrub cover, wind speed, and precipitation, emerged as the strongest predictors of wildfire occurrence. However, preparedness proxies also registered measurable influence: the proportion of the Los Angeles city budget allocated to the fire department ranked fourth in importance, and spatial access to fire infrastructure contributed modestly to model outputs. The results suggest that while preparedness cannot override climatic and ecological constraints, targeted investments in emergency capacity can alter fire risk in statistically significant ways. By integrating environmental realities with human dimensions of readiness, predictive modelling can inform not just forecasts of fire, but strategies to shape its impact.

## CONTENTS

## 1. Introduction

The growing number of destructive wildfires, particularly in wildland-urban interface (WUI) zones, has raised urgent questions about how we understand fire risk and preparedness. Most wildfire models prioritise environmental and climatic inputs such as weather patterns, vegetation, landscape characteristics (Hultquist et al., 2014; Collins et al., 2018; Ghimite et al., 2012). However, fewer models account for social and infrastructural factors like fire department funding, building density, or proximity to water sources. This gap is significant. As fires increasingly threaten developed and populated areas, the question of how statistically significant these parameters are in influencing a fire's behaviour demands exploration.

The objective of this study is to examine whether fire preparedness measures – represented by proxies such as local fire department budgets and access to emergency infrastructure – have a statistically significant influence on wildfire behaviour. The central hypothesis is that these preparedness proxies will exhibit measurable and notable statistical significance, but their predictive influence will remain weaker than that of natural processes such as wind, temperature, or fuel conditions. This question arises not only from academic curiosity, but also from public sentiment. In a historically destructive firestorm event, the January 2025 fires in Los Angeles collectively burned over 37,000 acres and 16,000 structures (CAL FIRE, 2025). The public response was one of despair and frustration; it was a perceived failure in public preparedness. Many pointed to budget cuts and a thinning fire response force, suggesting that the destruction was not solely the culmination of climatic and environmental conditions, but also a lack of investment in response infrastructure (*Los Angeles Times*, 2025). Yet, to date, the relationship between preparedness funding and wildfire behaviour has not been rigorously analysed through data-oriented modelling.

This study seeks to address that gap by applying a Random Forest classification model with SHapley Additive exPlanations (SHAP) value interpretation to assess the relative importance of social and infrastructural parameters in predicting wildfire hazard alongside their environmental counterparts. The focus area includes the western Santa Monica Mountains and its surrounding WUI neighbourhoods, as its landscape complexity, climatic variability, structure density, and history of destructive fires renders it a uniquely suitable region for this analysis (Dye et al., 2020; McGranahan et al., 2021). By incorporating both natural and human systems into the modelling framework from the years 5 to 2025, this study seeks to evaluate not only where fire is likely to occur, but why. Further, it seeks to reveal whether communities can reduce the risk of destruction through targeted investments in emergency preparedness.

The implications are potentially transformative. If anthropogenic variables are found to significantly influence fire outcomes, the findings could inform policymakers to make more strategic, data-informed decisions about where and when to allocate resources. For example, in years projected to bring dry, high-risk fire conditions, municipalities might avoid budget cuts to fire departments, or invest in water infrastructure expansion. As wildfires grow in intensity and frequency and encroach on built environments in the WUI, it becomes paramount not only to model the physical conditions under which fires burn, but also to understand anthropogenic choices that can prevent them from becoming catastrophic.

## 2. Literature Review

Wildfires have become an increasing global concern as their frequency, intensity, and seasonal duration escalate beyond historic norms (Schug et al., 2023). This trend is driven by the complex interplay between environmental, climatic, and anthropogenic factors that collectively amplify fire propagation and intensity. Climate change has intensified drought conditions and increased the occurrence of strong wind events, both of which are known to elevate fire spread potential (McGranahan et al., 2021). Globally, fire weather seasons have lengthened across 25.3% of the Earth's vegetated surface, resulting in an average 18.7% increase in the duration of fire-prone conditions (Jolly et al., 2015). This prolonged exposure to extreme fire weather increases the risk of catastrophic wildfires across a wider range of regions and for extended periods each year.

In the United States, fire suppression policies implemented since the mid-20th century have inadvertently intensified wildfire hazards; by preventing small-scale, low-intensity fires, these policies have led to excessive fuel accumulation in fire-adapted ecosystems, thereby creating conditions conducive to large-scale, high-intensity fires (Schug et al., 2023; David, 1995; McGranahan et al., 2021). Coupled with the ramifications of climate change and its sub-conditions of high drought intensity and irregular precipitation patterns, fire conditions have grown even more severe. California, in particular, has borne the brunt of this trend. It leads the nation on both the number of wildfire events and the total acreage burned in modern history (National Interagency Fire Center, 2025). The frequency of major fire events and their associated damage, including loss of life, acreage consumed, and structures destroyed, continues to rise. While fire remains a natural ecological cycle, the current fire regime is increasingly out of balance.

Wildfire behaviour is primarily driven by the interaction of three factors: topography, weather, and fuel. Each of these components operates across multiple spatial and temporal scales, contributing to the complexity of fire ignition, combustion, and spread (McGranahan et al., 2021). As McGranahan et al. (2021) explain, each broad category encompasses multiple subcomponents. For example, "weather" captures wind speed, precipitation patterns, and wind speed – all of which fluctuate daily and seasonally. "Fuel" refers not only to the presence of combustible material, but its characteristics, including moisture content and ratio of fine-to-coarse fuel. Given that these variables are highly dynamic and context-specific, modelling fire behaviour is inherently complex and must rely on current, high-resolution data to accurately assess fire risk and spread potential.

***Table 1: Fire indicators and their associated spatial and temporal dynamism. Source: McGranahan et al., 2021***

| Component | Spatially Dynamic | Temporally Dynamic | Manageability |
|---|---|---|---|
| Topography | Yes | No | None |
| Weather | Not very | Yes | Low |
| Fuels | Potentially | Yes | High |

Historically, many ecosystems in the Western United States have evolved with fire as a regular and necessary disturbance. They developed under steady fire regimes, marked by more consistent frequency, intensity, and seasonality; however, decades of fire suppression policies, combined with climate-driven changes in precipitation patterns and temperature extremes have disrupted natural cycles (Stephens et al., 2013; McGranahan et al., 2021). Research has revealed that these altered fire regimes are not only more intense, but less predictable. Westerling (2016) notes that trends in earlier snow melt, frequent periods of extreme heat, and extended dry seasons are all exacerbated by climate change. As these factors collectively lengthen the fire season and the likelihood of highly destructive fires, many landscapes that once experienced low-to-moderate-intensity fires are now increasingly subject to high-severity events that exceed historical norms.

Fire behaviour becomes significantly more complex when wildland area meets urban development in the wildland-urban interface (WUI). The WUI is defined as an area where wildland vegetation and buildings intermingle (Radeloff et al., 2005). This convergence introduces multiple complexities, as both the expansion of homes within the WUI and climate change are key factors controlling how many homes burn in a wildfire (Radeloff et al., 2023). Fires occurring in the WUI pose a heightened risk of property destruction and loss of human life as buildings become fuel sources themselves; they become part of the fuel path. Due to their mix in materials from plastic to fabric to foam, buildings enter the fuel path in a way that differs from vegetation loads; as their combustible materials like plastic and foam burn and release toxins, additional public health and environmental hazards are introduced (Averett, 2024; Gaudet et al., 2020). Despite these heightened stakes and the increasing frequency of such fire events, the broader implications of human-environment interactions that arise when fires devastate these areas remain under-studied.

Although the WUI accounts for just 4.7% of Earth's land surface, it is home to nearly half of the global population, at an estimated 3.5 billion people (Schug et al., 2023). Projections suggest that the WUI will continue to expand worldwide, driven by population growth and the appeal of living near natural amenities. In California, these trends mirror broader demographic patterns, as pressures for affordable housing have contributed to urban sprawl into fire-prone watersheds and landscapes with heavy fuel loads (Keeley and Syphard, 2021). Between 2010 and 2020 alone, California added 244,000 homes within the WUI (Radeloff et al., 2023).

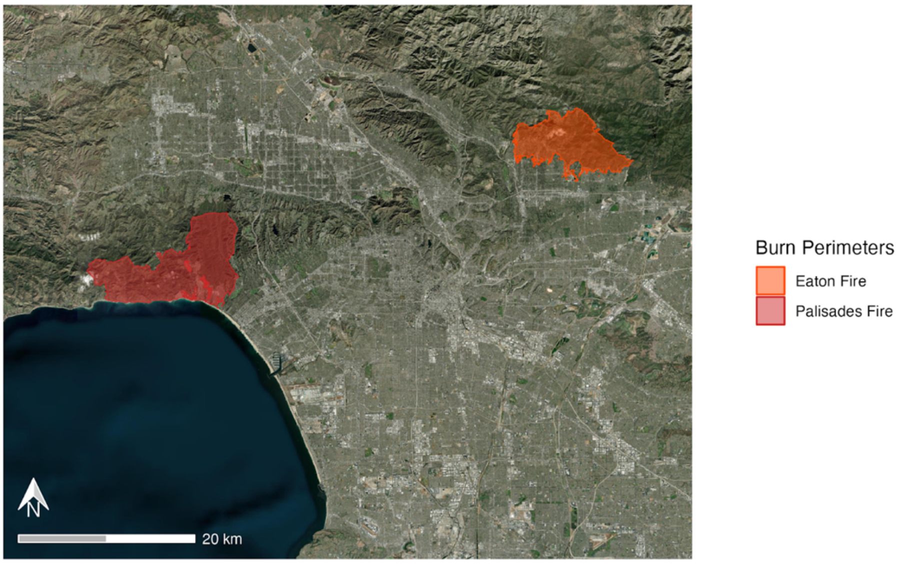
The dual growth of the WUI and of wildfire occurrence has sharply increased the risk of harm to human life and the destruction of infrastructure. In California, the number of houses located within wildfire perimeters has doubled since the 1990s, a trend driven by housing growth and the expansion of burned areas (Radeloff et al., 2023). As Radeloff et al. (2023) conclude, the number of homes within the WUI and climate change are the two most influential factors controlling how many homes are lost to fire. The continued expansion of the WUI worldwide, including in California, underscores that the rise of wildfire frequency and intensity is not solely an environmental issue; it is increasingly intertwined with human livelihood, housing stability, public health, and local economic resilience.

The Santa Monica Mountains and surrounding developed areas in Southern California fall within the Mediterranean ecosystem, an ecosystem classification characterised by wet winters, hot, dry summers, and distinct shrublands that are either evergreen or drought-deciduous (Muñoz-Gálvez et al., 2025). In this region, chaparral is the dominant evergreen shrubland (Barro and Conrad, 1991). These chaparral species are drought-tolerant, fire-adapted, and capable of recovering from infrequent crown fires. However, its low water content and needle-like leaves, compounded by species-specific high concentrations of flammable compounds like waxes and oils, render them highly flammable and particularly vulnerable to recurring, severe wildfires (Barro and Conard, 1991; Schug et al., 2023). In the Santa Monica Mountains, steep topography and the seasonal hot, dry Santa Ana winds further exacerbate wildfire risk, making the region one of the most fire-prone areas in Southern California (Dye et al., 2020; McGranahan et al., 2021).

With most of its surrounding wildlands classified under the Mediterranean ecosystem, Los Angeles has become somewhat synonymous with fire. Home to a population of about 3.8 million (U.S. Census Bureau, 2025), much of the city borders fire-prone landscapes like the Santa Monica Mountains to the west and San Gabriel Mountains to the east. As a result, smoke-induced air quality issues are commonplace, and wildfire preparedness remains a critical concern for residents.

Though Los Angeles is no stranger to fire, the Palisades and Eaton Fires of January 2025 shocked the region and raised concerns over the city's budgeting and preparedness. Collectively, the fires burned 37,000 acres, caused 30 direct fatalities, and led to an estimated 440 additional deaths attributed to indirect causes such as air pollution exposure and delays in healthcare services (CAL FIRE, 2025; Paglino et al., 2025).

***Figure 1: Map of the extent of the Palisades Fire and the Eaton Fire. The fires covered 23,445 acres and 14,021 acres, respectively. Sources: ESRI, CalFire, Google Earth.***




The Santa Ana winds, blowing consistently at 85 miles per hour, were a major driver of the firestorm. Another key factor was the accumulation of vegetation fuel, following two years of unprecedented heavy rainfall, which was then succeeded by severe drought (Hewitt, 2025). In the aftermath, public discourse focused heavily on perceived budget cuts to the Los Angeles Fire Department (LAFD). While the initial budget report indicated an overall reduction in the city's budget, including fire resources, it did not necessarily reflect a targeted cut to the fire department. In response to the criticism, however, city financial analysts clarified that LAFD's operating budget had actually increased by over 7% relative to the previous fiscal year, largely due to allocations for firefighter salary increases and new fire trucks (*Los Angeles Times*, 2025). The gap between the initial reports and the subsequent clarification drew heightened public attention; a fire budget that typically garnered little public interest in stable times had quickly become a focal point for residents, reflecting the wide scale concern for city preparedness in the face of disaster.

Whether every line item of the fire department's approved budget directly contributes to its operational capacity remains uncertain. Freddy Escobar, President of the United Firefighters of Los Angeles City Local 112, reflected that while he felt LAFD should have received more funding for staffing, no amount of money could have quelled flames fuelled by winds of 85 miles per hour (*Los Angeles Times*, 2025). Nonetheless, given established links between wildfire behaviour and climate variability, the sequence of two anomalously wet years followed by a severe drought should have indicated elevated fire risk in 2025. For city officials, this warrants consideration of proactive investment in fire infrastructure and emergency systems.

Modern wildfire modelling techniques have evolved to predict outcomes like ignition probability, spread, intensity, smoke impact, and exposure risk. These models typically rely on environmental and biophysical inputs such as fuel characteristics and weather conditions. They rarely integrate policy-based or infrastructural variables, such as municipal fire budgets, emergency service accessibility, or development indices density. Whether such social indicators should be more directly embedded in wildfire models remains an open question – one that this study seeks to explore.

Wildfire spread and intensity are simulated using a variety of well-established models, each designed with different inputs and use cases. The Fire-Area Simulator (FARSITE) is widely used to arrive at empirical spread projections by integrating fuel load, wind, slope, and weather conditions (Srivas et al., 2016). The Coupled-Atmosphere-Wildland Fire Environment (CAWFE) model specifically analyses the dynamic feedback between atmospheric conditions and fire behaviour, offering a more complex understanding of the relationship in real time (Baptise Filippi et al., 2009). These models were originally developed for wildland contexts, where the analysis lies squarely on physical and environmental variables. Consequently, they do not account for anthropogenic factors, which are increasingly relevant in today's fire-prone, human-modified landscapes.

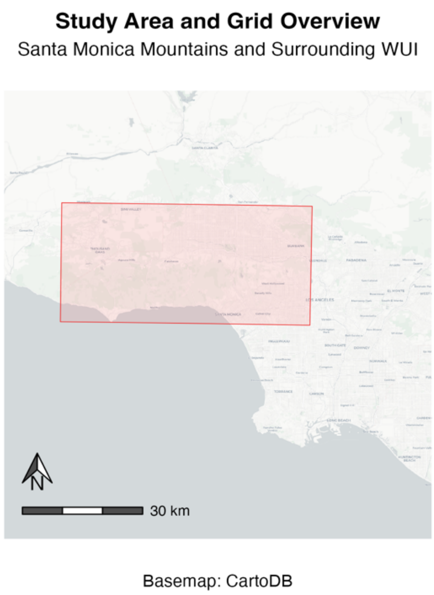
To model wildfire behaviour and risk across complex, multi-dimensional datasets, many recent studies have leveraged more flexible machine learning (ML) approaches such as Support Vector Machines (SVMs) and Random Forests (RF). While SVMs can be powerful with high-dimensional and small-sample datasets common in remote sensing, performance analysis reveals significant degradation when handling high-dimensional imagery and high-complexity input data (Kar et al., 2024). High complexity demands more pre-processing, rendering SVMs inefficient for this application. Previous studies have found RFs to be better suited for severity mapping than other ML classifiers, with consistently high accuracy of over 90% (Hultquist et al., 2014; Collins et al., 2018). It has been noted for its strength in handling both numerical and categorical data, a critical requirement as fire probability depends on a diverse range of variables. Introduced by Breiman (2001), RF is a widely used ML algorithm constructed by a series of decision trees that classify observations by recursively splitting the data into subsets based on the values of input variables, with each tree voting on the final predicted class or value. It is particularly applicable in spatial and remote sensing applications, as it can manage complex interactions between variables where there is high intra-class variability (Ghimite et al., 2012; Belgui and Drāgut, 2016). Wildfire modelling in heterogeneous landscapes with high intra-class variability, therefore, becomes well-suited for random forest analysis. Predictors like fuel variability, moisture, and development often exhibit nonlinear relationships and spatial variability to wildfire outcomes, and the ability to handle the complex interactions between the predictors underscores its strength (Ghimite et al., 2012).

Given these advantages, RF modelling was selected for this study to assess the statistical significance of wildfire drivers across the study area, in relation to whether a fire occurred in a specific grid cell. The method accommodates both categorical and continuous anthropogenic inputs, such as access to emergency services and public budget data. As wildfire behaviour in the WUI increasingly reflects the intersection of climate variability, land use, and socio-spatial dynamics, modelling tools like RF offer a promising avenue for capturing the complexity, and thereby improving risk prediction. To the author's knowledge, this represents a novel contribution to the field, as infrastructure investment in the form of city funding and relative access to emergency services have rarely been treated as equal input parameters in wildfire behaviour modelling.

## 3. Data Collection and Processing

The study area was defined by intersecting the Santa Monica Mountain Northern Recreation Area (SMMNRA) with Los Angeles County, ensuring the area fell under Los Angeles County jurisdiction. A rectangular grid covering the shape was then created, capturing surrounding neighbourhoods in the WUI and further-developed areas like Santa Monica and the San Fernando Valley. Each grid square represents 300 $m^2$ of land coverage. Analysis for the study was performed in California Albers (CRS 3310), suited for the region's scale.

***Figure 2: Map of study area over greater Los Angeles***




This region was selected for its high variation in input data and its well-documented wildfire history, including the Woolsey Fire (2018), Palisades Fires (2021, 2025), and Skirball Fire (2017). It includes both WUI zones along the SMMNRA and adjacent urbanised areas, enabling high-resolution analysis of spatial factors, like vegetation classification and development indices, which vary greatly across short distances in this terrain. The dataset spans 2015-2025 at yearly resolution. This range was selected to align with the availability and consistency of other social, climatic, and environmental input variables.

### 3.1 CAL FIRE Historic Fire Perimeters

Historic fire perimeters are made available by CAL FIRE's Fire and Resource Assessment Program. The dataset, which contains state-wide records wildfire incidents from years 1895 onwards, was filtered for years 2015-2025 and clipped to the study area's boundary. These perimeters serve as the outcome variable against which all input attributes are evaluated in the RF model. The modifications ensure that only fire events occurring wholly or partially within the defined spatial extent were retained. Each grid in the cell study area was assigned a binary value:

- 1 if the cell intersected with a fire perimeter during the study period (indicating at least one recorded burn event).
- 0 if the cell did not intersect with any recorded perimeter.

The binary classification serves as the outcome variable in the RF model, against which all predictor values were evaluated. By treating fire occurrence as a binary event at the grid-cell level, the analysis focused on distinguishing spatial patterns of burned verses unburned areas over the ten-year period.

### 3.2 National Oceanic and Atmospheric Association (NOAA) Climate Data

Climatic variables were obtained from NOAA's Divisional Time Series data service, which provides standardised long-term climate records for geographic divisions across the United States. Because the study area lies entirely within the South Coast Drainage Division, relevant climatic parameters were extracted for this division. Key parameters included the Palmer Hydrological Drought Index (PHDI), total precipitation, and average temperature. Severe wind conditions were represented by NOAA's fastest 2-minute wind speed metric (WSF2), obtained from the nearest weather station (ID USW00093134) located in downtown Los Angeles.

Initial calculations using annual mean values were found to obscure seasonal extremes, which are critical to understanding fire dynamics in Mediterranean ecosystems like Southern California. To address this, variable-specific temporal scopes were applied, aligning each climatic metric with the period most relevant to its influence on fire behaviour. Temperature and wind speed values were averaged across May-September, corresponding to California's peak "fire season". Rainfall and PHDI were instead averaged over the preceding wet season months of December-April, as precipitation during this period determines the growth of fuel.

Finally, the El Niño-Southern Oscillation (ENSO) phenomenon was incorporated as a binary predictor variable, labelling El Niño "on" years. In Southern California, El Niño years are typically associated with above-average winter precipitation, which influences fuel loads. By including ENSO phase in the model, the analysis accounts for a large-scale climatic driver that indirectly shapes fire risk through its global precipitation patterns.

### 3.3 Multi-Resolution Land Characteristics (MRLC) Consortium Land Classification

Land use and land cover classification data were obtained from the Multi-Resolution Land Characteristics (MRLC) Consortium, which provides standardised, high-resolution datasets across the United States. For this study, the 2016 National Land Cover Database (NLCD) raster was selected as a baseline land cover due to its proximity to the study period. The original NLCD dataset includes a wide range of detailed land cover classes, such as evergreen forest, mixed forest, various shrublands types, and multiple categories of developed land. To facilitate analysis within the modelling framework and reduce complexity, these classes were reclassified into three primary categories:

- Forest cover: including all evergreen, deciduous, and mixed forest categories.
- Shrub cover: including all shrub and scrub categories.
- Developed land: including all developed categories (low, medium, and high intensity, as well as developed open space).

The NLCD raster was resampled to match the study grid's spatial resolution of 300m x 300m. For each grid cell, the proportion of each reclassified land cover type was calculated. The dominant category within a cell was assigned a binary "1" for that category, with all other categories assigned a "0". This binary majority-rule approach ensured that each cell was classified according to its prevailing land cover type. This reclassification approach captures major landscape patterns and enhances interpretability, and, importantly, it does so without overcomplicating or overfitting the model.

### 3.4 Municipal, State, and Federal Data Portals

Geospatial layers for county boundaries, the SMMNRA, and water bodies were sourced from the Los Angeles City, Los Angeles County, and California State open data portals. Budget data for the Los Angeles Fire Department (LAFD) was sourced from the LA City Open Budget platform. To normalise across years and account for inflation, annual LAFD budget allocations were expressed as a proportion of the total city budget, with values in the grid represented as percentages. Using proportional allocations rather than raw dollar amounts ensures comparability across years by controlling for changes in the overall municipal budget, allowing the analysis to focus on relative investment priority rather than absolute spending levels.

### 3.5 OpenStreetMap Fire Infrastructure

OpenStreetMap queries provided the geospatial point data for fire stations, fire hydrants, and water tanks relevant to emergency firefighting operations. The dataset was filtered to ensure that only features with verified coordinates and correct attribute tags were retained. These infrastructure points were then processed to create proximity buffers representing optimal service areas for each resource type, with buffer distances informed by National Fire Protection Association guidelines.

Two buffer thresholds were applied:

- Fire hydrants and water sources: buffered at 100 metres, representing an operationally feasible hose-lay distance for firefighting contexts.
- Fire stations: buffered at 1 kilometre, approximating the maximum desirable distance for rapid response in the WUI.

The resulting buffered shapefiles were merged and unionised into a single shapefile, which later represents optimal fire infrastructure access in the grid as an indicator variable. Cells with a value of "1" contained either a fire station, a hydrant or water tank, or both within the designated buffer thresholds. Cells with a value of "0" had no such infrastructure in optimal proximity.

***Table 2. Summary of input variables influencing fire occurrence***

| Indicator Type | Indicator | Definition | Unit |
|---|---|---|---|
| **Climatic** | Palmer Hydrological Drought Index (PHDI) | Average PHDI value across "wet" months December – April | Index Value |
| | Precipitation | Average precipitation across "wet" months December – April | Inches |
| | Temperature | Average temperature across "dry" months May – September | F° |
| | Wind Speed | Average fastest 2-minute wind speed across "dry" months May – September | Metres/Second |
| | El Niño | If the year is classified as an El Niño year, the value is 1; if it is not, the value is 0. | / |
| **Land Use** | Water Bodies | If the grid square is intersected by a body of water, the value is 1; if it is not, the value is 0. | / |
| | Shrub Cover | If the grid cell is covered by majority forest, the value is 1; if it is not, the value is 0. | / |
| | Forest Cover | If the grid cell is covered by majority forest, the value is 1; if it is not, the value is 0. | / |
| | Developed Land Cover | If the grid cell is covered by majority developed land, the value is 1; if it is not, the value is 0. | / |
| **Emergency Service Access** | Fire Infrastructure Proximity | If the grid cell falls within 1 kilometre of a fire station, or 100 metres of a water source intended for fire services, the value is 1; if it does not, the value is 0. | / |
| | LAFD Budget | Percent of annual Los Angeles City budget allocated to LAFD | % |

# 4. Methods

## 4.1 SHAP Value Interpretation

To explain model predictions, the study employs SHAP values. SHAP values are a model-agnostic interpretability tool rooted in game theory, designed to assign a contribution value to each input feature based on its influence on the model's output (Lundberg and Lee, 2017). In this context, SHAP values were computed for each grid cell and associated feature, quantifying the direction (positive or negative) and magnitude of that feature's contribution to the predicted fire probability. A positive SHAP value will indicate a feature's contribution to increased fire risk, while negative values indicate a contribution to the suppression of risk. By summarising these values across all observations, the analysis identifies not only which features are most influential, but also how their effects vary across different spatial contexts.

## 4.2 Relative Importance Score

In addition to SHAP interpretation, traditional feature importance scores were computed using the impurity-based metric native to the RF algorithm. This approach evaluates the total decrease in Gini impurity contributed by each feature across all decision trees in the ensemble model (Breiman, 2001). The Gini importance, or the mean decrease in impurity, quantifies how much each feature, on average across all decision trees, reduces Gini impurity whenever it was used to split a node, thereby indicating its relative contribution to classification performance (Nembrini et al., 2018). By aggregating the impurity reductions at every split where a feature is used, the algorithm generates a relative importance score, which provides a ranked list of the most influential predictors in the model.

These scores are widely used due to their computational efficiency. However, they do not provide information on the directionality of a feature's influence, only the degree to which it helps reduce classification uncertainty (Strobl et al., 2007). As wildfire modelling in heterogenous landscapes includes a wide class of inputs, many of which share nonlinear relationships, directionality it a critical input. Therefore, these scores are best interpreted alongside techniques such as SHAP. The combination of these interpretability methods helps validate the logic of the model while revealing more information on variable influence across the dataset.

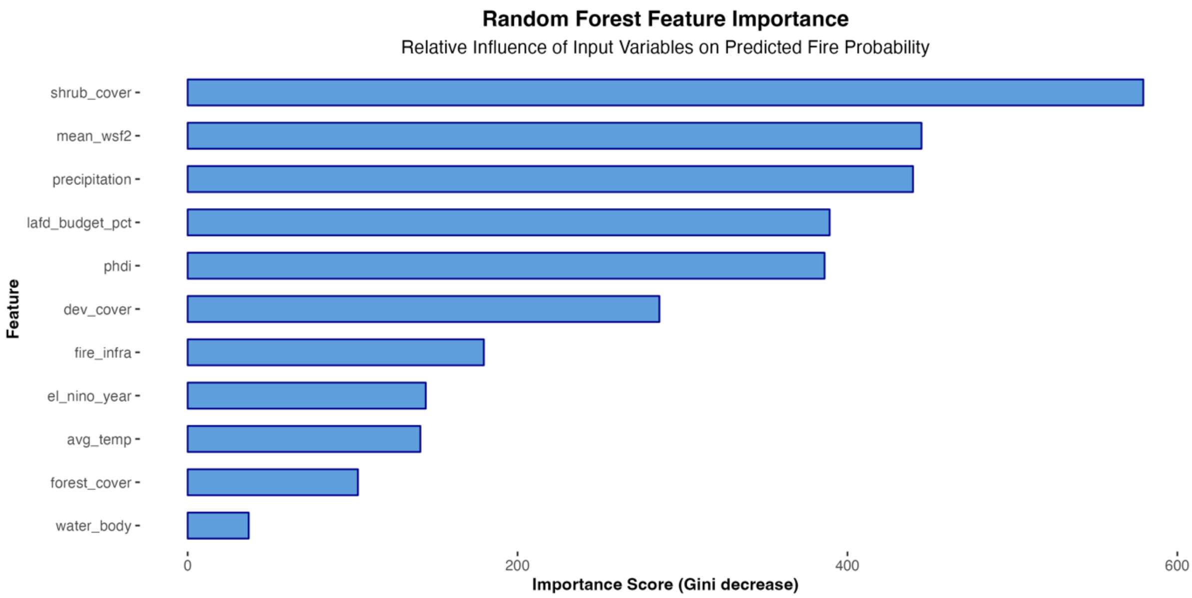
## 4.3 Evaluation Metrics

To assess the model's performance, several standard classification metrics were employed, consistent with common practice in ML and wildfire prediction research (Cutler et al., 2007). These include accuracy, precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). Accuracy provides an overall measure of the proportion of correctly classified instances but can be misleading in datasets with significant class imbalance – a common feature in wildfire occurrence data. Precision evaluates the model's ability to minimise false positives by quantifying the proportion of predicted fire cells that correspond to actual fires, while recall (or sensitivity) measures the proportion of actual fire cells correctly identified by the model, thereby capturing the rate of false negatives. The F1-score, as the harmonic mean of precision and recall, provides a single metric that balances these two performance dimensions (Cutler et al., 2007). Finally, the AUC-ROC reflects the model's ability to discriminate between fire and non-fire classes across a range of classification thresholds, with values closer to 1.0 indicating stronger discriminatory performance.

## 5. Results

To interpret the outputs of the RF classification model and evaluate the relative influence of each predictor, this section presents results from both feature importance rankings and SHAP values. Feature importance scores provide a global measure of each variable's contribution to reducing classification error within the ensemble, while SHAP values offer local interpretability by indicating both the direction and magnitude of each feature's effect on individual predictions. Together, these complementary approaches allow for a nuanced understanding of the models' decision-marking process, highlighting which environmental, climatic, and social factors most strongly influence wildfire occurrence predictions in the study.

***Figure 3: Random Forest Feature Importance, by decrease of Gini Impurity Value***

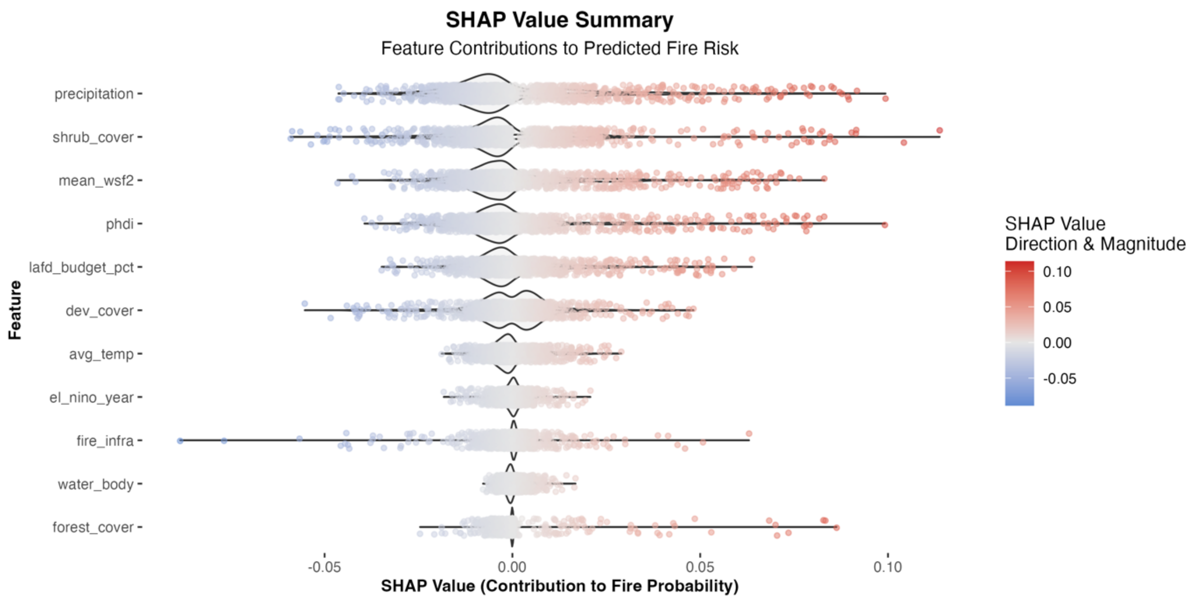



In Figure 2, the feature importance plot displays the relative contribution of each input variable to the model's overall predictive structure, calculated using the mean decrease in Gini impurity. This metric quantifies how frequently a variable is used to split decision trees and the extent to which those splits reduce classification error. Higher scores indicate greater influence on model performance. Figure 3 reveals that shrub cover has the highest Gini Importance Score among all predictors, indicating that the presence and extent of shrub-dominated landscapes was the single most dependable predictor of wildfire occurrence within the study area. This is consistent with established fire ecology in Southern

California, where dense chaparral vegetation provides abundant fine fuel loads, particularly under dry and windy conditions. The prominence of shrub cover in the model underscores the critical role of vegetation type in wildfire dynamics and suggests that fuel management strategies could be effective levers for reducing fire risk in WUI zones.

In contrast, the SHAP summary plot provides both the direction and magnitude of each feature's influence on individual fire probability predictions. In Figure 3, negative correlations are marked in blue, indicating features that tend to decrease predicted fire probability. For example, higher precipitation values are consistently associated with lower fire likelihood, reflecting their dampening effect on fuel flammability. Positive correlations, shown in red, indicate features that contribute to increased fire probability. While feature importance scores from the RF model quantify a variable's overall contribution to reducing classification error, SHAP values add an additional layer of interpretability by showing how, and in which direction, each feature affects predictions for specific observations. Taken together, these tools provide a more holistic understanding of model behaviour. Feature importance highlights global influence, whereas SHAP values highlight local effects and complex feature interactions.

***Figure 4: SHAP Value Summary***




The RF classification model was evaluated using five standard performance metrics: accuracy, precision, recall, F1 Score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). After lowering the classification threshold from 0.5 to 0.2, the model demonstrated improved sensitivity to fire occurrence events. Accuracy remained high at 96.6%, indicating that the model correctly classified the majority of grid cells overall. Precision was measured at 43.7%, meaning that of all predicted fires, roughly 44% corresponded to actual fire events. Recall, which had been near-zero before the threshold adjustment, improved substantially to 55.5%, reflecting the model's enhanced ability to correctly identify true positives. The F1 score, balancing precision, and recall, reached 48.9%. This represents a reasonable trade-off, given the significant class imbalance of fire occurrence in the dataset. Finally, the AUC value of 0.93 suggested a strong overall discriminative ability, indicating that the model is effective at distinguishing between fire and non-fire conditions.

***Table 3: Model Evaluation Metrics of Random Forest Model Performance***

| Metric | Value |
|---|---|
| Accuracy | 0.96551869 |
| Precision | 0.43715239 |
| Recall | 0.55547703 |
| F1 Score | 0.48926237 |
| AUC | 0.92850501 |

The threshold adjustment was necessary because the original model exhibited a strong bias toward predicting the majority class (0, no fire). Under the default 0.5 threshold, the model's recall for fire events was near zero, meaning it was failing to identify almost all true positive cases. In the context of wildfire prediction, this is particularly problematic: a model that consistently misses actual fire events is of limited practical value, especially for preparedness planning where false negatives (unpredicted fires) can have severe consequences. By lowering the classification threshold to 0.2, the model became more sensitive to potential fire occurrences, trading a degree of precision for a notable gain in recall. This trade-off is often acceptable in disaster risk modelling, where the cost of a false positive (predicting a fire that does not occur) is consistently far lower than the cost of a false negative (failing to predict a fire that does occur). In this case, the adjusted threshold allowed the model to flag more at-risk grid cells.

***Table 4: Feature Inputs and Respective Gini Importance Values and Absolute SHAP Values***

| Feature | Gini Importance | Mean Absolute SHAP |
|---|---|---|
| Shrub Cover | 579.36346 | 0.01195 |
| Peak Wind Speed (WSF2) | 444.87917 | 0.01089 |
| Precipitation | 439.73047 | 0.01386 |
| LAFD Budget Percentage | 389.20728 | 0.00813 |
| Drought Index (PHDI) | 386.11639 | 0.01051 |
| Development Cover | 285.98195 | 0.00747 |
| Access to Fire Infrastructure | 179.54882 | 0.00365 |
| El Niño Phase | 144.35506 | 0.0025 |
| Average Temperature | 141.06698 | 0.0045 |
| Forest Cover | 103.22068 | 0.00253 |
| Water Body Presence | 36.999631 | 0.00183 |

The results of the Random Forest classification model, supported by both Gini-based feature importance scores and SHAP value interpretation, underscore the dominant role of environmental factors in predicting wildfire occurrence. Vegetation density (specifically shrub cover), peak wind speeds, and precipitation emerged as the most influential features in the model. This finding aligns with existing literature, which has long emphasised the role of fuel availability and weather conditions in shaping fire behaviour (Collins et al., 2018; Hultquist et al., 2014).

Critically, the results also support the central hypothesis of this study: infrastructural preparedness measures do contribute meaningfully to fire risk prediction. Used as a proxy for emergency response capacity, the percentage of overall LA City funding allocated to LAFD annually ranked fourth overall in Gini importance, above several environmental features including development cover, El Niño phase, and average temperature. Its SHAP value, while more moderate, still indicates a consistent directional influence on fire probability. Similarly, the merged variable representing access to fire infrastructure (fire stations and water sources) demonstrated a weaker but non-negligible contribution to model output.

These results suggest that while environmental conditions remain the primary determinants of wildfire occurrence, strategic investment in emergency services and infrastructure may influence outcomes at the margins. The model's ability to capture the statistical relevance of budgetary and infrastructural inputs affirms that wildfire risk is not only a product of climatic and ecological conditions, but also of human decisions around preparedness and capacity-building.

## 6. Discussion

The results of this study reveal a nuanced picture of wildfire drivers in the western Santa Monica Mountains and adjacent WUI zones. Consistent with the body of wildfire science, climatic and environmental variables emerged as the most influential predictors of wildfire occurrence (Collins et al., 2018; Hultquist et al., 2014). In this study's area, shrub cover, peak wind speeds, and precipitation were the most statistically significant predictors; this reinforces decades of research underscoring the primacy of fuel and weather conditions in influencing fire propagation and intensity (McGranahan et al., 2021). In chaparral-dominated ecosystems such as this study area, the combination of high biomass and prolonged seasonal drying creates a baseline of flammability that is readily exacerbated under high-wind conditions.

Yet, the model's findings also support the central hypothesis that infrastructural preparedness measures have a measurable, though comparatively smaller, influence on wildfire outcomes. Among all variables, the proportion of LA City budget allocated annually to LAFD ranked fourth in Gini importance. It landed above several environmental features traditionally associated with fire probability, such as development cover and temperature averages across the dry season. This suggest that social and infrastructural investments, even when represented by broad proxies, can impact the statistical balance of wildfire risk models. While their predictive strength did not match that of vegetation structure or precipitation totals, their consistent inclusion in tree-splitting decisions and their directional influence in SHAP plots indicate that they are not merely statistical noise.

This finding aligns with recent arguments in the wildfire resilience literature that emphasise the coupled nature of human and natural systems (Calkin et al., 2013). Though human preparedness cannot eliminate the fundamental role of climatic drivers, it can inform preparation measures that can reduce the scale of loss. For example, in years with high fuel loads and forecasted Santa Ana wind activity, a better-funded fire department might position more crews, initiate earlier evacuations, or expand air support contracts in advance of critical weather days.

It is particularly notable that the combined "emergency infrastructure access" variable showed weaker, but still discernible, contributions to the model output. As it represented proximity to fire stations and hydrants, its limitations would be attributed to the scale and processing of the data; a grid cell's proximity to a fire station does not necessarily capture the availability of resources at that station, its staffing level, or its ability to deploy rapidly in the event of severe weather. As such, the variable may underestimate the true influence of spatial emergency service access, illuminating the need for richer operational datasets in future modelling efforts.

**6.1 Preparedness Significance in Context**

While the preparedness proxies in this study demonstrated statistical significance, their relative effect size was modest compared to natural processes. The relatively small-scale statistical significance in preparedness proxies may reflect the inherent nature of extreme wildfires rather than the irrelevance of preparedness measures. Wildfires, particularly under extreme conditions, are natural disasters that can quickly surpass the limits of human intervention (McGranahan et al., 2021). The presence of a well-equipped fire service and constant view of fire hydrants across cities fosters a social expectation that fire can, and perhaps should, always be contained. While this may be true for small-scale structural fires, it is not always the case for large-scale firestorms. One can liken it to fighting a hurricane; under conditions such as sustained 85 mph winds, aerial suppression becomes impossible, and ground crews cannot safely navigate steep terrain to confront a firestorm. These realities underscore that even the most capable firefighting force may be unable to quell a wildfire operating at an unmanageable scale, highlighting the need for both preventative measures and a realistic public understanding of fire behaviour.

In the case of this model, an important interpretive point arises: statistical modesty in preparedness variables should not be read as irrelevance. Rather, it reflects that preparedness is a necessary but insufficient condition for avoiding large-scale damage. In this sense, the finding that LAFD budget share matters at all is significant, because it suggests that human investment can shape outcomes, even if only at the margins, within the much larger envelope defined by climate and vegetation. Moreover, preparedness investments can lead to modernised equipment and improved fire mapping capabilities, all of which can create more resilient response systems for future seasons. These systemic benefits, though difficult to isolate statistically in a single modelling window, are a critical dimension of long-term wildfire adaptation in the WUI.

### 6.2 Policy and Management Implications

The policy relevance of these findings is most immediate for municipalities with jurisdiction over expanding WUI zones. In Los Angeles, where high-value infrastructure and dense population intersect with fire-prone landscapes, integrating models like the one developed in this study could support more strategic allocation of resources. For example, identifying years in which climatic predictors signal elevated risk could trigger pre-emptive planning and adjustments to departmental budgets or fuel treatment schedules. Similarly, outputs could inform the prioritisation of hydrant installations, routine water supply checks, or the pre-positioning of firefighting staff. While climatic variability is becoming increasingly difficult to predict in the wake of climate change, historical patterns indicate that severe drought often coincides with El Niño “off” years. By monitoring these indices, city and state agencies could anticipate elevated fire risk and respond proactively. Similarly, model outputs could guide spatially targeted interventions. Hydrant installations and water infrastructure updates could be prioritised in areas where both vegetation coverage and access to firefighting resources indicate elevated risk. Such measures do not prevent extreme events but can reduce the risk of harm to human life and structural damage.

Other WUI cities such as San Diego (CA), Santa Fe (NM), and Boulder (CO), face similar structural challenges and could adapt a preparedness-variable modelling approach to their own contexts. Even if preparedness variables rank below climatic drivers in overall predictive influence, their local significance could be greater in regions where environmental variability is less extreme, or where specific infrastructure deficits exist. The broader implication is that preparedness modelling can reveal jurisdiction-specific leverage points, allowing for targeted, data-informed intervention strategies.

### 6.3 Balancing Quantitative Methodology and Lived Experience

The quantitative scope of this model is partial. It is critical to remember, across all facets of data science but particularly with disaster modelling, that quantitative modelling captures only part of the larger question of preparedness, resilience, and recovery. Lived experiences and professional expertise remain essential components of understanding fire preparedness. After the 2025 LA Fires, Chief of LAFD Kristin Crowley asserted that the City of LA did not offer adequate monetary support, and that their response fell short due to budget and labour constraints (*Los Angeles Times*, 2025). As a life-long fire professional, perspectives such as Crowley's should be prioritised in planning strategies; they offer insight into operational bottlenecks and on-the-ground conditions that cannot be captured by satellite imagery or OpenStreetMap.

**6.4 Limitations**

The temporal scope of this study was constrained by several factors. First, the availability of citywide budget records in consistent public formats is slim. Variations in municipal reporting structures limit both the number of years available for analysis, and the comparability of data across those years; consequently, the model's social preparedness layer, central to the discussion of this study, relies on the broad budgetary proxy of proportion of annual budget allocated to the Los Angeles Fire Department (LAFD). The temporal scope was further narrowed by computational limitations, which necessitated scaling down the span of years to produce a functional model. While the model would greatly benefit from incorporating a longer historical record of wildfire occurrence in relation to all other input variables, such an expansion would require substantially greater processing power to manage the increased data volume and complexity.

Moreover, the spatial resolution of this model imposes inherent limitations on its findings. While the study originally aimed to cover a broader geographic area at a finer spatial resolution, constraints in computing power necessitated reductions across spatial and temporal extents and grid resolution to accommodate the integration of the wide range of input variables. The final model operates at a relatively high spatial resolution of 300 $m^2$, which allows for detailed representation of land use characteristics and access to physical emergency amenities. However, vegetation classification can vary greatly across small stretches of land, so the granularity of that classification was lost. Additionally, computational resource limitations required a reduction in the temporal granularity of input variables. For instance, month-by-month climatic data points were standardised as mean values across key months, rather than incorporated as time-specific variables. As emphasised in existing literature, wildfire behaviour is highly sensitive to both temporal and spatial variation (McGranahan et al., 2021); thus, the current model is not designed to capture the full spatiotemporal complexity of wildfire behaviour at finer analytical scales. With more computational power, NLCD classification values could be computed and represented on a year-to-year basis, allowing the model to account for changes in land use over time. Such temporal dynamism would likely improve prediction accuracy by reflecting more realistic fuel and development conditions – though it should be noted that the model's predictive performance was strong even within the current limitations.

Variable representation and the reliance on proxies for complex underlying processes introduces limitations to any data project. Even where data were available, several variables in this model are necessarily proxies for broader systems and conditions. For example, the proportion of budget allocated to LAFD does not directly measure operational readiness or equipment condition. The model would benefit greatly from more granular, policy-level expenditure data. Access to detailed budget line items, such as those related to staffing, equipment acquisition, or fire prevention initiatives, would enable a more targeted analysis of how specific investments influence wildfire behaviour and outcomes. Similarly, proximity to a fire station or hydrant does not capture whether those resources were accessible or adequately staffed during a fire event. These simplifications may underrepresent the true effect of preparedness measures, and the model would benefit greatly from more direct representative data.

Presently, other proxies for preparedness and policy investments such as brush-clearing programs or the timing and location of fuel break construction, are either undocumented or unavailable in a format suitable for analysis. Although the California Conservation Corps, a state-funded agency, is involved in firebreak construction, publicly available budgetary data are inconsistent, and there is no comprehensive record of the spatial or temporal distribution of their projects (California Conservation Corps, 2025). This data gap is significant, as such preventative measures play a critical role in reducing fire intensity and spread, particularly within chaparral ecosystems (McGranahan et al., 2021). The model likewise excludes certain private mitigation resources, including residential swimming pools, rooftop sprinklers, and contracted private firefighting services. All had been documented as influential during recent fire events like the 2025 LA Fires, with fire departments utilising pool water in their efforts to stop the flames from spreading from home to home (*Los Angeles Times, 2025).* Publicly accessible datasets, such as OpenStreetMap, do not consistently encode these private features, thereby limiting their integration into spatial analysis.

### 6.5 Future Considerations

While the present model falls short due to its limited temporal and spatial scope, it establishes a promising baseline for future studies. With greater computational capacity and improved access to detailed, time-and-space-sensitive datasets, future research could incorporate finer temporal resolution, dynamic land cover updates, and a more diverse suite of social and infrastructural variables. Expanding both the spatial and informational scope would enhance the model's ability to capture the model's ability to capture the complex, multi-scalar factors that influence wildfire occurrence and severity in the WUI.

## 7. Conclusion

This study applied a Random Forest classification framework, supported by both impurity-based feature importance scores and SHAP value interpretation, to examine the relative roles of environmental, climatic, and social variables in shaping wildfire occurrence in the western Santa Monica Mountains and surrounding WUI. By integrating traditional biophysical predictors with proxies for preparedness such as fire department budget share and proximity to firefighting infrastructure, the model offered a more holistic lens on wildfire risk, one that situates human decision-making within the broader ecological and climatic context. Importantly, it treats human involvement as an equal input variable to the fire outcome variable.

The modelling outputs reinforced a consistent theme in wildfire research: fuels and weather remain the dominant forces driving fire behaviour, particularly in chaparral ecosystems subject to seasonal drying and high-wind events. At the same time, the statistical presence of preparedness measures within the model's decision structure suggests that human investments in readiness can meaningfully influence outcomes, even if their effect size is more modest. The combined use of Random Forest importance rankings and SHAP plots made it possible to capture both the overall contribution of each variable and the nuanced, directional ways in which those variables interact at the local scale. This is a clear interpretive advantage that purely aggregate metrics would not have afforded.

By framing fire risk as the product of both natural processes and human agency, this work underscores the importance of incorporating social and infrastructural data into wildfire prediction efforts. While the preparedness proxies employed here are necessarily coarse, their measurable role in the model signals the potential for more refined, operational datasets to strengthen predictive performance and inform policy. As WUI communities continue to expand into fire-prone landscapes, predictive modelling that combines physical science with social metrics offers a path toward more strategic, data-informed resource allocation. The challenge moving forward lies in improving data resolution, temporal sensitivity, and social variable specificity so that models not only predict where fire is likely to occur, but also guide interventions that reduce its most destructive impacts.

## 8. References


Baptiste Filippi, Jean, et al. “Coupled atmosphere-wildland fire modelling.” *Journal of Advances in Modeling Earth Systems*, vol. 1, no. 4, Apr. 2009, https://doi.org/10.3894/james.2009.1.11.

Barro, S.C., and S.G. Conard. “Fire effects on California chaparral systems: An overview.” *Environment International*, vol. 17, no. 2–3, Jan. 1991, pp. 135–149, https://doi.org/10.1016/0160-4120(91)90096-9.

Belgiu, Mariana, and Lucian Drăguţ. “Random Forest in remote sensing: A review of applications and Future Directions.” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 114, Apr. 2016, pp. 24–31, https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Breiman, Leo. “Random Forests.” *Machine Learning*, vol. 45, no. 1, Oct. 2001, pp. 5–32, https://doi.org/10.1023/a:1010933404324.

Calkin, David E., et al. “How risk management can prevent future wildfire disasters in the wildland-urban interface.” *Proceedings of the National Academy of Sciences*, vol. 111, no. 2, 16 Dec. 2013, pp. 746–751, https://doi.org/10.1073/pnas.1315088111.

Collins, L., et al. “The utility of random forests for wildfire severity mapping.” *Remote Sensing of Environment*, vol. 216, Oct. 2018, pp. 374–384, https://doi.org/10.1016/j.rse.2018.07.005.

Cutler, D. Richard, et al. “Random forests for classification in ecology.” *Ecology*, vol. 88, no. 11, Nov. 2007, pp. 2783–2792, https://doi.org/10.1890/07-0539.1.

Davis, Mike. “The case for letting Malibu Burn.” *Environmental History Review*, vol. 19, no. 2, 1 June 1995, pp. 1–36, https://doi.org/10.2307/3984830.

Dye, Alex W., et al. “Spatial heterogeneity of winds during Santa Ana and non-Santa Ana wildfires in Southern California with implications for fire risk modeling.” *Heliyon*, vol. 6, no. 6, June 2020, https://doi.org/10.1016/j.heliyon.2020.e04159.

Gaudet, Benjamin, et al. “A review of post-incident studies for Wildland-Urban Interface fires.” *Journal of Safety Science and Resilience*, vol. 1, no. 1, Sept. 2020, pp. 59–65, https://doi.org/10.1016/j.jnlssr.2020.06.010.

Haggerty, Noah. “Will Personal Firefighting Devices Help or Hurt in Future Wildfires?” *Los Angeles Times*, Los Angeles Times, 15 July 2025, www.latimes.com/environment/story/2025-07-15/home- firefighting-technology-developed-by-palisades-resident.

Hewitt, Alison. “Floods, Droughts, Then Fires: Hydroclimate Whiplash Is Speeding up Globally.” *UCLA*, UCLA, 29 May 2025, newsroom.ucla.edu/releases/floods-droughts-fires-hydroclimate-whiplash-speeding-up-globally.

Jolly, W. Matt, et al. “Climate-induced variations in global wildfire danger from 1979 to 2013.” *Nature Communications*, vol. 6, no. 1, 14 July 2015, https://doi.org/10.1038/ncomms8537.

Kar, Ankan, et al. “Performance analysis of Support Vector Machine (SVM) on challenging datasets for forest fire detection.” *International Journal of Communications, Network and System Sciences*, vol. 17, no. 02, 2024, pp. 11–29, https://doi.org/10.4236/ijcns.2024.172002.

Keeley, Jon E., and Alexandra D. Syphard. "Large California wildfires: 2020 fires in historical context." *Fire Ecology*, vol. 17, no. 1, 25 Aug. 2021, https://doi.org/10.1186/s42408-021-00110-7.

Lundberg, Scott, and Su-In Lee. "A Unified Approach to Interpreting Model Predictions." *arXiv*, 2017, doi:10.48550/arXiv.1705.07874.

McGranahan, Devan Allen, and Carissa L. Wonkka. *Ecology of Fire-Dependent Ecosystems: Wildland Fire Science, Policy, and Management*. CRC Press, 2021.

Muñoz-Gálvez, Francisco J., et al. "Trait coordination and trade-offs constrain the diversity of water use strategies in Mediterranean Woody Plants." *Nature Communications*, vol. 16, no. 1, 2 May 2025, https://doi.org/10.1038/s41467-025-59348-3.

Nembrini, Stefano, et al. "The revival of the Gini Importance?" *Bioinformatics*, vol. 34, no. 21, 10 May 2018, pp. 3711–3718, https://doi.org/10.1093/bioinformatics/bty373.

Paglino, Eugenio, et al. "Excess deaths attributable to the Los Angeles wildfires from January 5 to February 1, 2025." *JAMA*, 6 Aug. 2025, https://doi.org/10.1001/jama.2025.10556.

Radeloff, V. C., et al. "The wildland–urban interface in the United States." *Ecological Applications*, vol. 15, no. 3, June 2005, pp. 799–805, https://doi.org/10.1890/04-1413.

Radeloff, Volker C., Miranda H. Mockrin, et al. "Rising wildfire risk to houses in the United States, especially in grasslands and shrublands." *Science*, vol. 382, no. 6671, 9 Nov. 2023, pp. 702–707, https://doi.org/10.1126/science.ade9223.

Santana, Victor M., et al. "Redirecting fire-prone Mediterranean ecosystems toward more resilient and less flammable communities." *Journal of Environmental Management*, vol. 215, June 2018, pp. 108–115, https://doi.org/10.1016/j.jenvman.2018.03.063.

Schug, Franz, et al. "The global wildland–urban interface." *Nature*, vol. 621, 19 July 2023, pp. 94–99, https://doi.org/10.1038/s41586-023-06320-0.

Srivas, Thayjes, et al. "Wildfire spread prediction and assimilation for FARSITE using ensemble Kalman filtering 1." *Procedia Computer Science*, vol. 80, 2016, pp. 897–908, https://doi.org/10.1016/j.procs.2016.05.328.

"Statistics ." *National Interagency Fire Center*, www.nifc.gov/fire-information/statistics. Accessed 13 Aug. 2025.

Stephens, S. L., et al. "Managing forests and fire in changing climates." *Science*, vol. 342, no. 6154, 4 Oct. 2013, pp. 41–42, https://doi.org/10.1126/science.1240294.

Strobl, Carolin, et al. "Bias in random forest variable importance measures: Illustrations, sources and a solution." *BMC Bioinformatics*, vol. 8, no. 1, 25 Jan. 2007, https://doi.org/10.1186/1471-2105-8-25.

Westerling, Anthony LeRoy. "Increasing western US forest wildfire activity: Sensitivity to changes in the timing of Spring." *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 371, no. 1696, 5 June 2016, p. 20150178, https://doi.org/10.1098/rstb.2015.0178.

Zahniser, David. “Did Mayor Karen Bass Really Cut the Fire Department Budget? The Answer Gets Tricky.” *Los Angeles Times*, Los Angeles Times, 10 Jan. 2025, www.latimes.com/california/story/2025-01-10/how-much-did-the-l-a-fire-department-really-cut-its-budget.

## Appendix

All coding scripts and data files (raw and modified) for this study were submitted in a compressed Zip file on Moodle. They are also uploaded on Dropbox. For reproducibility and consistency, users should replace the file paths at the beginning of each script with their preferred local file path.