FireScape: A Data-Driven Approach to Fire-Resilient Tree Planting in Los Angeles

June 5th, 2025

Authors:

UCLA
Hengyuan Liu - hengyuanliu@ucla.edu

Tufts University
Xinhui Luo - xinhui.luo@tufts.edu
Jiahui Xu - jiahui.xu@tufts.edu

ETH zürich
Li Yuan - leeyuan@ethz.ch

Harvard University
Zhilu Chen - lucchen@fas.harvard.edu

GitHub Repository:

https://github.com/GitData-GA/FireScape

Source Data:

[1] TreeKeeper Inventory Management System https://losangelesca.treekeepersoftware.com

[2] California Fire Perimeters https://services1.arcgis.com/jUJYIo9tSA7EHvfZ/ArcGIS/rest/services/California_Historic_Fire_Perimeters/FeatureServer/0

[3] Los Angeles Almanac https://www.laalmanac.com/weather/we04a.php

Image Source: Wikimedia

Abstract

This project explores how urban tree species influence wildfire risk across Los Angeles using a multinomial logistic regression model. By combining over 850,000 tree records with historical wildfire perimeters and climate data, we created a dataset linking tree locations to fire exposure outcomes. The model estimates the probability of a tree being exposed to different fire size categories including small, medium, large, or very large based on species and environmental context. Our findings reveal clear differences in fire risk across tree types, supporting targeted planting strategies to reduce wildfire vulnerability. This approach demonstrates the power of statistical modeling and open data to inform urban forestry decisions and enhance climate resilience in cities.

Introduction

Inspired by the major wildfire that struck the Los Angeles area earlier this year Jan. 2025, our project investigates the relationship between urban trees and fire risk. We aim to understand how factors such as tree species, location, and environmental conditions contribute to wildfire occurrence and severity, with the ultimate goal of guiding safer tree planting practices in fire-prone urban areas.

To do this, we aggregated and merged multiple datasets: street and park tree inventories, historical temperature records, and past wildfire perimeters. Each tree was associated with one or more fire events based on spatial overlap, allowing us to analyze patterns between current tree distribution and historical fire activity. In some cases, the same tree appeared more than once in the dataset, which we accounted for in our analysis.

Data Processing & Description

We obtained our tree dataset from the Los Angeles City Tree Inventory, accessible through the TreeKeeper platform [1]. The original tree dataset consisted of two separate sources: one containing trees located along streets and another covering trees in public parks. We merged these datasets to create a comprehensive view of urban vegetation in Los Angeles. After cleaning and combining the data, the final dataset included 848,194 records. Each record contains information on a single tree, including its geographical coordinates (Latitude and Longitude), species name, and diameter at breast height (DBH).

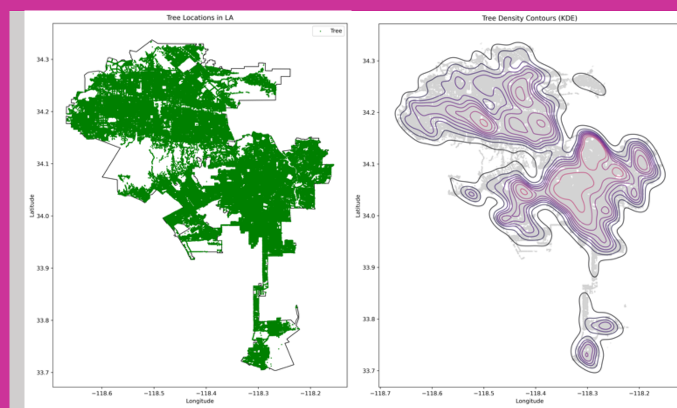
We obtained the fire perimeter data from the California Historic Fire Perimeters dataset, hosted by the California Department of Forestry and Fire Protection (CAL FIRE) and accessible via the ArcGIS REST service [2]. This dataset contains spatial and temporal information on historic wildfires across California, including attributes such as the fire name, year, cause, and total acreage burned. It provides detailed polygon perimeters for each fire event, which we used to analyze fire occurrence and severity patterns in relation to urban tree distribution within Los Angeles. We categorized wildfire sizes based on USDA standards, where fires under 100 acres are labeled "Small" (Classes A–C), those between 100 and 1,000 acres are "Medium" (Classes D–E), and fires over 1,000 acres are "Large" or "Very Large" (Classes F–G), with a special "Very Large" category for those exceeding 10,000 acres.

We obtained historical temperature data for Los Angeles [3]. This source provides annual average temperature records for downtown Los Angeles, covering multiple decades. The data includes yearly average temperatures, offering a general view of climate trends over time. Although the temperature data is not spatially granular, it provides useful contextual information for analyzing long-term environmental conditions that may influence fire activity and vegetation health in the region.

We first collected and combined data from three main sources: Los Angeles street and park tree inventories, historical wildfire perimeter records, and yearly temperature data. The tree data included information on species and geographic coordinates and we deleted non-tree plants like bushes and flowers. Each tree was assigned a unique ID to track and manage duplicates across datasets. The wildfire data provided perimeter boundaries and metadata such as fire size and year of occurrence. For environmental context, we incorporated historical yearly average temperature data, matched to each row based on the year.

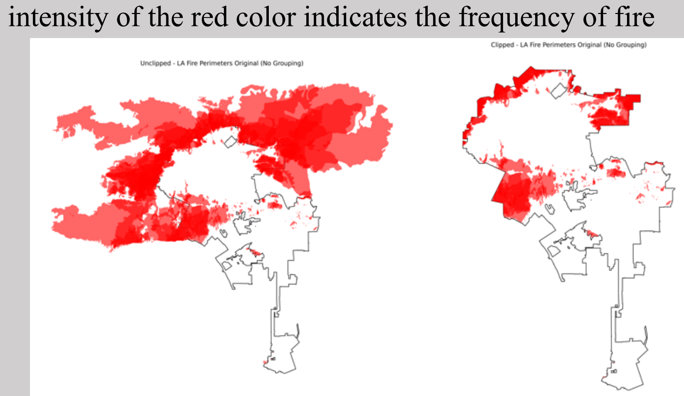
To integrate the datasets, we used spatial joins to determine whether each tree fell within the boundary of a recorded fire. A single tree could appear in multiple fire events if located within overlapping perimeters. We matched each tree-fire interaction to its corresponding fire severity based on fire size (in acres). Each row in our final dataset represented a tree-fire interaction and was labeled with the fire size category as the target variable.

Visualization



The top left plot shows individual tree locations across Los Angeles as green dots within the city boundary, revealing dense coverage in central and northern areas with some sparse inner-city zones. The top right plot uses kernel density estimation (KDE) to highlight tree density patterns, where darker contours represent regions with higher concentrations of trees. Together, they provide a comprehensive understanding of urban tree distribution across the city.

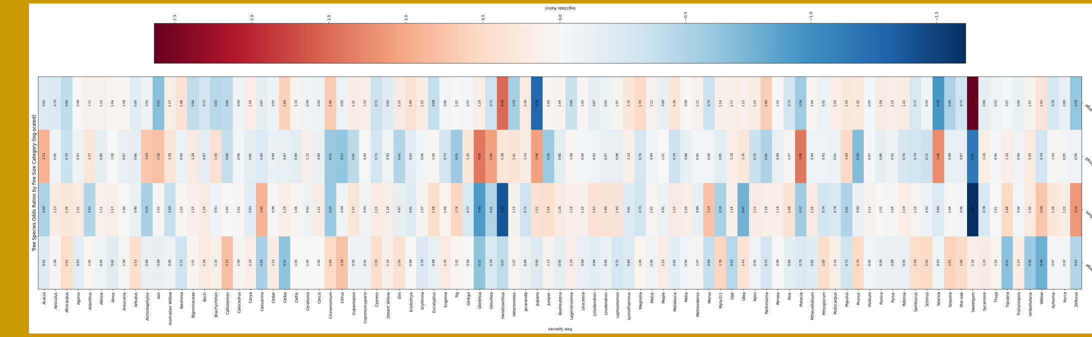
In the fire perimeter visualization, the intensity of the red color indicates the frequency of fire occurrences in a given area. Each fire perimeter is drawn with a semi-transparent red fill. As a result, areas where only a single fire has occurred appear lighter red, while regions where multiple fire perimeters overlap appear darker due to the stacking of transparent layers. This visual layering effectively highlights zones that have experienced repeated fire events over time.



Modeling

To model the relationship between these predictors and fire size outcomes, we applied a multinomial logistic regression. This statistical model is well-suited for categorical target variables with more than two unordered classes. Using "small fire" as the reference category, the model estimated the log-odds of a tree being associated with a medium, large, or very large fire based on the predictor variables. The model outputs odds ratios, which allow us to interpret how specific tree species, climate conditions, or canopy densities increase or decrease the likelihood of exposure to larger fires.
One of the major limitations of our approach is the lack of historical tree data. Our analysis assumes that the current tree distribution is representative of vegetation conditions during past fire events, which may not hold true in areas that experienced replanting or removal following a fire. Another limitation is the use of yearly average temperature data, which may obscure important short-term weather variations such as heatwaves or extreme droughts that significantly influence fire ignition and spread.

$$\hat{y}_i = \operatorname{argmax}_j \left(\frac{e^{\beta_{0j} + \beta_{1j} x_{1i} + \beta_{2j} x_{2i} + \ldots + \beta_{mj} x_{mi}}}{\sum_{k=0}^{N-1} e^{\beta_{0k} + \beta_{1k} x_{1i} + \beta_{2k} x_{2i} + \ldots + \beta_{mk} x_{mi}}} \right)$$



Analysis & Results

Table 1: Top 5 Tree Species Ranked by Fire Prevention Score.

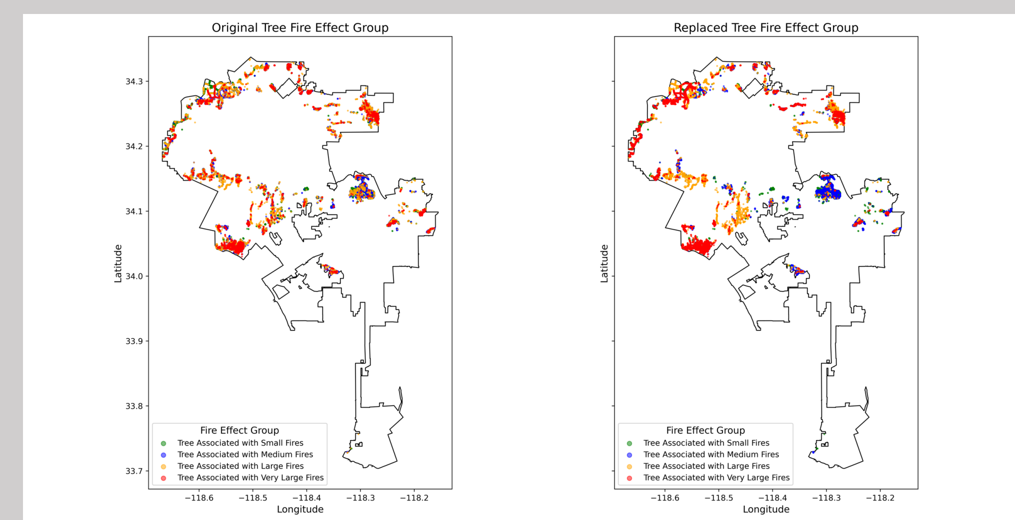
Small	Medium	Large	Very Large
Sweetgum (0.319)	Sweetgum (0.201)	Willow (0.421)	Juglans (0.280)
Prunus (0.509)	Handroanthus (0.250)	Tipuana (0.524)	Searsia (0.399)
Citrus (0.536)	Gleditsia (0.378)	Gleditsia (0.527)	Ash (0.539)
Juniper (0.541)	Olea (0.437)	Ceiba (0.534)	Pistacia (0.543)
Cinnamomum (0.554)	Pistacia (0.556)	Casuarina (0.545)	Zelkova (0.554)

Table 1 identifies tree species most associated with reduced fire risk. Sweetgum (Liquidambar styraciflua) consistently ranks first for small and medium fires, with odds ratios of 0.319 and 0.201, likely due to its high leaf moisture, dense canopy, and low flammability. Willow (Salix spp) shows the strongest negative link with large fires (0.421), benefiting from high water content and riparian habitat. Juglans (walnut) has the lowest odds for very large fires (0.280), possibly due to moderate fire resistance and sparse presence in high-risk zones. These findings highlight how species traits can shape fire exposure and guide risk-reducing planting strategies.

Table 2: Fire size category counts before and after replacing the tree group using model predictions.

Fire Category	Original Count	After Replacement	Change
Small	4,550	1,470	-3,080
Medium	20,362	23,269	+2,907
Large	9,586	9,990	+404
Very Large	17,398	17,167	-231

After replacing tree groups based on model predictions, small fires decreased by ~3,080 and very large fires dropped by 231. Meanwhile, medium fires rose by 2,900 and large fires by 404. This redistribution reflects a shift from extreme fire categories (small and very large) toward more moderate ones (medium and large).



Conclusion & Suggestion

This project demonstrates how urban tree species influence wildfire risk in Los Angeles. Using a multinomial logistic regression model, we identified species like Sweetgum, Willow, and Juglans as being negatively associated with fire occurrence across different severity levels. By simulating a replacement strategy guided by model predictions, we observed a shift away from extreme fire categories, reducing both small and very large fires while increasing medium and large ones. These findings suggest that strategic tree planting can play a meaningful role in urban fire resilience. Future work could incorporate more granular environmental variables, such as soil moisture, wind speed, and real-time weather conditions. Additionally, using remote sensing data and more advanced models like spatial-temporal machine learning could improve prediction accuracy. Partnering with city agencies may help integrate these findings into public tree planning and fire prevention strategies.