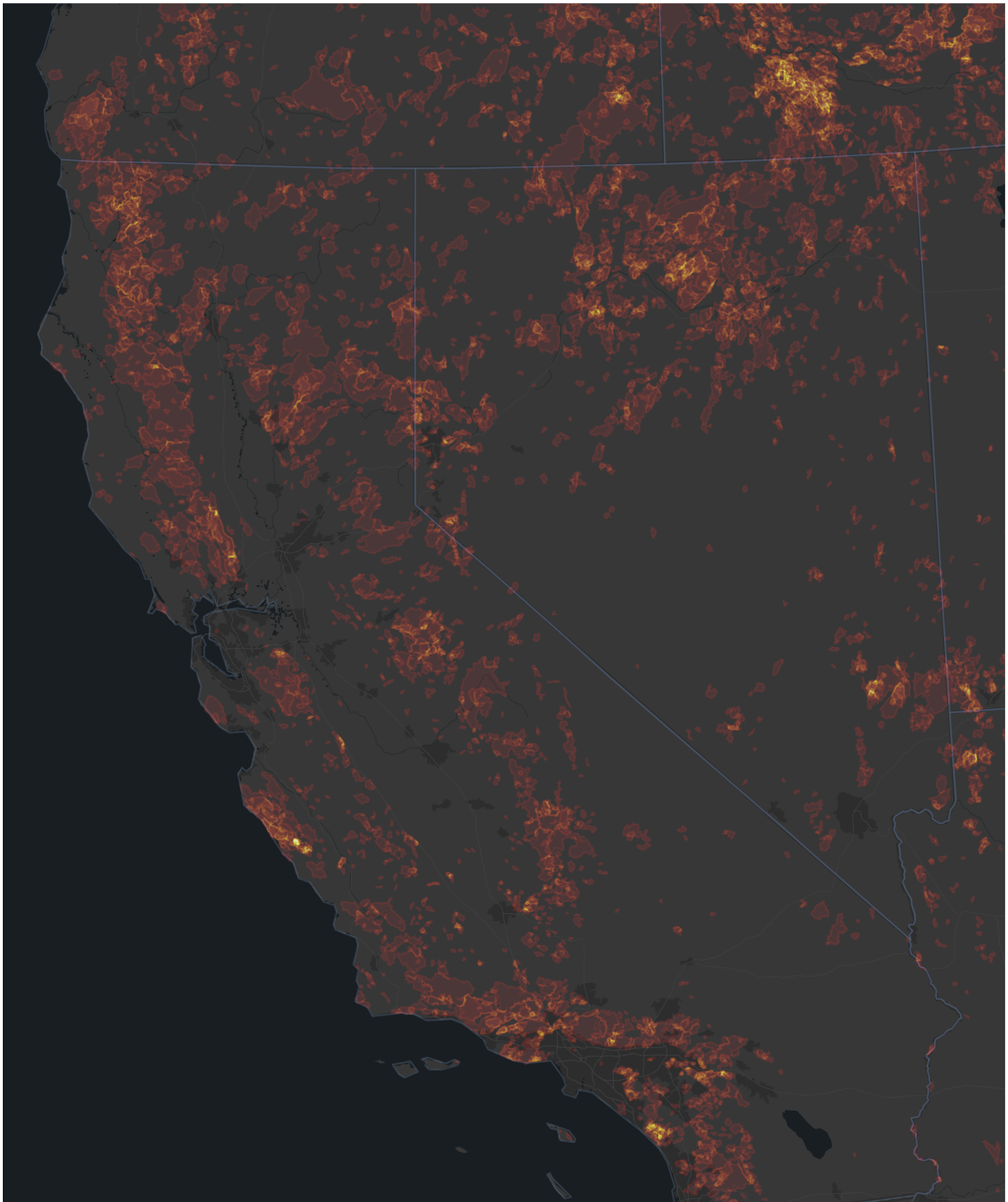


California Wildfires

Exploring the geospatial characteristics of California wildfires between 1984 and 2024

Nicole Mizrahi | LAPC GIS 040 | May 2025
May 30, 2025

Introduction

New plant growth after a fire.

For millions of years, fire has been a vital element in sustaining Earth's biological life. A variety of plants, animals, and insects rely on the natural phenomenon of fire to maintain an ecological balance and provide the necessary components for long-term survival. Fires burn away old vegetation and overgrowth that would otherwise fuel more intense flames, fertilizing the soil with nutrients from the ash. Some plants and insects depend on fire to promote regrowth by eliminating their competition, releasing seeds, clearing brush and debris to increase solar exposure, and killing off destructive pests and plant diseases. Early humans discovered how to control fire and utilize it for light, warmth, protection, and a new method of cooking, which scientists believe contributed to the growth of our brains and a significant improvement in our cognitive abilities. Much of modern society would be unable to operate without the extensive use of fire.

A large smoke cloud rising into the atmosphere.

Although life evolved alongside it, rapid and uncontrolled fires remain a significant danger to the environment. Experts say, "Wildland fires are a force of nature that can be nearly as impossible to prevent, and as difficult to control, as hurricanes, tornadoes, and floods. Wildland fire can be a friend and a foe. In the right place at the right time, wildland fires can create many environmental benefits… In the wrong place at the wrong time, wildfires can wreak havoc, threatening lives, homes, communities, and natural and cultural resources." (USDA) Smoke and ash from large fires can severely impact air quality for hundreds of miles and contaminate water with toxic materials, as well as release considerable amounts of carbon dioxide and other greenhouse gases into the atmosphere. With steadily increasing global temperatures and shifting weather patterns, the risk of more extreme wildfire events is also on the rise.

The United States Department of Agriculture (USDA), Forest Service provides incredible resources and valuable datasets that describe many environmental issues nationwide, including a comprehensive collection of fire event data from the Monitoring Trends in Burn Severity (MTBS) Program. The *MTBS Fire Occurrence Points* and *MTBS Burn Area Boundary* datasets contain information on over 30,000 large wildland fires throughout the contiguous United States, Alaska, Hawaii, and Puerto Rico, occurring between 1984 and 2024. Large fires are defined as burning 1000+ acres in the western U.S., 500+ acres in the eastern part of the country, and are categorized into four types: Prescribed Fire, Wildfire, Wildland Fire Use, and Unknown. Given the inherent danger and unpredictability of wildfires, I will assess only this type of fire from the MTBS datasets.

The burn areas have a faint yellow outline to highlight where multiple fires overlap. The brighter areas represent where wildfires have repeatedly burned.

The datasets contain approximately 16,000 U.S. wildfires, which make up over half of the most extreme fires recorded in the last 40 years.

To evaluate ***the geospatial characteristics of U.S. wildfires***, let's start by exploring ***California***.

Where were the wildfires located?

The first step in a spatial analysis is to define the study area. Our area of interest is the state of California and the wildfires that occurred between 1984 and 2024.

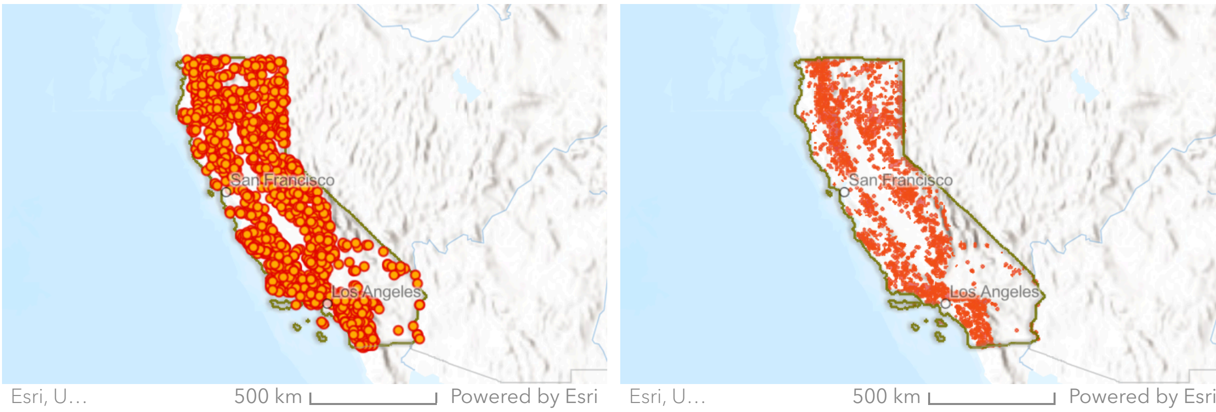
I've used ArcGIS Pro's ***Select Layer By Attribute*** in the Data Management Tools to identify only those fires in the MTBS data where the fire type is labeled as "Wildfire." This includes all of the U.S. wildfires shown on the map above. At the time of access, the MTBS datasets included wildfire information for January 1984 to July 2024. Any large fires occurring during the last half of 2024 are not included in this study.

Next, I applied ***Clip*** in the Analysis Tools to isolate only the wildfires that fall within the California state boundary. The output features from the clipped *MTBS Fire Occurrence Points* and *MTBS Burn Area Boundary* datasets were projected using the NAD 1983 California (Teale) Albers (US Feet) coordinate system. This system uses the Albers equal area conic projection, which is ideal for measuring area on mid-latitude land masses.

Visualizing California's wildfires as both point locations and burn areas provides two distinct yet valuable perspectives on their spatial distribution.

Between January 1984 and July 2024, California experienced over 1,800 wildfires that burned approximately 25,000,000 acres.

At first glance, there appears to be a pattern in the wildfire locations, especially along the mountain ranges that surround California's Central Valley.



Swipe the arrow on the interactive map to compare the fire point locations with their corresponding burn areas. Click on an object to see details about that fire.

Mountains play a major role in how quickly a fire can spread. When flames are moving uphill, the rising heat dries out the vegetation above it, making it easier to ignite. Mountain slopes can also push wind downhill, which causes air to increase in speed and temperature as it falls. Stronger winds can carry embers much farther, and erratic gusts can send them in unexpected directions. The combination of slopes, winds, hot temperatures, and dry forest vegetation is what makes mountain forest fires so intense.

A diagram of how mountains can affect wind. This example explains the formation of the Santa Ana winds.

A spatial analysis that plots locations is interesting, but it does not provide the whole picture. To gain a better understanding of California wildfires, we need to examine how they vary with time.

When did the wildfires occur?

Exploring data with different charts and tables can provide important insights into any temporal patterns that are not easily discernible from a location-based reference map.

Bar Charts:

First, let's look at a few bar charts using the *MTBS Burn Area Boundary* data. This dataset already had the month and year for each fire separated into individual fields, allowing for a quick breakdown of their distributions.

A bar chart showing the number of California wildfire events, separated by the month each fire started.

A bar chart showing the sum of acres that burned in California wildfires, separated by the month each fire started.

The distributions for both the monthly number of fires and the monthly sum of burned acres are approximately normally distributed, with the highest values occurring during the middle of the year. These months correspond to California's summer and fall seasons, which tend to have more extreme weather conditions.

A bar chart showing the number of California wildfire events, separated by the year each fire occurred.

Unlike the monthly totals, the annual number of fires exhibits a multimodal distribution, showing several high peaks across the time period. This pattern could be linked to the El Niño-Southern Oscillation (ENSO) cycle, which recurs every 2 to 7 years and significantly influences California's weather.

A bar chart showing the sum of acres that burned in California wildfires, separated by the year each fire occurred.

For better readability, the yearly totals are presented in the following table:

A data table showing the sum of acres that burned in California wildfires, separated by the year each fire occurred.

This distribution is heavily skewed, with a massive increase in the number of acres that burned in both 2020 and 2021. This distinct difference is a clear sign to take a closer look at the data.

A screenshot of the attribute table from ArcGIS Pro
showing the top 10 largest wildfires in descending order.

The attribute table shows that half of the top ten largest wildfires in California occurred in 2020, including the largest single wildfire in U.S. recorded history, the August Complex fire. Interestingly, 80% of these 2020 wildfires occurred in August.

Data Clocks:

Another way of viewing temporal distributions is with data clocks. Here we'll use the *MTBS Fire Occurrence Points* because it has the proper date format for this chart tool. Each ring in the data clock represents a year, and each wedge corresponds to a month. To provide a snapshot of the data's range, the sum totals are presented in logical divisions and distinguished by color intensity. In this design, we can quickly identify which blocks of time had the least and most extreme values, and if there are any cyclical patterns.

Data clocks that represent how the number of fires and sum
of burnt acres varied through different months and years.

Matrix Heat Charts:

Matrix heat charts are very similar to data clocks, except they present the data in a grid instead of a circle. Both the matrix heat charts and the data clocks emphasize the concentration of extreme wildfire activity during the middle of the year.

Matrix heat charts that show how the data varies throughout the months and years in our study window.

Another useful function of the charts on ArcGIS Pro is their ability to quickly select all of the data points within a certain time block. The example below shows how both charts can be used to select the wildfires that occurred in August of 2020.

With the selection outlined in teal, it is easier to see how these graphs correspond to each other and the location map on the left.

From this initial data exploration using a mix of reference maps, charts, and tables, we have seen some patterns emerge in space, time, and magnitude of California's wildfires. Now we will use geospatial statistics to see if there are any significant hot spots, cold spots, or outliers in the clustering of wildfire occurrence locations and their burn size.

What is the *spatial significance* of the wildfires?

Heat Maps:

A series of what heat maps can reveal. The map on the left plots the point locations. The middle map shows the density of location points only. The map on the right has been weighted for acres.

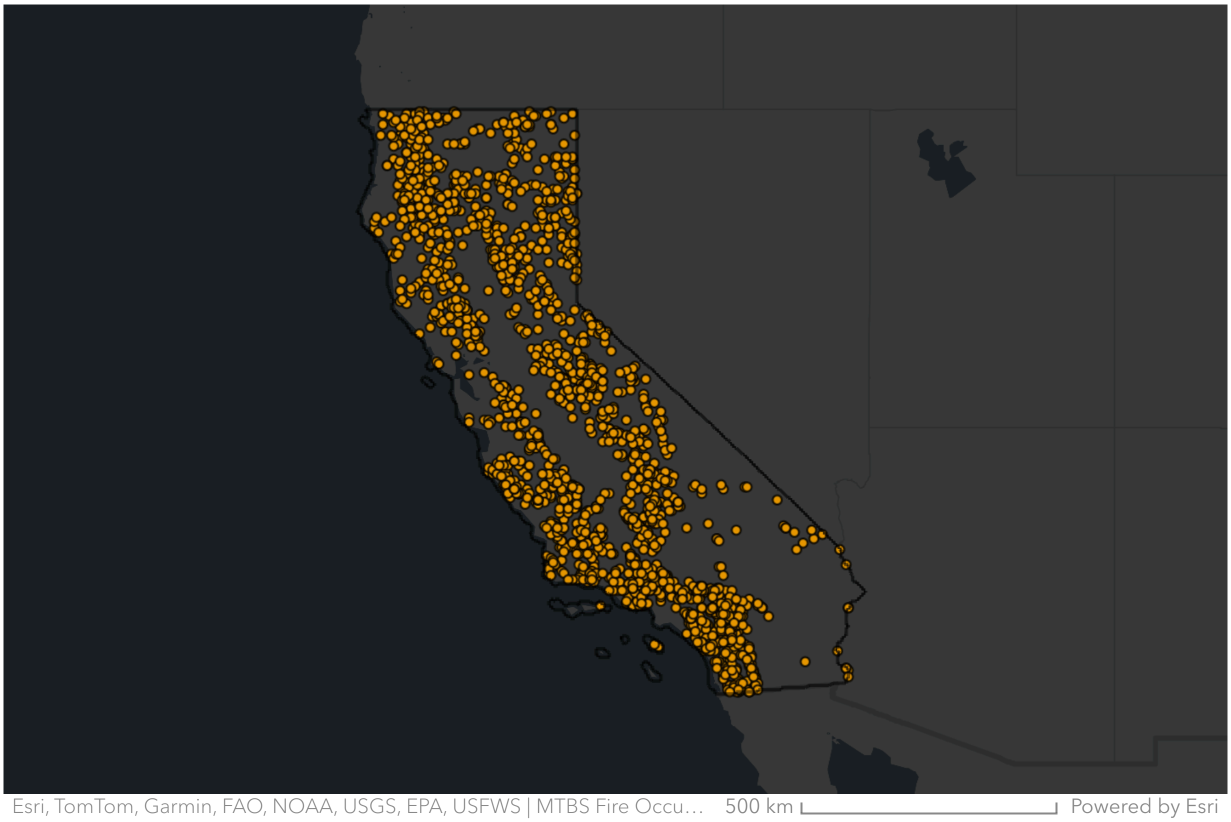
These heat maps reveal that there is a difference between the density of wildfire locations, and the concentration of fires that burned a higher number of acres.

Optimized Hot Spot Analysis

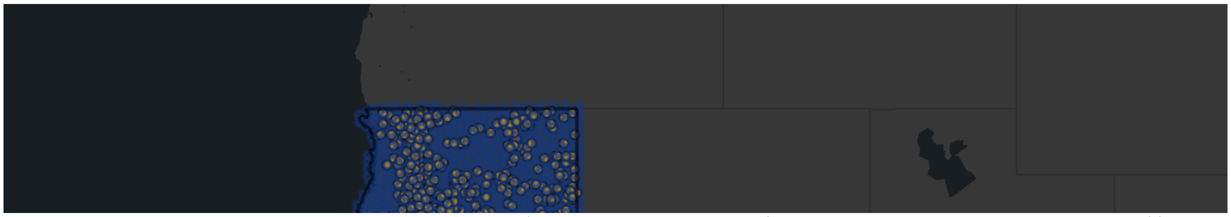
The ***Optimized Hot Spot Analysis*** from the Spatial Statistics Tools uses the Getis-Ord Gi* statistic to measure spatial autocorrelation, or more specifically, the degree to which features with similar values are clustered together when compared to the whole study area. This analysis tool identifies statistically significant hot spots (clusters of high values) and cold spots (clusters of low values) that are unlikely to have happened just by random chance.

Optimized Hot Spot Analysis is most effective on point data that is aggregated into polygons because it detects a more meaningful and relevant spatial pattern rather than just identifying a cluster of individual points. Aggregating the California wildfire occurrence points into a hexagon grid allows the tool to find the hot and cold clusters of both the density of fire events and the weighted sum of acres burned.

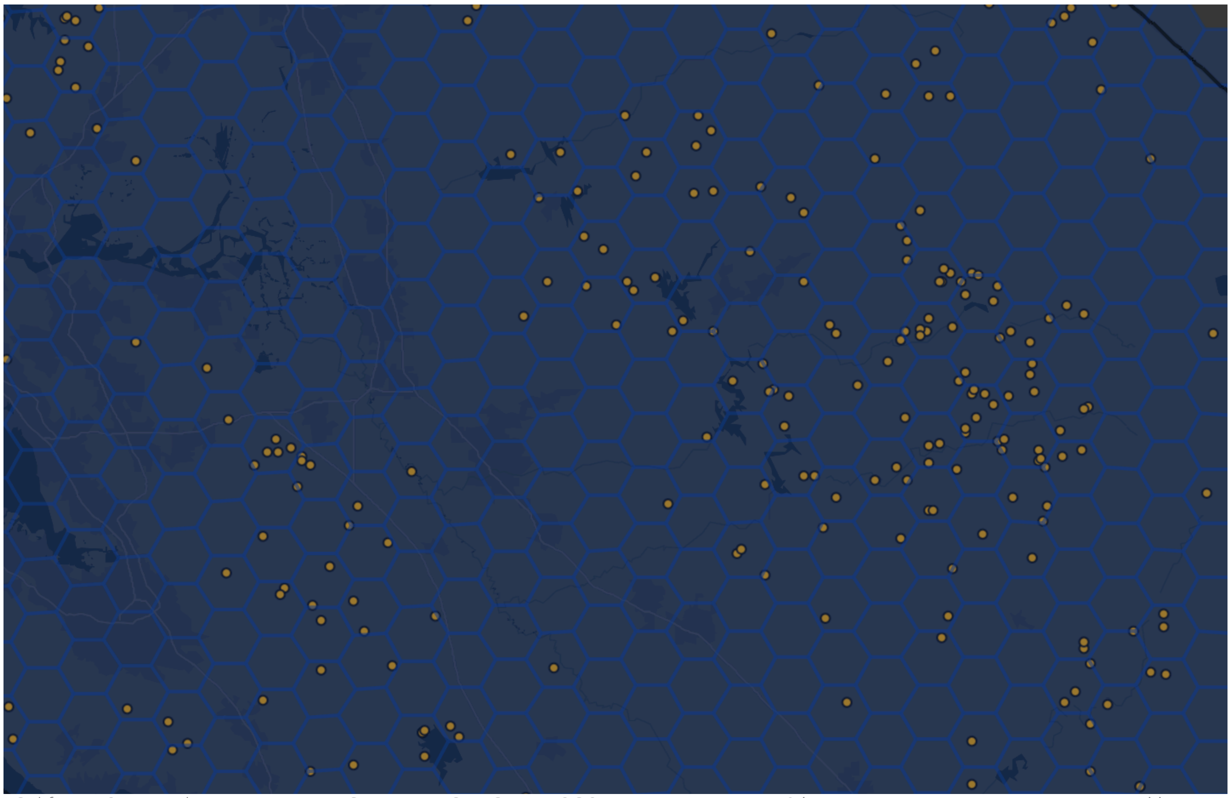
The step-by-step process of aggregation and running the Optimized Hot Spot Analysis is detailed in the map below. If necessary, scroll through the map slides slowly to allow the data to fully load.



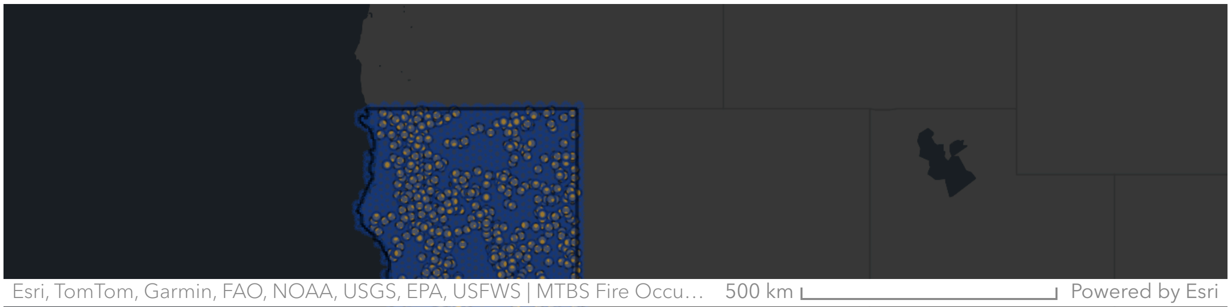
Step 1: Add the California Wildfire Points. These individual locations will be grouped along a uniform grid to get the most meaningful result from the geostatistical analysis.



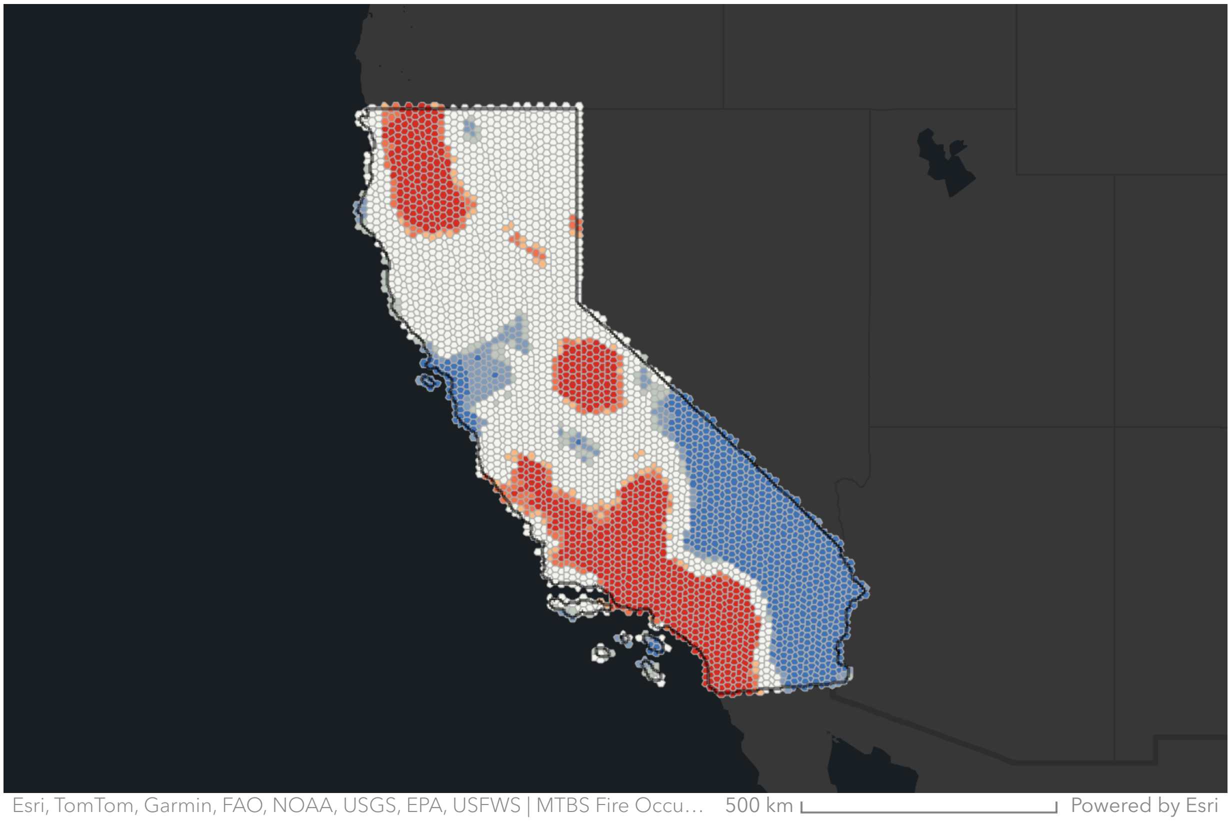
Step 2: Use the ***Generate Grids and Hexagons*** in the Business Analyst Tools. Each hexagon is called a bin with a cell size of 50 square miles and extends to the California state boundary.



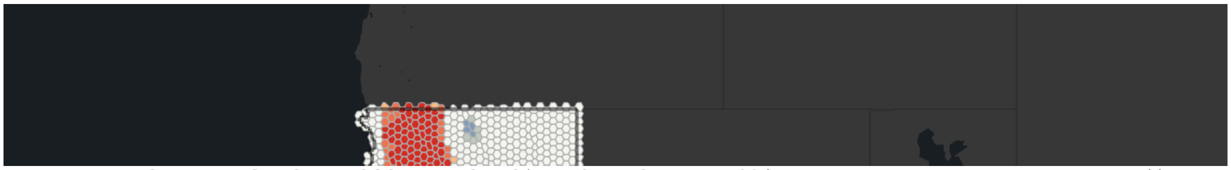
Step 3: Apply ***Summarize Within*** from the Analysis Tools. This will aggregate the wildfire locations into their corresponding hexagon bin, and then summarize the data points by counting the number of fires and the sum of their burned acres for each bin. Keep all empty bins because the lack of fire in those locations is still a meaningful measurement.



The aggregated hexagon grid of CA wildfire points is now ready for geostatistical testing.

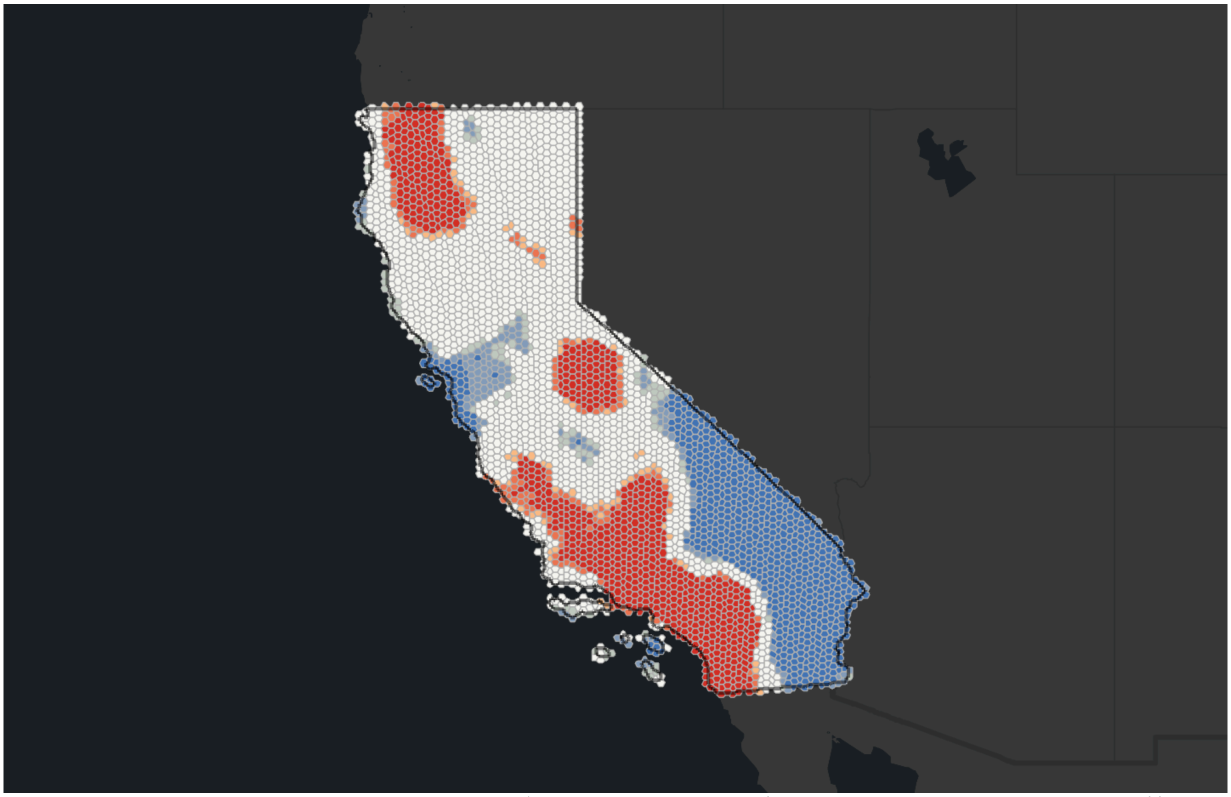


Step 4: The first run of this test looked at statistically significant hot and cold spots for the number of fire occurrence points in each local neighborhood.



Esri, TomTom, Garmin, FAO, NOAA, USGS, EPA, USFWS | MTBS Fire Occu… 500 km Powered by Esri

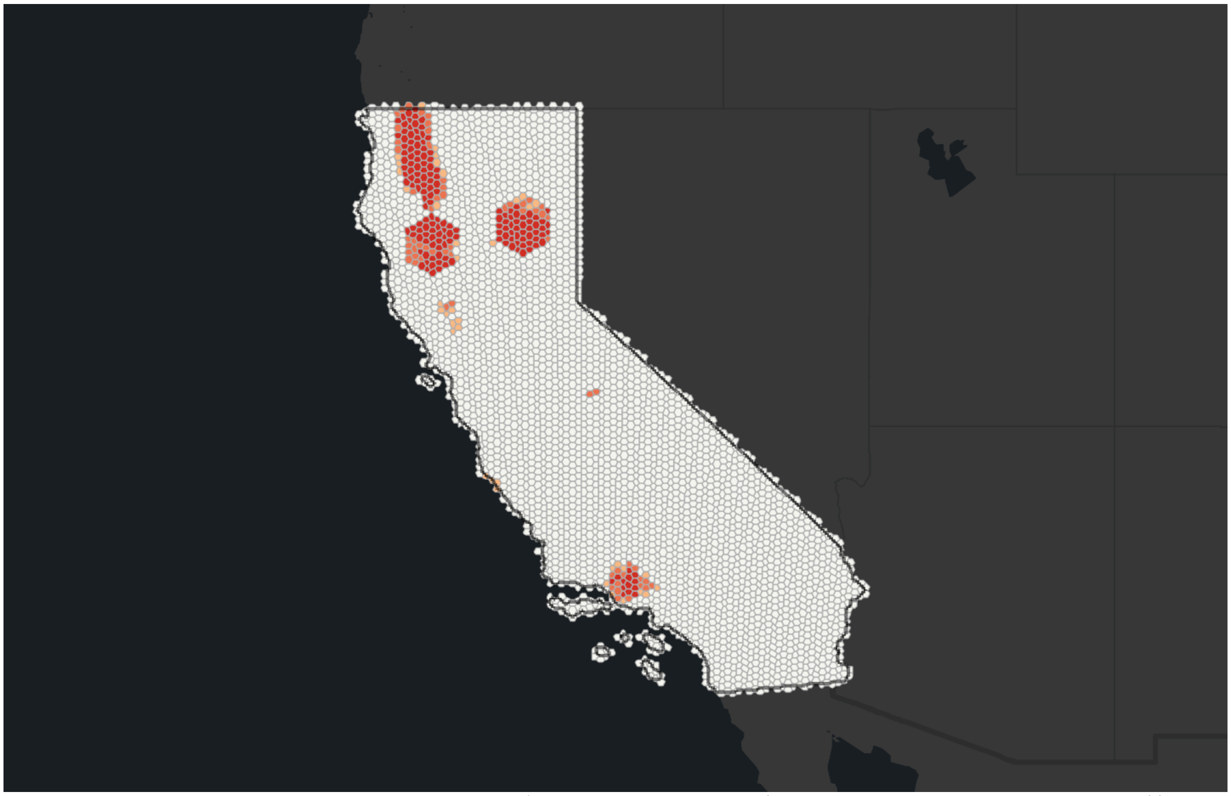
The red areas were found to have statistically significant high values of wildfires occurring in those neighborhoods. These areas are considered "Hot Spots" for wildfire occurrences, and are very unlikely to have clustered together just by random chance.



Esri, TomTom, Garmin, FAO, NOAA, USGS, EPA, USFWS | MTBS Fire Occu… 500 km Powered by Esri

The white areas were found to have no statistically significant difference from what would be expected if the values were randomly distributed across the entire study area.

The blue areas were found to have statistically significant low values of wildfires occurring in those neighborhoods. These areas are considered "Cold Spots" for wildfire occurrences, and are very unlikely to have clustered together just by random chance.



Step 5: The second run of this test looked at statistically significant hot and cold spots for the sum of acres burned in each local neighborhood.

The results of both versions of the Optimized Hot Spot Analysis were somewhat similar to the density heat maps, but with an added statistical level of detail. These statistically significant areas offer deeper insights into the geospatial characteristics of California wildfires, potentially benefiting urban planning, local communities, and wildlife conservation efforts.

Conclusion

Using the suite of visual and analysis tools available through ArcGIS Pro, our study on the geospatial characteristics of California wildfires has uncovered a wealth of interesting information. The spatial and temporal patterns that were revealed through the combination of reference maps, charts, tables, density heat maps, and hot spot analysis has shown the areas that have been most affected by wildfires in the past and still remain at high risk for future large-scale fire events. The opportunities for further research into the wildland-urban interface, climate and weather patterns, topography, plant life, and many other critical components that influence wildfire activity around the globe are seemingly limitless. Hopefully this study will inspire others to learn more about the geospatial characteristics of wildfire, and perhaps eventually we can discover a way to prevent more extreme fire events from occurring.

Credits

USDA, Forest Service: MTBS Burn Area Boundary; MTBS Fire Occurrence Points	https://data.fs.usda.gov/geodata/edw/datasets.php
CA Geographic Boundaries	https://catalog.data.gov/dataset/ca-geographic-boundaries
Monitoring Trends in Burn Severity (MTBS)	https://www.mtbs.gov/
Esri ArcGIS Pro Documentation	https://pro.arcgis.com/en/pro-app/3.3/tool-reference/main/arcgis-pro-tool-reference.htm
Esri Spatial Statistics Resources	https://spatialstats-analysis-1.hub.arcgis.com/
CAL FIRE	https://www.fire.ca.gov/our-impact/statistics
National Park Service	https://www.nps.gov/subjects/fire/fire-in-depth.htm
Media Photographs	https://unsplash.com/s/photos/fire
CNY Central	https://cnycentral.com/weather/weather-wisdom/three-reasons-the-california-wildfires-are-so-severe-plus-explaining-santa-ana-winds