Are California Wildfires Getting more Intense in the North or South?

true

11-29-2021

Contents

Motivate the question / Introduction

Growing up in the Antelope Valley, located in the Mojave Desert at the northern edge of LA County, I experienced wildfires at a very young age. As I grew up, orange haze outside, flying ash, and evacuation announcements on the news were not out of the ordinary during the later summer months. I will never forget taking a day trip to the beach with my family one summer only to come home to barricades outside our neighborhood because a wildfire had jumped the nearby ridge deeming it unsafe to enter. While our house, and my dog inside, was thankfully intact, we were reminded of the event for months by the charred hill next to our house.

The statistics describing the frequency, intensity, and destructiveness of California wildfires are alarming to say the least, and are all likely to be further exacerbated by the climate crisis in the years to come. As temperature and droughts increase with elevated greenhouse gas emissions, it is probable that we can expect more wildfires in the future, especially with the fire seasons getting longer.

In the past couple decades, California's wildfire season has progressively gotten worse. A recent study by researchers at UC Irvine found that the annual burn season has lengthened since 2000 and that the yearly peak has shifted from August to July. In recent years especially, fires seem to be getting both more intense and more frequent. For example, the 2017 season was the most destructive wildfire season on record at the time, with 1.2 million acres burned. The largest fire during that season was the Thomas Fire in Santa Barbara County, which was California's largest modern wildfire at the time. In 2020, 4.2 million acres were burned, amounting to more than 4% of the state's total land. This resulted in 2020 being the largest wildfire season recorded in California's modern history. The August 2020 Complex Fire alone burned more than 1 million acres, making it the first "gigafire" on record.

Each year, I hear and read about wildfires all over California. Just this year, I remember talking with some friends about hazy air quality in Santa Barbara. We were wondering if it was due to smoke from fires in the north or the south because there were many wildfires burning at the time.

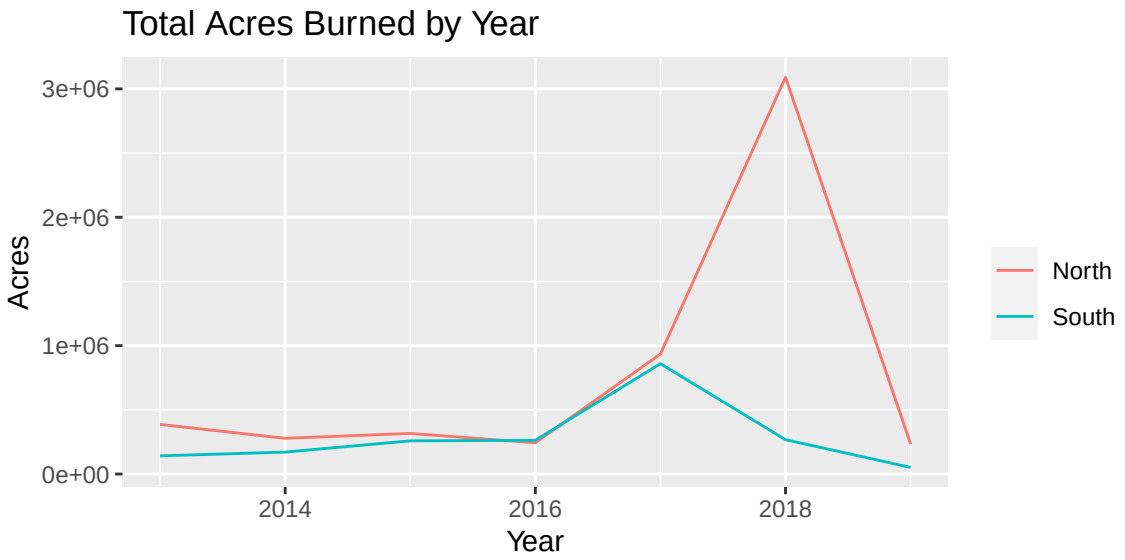


Figure 1: Total Acres Burned (2013-2019)

I am interested in how California wildfires have changed over time, and if this change is different in the north than in the south. Are California wildfires getting more intense in the northern part of the state or in the south? Is there a difference in trends between start date and acres burned for fires in the north versus fires in the South?

From my preliminary research, there does not seem to be existing evidence on this question. While many articles discuss trends in wildfire season as a whole and factors that contribute to the growing wildfires in California, there is not much to be said about how geographic location is potentially correlated with changes in wildfires over time.

Data Description

To explore changes in California's wildfire season and compare trends in the North to the South, I will be using a dataset which contains over 1600 wildfire events in the state between 2013 and 2020. This dataset is made available on Kaggle.com and was originally scraped from records on the CAL FIRE website. For each fire, this dataset contains 40 variables; for the purposes of this analysis, the most relevant variables include latitude and longitude, start date, extinguished date, and acres burned. A potential limitation of this dataset is that it only includes wildfires responded to by CAL FIRE; many wildfire events that contribute to the overall trends in the north and the south may not be included in this dataset.

Analysis Plan

To assess the effects of wildfire start date and location on the number of acres burned, I will be using the following interaction model:

$$acres\_burned_i = \beta_0 + \beta_1 \cdot started_i + \beta_2 \cdot north\_south_i + \beta_3 \cdot started_i \cdot north\_south_i + \varepsilon_i$$

The variable `north_south` takes on a value of "North" for all wildfires occurring above 36 degrees latitude, and everything below this is assigned "South". I have chosen this interaction model in order to see if wildfire trends differ in the north and south.

After running this regression, I will determine if the slope coefficients are statistically significant by looking at the associated p-values.

Results

From the regression output in Table 1, we can see that the intercept for fires that occur in the South is 24,345 acres above the intercept for fires in the South. The slope of the regression line for fires in the South has a slope that is 1.56 less than that of the regression line for fires in the North. In other words, the trend line for northern fires is steeper.

Table 1: Multiple Linear Regression Output

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-21405.07	22705.70	-0.94	0.35
started	1.57	1.31	1.19	0.23
north_southSouth	24344.65	34590.61	0.70	0.48
started:north_southSouth	-1.56	2.01	-0.78	0.44

While this model does show a difference in slopes, because the p-values are so high, these results are not statistically significant. We do not have sound evidence that there is a difference in wildfire intensity over time between the North and South. This brief analysis does not help to answer the original question. More data would be needed to perform a more robust and comprehensive analysis to accurately answer the question "Are California wildfires getting more intense in the northern part of the state or in the south?"

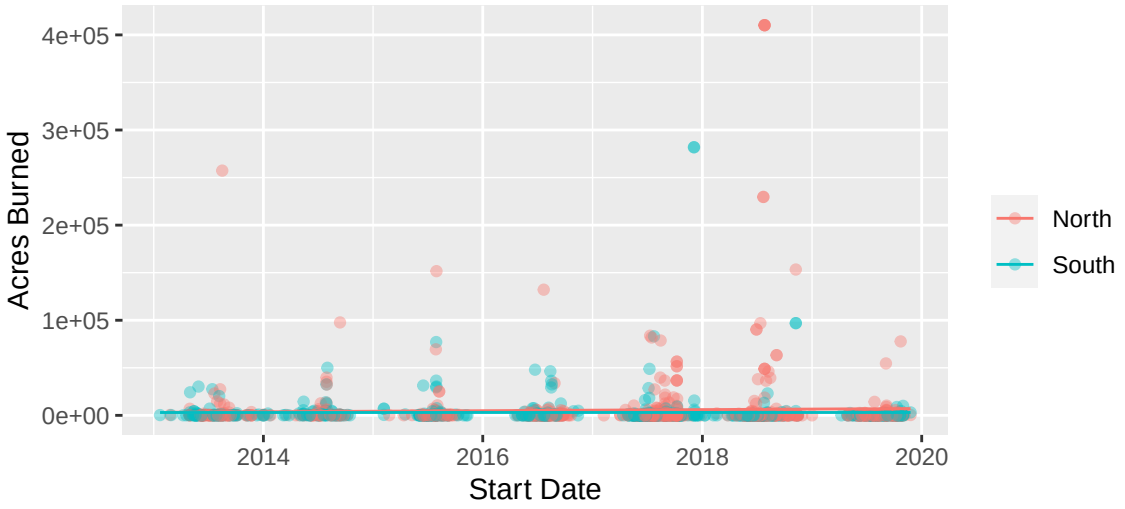


Figure 2: Acres Burned vs. Start Date

Further Analysis

- more years of fire data
- 7 years is not enough for a robust analysis of the question posed
- 2020 fire season is not included
- time series analysis/predictions

In order to accurately answer this question, much more data is needed. This dataset only includes wildfire events during a seven year span. To get a more thorough picture of the trends in acres burned for northern and southern wildfires in California over time, a much larger temporal range is needed beyond the 7 years provided in this dataset. Additionally, this dataset does not include 2020 data which means observations from the largest fire season on record are not taken into account in this preliminary analysis. With more data points spanning a 20-30 year period, I suspect much more interesting and informative results.

Beyond doing a regression analysis to explore differences in geographic trends in wildfire intensity over time, doing time series analyses to produce predictions of wildfire intensity would also be useful. This type of analysis would be of interest to government organizations, first responders in California, and the general public living in fire-prone areas of California.

Lastly, looking at seasonal trends over time can help identify if and how the wildfire season itself is changing. This type of analysis could help answer questions such as: Is the wildfire season in California getting longer? Are more intense wildfires happening at a certain point in the season?

References

California Department of Forestry and Fire Protection (CAL FIRE). "2018 Incident Archive." Cal Fire Department of Forestry and Fire Protection, https://www.fire.ca.gov/incidents/2018/.

California Department of Forestry and Fire Protection (CAL FIRE). "2020 Incident Archive." Cal Fire Department of Forestry and Fire Protection, https://www.fire.ca.gov/incidents/2020/.

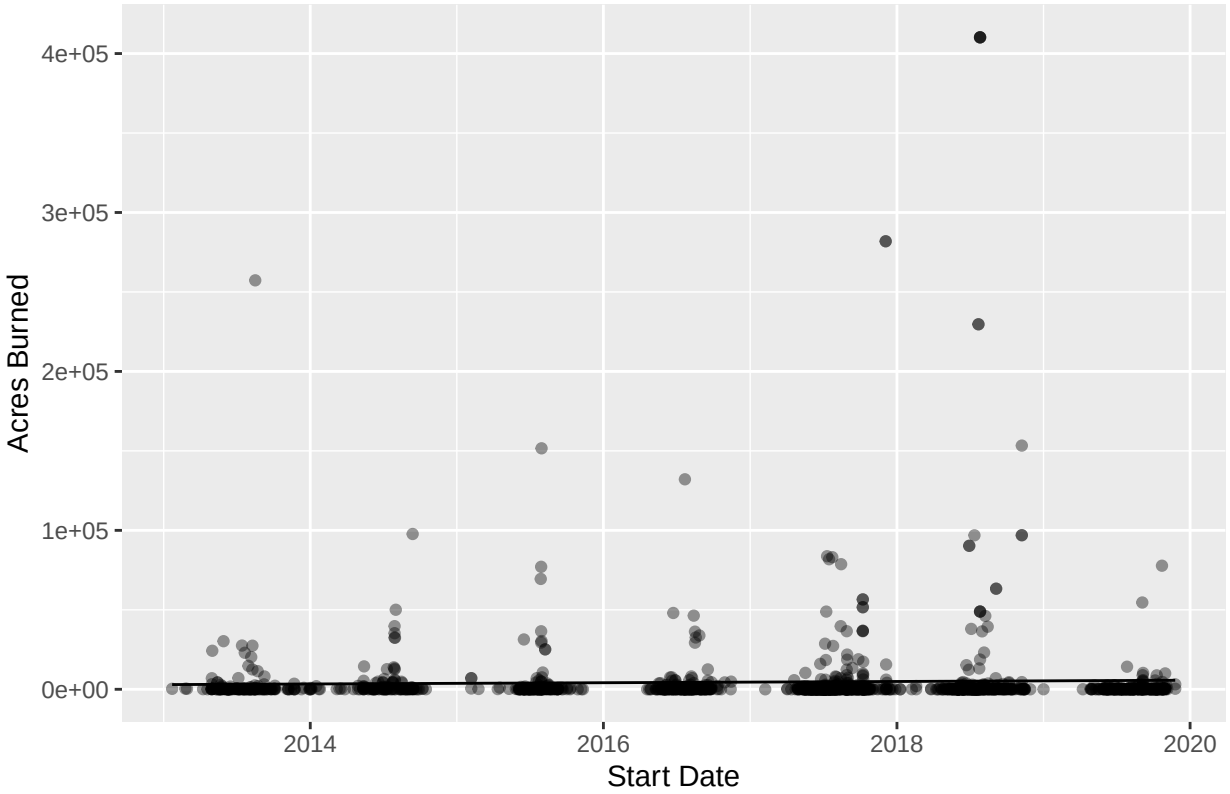
"One Reason for Northern California's Terrible Fire Season: Less Rain than Southern California." Los Angeles Times, Los Angeles Times, 1 Oct. 2020, https://www.latimes.com/california/story/2020-10-01/northern-california-fire-season-less-rain-than-southern-california.

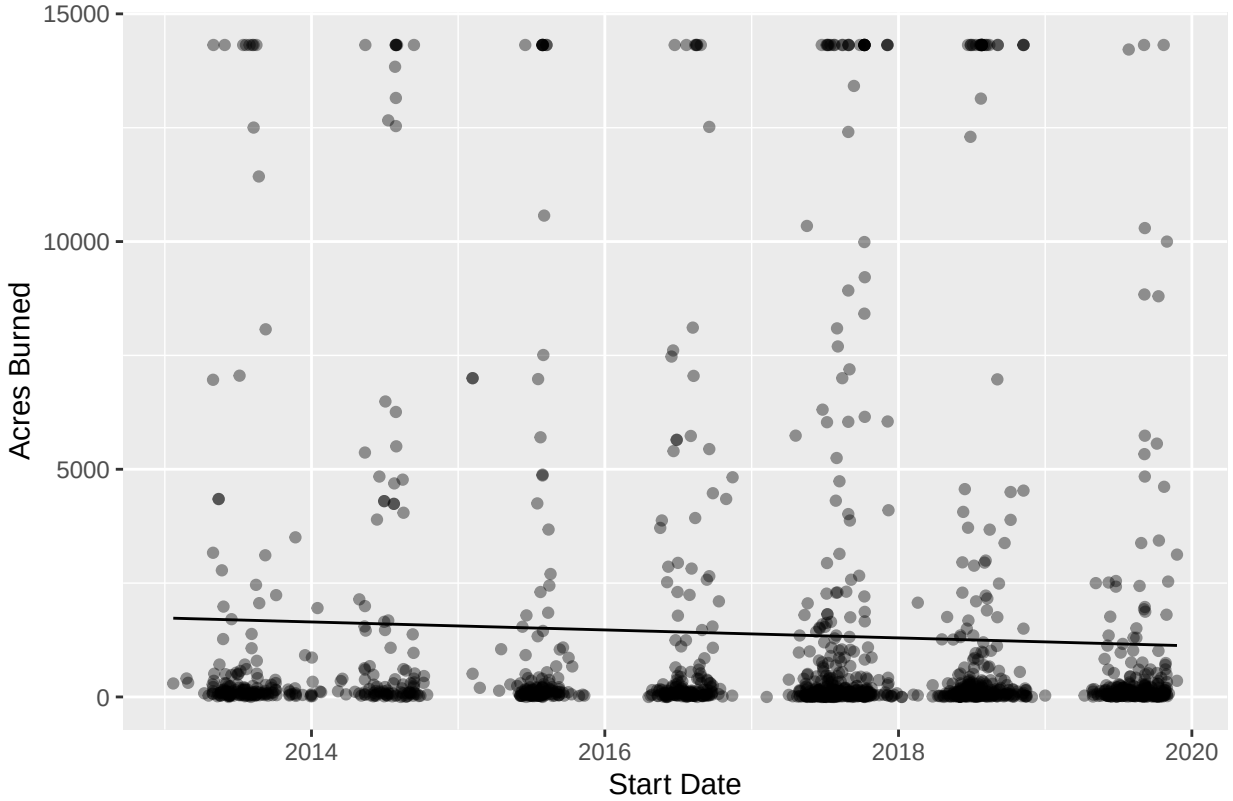
"California's Wildfire Season Has Lengthened, and Its Peak Is Now Earlier in the Year." (GSTDTAP): California's Wildfire Season Has Lengthened, and Its Peak Is Now Earlier in the Year, 22 Apr. 2021, http://resp.llas.ac.cn/C666/handle/2XK7JSWQ/323227.

First/general regression

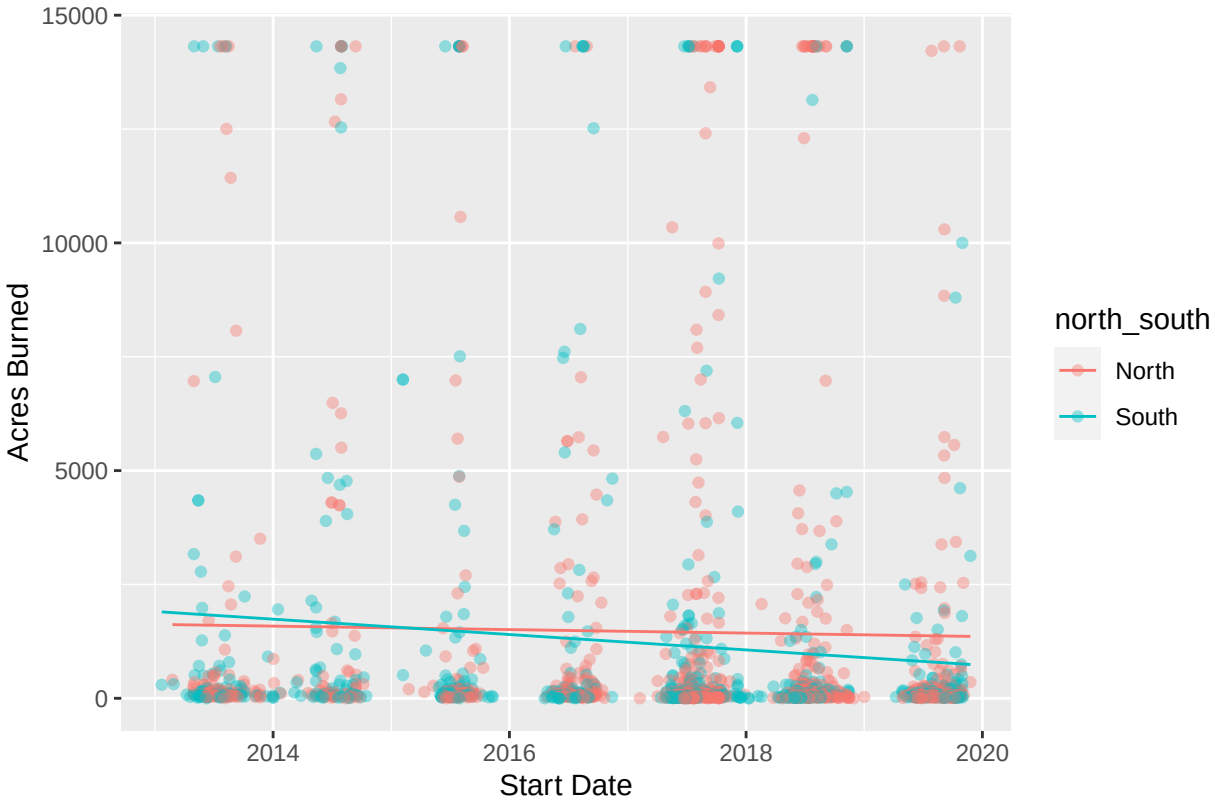
Table 2: Linear Regression Output

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-13959.94	17060.52	-0.82	0.41
started	1.08	0.99	1.09	0.28





Winsorization


15000
10000
5000
0
Acres Burned
2014
2016
2018
2020
Start Date
north_south
North
South