CESifo WORKING PAPERS

10679
2023

September 2023

Forest Fires: Why the Large Year-to-Year Variation in Forests Burned?

Jay Apt, Dennis Epple, Fallaw Sowell

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10679*

Forest Fires: Why the Large Year-to-Year Variation in Forests Burned?

Abstract

Quantifying factors giving rise to temporal variation in forest fires is important for advancing scientific understanding and improving fire prevention. We demonstrate that eighty percent of the large year-to-year variation in forest area burned in California can be accounted for by variation in temperature, precipitation, housing construction, electricity transmission, and ocean surface temperatures in the North Atlantic, North Pacific, and Equatorial Pacific. California is of particular interest because of its large acreage burned and proximity of fires to human populations. We believe our model is the first unified treatment of climatic factors and human activities that affect forest area burned.

JEL-Codes: H, Q200, Q500.

Keywords: forest fires, climate, human activities, ocean surface temperatures.

Jay Apt
Carnegie Mellon University
Pittsburgh / PA / USA
apt@cmu.edu

Dennis Epple
Carnegie Mellon University
Pittsburgh / PA / USA
epple@cmu.edu

Fallaw Sowell
Carnegie Mellon University
Pittsburgh / PA / USA
fs0v@andrew.cmu.edu

Forest Fires:
Why The Large Year-to-Year Variation in Forests Burned?

Jay Apt Dennis Epple Fallaw Sowell
Carnegie Mellon University

September 2023

Contents

1 Abstract

Quantifying factors giving rise to temporal variation in forest fires is important for advancing scientific understanding and improving fire prevention. We demonstrate that eighty percent of the large year-to-year variation in forest area burned in California can be accounted for by variation in temperature, precipitation, housing construction, electricity transmission, and ocean surface temperatures in the North Atlantic, North Pacific, and Equatorial Pacific. California is of particular interest because of its large acreage burned and proximity of fires to human populations. We believe our model is the first unified treatment of climatic factors and human activities that affect forest area burned.

2 Introduction[1]

Fires destroy millions of acres of forest annually in the US and around the world.[2] The economic impacts are extensive, including damage to human health from air pollution, destruction of forests and homes, and loss of wildlife. As manifest by recent fires in North America, air pollution can reach far beyond the areas where fires are burning, including reaching across national boundaries. Moeltner, et. al. (2013) study the effect of Nevada forest fires on hospital admissions, finding that patient counts can be causally linked to fires as far as 200–300 miles from the impact area. We provide a unified treatment of factors impacting the amount of forest area burned annually in California over a period of 36 years. California is of particular interest because of its large acreage burned and proximity of area burned to human populations. California has a large forested area, 33 million acres, comprising one third of the state's total land area. That fraction is broadly representative of the United States as a whole; 34% of US land area is forested (World Bank, 2022).

Over the period of our sample, forest area burned in California averaged 13% of total acreage burned nationwide. Over the last decade, the average percentage has been 14%, with individual years as low as 1% and as high as 41% (US National Interagency Fire Center, 2022). There is much year-to-year variation in both climate and local human activities in California. This permits relatively precise estimation of the parameters of the model we develop. As we noted in our abstract, we believe that the model we develop is the first to provide a unified treatment of climatic factors and human activities. Also, we find significant evidence of a structural break between years 2020 and 2021. We discuss the potential implications of this break in the presentation of our findings.

3 Literature Review

There is a rich literature on the predictors of forest fires. At millennial time scales, changes in insolation were found to be the main driver of changes in the number and extent of fires during the past 17,000 years in Yellowstone National Park (Millspaugh, Whitlock, and Bartlein, 2000). More forest fires were recorded during the warmer and drier early Holocene than at the present day. The availability of fire data processed from the Terra satellite Moderate Resolution Imaging Spectroradiometer (MODIS)[3] has allowed several groups to examine changes in fire extent and frequency in recent times to examine the effect of temperature, rainfall, dry days, fuel status, and anthropogenic disturbances. For example, Kale, Ramachandran, and

[1] We are indebted to Marshall Burke and Sam Heft-Neal, whose work inspired ours. https://siepr.stanford.edu/publications/policy-brief/managing-growing-cost-wildfire. We thank them for guiding us to data for forest area burned and precipitation for California.

[2] The scourge of wildfires is truly worldwide. A BBC report on a recent wildfire in Greece was headlined: "Inside the horror of Europe's biggest wildfire". https://www.bbc.co.uk/news/extra/ifgej14zt1/greece-wildfire

[3] MODIS Active Fire and Burned Area Products, https://modis-fire.umd.edu/

Pardeshi (2017) found that in India fires are correlated with temperature, including higher temperatures driven by El Niño. The number of other types of fires have also been shown to be temperature-related. Using 7 years of recent data, Xu, Liu, Yan (2021) find that an increasingly warm climate is likely to very slightly increase the annual non-forest fire frequency over the next fifty years.

Guyette (2012) develops a physical chemistry-based model to predict the mean interval between forest fires using data prior to the year 1850. The significant variables in this model are annual mean maximum temperature, annual mean precipitation, and the partial pressure of oxygen as represented by elevation. He reports: "Although it may seem intuitive that fire frequency is relatively lower in landscapes where annual precipitation is high, this relationship is not well supported by data and literature. Increases in precipitation have a negative influence on fire frequency due to influences such as increased fuel moisture and relative humidity. In contrast, increased precipitation generally has a positive influence on fuel production."

This fuel production effect of precipitation has been examined by a number of researchers. In their study of an 11-year period in India, Kale, Ramachandran, and Pardeshi (2017) found that the large fire year of 2009 had both record heat and three contiguous prior years of heavy rain. In their study area, La Niña causes high rainfall, while El Niño causes reduced rainfall. They had two high fire years in their data set; both were El Niño years preceded by strong La Niña events. Quantitative analyses of the lagged relationship between rainfall in a given year and forest fire area burned or number of forest fires have been undertaken for particular locations. In India's Western Ghats, Quentin, et. al. (2012) found that the monsoon season prior to a fire year controlled fuel moisture content during the fire year. Forkel, et. al. (2012) found that 1-year lagged precipitation was a good predictor of burned area for particular types of forests in Siberia that rest on permafrost. Koutsias et. al. (2012) used an 1894-2010 data set for Greece, finding that 2-year lagged annual precipitation was positively correlated with fire risk (although the dominant rainfall effect was that increased precipitation decreases fire risk in that year by making fuel moist). Two studies, Pausas (2004) and Turco, et al. (2013), examined fire number and area burned in the Iberian Peninsula, both finding a positive correlation with 2-year lagged precipitation. Using a 29-year data set in southeastern Arizona, Crimmins and Comrie (2004) found a positive association between large fires at low elevation with 1-year lagged rainfall, and between large fires at high elevations with wet years up to three years prior to the fire year. Brooks and Matchett (2006) studied wildfires in the Mojave Desert (portions of which are in California, Nevada, and Arizona), finding that fire size increased for middle elevation shrublands in their 25-year study period in a year after a year of high rainfall, due to increased growth of annual grasses.

Winds can propagate wildfires. As Li, Paek, and Yu (2016) explain "The Santa Ana winds (SAW) are a weather phenomenon in Southern California that occur most often during late autumn to early spring. SAW events are characterized by high wind speeds, low relative humidity and high temperatures, and are well known for their ability to exacerbate fire conditions and aid in the spread of wildfire." Santa Ana Winds in turn are influenced by sea surface temperature and associated atmospheric pressure patterns. Measures of North Atlantic, North Pacific, and Equatorial Pacific sea surface temperatures are Atlantic Multi-decadal Oscillation (AMO), Pacific Decadal Oscillation (PDO) and El Niño–Southern Oscillation (ENSO). Using data for 1960-2010, Li, Paek, and Yu (2016) find that a regression of mean annual Santa Ana Wind Days (SAD) on mean annual PDO and mean annual AMO can account for 76% of the variation in SAD.[4] Cardil et al. (2021) investigate Atlantic and Pacific ocean influences on drought, Santa Ana winds and wildfires in southern California. As discussed above, there is evidence that weather conditions associated with El Niño impact forest fires in the Southern Hemisphere. The impact of El Niño on southern California has been a subject of some debate, with the debate centered on whether El Niño impacts precipitation in southern California and Baja, New Mexico (Minnich, Vizcaíno, and Dezzani, 2000).

Anthropogenic effects on wildfires have attracted the attention of a number of authors. Westerling et al.(2011) found that higher population density in California increased fire activity up to a point, but that in a few very densely-populated areas fire activity was rather low. Globally, Knorr et al. (2014) found that the effect of increasing population is mainly to decrease wildfire frequency, and that only for areas with more than about 0.1 people per km^2 does fire frequency increase (by 10-20%). Since many fires are sparked by humans in areas adjacent to wild areas, the wildland-urban interface (WUI) has been defined for two decades as

[4]They report that the correlation between SAD predicted from their regression and observed SAD is .87. The R-squared value, .76, is the square of the correlation.

an important study area (Glickman and Babbitt (2001). Radeloff et al. (2005) performed the first detailed national mapping of the WUI in the United States, noting that in the year 2000 almost 40% of all US housing units are located on the WUI. Price and Bradstock (2014) studies 38 years of data on the WUI near Sydney, Australia. They found that "fires are most common where un-vegetated land in urban areas has been little other than to accommodate the houses." Farmed areas reduce fuel, while homes in forest areas are correlated with people sparking abundant fuel. For California, Syphard et al. (2007) found that the number of fires per unit time, or per unit area, from 1960-2000 is related to population density, distance to the WUI, and vegetation type. They observe that humans caused roughly 95% of the number of fires in California.

The potential for electrical transmission lines to create sparks that ignite wildfires has been well know for many years. The regulatory authority in California with primary responsibility for wildfire prevention is the California Department of Forestry and Fire Protection, known by the shorthand CalFire. For many years, CalFire has promulgated and annually updated regulations and guidelines with the goal of reducing fires caused by power lines.[5]

In July 12, 2019, the California state government passed and signed into law Assembly Bill No. 1054. Page 3 of the bill states "Existing law requires each electrical corporation to annually prepare and submit a wildfire mitigation plan to the commission for review and approval. Existing law requires the commission to consider whether the cost of implementing an electrical corporation's plan is just and reasonable in the electrical corporation's general rate case. This bill would require the plan, in calendar year 2020 and thereafter, to cover at least a 3-year period." Section 2(b) specifies "The state's electrical corporations must invest in hardening of the state's electrical infrastructure and vegetation management to reduce the risk of catastrophic wildfire." and Section 2(g) requires "The first $5 billion in safety investments in the aggregate by the large electrical corporations must be made under this act without return on equity that would have otherwise been borne by ratepayers."

In addition to deaths, injuries, property damages, and harm to wildlife, fires create pollutants with adverse health effects that extend well beyond the perimeter of the burned area. These adverse health impacts are manifest across the life cycle. Burke et al. (2021) combine satellite-based fire and smoke data with information from pollution monitoring stations to investigate how changes in wildfire activity affect air pollution and related health outcomes. They develop a statistical model and demonstrate that fuel management interventions could have large beneficial health benefits. Rosales-Rueda and Triyana (2019) study the effects of early-life exposure to air pollution from Indonesian forest fires, finding delayed attainment of normal height of children exposed to the fires, lower lung capacity 10 years after exposure, and shorter stature of children who were exposed in utero at 10 and 17 years after exposure. Analyzing agricultural fires, Rangel and Vogl (2019) find that late-pregnancy smoke exposure decreases birth weight, gestational length, and in utero survival. They go on to say "Fires less associated with smoke exposure predict improved health, highlighting the importance of disentangling pollution from its economic correlates." Pullabhotla and Souza (2022) find significant evidence of the effect of upwind fires on hypertension. Zhang, et. al. (2023) demonstrate association between small airborne particulates and dementia. Bishop, Ketcham, and Kuminoff (2023) exploit changes in regulation by the Environmental Protection Administration to provide causal evidence that exposure to small airborne particulates increases the probability of dementia among the elderly. Forest fires are among prominent sources of small airborn particulates.[6]

4 Data Overview

As summarized above, an impressive body of research has identified factors that affect the amount of forest area burned and attendant damages. Because the factors that have been identified can act in combination, we develop a model to investigate the combined effects of environmental factors and human activities on the extent of forest area burned annually over the past 36 years in the US state of California.

[5]The importance attached to this effort is reflected in the signatories of CalFire publications. For example, the 2008 report carried the names of the governor, Arnold Schwarzenegger, The Secretary for Resources, Mike Chrisman, the Director of CalFire, Ruben Grijalva, and the state Fire Marshall, Kate Dargan. https://osfm.fire.ca.gov/media/8482/fppguidepdf126.pdf
For the 2021 edition, see Porter et. al. (2021)

[6]See also https://www.epa.gov/wildfire-smoke-course/why-wildfire-smoke-health-concern

Based on the evidence discussed in our literature review, we investigate the effects of the variables enumerated below. We anticipate positive algebraic signs for the first five variables. As explained below, we use stepwise regression to select the sea surface temperature variables that are included in the model.
1. Housing Starts: Housing construction entails a variety of activities that can inadvertently ignite forest fires, including transportation of materials, extension of power lines and roadways, operation of both gasoline and electrically powered equipment, and, in some instances, cigarette smoking.
2. Electricity Imports: Electrical wires rubbing against trees can create sparks that ignite wildfires.
3. Average Daily Maximum Temperatures: Higher temperatures increase the rate with which fires spread.
4. Lagged Precipitation: Precipitation promotes growth of weeds, small plants, and shrubs that die off during the winter, providing fuel for fires during the succeeding year.
5. Vapor Pressure Deficit: Dry air (high VPD) promotes burning.
6-8. Sea Surface Temperatures: Atlantic Multidecadal Oscillation (AMO), Pacific Decadal Oscillation (PDO), and El Niño Southern Oscillation (ENSO) are known to impact climate in the western United States: We use stepwise regression to estimate the timing, algebraic signs, and magnitudes of their impacts on forest area burned.

4.1 Notation

Data for forest acreage burned annually in California have been collected on a consistent basis dating back to 1987. Data for all of our explanatory variables are available over that time period and extend back several years prior to 1987. Our units of measure are the following: Area burned is in millions of acres, temperature is in Fahrenheit, and precipitation is in inches. New housing construction is for the western Census region of the US and is measured in thousands of units.[7] Net electricity imports to California are measured in terawatts. Ocean surface temperatures are measured in degrees Fahrenheit. We use F, T, P_{-1}, H, and E to denote respectively forest area burned, maximum temperature, lagged precipitation, and new private home construction, and net electricity imports. In the text, subscript $-k$ denotes the k-year lag of a variable. In the tables, suffix $\_k$ denotes the k-year lag of a variable.

We obtained for each year the highest and lowest monthly values of Pacific Decadal Oscillation (PDO). We denote these PDO_H and PDO_L respectively. We calculated the range of PDO as the difference between PDO_H and PDO_L and denoted the result PDO_R. We did the same for AMO, denoting the highest, lowest, and range as AMO_H, AMO_L, and AMO_R. Similarly for El Niño Southern Oscillation, we obtained $ENSO\_H$, $ENSO\_L$, and calculated $ENSO\_R$.[8]

We investigated availability of measures of the vapor-pressure deficit (VPD). We obtained county-level measures of the VPD for two counties that have experienced extensive fire damage, Riverside County and San Diego county. The measures we obtained are the mean of annual daily maximum values of the VPD for Riverside and San Diego counties. We denote these respectively VPD_RS and VPD_SD. As shown in the Figure 2(h), VPD_RS is much higher than VPD_SD.

We list below the variables in the form, linear or logarithmic, in which they appear in our models. In our subsequent presentation, we provide the rationale for choosing the logarithmic form for some variables. We use suffix $\_k$ to denote the k-year lag of a variable. For example, LP_1 is the one-year lag of precipitation.

LF: The logarithm of forest area burned. Area burned is expressed in millions of acres.
T: Average daily maximum temperature.
LP: The logarithm of precipitation.
LH: The logarithm of new private houses.
LE: The logarithm of electricity imports to California.
R_AMO: The range of Atlantic Multidecadal Oscillation

[7] We submitted an inquiry to the Census Bureau requesting data for housing construction for the state of California and were informed that state-level data are not available. California had 55 percent of the population of the western Census region in the 1980 decennial census and 50% in 2020. Thus, the pattern of change in housing construction in California can be expected to exhibit a similar pattern as the western Census region. Moreover, national market factors such as mortgage rates and the financial crisis can be expected to have similar proportionate impact on housing starts in California as in the western region.

[8] Appendix B details the sources of data.

R_PDO: The range of Pacific Decadal Oscillation
R_ENSO: The range of El Niño–Southern Oscillation
$LVPD\_RS$: The logarithm of the vapor pressure deficit in Riverside County California

4.2 Summary Statistics

Summary statistics for our main variables follow.

Table 1: Summary Statistics

	Mean	Median	Max	Min	Stdev
F	0.77	0.49	4.30	0.04	0.83
T	71.58	71.60	74.20	68.20	1.20
P	21.30	19.84	36.41	7.93	6.57
H	242.21	239.15	463.10	87.00	85.96
E	81.14	82.92	102.53	58.29	12.29
R_AMO	0.35	0.34	0.65	0.16	0.11
R_PDO	2.13	2.14	3.75	0.70	0.70
R_ENSO	1.47	1.30	3.80	0.20	0.89
VPD_RS	29.04	29.14	32.51	25.43	1.51

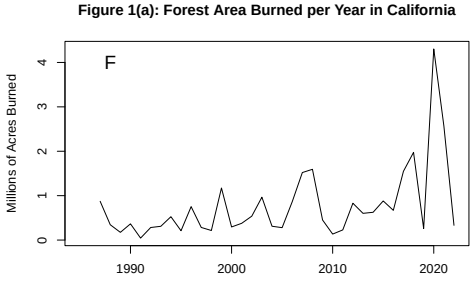
The correlation matrix for our variables is shown below. For brevity, we include the range of AMO, PDO, and ENSO but not the high and low values. It is of interest to note the relatively high correlation, .43, between maximum temperature and net electricity imports. This is intuitive. During hotter weather, households are likely to make more extensive use of air conditioning. The correlation between temperature and VPD_RS is also noteworthy. We explore this further in the analysis that follows.

Table 2: Correlation Matrix

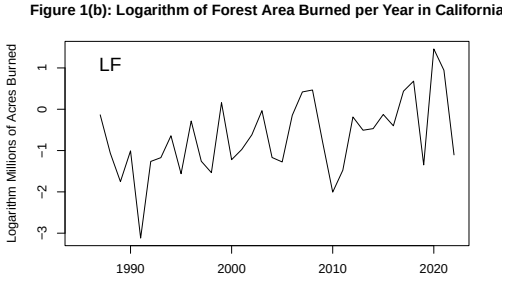
	F	T	P	H	E	R_AMO	R_PDO	R_ENSO	VPD_RS
F	1.00	0.54	-0.32	0.40	0.26	-0.05	-0.35	-0.04	0.65
T	0.54	1.00	-0.49	0.34	0.43	0.16	-0.44	-0.18	0.82
P	-0.32	-0.49	1.00	-0.10	-0.23	-0.08	0.43	0.30	-0.41
H	0.40	0.34	-0.10	1.00	0.01	0.23	-0.22	-0.20	0.26
E	0.26	0.43	-0.23	0.01	1.00	0.06	-0.11	0.07	0.33
R_AMO	-0.05	0.16	-0.08	0.23	0.06	1.00	-0.14	-0.08	0.03
R_PDO	-0.35	-0.44	0.43	-0.22	-0.11	-0.14	1.00	0.26	-0.26
R_ENSO	-0.04	-0.18	0.30	-0.20	0.07	-0.08	0.26	1.00	-0.18
VPD_RS	0.65	0.82	-0.41	0.26	0.33	0.03	-0.26	-0.18	1.00

4.3 Plots of the Variables

Figure 1(a) shows our dependent variable. An upward tendency in area burned is evident, along with increasing year-to-year variation in area burned. Figure 1(b) shows the logarithm of area burned. We return to the plot at the right in our discussion of model specification.



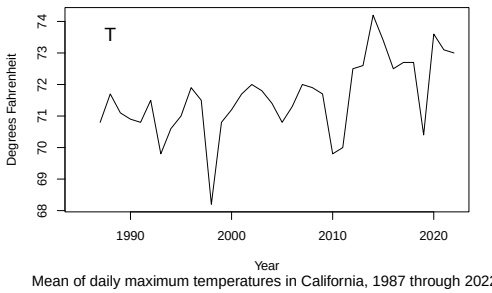
Forest area burned in California, 1987 through 2022.



Logarithm of Forest area burned in California, 1987 through 2022.

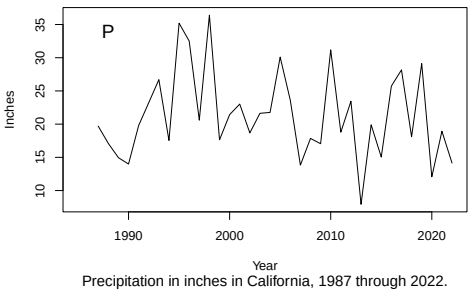
Figures 2(a) through 2(h) display our explanatory variables. The large variation in new housing construction, Figure 2(c), is noteworthy. New housing construction peaked prior to the financial crisis and then plummeted in 2007. Starts neared their pre-crisis peak in 2020 and exceeded the pre-crisis peak in 2022. It hardly needs saying that the financial crisis imposed great costs. From a statistical perspective, however, this large variation in new housing construction aids in obtaining a precise estimate of the effects of housing construction on forest fires. The plots also reveal much year-to-year variation in the other explanatory variables, enhancing estimation of the effects of these variables on forest area burned.

Figure 2(a): Maximum Temperature in California by Year



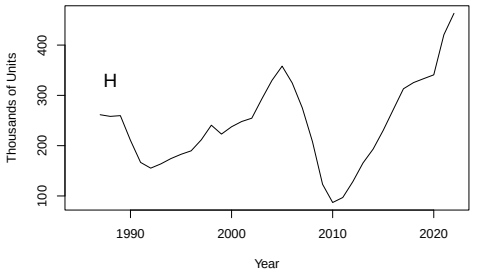
Mean of daily maximum temperatures in California, 1987 through 2022.

Figure 2(b): Precipitation per Year



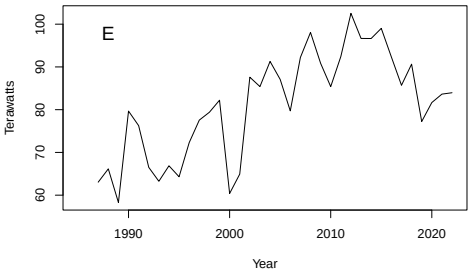
Precipitation in inches in California, 1987 through 2022.

Figure 2(c): New Housing Construction by Year



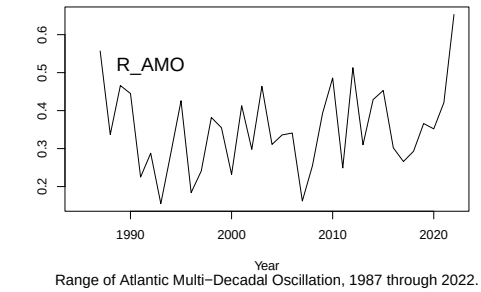
New housing construction (1,000's) in California, 1987 through 2022.

Figure 2(d): Net Electricity Imports to California by Year



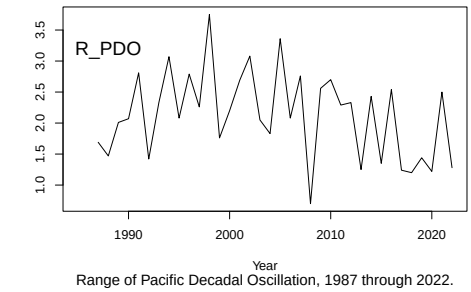
Terawatts of electricity imported to California, 1987 through 2022.

Figure 2(e): Range of Atlantic Multi−Decadal Oscillation



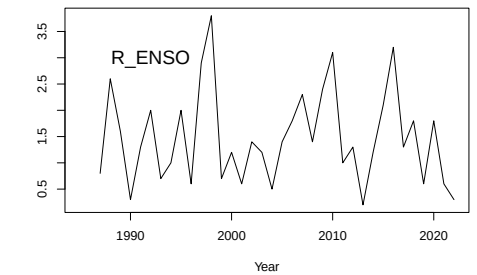
Range of Atlantic Multi−Decadal Oscillation, 1987 through 2022.

Figure 2(f): Range of Pacific Decadal Oscillation



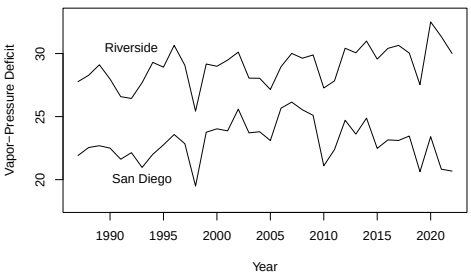
Range of Pacific Decadal Oscillation, 1987 through 2022.

Figure 2(g): Range of El Niño−Southern Oscillation Oscillation



Range of El Niño−Southern Oscillation, 1987 through 2022.

Figure 2(h): Vapor−Pressure Deficit: Riverside and San Diego Counties



Vapor−Pressure Deficits in San Diego and Riverside Counties, 1987−2022

5 Models of Forest Area Burned

We first estimated a linear regression with F as the dependent variable and T, P_{-1}, H, and E as independent variables. A Ramsey test of functional form yielded p-values of .005 and .004 with one and two fitted terms respectively. Column (1) is the counterpart model with variables other than temperature expressed in logarithms. Ramsey tests for this functional form yield p-values of .81 and .84 with one and two terms respectively. These high p-values are favorable to this functional form. For all models reported below, we use logarithms of forest area burned, precipitation variables, housing, and electricity imports.[9] Ramsey tests with one and two fitted terms for the regression in Column (2) of Table 3 yield p-values of .96 and .59 respectively. These are favorable to this functional form. Indeed, among the twelve Ramsey tests conducted for the regressions in Table 3, the lowest Ramsey test p-value is .29; this is for the test with two fitted terms for the regression in Column (6) of Table 3. Thus, all p-values are quite favorable for the functional forms for all regressions in Table 3.

In our initial work on this model, data were available for 1987 through 2020. Subsequently, data for years 2021 and 2022 became available. We noted non-negligible changes in the magnitudes of some of the coefficients when the two additional years of data were added. Hence, using the Chow forecast test, we tested for a structural break for each of the models in Columns (2), (4), and (6) of Table 3. We used 1987-2020 as the pre-break period and 2021-2022 as the post-break period. We obtained .02, .04, and .02 respectively. Hence, we strongly reject the hypothesis of no structural break for all three regressions. Therefore, in Table 3, we report in Columns (1), (3), and (5) regressions with data through 2020 for comparison with the regressions with data through 2022 in Columns (2), (4), and (6).

Columns (1) and (2) of Table 3 have four key explanatory variables, temperature, lagged precipitation, housing construction, and electricity imports. Two of these are climate variables and two are measures of human activities. Columns (3) and (4) add another climate variable, the logarithm of the vapor-pressure deficit. Columns (5) and (6) include the ocean temperature variables found to be significant in our stepwise analysis.[10] From the adjusted R^2, we see that the four variables in Column (1) account for 58% of the variation in forest area burned over the 1987-2020 period. The first three variables are significant at the 5% level while the logarithm of electricity imports is significant at the 10% level. Comparing Column (2) to Column (1), we see that coefficients of all except the temperature variable are smaller in Column (2), and housing starts and electricity imports are less significant in Column (2) than Column (1). These differences between Columns (1) and (2) comport with the finding of a significant structural break between 2020 and 2021. While we do not assert causality for the models in these columns, we note that these results are consistent a causal interpretation.

In Columns (3) and (4), we include the logarithm of the vapor pressure deficit for Riverside County. The estimates are consistent with the expectation that fires spread faster when humidity is low. We see as well that the coefficients of temperature in both regressions are small and insignificant. This clearly signals the presence of multicollinearity and strongly suggests that a channel by which temperature promotes forest fires is by drying the air. We explore this further in Table 4. Interestingly, while its coefficient declines, lagged precipitation remains significant in Column (3) when the vapor pressure deficit is introduced. We explore further in Table 4 whether precipitation may also be a channel affecting the vapor pressure deficit.

The results in Columns (1) through (4) can be thought of as estimates of the predicted effects of local climate variables on forest area burned. By local variables, we mean measures specific to California or, in the case of the vapor pressure deficit, a county in California. From the adjusted R^2 values for Columns (3) and (4), we see that local climate variables and human activities can account for approximately 60% of the annual variation in forest area burned.

The ocean temperature variables can be thought of as global climate variables in the sense that their effects

[9] We do not use the logarithm of temperature variables, T, R_AMO, R_ENSO, R_PDO, because temperature is not a cardinal variable. If the logarithm of temperature were used, spurious differences in results would arise based on the measure used, e.g., centigrade, Fahrenheit, or Kelvin. This does not arise with temperature in unlogged form.

[10] We performed the Jarque-Bera normality test for the residuals for the six regressions in Table 3, obtaining, in order the regressions appear in the table, p-values of .57, .62, .54, .80, .52, and .62. These high p-values provide strong evidence for the validity of the t-distribution for hypothesis tests of the coefficients of the equations in Table 3.

are much more far reaching than the effects of variables in Columns (3) and (4). We next turn to discussion of the effects of these global climate variables. In columns (5) and (6), we include results from our stepwise analysis of ocean surface temperature variables. We see that the coefficients of R_AMO_1 and R_AMO_2 are negative and highly significant. Thus, an increase in the difference between maximum and minimum Atlantic sea temperatures in a given year is followed by a predicted reduction in forest area burned in the subsequent two years. By contrast, an increased range of variation of ocean surface temperatures near the equator is followed by a predicted increase in forest area burned the following year. The logarithm of the vapor pressure deficit is not significant in Columns (5) and (6). However, a joint test of the hypothesis that the coefficients of LP_1 and LVPD_RS are both zero yields a p-values of .02 and .03 for the regressions in Column (5) and (6) respectively. Thus, the variables are jointly significant at the .05 level. From the adjusted R^2 values, we see that the regressions in Columns (5) and (6) predict more than 75% of the variation in forest area burned in California over the time period of our sample.

Readers might wonder whether the use of range of ocean temperatures is appropriate, or whether annual maximum and annual minimum values, appropriately lagged, should instead be used. Appendix Table A1 contains the regression with maximum and minimum values instead of the ranges. With that regression, we tested the constraints that are implied by use of ranges of the AMO and ENSO variables. This is a standard linear hypothesis test. The test yields a p-value of .92 with data through 2020 and a p-value of .76 with data through 2022. We also did the same test with a model that did not include the logarithm of the vapor pressure deficit and got p-values of .91 and .97 respectively. These high p-values support the conclusion that the ranges of the AMO and ENSO variables capture their effects quite well.

In Table 3, the coefficients of $ln(P_{-1})$, $ln(H)$, and $ln(E)$ are elasticities. The coefficients of all other variables are semi-elasticities. Hence, the coefficients in Column (6) of Table 1 can be interpreted as follows. Holding constant other variables in the regression, a 0.1 degree increase in maximum temperature increases predicted area burned by 4.2%. Holding constant other variables in the regression, 1% increase in precipitation in a given year increases predicted area burned the following year by .50%. A 1% increase in new housing construction predicts a .59% increase in predicted area burned.[11] A 1% increase in net electricity imports predicts a 1.00% increase in predicted area burned. A .1 degree increase in R_AMO predicts a .32% decrease in area burned one year later and a .23% decrease in area burned two years later. A .1 increase in R_ENSO predicts a 3.6% increase in area burned one year later.

While the preceding interpretations are correct, they do not convey a sense of the relative magnitudes of the predicted impacts of the variables over the period of our sample. For this purpose, standardized coefficients are more informative. We present and discuss these following Table 3.

[11] The proviso "holding constant other variables" applies to the interpretation of each coefficient.

	Table 3					
	LF					
	(1)	(2)	(3)	(4)	(5)	(6)
T	0.398	0.393	0.077	-0.003	0.439	0.411
	p = 0.00001	p = 0.00002	p = 0.685	p = 0.988	p = 0.0002	p = 0.001
LP_1	1.139	0.983	0.782	0.624	0.622	0.498
	p = 0.002	p = 0.013	p = 0.028	p = 0.072	p = 0.028	p = 0.067
LH	0.595	0.552	0.715	0.647	0.657	0.588
	p = 0.017	p = 0.075	p = 0.005	p = 0.025	p = 0.0003	p = 0.008
LE	1.211	1.170	1.324	1.314	1.052	0.996
	p = 0.062	p = 0.068	p = 0.059	p = 0.068	p = 0.031	p = 0.054
LVPD_RS			8.165	9.966	3.310	4.224
			p = 0.067	p = 0.032	p = 0.178	p = 0.105
R_AMO_1					-2.632	-2.832
					p = 0.0003	p = 0.0001
R_AMO_2					-2.045	-2.051
					p = 0.015	p = 0.020
R_ENSO_1					0.312	0.357
					p = 0.002	p = 0.0003
Constant	-49.518	-47.973	-54.874	-54.233	-59.419	-59.116
	p = 0.000	p = 0.000	p = 0.000	p = 0.000	p = 0.000	p = 0.000
Observations	34	36	34	36	34	36
R2	0.626	0.551	0.689	0.643	0.839	0.806
Adjusted R2	0.575	0.493	0.633	0.583	0.787	0.749
Residual Std. Error	0.583	0.650	0.541	0.589	0.412	0.458
F Statistic	12.156***	9.507***	12.398***	10.796***	16.233***	14.050***

5.1 Standardized coefficients

Standardized coefficients (beta weights) are scale-free and provide further insight into the relative impacts of the variables. The standardized coefficients for the models in Columns (2), (4), and (6) of Table 3 are presented below. A standardized coefficient is interpreted as follows. Suppose an explanatory variable increases by one standard deviation while all other explanatory variables are unchanged. The standardized coefficient of that explanatory variable is the number of standard deviation by which the dependent variable changes. Hence, we see that the estimates from Column (2) imply that a one standard deviation increase in temperature increases the logarithm of predicted area burned by .52 standard deviations. The standardized coefficient, .54, of temperature in Column (6) is exceeding close to that for Column (2). By contrast, temperature is insignificant in Column (4) and its standardized coefficient is negligibly small. A one standard

deviation increase in the logarithm of lagged precipitation increases predicted logarithm of area burned in Columns (2), (4), and (6) by .34, .21, and .17 standard deviations respectively. A one standard deviation increase in the logarithm of new housing construction increases predicted logarithm of area burned by .23, .27, and .25 standard deviations respectively in Columns (2), (4), and (6). Thus, the estimated effects of this human activity are remarkably similar across the three regressions. A one standard deviation increase in the logarithm of net electricity imports increases predicted logarithm of area burned by .20, .23, and .17 respectively in Columns (2), (4), and (6).

A one standard deviation increase in the range of AMO in a given year reduces the predicted logarithm of area burned by .32 standard deviations one year later and by .23 standard deviations two years later. A one standard deviation increase ENSO increases predicted area burned one year later by .34 standard deviations. The standardized coefficients for Column (6) show that each of the variables had a substantial predicted effect on area burned. It should be noted that the standardized coefficients are not additive across the variables because the variables are correlated; each standardized coefficient is the effect of variation in the associated variable with all other variables held constant.

Standardized Coefficients Column (2) of Table 3

	Coefs	Std_Coefs
T	0.3928	0.5184
LP_1	0.9826	0.3384
LH	0.5524	0.2311
LE	1.1702	0.2014

Standardized Coefficients Column (4) of Table 3

	Coefs	Std_Coefs
T	-0.0032	-0.0042
LP_1	0.6238	0.2148
LH	0.6469	0.2706
LE	1.3143	0.2262
LVPD_RS	9.9655	0.5737

Standardized Coefficients Column (6) of Table 3

	Coefs	Std_Coefs
T	0.4105	0.5418
LP_1	0.4984	0.1716
LH	0.5881	0.2460
LE	0.9957	0.1714
LVPD_RS	4.2237	0.2432
R_AMO_1	-2.8324	-0.3194
R_AMO_2	-2.0514	-0.2303
R_ENSO_1	0.3568	0.3406

6 Implications for Policy

We noted above the finding of a significant structural break break between 2020 and 2021 for the three equations in Table 3 that are estimated with data for 2022. This is potentially very important when we consider these results in light of California Assembly Bill No. 1054 passed July 12, 2019. As summarized in our literature review, this law states "The state's electrical corporations must invest in hardening of the state's electrical infrastructure and vegetation management to reduce the risk of catastrophic wildfire." Moreover, a heavy financial burden is placed on electrical corporations to incentive them to undertake the hardening of infrastructure. Such hardening could include better insulation of power lines or burying them underground, removing tree branches that might ignite a fire by rubbing against above-ground electrical wires in windy conditions, and removing vegetation in the vicinity of such power lines so that there is less fuel to feed a fire if one should be ignited. Such changes would, in turn, be expected to reduce the estimated magnitude of coefficients of the coefficients of the first five variables in the model in Column (6) relative to the coefficients in Column (5). We see that the coefficients of the first four variables are in fact smaller in Column (6) than in Column (5). The coefficient of the fifth variable, the vapor pressure deficit, increases, but is not significant in either regression. One would expect that, controlling for the first five variables, hardening would have little effect on the coefficients of the global variables, and we find little difference in those coefficients between Columns (5) and (6). We also wish to emphasize that, while the differences in the first four coefficients are consistent with improved hardening mandated by Assembly Bill No. 1054, the differences are by no means definitive. Hence, the policy change made in Assembly Bill No. 1054 may have given rise to the significant structural break that we find, but more definitive analysis must await additional data.

Also, from a policy perspective, the local climate variables are more readily monitored on an ongoing basis than the global climate variables. The contrasting results in Columns (2) and (4) argue for further analysis to obtain a deeper understanding of the interdependence of the local climate variables. In particular, the contrast of the coefficients of temperature in Columns (2) and (4) strongly suggest that the vapor pressure deficit is the major channel by which temperature exacerbates forest area burned.

7 Modeling the Vapor Pressure Deficit

The analysis of factors affecting wildfires is our primary objective. Analysis of the vapor pressure deficit contributes to this objective, and is also of independent scientific interest. The first four columns of Table 4 utilize data for 1987-2022. The fifth utilizes data for 1987-2020. In Column (1) of Table 4, we find that temperature alone can account for two thirds of the variation in LVPD_RS over the course of our sample. Contemporaneous precipitation would be expected to increase humidity, thereby reducing the vapor pressure deficit. Column (2) adds the logarithm of contemporaneous precipitation. It has the anticipated algebraic sign, but the coefficient is quantitatively small and far from significant. Column (3) includes instead the year to year change in the logarithm of precipitation. This proves to be significant at the 10% level and is of the anticipated sign. Standardized coefficients of the two variables are .79 and -.16 respectively. Hence, temperature is the dominant factor in this regression. In Column (4), we include ocean surface temperature variables, finding the first and second lags of R_AMO and R_ENSO to be significant as well as contemporaneous R_PDO. We investigate whether there was a structural break between 2020 and 2021, obtaining a p-value of .85. Hence, there is no evidence of a structural break. This is as expected and is quite reassuring. There is no reason why a change in policy regarding electricity transmission should have any effect on the phenomena that govern weather in California. The regressions in Columns (1) through (4) use data for 1987-2022. The regression in Column (5) uses data from 1987-2020, and is included for comparison to the regression in Column (4). The coefficients are nearly identical across the models. This is as expected given the high p-value of the test for a structural break for the model in Column (4). Interestingly, the coefficient of temperature increases substantially when we include the ocean temperature variables as can be seen by comparing the coefficients of temperature between Columns (3) and (4). We next present the scaled coefficients, followed by an analysis of autoregressive errors.

Table 4

	LVPD_RS				
	(1)	(2)	(3)	(4)	(5)
T	0.036	0.035	0.035	0.049	0.048
	p = 0.000	p = 0.000	p = 0.000	p = 0.000	p = 0.000
LP		-0.005			
		p = 0.727			
dLP			-0.019	-0.016	-0.017
			p = 0.031	p = 0.073	p = 0.061
R_AMO_1				-0.070	-0.070
				p = 0.017	p = 0.019
R_AMO_2				-0.113	-0.110
				p = 0.055	p = 0.065
R_ENSO_1				0.017	0.017
				p = 0.0002	p = 0.001
R_ENSO_2				0.012	0.012
				p = 0.031	p = 0.028
R_PDO				0.024	0.023
				p = 0.001	p = 0.003
Constant	0.803	0.867	0.891	-0.203	-0.128
	p = 0.002	p = 0.009	p = 0.0001	p = 0.505	p = 0.692
Observations	36	36	36	36	34
R2	0.674	0.675	0.700	0.853	0.843
Adjusted R2	0.665	0.656	0.682	0.817	0.800
Residual Std. Error	0.030	0.031	0.030	0.023	0.023
F Statistic	70.420***	34.307***	38.464***	23.252***	19.892***

Standardized coefficients for the regression in Column (4) of Table 4 are presented below. We see that temperature has an extremely large effect on the VPD. Holding other variables constant, a one standard deviation increase in temperature increases the logarithm of VPD by 1.13 standard deviations. A standard deviation increase in precipitation has an effect roughly one eighth the magnitude of a one standard deviation in temperature. Precipitation, whether entered contemporaneously or as the change from the prior year, has an effect of the anticipated sign but that effect is relatively modest compared to the effects of other variables in the regression. An increase in either the one-year or two-year lag of R_AMO has a substantial effect in reducing the VPD while an increase in either the one-year and two-year lag of R_ENSO increases the VPD. A contemporaneous increase of PDO increases the VPD.

Standardized Coefficients Column (4) of Table 4

	Coefs	Std_Coefs
T	0.0494	1.1326
dLP	-0.0161	-0.1387
R_AMO_1	-0.0700	-0.1370
R_AMO_2	-0.1127	-0.2197
R_ENSO_1	0.0175	0.2900
R_ENSO_2	0.0117	0.1937
R_PDO	0.0236	0.3129

The results above suggest that temperature impacts forest area burned primarily through the vapor pressure deficit. As shown in Columns (3) and (4) of Table 3, the coefficients of temperature are small and insignificant when the vapor pressure deficit is introduced. By contrast, precipitation operates through two channels. From Columns (3) and (4) of Table 3 and the associated standardized coefficients, we see that lagged precipitations has a substantial effect in increasing forest area burned. From Column (3) of Table 4, we see that an increase in precipitation reduces the vapor pressure deficit, thereby contributing to a reduction in area burned. The former effect far outweighs the latter. Hence, the overall effect of precipitation is to increase forest area burned.

Below, we provide a correlation matrix of the explanatory variables in the models in Table 4. There is remarkably little correlation across the sea surface temperature variables. These low correlations provide insight into why the coefficients of the sea surface temperature variables in Column (4) are relatively precisely estimated.

Table 6: Correlation Matrix of Explanatory Variables in Table 4

	T	dLP	R_AMO_1	R_AMO_2	R_ENSO_1	R_ENSO_2	R_PDO
T	1.00	-0.17	0.15	0.31	-0.39	0.02	-0.44
dLP	-0.17	1.00	-0.20	0.18	-0.04	-0.23	0.34
R_AMO_1	0.15	-0.20	1.00	-0.09	0.01	0.03	-0.08
R_AMO_2	0.31	0.18	-0.09	1.00	-0.11	0.06	0.01
R_ENSO_1	-0.39	-0.04	0.01	-0.11	1.00	0.02	-0.07
R_ENSO_2	0.02	-0.23	0.03	0.06	0.02	1.00	-0.13
R_PDO	-0.44	0.34	-0.08	0.01	-0.07	-0.13	1.00

8 Further Analysis

8.1 Robustness

To investigate robustness, we estimate in Appendix Table A1 five regressions. All include the variables from the model in Column (6) of Table 3. In addition, the five regressions include respectively the logarithm of precipitation, lagged mean maximum temperature, mean minimum temperature, mean temperature, and California population growth. All five have p-values greater than .3. Similarly, regressions in Table A2 include the variables in our preferred model and, respectively, the Palmer Drought Severity Index (PDSI), the one-year lag of the Palmer Drought Severity Index (PDSI_1), the mean of daily maximum summer temperatures, and the logarithm of total electricity generation in California. The p-values are above .15 for all of these variables. Thus, none of the additional variables approach significance at the 5% or 10% significance levels. Moreover, the coefficients of the variables in our preferred model are relatively little affected by the inclusion of variables enumerated above.

8.2 Estimates with Autoregressive Errors

A potential concern is that the variables in our data might be cointegrated, potentially leading to spurious results. Dickey-Fuller unit root tests for LF, T, LP, LH, R_AMO, R_ENSO and R_PDO yield p-values of .001, .014, .000, .024, .000, .000, and .000. Hence, we reject the null hypothesis of a unit root for these seven variables. The unit root test for LE yields a p-value of .146. Hence, we do not reject a unit root for LE. However, this is not cause for concern. McCallum (2010) makes a compelling case that, if present, the spurious regression problem will be detected by estimating the model allowing for autocorrelation of the error terms. If the spurious regression problem is present, the AR(1) coefficient will be close to one. We estimated the model in Column (6) of Table 3 with autoregressive errors and found the first- and second-order coefficients to be significant. The resulting model is presented below. The ar1 coefficient is far below one, indeed significantly negative. Moreover, the estimated coefficients are similar to those in the corresponding model without autoregressive errors.

Table 6: Forest Fires Model with Autoregressive Errors

	Coef	pvalue
T	0.3580	0.0070
LP_1	0.3296	0.0781
LH	0.6876	0.0000
LE	1.1753	0.0013
LVPD_RS	3.5542	0.0839
R_AMO_1	-2.7786	0.0001
R_AMO_2	-1.9923	0.0070
R_ENSO_1	0.2700	0.0019
intercept	-55.0622	0.0000
ar1	-0.4293	0.0137
ar2	-0.4278	0.0217

We did a similar investigation for the vapor pressure deficit. The following is the counterpart to Column (5) of Table 4. The third order autoregressive coefficient is significant; the first- and second-order terms are not. While the coefficients are somewhat different from those in Column (5), there are no major changes in the coefficients. The significance levels change, with all variables now having p-values well below .05. Note the magnitude and significance of the coefficient of temperature are little changed. This further ratifies the importance of temperature in affecting the vapor pressure deficit.

Table 7: VPD Model with Autoregressive Errors

	Coef	pvalue
intercept	0.0926	0.3361
T	0.0452	0.0000
dLP	-0.0197	0.0054
R_AMO_1	-0.0725	0.0056
R_AMO_2	-0.0848	0.0041
R_ENSO_1	0.0196	0.0000
R_ENSO_2	0.0125	0.0008
R_PDO	0.0212	0.0000
ar3	-0.5958	0.0005

9 Conclusion

California has 4% of the land area of the United States, but over the 36-year period of our sample (1987-2022) California averaged 13% of the total US forest area burned. This underlines the importance of understanding the factors that give rise to forest fires in that state. Moreover, understanding the factors impacting forest fires in California is of broad economic and, more generally, scientific interest. We use fire area burned data from the state government, statewide temperature and precipitation data from the US National Oceanographic and Atmospheric Administration, new housing construction from the US Census Bureau, electric power import data from the California Energy Commission, and measures of AMO, PDO, and ENSO from the National Oceanic and Atmospheric Administration. We find that 75 percent of the variability in forest area burned can be accounted for by variation in six variables: the mean of maximum annual temperatures, prior year precipitation, new housing construction, net electricity imports, and variation in AMO and ENSO. The latter two variables have been shown to influence Santa Ana Winds which, in turn, propagate forest fires. From the standardized coefficients we find that all variables have quantitatively large effects on area burned with temperature having the largest effect. From a policy perspective, local human activities are the only variables that state and local authorities can potentially change. Reporting on the recent fires in Hawaii, Sacks (2023) "Before the Maui wildfires, Hawaiian Electric did not have a plan — adopted widely in California and other states — to shut off power in certain lines in advance of dangerous winds." We noted the correlation between electricity imports and temperature. This poses a particular challenge for regulatory authorities because electricity imports increase during times of high temperature and, hence, during times of high potential for forest fire severity.

9.0.1 Data And Software Availability

Data will be made available as well as R Markdown code used to analyze the data and format output.

9.1 References

Bishop, Kelly C., Jonathan D. Ketcham, and Nicolai V. Kuminoff, "Hazed and Confused: The Effect of Air Pollution on Dementia," *Review of Economic Studies*, (2023) 00, 1–27.

Brooks, M.L. and J.R. Matchett, Spatial and temporal patterns of wildfires in the Mojave Desert, 1980–2004, *Journal of Arid Environments*, Volume 67, Supplement, 2006, Pages 148-164, https://doi.org/10.1016/j.jaridenv.2006.09.027.

Burke, M., Driscoll, A., Heft-Neal, S., Xue, J., Burney, J., and Wara, M. (2021) The changing risk and burden of wildfire in the United States. *Proc. Natl. Acad. Sci.*, 118(2), pp. 1-6. https://doi.org/10.1073/pnas.2011048118

Cardil, Adrián, Marcos Rodrigues, Joaquin Ramirez, Sergio de-Miguel, Carlos A. Silva, Michela Mariani, Davide Ascoli. Coupled effects of climate teleconnections on drought, Santa Ana winds and wildfires in southern California. *Science of the Total Environment* 765 (2021) 142788.

Crimmins Michael A., Comrie Andrew C. (2004) Interactions between antecedent climate and wildfire variability across south-eastern Arizona. *International Journal of Wildland Fire* 13, 455-466. https://doi.org/10.1071/WF03064

Environmental Protection Administration, https://www.epa.gov/wildfire-smoke-course/why-wildfire-smoke-health-concern

Glickman, D., and B. Babbitt. "Urban wildland interface communities within the vicinity of federal lands that are at high risk from wildfire." *Federal Register* 66.3 (2001): 751-777.

Guyette, R.P., Stambaugh, M.C., Dey, D.C. et al. Predicting Fire Frequency with Chemistry and Climate. *Ecosystems* 15, 322–335 (2012). https://doi.org/10.1007/s10021-011-9512-0.

Kale, M.P., Ramachandran, R.M., Pardeshi, S.N. et al. Are Climate Extremities Changing Forest Fire Regimes in India? An Analysis Using MODIS Fire Locations During 2003–2013 and Gridded Climate Data of India Meteorological Department. *Proc. Natl. Acad. Sci., India, Sect. A Phys. Sci.* 87, 827–843 (2017). https://doi.org/10.1007/s40010-017-0452-8.

Knorr, W., Kaminski, T., Arneth, A., and Weber, U.: Impact of human population density on fire frequency at the global scale, *Biogeosciences*, 11, 1085–1102, https://doi.org/10.5194/bg-11-1085-2014, 2014.

Koutsias Nikos, Xanthopoulos Gavriil, Founda Dimitra, Xystrakis Fotios, Nioti Foula, Pleniou Magdalini, Mallinis Giorgos, Arianoutsou Margarita (2012) On the relationships between forest fires and weather conditions in Greece from long-term national observations (1894–2010). *International Journal of Wildland Fire* 22, 493-507. https://doi.org/10.1071/WF12003

Li, A.K., Paek, H., Yu, J.Y., 2016. "The changing influences of the AMO and PDO on the decadal variation of the Santa Ana winds. *Environ. Res. Lett.* 11, 064019.

Matthias Forkel et al., Extreme fire events are related to previous-year surface moisture conditions in permafrost-underlain larch forests of Siberia, 2012 *Environ. Res. Lett.* 7 044021. doi:10.1088/1748-9326/7/4/044021.

McCallum, Bennett T., "Is the spurious regression problem spurious?" *Economics Letters*, Volume 107, Issue 3, June 2010, Pages 321-323.

Millspaugh, Sarah H., Cathy Whitlock, and Patrick J. Bartlein. "Variations in fire frequency and climate over the past 17 000 yr in central Yellowstone National Park." *Geology* 28.3 (2000): 211-214.

Minnich, Richard A., Ernesto Franco Vizcaíno, Raymond J. Dezzani, "The El Niño/Southern Oscillation and Precipitation Variability in Baja California, Mexico," Atmósfera vol.13 no.1 Ciudad de México ene. 2000.

MODIS Active Fire and Burned Area Products, https://modis-fire.umd.edu/

Moeltner. K., M.-K. Kim, E. Zhu, W. Yang, "Wildfire smoke and health impacts: A closer look at fire attributes and their marginal effects," *Journal of Environmental Economics and Management*, 66 (2013), 476-96.

Owen Price, Ross Bradstock, Countervailing effects of urbanization and vegetation extent on fire frequency on the Wildland Urban Interface: Disentangling fuel and ignition effects, *Landscape and Urban Planning*, Volume 130, 2014, Pages 81-88, https://doi.org/10.1016/j.landurbplan.2014.06.013.

Pausas, J.G. Changes in Fire and Climate in the Eastern Iberian Peninsula (Mediterranean Basin). *Climatic Change* 63, 337–350 (2004). https://doi.org/10.1023/B:CLIM.0000018508.94901.9c

Porter, Thomas, Director CAL FIRE, Mike Richwine, State Fire Marshal, Marybel Batjer, President CA Public Utilities Commission, *California Power Line Fire Prevention Field Guide*,2021.

Pullabhotla, Hemant K. and Mateus Souza, "Air pollution from agricultural fires increases hypertension risk," *Journal of Environmental Economics and Management*, 11 (2022) 102723. https://doi.org/10.1016/j.jeem.2022.102723

Radeloff, V.C., Hammer, R.B., Stewart, S.I., Fried, J.S., Holcomb, S.S. and McKeefry, J.F. (2005), The Wildland–Urban Interface In The United States. *Ecological Applications*, 15: 799-805. https://doi.org/10.1890/04-1413

Rangel, Marcos A. and Tom S. Vogl, Agricultural Fires and Health at Birth, *The Review of Economics and Statistics*, October 2019, 101(4): 616–630.

Renard Quentin, Pélissier Raphaël, Ramesh B. R., Kodandapani Narendran (2012) Environmental susceptibility model for predicting forest fire occurrence in the Western Ghats of India. *International Journal of Wildland Fire* 21, 368-379. https://doi.org/10.1071/WF10109

Rosales-Rueda, Maria and Margaret Triyana, "The Persistent Effects of Early-Life Exposure to Air Pollution", Evidence from the Indonesian Forest Fires, *Journal of Human Resources*, Volume 54, Number 4, Fall 2019, pp. 1037-1080.

Sacks, Brianna, "Hawaii utility faces scrutiny for not cutting power to reduce fire risks," August 12, 2023. https://www.washingtonpost.com/climate-environment/2023/08/12/maui-fire-electric-utility

Syphard, A.D., Radeloff, V.C., Keeley, J.E., Hawbaker, T.J., Clayton, M.K., Stewart, S.I. and Hammer, R.B., 2007. Human influence on California fire regimes. *Ecological applications*, 17(5), pp.1388-1402.

Turco, M., Llasat, M.C., von Hardenberg, J. et al. Impact of climate variability on summer fires in a Mediterranean environment (northeastern Iberian Peninsula). *Climatic Change* 116, 665–678 (2013). https://doi.org/10.1007/s10584-012-0505-6

US National Interagency Fire Center Wildland Fire Summaries, https://www.nifc.gov/fire-information/statistics, accessed 2022 January 24.

Westerling, A. L., Bryant, B. P., Preisler, H. K., Holmes, T. P., Hidalgo, H. G., Das, T., et al.(2011). Climate change and growth scenarios for California wildfire. *Climatic Change*, 109, 445–463.

Williams, A. P. , J. T. Abatzoglou, A. Gershunov, J. Guzman-Morales, D. A. Bishop, J. K. Balch, and D. P. Lettenmaier (2019). Observed impacts of anthropogenic climate change on wildfire in California. *Earths Future*, 7, 892–910.

World Bank, https://data.worldbank.org/indicator/AG.LND.FRST.ZS?locations=US, accessed 2022 January 22.

Zhang, Boya, Jennifer Weuve, Kenneth M. Langa, Jennifer D'Souza, Adam Szpiro, Jessica Faul, Carlos Mendes de Leon, Jiaqi Gao, Joel D. Kaufman, Lianne Sheppard, Jinkook Lee, Lindsay C. Kobayashi,

Richard Hirth, Sara D. Adar, "Comparison of Particulate Air Pollution From Different Emission Sources and Incident Dementia in the US," *Journal of The American Medical Association*, August 14, 2023. https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/2808088?guestaccesskey=41afaad4-70b6-4cac-a49b-1b6de01b83d4&utm_source=for_the_media&utm_medium=referral&utm_campaign=ftm_links&utm_content=tfl&utm_term=081423

Zhisheng Xu, Dingli Liu, Long Yan, Temperature-based fire frequency analysis using machine learning: A case of Changsha, China, Climate Risk Management, 31, 100276 (2021) https://doi.org/10.1016/j.crm.2021.100276.

10 Appendix A: Supplementary Analyses

10.1 Analysis of Global Variables

In our regressions in Table 3, we used the ranges of AMO and ENSO. It is of interest to investigate whether the ranges are sufficient or whether the maximum and minimum values should instead be used. Column (1) below reproduces the regression from Column (3) of Table 3. Column (2) below instead uses the maximum and minimum values. For example, Column (2) uses AMO_H_1 and AMO_L_1 rather than R_AMO_1. From Column (2) we see that the coefficients of AMO_H_1 and AMO_L_1 are -2.83 and 2.64 respectively. Thus, they are opposite in sign and similar in magnitude. The same is true for the coefficients of AMO_H_2 and AMO_L_2, and the same is true for the coefficients of ENSO_H_1 and ENSO_L_1. We did a single joint test of the null hypothesis that the following three conditions hold: the population coefficients of AMO_H_1 and AMO_L_1 are equal in magnitude and opposite in sign and the same for the population coefficients AMO_H_2 and AMO_L_2 and the same for the population coefficients of ENSO_H_1 and ENSO_L_1. We obtained a p-value=.91. This test then establishes that the regression in Column (1) that utilizes ranges is capturing the impact of movements in AMO and ENZO in impacting forest fires in California. A similar test applied to the model in Column (5) of Table 4 yields a p-value=.24. Hence, the range variables capture the impact of the AMO, ENSO, and PDO variables on the vapor pressure deficit.

Table A1

	LF	
	(1)	(2)
T	0.605	0.601
	p = 0.000	p = 0.000
LP_1	0.597	0.582
	p = 0.034	p = 0.067
LH	0.547	0.628
	p = 0.016	p = 0.029
LE	0.871	1.258
	p = 0.073	p = 0.221
R_AMO_1	-3.017	
	p = 0.00002	
R_AMO_2	-2.389	
	p = 0.003	
R_ENSO_1	0.410	
	p = 0.00003	
AMO_H_1		-3.146
		p = 0.0001
AMO_L_1		3.081
		p = 0.017
AMO_H_2		-2.418
		p = 0.006
AMO_L_2		1.842
		p = 0.103
ENSO_H_1		0.400
		p = 0.001
ENSO_L_1		-0.435
		p = 0.012
Constant	-57.409	-61.870
	p = 0.000	p = 0.00001
Observations	36	36
R2	0.793	0.797
Adjusted R2	0.741	0.716
Residual Std. Error	0.465	0.486
F Statistic	15.315***	9.841***

10.2 Investigation of Potential Effects of Other Variables

Each of the regressions in Table A2 below adds one variable to the regression in Column (3) of Table 3. The regressions include, respectively, the log of contemporaneous precipitation, lagged maximum temperature, minimum temperature, mean temperature, California population growth, and global surface temperature. The p-values are above .3 for all of these variables. Hence, we conclude that these variables do not have significant additional impact on forest area burned.

Each of the regressions in Table A3 below adds one variable to the regression in Column (3) of Table 3. The regressions include, respectively, the range of PDO lagged one year, the mean of daily maximum summer temperatures, and the logarithm of total electricity generation in California. The p-values are above .13 for all of these variables. Hence, we conclude that these variables do not have significant additional impact on forest area burned.

Table A2

	LF				
	(1)	(2)	(3)	(4)	(5)
T	0.570	0.719	0.858	0.615	0.622
	p = 0.000	p = 0.000	p = 0.00002	p = 0.000	p = 0.000
LP_1	0.561	0.548	0.553	0.608	0.552
	p = 0.046	p = 0.047	p = 0.047	p = 0.048	p = 0.078
LH	0.568	0.606	0.600	0.588	0.673
	p = 0.019	p = 0.008	p = 0.010	p = 0.004	p = 0.007
LE	0.872	0.880	0.850	1.070	1.327
	p = 0.096	p = 0.103	p = 0.117	p = 0.105	p = 0.089
R_AMO_1	-3.216	-3.323	-3.253	-3.142	-3.039
	p = 0.00001	p = 0.00000	p = 0.00000	p = 0.00002	p = 0.00001
R_AMO_2	-2.336	-2.370	-2.386	-2.512	-2.551
	p = 0.007	p = 0.007	p = 0.007	p = 0.002	p = 0.002
R_ENSO_1	0.417	0.415	0.417	0.411	0.425
	p = 0.00003	p = 0.00004	p = 0.00005	p = 0.0001	p = 0.00004
LP	-0.274				
	p = 0.300				
T_min		-0.189			
		p = 0.093			
T_mean			-0.315		
			p = 0.154		
Pop_Gro				7.952	
				p = 0.618	
Global_Surface_Temp					-0.319
					p = 0.358
Constant	-54.027	-56.955	-56.772	-60.591	-45.695
	p = 0.000	p = 0.000	p = 0.000	p = 0.000	p = 0.007
Observations	36	36	36	36	36
R2	0.799	0.807	0.803	0.796	0.800
Adjusted R2	0.740	0.750	0.745	0.735	0.740
Residual Std. Error	0.466	0.456	0.461	0.470	0.465
F Statistic	13.441***	14.133***	13.772***	13.150***	13.468***

Table A3

	LF		
	(1)	(2)	(3)
T	0.598	0.521	0.606
	p = 0.000	p = 0.000	p = 0.000
LP_1	0.671	0.608	0.595
	p = 0.027	p = 0.017	p = 0.038
LH	0.533	0.369	0.553
	p = 0.027	p = 0.113	p = 0.015
LE	0.923	0.569	1.010
	p = 0.062	p = 0.241	p = 0.350
R_AMO_1	-3.027	-2.861	-3.044
	p = 0.00004	p = 0.00002	p = 0.00004
R_AMO_2	-2.321	-2.201	-2.395
	p = 0.005	p = 0.003	p = 0.004
R_ENSO_1	0.414	0.391	0.413
	p = 0.00004	p = 0.00002	p = 0.0001
R_PDO_1	-0.085		
	p = 0.551		
SummerMaxTmp		0.114	
		p = 0.048	
log(TotElecGen)			-0.306
			p = 0.867
Constant	-57.471	-57.218	-55.176
	p = 0.000	p = 0.000	p = 0.0002
Observations	36	36	36
R2	0.796	0.809	0.793
Adjusted R2	0.736	0.752	0.732
Residual Std. Error	0.469	0.454	0.473
F Statistic	13.175***	14.284***	12.942***

11 Appendix B: Data Sources

Data for California fire area burned for 1987 through 2018 are from: https://www.fire.ca.gov/media/11397/fires-acres-all-agencies-thru-2018.pdf Data for subsequent years are from: https://www.fire.ca.gov/stats-events/

California temperature and precipitation data are from: https://www.ncdc.noaa.gov/cag/statewide/time-series/4/tavg/12/12/1895-2021?base_prd=true&begbaseyear=1901&endbaseyear=2000

Data for new housing construction in the west are from: Housing Units Under Construction at End of Period in the West Census Region: https://www.census.gov/construction/nrc/historical_data/index.html

Data for net electricity imports are from: https://www.energy.ca.gov/data-reports/energy-almanac/california-electricity-data/california-electrical-energy-generation

Monthly AMO, PDO, and ENSO data are from:
https://psl.noaa.gov/data/correlation/amon.us.data
https://www.ncei.noaa.gov/pub/data/cmb/ersst/v5/index/ersst.v5.pdo.dat
https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php

Data for the Vapor Pressure Deficit for Riverside and San Diego Counties, VPD_RS and VPD_SD are from https://prism.oregonstate.edu/explorer