5-2022

A Contribution to the Statistical Analysis of Climate-Wildfire Interaction in Northern California

Adam Diaz
diaz8@clemson.edu

A Contribution to the Statistical Analysis of Climate-Wildfire Interaction in Northern California.

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Mathematical Sciences

---

by
Adam Diaz
May 2022

Accepted by:
Dr. Whitney Huang, Committee Chair
Dr. Brook Russell
Dr. Shyam Ranganathan

Abstract

Wildfires are extreme weather events that exist at the interface of atmospheric, ecological, and human processes. Ongoing anthropogenic climate change is expected to impact the distribution, frequency, and behavior of wildfires on a grand scale, however the exact nature of this change remains shrouded in a great deal of uncertainty. This study takes a statistical approach to the question over the fire-prone Northern California region of the western United states. Climate model projections are analyzed to investigate changes in a major driver of fire weather in the region. The relationship between wildfire severity and climate factors is then explored separately, utilizing a historical data set of California wildfires and climate reanalysis data to analyze the impact of environmental factors on the burned area associated with historical wildfires.

Acknowledgments

The completion of this thesis would not have been possible with the help of my advisor, Dr. Whitney Huang, who has provided fantastic guidance in putting together this study, and in general throughout my time at Clemson. He has been endlessly supportive and patient, and his passion for research has been a source of constant inspiration.

I would also like to thank Dr. Brook Russell for his help and support, both in starting me down the path of research environmental statistics, and continuously since.

A great debt of gratitude is also owed to Dr. Jiali Wang of Argonne National laboratory for both her contributions to this work, and the opportunities that made it possible in the first place. In a similar vein, I would also like to acknowledge Dr. Rao Kotamarthi of ANL, and the Pacific Gas and Electric Corporation.

Finaly, I would like to thank Dr. Shyam Ranganathan for agreeing to serving on this committee, and for his infectious passion and relentless positive energy.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background: Wildfires

Wildfires are unplanned, unwanted, and uncontrolled fires in areas of combustible vegetation [22]. These fires can be ignited by a variety of natural and man-made sources and, once burning, their behavior is dictated by a number of factors including the type, quantity, and moisture of the vegetation available as fuel, atmospheric factors like wind and temperature, topography, and human suppression attempts [36].

These events pose a significant threat to infrastructure and human life, both directly and as a public health risk through associated air and water pollution [69]. Recent years have seen an increasing trend in the frequency, duration, and severity of wildfires in many parts of the world, including the American west, as shown in figure 1.1 [68].

Understanding the physical dynamics that govern the spread and behavior of

h

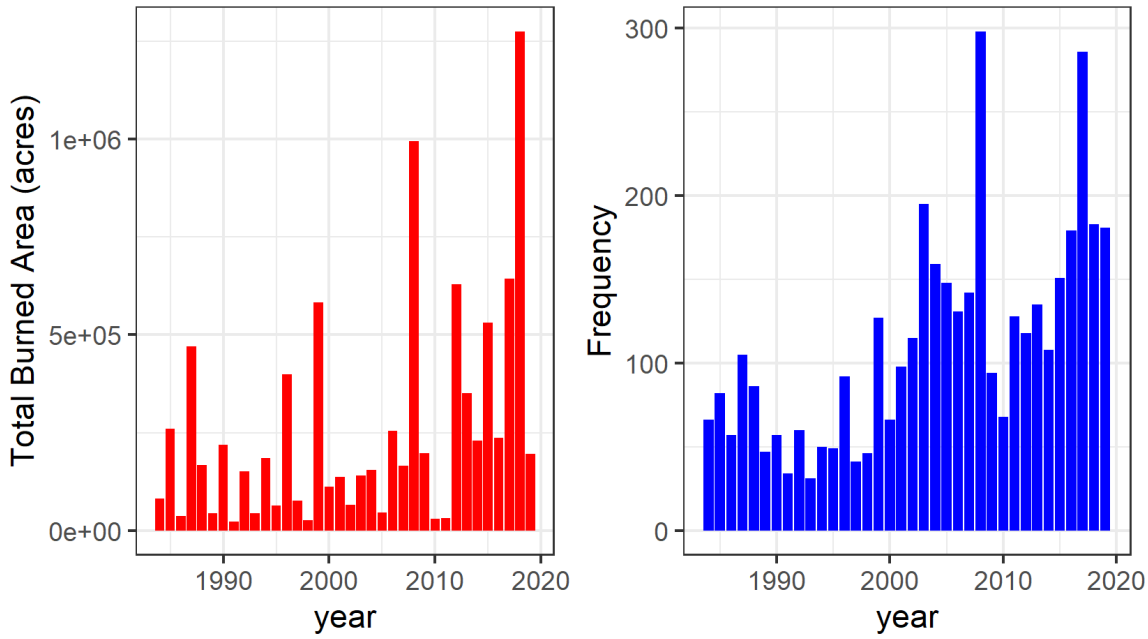

Figure 1.1: California Wildfires by year.

individual wildfires is one of the primary goals of fire science, and an active area of research [65]. These physical properties are the ultimate determinant of fire behavior, and a basic understanding is necessary before any serious modeling effort can take place.

The shape of a fire is complex, and is closely tied to the fire's rate of spread. Fast-moving fires generally have a high area/perimeter ratio [3]. Fires spread primarily along fronts, which are the boundaries between flames and unburned fuel sources. In windy conditions, large fires also spread via a process called spotting. A mass of burning plant matter, called a firebrand, is blown from the main fire into an unburned area, which ignites another small fire. Depending on the surrounding conditions, this spot can burn backwards towards the front, or spread on its own accord. Spotting in

exceptionally large fires has been observed more than five miles away from the front [42].

Starting from the point of ignition, fires can be broadly classified based on fuel type and flame height relative to the surrounding environment [11]. This is a convenient classification system as the mechanisms of spread differ depending on the height of the fire. Flame height is also closely related to the intensity of the fire, and the influence of atmospheric factors on fire behavior is strongest mid-flame [3].

- **Ground fires** are subterranean, and consist of smoldering organic matter in the oxygen-permeable upper layers of soil. These fires are often ignited as a byproduct of larger fires, and can burn for months undetected. [90]
- **Surface fires** are low-temperature fires that burn low-lying vegetation, including grasses and shrubbery. These fires are generally slow-moving, but are easily exacerbated by atmospheric drivers [3].
- **Crown fires** are large enough to ignite fuels in the canopy layer of a forest. Fires at this stage are unpredictable and difficult to control, and produce enough heat to influence local atmospheric conditions [19] [3].

Fuels that are taller than the shrubs and grasses burned by surface fires but shorter than the treetops consumed in crown fires are referred to as ladder fuels. These provide an intermediary fuel source that enables the transition between surface and crown fires, giving the flames a path to the treetops. Ladder fuels often consist of invasive plants, such as kudzu vine in the western US [59].

1.2 Statistical Modeling of Wildfire Behavior

Past approaches to modeling wildland fires can be loosely categorized into physical and empirical paradigms. Physical models represent the interaction between a fire and its environment as a physical system then conduct numerical simulation to analyze specific aspects of behavior [65], while empirical models use data collected from past wildfires to try and understand the behavior of these events more generally [73]. More recent modeling approaches have focused on combining these two frameworks to construct models that are informed by both data and physical principles [54].

A purely statistical approach - and not in the sense of statistical thermodynamics - falls decidedly under the empirical banner, and is distinguished from deterministic empirical models by the use of probabilistic assumptions and emphasis on uncertainty quantification [98]. This approach is particularly useful in problems where data quality and availability are potential issues or where the modeling goal is inherently probabilistic, e.g. spatial estimation of burn probabilities as in [68].

[75] and [98] provide a fairly comprehensive survey of the statistical approach to wildfire modeling, with the former highlighting a number of important historical papers, and the latter giving of the more modern view of the field. Modeling the behavior of specific wildfires - factors like their shape, spread, and intensity - is generally a task better suited for physical models. Statistical modeling strategies are better suited for investigating general trends across multiple wildfires, such the relationship between ignition source in eventual fire behavior as in [9], and can be further subdivided by which aspect of wildland fire behavior that they attempt to capture.

Fire occurrence models estimate the probability of ignition over a spatial do-

main. Popular methods for this sort of modeling include the spatio-temporal point process framework as in [85] and [66], and logistic approximations as in [96].

A more recent area of study is in modeling duration, or time until containment. Wildland fires generally go undetected in the ground fire phase, only attracting attention some time after they emerge as surface fires, which results in observational data on wildfire duration being almost universally left-censored. Other factors like the time between reports and suppression efforts, and the surrounding environmental conditions are also important determinants of burn time. All of these factors can be accounted for using the tools of survival analysis [39], such as proportional hazards regression [46] and accelerated failure time models [91]. Notable examples of work in this area include [4] and [31].

Another modeling direction, and the focus of this work, is the on the size of fires. This can be further subdivided into modeling size at the time of containment using survival analysis methods [84] and modeling total burned area as a proxy for fire size, as in [13] and [76]. We are specifically interested in the latter approach.

1.3 Pyrogeography

The behavior of a fire, no matter how we choose to model it, is ultimately constrained by environmental factors. Wildfire regimes do not exist in a vacuum. Fires perform an important ecological role (DellaSalla and Hanson), with some plants and animals being completely dependent on fires. As an example, the cedar wasp of California exclusively lays eggs in smoldering cedarwood [93]. This broad earth systems perspective of wildfires has been termed pyrogeography [11]. Research in pyrogeog-

raphy is often centered around the large-scale spatial distribution of wildfires and their environmental drivers. This includes anthropogenic factors [14] [17], as well as natural factors related to fuel and atmospheric conditions [44].

This modeling approach typically differs from the aforementioned methods in its scale, generally looking at fires across large spatial regions and temporal spans, and in its aim. The goal of this grand-scale modeling is not to understand the behavior of any particular fire, but to understand the relationship between fire activity and the environment within a given region.

1.4 Climate-Wildfire Interaction

The three environmental factors that are considered most important in controlling a fire's behavior are topography, fuel availability, and atmospheric conditions [67]. The most significant atmospheric conditions in this regard include temperature, humidity, precipitation, and wind.

Wind is one of the primary climate factors that drives the spread of wildland fires. In addition to physically pushing the fire along a course and transporting burning material to unburned fuel, strong winds transfer heat energy from a fire to potential fuel sources. This accelerates the evaporation of moisture within the fuel, and supplies the oxygen required to maintain the combustion reaction. Wind can also confound attempts to contain a fire, as workers may not be prepared for sudden shifts in atmospheric conditions. Topography can exacerbate all of these effects, which can lead to fires growing rapidly out of control. [45, 50].

Unlike wind conditions, the effects of temperature and precipitation are more dependent on long-term trends. Periods of prolonged drought create a massive stockpile of flammable material, which greatly increases the risk of fires getting out of hand.

1.5 Extreme Fires

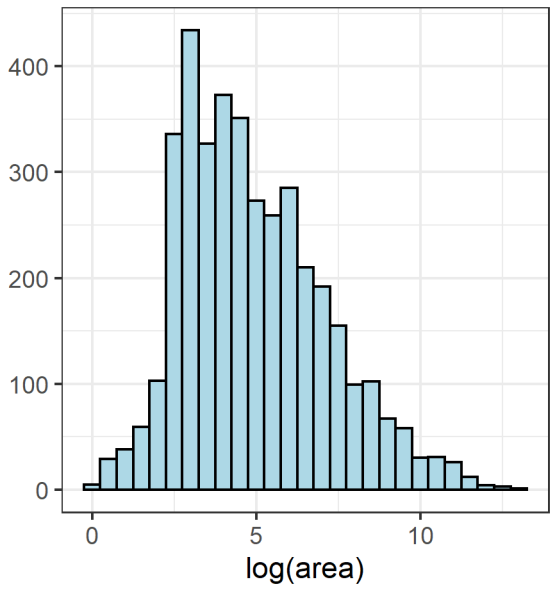

Figure 1.2: Log distribution of burned area: Northern California fires, 1984-2019

One common factor across most threads of statistical fire modeling research is the issue of extremes. The distributions of burned area heavy tailed and exhibits power law behavior, as shown in figure 1.2 [25]. This often necessitates the application of methods from extreme value theory [21] to accurately capture the tail behavior any of these quantities. [26] [76]

This empirical power law behavior suggests that the growth of a fire is self-

reinforcing. Research into the physical dynamics of extremely large fires [20] suggests that this is indeed the case, and offers an explanation. When the heat generated by a fire reaches a certain point it begins to generate convection currents in the surrounding atmosphere, which can in turn lead to extremely strong winds being generated by the fire itself. Fires at this scale are sometimes referenced as crown-convective fires, and spread rapidly across the canopy layer of a forest [18]. Fuel moisture is also less of a concern for fires on these scales, since any unburnt fuel is rapidly dessicated by the extreme heat and wind conditions generated by these megafires.

All of this indicates that extremely large fires have a different relationship with their environment than fires that haven't achieved such a disastrous scale. [20] highlights the difficulties faced in modeling these crown fires from both a physical and statistical perspective. Physical models often fail to account for these convection currents, and these extremely localized effects are not factored into the large-scale climate data used in statistical modeling.

1.6 California's Fire Regime and the Effects of Climate Change

California is the most populous US state, and one of the world's major economic centers. California's 2020 regional GDP was over 3 trillion dollars. If California were a sovereign nation, this would amount to the world's 5th largest economy. It is also one of the most geographically diverse regions of the country, and the most affected by wildfire activity.

The 2020 California fire season saw 9,917 fires burn over four million acres of land, making it the most destructive fire season for the state in over 200 years [15].The 2021 season resulted in the destruction of over 2.5 million acres.

The scale of the 2020 and 2021 fire seasons is indicative of an increasing trend in the severity of California's fire seasons, a trend which has been related to anthropogenic climate change [95]. This trend is expected to continue into the future as these effects become more significant, both in California and in other fire-prone parts of the world [99, 87]. The relationship between wildfires and climate drivers is an area of active research, as are the mechanisms through which climate change could influence this relationship [53]. One well-studied effect of climate change that might impact the increasing trend in fire severity is a general decrease in relative humidity levels over land [94]. California features several regions with high fuel availability and high levels of fuel moisture, for which this kind of climate shift poses a particularly severe risk [60].

1.7 Background: Extreme Value Theory in Brief

Traditional methods of statistical inference often focus on "what usually happens" - building sophisticated estimates of means, variances, and other measures of central tendency. In many problems that arise in the sciences, and particularly in engineering, these methods are effectively useless. Skyscrapers aren't built to survive an average windy day, they need to stand up against the strongest winds that are ever going to batter them.

Let $Y_1, ..., Y_n$ be iid random variables with mean μ, finite variance σ^2, and CDF $F_Y(y)$. The central limit theorem tells us that

$$\lim_{n\to\infty} \sqrt{n}(\bar{Y} - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2),$$

where $\bar{Y}$ is the sample mean. This is a very powerful tool for performing inference on the mean, but what if we're interested in something along the liens of $Y_{(n)} = \max(Y_1, ..., Y_n)$? Luckily there exists an analogous results for the asymptotic distribution of sample maxima. First note that

$$P(Y_{(n)} < x) = P(Y_1 < x) \cdot P(Y_2 < x) \cdots P(Y_n < x) = (F_Y(y))^n$$

then the Fisher–Tippett–Gnedenko theorem states that

$$\lim_{n\to\infty} F_Y^n(a_n y + b_n) \xrightarrow{d} \mathcal{G}(y)$$

Where a_n and b_n are normalizing constants, and $\mathcal{G}(y)$ is the generalized extreme value (GEV) distribution

$$\mathcal{G}(x \mid \xi, \mu, \sigma) = \exp\left(-\left(1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi}\right),$$

with location parameter μ, scale parameter σ, and shape parameter ξ [21].

To apply this result and estimate the distribution of extremes we start with a random sample $X_1, ..., X_n$, then partition that sample into k 'blocks' $B_1, ..., B_k$ with $k < n$. In many applied cases we're looking at time series data, and blocks often corresponding to months or years. We then consider the maximum value in each

block, and use this information to fit a GEV model. This is often termed the block maxima approach.

Another strategy for modeling extreme values is the peaks over threshold method. Rather than considering $P(X_{(n)} < x)$, we fix a threshold μ and condition on the exceedence of that threshold: $P(X < x|X > \mu)$. Asymptotically, this results in a distribution in the generalized pareto (GP) family, with the following CDF:

$$\mathcal{G}(x) = \begin{cases} 1 - \left(1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi} & \xi \neq 0 \\ 1 - \exp\left(\frac{-(x-\mu)}{\sigma}\right) & \xi = 0 \end{cases}$$

Peaks over threshold methods offer a number of advantages over block maxima [8] , namely that a larger number of extreme events can be considered. The main problem with peaks over threshold is choosing a good threshold. There's a bias-variance trade-off: Too high of a threshold and the sample size is too small to conduct sound inference. Too low and we run into convergence issues.

The GPD and GEV are members of the location-scale family, and exhibit some "nice" stability properties. If X is a collection iid GEV random variables then $\max(X)$ is a GEV random variable with the same shape parameter and shifted location and scale. Similarly, conditioning a collection of iid GP random variables on the exceedence of a threshold results in a GP random variable with the same shape parameter.

Chapter 2

The Diablo Winds of Northern California

There are several wind patterns over California that have long been associated with wildfires in the state. The most notable of these winds are generated by air from the high pressure system over the great basin interacting with the topography of the region, eventually manifesting as a hot, dry northeasterly wind. The winds generated by this system are given different names depending on where they occur and include the El Norte winds in Northwestern Mexico, Santa Anas in southern California, and Diablo winds in northern California [1].

The extremely low humidity of Diablo wind (DW) events, coupled with their tendency to blow downslope, create atmospheric conditions that accelerate the growth of fires. In 1923, dry northeasterly winds were cited as the driving force behind a fire that raged through the city of Berkeley, incinerating over 600 structures near the UC Berkeley campus [2]. In 1991, dry northeasterly winds with gusts in excess of 65 mph fanned a hillside grass fire into an out of control conflagration that spread into the

Oakland urban area. This so-called "tunnel fire" resulted in 25 deaths and over 1.5 billion dollars in damage to infrastructure.

Diablo winds are shaped and driven by the topography of the region. The strong northeasterly winds pass over the Sierra Nevada range in eastern California and down its western slopes, heating up and losing humidity before passing through the central valley area where high inland pressure drives the air mass towards the sea. The air is then forced over the rugged coastal range before sinking down to sea level on the western slopes, where the compression and loss of moisture produces intense, dry, downslope winds. This physical behavior is in contrast to the Santa Ana winds in the southern region of the state which are primarily driven by gravity rather than the action of sinking air masses, and usually reach their peak intensity passing through canyons rather than downslope [19].

The Diablo winds have been a subject of much recent study. [56] and [27] conducted case studies of specific fire events, and noted the role of these strong downslope winds. [57] looked at the climatology of the Diablo winds and identified factors that are associated with the frequency of these events. [80] analyzed the wind pattern based on station data and concluded that the events tend to display a low gust factor.

[1] studied downslope winds in a more general context, and found an association between these events and the southern Pacific oscillation on a global scale. [50] used reanalysis data to look for trends in these wind events over a 40 year historical period and found no trend in frequency, but a downward trend in the relative humidity associated with the Diablo winds. [10] used numerical simulation data to investigate the synoptic trends associated with these events.

[49] used one single global climate model Community Atmospheric Model (CAM) projections to look at potential end of century (2106-2115) Diablo and Santa Ana wind activity under a number of potential warming conditions, with the goal of evaluating future fire risk. Their results indicated an increase in the frequency of Diablo and Santa Ana wind events under more moderate warming scenarios - a global mean temperature increase of 1.5°C or 2°C over the next century - while the most extreme scenario they considered (GMT +3°C) shows a decrease in the frequency of these wind events. They postulate that this reversal of trend is caused by the non-linear response of the underlying pressure systems to changes in temperature.

Wind speed, one of the primary climate factors associated with very large, infrequent fires, are expected to slow down in general ([12, 70]), however climate models have indicated an increasing risk of these massive, out of control blazes [92]. Our investigation of Diablo winds is partially motivated by a desire to reconcile these seemingly paradoxical results. By examining the possible changes in these wind patterns, we can glean some insight into the climate factors that drive these projections of worsening fire seasons.

In this work we consider the output of multiple high resolution regional climate models (RCMs, see Sec. 2.1) to study changes in the frequency of Diablo wind events and in the associated wind speeds, event duration, and humidity, as well as changes in the upper atmospheric pressure systems that generate these winds. The use of multiple RCMs also allows us to account for the intermodel uncertainty of climate change projection.

2.1 Dataset

2.1.1 Model simulations

The diablo winds were extracted for historical (1995–2004), and end of century periods (2085–2094) under representative concentration pathway (RCP) 8.5, a pathway that assumes very high levels of greenhouse gas emissions by 2100 with an effective radiative forcing increase of 8.5 W/m^2 due to large populations and little technology improvement [72]. The input data for defining the diablo wind (see Ch. 2.2) were obtained from Weather Research and Forecast model [WRF, 79]—at a spatial resolution of 12 km [88], driven by three earth systems models (ESMs). They are Community Climate System Model version 4 [CCSM4, 35], Geophysical Fluid Dynamics Laboratory Earth System Model with generalized ocean layer dynamics component [GFDL-ESM2G, 28], and the Hadley Centre Global Environment model, version 2-earth system [HadGEM2-ES, 40]. These three ESMs represent high (HadGEM2-ES), median (CCSM4), and low sensitivity (GFDL-ESM2G), respectively, of global average air temperature response to the doubling of CO2 [78]. See [104] for detail evaluation of the three WRF simulations. The WRF simulations driven by the three ESMs are named WRF_CCSM, WRF_GFDL, and WRF_HadGEM, respectively, in this study.

2.1.2 Historical fire data

Our historical fire data is derived from satellite imagery, and contains information on the date, location, and burned area associated with Northern California wildland fires between the years 1984 and 2019 that burned a land area in excess of 1000 acres. This data set is derived from satellite remote sensing data from the Monitoring Trends in Burn Severity Project (MTBS; [29]). MTBS includes fires larger

than 1,000 acres in the western CONUS;

2.1.3 Reanalysis climate data

For information on the climate conditions surrounding each fire we turned to the ERA5-LAND reanalysis data set [38]. This gridded data set is laid out on a $0.1^\circ \times 0.1^\circ$ grid, with variables reported at an hourly temporal resolution. The identification of Diablo wind events within the ERA5 data consisted of the same algorithm, however it did require some slight adjustment to the criteria. In keeping with [80] we adjusted our wind speed threshold down from 8 m/s to 5 m/s in order to account for the lower average wind speeds and coarser spatial resolution of the ERA5 data.

2.2 Diablo Wind Identification

The first step in our analysis was finding a set of criteria to distinguish Diablo winds from other atmospheric events, and constraining our search to a geographic region surrounding Northern California, as the winds of interest are endemic to this region. The data used for event identification purposes consisted of 2 meter relative humidity and 10 meter zonal and meridional wind components on a grid spacing of 12 km with 3 hourly temporal resolution. Diablo winds are characterized by their intensity, long duration, low moisture content, and northeasterly direction [10]. We determined that a Diablo wind had occurred in a given grid cell at a certain time step if all of the following conditions were met:

- Northeasterly direction: Meteorological wind direction between $0 and 90$
- Intensity: Wind speeds exceeding 8m/s

- Relative humidity below 25%
- Duration: These conditions must be present for a minimum of 2 consecutive time steps (6 hours) within the same grid cell.

The choice of threshold was based on previous work by (cite PG&E). [49] and [80] defined these wind events using a higher threshold for relative humidity and allowed for more northerly winds, in addition to requiring a minimum of a 24 hour gap between events within the same grid cell. This gap between events is excluded from our criteria.

To assess the sensitivity of the identified Diablo winds to these thresholds, sensitivity analysis to threshold selection was conducted by replicating our analysis across a number of alternative thresholds for Diablo wind identification. For example, relative humidity thresholds ranging from 10 to 35 percent, wind speeds ranging from 5 to 10 m/s, and the introduction of an inter-event time criteria in the vein of [49] and [80].

To understand the synoptic processes driving the phenomenon and its changes under the future scenario, we examined the 500m geopotential height over the Northwest hemisphere on days corresponding to DW activity. These events are associated with a high pressure ridge that circulates over the Great Basin, and they travel along the steep pressure gradient between this inverted trough and the lower pressure over the eastern Pacific. With all other factors held constant, any change in this pressure gradient should have a major impact on the frequency and behavior of DW events.

2.2.1 Event Identification Algorithm

The algorithm used to identify Diablo wind events and compute counts and summary statistics is as follows:

Algorithm 1

Inputs: 3 hourly relative humidity (rh) and UV wind vectors on a $n \times m$ grid recorded at T time steps.

1. Store the input as vectors indexed according to their spatial location s and time point t. The spatial locations correspond to grid cells.

2. Calculate wind speed and meteorological wind direction from u and v components. Filter out all entries that don't meet the wind speed, direction, and humidity criteria.

3. Create a list with an entry for each remaining location, populate it with the remaining entries plus their original time index. Loop over the list and identify arithmetic progressions in the index set with common difference 1 and length at least two. Each of these progressions corresponds to a DW event starting at the initial time index, lasting a number of time steps equal to the length of the progression.

4. Remove every entry with an index that does not belong to one of these progressions. Average the RH and wind speed across the entries of every remaining sequence, record its length as the event duration, and retain the time index of the first step.

Outputs:

- The number of rows in each list entry corresponds to the event count in that cell.
- Event-specific wind speed, relative humidity, and duration are all contained in the entries matched with each cell.

2.3 Results

2.3.1 Historical Association between Wildfires and Diablo winds

This analysis aims to establish a historical association between fires and Diablo winds, a relationship that underscores the importance of studying this atmospheric phenomenon and motivates our work with climate projections.

As discussed in the 2.1 section, this analysis is carried out using historical wildfire observations and ERA5 reanalysis climate data. The fire data does not include any information on burn duration or on the start/end of each event. Since a DW event could influence a fire at any point during the fire's life cycle, we designated a fire as being associated with a DW event if such an event was indicated in the reanalysis data within two weeks of the date associated with the fire with the same 0.1×0.1 degree grid cell.

We found that, while Diablo winds are not directly associated with the largest fires in Northern California, they are very clearly associated with the largest fires that occur in the vicinity of the heavily populated San Francisco Bay region. If we relax the duration criteria to 3 consecutive hours and allow our winds to vary between 340 and 110 degrees, then the pattern becomes even more apparent.

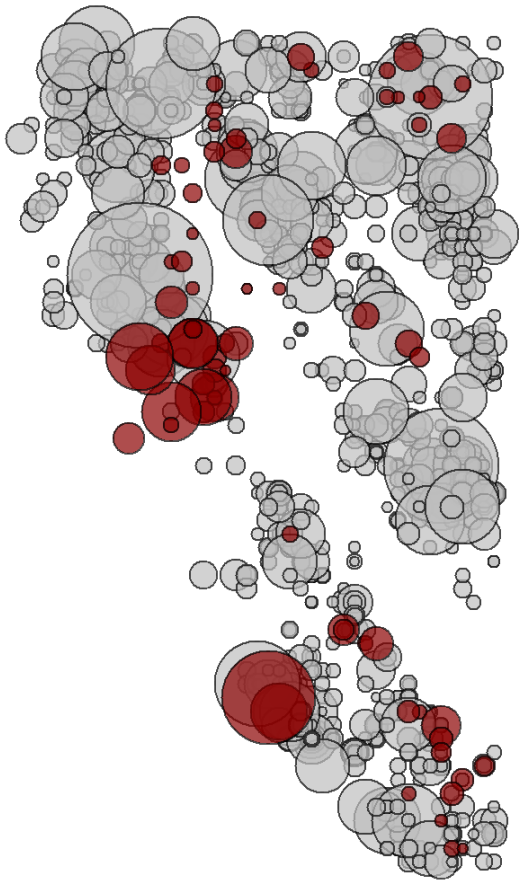

Figure 2.1: Northern CA wildfires: 1984-2019

Figure 2.1 shows the location every Northern California wildfire between 1984 and 2019 that burned over one thousand acres of land. The points are scaled by fire size, with larger points representing larger fires. Points that are colored red indicate that the fire was associated with one or more Diablo wind events.

This spatial relationship can be explained by the direction of the winds relative to the topography of the region. After passing over the coastal ranges. The western slopes of the mountain range are steep and largely W/SW facing, which results in the winds blowing downslope and parallel to the aspect of the terrain. This is by far the most populous area within the study region and the most affected by any change

in DW activity, so it will be of particular concern in analyzing our results.

2.3.2 Future change in Diablo winds

Now we turn our attention to the regional climate simulation and their future projections. The first statistic we consider is the frequency of Diablo wind events within the region, summarized in Fig. 2.2 below:

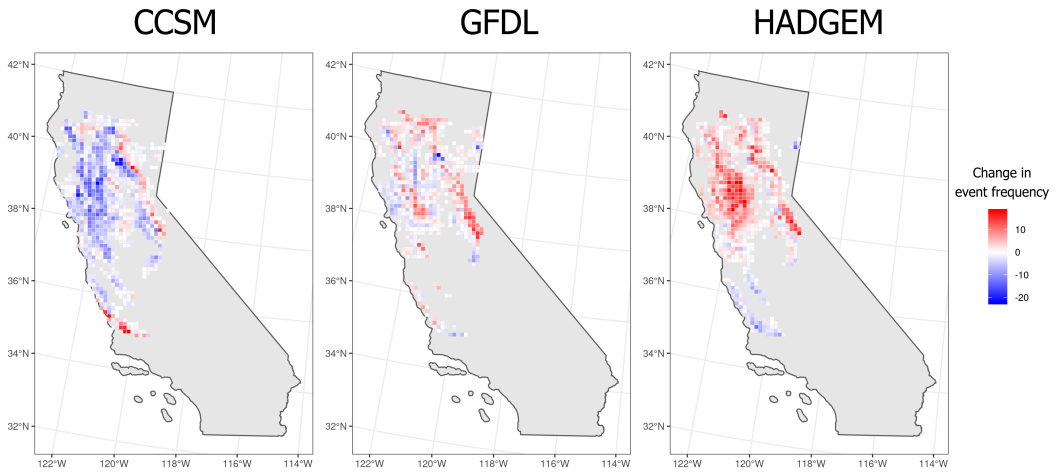

Figure 2.2: Difference in event frequency: EoC - Historical

HADGEM and GFDL both indicate a general pattern of increasing frequency, particularly in the western central valley and north bay regions under HADGEM. GFDL shows increases along the northern coastal and sierra nevada mountain ranges, but a general trend of decreasing frequency almost everywhere else.

Conversely, the CCSM model shows a pronounced trend of decreasing event frequency, but with a few notable areas that show either an increase or relatively little change. This downward trend is is less pronounced in coastal areas, and along the sierra nevada range, and we can also see a small region of increased frequency north

of Santa Barbara.

Frequency is not the only important consideration regarding the future behavior of these wind events. The relationship between these events and local fire regimes is complex, and varies across space and time. Other factors that make Diablo winds a major contributor to fire weather include their long duration, high sustained wind speeds, and the low humidity of the air they carry. It is only natural that we consider how these qualities might change over time.

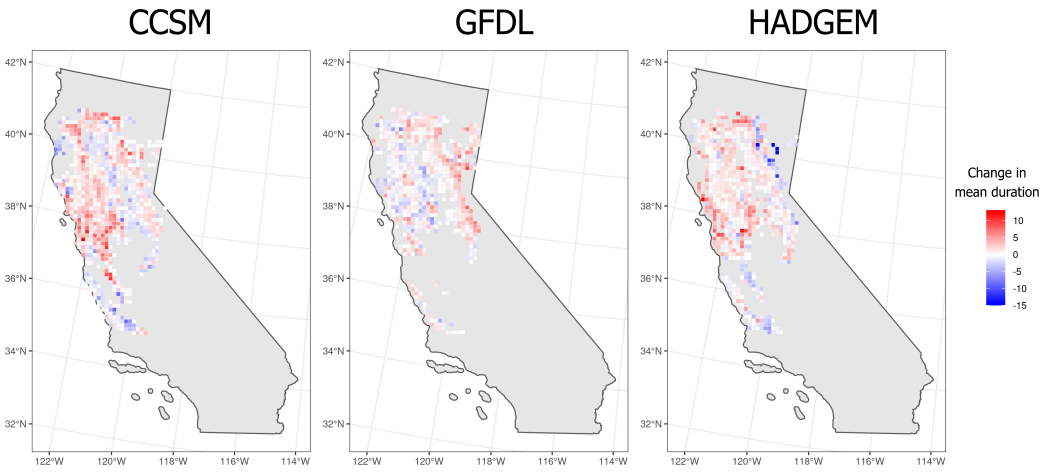

Figure 2.3: Difference in average event duration (hours)

In Fig.2.3 we see the projected change in average event duration in each grid cell. All three models show an increase in the average duration of events in the vicinity of the highly populated San Francisco bay area, with the CCSM and HADGEM models indicating a more extreme shift that extends over large sections of the coast and adjacent mountain ranges. GFDL shows the least overall change in mean duration, and the spatial trend is somewhat reversed from what we see in the other two models, with durations increasing in the east and decreasing towards the west.

The duration of these winds is particularly notable with regards to their effect on fire weather. In addition to their role as a driver of wildland fires, persistent heavy winds limit the scope and efficacy of wildland fire containment strategies. Aerial suppression strategies are particularly affected. The operation of aircraft in these conditions is hazardous. The efficacy of aerial dispersal of flame retardant chemicals is greatly reduced, and is considered ineffective in conditions where wind speeds exceed 9m/s. Backfiring - the practice of setting small fires along the edge of a fireline in order to influence the direction of a larger fire - is also far riskier under these conditions. [63].

2.3.3 Atmospheric conditions associated with Diablo wind events

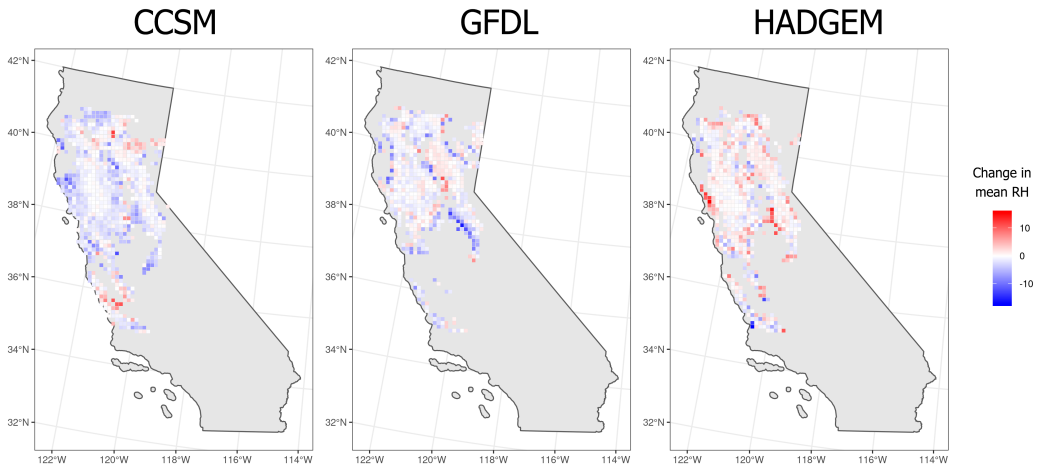

Figure 2.4: Difference in average event-associated 2m relative humidity (%)

The CCSM model shows a downward trend in relative humidity associated with Diablo winds, which is in line with the findings of [49] in their analysis of CAM output. CCSM and GFDL show a fairly similar spatial distribution of differences, although CCSM depicts a more extreme drying trend along the coast north of Big Sur. Of the three models, GFDL indicates the lowest event-associated humidity across both time periods, so the large areas of very little change still represent very dry events in general.

The HADGEM model appears to be a notable outlier in 2.4, showing an increase in the relative humidity of DW events along the pacific coast, particularly near the bay area. It is important to note that differences in the relative humidity associated with these events could only be computed for grid cells where DW events occurred in both the historical and end of century periods. More detailed maps showing each time period separately are available as supplementary materials. The HADGEM projections for both periods are presented in 2.5 below.

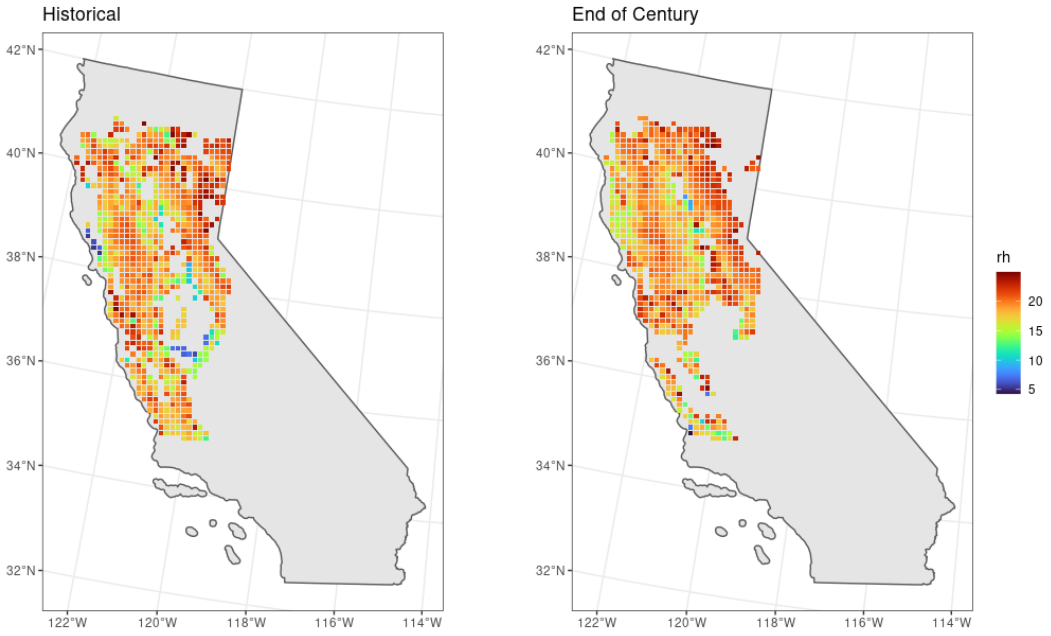

Figure 2.5: Event-associated relative humidity (%) for HADGEM projections

The humidity associated with events in the north bay region was significantly lower in the historical HADGEM projections compared to the other two models. In addition, HADGEM showed almost no events in the heavily forested Mendocino county north of the bay area.

The end of century projections for event-associated relative humidity in the north bay area are still low compared to most of the study area, and is roughly in line with the end of century projections from the GFDL and CCSM models (see supplementary materials). In addition, we see a number of events affecting cells in the northern coastal region that were not affected in the historical period, for which the associated mean relative humidity generally falls below the 15% mark.

2.3.3.1 Wind Speed

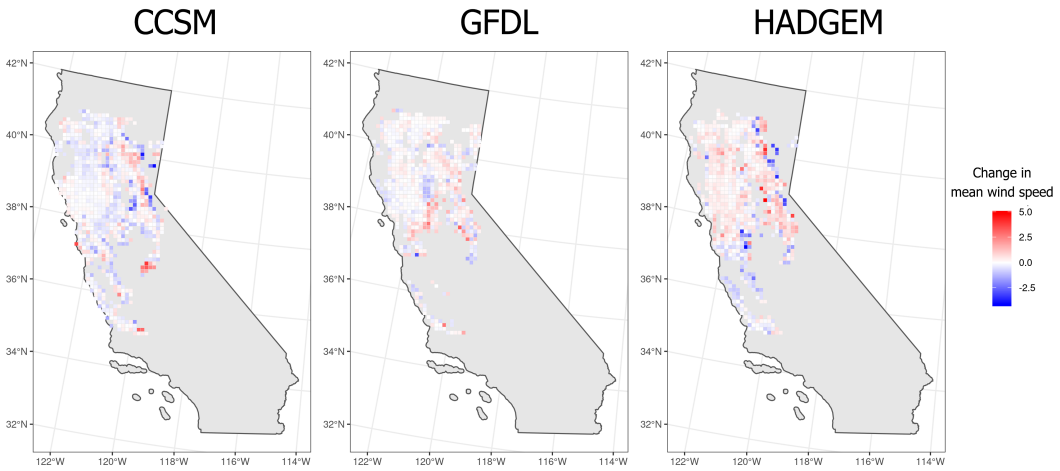

Figure 2.6: Difference in average event-associated 10m wind speed (m/s)

Fig. 2.6 shows the change in average wind speed associated with DW events. The threshold wind speed of 8 m/s represents a fairly major anomaly compared to usual atmospheric conditions, and most models show fairly little change between the two periods.

Of the three projections, HADGEM shows the most pronounced increase in mean wind speed in the vicinity of the bay area and capital region, while CCSM shows relatively litle change in the region. CCSM shows a weakening trend over the sierra Nevada range, while HADGEM displays a lot of variability over this area. GFDL depicts fairly consistent wind speeds across both decade, with a slight upward trend in the central valley region and very little change elsewhere.

2.3.3.2 Geopotential height at 500hPa

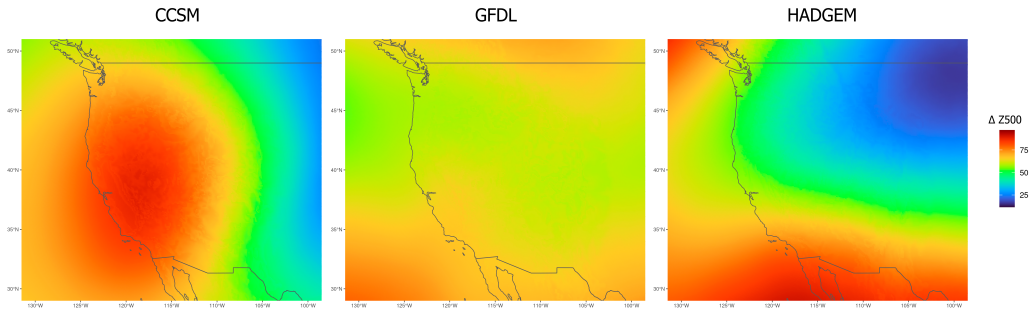

Figure 2.7: Projected difference in event-associated 500hPa geopotential height (m)

All three models project an increase - both in general, and in association with DW events as shown in figure 6. This is consistent with previous findings on the subject [16] and with the widely acknowledged upward trend in global mean temperature.

The GFDL model is the outlier here, indicating very little change in the pressure gradient across the region of interest, however it does show a large and uniform increase in the height of pressure systems.

HADGEM and CCSM project a dramatic difference in the distribution of pressure systems on days that correspond to DW wind activity, showing lower pressure over the continent and higher pressure over the coast. The biggest point of contention between the two models is over the western end of the study area, where the CCSM projection indicates an increase in pressure zone that extends well into the eastern pacific.

This change in the pressure gradient that we observe in CCSM and HADGEM but not GFDL offers us some insight into why these models differ in their DW pro-

jections.

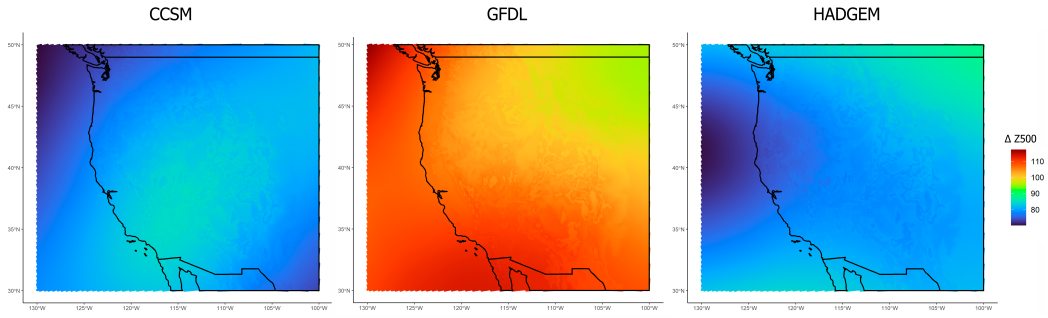

Figure 2.8: Projected difference in average 500hPa geopotential height (m)

Figure 2.8 shows the difference in geopotential height between the periods averaged across the entire decade, giving us an idea of the projected change in the spatial distribution of upper atmospheric pressure systems outside of DW events. CCSM and HADGEM are extremely similar in this regard, showing a relative increase in pressure over the continent compared to offshore levels.

The higher offshore pressure during DW events in the HadGem model indicates a high pressure airmass moving offshore towards the S/SW, which is consistent with the behavior of modern Diablo winds. The intensification of this pattern in the HADGEM model may account for the increased frequency of winds in its end of century projections.

GFDL shows the lowest level of change in the pressure gradient of the three models models, however it does feature the most extreme upward shifts in overall pressure. This large, uniform increase in pressure and relatively little apparent change in existing pressure systems.

The CCSM model predicts a sharp drop in DW event frequency across most of the study region, coupled with a significant uptick in average event duration on the west coast of the state. The upper atmospheric conditions associated with DW events under this model show significant high pressure lingering over the great basin, which lines up with the projection of less frequent and significantly longer lasting wind events. This changes could indicate a stronger but slower moving pressure system driving the DW events.

2.4 Summary and Discussion

In this study we examined the potential effect of anthropogenic climate change on the incidence and characteristics of these wind events. There are several points of consensus between the three models. We see an increase in DW-like events on the eastern side of the state over the Sierra Nevada mountain range, an increase in mean event duration near the coastal range and bay area, and relatively little change in event-associated wind speed and moisture over the central valley. There are also a number of noteworthy differences.

Previous work in fire science [83, 37] has indicated that certain fire regimes in the Western US may experience a shift towards less frequent, more intense fires. The CCSM model suggests a similar shift in the behavior of the Diablo winds. The CCSM results agree with previous work by [50] under different global mean temperature conditions, which was conducted using the atmospheric component of the CCSM model.

The GFDL model shows the least overall change between the two periods in terms of both DW events and the associated ridge. This model also depicted these wind conditions affecting a smaller spatial region, and with lower frequencies across both time periods. The central valley region sees a notable increase in the frequency and severity of DW events under this projection. This is a fairly positive outlook, as these wind events appear to be less influential over fire weather in this region.

Of the three models, HADGEM to indicate an increase in the frequency, duration, and intensity of DW events in the area surrounding San Francisco bay. It does indicate an upward trend in the relative humidity of the winds, however that can largely be explained by the exceptionally dry historical projections. This projection also indicates a pronounced increase in frequency spanning most of the study region, with the exception of the most southerly areas under consideration. This model also shows the most pronounced locational shift in the distribution of DW events, with the end of century projections indicating far fewer events in the eastern and southern portions of the study area compared to the historical period.

The results of our analysis on historical fire data indicate that these wind are most associated with fires that occur on the western slopes of the Coastal mountain ranges, which is precisely what the fire science literature would suggest. With this in mind, the HADGEM and CCSM projections present far worse scenarios in terms of DW-associated fire weather.

Chapter 3

Analysis of Historical California Fires

We are interested in modeling the distribution of burned area associated with fires over a large and geographically diverse spatial domain. Each observation of burned area y_i is associated with some spatial point s_i and a set of covariates $\boldsymbol{x}_i = (x_{1,i}, \cdots, x_{p,i})^T$. The goal of this modeling exercise is to study the empirical relationship between the response y and the covariates $\boldsymbol{x}$, while accounting for spatial dependence via the inclusion of a random effect term [5], with particular attention to how this might differ towards the upper tail of the distribution.

Spatial modeling of large-scale wildfire behavior is complicated by the issue of non-stationarity in both mean and covariance [43], [48]. Previous attempts to account for this non-stationary behavior have included the application of weighted regression models that avoid explicit modeling of the covariance [62], and the use of Gaussian random field models with non-stationary covariance functions [102].

3.1 Data

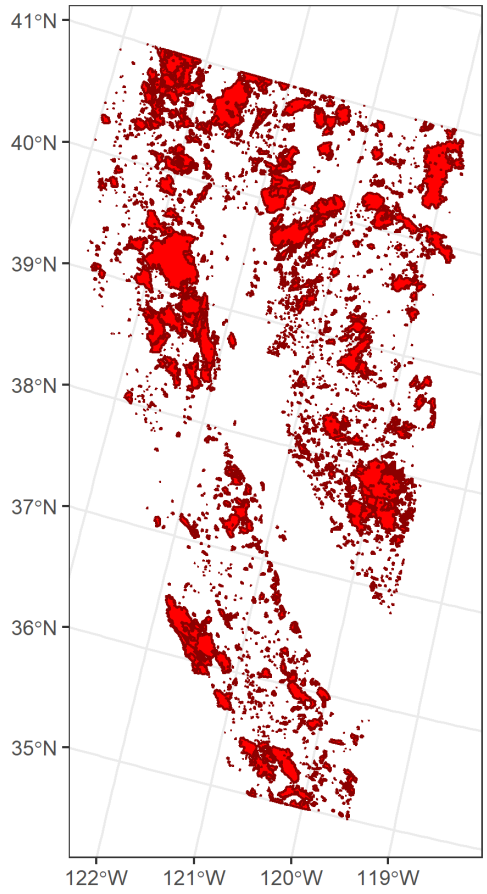

Figure 3.1: Area Burned by Northern California Wildfires

The data used in this analysis differs slightly from the historical data used in chapter 2. The fire observations are taken from a publically available dataset provided by the state of California through their Fire and Resource Assesment program. The data ranges from 1984 to 2019, and the fields that we utilize in this analysis include burned area, date of discovery, and a shapefile indicating the extent of the burned area, as shown in figure [15].

Lower atmospheric covariates are derived from a finer scale ERA5 reanalysis

data product, available at a 0.1×0.1 degree spatial resolution, and hourly temporal resolution, while upper atmospheric covariates are available at a 0.25×0.25 spatial resolution, again with hourly temporal resolution (Copernicus). Topographic data is derived from digital elevation models retrieved from Amazon web services open terrain project (AWS).

One data quality issue of note in the fire data is a very large number of missing values regarding date of containment. This makes it difficult to gauge the duration of a fire or construct covariates that range across a fire's lifespan. The gridded atmospheric data is at a much coarser resolution than the polygonal fire observations, and the covariates associated with any polygon that extends across multiple grid cells are averaged across those cells.

3.2 Covariate Construction

A common issue that arises in dealing with atmospheric covariates is that of multicollinearity [51]. Much of the previous literature on regression analysis of wildfire behavior has employed dozens of covariates in the interest of more accurate prediction [89], [102], [30], however our interest is more in understanding the relationship between these variables and the size of fires, which makes this a highly relevant issue. Over thirty covariates were considered for inclusion, but the majority of these were eliminated due to strong linear relationships with one or more other factors.

Another important consideration is the process of summarising gridded spatio-temporal data into covariate vectors. The fire data set contains information on the shape and extent of each fire rather than point locations, which simplifies the process

of aggregating across space. If a fire extends across more than one covariate grid cell, we consider a weighted average based on the area of intersection with each cell relative to the total area of the fire.

The more difficult problem is aggregating these variables across time. We have the dates and times of discovery and containment for each fire event. This interval is likely to be left-censored, particularly for fires that occur further from densely populated areas, and there are some potential issues with data quality. Several entries were manually corrected by referencing the names and identification numbers in official records, but verifying the information for every event is not feasible.

Different factors are also likely to have a greater impact at different points in the fire's life cycle. Precipitation, for example, is most impactful in terms of a long-run trend preceding the actual fire (find the citation). Wind, on the other hand, is most meaningful during the surface and crown fire phases (that one 70s book). The longer duration of large fires also has a smoothing effect on time-varying predictors, since they have much more room to vary across the lifespan. A selection of temporal averages were considered for each covariate to account for these effects, and final selections were made based on fire science literature, and on their empirical relationship with the response.

A final consideration in covariate selection is that the relationship between covariates changes across space, time, and the quantiles of the response. This leads to a sort of Simpson's paradox, where correlations change at different scales of spatial and temporal grouping. In an attempt to account for this effect across time, multicollinearity was assessed separately with the data separated into seasonal, monthly,

annual, and five-year groups. The spatial effect is handled by the inclusion of spatial covariates related to fuel models and vegetation coverage.

3.2.1 Surface Atmospheric Variables

Total precipitation, 10m wind speed, 2m temperature, 2m relative humidity, 10 hPa geopotential height, surface pressure, and surface convective potential were considered as lower atmospheric covariates. These were derived from ERA5 reanalysis data with precipitation, wind speed, temperature, and RH available at a 0.1×0.1 degree spatial resolution, the pressure variables at a 0.25×0.25 degree resolution, and both at an hourly temporal resolution. 10m wind speed, and precipitation were selected for inclusion in the final model.

Precipitation was averaged over a 60-day period prior to the start of the fire. The averages across this period were highly correlated with levels during fire's lifespan. The fire science literature suggests that temperature is most impactful as a long-run trend, and during the initial phases of a fire's life-cycle (that 70s book again). Wind speed was averaged over the lifespan of the fire from discovery date to containment date.

The other variables were excluded due to collinearity. Relative humidity was heavily correlated with precipitation and temperature, particularly after adjusting for seasonal and spatial effects. It also showed a more linear relationship with the upper atmospheric covariates than long-run temperature or precipitation, and was removed. The pressure-related variables were all removed based on linear relationships with topographical covariates and upper atmospheric variables. Temperature was found to

be negatively correlated long-term precipitation trends and with wind speed.

3.2.2 Ecological and Topographic

Soil moisture level, slope, high vegetation coverage, low vegetation coverage, terrain aspect, NFDRS fuel load, and the angle of average wind direction relative to aspect were all considered. The topographic covariates were derived from the USGS national map 3d elevation program, and give some indication of the terrain underlying the wildfire event. Mean slope gives a rough measure of steepness, while mean aspect gives the direction of the slope. The incident of wind and aspect was computed based on the difference in angle of meteorological wind direction and terrain aspect on a 0-1 scale, with 0 indicating upslope winds and 1 indicating downslope. Vegetation coverage was derived from ERA-5 reanalysis, and fuel model classification was drawn from USFS wildfire assessment system data.

Fuel classification, wind-aspect incidence, and high/low vegetation were ultimately selected as covariates. Slope was tightly correlated with wind speed and vegetation coverage, while soil moisture was explained by long-term trends in temperature and precipitation. Aspect is only important in its relation to wind direction, particularly once spatial relationships are accounted for, so this was also excluded.

3.2.3 Upper Atmospheric

The upper atmospheric covariates included in the model are geopotential height and potential vorticity measured at 500hPa, both averaged across the lifespan of the fire at a daily temporal resolution and a 0.25×0.25 degree spatial resolution.

These effects are hypothesized to be particularly relevant to larger fires ([103]. Potential vorticity (PV) is effectively a measure of the air's ability to move at a particular pressure level, in this case 500 hPa, which gives us some idea of how heavily convection currents can churn over a large fire. Geopotential height gives us an indication of both the pressure system directly overlying the fire, and of the vertical flame height required to take advantage of the free-moving air in high-PV conditions [77].

3.3 Exploratory Analysis

The heavy-tailed nature of the burned area distribution is immediately apparent - see 2.1, with the largest fires accounting for the vast majority of burned area. The data set contains records of 280 fires that burned in excess of 5000 acres, accounting for roughly 7% of the observations and over 84% of the total burned area.

A seasonal trend with regards to both frequency and size of fires is also present, with the vast majority of fires starting in the summer months, or in late spring and early fall. No fires burning in excess of 5000 acres began during winter months. It is also clear that fires that start outside of California's fire season - roughly April through October - show some very different behavior, and account for less than 10% of the data. In the interest of getting clear results, these have been excluded from the analysis. Covariates and their relationships also change within the fire season. Potential vorticity and Z500 appear more heavily correlated in fires that started outside of the autumn months, with the extent of that correlation experiencing a reasonable amount of inter-annual fluctuation.

One area of interest is in how the covariates appear to differ when we separate out the extreme fire observations from the rest.

3.4 Model Specification

We're interested in evaluating a relationship that varies across space, so a spatially varying regression model is an appropriate choice. The most interesting quantities within the scope of this analysis are the regression coefficients, and particularly how they vary in the tail as opposed to the bulk of the distribution. In this vein, we'll consider two separate models - one model examining the upper end of the burned area distribution, and one fit to the remaining observations.

For the bulk distribution, given a vector of observations $\{y_1, ..., y_n\}$ associated with locations $\{s_1, ..., s_n\}$ and covariate vectors $\{\boldsymbol{x}_1, ..., \boldsymbol{x}_n\}$,

$$y_i|\boldsymbol{x}_i, s_i, \boldsymbol{\theta} \sim \text{Lognormal}(\mu_i, \sigma^2)$$

$$\mu_i = \boldsymbol{x}_i^T\boldsymbol{\beta} + W(s_i) + \epsilon$$

$$S \sim GP(0, \Sigma),$$

where $\boldsymbol{\theta}$ is a vector of hyperparameters, W is a spatial random effect modeled as a zero mean Gaussian spatial random field with stationary Matérn covariance function

$$C(d) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu} d\rho^{-1}\right)^{\nu} K_\nu \left(\sqrt{2\nu} d\rho^{-1}\right)$$

where d is the distance between two spatial points, ρ is a range parameter, σ is the standard deviation, and ν is a smoothness parameter [24]. The SPDE formulation of the INLA model leads to a slightly offbeat parameterization of the Matérn, where $\nu = \alpha - d/2$ for some integer α, in this case we take $\alpha = 2$ [47].

The extremes will be treated with a spatially varying generalized Pareto model in the vein of [23], where the scale parameter is allowed to vary across space with a GP prior. given a vector of observations $z_1, ..., z_m$ exceeding a threshold u, selected via examination of diagnostic plots [21] we can write this model as

$$z_i|\boldsymbol{x}_i, s_i, \nu \sim GPD(\mu, \sigma(s_i), \xi)$$

$$\sigma = \boldsymbol{x}_i\boldsymbol{\beta}_e + W_e(s_i) + \epsilon_e(s_i)$$

$$S_e \sim GP(0, \Sigma_e)$$

$$\log(\xi) \sim \text{Gamma}(1, 12)$$

We're interested in comparing the beta terms $\boldsymbol{\beta}_e$ and $\boldsymbol{\beta}$ so the covariates have the same structure between the two models. The spatial effects W and W_e and the nugget terms are estimated separately, and we also have a separate vector of hyperpriors ν. The GPD shape parameter is equipped with a loggamma prior in order to constrain it within reasonable bounds.

Priors for range and standard deviation parameters of the Mateèrn covariance function were set based on examination of semivariograms [64], which suggest a very small range and standard deviation. Minimally informative penalized complexity priors [34] were selected such that the probability of the range parameter exceeding 1

or the SD parameter exceeding 3.5 is 0.10. The same priors were employed for the spatial effect term in both models for the sake of easy comparison.

3.5 Seperating Extreme observations

The first step in fitting a GPD model is determining the threshold above which observations are considered extreme. We assume that this threshold is constant across time and space, and can be approximated by examination of mean residual life (MRL) and threshold stability plots [21].

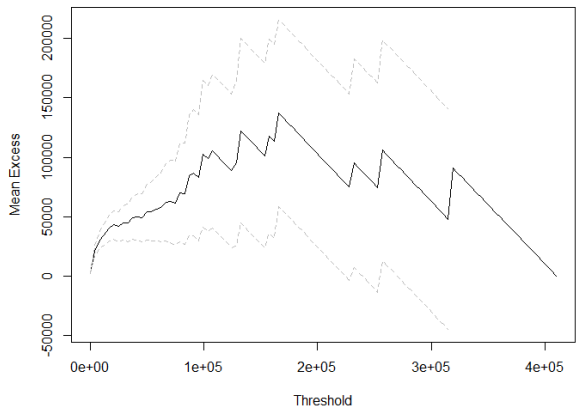

Figure 3.2: Mean Residual Life Plot of observed burned area

The MRL plot above suggests a threshold of around 10,000 acres to ensure stability of the shape parameter estimate, roughly the 97^{th} percentile of the distribution. This leaves us with 180 observations in the extreme data set and 3697 fires in the bulk model.

3.6 Results

The models were fit via integrated nested Laplace approximation [74], a method of approximate Bayesian inference for Gaussian random field models that offers significant computational advantages over traditional probabilistic methods, albeit with some disadvantages - A required Markov conditional independence assumption on the underlying field, and restriction to marginal posterior inference.

The distributions of the resulting regression coefficients are summarised below. In general, the results indicate that vegetation coverage, upper atmospheric variables, and total precipitation in the months leading up to the fire have a more pronounced effect on extremely large fires, while average slope and wind speed are more influential over the rest of the distribution.

	Estimate (Bulk)	95% CI (Bulk)	Estimate (Tail)	95% CI (Tail)
Wind Speed	0.024	-0.047, 0.096	-0.018	-0.258, 0.229
Precipitation	0.073	-0.002, 0.147	-0.268	-0.5, -0.028
Slope	0.266	0.184, 0.348	0.170	-0.093, 0.428
Incidence	0.003	-0.066, 0.072	0.033	-0.18, 0.239
High Vegetation	0.007	-0.069, 0.083	-0.197	-0.403, -0.018
Low Vegetation	0.066	-0.009, 0.142	0.223	0, 0.406
Z500	0.039	-0.032, 0.110	0.097	-0.104, 0.291
Potential Vorticity	0.058	-0.013, 0.128	0.157	-0.037, 0.347

Table 3.1: Standardized Regression Coefficients

The GPD shape parameter for the extreme observations was estimated at 0.04, with a 95% credible interval of $[0.009, 0.11]$. This decay towards zero is in keeping with the tail behavior observed in [76] for the distribution of burned area in Los Angeles county between 1950 and 2000.

Coefficient estimates for fuel model designation are omitted for the sake of brevity, however fires in the bulk distribution were positively associated with pine grass savannah, while larger fires were positively associated with sagebrush grass and negatively associated with tundra. Fires on both ends of the spectrum showed a positive association with intermediate brush.

These results are largely in keeping with what the science suggests. The association with extreme fires and vegetation coverage is particularly notable, and may reflect a pattern of fires burning across areas of lower vegetation, gaining height and intensity before spreading over dense forests with high canopies. Upper atmospheric covariates, particularly vorticity, are also more strongly associated with the area consumed by larger fires.

The relatively minor effect of wind-slope incidence is also noteworthy, as strong downslope winds are widely accepted as a significant driver of fire activity. The fact that this effect is not reflected in the results is likely indicative of the fact that wind was averaged over the duration of each event, smoothing out the potential effects of brief but intense downslope gusts.

There appears to be a general lack of strong associations with burned area among the smaller fires in the model, which likely indicates a great deal of variety in how these fires spread and behave. Larger fires, on the other hand, appear to be consistently associated with extended periods of low precipitation preceding the event itself along with upper atmospheric conditions that are conducive to convection.

Chapter 4

Conclusions and Discussion

Looking at changes in the Diablo winds gives us an important glimpse at the future of Northern California's fire weather. Looking at the history of these fires from a quantiative perspective helps us contextualize these observations, and also reveals some of the inherent difficulties inherent in examining this natural processes through a statistical lens.

The juxtaposition between the important role of Diablo winds in dictating the region's fire regime and the relatively insignificant contribution of wind conditions in the spatial regression model highlights one of the major issues in pursuing this line of study. As briefly discussed in section 1.3, every stage of a wildland fire, from ignition to containment, is influenced by physical, ecological, and human factors. Some factors - like the Diablo winds - have an obvious and immediate impact and a more subtle, long-term impact. Centuries of atmospheric conditions that lend themselves to fireweather have cultivated fire-reliant ecosystems, which have in turn been impacted and altered by recent human settlement and development [58] [86].

Similarly, the seasonality of wildfires across time is intertwined with the seasonal trends in of their numerous atmospheric and ecological drivers [42]. This leads to severe issue of multiseasonality that make temporal forecasting extremely difficult [61].

The goal of the modeling effort in this paper was to conduct inference on the factors that drive the extent Northern California's fires in broad strokes. An explicit effort was made to minimize the confounding effects of multicollinearity, and the latent Gaussian model allowed us to develop a covariance-stationary approximation to the spatial dependence structure. Some temporal effects were accounted for in covariate construction, with different quantities being averaged across relevant time frames.

A natural question is why this choice was made over a more complex model with a non-stationary covariance function and a temporal random effect as in [102]. This decision was made in order to minimize the potential error that would result from the misspecification of non-stationary random effects. The asymptotic robustness of Gaussian process regression with misspecified covariance has been well studied both theoretically [82] [81] and empirically [97]. The finite sample case is not as well studied. Simulation studies as in [7] and [101] often focus on the issue of misspecified parameters, rather than the function itself. There has also been some work in the machine learning literature on finding bounds for the error induced by covariance misspecification as in [32]. In the case of this study, the question is whether a stationary random effect or a misspecified non-stationary random effect works as a better approximation to the true non-stationary covariance structure. The Bayesian framework and relatively small data set adds another layer of complication, creating a scenario where the consistency of our estimates could depend heavily on our choice

of priors. Overfitting is a major concern with non-stationary modeling as explored in [33] and [41].

What ultimately makes the choice of a stationary covariance function more appealing for this application is the sparseness of observations. It is very difficult to get a sense of what local covariance should look like when dealing with events as rare and complicated as wildland fires. There are 180 or so fires in this data set that exceed the extremal threshold, and slightly over 2800 in the bulk model. These are rare and spread out events. Making the assumption of covariance stationarity and including some of the major influencers of local behavior in the fixed effect term to account for mean non-stationarity allows us to recover a somewhat-reasonable approximation to the underlying process.

The choice of a purely spatial model is another point worth addressing, and again it boils down to the lack of a clear specification for a temporal effect. The seasonality of wildfires across time is intertwined with the seasonal trends in atmospheric and ecological drivers, many of which move on scales that are beyond the scope of our data [42]. This leads to severe issue of multiseasonality, which is particularly confounding with regards to the extreme observations [52]. Outside of seasonality concerns, the anthropogenic warming trend present throughout the study period and the complex relationship of global temperature and regional atmospheric effects is another significant barrier to specifying a coherent temporal random effect.

Predictive methods such as deep learning have recently become popular in wildfire modeling [71] [6], however it is worth noting that non-stationarity is as much a matter of concern in these approaches as it is in statistical inference [100]. The

problem of modeling wildfire behavior on any large spatial or temporal scale is ultimately constrained by the sheer number of moving parts involved, and the difficulty of attempting to account for every contributing factor.

Data-related issues of scale and availability are limitations to any purely data-driven approach, while physical approaches are limited by the reliance on idealized models that may not reflect reality. An interesting line of research in modern fire science is data assimilation, incorporating physical and statistical methods into the same model to draw on the strengths of both approaches [55].

Bibliography

[1] John T. Abatzoglou, Benjamin J. Hatchett, Paul Fox-Hughes, Alexander Gershunov, and Nicholas J. Nauslar. Global climatology of synoptically-forced downslope winds. 41(1):31–50, 2021. _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.6607.

[2] AIA. Berkeley conflagration of 1923, 1923.

[3] Frank A Albini. *Estimating wildfire behavior and effects*, volume 30. Department of Agriculture, Forest Service, Intermountain Forest and Range . . . , 1976.

[4] Amy A. Morin, Alisha Albert-Green, Douglas G. Woolford, and David L. Martell. The use of survival analysis methods to model the control time of forest fires in Ontario, Canada. *International journal of wildland fire*, 24(7):964–973, 2015. PubAg AGID: 4358753.

[5] Sudipto Banerjee, Bradley P Carlin, and Alan E Gelfand. *Hierarchical modeling and analysis for spatial data*. Chapman and Hall/CRC, 2003.

[6] John Ray Bergado, Claudio Persello, Karin Reinke, and Alfred Stein. Predicting wildfire burns from big geodata using deep learning. *Safety science*, 140:105276, 2021.

[7] Moreno Bevilacqua and Tarik Faouzi. Estimation and prediction of gaussian processes using generalized cauchy covariance model under fixed domain asymptotics. *Electronic Journal of Statistics*, 13(2):3025–3048, 2019.

[8] Nejc Bezak, Mitja Brilly, and Mojca Šraj. Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. *Hydrological Sciences Journal*, 59(5):959–977, 2014.

[9] Michael Bowden, Jeffrey Prestemon, Todd Hawbaker, John Carpenter, Maureen Brooks, Karen Abt, Ronda Sutphen, and Samuel Scranton. Wildfire ignitions: A review of the science and recommendations for empirical modeling. 01 2013.

[10] Carrie Bowers. The diablo winds of northern california: Climatology and numerical simulations. Master's thesis, San José State University, 2018.

[11] David M. J. S. Bowman, Jennifer K. Balch, Paulo Artaxo, William J. Bond, Jean M. Carlson, Mark A. Cochrane, Carla M. D'Antonio, Ruth S. DeFries, John C. Doyle, Sandy P. Harrison, Fay H. Johnston, Jon E. Keeley, Meg A. Krawchuk, Christian A. Kull, J. Brad Marston, Max A. Moritz, I. Colin Prentice, Christopher I. Roos, Andrew C. Scott, Thomas W. Swetnam, Guido R. van der Werf, and Stephen J. Pyne. Fire in the Earth System. *Science*, 324(5926):481–484, April 2009. Publisher: American Association for the Advancement of Science.

[12] Paul B Breslow and David J Sailor. Vulnerability of wind power resources to climate change in the continental united states. *Renewable energy*, 27(4):585–598, 2002.

[13] Benjamin Bryant and A. Westerling. Scenarios for future wildfire risk in California: Links between changing demography, land use, climate, and wildfire. *Environmetrics*, 25, September 2014.

[14] MP Calef, AD McGuire, and FS Chapin III. Human influences on wildfire in alaska from 1988 through 2005: an analysis of the spatial patterns of human impacts. *Earth Interactions*, 12(1):1–17, 2008.

[15] CalFire. 2020 incident archive, 2021.

[16] Nikolaos Christidis and Peter A. Stott. Changes in the geopotential height at 500hpa under the influence of external climatic forcings. *Geophysical Research Letters*, 42(24):10,798–10,806, 2015.

[17] Emilio Chuvieco, Louis Giglio, and Chris Justice. Global characterization of fire activity: toward defining fire regimes from Earth observation data. *Global Change Biology*, 14(7):1488–1502, 2008. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2486.2008.01585.x.

[18] Terry L. Clark, Larry Radke, Janice Coen, and Don Middleton. Analysis of Small-Scale Convective Dynamics in a Crown Fire Using Infrared Video Camera Imagery. *Journal of Applied Meteorology and Climatology*, 38(10):1401–1420, October 1999. Publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology.

[19] Janice L. Coen, Wilfrid Schroeder, and Brad Quayle. The generation and forecast of extreme winds during the origin and progression of the 2017 tubbs fire. 9(12):462, 2018. Number: 12 Publisher: Multidisciplinary Digital Publishing Institute.

[20] Janice L. Coen, E. Natasha Stavros, and Josephine A. Fites-Kaufman. Deconstructing the King megafire. *Ecological Applications*, 28(6):1565–1580, 2018. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/eap.1752.

[21] Stuart Coles. In Stuart Coles, editor, *An Introduction to Statistical Modeling of Extreme Values*, Springer Series in Statistics, pages 45–73. Springer, 2001.

[22] Sean C.P. Coogan, Lori D. Daniels, Den Boychuk, Philip J. Burton, Mike D. Flannigan, Sylvie Gauthier, Victor Kafka, Jane S. Park, and B. Mike Wotton. Fifty years of wildland fire science in canada. *Canadian Journal of Forest Research*, 51(2):283–302, 2021.

[23] Daniel Cooley, Douglas Nychka, and Philippe Naveau. Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, 102(479):824–840, 2007.

[24] Noel Cressie and Hsin-Cheng Huang. Classes of nonseparable, spatio-temporal stationary covariance functions. *Journal of the American Statistical association*, 94(448):1330–1339, 1999.

[25] S G Cumming. A parametric model of the fire-size distribution. *Canadian Journal of Forest Research*, 31(8):1297–1303, August 2001. Publisher: NRC Research Press.

[26] A. A. Cunningham and D. L. Martell. A Stochastic Model for the Occurrence of Man-caused Forest Fires. *Canadian Journal of Forest Research*, 3(2):282–287, June 1973. Publisher: NRC Research Press.

[27] Philip E. Dennison, Simon C. Brewer, James D. Arnold, and Max A. Moritz. Large wildfire trends in the western united states, 1984–2011. 41:2928–2933, Apr 2014.

[28] Leo J Donner, Bruce L Wyman, Richard S Hemler, Larry W Horowitz, Yi Ming, Ming Zhao, Jean-Christophe Golaz, Paul Ginoux, S-J Lin, M Daniel Schwarzkopf, et al. The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component am3 of the gfdl global coupled model cm3. *Journal of Climate*, 24(13):3484–3519, 2011.

[29] Jeff Eidenshink, Brian Schwind, Ken Brewer, Zhi-Liang Zhu, Brad Quayle, and Stephen Howard. A project for monitoring trends in burn severity. *Fire ecology*, 3(1):3–21, 2007.

[30] Nicolas Faivre, Yufang Jin, Michael L. Goulden, James T. Randerson, Nicolas Faivre, Yufang Jin, Michael L. Goulden, and James T. Randerson. Controls on the spatial pattern of wildfire ignitions in Southern California. *International Journal of Wildland Fire*, 23(6):799–811, July 2014. Publisher: CSIRO PUBLISHING.

[31] Paulo Fernandes, Abilio Pacheco, Rui Almeida, and João Claro. The role of fire suppression force in limiting the spread of extremely large forest fires in Portugal. *European Journal of Forest Research*, 135:253–262, April 2016.

[32] Christian Fiedler, Carsten W Scherer, and Sebastian Trimpe. Practical and rigorous uncertainty bounds for gaussian process regression. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 7439–7447, 2021.

[33] Geir-Arne Fuglstad, Daniel Simpson, Finn Lindgren, and Håvard Rue. Does non-stationary spatial data always require non-stationary random fields? *Spatial Statistics*, 14:505–531, 2015.

[34] Geir-Arne Fuglstad, Daniel Simpson, Finn Lindgren, and Håvard Rue. Constructing priors that penalize the complexity of gaussian random fields. *Journal of the American Statistical Association*, 114(525):445–452, 2019.

[35] Peter R Gent and Gokhan Danabasoglu. Response to increasing southern hemisphere winds in ccsm4. *Journal of climate*, 24(19):4992–4998, 2011.

[36] Russell T Graham, Sarah McCaffrey, and Theresa B Jain. *Science basis for changing forest structure to modify wildfire behavior and severity*. United States Department of Agriculture Forest Service, Rocky Mountain ..., 2004.

[37] Erin J. Hanan, Jianning Ren, Christina L. Tague, Crystal A. Kolden, John T. Abatzoglou, Ryan R. Bart, Maureen C. Kennedy, Mingliang Liu, and Jennifer C. Adam. How climate change and fire exclusion drive wildfire regimes at actionable scales. 16(2):024051, 2021. Publisher: IOP Publishing.

[38] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, Saleh Abdalla, Xavier Abellan, Gianpaolo Balsamo, Peter Bechtold, Gionata Biavati, Jean Bidlot, Massimo Bonavita, Giovanna De Chiara, Per Dahlgren, Dick Dee, Michail Diamantakis, Rossana Dragani, Johannes Flemming, Richard Forbes, Manuel Fuentes, Alan Geer, Leo Haimberger, Sean Healy, Robin J. Hogan, Elías Hólm, Marta Janisková, Sarah Keeley, Patrick Laloyaux, Philippe Lopez, Cristina Lupu, Gabor Radnoti, Patricia de Rosnay, Iryna Rozum, Freja Vamborg, Sebastien Villaume, and Jean-Noël Thépaut. The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020.

[39] Stephen P Jenkins. Survival analysis. *Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK*, 42:54–56, 2005.

[40] CDea Jones, JK Hughes, Nicolas Bellouin, SC Hardiman, GS Jones, Jeff Knight, Spencer Liddicoat, FM O'connor, Robert Joseph Andres, Christopher Bell, et al. The hadgem2-es implementation of cmip5 centennial simulations. *Geoscientific Model Development*, 4(3):543–570, 2011.

[41] Kenneth Jung and Nigam H. Shah. Implications of non-stationarity on predictive modeling using ehrs. *Journal of Biomedical Informatics*, 58:168–174, 2015.

[42] Jon E. Keeley and Alexandra D. Syphard. Large california wildfires: 2020 fires in historical context. 17(1):22.

[43] Maureen C. Kennedy, Ryan R. Bart, Christina L. Tague, and Janet S. Choate. Does hot and dry equal more wildfire? contrasting short- and long-term climate effects on fire in the sierra nevada, ca. *Ecosphere*, 12(7):e03657, 2021.

[44] Meg A. Krawchuk, Max A. Moritz, Marc-André Parisien, Jeff Van Dorn, and Katharine Hayhoe. Global Pyrogeography: the Current and Future Distribution of Wildfire. *PLOS ONE*, 4(4):e5102, April 2009. Publisher: Public Library of Science.

[45] Meg A. Krawchuk, Max A. Moritz, Marc-André Parisien, Jeff Van Dorn, and Katharine Hayhoe. Global pyrogeography: the current and future distribution of wildfire. *PLOS ONE*, 4(4):1–12, 04 2009.

[46] D. Y. Lin and L. J. Wei. The robust inference for the cox proportional hazards model. *Journal of the American Statistical Association*, 84(408):1074–1078, 1989.

[47] Finn Lindgren and Håvard Rue. Bayesian spatial modelling with r-inla. *Journal of statistical software*, 63:1–25, 2015.

[48] Jeremy S. Littell, Donald McKenzie, Ho Yi Wan, and Samuel A. Cushman. Climate Change and Future Wildfire in the Western United States: An Ecological Approach to Nonstationarity. *Earth's Future*, 6(8):1097–1111, 2018. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018EF000878.

[49] Yi-Chin Liu, Pingkuan Di, Shu-Hua Chen, Xue Meng Chen, and Jeremy Avise. Changes in fire potential and extreme downslope winds in california under global warming projections of 1.5 °c, 2.0 °c, and 3°c. AGU, 2020.

[50] Yi-Chin Liu, Pingkuan Di, Shu-Hua Chen, XueMeng Chen, Jiwen Fan, John DaMassa, and Jeremy Avise. Climatology of diablo winds in northern california and their relationships with large-scale climate variabilities. 56(3):1335–1356, 2021.

[51] A. C. Lute and John T. Abatzoglou. Best practices for estimating near-surface air temperature lapse rates. *INTERNATIONAL JOURNAL OF CLIMATOLOGY*, 41(S1):E110–E125, 2021.

[52] Ed Mackay and Philip Jonathan. Assessment of return value estimates from stationary and non-stationary extreme value models. 207:107406.

[53] Ashima Malik, Megha Rajam Rao, Nandini Puppala, Prathusha Koouri, Venkata Anil Kumar Thota, Qiao Liu, Sen Chiao, and Jerry Gao. Data-driven wildfire risk prediction in northern california. 12(1):109, 2021. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.

[54] J. Mandel, J. D. Beezley, and A. K. Kochanski. Coupled atmosphere-wildland fire modeling with wrf 3.3 and sfire 2011. *Geoscientific Model Development*, 4(3):591–610, 2011.

[55] Jan Mandel, Lynn S Bennethum, Jonathan D Beezley, Janice L Coen, Craig C Douglas, Minjeong Kim, and Anthony Vodacek. A wildland fire model with data assimilation. *Mathematics and Computers in Simulation*, 79(3):584–606, 2008.

[56] Clifford F. Mass and David Ovens. The synoptic and mesoscale evolution accompanying the 2018 camp fire of northern california in: Bulletin of the american meteorological society volume 102 issue 1 (2021). 2021.

[57] Brandon T. McClung and Clifford F. Mass. The strong, dry winds of central and northern california: Climatology and synoptic evolution. *Weather and Forecasting*, 35:2163–2178, 2020.

[58] Kendra K. McLauchlan, Philip E. Higuera, Jessica Miesel, Brendan M. Rogers, Jennifer Schweitzer, Jacquelyn K. Shuman, Alan J. Tepley, J. Morgan Varner, Thomas T. Veblen, Solny A. Adalsteinsson, Jennifer K. Balch, Patrick Baker, Enric Batllori, Erica Bigio, Paulo Brando, Megan Cattau, Melissa L. Chipman, Janice Coen, Raelene Crandall, Lori Daniels, Neal Enright, Wendy S. Gross, Brian J. Harvey, Jeff A. Hatten, Sharon Hermann, Rebecca E. Hewitt, Leda N. Kobziar, Jennifer B. Landesmann, Michael M. Loranty, S. Yoshi Maezumi, Linda Mearns, Max Moritz, Jonathan A. Myers, Juli G. Pausas, Adam F. A. Pellegrini, William J. Platt, Jennifer Roozeboom, Hugh Safford, Fernanda Santos, Robert M. Scheller, Rosemary L. Sherriff, Kevin G. Smith, Melinda D. Smith, and Adam C. Watts. Fire as a fundamental ecological process: Research advances and frontiers. *Journal of Ecology*, 108(5):2047–2069, 2020.

[59] Kurt M Menning and Scott L Stephens. Fire climbing in the forest: a semiqualitative, semiquantitative approach to assessing ladder fuel hazards. *Western Journal of Applied Forestry*, 22(2):88–93, 2007.

[60] Andrea Meyn, Peter White, Constanze Buhk, and Anke Jentsch. Environmental drivers of large, infrequent wildfires: The emerging conceptual model. 3:287–312, 2007.

[61] Russell G. Miller, Ryan Tangney, Neal J. Enright, Joseph B. Fontaine, David J. Merritt, Mark K.J. Ooi, Katinka X. Ruthrof, and Ben P. Miller. Fire seasonality mechanisms are fundamental for understanding broader fire regime effects. 35(10):869–871. Publisher: Elsevier.

[62] Norma Angélica Monjarás-Vega, Carlos Ivan Briones-Herrera, Daniel José Vega-Nieva, Eric Calleros-Flores, José Javier Corral-Rivas, Pablito Marcelo López-Serrano, Marín Pompa-García, Dante Arturo Rodríguez-Trejo, Artemio Carrillo-Parra, Armando González-Cabán, Ernesto Alvarado-Celestino, and William Matthew Jolly. Predicting forest fire kernel density at multiple scales with geographically weighted regression in Mexico. *Science of The Total Environment*, 718:137313, May 2020.

[63] NWCG. Wildland fire suppression tactics reference guide. 1996.

[64] Ricardo A Olea. A six-step practical approach to semivariogram modeling. *Stochastic Environmental Research and Risk Assessment*, 20(5):307–318, 2006.

[65] Elsa Pastor, Luis Zarate, Eulalia Planas, and Josep Arnaldos. Mathematical models and calculation systems for the study of wildland fire behaviour. *Progress in Energy and Combustion Science*, 29:139–153, 12 2003.

[66] Roger D Peng, Frederic Paik Schoenberg, and James A Woods. A space–time conditional intensity model for evaluating a wildfire hazard index. *Journal of the American Statistical Association*, 100(469):26–35, 2005.

[67] Trent D. Penman, Luke Collins, Alexandra D. Syphard, Jon E. Keeley, and Ross A. Bradstock. Influence of Fuels, Weather and the Built Environment on the Exposure of Property to Wildfire. *PLOS ONE*, 9(10):1–9, October 2014. Publisher: Public Library of Science.

[68] Haiganoush K. Preisler and Anthony L. Westerling. Statistical model for forecasting monthly large wildfire events in western united states. *Journal of Applied Meteorology and Climatology*, 46(7):1020 – 1030, 2007.

[69] Caitlin R. Proctor, Juneseok Lee, David Yu, Amisha D. Shah, and Andrew J. Whelton. Wildfire caused widespread drinking water distribution network contamination. *AWWA Water Science*, 2(4):e1183, 2020.

[70] SC Pryor, RJ Barthelmie, and Justin T Schoof. Past and future wind climates over the contiguous usa based on the north american regional climate change assessment program model suite. *Journal of Geophysical Research: Atmospheres*, 117(D19), 2012.

[71] David Radke, Anna Hessler, and Dan Ellsworth. Firecast: Leveraging deep learning to predict wildfire spread. In *IJCAI*, pages 4575–4581, 2019.

[72] Keywan Riahi, Shilpa Rao, Volker Krey, Cheolhung Cho, Vadim Chirkov, Guenther Fischer, Georg Kindermann, Nebojsa Nakicenovic, and Peter Rafaj. Rcp 8.5—a scenario of comparatively high greenhouse gas emissions. *Climatic change*, 109(1):33–57, 2011.

[73] Carlos Rossa and Paulo Fernandes. Empirical modeling of fire spread rate in no-wind and no-slope conditions. *Forest Science*, 64, 04 2018.

[74] Håvard Rue, Sara Martino, and Nicolas Chopin. Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392, 2009.

[75] Chettouh Samia and Hamzi Rachida. Statistical Fire Models : Review. *International Journal of Engineering Research & Technology*, 2(12), December 2013. Publisher: IJERT-International Journal of Engineering Research & Technology.

[76] Frederic Paik Schoenberg, Roger Peng, and James Woods. On the distribution of wildfire sizes. *Environmetrics*, 14(6):583–592, September 2003.

[77] Jason J. Sharples and James E. Hilton. Modeling vorticity-driven wildfire behavior using near-field techniques. *Frontiers in Mechanical Engineering*, 5, 2020.

[78] Steven C Sherwood, Sandrine Bony, and Jean-Louis Dufresne. Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature*, 505(7481):37–42, 2014.

[79] William C Skamarock, Joseph B Klemp, Jimy Dudhia, David O Gill, D Barker, Michael G Duda, X-Y Huang, Wei Wang, and Jordan G Powers. A description of the advanced research wrf version 3 (no. ncar/tn-475+ str), university corporation for atmospheric research. *National Center for Atmospheric Research Boulder, Colorado, USA, https://doi. org/10.5065/D68S4MVH*, 2008.

[80] Craig Smith, Benjamin J. Hatchett, and Michael Kaplan. A surface observation based climatology of diablo-like winds in california's wine country and western sierra nevada. 1(2):25, 2018. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute.

[81] Michael L. Stein. Bounds on the efficiency of linear predictions using an incorrect covariance function. *The Annals of Statistics*, 18(3):1116–1138, 1990.

[82] Michael L. Stein and Mark S. Handcock. Some asymptotic properties of kriging when the covariance function is misspecified. 21(2):171–190.

[83] Alan Taylor and Carl Skinner. Spatial patterns and controls on historical fire regimes and forest structure in the klamath mountains. *Ecological Applications - ECOL APPL*, 13:704–719, 06 2003.

[84] Pier-Olivier Tremblay, Thierry Duchesne, and Steven G. Cumming. Survival analysis and classification methods for forest fire size. *PLOS ONE*, 13(1):1–16, 01 2018.

[85] Rolf Turner. Point patterns of forest fire locations. *Environmental and Ecological Statistics*, 16(2):197–223, June 2009.

[86] Jan W. van Wagtendonk. The history and evolution of wildland fire use. 3(2):3–17.

[87] Giovanni Di Virgilio, Jason P. Evans, Stephanie A. P. Blake, Matthew Armstrong, Andrew J. Dowdy, Jason Sharples, and Rick McRae. Climate change increases the potential for extreme wildfires. 46(14):8517–8526, 2019. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019GL083699.

[88] Jiali Wang and Veerabhadra R Kotamarthi. High-resolution dynamically downscaled projections of precipitation in the mid and late 21st century over north america. *Earth's Future*, 3(7):268–288, 2015.

[89] Sally S.-C. Wang and Yuxuan Wang. Quantifying the effects of environmental factors on wildfire burned area in the south central US using integrated machine learning techniques. *Atmospheric Chemistry and Physics*, 20(18):11065–11087, September 2020. Publisher: Copernicus GmbH.

[90] Adam C. Watts and Leda N. Kobziar. Smoldering combustion and ground fires: Ecological effects and multi-scale significance. 9(1):124–132.

[91] L. J. Wei. The accelerated failure time model: A useful alternative to the cox regression model in survival analysis. *Statistics in Medicine*, 11(14-15):1871–1879, 1992.

[92] A. L. Westerling and B. P. Bryant. Climate change and wildfire in california. 87(1):231–249, 2008.

[93] Boyd E. Wickman. Life History of the Incense-Cedar Wood Wasp, Syntexis libocedrii (Hymenoptera: Syntexidae). *Annals of the Entomological Society of America*, 60(6):1291–1295, November 1967.

[94] Katharine M. Willett, Nathan P. Gillett, Philip D. Jones, and Peter W. Thorne. Attribution of observed surface humidity changes to human influence. 449(7163):710–712, 2007. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7163 Primary_atype: Research Publisher: Nature Publishing Group.

[95] A. Park Williams, John T. Abatzoglou, Alexander Gershunov, Janin Guzman-Morales, Daniel A. Bishop, Jennifer K. Balch, and Dennis P. Lettenmaier. Observed impacts of anthropogenic climate change on wildfire in california. 7(8):892–910, 2019. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019EF001210.

[96] B M Wotton and David L Martell. A lightning fire occurrence model for Ontario. *Canadian Journal of Forest Research*, 35(6):1389–1401, June 2005. Publisher: NRC Research Press.

[97] George Wynne, François-Xavier Briol, and Mark Girolami. Convergence guarantees for gaussian process means with misspecified likelihoods and smoothness. *Journal of Machine Learning Research*, 22, 2021.

[98] Dexen D.Z. Xi, Stephen W. Taylor, Douglas G. Woolford, and C.B. Dean. Statistical models of key components of wildfire risk. *Annual Review of Statistics and Its Application*, 6(1):197–222, 2019.

[99] Dexen D.Z. Xi, Stephen W. Taylor, Douglas G. Woolford, and C.B. Dean. Statistical models of key components of wildfire risk. 6(1), 2019. _eprint: https://doi.org/10.1146/annurev-statistics-031017-100450.

[100] Annie Xie, James Harrison, and Chelsea Finn. Deep reinforcement learning amidst lifelong non-stationarity. *arXiv preprint arXiv:2006.10701*, 2020.

[101] Alexey Zaytsev, Evgenya Romanenkova, and Dmitry Ermilov. Interpolation error of Gaussian process regression for misspecified case. 91:83–95, 11–13 Jun 2018.

[102] Zhongwei Zhang, Elias Krainski, Peng Zhong, Håvard Rue, and Raphaël Huser. Joint Modeling and Prediction of Massive Spatio-Temporal Wildfire Count and Burnt Area Data with the INLA-SPDE Approach. *arXiv:2202.06502 [stat]*, February 2022. arXiv: 2202.06502.

[103] Tarisa K. Zimet and Jonathan E. Martin. A Model Based Analysis of the Role of an Upper-Level Front and Stratospheric Intrusion in the Mack Lake Fire. *In: 2d International Wildland Fire Ecology and Fire Management Congress: 5th Symposium on Fire and Forest Meteorology; 2003 November 16-20; Orlando, FL. Boston, MA: American Meteorological Society: 1.3.*, 2003.

[104] Zachary Zobel, Jiali Wang, Donald J Wuebbles, and V Rao Kotamarthi. Evaluations of high-resolution dynamically downscaled ensembles over the contiguous united states. *Climate Dynamics*, 50(3):863–884, 2018.