It is a Dry Heat: Econometric Model of Historic Fires

IT IS A DRY HEAT: ECONOMETRIC MODEL OF HISTORIC FIRES

by

Taylor James Dew

A Thesis Submitted to the Faculty of the

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

WITH A MAJOR IN APPLIED ECONOMICS AND DATA ANALYSIS

In the Graduate College

THE UNIVERSITY OF ARIZONA

2021

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Master's Committee, we certify that we have read the thesis
prepared by: Taylor James Dew
titled: It Is a Dry Heat: Econometric Model of Historical Fires

and recommend that it be accepted as fulfilling the thesis requirement for the Master's Degree.

George Frisvold
George Frisvold — Date: Dec 17, 2021

Gary Thompson
Gary Thompson — Date: Dec 17, 2021

Daniel Scheitrum — Date: Dec 17, 2021

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to the Graduate College.

I hereby certify that I have read this thesis prepared under my direction and recommend that it be accepted as fulfilling the Master's requirement.

George Frisvold
George Frisvold — Date: Dec 17, 2021
Thies Commmittee Chair
Department of Agricultural & Resoure Economics

Acknowledgements

This thesis would have never been possible without the mentors, friends, and family that have helped guide this paper to completion.

I would first like to thank my supervisor, Professor George Frisvold. Without his understanding, knowledge, and kindness this thesis would not have been possible. As he suggested, the theory behind mastery is not the absence of mistakes but the ability to prove skill.

Additionally, I would like to thank the Agriculture & Resource Economics Department for the education that I have received over the last four years. I would particularly like to single out Danielle Burrow for all the hard work she has done since we first met. The kindness and upbeat attitude have been truly a benefit to my academic career.

I would like to thank my family for consistently asking if I had finished my thesis and why I was not finished yet. Without the constant persistence I would have not been as motivated to work consistently even through the early hours of the morning.

Finally, I would like to thank Markanna Moore, who provided countless discussions about almost every aspect of the thesis and was instrumental to the completion of the paper.

Table of Contents

FIGURES:

TABLES:

Abstract:

Several studies have applied regression analysis to measure factors contributing to larger wildfire suppression costs. They often include acres burned, variables that are functions of acres burned, or both. This can create problems of simultaneity bias. While it is common for studies to use instrumental variable methods to address simultaneity, they in general do not evaluate the strength or weakness of their instruments. Another drawback of using acres burned as an explanatory variable is that regression models have limited value in forecasting suppression costs ahead of time, because suppression and burning occur at the same time.

This study takes a different approach, relying on variables that can be used as soon as fire starts. It attempts to answer the question, given that a fire has started, what accounts for it having higher suppression costs and more burned acres? Data from the Burned Area Emergency Response (BAER) reports are combined with other geo-coded variables to examine wildfires in Arizona's national forests from 2002-2019.

Regressions were run for three different variables: (a) natural log of suppression costs, (b) natural log of acres burned, and (c) a binary variable that equaled one if the fire was greater than 30,000 acres and zero otherwise. The regression results suggest that Arizona wildfires that start in May and June are positively associated with higher suppression costs and more acres burned. This suggests benefits of increased vigilance of fire managers during these months. This variable was less able to predict the occurrence of the very largest fires, however. The amount of land in the Wildland Urban Interface (WUI) was negatively associated with fire suppression costs and not a significant predictor of fire size. Past empirical results regarding the WUI have been mixed. Average relative humidity was a significant (negative) predictor of both suppression costs and of very large fires. This variable has not been much used in previous studies and may become important if aridity in Arizona increases with climate change.

Introduction and Motivation

As this thesis is being written, a wildfire named the Bootleg Fire (in Oregon) is the largest of the 2021 wildfire season at over 200,000 acres, with 21 homes burned and thousands of people evacuated. This fire, and others like the Bootleg Fire, have been drawing interest in the news over the past 40 years. *The Guardian* illuminated the devastation in July of 2021, “more than 60 wildfires were burning across at least 10 states in the parched American West” (*The Guardian 2021*). Another article describes an effort to save a landmark, where “firefighters wrapped the base of the world’s largest tree in a fire-resistant blanket as they tried to save a famous grove of gigantic old-growth sequoias from wildfires burning in California’s rugged Sierra Nevada” (*Associated Press 2021*). Wildfires have the media speculating about a large variety of potential causes, influences, and costs. The possible causes that are suggested range from climate change (Westerling & Bryant 2007), agency management (Stephens & Ruth, 2005), or natural changing ecosystems (Westerling & Bryant 2007). An area of scientific consensus, though, is that these large fires have severe consequences for western United States. The negative externalities include environmental health, respiratory health, water health, ecosystem services, tourism, and air travel (Richardson et al. 2012). Some of these impacts are easily measurable, but others, such as ecosystem services, which are commonly defined as benefits people obtain from the ecosystem, are extremely difficult to measure.

This study examines wildfire suppression costs and acres burned on fires originating on U.S. Forest Service lands in Arizona over last two decades. One strand of empirical literature attempts to predict suppression costs or acres burned at a larger regional level (Abt et al. 2009, Prestemon & Donovan 2008, Gebert 2007). For example, this might be at the level of multi-state Forest Service regions. Another strand of literature attempts to estimate factors affecting suppression costs for individual fires (Lui et al 2015, Donovan et al 2004, Katuwal et al 2015).

This present study follows this second approach. As will be discussed below, a relatively small number of mega-fires account for a large share of suppression costs and acres burned in Arizona. For example, just two fires, the Wallow Fire, and the Rodeo/Chediski Fire alone account for 35% of acres burned and 25% of suppression costs in Arizona from 2002-2019. The "top 10 fires" account for two-thirds of acres burned and more than half of suppression costs. The research question here then is, given that a fire starts, can we identify which factors contribute to a fire becoming a mega-fire? It is hoped that being able to identify whether a fire is likely to be a mega-fire early will improve resource allocation for fire suppression.

Several previous studies use acres burned as a major explanatory variable in regression equations explaining suppression costs for individual fires. Yet, as illustrated below, these variables are simultaneously determined. More recent research has attempted to test for and correct for simultaneity bias using instrumental variable methods (Gebert 2007). Yet, to date, studies have not reported tests of the strength or weakness of the instrument chosen, and in some cases, use instruments that are ratios, where acres burned (the endogenous variable), is the denominator. Thus, the instruments themselves are functions of the endogenous variable. Aside from potential estimation bias, using simultaneously determined variables limits the ability of these models to forecast suppression costs. In contrast, this study relies on reduced-form equation estimation using predetermined variables, instead of simultaneously determined variables. Here, there is less scope for simultaneity bias and more scope for developing useful forecasts.

Chapter 1: Background on Wildfire Management

Comparing the costs of wildfires to other natural disasters helps put into perspective the money allocated to wildfires in relation to other natural disasters. The federal budget for natural disasters is something that is difficult to allocate as it funds many different governing agencies and management structures. The National Oceanic and Atmospheric Administration (NOAA) compiled multi-agency data for economic impact of national disasters that cost more than a one billion dollars from 1980 to 2011 (Smith, 2013). The natural disasters with the greatest number of economic damages over one billion dollars were tropical cyclones with damages of roughly $417.9 billion. Wildfires were sixth on the list. From 1980 to 2011, there were 11 wildfires that that cost over one billion dollars, with more than $22.2 billion spent in total (Smith, 2013). Notably, the NOAA report estimates that droughts/heat waves, also common in the western United States, cost $210 billion over the same period (Smith, 2013). Unfortunately, Smith (2013) only looks at direct economic losses from natural disaster events, while wildfires have major secondary economic losses as well as ecosystem services losses, which were not calculated in the study. However, these estimates of wildfire costs give us some perspective of how they compare to different types of natural disaster. A little lower on the national disasters list of economic impacts, wildfires are still one of the most talked-about issues of our time and have large consequences for the federal government in the future.

A Congressional Research Service report (Hoover, 2017) defined wildfires as unplanned, unwanted wildland fires, including lightning caused fires, unauthorized human caused fires, and escaped prescribed fire projects. Responsibilities for these wildfires are dependent on fire start location and therefore can either be a non-federal or federal response (Hoover, 2017). The term wildfire suppression is a broad term that describes all work associated with extinguishing or confining a fire (Hoover, 2017). From 2007 to 2016, an average of just over 70,000 wildfires have burned an average of 6.6 million acres (Hoover, 2017). From 1994 to 2014, in comparison, the variation in wildfires and acres burned fluctuated significantly (Hoover et al, 2015). Data on wildfires during the twenty year period shows that 2013 had the lowest number of fires at 46,579 compared to the highest number which was in 2006 at just under 100,000 fires (Hoover et al, 2015). Acres burned during the period differed by 8 million acres with the most acres burned in

2006 at just under 10 million, while in 1998 there were under 2 million acres (Hoover et al, 2015). The Congressional Research Service report update for 2020 also indicates that this fluctuation continues through current day (Hoover and Hanson, 2020). The report suggests that the acres burned are steadily increasing but the number of fires is steadily decreasing from roughly 90,000 to 60,000 each year (Hoover and Hanson, 2020). This indicates that fires are getting larger as more acres are burned with lower numbers of fires. From the CRS reports that sourced all the data from the National Interagency Fire Center (NIFC), the acres burned fluctuates significantly from less than two million in 1998 to just under 10 million in 2017, with little trends from one year to another (Hoover and Hanson, 2020). Looking at the distribution of wildfires, 60% burned on federal lands compared to 40% burned on state, local, or private owned lands in 2014 (Hoover et al, 2015).

Additionally, the report suggests that in 2014, just over two thirds of registered fires were registered in eastern states, three fourths of the total acres burned are in the western states (Hoover et al, 2015). The NIFC combines data from all western states (Arizona, Alaska, Colorado, California, Hawaii, Oregon, Nevada, New Mexico, Montana, and Wyoming) (Hoover et al, 2015). Another difference between western states and eastern states is that about 50% of the fires in western states are on federal land compared to only seven percent of the fires are on federal land in the east (Hoover et al, 2015). The Wildland Urban Interface (WUI) – defined as a zone of transition between unoccupied land and human development (USFS website) – has become a large topic of discussion of the increased costs for fire suppression. More than one third of all houses developed in the United States in 2015 were located within the WUI (Hoover et al, 2015). It is estimated that nearly 900,000 homes in the West were at high or very high risk of wildfire damage in 2015 (Hoover et al, 2015). There has been increased debate about how much the federal government should pay to protect these higher-risk developments (Hoover, 2017). US federal wildfire policy prioritizes ecological, social, and legal actions first, with economic costs last (Hoover et al, 2015). The debate is whether this directive gives the land management agencies a blank check to protect these assets with little accountability for increased costs (Hoover, 2017).

The primary federal land management agencies tasked with controlling wildfires are almost exclusively in two departments of the federal government: the Department of the Interior (DOI) and the Department of Agriculture (USDA). Within Interior, they include the Bureau of Land Management (BLM), Bureau of Indian Affairs (BIA), Fish and Wildlife Service (FWS), and the National Park Service (NPS). Within USDA, it is primarily the Forest Service (FS). This present study will focus on just the Forests Service (FS). The FS has nearly 193 million acres compared to the DOI which has more than 407 million acres (Hoover and Hanson, 2020). The Department of Interior has more than double the amount of land than the Forest Service. This would suggest that the Department of Interior would receive roughly double the amount of appropriation than the Forest Service. This holds true as the total DOI appropriations are $111.5 billion (nominal) over 11 years (FY2008 – 2017) over 11 years, roughly twice Forest Service appropriations of $53.9 billion over the same period (Hoover, 2017).

A more important comparison, however, is the difference between fire suppression spending between the two departments. The Forest Service spent on average of 47% of its annual budget over the 11 years on wildfire suppression (Hoover, 2017). In contrast, the Department of Interior, on average, spent just over 7% of its budget on wildfires over the same period (Hoover, 2017). To normalize the differences in budget allocations towards wildfire suppression, the DOI spends on average $19.15 per acre versus the Forest Service, which spends $131.29. This discrepancy in spending is the reason why this study is focusing on the Forest Service. Another reason, discussed below, is data availability.

Federal agencies take the lead on any wildfire that starts on federal lands. Because federal lands account for 46% of the land area in the western United States, the likelihood is high that the federal government will be responsible for putting out wildfires (Hoover and Hanson,2020). With so much overlap between federal government and state local and municipality groups, multiple teams may respond to a potential fire. In that case, the National Interagency Fire Center (NIFC) coordinates the mobilization of resources for wildfires and other incidents throughout the United States. The NIFC was known as the Boise Interactive Agency Fire Center (BIFC). It was created in 1965 because the USDA Forest Service, the Bureau of Land Management, and the National Weather Service saw the need to work together to prevent duplication (NIFC 2021). Multiple other agencies joined the BIFC, and in 1993, the center changed its name to NIFC to reflect its national mission. The NIFC keeps information both on incidences currently underway as well as historical fire data. The NIFC currently has two public datasets with observations from the Wildland Fire Interagency Geospatial Service (WFIGS) with observations from all over the United States. The WFIGS, a new program as of June 7th, 2021, provides the geospatial data to the public and internal departments.

Both the Department of Interior and Forest Service have two methods of paying for wild suppression costs: (a) the Wildland Fire Management (WFM) account from appropriations managed by the Office of Wildland Fires and (b) the Federal Land Assistance Management and Enhancement (FLAME) Act. Both these source use appropriations from Congress to better fight fires. There is also an additional account using the Federal Emergency Management Agency (FEMA) to provide a disaster relief for nonfederal government fires (Hoover et al, 2015). The WFM account was the only source of appropriation until the FLAME act was passed and was used for the preparedness and suppression of wildfires. As of 2017 the WFM account has the largest share of funds for the fire operations, accounting for 64% of appropriation on average over the last ten years (FY2016-FY2006) (Hoover, 2017). The FLAME act of 2009 was passed by Congress in response to federal land agencies consistently going over the WFM budget appropriations during the late 1990s and significantly after fiscal year 2000 (Hoover, 2017).

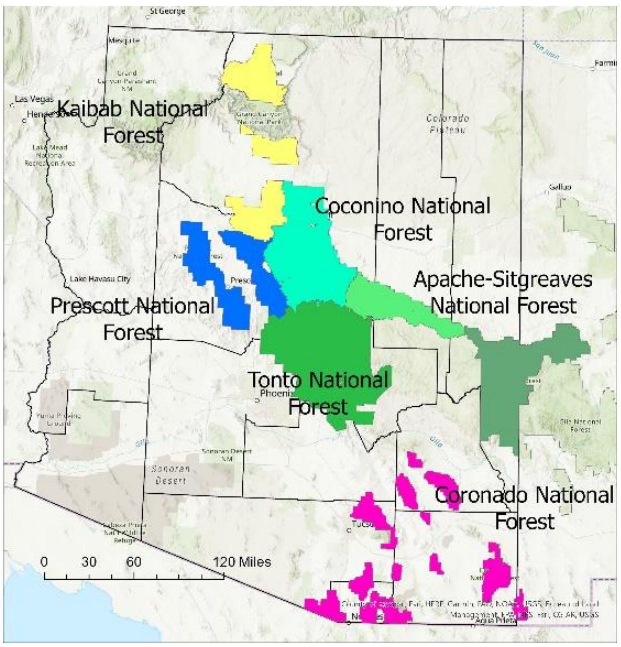
FLAME had three elements: restoring and maintaining resilient landscapes, creating fire adoption communities, and wildfire response (WFLC, 2009). FLAME is supposed to prevent Congress from making emergency appropriations for federal government fire suppression. The last type of account that helps with federal fire appropriations is FEMA, which helps after a governor determines that a fire is out of control and declares a Fire Management Assistance Grant (FMAG). FMAG was barely used from the first occurrence in 1970 until around the 1990s, but spending increased dramatically in the late 1990's and reached an all-time high in 2012 with over 100 FMAG requests by states (Hoover et al, 2015).

How much is requested by the agencies from Congress for the WFM and FLAME accounts are based on statistical models which are used to help predict costs in the next year. The standard practice for the budget process was based on a rolling 10-year suppression obligation average, calculated two fiscal years previously (Hoover et al, 2015). Over the history of suppression cost appropriation, the 10-year rolling average was a good predictor of costs for the next year, until the mid to late 1990s. From 2004 to 2014, a 10-year rolling average underestimated total expenditure for 8 out of the 10 years. Underestimates ranged from 100 million dollars (13%) to $1.26 billion dollars (138%) (Hoover et al, 2015). Given the need to predict suppression costs more accurately for both Congress and the Forest Service, the need for better methods of forecasting became critical.

With the increase in wildfire suppression expenditures since the late 1990s, the lack of predictive ability by government agencies became apparent. The Office of Management and Budget (OMB) required monthly requests from Fire and Aviation Management (FAM). Starting in June of each year since 1970, FAM had to keep up-to-date predictions of total fire suppression costs per the physical year. Predictive algorithms beginning in 1970 used simple moving averages of spending. By 1999, it was becoming obvious to fire management officials and to

OMB, that these moving averages were not a reliable predictor. FAM requested the development of a tool for predicting fire suppression expenditure, developed by Gebert and Schuster (1999). Their chapter introduced a linear regression model, which was stated to be easily interpreted and could be then transferred easily to an excel spread sheet. Additional research by several employees in the Forest Service expanded Gebert and Schuster's work, forecasting into the future by years instead of just a month. This work was provided to the Forest Service and Congress to help make appropriation decisions to the agency for wildfire suppression, though according to the CRS reports the FS still uses moving averages for budget projection requests.

Figure 1. Arizona National Forests

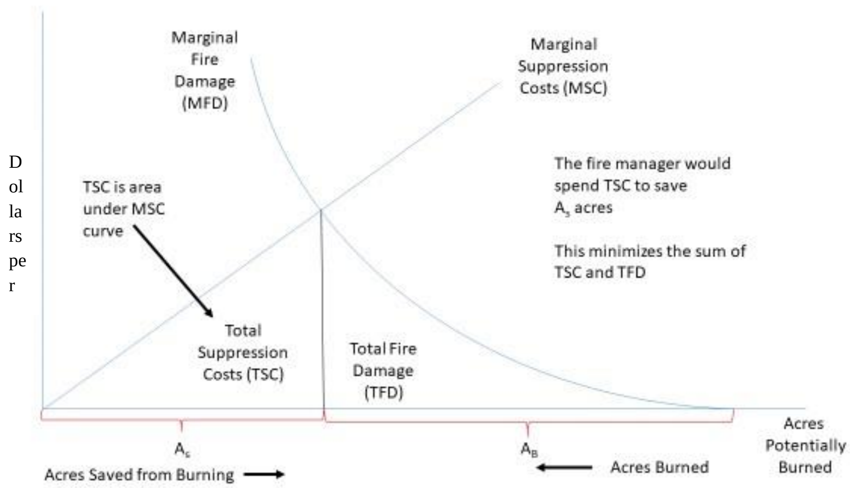


This thesis will contribute to a better understanding of variables that influence suppression costs and acres burned. Specifically, it tries to assess which factors influence the number of acres burned and suppression costs for individual fires, once they have started. For these variables, the thesis relies on Burned Area Emergency Response (BAER) database of the U.S. Forest Service, focusing on national forest lands in Arizona (Figure 1).

Chapter 2: Models of Optimal Fire Suppression

An Economist named William N. Sparhawk introduced an economic model called the Least Cost plus Loss model (C + L) in 1925 (Sparhawk, 1925). Sparhawk's goal was to, "determine how much money can justifiably be spent for fire protection on national forests." Fires created losses, L, through damage while pre-suppression actions and fire suppression entailed costs. Sparhawk introduced the objective of minimizing the sum of fire control costs and the economic value of damage (C + L). Sparhawk did not write in these terms, but one can view his objective in terms of marginal damage and marginal cost curves.

Figure 2. Marginal Costs and Benefits of Fire Suppression



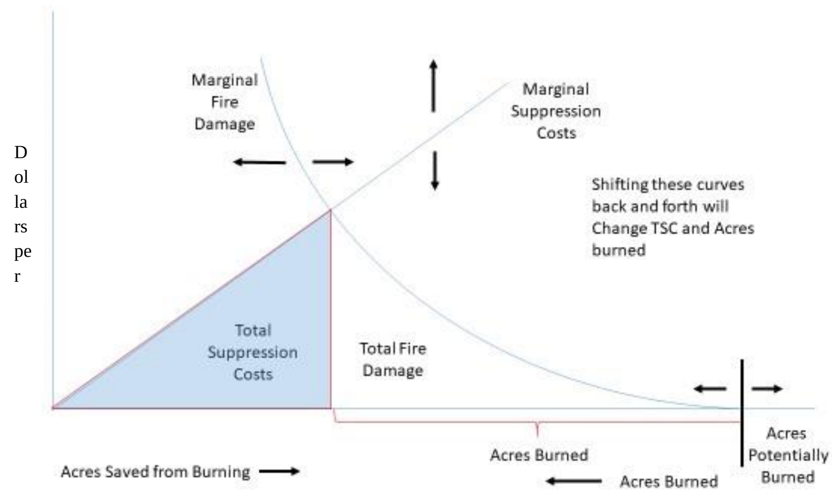
The x-axis is the total number of acres that are saved from burning, while the y axis measures the marginal cost of saving acreage. Reading the x-axis right to left, it measures the converse of acreage saved, acreage burned. Figure 2 also shows the marginal fired damage (MFD) of each additional acre burned. Marginal damages increase as acres burned move from more remote wildlands to encroach on inhabited areas, damaging structures, disrupting economic activity and threatening lives. To fit figure 2 to Sparhawk's scheme, the goal is to minimize the sum of total suppression costs (TSC) and total fire damages (TFD). This occurs by saving A_s acres, the point where Marginal Suppression Costs (MSC) equal Marginal Fire Damage (MFD).

Sparhawk's model considered both fire suppression (putting out fires that have started) and pre-suppression costs (fire prevention spending). To simplify things, we can assume there are two periods, fire season and pre-season. Pre-suppression spending occurs before the fire season starts (pre-season), while fire suppression (putting out fires) occurs in period two. Figure 2 above can then be treated as the optimal amount of fire suppression during fire season. All pre-season activities can then be treated as exogenous and pre-determined.

Sparhawk's model introduced a new field of study described as fire economics and his work went onto evolve into what fire economist use today called the Cost plus Net Value Change (C + NVC) model (Simard, 1976; Donovan and Rideout, 2003). This newer approach accounts for the fact that fires do not necessarily only cause damage but can have beneficial effects (such as ecological benefits or preventing more damaging fires in later periods). The goal of this model is to minimize *C* + *NVC*. A Cost plus Net Value Change (C + NVC) model is the foundation for current fire economic theory where the fire managers face two different costs the suppression cost (C) and net fire related damages (NVC) (Donovan and Rideout, 2003). Donovan and Rideout (2003) argued that the min (C +NVC) objective *could* be used in the National Fire management analysis system used by the US Forest Service in a computerized fire budgeting and planning tool (USDA Forest Service 1995). Although currently this not how the Forest Service request funds from Congress, it could be used in the future for that purpose if the estimations were accurate.

This present study is looking specifically at fire suppression costs. While pre-suppression expenditure is important and should be analyzed in additional literature, this study decided to look only at suppression costs.

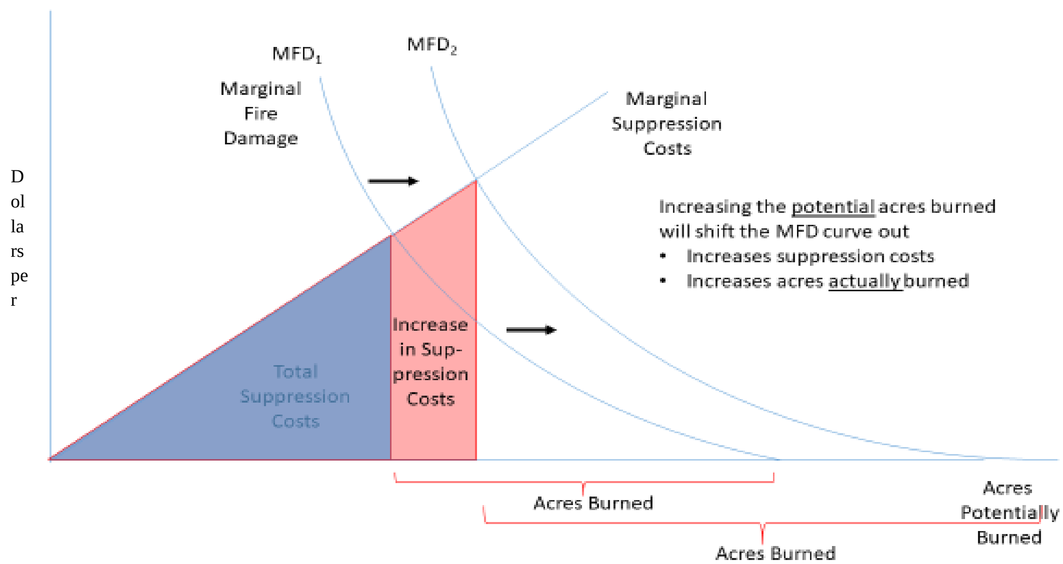
Figure 3. Observable variables in optimal fire suppression



Several empirical studies have focused on estimating the role of different factors in explaining fire suppression costs. Figure 3 shows why these variables are frequently studied. Even though they drive optimal fire suppression decision, neither marginal fire damage (MFD) or marginal suppression costs (MSC) are directly observed. Nor do we observe the total acres that could be potentially burned if there were no fire suppression. What we do observe (or at least have estimates for) are acres burned and total fire suppression costs. These are the variables that applied researchers have tried to explain. Figure 3 shows how shifts in the marginal curves could affect acres burned and total suppression costs. Factors increasing suppression costs (shifting the MSC curve upward) would increase acres burned. Effects on total suppression costs are less clear and will depend on how the MSC curve shifts. Shifts in the marginal damage of fires would affect suppression costs and acres burned. If the MFD curve shifts to the right (as potentially higher valued areas are threatened) acres protected and suppression costs would increase.

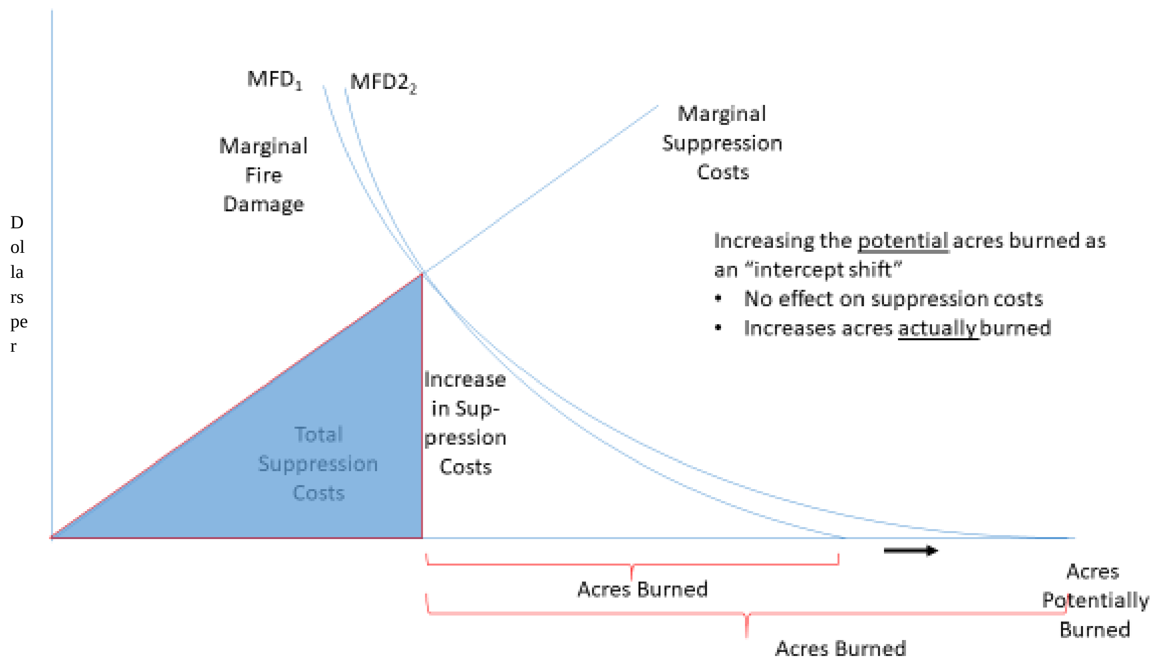
As the potential area burned increases, the x-intercept of the MFD curve would shift out. All else equal, this would increase acres burned. Its effect on suppression costs is ambiguous. If it is a simple intercept shift, then total suppression costs would also increase. Figure 4 shows a case where the potential acres increase shifting the marginal fire damage curve from MFD_1 to MFD_2. The nature of the curve shift increases the optimal level of fire suppression. In this case both acres burned, and optimal suppression costs increase.

Figure 4. Marginal fire damage curve shift increases both acres burned and suppression costs



One can construct a counter example, where the acres potentially burned increases, but optimal suppression costs remain unchanged (Figure 5). Even though there is not a direct way to observe MFD, MSC, acres saved, and potential acres burned, we can hypothesize how different factors might cause these curves and variables to shift. Possible ways these curves might shift are with variables such as number of people, houses, businesses, road density, endangered species, or cultural significance. Variables that might shift the marginal fire damage or the marginal suppression costs to create new equilibriums is extremely important and improving the cost plus net value change modeling.

Figure 5. Marginal fire damage curve shift only increases acres burned but not suppression costs



Factors Affecting the Rise in Wildfire Suppression Costs

Although the Office of Management and Budget (OMB) required the development of a new tool to predict suppression costs in the late 1990s, Congress did not change allocation systems until 2009. In 2009, Congress passed the Federal Land Assistance Management and Enhancement (FLAME) Act which had three major objectives: restore and maintain resilient landscapes, create fire adapted communities, and for wildfire response (WFLC, 2009). Before 2009, the only funding that Congress provided to both the DOI, and FS was via the Wildland Fire Management (WFM) account which was showing signs of inadequacy by 2009 (Hoover et al, 2015). However, Congress required that the FLAME account could only be used after certain requirements were completed. One of these requirements is that the Wildland Fire Management (WFM) account is depleted for the fiscal year, and the management team needs more resources. The DOI and the FS use the WFM accounts for preparedness and suppression, but the FS additionally uses the account for emergency stabilization and initial site rehabilitation activity, while the DOI primarily uses other fire operations programs to fund stabilization activities (Hoover, 2017). This process is called the Burned Area Emergency Response (BAER), the program pursues many activities, including post-fire analysis of suppression costs. Congress, with the management agencies, has been trying to address the observable change in wildfire events for the last 30 years. Although much work has been done to help with fire management, some literature (Baylis & Boomhower, 2019) indicates federal subsidies for home damages are driving the increase in suppression costs. Protection for private homes account for a large portion of the spending by the federal and state governments. Arguments have been made that development in the Wildland Urban Interface (WUI) by local jurisdictions result in externalities for federal and state suppression expenditures, and therefore, homeowners should help mitigate the costs (Baylis & Boomhower, 2019).

Martin et al. (2011) conducted interviews of policy experts, the public, and informed lay public about the need for active land management and the treatments that could be used. Interviewees were from Forest Service Region 3, which includes Arizona, New Mexico, and a small part of Texas. Martin et al. (2011) concluded the common agreeance that the national forests should be actively managed but differ significantly how to make these treatments (Martin et al. 2011). Another common theme in Region 3 is mistrust between stakeholders and the Forest Service (Martin et al. 2009, Lien et al 2021). Lines of communication and educational programs could make a significant impact in providing new alliances to help combat regional fire suppression issues. Land management agencies, such as the Forest Service, need good relationships with the people in their area to help combat the various issues that arise especially with fighting fires. Mistrust leads to individuals not following the Forest Service recommendation on best practices or helping fire resource management with fuel reduction around their residence.

Finkelestein et al. (2005) found evidence that fires in the West are common in geological history. Sedimentological evidence suggests the area has gone through periods of seasonal aridity, which increases the fires. Land management agencies have an objective to find a balance between maintaining natural fire cycles and protecting vulnerable areas. Land management agencies have adopted active management policies to find an equilibrium between these two objectives. Several studies have identified that there is a need for fuel reduction to help reduce fire severity and therefore fire suppression costs (Calkin 2013). In the past, fire suppression was aggressive by immediately putting out fires as quickly as possible, which resulted in larger fuel reserves to create larger fires. Calkin (2013) discusses the benefits of allowing fires to burn to reduce fuels that could create even larger fires in the future. The theory suggests that allowing fires to burn will decrease suppression resources required and therefore reduce costs and allow for natural fire cycles as part of a healthy ecosystem (Calkin, 2013). Allowing the fires to burn is likely the cheapest option, because it requires no effort by Forest Service teams, but has the most uncertain outcome. An alternative method was suggested to help reduce wildfire suppression costs through hazardous fuel reduction treatments (Thompson et al, 2012). In the dryer Ponderosa Pine forests in the Northern Arizona landscapes where fire characteristics are highly dependent on fuel (Fitch et al, 2017), having fuel reduction teams that take an active management approach provides more certainty of outcomes by reducing fuel loads. But, this requires upfront capital expenditures. These methods use tactics that are physically preventative to reduce fuel loads in the national forests to bring down costs.

Physical attributes are not the only reason for the rise in suppression costs as several studies investigated how management agencies can be more efficient with their resource allocation. Optimizing resource location to stations with high dispatch frequencies increases resource availability on high fire days (Lee et al, 2012). By prestaging resources in dispatch areas, less money would be required to move these resources when fires start. Strategic equipment placement with more upfront capital expenditures in locations with high wildfire probability can also decrease the overall costs. More efficient use of resources can also reduce overall suppression costs. Controlling fire lines with bulldozers and fire engines on natural breaks as an alternative to fire crews has a statistically significant impact on the success of suppression (Katuwal et al. 2015). Properly using resources such as boulders and fire engines on fire lines rather than firefighting crews would reduce costs by being more efficient than paying for manpower that is not as effective. Optimizing physical attributes and management resources are possible strategies for reducing suppression costs.

Chapter 3: Empirical Studies to Explain Suppression Cost

Suppression cost itself has many variables that might affect it, and therefore, there are several publications on what might affect suppression costs (Donovan 2011, Prestemon & Donovan 2008, Yoder & Gebert 2012). A large focus of the literature are the effects of homes near fires and in the Wildland Urban Interface (WUI), studies investigating these effects in different regions of the West find contradicting results (Liu et al. 2015, Donovan et al, 2004). Wildfires in the Sierra Nevada areas of California had a statistically significant relationship between suppression costs and houses within 6 miles of initial start location (Gude 2013). In contrast, a study of wildfires in Oregon and Washington found no statistically significant relationship between suppression costs and the total number and density of houses (Donovan, 2004). Donovan et al. 2011 investigated the effects of news coverage and political pressures on suppression costs. They conclude that spending increases with news coverage due to possible personal liabilities of fire managers for their actions (Donovan et al., 2011). Both biophysical and non-biophysical variables seem to have as significant impact on fire suppression costs

(Donovan et al., 2011). Biophysical variables included fuel loads and weather while non-biophysical variables were newspaper coverage and political pressures (Donovan et al., 2011).

Gebert and Schuster (1999) considered basing estimates of suppression costs on regression analysis instead of a moving average. The linear regression model became a standard by both the US Forest Service and the Department of Interior during the height of fire season (Abt et al. 2009). The linear regression modeling that was designed greatly improved fire suppression cost estimates by each month, but significantly lacked the ability to predict fire suppression costs farther into the future. In the late 1990's, Gebert and Schuster (1999) conducted regression analysis, using national monthly fire suppression expenditure as the dependent variable, National Fire activity variables consisted of regional, fire activity data collected from Incident Management Situation Reports (SIT Reports) as independent variables, but did not include firefighting resource variables. They argued that the regression models that were developed would be better at predicting suppression costs with an adjusted R-squared value ranging from 0.696 to 0.969 (Gebert and Schuster, 1999).

Subsequent work expanded on Gebert and Schuster's (1999) work. Prestemon et al. (2008) extended the forecast time to a year in advance, showing improvement over using the .10-year moving average of suppression costs. Abt et al. (2009) developed a forecast that started in November. Based on model results, they concluded suppression costs for FY 2009 would significantly exceed the 10-year moving average. Therefore, Congress would have to allocate additional resources for wildfire suppression. The prediction was born out as true as Congress appropriated additional funds for wildfire suppression (Hoover, 2017). Gebert et al. (2006) estimated the determinants of suppression costs, analyzing data from over 1550 fires. Major explanatory variables included log of acres burned, aspect (sine), fire intensity level, an energy release component, and WUI, all of which were statistically significant. Housing values were not statistically significant.

The most common variables discussed in the fire suppression cost literature are physical attributes of the area burned, such as slope, aspect, elevation, drought conditions, fuel type, distances from roads etc. These variables are likely to increase marginal suppression costs and/or marginal fire damage. For example, slope with significant distance from roads may increase the marginal suppression costs because air units would be the only way of putting out fires. These units are significantly more expensive than ground based units. Marginal fire damage theoretically will increase with high fuel loads becoming more flammable as extreme droughts conditions become more common in the Western United States. Physical attributes are not the only possible reason for shifting these two curves. Socioeconomic, political and management factors also may influence the shift. For example, management practices such as not having adequate resources in strategic areas can result in higher transportation costs. Additionally, political pressures theoretically increase suppression costs by fire management teams feeling personally liable for perceived mistakes during firefighting efforts. Because of limited data availability, this study focuses on physical attributes.

Data

For this project, the initial plan was to gather data based on a review of empirical studies. This would be done to identify commonly used variables in previous research, focusing on those with the most potential to aid the Forest Service in predicting large fires. Relying on past studies was less helpful than hoped because most studies use ICS-209 reports. Incident Status Summary (ICS-209) reports are maintained as part of the US National Fire and Aviation Management Web Application program. They report on large wildfires fires and other natural disasters on federal lands (Short, 2014; Mangeon et al. 2015). The ICS-209 reports submitted by fire management teams daily during a confrontation with a fire. Therefore, assessments of what is happening on the ground have much uncertainty as events shift rapidly. Final reports are required to be submitted quickly. Data from these reports may not be consistent with satellite imagery data (Mangeon, et al., 2015). Also, some data available to federal agency staff through the ICS-209 program are not readily available to the general public. Given these issues, a different data source was sought for this study as an alternative to the ICS-209 reports.

Another common database used in the literature was the National Interagency Fire Management Integrated Database (NIFMID), which was suggested by the Congressional Research Service and several other sources. The NIFMID database, however, seems to be going through renovations. Web portals to publically access the data do not appear to be active. This hampers the ability of this present study use data from previous studies carried by federal agency staff. For example, the NIFMID reported on fire height, a variable not available from other publicly available sources.

National Interagency Fire Center (NIFC) website maintains data under Wildland Fire Interagency Geospatial Services (WFIGS). This new database was created on June 7th, 2021. NIFC previously maintained the NIFMID database. Table 1 provides an example of publically available data from the WFIGS system. Data include total acres burned as well as acres burned by landownership type (e.g. federal, state, county, city, etc.) and by federal agency managing the federal lands affected.

Unfortunately, data are not available for suppression costs for fires. Additional problems were found when the two datasets available through this system had a significant number of missing data entries for several variables. The first dataset called “wildlife fire locations” does not contain complete or relevant data for the years used in this study. The second dataset called “NIFC dataset full history” has only a few variables including spatial information, sources of fire location and acres burned. Again, this dataset has no suppression costs for each individual fire and although the dataset is complete with very little missing values it just does not have anything extremely useful for this analysis. This paper instead used data from multiple different agencies and private companies such as BAER, Zillow, Forest Service region three, United States Geological Survey (USGS), ArcGIS, western region climate center (WRCC) and the US Drought Monitor.

Table 1. Example of WFIGS Data

Ownership	
Total	189886
Blank	67900
"null"	4
#	38
ANCSA	290
BIA	24643
BLM	14369
BOR	93
City	286
County	294
DOD	405
DOE	36
Foreign	5
NPS	3978
OthLoc	31
Private	30010
State	3059
Tribal	3198
USFS	39396
USFWS	1851

Suppression costs and acres burned are the most important variables in this analysis (specifically for each individual fire). The only database with easily accessible data available to the public for both these variables was the Burned Area Emergency Response (BAER) database, maintained by the U.S. Forest Service. The goal of BAER is to assess post-fire impacts to watersheds and land area affected by wildfires. BAER reports are used to assess watershed conditions and to determine the post-fire level of risk to human's property, critical natural areas and cultural sites. These may include habitat deterioration, reduced water quality or risks from flooding or mudslides. Reports are used to determine whether it is necessary to enact emergency stabilization measures. The downside of the BAER reports is that they are based on large teams staffed with hydrologists, economists, engineers, geologists, biologists, botanists, archeologists, GIS mapping specialists, and recreation specialists, and are so, relatively expensive. As a consequence, BAER reports are only developed and published for relatively large fires. Although the database does not report on smaller wildfires, it does include historical information on both acres burned and suppression costs for individual fires.

After determining the data source for the dependent variables, the next step was to identify the independent variables to be used in regression analysis. Physical attributes of terrain, location, time, vegetation, drought, and housing were the primary variables that would be desired for the analysis. The first variables found were in the United States Geological Survey (USGS).

The USGS provides scientific analysis and information about the water, energy, minerals, and other natural resources. The USGS maintains downloadable digital terrain models (DTM) from its 3D Elevation Program (3DEP). The models are formed from the National Enhanced Elevation Assessment results which is from Light Detection and Ranging (LIDAR) imagery from satellites. The DTM is a computer graphic representation of terrain and therefore can be used to find elevation, slope, and aspect. Slope and altitude are commonly discussed variables in everyday vocabulary, but aspect is the cardinal direction that the terrain is facing. For example, if an individual is hiking in the mountains and is facing exactly south on the slope, the aspect would be 180 degrees. This study developed variables from this data specific to fires in Arizona. The 3DEP product that was used in this analysis was a 1/3 arc second, DTM which has an approximate 10-meter resolution. In Arizona there were 78 images that were exported from the USGS website.

The variables were then transformed into 10 mile radius circles around the start locations so the variable would be independent of final acres burned. Hand et al. (2016) argued that fire suppression costs were often functions of the broader geographical surrounding the initial fire ignition site. Further, they argued that analysis that only consider land attributes right at the ignition site would fail to account for their features. They acknowledge that suppression costs and acres burned could be simultaneously determined and find evidence of endogeneity of acres burned their initial suppression cost regressions. They use instrumental variable (IV) techniques to account for the endogeneity of their acres burned variable. Yet, many of their other variables are determined over the acres burned. This means that the physical attribute variables they use are themselves functions of acres burned and quite possibly endogenous too. This construction has two drawbacks. First, there is the potential for simultaneity bias in regression estimation. Second, because these variables depend on final acres burned they cannot be used for forecasting suppression costs as costs and acres burned are determined simultaneously. Rather than use physical attribute variables that are functions of acres burned, this study considers physical attributes around a fixed radius.

The reason to add these variables is because if the medians of these different variable are high it could possibly suggest that the terrain is difficult and likely would require helicopters and other aerial firefighting equipment increasing the costs. Another possible reason to add these variables is because altitude and aspect help determine vegetation type and density. For example, vegetation at 5,000 feet and westward facing has significantly different vegetation then northern facing vegetation at 2,000 feet. These variables would have a significant difference in the firefighter’s strategy and likely the suppression costs of putting out fires in this region. The theory for these variables is that if altitude is higher, the suppression costs will be higher. This is especially true around 5,000 feet where the junipers, which are difficult to extinguish, thrive. Higher slopes can increase suppression costs because they can prevent vehicles and personnel to reach these areas. Lastly, aspect can determine if areas are more southern and western facing, affecting the amount of sunlight and rain they receive.

The next set of variables that were developed characterized the type of vegetation, road system and Wildfire Urban Interface (WUI) surrounding the point of fire ignition. These variables were located within the Forest Service (FS) Region 3 geospatial database which, also had historical GIS data on the perimeter and fire start location. The topics in the feature class range from aerial photography to forest management and planning and was extremely beneficial to this study. This research took a small amount of information from the FS website and downloaded the feature classes that were vegetation, roads, and WUI to get several variables. Physical variables were measured over a 10 mile radius around the start location.

The Region 3 database had a feature class on vegetation which was created by Oklahoma State University Institute of Natural Resources Existing Vegetation (INREV). The existing vegetation map provides basic information on current conditions of vegetation structures and composition from 2004 to present. The variables gathered from this feature class was “tree canopy cover,” “shrub cover”, and “herb cover” (cover of herbaceous plants) in percentages for each individual area in Arizona. The default variable was land covered by bare soil. The data gathered was by specific vegetation regions such as the Sonoran Mojave, Sky Islands, Mogollon Rim, Colorado Plateau Highlands, and the regular Colorado Plateau. The reason for this subdivision is because the dataset is massive and when importing into ArcGIS pro the INREV is so large, ArcGIS pro struggles to even render it. However, once the five different datasets are imported into ArcGIS, they can be merged using ArcGIS geospatial tool merge to create a

INREV dataset for all of Arizona. Again, to make these cover variables independent of acres burned, the area chosen was a fixed circular radius of 10 miles around the start location.

Region 3 also had two feature classes with data on Arizona roads and trails. The difference between the road data versus the trails data is that the road database is classified as a motor vehicle travel way over 50 inches wide, while trails were everything else including off highway vehicles (OHV) roads. In this study, only the roads database was used. The logic behind excluding trails is that they would not be useful for moving fire equipment. Unlike roads, they would not allow vehicles passage for fire suppression. The road feature class was downloaded into ArcGIS pro. Again, to have the variable be independent of the acres burned, the denominator (area) of a road density variable was a circular radius around the start location of 10 miles. The reason this paper chose to include roads in this analysis is there an intuition that if fires are easily accessible then they're going to be easier and cheaper to put out.

The last variable to be generated from the FS Region 3 database was the Wildland Urban Interface (WUI). The reason to include the WUI especially with a 10-mile radius of the starting point serves as a proxy for resources at risk and wildlife managers' possible reactions to protect them. A potential theory is that if a substantial amount of area were WUI within a 10-mile area of a fire starting the amount of money spending to either suppress the fire or control the fires movement would increase. However, another possibility is that if a fire is not close to wildland urban interface and it is in a much more remote place the fire manager might suggest letting the fire burn and not spend time and resources trying to manage it. The WUI is defined by the United States Forest Service as any area within or adjacent to an "at-risk Community" (US forest service). The WUI dataset managed by the Forest Service in Region 3 differs from this definition by digitizing sources only from NEPA status or on potential status as interpreted by fire analysts. In the WUI dataset NEPA status are from 2001, while fire analysts make up the other portion of the feature class. With both the WUI and road feature classes in ArcGIS pro the discussion of the correlation between the size of the fire and how much area is a road or WUI would likely skew the results. To mitigate this issue, it was necessary to specify a standard area around a fire's starting point. The FS Region 3 database had historic fire start locations. This was used to create a circular buffer zone around each fire location to get a better representation of the area around the initial fire. The feature class of fire history occurrences was also downloaded into ArcGIS pro. Again, to prevent the variable to have influence due to the acres burned in the denominator the area was just a circular radius around the start location. Within the 10-mile radius of each of the fires start location only 51 out of 127 of the fires had any area within the WUI. Indicating that a majority of the fires did start within a 10-mile radius of a WUI.

Population data for U.S. zip codes was available from ArcGIS pro online packages. The ArcGIS package layer is called USA Zip Code Areas, which has five-digit zip codes for all the US postal service locations the dataset has zip codes divided the United States into 10 large groups and are numbered from zero to 9 (ESRI Website). The USA Zip Code Areas only have population for the fiscal year 2020. These digits in conjunction with the first digits represent a section centered facility or a mailing processing facility area. ESRI provides this data with zip codes, postal district names, population per square mile for 2020, in the zip code in the United States. The zip codes were then clipped with an Arizona county map to just get the zip codes in Arizona instead of the whole United States and then imported into ArcGIS pro. Although all the variables were included from the USA Zip Code Areas only population per square mile was used in the final analysis.

Weather data were obtained from the Western Region Climate Center, which has an interactive map of all the weather stations in the Western United States. The center has a specific subsection of Arizona. Key weather station locations that were relevant and relatively close to the fire locations were identified by visual inspection. Data used to develop explanatory variables included s solar radiation, wind speed, air temperature, relative humidity, and precipitation. All these attributes can potentially affect how difficult it can be to contain a fire. These data are reported in terms of longer-term averages and so may reflect more long-term climate effects than shorter term weather effects.

The United States Drought Monitor Center jointly with other national agencies such as NOAA and Department of Agriculture releases data once a week that identifies areas of drought by county. The data is classified into five categories: abnormally dry(D0), moderate(D1), severe(D2), extreme (D3) and exceptional (D4) drought. When multiplied by a series from 1 to 5 with relation to the severity and then added together, the classes make up the drought severity and coverage (DSC) index. The DSC index is used instead of the classes as it is continuous, helps indicate the total area drought severity, and summarizes drought effects in just a single variable. The DSC index is summed over two different periods. The first measures drought severity 30 days before a fire has started and the second measures drought conditions the seven days after the fire starts.

Finally, data on the value of homes around the point of fire ignitions was obtained from the real estate's locator company, Zillow. Zillow data was to measure home values at the zip code level. To access this data a clip geoprocessing tool was used in ArcGIS with the county layer so that only Arizona postal zip codes would be created in a new layer. The data from Zillow is from their Zillow Home Value Index (ZHVI) which is a smoothed, seasonally adjusted measure of the typical home value and market changes across a given region and housing type. The ZHVI was gathered for the whole time period of the fire occurrence from 2002 to 2020 and the median value was taken over the time period. The index reflects the typical value for homes in the 35th to 65th percentile range.

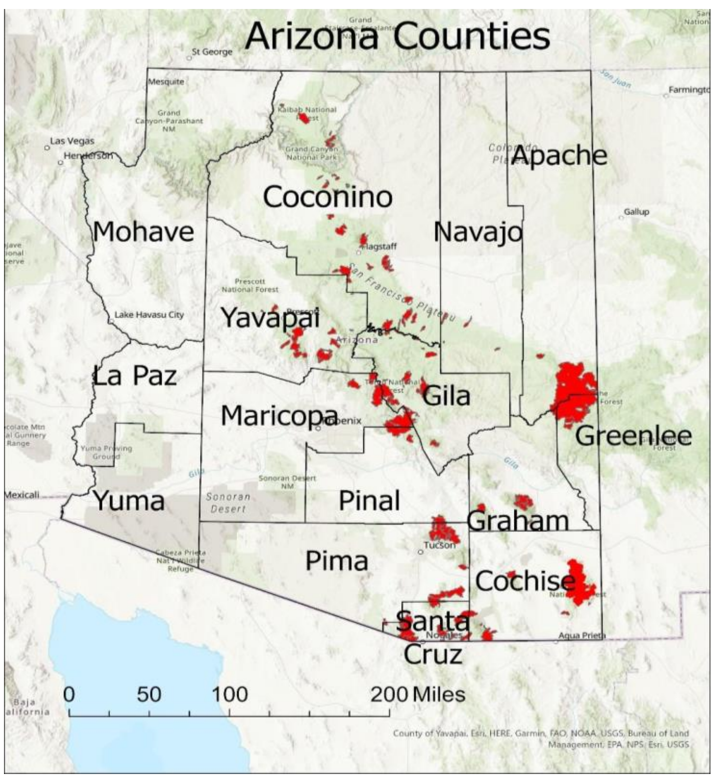
Table 2. Complete list of Variables and Locations in Research

Variables	Variable Definition	Units of measure	Source
Year the Fire Started	The year the fire burned	Years	BAER
Fire Name	The official name given to the fire	Name	BAER
Suppression Cost	The amount of money to put out a given fire	Real US dollars deflated using the GDP price deflator	BAER
Median Altitude	the median altitude of each fire 10-mile buffer around fire start points	Meters	UDGS
Median Slope	the median slope of each fire 10-mile buffer around fire start points	Degrees	UDGS
Median Aspect	the median aspect of each fire10-mile buffer around fire start point	magnetic heading	UDGS
road length	roads 10-mile buffer around fire start point	miles	USDA FS
Wildland Urban Interface Area	WUI area 10-mile buffer around fire start point	square miles	USDA FS
Tree Canopy Cover	Total Tree canopy cover within 10-mile buffer around fire start point	percentage	USDA FS
Plant Shrub Cover	Total Shrub cover within 10-mile buffer around fire start point	percentage	USDA FS
Plant Herb Cover	Total Herb cover within 10-mile buffer around fire start point	percentage	USDA FS
Mean Population from 2020	Mean population by zip code from 2020	Humans	ArcGIS pro
Mean Population per Square mile 2020	Mean population per square mile by zip code	Humans	ArcGIS pro
Mean Solar Radiation	Mean solar radiation 7 days after the start of the fire	Total light	WRCC
Mean Wind Speed	Mean wind speed 7 days after the start of the fire	Miles per hour	WRCC
Average Max Air Temperature	Average max air temperature 7 days after the start of the fire	Fahrenheit	WRCC
Average Relative Humidity	Average relative humidity 7 days after the start of the fire	Percentage	WRCC
Precipitation	Total Precipitation 7 days after the start of the fire	inches	WRCC
Drought Severity and Coverage Index month before	Sum of Drought Severity and Coverage Index a month before the fire occurs	index no units	USDM
Drought Severity and Coverage Index	Sum of Drought Severity and Coverage Index 1 week during every Fire	index no units	USDM
Average Median Home Value	Average home values by zip code with ZHVI	US dollars	Zillow
Big fires	Fires bigger than 30,000 acres	Dummy	Calculated
Cost per Acre	The suppression cost of each fire divided by the acres burned	US dollars per acre	Calculated
Natural Log of Acres Burned	The natural log of acres burned variable	acres	Calculated
Natural Log of Suppression Costs	The natural log of Suppression Cost variable	US dollars	Calculated
Month Dummy	A one for the month a fire occurred	Dummy	Calculated
Arizona Forest Dummy	A one for the forest a fire occurred on	Dummy	Calculated

The last set of variables were dummy variables that were used created from data that was already in the BAER & Region 3 datasets. The first was month the fire started. The second set of dummy variables identifying the national forest where individual fires The BAER data set only reports on fires starting in national forests. There are six different national forests in Arizona: Apache-Sitgreaves, Coronado, Coconino, Kaibab, Prescott and Tonto.

Table 2 shows the variables that were used in the study their units of measure and the source location as described above. The process of making a complete dataset from all the raw data is described in the section below and the Appendix. The database includes information on acres burned and suppression costs for 127 Arizona wildfires occurring between the years 2002 and 2019.

Figure 6. Final BAER dataset in Arizona Counties



Chapter 5: Characterizing Arizona Wildfires

Figure 6 shows the area of the fires in the data set. The 127 wildfires with acreage burned and suppression cost data represent just a small fraction of the total number of wildfires that occurred in Arizona from 2002 to 2019. It is important then to consider how the wildfires in the sample differ from the entire population of wildfires. BAER reports are only conducted for fires above a certain size, so we already know that fires in the BAER reports will be on average large than the entire population and that information on fires will be truncated from below, with no observations for smaller fires.

Figure 7. FS Region 3 Dataset with 1,222 Fires

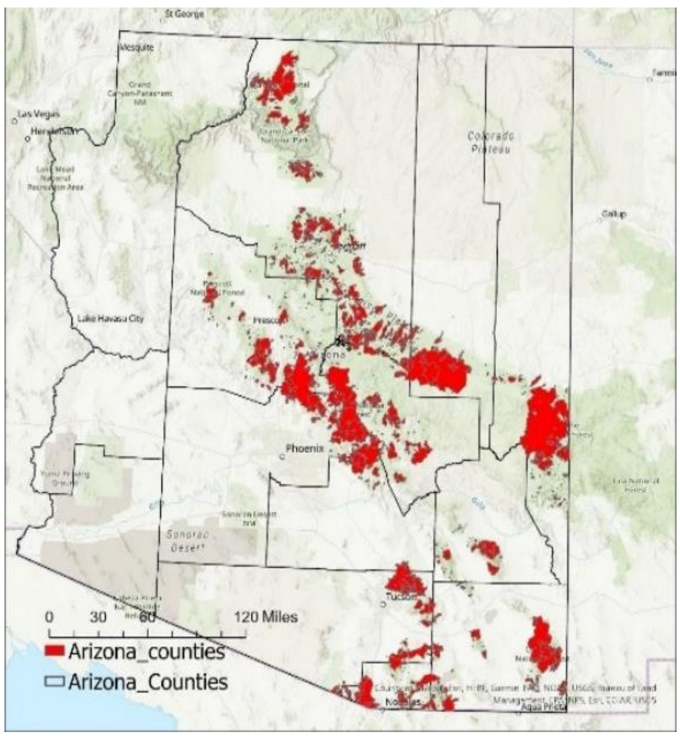


Figure 7 shows the geographic distribution of Arizona wildfires from the US Forest Service Region 3 geodatabase from Fires are primarily in the Mogollon Rim, Grand Canyon region, and the Sky Island region in the southern part of the state.

Comparing the Region 3 dataset at just under 1,222 observations to the BAER dataset at 127 observations, the distribution of the fires geographically is similar. Figure 8 shows the distribution of fires in from the BAER report database. One can see that the general pattern of where fires are occurring is like the overall patterns from the Regions 3 data in Figure 7. Comparing the Figures 7 and 8, the BAER report data retain a good portion of the larger fires.

Figure 8. BAER Reports with 127 Fires

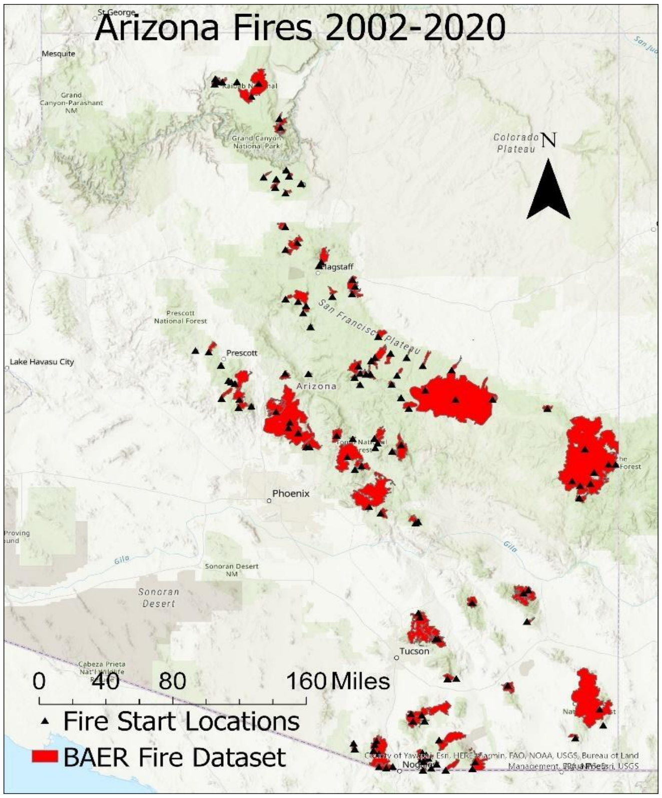


Table 3 compares data from the BAER reports from 2002-2020 with the large Region 3 database. The BAER data do not include any wildfires smaller than 326 acres. The median fire size in the Region 3 dataset is 106 acres, indicating that half the wildfires in total are relatively small in area. For some variables – tree cover, shrub cover, and altitude – the BAER data seem relatively close to the entire Region 3. In contrast, the BAER data seem to exclude (smaller) wildfires that started on steeper slopes. The BAER data also appear to have a smaller proportion of acreage in herb cover. Because the BAER data has less variation in some explanatory variables, this may affect its ability to find factors predicting large fires.

Table 3. Comparing Descriptive Statistics between Region 3 dataset and BAER Dataset

Variables	Region 3 dataset				BAER Dataset			
	Mean	Median	Max	Min	Mean	Median	Max	Min
Area Burned	3,350	106	538,052	-	22,324	4,992	534,639	326
Average Slope	12.3	11.2	44.8	0.3	14.0	14.2	28.3	2.5
Median Slope	11.5	9.8	44.4	-	12.9	12.7	27.9	1.6
% Tree Cover	21.5	22.0	80.0	-	20.0	19.4	79.5	-
% Shrub Cover	13.2	6.6	91.0	-	16.8	10.0	90.0	-
% HERB Cover	21.1	21.5	100.0	-	20.3	16.0	77.4	-
Mean Altitude	1,838.8	1,941.3	2,857.1	464.2	1,794.4	1,800.0	2,646.1	875.6
Median Altitude	1,840.2	1,946.5	2,856.3	418.7	1,787.1	1,776.0	2,652.4	887.1
Aspect MEAN	179.0	179.9	333.6	25.7	172.9	174.5	238.5	106.0
Aspect MEDIAN	179.8	180.1	349.9	17.0	165.3	168.5	274.1	77.4

The BAER reports data reveal that just a small number of fires account for the bulk of suppression costs and acres burned. Just two fires, the Wallow Fire and the Rodeo-Chediski Complex Fire, accounted for one quarter of all suppression costs (among larger wildfires) over nearly two decades (Table 4). Nine fires accounted for half of suppression costs, while the largest 25 fires accounted for 75% of suppression costs in the sample. Suppression costs for the Wallow Fire >$126 million were more than six times larger the Bullock Fire, which was 9th out of 127 fires.

There is a similar concentration for acres burned. Again, the two largest fires, the Wallow Fire and the Rodeo-Chediski Complex Fire, accounted for one quarter of all acres burned over nearly two decades (Table 5). Just five fires accounted for half of acreage burned, while 19 fires accounted for 75% of all acres burned. Given that a relatively small number of fires account for the bulk of acres burned and suppression costs, it would be useful for fire management to be able to predict what turns a 5,000-acre fire into a 30,000-acre fire or a 30,000-acre fire into a more than 100,000-acre mega-fire.

Table 4. Fires making up a percentage of total Suppression cost

Suppression Costs				
#	**Name**	**25% of total dataset**	**50% of total dataset**	**75% of total dataset**
1	Wallow Fire	$126,193,227.66	$126,193,227.66	$126,193,227.66
2	Rodeo/Chediski Complex	$64,520,129.34	$64,520,129.34	$64,520,129.34
3	Horseshoe 2		$59,160,311.32	$59,160,311.32
4	Frye		$27,423,946.84	$27,423,946.84
5	Woodbury		$25,301,440.86	$25,301,440.86
6	Monument		$23,559,928.28	$23,559,928.28
7	Aspen		$22,420,119.82	$22,420,119.82
8	Cave Creek Complex		$21,541,480.71	$21,541,480.71
9	Bullock		$20,135,890.80	$20,135,890.80
10	Gladiator			$16,819,904.00
11	Juniper			$16,121,808.21
12	Highline			$15,821,507.79
13	Goodwin			$15,469,087.92
14	Slide			$12,390,112.48
15	Schultz			$11,108,824.32
16	Fuller			$11,048,490.12
17	Florida			$10,520,076.80
18	Warm			$10,079,198.26
19	Pinal Fire			$10,020,288.27
20	Boundary			$9,492,904.67
21	Burro 2			$9,438,525.10
22	Museum Fire			$9,108,518.71
23	Sawmill Fire			$8,649,090.93
24	Pk Rat Complex			$8,380,603.76
25	Doce Fire			$8,079,954.96

Table 5. Fires making up a percentage of total acres burned

Fire acres				
#	**Name**	**25% of total dataset**	**50% of total dataset**	**75% of total dataset**
1	Wallow Fire	534,639	534,639	534,639
2	Rodeo/Chediski Complex	462,614	462,614	462,614
3	Cave Creek Complex		246,714	246,714
4	Horseshoe 2		222,954	222,954
5	Woodbury		126,591	126,591
6	Aspen			84,750
7	Edge Complex			71,986
8	Murphy Complex			68,079
10	Warm			58,568
12	Frye			48,302
13	Sawmill Fire			46,991
14	Ryan 2002			38,179
15	Brooklyn			33,550
17	Juniper			30,631
18	Bullock			30,563
19	Monument			30,328

Table 6 shows descriptive statistics for variables to be used in the regression analysis, with data matched to the BAER report data for acres burned and suppression costs for 127 large Arizona wildfires from 2002 to 2019. The mean of area burned is just over 22,300 acres and the median is under 5,000 acres. This suggests that very large fires are influencing the mean. A similar situation is happening with suppression costs as the mean is just over $6 million and the median is $1.45 million. The suppression costs in the research were deflated with the GDP deflator calculation with the base year being in 2020. The reason to deflate the values was to put all the suppression costs in relation to one year. That would allow for comparison of each fire through time as well as giving people a logical reference while interoperating the costs. This also suggests that very costly fires are driving the mean higher, skewing the distribution. The next variable included was the cost per acre and not surprisingly the mean is skewed higher at $486

Table 6. Descriptive Statistics of BAER dataset

Descriptive Statistics of Total Fire Dataset				
N = 127	**Mean**	**Median**	**Minimum**	**Maximum**
Suppression Costs Base year 2020	6,006,509	1,451,783	2,598	126,193,228
Cost per Acre	486.10	250.00	2.26	4,589.50
Mean Solar Radiation	629.50	652.69	190.45	852.05
mean wind speed	5.36	4.90	0.01	11.68
average max air temperature	84.58	83.88	57.13	108.25
Average Relative Humidity	29.84	28.00	10.38	71.63
Precipitation	31.72	29.38	0.00	74.02
MEDIAN altitude	5,473.87	5,253.19	2,676.98	8,719.45
MEDIAN slope	9.34	8.40	1.31	19.53
MAJORITY aspect	4.50	4.00	1.00	8.00
Average Median Home values 2002-2020	197,855.90	182,724.00	0.00	467,006.50
Population from 2020	10,062	9,650	52	40,507
Tree Canopy Cover within fires	20.01	19.40	0.00	79.50
Shrub Cover within fires	16.76	10.00	0.00	90.00
Herb Cover within	20.26	16.00	0.00	77.43
MEAN_POP20_SQMI	49	19	0.04	1,172
Road Length	1.07	1.04	0.00	6.09
WUI area	2.47	0.00	0.00	28.80
APR	0.08	n/a	n/a	n/a
MAY	0.21	n/a	n/a	n/a
JUN	0.40	n/a	n/a	n/a
JUL	0.17	n/a	n/a	n/a
AUG	0.08	n/a	n/a	n/a
SEP	0.01	n/a	n/a	n/a
Apache-Sitgreaves	0.13	n/a	n/a	n/a
Kaibab	0.10	n/a	n/a	n/a
Coronado	0.33	n/a	n/a	n/a
Coconino	0.15	n/a	n/a	n/a
Prescott	0.09	n/a	n/a	n/a
Tonto	0.20	n/a	n/a	n/a
DSCI one week during fire	213.05	200.00	0.00	1,000.00
DSCI month before	806.13	800.00	0.00	2,500.00

per acre compared to the median of $250 per acre. An additional note with costs per acre is that the minimum costs per acre for this dataset is $2.25 which would suggest very little suppression with fire management on some of these fires compared to other fires with the maximum cost per acre being $4,589. This suggests a lot of variation with suppression expenditure per acre indicating different management techniques for different fires.

Weather data include mean solar radiation, mean wind speed, average max air temperature, average relative humidity, and precipitation. The mean solar radiation for this dataset is 606 Langley, the median roughly being the same at 620 Langley where the range goes from 196 to 887 Langley. This suggests that there are no days in fires that did not get at least some type of solar radiation. The mean and median wind speed are relatively the same at 5.2 and 4.6 mph and the range is from basically calm at 0.7 to 13 mph. What is interesting about this statistic is that these wind speeds are relatively calm in comparison to most weather patterns in AZ. it might be suggested that these weather stations are in more sheltered areas making winds appear calmer than they are. The average max air temperature was a surprise at 84.6 degrees Fahrenheit with the median being at 83.8 degrees Fahrenheit in the range going from 54.5 to 106.8 degrees Fahrenheit. The range is the most surprising with fires being in relatively cool weather all the way down to 54 degrees and having very little skewness in distribution. The average relative humidity was 31.5 percent which seems a little high although most of the fires were in monsoon season, a possible explanation for this type of mean. The main precipitation was 0.81 inches and the medium being 0.07 inches this suggests that most of the values are zero. This suggests that a lot of the time during fires there is no precipitation in the dataset. However, the range from 0 to 8.32 inches suggests that there are a few fires that have significant rainfall during the time of the fire. This suggests that the variable might have some ability to predict suppression costs.

The next variables are altitude, slope, and aspect with the first set being the mean values and the second set being the median values. Looking at the mean of altitude versus the median of altitude there is a difference of only 7 meters with mean altitude at 1794.4 meters versus median altitude at 1787.1 meters, The mean altitude at just under 1,800 meters or around 5,900 feet. This suggests that on average the fires are in the Mogollon Rim region of vegetation according to the US Forest Service. This area is predominantly Ponderosa Pines, Gable oaks, White Douglas fir, Engelmann Spruce, and Cork Bark Fir suggesting moderate to high tree canopies resulting in more fuel for fires. The next variable is slope, with the mean and median only differing by 1.1 degrees. The mean slope value of the fire area is 14 degrees. This indicates that most of these areas would be very rugged for ground fire crews. If we look at the last variable in terrain called aspect, the mean value was 172.9 magnetic degrees., Looking at the median value of 174.5 degrees indicates that many of the slope faces are pointing to the south. Southern facing slopes can get more direct sun, which could result in more and drier vegetation, compared to north facing slopes.

The next variables are population and median home value. Both are meant to capture fire service management's likelihood of intervening in more populated areas with more expensive homes. Home values for each zip code were found through Zillow for the datasets range which was from 2002 to 2020. The Median Value was taken over this time as household values vary wildly dependent on US economy and would greatly impact the end values. Given that zip codes overlap with the buffer zones an average of all zip codes within the region were taken to find the average of the areas median home values from 2002 to 2020. The average median home value is just under $200,000 with the range being from $0.00 (indicating there are no houses within the perimeter each fire) to $467,000. Comparing that to the Arizona market where the median home value is listed at $225,500 from 2015-2019, suggests areas around the fires in the dataset less valuable on average (US Census Bureau). The next variables are mean population and mean population per square mile. These variables can be used to test the hypothesis is that if there are more people in the area, there is a greater likelihood that the fire management team will have more incentive to spend money suppressing the fire. The mean population per square mile is 50 compared to the median at 19 suggesting there are certain fires closer to populated areas, which skews this variable. Looking at the mean population in the areas where fires occur, there are only about 10,000 people with a median of 9,600 people, suggesting these areas are not very highly populated.

The next variable is the drought severity and coverage index, which monitors drought conditions. It has a mean of 580.5 and a medium 291, indicating that high numbers are pushing up the average. For example, the maximum value is 4175.

The last three variables deal with vegetation cover. This can be important because it indicates what type of fire the managers are dealing with. For example, a shrub or herb fire is much different than a tree fire and will likely affect suppression costs in a different manner. The default variable is bare soil for all the percentages of coverage for tree, shrub, and herb. This indicate if the satellites don't identify the three coverage types it classifies as base soil. The mean of tree canopy cover is 20% with the median being 19.4%, with a maximum of 79.5%. Shrub cover has a mean of 16.8%, a median of 10% and a max to 90%. Lastly, herb cover has a mean of 20%, a median of 16%, and a max 77.4%. The coverage of a certain type of plant might affect suppression costs and acres burned by it affecting the fire's size and travel ability.

Chapter 6: Regression Specifications

Regressions were run for three different variables: (a) natural log of suppression costs, (b) natural log of acres burned, and (c) a binary variable that equaled one if the fire was greater than 30,000 acres and zero otherwise. The first two dependent variables were estimated using least squares methods, while the third was estimated using a linear probability model (OLS applied to the binary dependent variable.

The BAER fire report only 16 out of the 127 fires were above 30,000 acres burned. Although the United States Forest Service has different categories for acres burned with a major category being 5,000 acres and up. The reason this study chose 30,000 acres is because the

BAER reports are already of large fires, making the USFS classifications unrealistic for this study. Fires larger than 30,000 acres accounted for 75% of acres burned in the sample and more than half of total suppression costs (Tables 4 and 5). Given a small number of large fires account for the bulk off costs and damage, it might be useful to be able to identify what contributes to have such mega-fires.

For each regression, the first specification was a "naïve" model that estimated the variables simply as functions of the month the fire started and the specific national forest where the fires started. The default for national forests was Apache-Sitgreaves National Forest. For months, dummies were included for March through September with October through February acting as a default, low fire risk season.

Next, specifications included, in addition to month and national forest variables: (a) physical variables (slope, elevation, and aspect), (b) weather variables, (c) vegetable cover variables, and (d) population, WUI, development variables. Many groupings of variables were not statistically significant.

Chapter 7: Regression Results

The first three models presented below are "naïve" in the sense that they attempt to explain suppression costs and fire size only with information about the month the fire started and the national forest where the fire started. None of these performed especially well and are not reported here. The R-square for the log suppression cost regression was 0.17, while it was 0.12 for the log acres burned regression. For the big fire linear probability model it was only 0.06. More troubling, none of the month of fire start or national forest variables were statistically significant. In an experiment, the number of month dummies in the regression was reduced and it was found that when only May and June were included, both were highly significant, with positive regression coefficients that were quite close to each other. Subsequent regressions include a dummy variable that equaled one if the fire started in either May or June.

Table 7 reports regression results for log of suppression costs. For this, as well as the acres burned big fires, variables for vegetation cover, drought, population, home values and road density, were always statistically insignificant. As such, results with different combinations of these variables are not presented below.

The R-square for the suppression cost regression was 0.385, with an adjusted R-square of 0.302, indicating a large number of insignificant explanatory variables. The variable South Aspect is a binary variable if the aspect is larger than 150 and smaller than 210, indicating a southward facing slope. As hypothesized, the coefficient is positive and significant. Altitude is also positive and significant, suggesting that fires starting at higher elevations lead to higher suppression costs. The WUI variable is significant but negative. One hypothesis is that fire managers would spend more to protect areas with potential more development of value to protect. There are other reasons why the WUI coefficient might be negative, however. First, a low WUI measure may be a signal of remoteness. Remote fires may involve more costs to get to. Second, in more developed areas, fire managers may act quickly to put out fires, so that they have less time to grow large and require much spending.

Table 7. Least Squares Regression: Log of Suppression Costs for Arizona Wildfires, 2002-2019
Dependent Variable: log of Suppression Costs

Regression Statistics	
R Square	0.385
Adjusted R Square	0.302
Standard Error	1.526
Observations	127

ANOVA

	df	*SS*	*MS*	*F*
Regression	15	162.0112	10.80075	4.637154
Residual	111	258.5385	2.329176	
Total	126	420.5497		

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	10.786	2.409	4.48	0.00
South Aspect	0.882	0.361	2.44	0.02
Altitude_MEDIAN	0.001	0.000	3.89	0.00
Slope_MEDIAN	0.043	0.036	1.20	0.23
WUI Area	-0.085	0.033	-2.63	0.01
Mean Solar Radiation	-0.002	0.002	-1.17	0.25
Mean Wind Speed	-0.188	0.226	-0.83	0.41
Average Max air temperature	0.002	0.014	0.16	0.87
Average Relative Humidity	-0.047	0.015	-3.18	0.00
Precipitation	0.021	0.035	0.62	0.54
May-June Fire Start	1.074	0.314	3.42	0.00
Kaibab	0.395	0.614	0.64	0.52
Coronado	1.270	0.563	2.25	0.03
Coconino	0.686	0.546	1.26	0.21
Tonto	1.566	0.627	2.50	0.01
Prescott	1.537	0.700	2.19	0.03

The only significant weather variable was Average Relative Humidity, which had a negative coefficient. This suggests that more arid (less humid) conditions contribute to more suppression spending. The May-June Fire Start variable was positive and significant. To fire managers, this suggests it might make sense to pay special attention to fires starting in these months. Finally, once combined with other variables, there appear to be significant national forest-specific fixed effects.

Table 8. Least Squares Regression: Log of Acres Burned for Arizona Wildfires, 2002-2019

Dependent Variable: Log of Acres Burned

Regression Statistics	
Multiple R	0.451
R Square	0.203
Adjusted R Square	0.095
Standard Error	1.451
Observations	127

ANOVA

	df	*SS*	*MS*	*F*
Regression	15	59.553	3.970	1.886
Residual	111	233.651	2.105	
Total	126	293.204		

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	6.7105	2.29	2.93	0.00
South Aspect	0.6635	0.34	1.93	0.06
Altitude_MEDIAN	0.0001	0.00	0.71	0.48
Slope_MEDIAN	0.0625	0.03	1.84	0.07
WUI area	-0.0118	0.03	-0.38	0.70
Mean Solar Radiation	-0.0005	0.00	-0.32	0.75
Mean wind speed	0.0244	0.21	0.11	0.91
Average max air temperature	0.0045	0.01	0.33	0.74
Average Relative Humidity	-0.0182	0.01	-1.28	0.20
Precipitation	-0.0048	0.03	-0.15	0.88
May-June Fire Start	0.8586	0.30	2.87	0.00
Kaibab	0.1380	0.58	0.24	0.81
Coronado	0.1495	0.54	0.28	0.78
Coconino	0.5223	0.52	1.01	0.32
Tonto	0.1430	0.60	0.24	0.81
Prescott	-0.2799	0.67	-0.42	0.67

Table 8 presents results of the log of acres burned regression. The model fit was not as good that for the suppression cost regression. The adjusted R-square was only 0.95. In addition, far fewer variables were statistically significant. The May-June Fire Start variable was significant (p = 0.005) and positive. South Aspect and slope had the expected positive signs, but were only marginally significant (10% level). For acres burned there do not appear to be significant national park specific fixed effects.

Table 9. Descriptive Statistic for Large Fires vs All others in BEAR dataset

Descriptive Statistics for Large Fires vs All other fires				
	Large fires		All other fires	
Variables	Mean	St Dev	Mean	St Dev
Number of Days a fire occurs	26.31	16.13	20.31	23.47
Mean Solar Radiation	632.72	110.69	603.11	128.86
Mean Wind Speed	5.38	2.66	5.19	2.43
Average Max Air Temperature	87.54	10.66	84.17	10.70
Average Relative Humidity	24.88	11.47	32.44	13.27
Precipitation	0.89	2.08	0.80	1.40
Mean Altitude of fire	1697.05	440.18	1808.48	365.49
Mean Slope of fire	16.35	6.05	13.68	6.24
Mean Aspect of fire	173.34	16.88	172.88	25.43
Average Median Home Value (2002-20)	183,621.	75,564	199,945	104,505
Population from 2020 per square mile	121.72	285.69	38.91	55.19
Drought Severity and Coverage Index	720.69	868.53	560.34	774.58
Tree Canopy Cover within fires	12.63	16.22	21.07	17.48
Shrub Cover within fires	13.80	12.28	17.19	19.60
Herb Cover	18.30	19.14	20.54	20.79
Road length	0.34	0.20	0.64	0.90
WUI area	80.32	80.39	237.22	236.03

Table 9 compares means and standard deviations of different variables for big fires (>30,000 acres) and other fires in the BAER reports data. The Descriptive statistics show a few differences in the variables between the two subsets. The larger fires some intuitive variables such as the number of days a fire occurs, solar radiation, terrain, DCSI and average relative humidity. Some variables that were not intuitive are precipitation, foliage covering and mean population per square mile. Yet, a visual inspection of the standard errors relative to variable means indicates that there does not appear to be much difference between big and other fires. This is largely borne out in the linear probability model (Table 10).

Table 10. Linear Probability Model: Large Wildfires Fires in Arizona, 2002-2019

Dependent Variable: Binary, =1 if fire larger than 30,000 acres, = otherwise

Regression Statistics	
R Square	0.123
Adjusted R Square	0.005
Standard Error	0.332
Observations	127

ANOVA

	df	*SS*	*MS*	*F*
Regression	15	1.726	0.115	1.042
Residual	111	12.258	0.110	
Total	126	13.984		

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	0.585	0.525	1.12	0.27
South Aspect	0.025	0.079	0.32	0.75
Altitude_MEDIAN	0.000	0.000	-0.33	0.74
Slope_MEDIAN	0.013	0.008	1.66	0.10
WUI area	-0.001	0.007	-0.14	0.89
Mean Solar Radiation	0.000	0.000	-0.99	0.33
Nean wind speed	-0.018	0.049	-0.37	0.71
Average max air temperature	-0.001	0.003	-0.37	0.71
Average Relative Humidity	-0.007	0.003	-2.17	0.03
Precipitation	0.003	0.008	0.36	0.72
May-June Fire Start	0.041	0.068	0.61	0.55
Kaibab	0.046	0.134	0.35	0.73
Coronado	0.042	0.123	0.34	0.73
Coconino	0.000	0.119	0.00	1.00
Tonto	-0.009	0.137	-0.06	0.95
Prescott	-0.224	0.153	-1.47	0.15

The model has a poor fit. This isn't surprising given the low R-square of the acres burned regression and the binary dependent variable. The humidity variable is significant and negative at the 3% level, while slope is positive and significant at the 10%. Interestingly, the May-June Fire Start date variable is not significant, although the coefficient is positive.

Chapter 8: Discussion

Several studies have applied regression analysis to measure factors contributing to larger wildfire suppression costs. They have tended to include acres burned, variables that are functions of acres burned, or both. This can create problems of simultaneity bias. While it is common for studies to use instrumental variable methods to address simultaneity, they in general do not evaluate the strength or weakness of their instruments. Another drawback of using acres burned as an explanatory variable is that regression models have limited value in forecasting suppression costs ahead of time.

The approach taken here is different. It relies on variables that can be used before a fire starts or soon after it does. The models reported above do not "fit" as well as ones that include burned acres. Comparatively, the R-square of the suppression costs equation is low. But the modeling approach here has more potential to forecast suppression costs as a fire starts, instead of "predicting" them after the fact in an ex post regression specification.

Chapter 9: Conclusions

The regression results suggest that Arizona wildfires that start in May and June are positively associated with higher suppression costs and more acres burned. This variable was less able to predict the occurrence of the very largest fires, however. The amount of land in the Wildland Urban Interface (WUI) was negatively associated with fire suppression costs and not a significant predictor of fire size. Past empirical results regarding the WUI have been mixed.

One possible explanation for the negative sign is that it is a proxy for remoteness. Another is that fire managers more aggressively control fires with more WUI area so that suppression costs and burned acres are limited. The association between WUI area and burned was negative, but not significant. Average relative humidity was a significant (negative) predictor of both suppression costs and of very large fires. This variable has not been much used in previous studies and may become important if aridity in Arizona increases with climate change.

Appendix: Database Construction

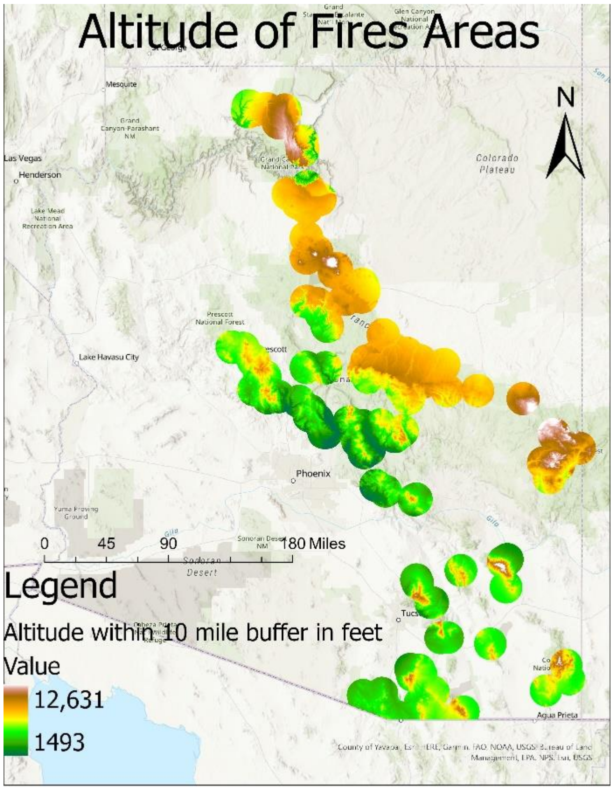
The first step in creating the final dataset was to transfer all the Burned Area Emergency Response (BAER) data from the website onto an excel spreadsheet that had acres burned and suppression costs for each individual fire. There were 140 fires in the BAER dataset that were usable and had the acres burned and suppression costs for the time between 2002 and 2019. The BAER reports had data from previous fires all the way back into the early 80s but the reports had very little information and did not have a lot of structure so those were not included in analysis. The next step was to identify the geospatial area of each of these fires so that physical variables could be calculated. Each BAER report had a photomosaic .tiff image of the fire's perimeter and could be put into ArcGIS pro to create a shapefile of each individual fire. This would be a very laborious task and instead this study decided to use a different dataset called FS Region 3 GIS database, available on the United States Forest Service website. Once downloaded to ArcGIS pro, an excel table was exported so that the suppression cost could be added to the GIS database through formulas. This paper used the data from the BAER reports which had 140 fires and then used a Vlookup function in excel to match fire names of the 1,222 fires from the region three database so that each fire from the BAER reports corresponded with a shapefile in ArcGIS pro. A verification was required to make sure that the BAER reports and the region three database matched up correctly, to do a verification the year the fire was matched between the two datasets and then acres burned were verified for each of the 140 different fires. Going through the verification process there were around 20 that were not the correct year meaning the name of the fire in the dataset region three repeated multiple times. To address this required manually going through the data to find the name of each of those fires in the Region 3 dataset and corrected the error to complete a preliminary dataset. The preliminary dataset had shapefiles of each fire, the acres burned suppression cost, the size of the fire by class, the forest that it was in, and the county that it was in. The number of observations for the preliminary dataset was reduced to 128 unique fires because 7 fires were not in the Region 3 dataset and five had incorrect suppression costs. Once all that was completed the preliminary dataset was imported back into ArcGIS pro so that additional variables could be created for this analysis.

The next variable that needed to be calculated was the natural logs of suppression costs and acres burned. The reason for doing this is because in our sample with such a small number of observations large suppression costs and acres burned will have a significant effect on regression analysis. The natural logs of the two dependent variables will limit the effects that huge values have on the results and an additional bonus is that it creates a different type of regression output for each of the estimated beta coefficients.

Slope, altitude, and aspect variables

The three independent variables from the USGS that were used for this analysis is slope, altitude, and aspect for each individual fire. Each of the 78 images were imported into ArcGIS pro and then merge with the geoprocessing tool into one digital terrain model (DTM). Once the images were merged in ArcGIS pro, a buffer zone need to be created that was independent of acres burned. A 10 mile buffer zone around the start locations of each individual fire were created and then the two feature classes were added together. This was done using a geoprocessing tool called clip to get a DTM for each individual fire. Once there was a DTM for each individual fire the DTM had to be referenced to the name of that fire with the geoprocessing tool called spatial join. The spatial join references the name of the fire to this corresponding DTM which then can be analyzed. The last step for terrain data was to summarize to get the median and mean values for altitude, slope, and aspect for each fire using the geoprocessing tool called “summarize attributes.”

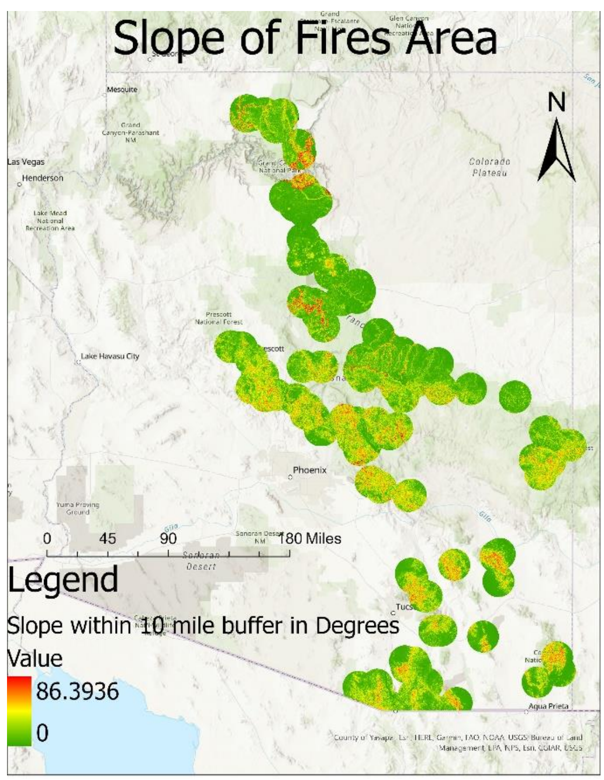
Figure 9. BAER dataset of each fires Altitude



The fire altitudes shown in Figure 9 illustrate that many of the fires are at or above 2,000 meters or around 6,000 feet. This is illustrated by the large number of fires shown in a light orange color, while a much smaller number are in the lower areas of the state. The fires in the lower altitudes are identified by the yellow and green colors and are relatively close geographically. These fires also share the proximity to the Phoenix metro area. The higher altitude fires are more common in the Mogollon Rim or Grand Canyon regions. Finally, the fires in the Sky Island regions are relatively similar with massive elevation gain ranging from 500 meters or 1,600 feet to almost 2,700 meters. These indicate three different styles of fires in Arizona, with the possibility of different management techniques for the different areas.

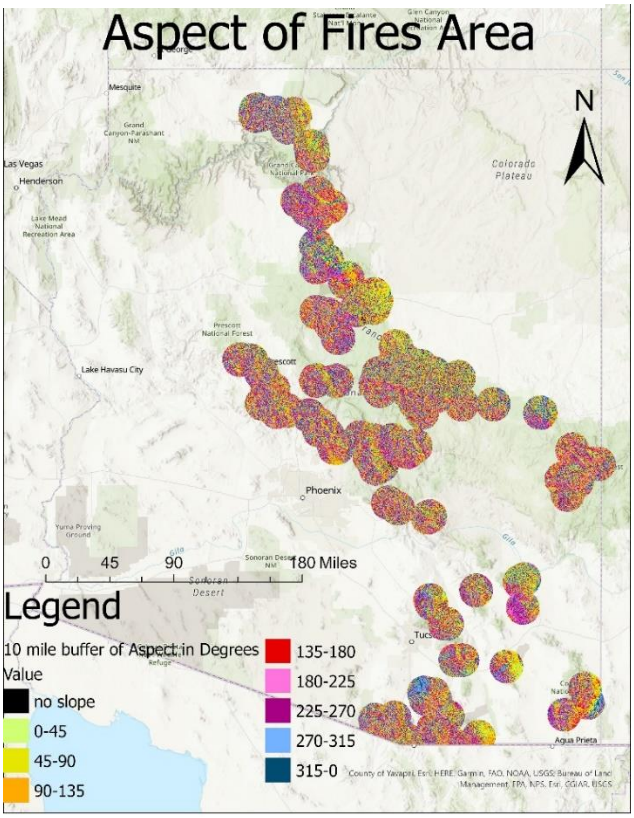
The fires' slopes for the BAER reports are shown in Figure 10. The slopes range from 0 to 84.6 degrees with the lower slopes being represented in green while the larger slopes are red. The larger fires burned in

Figure 10. BAER dataset of each fires Slope



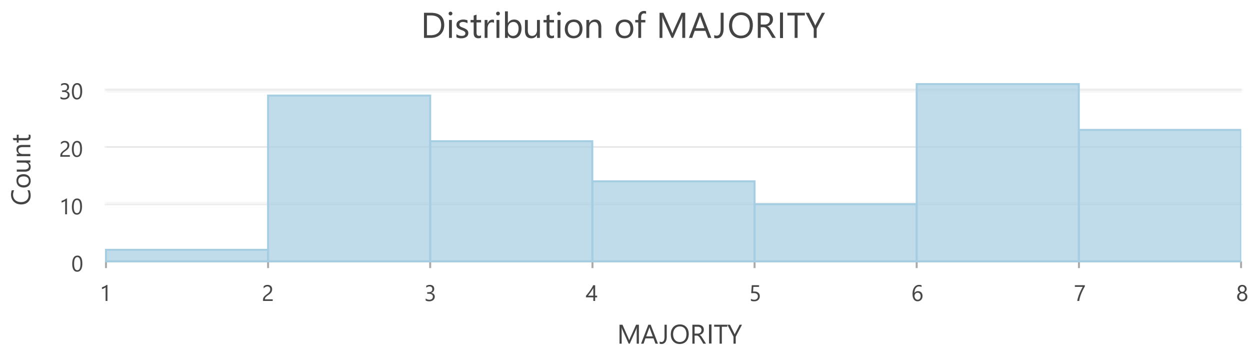
relatively low sloped regions especially in the Mogollon rim or Grand Canyon regions. The more sloped fires occur around the Phoenix area and in the sky island regions in the southeast of Arizona.

Figure 11. BAER dataset Aspect Visual



The fire aspect as described above in methodology is the cardinal direction that the slope is facing. The figure 11 show all the fires' aspects which go from 0 to 360 degrees. The aspect however cannot be continues as the result would always leave the median value to be 180 degrees which is south. Instead the aspect which was created by DTM were categorized into cardinal direction. These directions were North, North-east, East, South-East, South, South-West, West and North-West. The legend shows the aspect in cardinal degrees for example north-east are the degrees of 45-90. Categorizing the variable allows for ArcGIS pro to find the majority in the category which was then used for each fire to determine the majority slope. The majority slope will hopefully show most of the direction for the 10 mile area around each fire for the analysis. Looking at the distribution of the aspects for all the fires in the BAER dataset shown in figure 12.

Figure 12. Distribution of aspect of the Fires in the BAER dataset

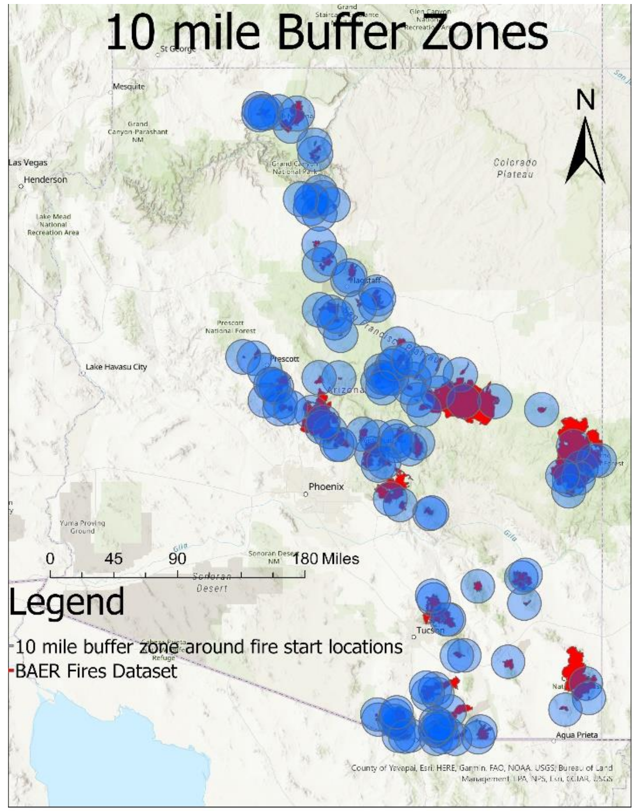


The distribution is basically the same for all different aspects of the fires, meaning that no side of a slope was burned more than another. This observation is also seen in Figure 11, the fires do not have a predominant color. This suggests that the analysis for the variable of aspect will not be statistically significant in this research due, to the lack of variation that might explain the differences in the dependent variables. A suggested assumption was assumption was suggested that the south and southwestern facing slopes would receive more sun exposure resulting in an area that might be easier to burn. It might be interesting in addition research to see if fires start locations in more location with a particular aspect.

WUI and road density variables

After terrain, the variables this paper

Figure 13. 10-mile Buffer Zone around each Fire shown



variables include wildland urban interface (WUI), the number of roads in a fire, and the distance of each road in the fire. However, to prevent a strong correlation between the size of the fire and how much of these three different variables are in the fire, this paper decided to find a buffer zone from each individual fire's identified starting point and find out how much of these three variables were within the 10-mile buffer of the starting location (figure 13). The Region 3 dataset in the wildland fire perimeters has fire history occurrences which is a point layer that is maintained at the forest district level to track the occurrence and the origins of each individual fire. This database has 2,894 observations for all Region 3, which includes Arizona and New Mexico and needs to be refined to just the 131 observations that have suppression costs. To do this, the suppression cost dataset was exported from ArcGIS pro and the occurrences dataset was downloaded into excel and a VLOOKUP function was created to verify it was correct fire years and total acres were compared. Once the verification was found the latitude and longitude of each location of the fire-starting points that were created and imported back into ArcGIS pro to verify on a map that the fire perimeters and fire occurrence points were spatially correlated. There were around 10 observations that needed to be manually maneuvered due to missing data. However, since the fires of every one of the 10 observations was under 1,000 acres putting the fire occurrence anywhere within the boundaries of the perimeter, the buffer would easily encompass the entirety of the fire. Once the fire location points were verified, the 10-mile buffer was created around each point with the buffered geodatabase tool in ArcGIS pro.

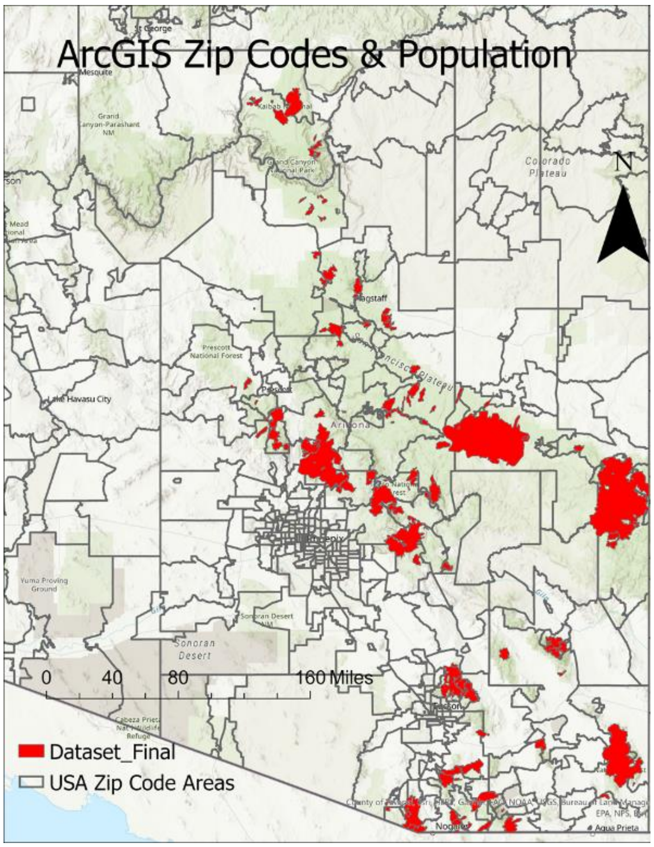
The next step is to use that buffer to clip the road and WUI data that was imported from region three GIS database of Arizona and New Mexico. To get the total length of road in the fire after clipping in ArcGIS, a spatial join was used to connect each fire to the clipped area of roads so that the roads could be identify to the specific fire.

This paper did almost the identical processing for the wildland urban interface (WUI) variable. Again, a buffer of 10 miles was created around each starting location. That buffer was used as a clip for the WUI, that clip was spatially joined to the fire location buffer, and a summarization of that spatial join was created. As explained above the WUI is defined by the United States Forest Service as any area within or adjacent to an "at-risk Community" (US forest service). The WUI dataset managed by the forest service in region 3 differs from this definition by digitizing sources only from NEPA status or on potential status as interpreted by fire analysts. In the WUI dataset NEPA status are from 2001 while fire analysts make up the other portion of the feature class. The result is the amount of WUI area in square meters that was within the 10-mile radius.

Population and home value data

The data in Figure 14 shows the fires that intersect the different zip codes which have population and household information. The household dataset started in 2002 and ended in 2020 which was the total time period of this study. To get only one value for the household values the median home value per zip code was taken. The population for each zip code was unfortunately only found in the ArcGIS pro database and only 2020 population data was available. Some of the fires are in multiple zip codes which causes difficulties. Therefore, the study took the average of the median household value and average population for the different zip codes. Some of the zip codes do not have any population or household values because the zip code is completely in the national forest. The zip codes are outlined in grey, and the fires are shown in red. Small zip code polygons indicate an area of higher population such as Phoenix or Tucson. Other areas have zip code polygons larger than all the fires combined, indicating there is almost no population in those areas.

Figure 14. Zip Codes overlay with BAER dataset



The study utilized data from Zillow to measure the value of homes surrounding fire ignitions. A five-mile buffer was then created around each individual fire to identify the population that might be affected by the potential wildfires. The last thing to do was to spatially join the buffered areas with the zip code and summarize the dataset based on each individual fire. This dataset gives median population values for both total and per square mile.

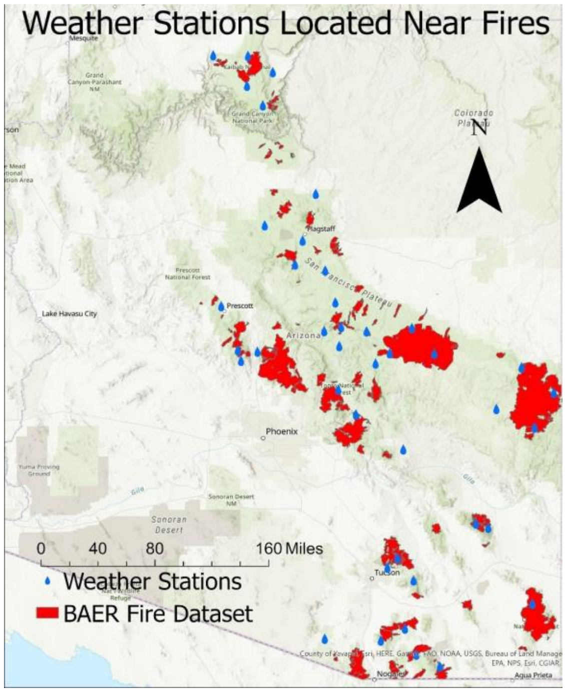
To get medium home values data, Zillow home value index (ZHVI) was used. The index is a smoothed, seasonally adjusted, and is meant to represent the "typical" value of homes in the 35th to 65th percentile range. Excel was used to match zip codes from ZHVI to Arizona zip codes to create a dataset that had all median home prices by zip code. This dataset was then imported into ArcGIS pro where it was joined to the Arizona zip codes with the geoprocessing tool join fields. Once the two tables were combined, a geoprocessing tool called spatial join could be used to join the buffered fire dataset with the Arizona zip codes to create a new dataset that had the zip codes within the buffer zone. To get one value per fire for regression analysis, the mean value was taken for each fire of the median home values in buffered area.

Weather data

Once the suppression costs acres burned terrain data, and human data were found the next step was to identify weather information. The locations of 44 weather stations from the WRCC were used create points in ArcGIS that were associated with each weather station's coordinate data. A spatial join was used with the object being set to closest, which found the weather station closest to each individual fire. Although there were 44 weather stations implemented into the ArcGIS modeling tool, only the closest 40 were selected. Once that information was found, the next step was to identify the dates that each individual fire occurred. The BAER fire report had both the start date and the date contained for all the fires in the dataset which was used to identify necessary dates for weather data. To make sure that the data was accurate, the full year of the fire was imported for each weather station. In excel there was a table for each weather station and each weather station had data for years with fires in its region. The dates were narrowed down to get weather information for the days each fire occurred. Once all the data was implemented into Excel a complex formula of an average ifs statement and indirect formula was used to identify the important variables for weather in this study. The complex formula gave one value for each of the five weather variables that was used in this study so that the regression can be used later once all the variables were created.

Figure 15 shows the weather stations and their proximity to each individual fire. For some fires, the weather stations were in the burned areas so the data would have a good representation of the burned area. Other weather stations were significant distances a way, with some weather stations up to 20 miles away indicated by the figure which could be misrepresenting the weather conditions. This is the best and only data for historic weather conditions this study could find. One weather station was used for each fire, but it could be possible to have multiple for the larger fires in future studies. Several weather stations were established after a given fire or were shut down before a fire began. Given the relatively inexpensive cost for setup and low maintenance cost for the FS, the possible future investment for weather stations throughout the Arizona forests could be beneficial to the overall health of the National Forests.

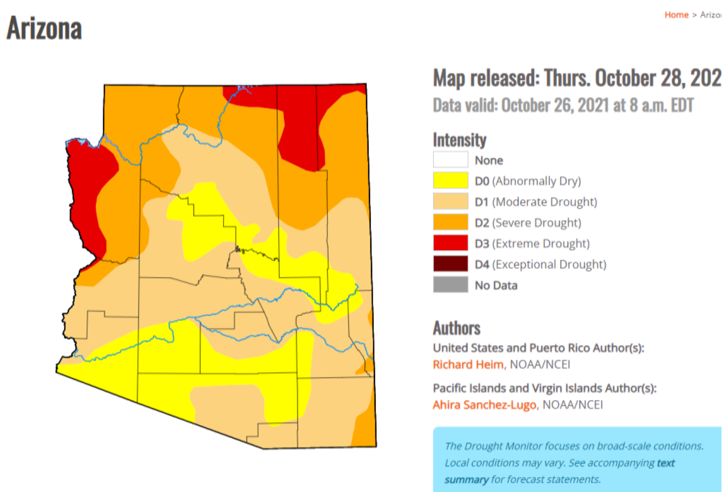
Figure 15. Weather Stations near BAER



Drought Monitor data

Although not weather, another variable was also added to this analysis from the United States Drought Monitoring Center. Figure 16 is the drought severity map of Arizona retrieved from the Drought monitoring website.

Figure 16: Drought Severity map



The map shows intensity of drought for Thursday October 28, 2021, as an example. This study used the dates that the fire was active and the county in which the fires were primarily located. The equation DSCI score is calculated with the equation shown below:

$$1(\mathrm{D0}) + 2(\mathrm{D1}) + 3(\mathrm{D2}) + 4(\mathrm{D3}) + 5(\mathrm{D4}) = \mathrm{DSCI}$$

Then the DSCI was added together for each of the days each individual fire burned giving one value for the regressions. The study's reasoning for adding them together instead of finding the average is that the Drought monitoring center does not recommend the finding the average as it results in less information loss. The ADCSI is the result from this summation and is observed for each individual fire. The study later found GIS information that should be used in future research if the ADCSI needed to be used again. The current process is identifying fires by the county boundary instead of the fires exact locations which could be found in the GIS version of the DSCI. The ADCSI values do have significant variation in for each individual fire with several being zero while others are over 4,000.

Vegetation cover

The substantial piece of information to round out the dataset for a regression analysis was vegetation. The reason to include vegetation, is especially in Arizona, there are areas with lots of juniper trees versus areas with cactus and buffelgrass. The INREV Arizona dataset was clipped to each individual fire and then spatially joined to the fires just for the three vegetation type covers.

References

Abt, K. L., Prestemon, J. P., & Gebert, K. M. (2009). Wildfire Suppression Cost Forecasts for the US Forest Service. *Journal of Forestry*, 6.

Abt, K. L., Prestemon, J. P., & Gebert, K. M. (n.d.). *An Answer to a Burning Question: What Will the Forest Service Spend on Fire Suppression This Summer?* 3.

Arizona Department of Health Services *ADHS-wildfire-emergency-response-plan.pdf*. (n.d.). Retrieved July 10, 2021, from https://www.azdhs.gov/documents/preparedness/epidemiology-disease-control/extreme-weather/wildfires/adhs-wildfire-emergency-response-plan.pdf

BAER Imagery Support Data Download | Burned Area Emergency Response (BAER) Imagery Support. (n.d.). Retrieved March 21, 2021, from https://fsapps.nwcg.gov/baer/baer-imagery-support-data-download

Baylis, P., & Boomhower, J. (2019). *Moral Hazard, Wildfires, and the Economic Incidence of Natural Disasters* (Working Paper No. 26550; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w26550

Chuo, Y. H. (1992). Management of wildfires with a geographical information system. *International Journal of Geographical Information Systems*, *6*(2), 123–140.

Wildland Fire Leadership Council (F) (2009). The Federal Land Assistance, Management and Enhancement Act of 2009: Report to Congress. From https://www.forestsandrangelands.gov/documents/strategy/reports/2_ReportToCongress03172011.pdf

Cova, T. J., Dennison, P. E., Kim, T. H., & Moritz, M. A. (2005). Setting Wildfire Evacuation Trigger Points Using Fire Spread Modeling and GIS. *Transactions in GIS*, *9*(4), 603–617.

Donovan, G. H., Noordijk, P., & Radeloff, V. (n.d.). Estimating the impact of proximity of houses on wildfire suppression costs in Oregon and Washington. *In: 2004. Proceedings of 2nd Symposium on Fire Economics, Planning and Policy: A Global View*.

Donovan, G. H., Prestemon, J. P., & Gebert, K. (2011). The Effect of Newspaper Coverage and Political Pressure on Wildfire Suppression Costs. *Society & Natural Resources*, *24*(8), 785–798

Ellison, A., Moseley, C., & Bixler, R. P. (n.d.). *Drivers of Wildfire Suppression Costs*. 40.

ESRI. *USA ZIP Code Areas—Overview*. (n.d.). Retrieved November 23, 2021, from https://www.arcgis.com/home/item.html?id=8d2012a2016e484dafaac0451f9aea24

Finkelstein, D. B., Pratt, L., Curtin, T., & Brassell, S.. (2005). Wildfires and seasonal aridity recorded in Late Cretaceous strata from south-eastern Arizona, USA. *Sedimentology*, *52*(3), 587–599.

Fire and Aviation Management Website *Fires_acres16.pdf*. (n.d.). Retrieved July 10, 2021, from https://www.predictiveservices.nifc.gov/intelligence/2016_Statssumm/fires_acres16.pdf

Fire and Aviation Management Website *Fires_acres20.pdf*. (n.d.). Retrieved July 10, 2021, from https://www.predictiveservices.nifc.gov/intelligence/2020_statssumm/fires_acres20.pdf

Fire and Aviation Management Website *FIRESTATUserGuide.pdf*. (n.d.-a). Retrieved July 2, 2021, from https://fam.nwcg.gov/fam-web/firestat/FIRESTATUserGuide.pdf

Fire and Aviation Management Website *FIRESTATUserGuide.pdf*. (n.d.-b). Retrieved July 10, 2021, from https://fam.nwcg.gov/fam-web/firestat/FIRESTATUserGuide.pdf

Fitch, R. A., Kim, Y. S., Waltz, A. E. M., & Crouse, J. E. (2018a). Changes in potential wildland fire suppression costs due to restoration treatments in Northern Arizona Ponderosa pine forests. *Forest Policy and Economics*, *87*, 101–114.

Gebert, K. M., Calkin, D. E., & Yoder, J. (2007). Estimating Suppression Expenditures for Individual Large Wildland Fires. *Western Journal of Applied Forestry*, *22*(3), 188–196.

Gebert K.M., Schuster, E./G (1999). Predicting national fire suppression expenditures. In ‘Proceedings of Symposium on Fire Economics, Planning, and Policy: Bottom Lines’. (Eds A Gonza´lez-Caba´n, PN Omi) USDA Forest Service, Pacific Southwest Research Station, General Technical Report PSW-GTR-173, pp. 21–30. (Albany, CA)

Gonzalez-Caban, A., & Omi, P. N. (n.d.). *Proceedings of the Symposium on Fire Economics, Planning, and Policy: Bottom Lines*. 314.

Gude, P. H., Jones, K., Rasker, R., & Greenwood, M. C. (2013). Evidence for the effect of homes on wildfire suppression costs. *International Journal of Wildland Fire*, *22*(4), 537.

Hand, M.., Thompson, M., & Calkin, D. (2016). Examining heterogeneity and wildfire management expenditures using spatially and temporally descriptive data. Journal of Forest Economics, 22, 80-102.

Hesseln, H. (2001). Refinancing and Restructuring Federal Fire Management. *Journal of Forestry*, *99*(11), 4–8.

Hoover, K. (2015, September). Federal assistance for wildfire response and recovery. R41858,

Hoover, K. (2017). Wildfire Management Funding: Background, Issues, and FY2018 Appropriations. 18. Congressional Research Service, Report R45005, Washington, DC.

Hoover, K., & Hanson, L. A. (2021). Wildfire Statistics. Congressional Research Service, Washington, DC.

Hoover, K., & Lindsay, B. R. (2017). Wildfire Suppression Spending: Background, Issues, and Legislation in the 115th Congress. Congressional Research Service, R44966, Washington, DC.

Hoover, K., Lindsay, B. R., McCarthy, F. X., & Tollestrup, J. (2017). *Wildfire Spending: Background, Issues, and Legislation in the 114th Congress*. Congressional Research Service, R44082, Washington, DC.

Houtman, R. M., Montgomery, C. A., Gagnon, A. R., Calkin, D. E., Dietterich, T. G., McGregor, S., & Crowley, M. (2013). Allowing a wildfire to burn: Estimating the effect on future fire suppression costs. *International Journal of Wildland Fire*, *22*(7), 871.

Jhariya, M. K., & Raj, A. (2014). Effects of wildfires on flora, fauna and physico-chemical properties of soil-An overview. *Journal of Applied and Natural Science, 6*(2), 887–897.

Katuwal, H., Calkin, D. E., & Hand, M. S. (2016). Production and efficiency of large wildland fire suppression effort: A stochastic frontier analysis. *Journal of Environmental Management*, *166*, 227–236.

Lee, Y., Fried, J. S., Albers, H. J., & Haight, R. G. (2013). Deploying initial attack resources for wildfire suppression: Spatial coordination, budget constraints, and capacity constraints. *Canadian Journal of Forest Research*, *43*(1), 56–65.

Li, D., Cova, T., & Dennison, P. (2019). Setting Wildfire Evacuation Triggers by Coupling Fire and Traffic Simulation Models: A Spatiotemporal GIS Approach. *Fire Technology*, *55*(2), 617–642.

Lien, A. M., Dew, T., Ruyle, G. B., Sherman, N. R., Perozzo, N., Miller, M., & López-Hoffman, L. (2021). Trust is Essential to the Implementation of Adaptive Management on Public Lands. *Rangeland Ecology & Management*, *77*, 46–56.

Liu, Z., Wimberly, M. C., Lamsal, A., Sohl, T. L., & Hawbaker, T. J. (2015). Climate change and wildfire risk in an expanding wildland–urban interface: A case study from the Colorado Front Range Corridor. *Landscape Ecology*, *30*(10), 1943–1957.

Mangeon, S., Field, R., Fromm, M., McHugh, C., & Voulgarakis, A. (2016). Satellite versus ground-based estimates of burned area: A comparison between MODIS based burned area and fire agency reports over North America in 2007. *The Anthropocene Review*, *3*(2), 76–92.

Martin, I. M., Martin, W. E., & Raish, C. B. (2011). *A qualitative and quantitative analysis of risk perception and treatment options as related to wildfires in the USDA FS Region 3 National Forests* (RMRS-GTR-260; p. RMRS-GTR-260). U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. https://doi.org/10.2737/RMRS-GTR-260

Melvin, A. M., Murray, J., Boehlert, B., Martinich, J. A., Rennels, L., & Rupp, T. (2017). Estimating wildfire response costs in Alaska's changing climate. *Climatic Change*, *141*(4), 783–795.

Monitoring Trends in Burn Severity *Data Explorer | MTBS*. (n.d.). Retrieved October 10, 2021, from https://www.mtbs.gov/data-explorer

More than 60 wildfires rage across US west – including blaze bigger than Portland. (2021, July 14). *The Guardian*. https://www.theguardian.com/us-news/2021/jul/13/us-wildfires-california-oregon-washington

Nation Park Service & United States Department of Agriculture *Fire Perimeter*. (n.d.). Retrieved July 10, 2021, from https://www.fs.fed.us/r3/gis/gisdata/FirePerimeter.html

Nation Park Service United States Department of Agriculture *Fire Terminology*. (n.d.). Retrieved April 14, 2021, from https://www.fs.fed.us/nwacfire/home/terminology.html

National Interagency Fire Center (NIFC). *National Fire News | National Interagency Fire Center*. (n.d.). Retrieved April 14, 2021, from https://www.nifc.gov/fire-information/nfn#Arizona

Preisler, H. K., & Ager, A. A. (2013). Forest-Fire Models Based in part on the article "Forest fire models" by Haiganoush K. Preisler and David R. Weise, which appeared in the Encyclopedia of Environmetrics. In *Encyclopedia of Environmetrics*. American Cancer Society. https://doi.org/10.1002/9780470057339.vaf010.pub2

Preisler, H. K., & Westerling, A. L. (2007). Statistical Model for Forecasting Monthly Large Wildfire Events in Western United States. *Journal of Applied Meteorology and Climatology*, *46*(7), 1020–1030.

Press, T. A. (n.d.). *World's largest tree wrapped in fire-resistant blanket due to threat of California wildfires*. NBC News. Retrieved October 4, 2021, from https://www.nbcnews.com/news/us-news/world-s-largest-tree-wrapped-fire-resistant-blanket-due-threat-n1279420

Prestemon, J. P., Abt, K., & Gebert, K. (2008). Suppression cost forecasts in advance of wildfire seasons. *Forest Science*, 54(4), 381-396

Prestemon, J. P., & Donovan, G. H. (2008). Forecasting Resource-Allocation Decisions under Climate Uncertainty: Fire Suppression with Assessment of Net Benefits of Research. *American Journal of Agricultural Economics*, *90*(4), 1118–1129. .

U.S. Forest Service. *Region 3—Geospatial Data*. (n.d.). Retrieved March 21, 2021, from https://www.fs.usda.gov/detailfull/r3/landmanagement/gis/?cid=stelprdb5201889&width=full

Richardson, L. A., Champ, P. A., & Loomis, J. B. (2012). The hidden cost of wildfires: Economic valuation of health effects of wildfire smoke exposure in Southern California. *Journal of Forest Economics*, *18*(1), 14–35. https://doi.org/10.1016/j.jfe.2011.05.002

Ritchie, M. W. (1999). *A Compendium of Forest Growth and Yield Simulators for the Pacific Coast States*. U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station.

Shared-funding-interagency-agreement.pdf. (n.d.). Retrieved July 16, 2021, from https://www.nwcg.gov/sites/default/files/docs/shared-funding-interagency-agreement.pdf

Short KC (2014) A spatial database of wildfires in the United States, 1992–2011. Earth System Science Data 6: 1–27.

Smith, D. M. (2017). Sustainability and wildland fire: The origins of Forest Service wildland fire research. *FS-1085. Washington, DC: US Department of Agriculture, Forest Service. 120 p*

Sparhawk, W.N. 1925. The use of liability ratings in planning forest fire protection. J. Agric. Res. 30(8):693–792.

Stephens, S. L., & Ruth, L. W. (2005). Federal Forest-Fire Policy in the United States. *Ecological Applications*, *15*(2), 532–542.

Suppression Costs | National Interagency Fire Center. (n.d.). Retrieved July 10, 2021, from https://www.nifc.gov/fire-information/statistics/suppression-costs

Thomas, D., Butry, D., Gilbert, S., Webb, D., & Fung, J. (2017). *The Costs and Losses of Wildfires: A Literature Review*. https://doi.org/10.6028/NIST.SP.1215

Thompson, M. P., Haas, J. R., Finney, M. A., Calkin, D. E., Hand, M. S., Browne, M. J., Halek, M., Short, K. C., & Grenfell, I. C. (2015a). Development and application of a probabilistic method for wildfire suppression cost modeling. *Forest Policy and Economics*, *50*, 249–258.

U.S. Census Bureau QuickFacts: Arizona. (n.d.). Retrieved November 23, 2021, from https://www. census.gov/quickfacts/AZ

United States Department of Agriculture Forest Service *RMRS - BAER Burned Area Reports DB*. (n.d.). Retrieved March 21, 2021, from https://forest.moscowfsl.wsu.edu/cgi-bin/ BAERTOOLS/baer-b/index.pl?totals=true&loc_type=state&state=AZ&exp=10&pp=1000&

United States Department of the Interior *Budget*. (2015, July 17). https://www.doi.gov/wildl andfire/budget

United States Drought Monitor Website *Data Tables | United States Drought Monitor*. (n.d.). Retrieved April 15, 2021, from https://droughtmonitor.unl.edu/Data/DataTables.aspx

U.S. Forest Service. *Los Padres National Forest—Home*. (n.d.). Retrieved July 10, 2021, from https://www.fs.usda.gov/detail/lpnf/home/?cid=fseprd570093

Westerling, A. L., & Bryant, B. P. (2008). Climate change and wildfire in California. *Climatic Change*, *87*(1), 231–249.

Yoder, J., & Gebert, K. (2012). An econometric model for ex ante prediction of wildfire suppression costs. *Journal of Forest Economics*, *18*(1), 76–89.