Essays on Simulation Modeling and Empirical Examination of Wildfire as an Outcome of a Coupled Human Natural System

Farshad Farkhondehmaal

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
In
Industrial and Systems Engineering

Navid Ghaffarzadegan

Shyam Ranganathan

Anne-Lise K. Velez

Manish Bansal

December 16, 2021

Blacksburg, VA

Keywords: Wildfire, coupled human natural system, fire ignition, burned area, system dynamics, simulation, hierarchical modeling, socioeconomic system, environment, climate change

Essays on Simulation Modeling and Empirical Examination of Wildfire as an Outcome of a Coupled Human Natural System

Farshad Farkhondehmaal

Abstract

Wildfire activity has increased in recent years in the United States, endangering both environment and society. Appropriate management of this phenomenon is only achievable with a thorough understanding of the critical factors influencing wildfire activity in each region. In three essays, I use statistical and mathematical models to examine wildfires and propose solutions to mitigate their impact on society. In the first essay, I focused on building a systematic framework for modeling wildfire as a coupled human-natural system. I employ system dynamics modeling, which was previously applied in various fields, including healthcare, sustainability, and disaster mitigation. I show how, in the absence of exogenous factors such as temperature or lightning, the human perception of fire danger may establish a feedback loop that can yield significant trends such as fluctuation or even fluctuation with rising amplitude when linked with the natural system. This conclusion is counter-intuitive, given that the human contribution to wildfire is typically described in the literature using constant or semi-constant variables. Additionally, I analyzed the impact of three important fire protection measures on reducing burning rates (prescribed burning, enhancing immediate suppression accomplishment, and regulating the rate of WUI growth). The research concludes that appropriately integrating several policies can result in a synergistic effect that is greater than the sum of the effects of the individual policies. The second essay calibrates the model built in the first essay and examines wildfire trends across the contiguous United States. The simulation results closely match the real data, and the model serves as a foundation for data-driven policy research. To be more precise, I fit the model to each state separately and then compare the model's goodness of fit. Following that, I examine the influence of various policies and scenarios on wildfire behavior. In the scenario, I examine the effect of maintaining constant temperatures and precipitation levels relative to the average values for these variables over the last century. For the policy analysis, I examine the influence of three policies on each state (prescribed burning, increasing immediate suppression achievement, and regulating the rate of WUI development). Here, I provide state-specific suggestions about the primary factors that contribute to wildfires and the most effective policies for each state. In the third essay, I have implemented the Oregon wildfire history dataset and integrated it with two other aerial datasets, including meteorological data gathered by weather stations located around the state and counties. Then, using hierarchical modeling on over 10,000 wildfire ignitions, I developed a classification system for determining if a given fire has the potential to grow major or not. However, utilizing a huge dataset and a variety of resources presents several obstacles, such as the presence of missing data. I imputed the missing numbers using a sophisticated mathematical approach called "Predictive Mean Matching".

Essays on Simulation Modeling and Empirical Examination of Wildfire as an Outcome of a Coupled Human Natural System

Farshad Farkhondehmaal

General Audience Abstract

Wildfire activity has increased in recent decades in the United States, which put many people in danger. Climate change, the Settlement of people in the Wildland Urban Interface, and an increase in vegetation density each play a role in this increase. In this dissertation, we discuss the wildfire in the United States in three essays. In the first essay, we develop a mathematical model to show how humans and nature affect wildfire activity in any area. We then test different major wildfire management policies on the hypothetical situation to compare the outcome of these policies together. In the second essay, we use the model developed in the essay (with some minor changes) to model the wildfire activity in 11 states of the U.S. which has the most wildfire activity in recent years. First, we show that our model can replicate the wildfire activity in each state. Second, we test the effect of wildfire mitigation policies on each state. This essay proposes state-specific policy recommendations based on the main reasons for the increase in wildfire activity in each state. Finally, in the third essay, we develop a statistical model to predict the existence of large wildfires in the next month in Oregon counties. We use climate, land, and fire history data to develop a warning system. Policymakers can use this system to move the fire suppression resources to counties with a high probability of experiencing large wildfires over the next month. Finally, all essays aim to enhance our understanding of the reasons for the increase in wildfire activity in recent years and suggest finding the appropriate way to deal with this change to reduce the effect of wildfire on human life.

Acknowledgment

I have gotten a great lot of help and guidance throughout the process of creating and writing my dissertation. I'd want to use this opportunity to express my gratitude to the fantastic people that made this difficult journey possible.

To begin, I would like to give special thanks to my advisor, Dr. Navid Ghaffarzadegan, for all of his assistance, support, and patience. He is an exceptional counsel, a clever professor, an accomplished scholar, and an excellent role model. I had a fantastic time working with him, and our collaboration helped me greatly. He has been an excellent mentor and adviser throughout my graduate studies, and without him, I would not have completed my PhD degree.

I'd also want to thank my committee members, Dr. Shyam Ranganathan, Dr. Anne-Lise Velez, and Dr. Manish Bansal. It's been a pleasure collaborating with you all. Thank you for your encouragement and useful criticism during the many stages of my study, which opened the path for me to deal with the more difficult aspects of my research. Also, I would like to thank all the individuals in the ISE department at Virginia Tech, notably Dr. Eileen M. Van Aken, Hannah Parks, Rhonda Hawley, Nicole Conner, and Akiko Nakata. Everyone here has always helped me feel welcomed and supported, and I appreciate all of their efforts.

I am also very grateful to my lovely wife, Nastaran Khalili, for her support throughout these years. I'd want to express my gratitude for being my best friend, coworker, and lab mate. Without your unending love, assistance, and support, I would be unable to complete this period of my life.

Additionally, I'd want to thank my parents and sister for their support during my education. I appreciate your constant love, encouragement, and unending support. Although I wish they could have been here with me during these years, their constant communication and support from thousands of miles away allowed me to focus more on my study and navigate the difficult immigration periods.

I also want to thank my wonderful lab mates, Hesam Mahmoudi, Maede Maftouni, Alireza Ebrahimvandi, Negar Darabi, Ran Xu, and Marcus Miller. Your presence, support, and valuable feedbacks on my research and presentations assisted me in thinking more deeply.

I could not have completed this dissertation without the support of my friends. I would like to thank my friends in Blacksburg who helped me have lots of enjoyable moments: Masoumeh Riazi, Dr. Masoud Agah, Dr. Leyla Nazhandali, Maede Maftouni, Ali Safiabadi, Ali Hosseini Fahraji, Kiana Moghimi, Ali Talebzadeh, Morteza Banagar, Behnaz Rezvani, Noushin Omidvar, Dr. Hussein El Hajj, Mona Safari, Dr. Ahmad Radmehr, Mehdi Bahrampoori, Darya Montazeri, Esmat Mohammadi and Mohsen Hasani.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

The recent increase in wildfire activity in the United States, as well as the monetary and human cost associated with it, has sparked public interest in determining the fundamental factors that influence this phenomenon (Westerling et al., 2003; Malamud et al., 2005; Burke et al., 2021). In 2020, more than 52000 wildfires occurred in the United States that have burned 8,889,297 acres nationwide, almost double the burning rate in 2019 and 2.3 million acres more than the 10-years average. Furthermore, the size of wildfires is also increasing as in 2010 California experienced five out of six of its largest wildfires since 1932 in 2020 (Philantropy, 2020). The problem is not limited to the United States: wildfire is a global challenge affecting different regions all around the world, with recent catastrophic events in countries such as Australia, Brazil, Greece, Algeria, France, Turkey, and Indonesia[1]. Given the trends, the problem of wildfires and their increasing catastrophic consequences are of immense policy relevance. This dissertation is a response to this growing global and regional concern.

Our **objective** in this dissertation is to provide a systematic understanding of wildfire as a complex phenomenon, and contribute to the literatures of wildfire management, as well as modeling complex systems. In past studies, three primary elements are believed to be the most significant drivers in the growth of wildfire activities (Westerling, 2016; Nagy et al., 2018; Steel et al., 2015). First, warmer climate, particularly in the western United States, is one significant mechanism being studied as a potential cause of this increase. For example, one study shows that the warmer spring and summer months are the reason that wildfire activity has increased in the western United States (Westerling et al., 2006). On the other hand, several research studies focus on the human role in wildfires. For example, recent decades of development of the Wildland Urban Interface (WUI), the place that was part of the forest and now occupied by housing units, led to an increase in the number of human ignitions and introduced fire to places where naturally fire rarely happened (Radeloff et al., 2018). Considering that humans are responsible for 84% of wildfire ignition nationwide, this is also an important factor to examine (Balch et al., 2017). The third group of factors focuses on the change in vegetation density in US forests, which may be the legacy of fire suppression policies implemented over the last century (Johnson et al., 2001). Hence, there is no unanimous consent among researchers on what is the key driving factor of the current wildfire increase.

Our research **approach** is quantitative, and based on system dynamics modeling, simulation, and statistical analysis. Dynamic modelers have contributed to understanding the causes and consequences of wildfire, policy analysis to mitigate such events or associated costs, as well as developing predictive tools. We can classify wildfire modeling into two types based on their desired purpose. Table1.1 shows these categories, which are short-term and long-term wildfire modeling. First, there is strong literature about predictive modeling (Dennison et al., 2006; Mandel et al., 2011). Several studies look into the topological, fuel, and meteorological factors to predict the behavior of an ongoing wildfire (Lamorlette et al., 2018). This information can help authorities take necessary actions, including evacuation orders, and redefine the firefighter's approach. Furthermore, a substantial body of literature examines wildfire management from a larger perspective, including multiple stakeholders and decision-making under time pressure (Steelman and Kunkel, 2004). For example, one study examined 21 large wildfires in the Midwest United States and found out that almost half of the post incidents networks have

[1] https://www.aljazeera.com/news

different important actors compared to pre-incident ones, generating potential for delay or even disconnection in the flow of the information during a wildfire (Faas et al., 2017).

On the other hand, another type of modeling aims to predict important variables contributing to wildfire (Balch et al., 2017; Fischer et al., 2016). The main role of these models is to understand factors shaping wildfire behavior for any region. These broad categories of research fall into two branches. First, there is a research about likelihood of fire ignition in different regions. Similarly, some research focuses on estimating the number of potential fires in a specific time period. The Santa ana index is developed to predict the ignition of any significant wildfire in the next six days (Rolinski et al., 2016; Kochanski et al., 2013). Additionally, other researches focus on the magnitude of the fire, which is important in case of estimating the wildfire cost to the communities (Westerling et al., 2006). Several studies suggest that environmental and socioeconomic factors play an important role in the burning rate, especially in the wildland-urban interface (Mercer and Prestemon, 2005; de Torres Curth et al., 2012).

Table 1.1- The classification of wildfire modeling based on the purpose of the study

Wildfire modeling			
Long term		Short term	
Spatial prediction	Information and command system	Burning rate	Wildfire ignition
• (Dennison et al., 2006) • (Mandel et al., 2011) • (Kochanski et al., 2013; Nagy et al., 2018; Steel et al., 2015) • (Lamorlette et al., 2015; Steel et al., 2015) • (El Houssami et al., 2016) • (Lamorlette et al., 2018) • (El Houssami et al., 2018) • (Finney, 1998)	• (Cutter et al., 2000) • (Steelman and Kunkel, 2004) • (Kapucu, 2005; Nagy et al., 2018; Steel et al., 2015) • (Kapucu, 2008a; Nagy et al., 2018; Steel et al., 2015) • (Kapucu, 2008b; Steel et al., 2015) • (Kapucu et al., 2010; Steel et al., 2015) • (Steelman and McCaffrey, 2013) • (Ai et al., 2016; Kochanski et al., 2013) • (Nowell et al., 2018; de Torres Curth et al., 2012) • (Comfort et al., 2020)	• (Riaño et al., 2002) • (Yebra et al., 2013; Nagy et al., 2018; Steel et al., 2015) • (Haas et al., 2013) • (Dennison et al., 2014; Steel et al., 2015) • (Calkin et al., 2015) • (Pimont et al., 2016) • (McWethy et al., 2017; Kochanski et al., 2013)	• (Schoennagel et al., 2017) • (Santamaria et al., 2015) • (Nagy et al., 2018) • (Keeley and Syphard, 2018) • (Romero-Calcerrada et al., 2008) • (Syphard and Keeley, 2015)

While the role of humans in a wildfire is evident in different regions, the literature lacks an in-depth examination of wildfire as outcome of human and nature interactions. Our **main contribution in this dissertation** stems from merging the natural and human sides of the problem of wildfire. Looking at wildfire as an outcome of a coupled human-natural system is the main novelty of this dissertation. Although it is not common in studies of wildfire, we are not the first to take this systematic approach. Recently, more attention is being paid to establishing wildfire as a coupled human-natural system (CHNS) (Fischer et al., 2016). This new approach implies that wildfire, which has been viewed as a natural event for centuries, should be viewed as a result of natural and human interaction (Kline et al., 2017). Table1.1 illustrates the interaction of these two key systems, human and natural, and how they affect each other. This figure serves as a foundation for demonstrating how human and environmental systems both influence and are influenced by wildfires. For example, people alter the vegetation's dynamic through dwelling expansion and fire suppression, by removing trees or preventing fires from burning naturally (Carleton and MacLELLAN, 1994; Theobald and Romme, 2007). On the other side, vegetation dynamic influences humans' perceptions of wildfire risk since humans perceive the dry and dense vegetation as a sign of possible high wildfire risk. We believe that modeling wildfires must incorporate these principles in order to have a full understanding of this phenomenon.

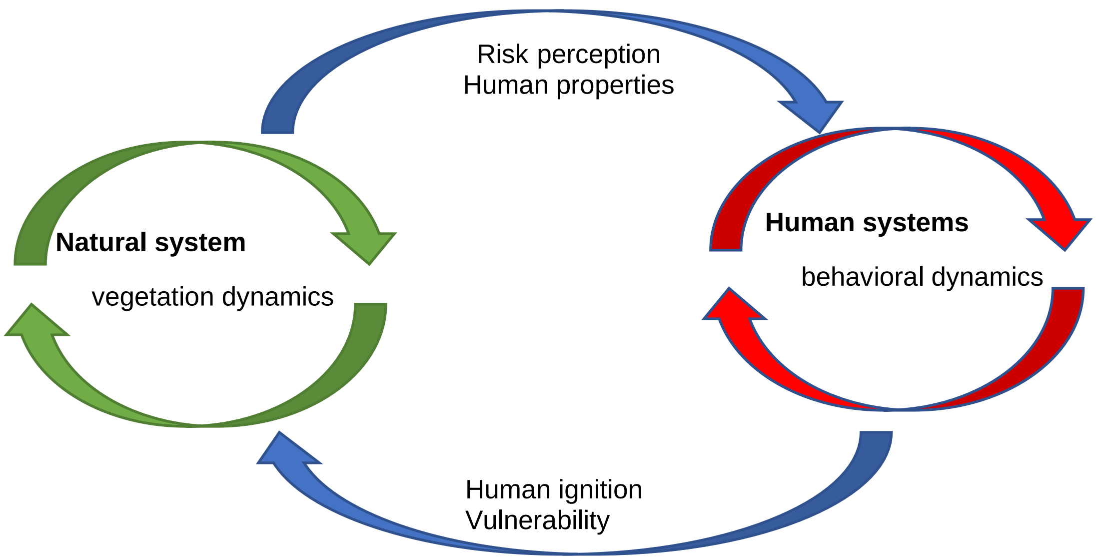


Figure 1.1- Wildfire as an outcome of a coupled human-natural system (adapted from https://www.nsf.gov/pubs/2018/nsf18503/nsf18503.htm and adjusted for the case of wildfires).

This dissertation is structured around multiple essays. We examine and develop this theoretical foundation in a three-essay dissertation that includes various dynamic modeling and statistical techniques. The overview of my essays are as follows: the first paper aims to develop a systematic theoretical model that shows the human and natural systems' coupling effect in shaping wildfire behavior. I used the system dynamics approach to develop a model that endogenously show how humans' perception of wildfire risk and change in their activities, accordingly, can be source of fluctuation in the wildfire ignition and total burning rate. The result of the paper is a general analytic tool to model wildfire activity. I test four wildfire mitigation policies and compare their effects on several outcome measures such as burning rate. In the second essay of my dissertation, I used the model from the first essay (with minor changes to include exogenous factors) to replicate the wildfire activity in the states of the conterminous United States. I checked the role of humans, vegetation, and climate on wildfire activity on each state and offered state-specific policy recommendations. The model is calibrated to data from 11 states of the United States. In the third essay, I used hierarchical regression modeling to develop a warning system to predict the probability of an ignition becoming a large wildfire (burning larger than 500 acres). County-level data from Oregon were used for this analysis.

These three essays are related to each other and have provided insights into designing and investigating the next study or validating a previous one. Altogether, and as a whole, **this dissertation provides a novel theoretical and empirical approach to analyzing wildfire**

where human factors are coupled with natural factors forming feedback loops. This is a major step towards not only looking at wildfire as a complex coupled system problem but also changing our perspectives about other major environmental challenges.

Details of contributions of the three essays

Essay 1

In the first study, I focused on developing a systematic framework for modeling wildfire as a coupled human-natural system. I use system dynamics modeling approach (Richardson 1991, Sterman 2000), which previously used for systematic modeling in different domains including healthcare, sustainability and disaster mitigation (Deegan 2007, Stave 2010). I demonstrate how, in the absence of any exogenous variables like temperature or lightning, the human perception of the risk of fire can create a feedback loop that, when coupled with the natural system, can generate important trends such as fluctuation or even fluctuation with increasing amplitude. This result is counter-intuitive since the human role is usually modeled with constant or semi-constant variables in the literature (de Torres Curth et al., 2012; Romero-Calcerrada et al., 2010). Additionally, I evaluated the effectiveness of three critical fire prevention policies (prescribed burning, improving the immediate suppression achievement, and controlling the WUI development rate) on burning rate reduction. The conclusion of this research is that properly combining numerous policies may leads in a synergistic effect, which is superior to the sum of the effect of separate policies.

This study made several contributions to wildfire modeling. To begin, it sheds new light on the relationship between human and ecological mechanisms that influence wildfire behavior. Understanding the factors contributing to wildfire behavior may aid in a more effective response to a wildfire outbreak. Many studies have focused on a single side of the wildfire. However, we consider both human and natural systems. In addition, this study answers the question of how the same people can have different reactions to the wildfire and ignite different numbers of fires in the short term, even though none of the factors are mainly considered in the literature to model human role change. Broader impacts of this study are related to its contribution to policy debates around understanding complex socio-environmental systems and saving natural resources (Davis et al. 2020) and providing another example of dynamic and predictive modeling of complex human systems for policy analysis (Ebrahimvandi 2020; Andalib 2018; Azadeh Fard 2016).

Essay 2

An outcome of the first essay is a system dynamics model to examine the effectiveness of different policies on wildfire behavior. In this essay, I calibrate the model to examine wildfire patterns for different states of the conterminous United States. The simulation results fairly replicate the data, and the model works as a platform for data-informed policy analysis. Specifically, we first fit the model to each state and then compare the model's goodness of fit. We then check the effect of policies and scenarios on wildfire behavior. In the case of the scenario, we test the effect of not having a change in temperature and precipitation compared to the average value of these variables in the past century. For the policy analysis, we test the effect of four policies (prescribed burning, clear-cutting, improving the immediate suppression achievement, and controlling the WUI development rate) for each state. Here we can provide state-specific recommendations on the main driving forces of wildfire and the most effective policy for each state.

Methodologically, this study provides another example of taking a model that represents a family of a problem and calibrating it with a range of contexts. Techniques of model calibration have been a major topic in modeling communities, including the system dynamics community (Rahmandad and Spiteri 2015; Hosseinichimeh et al. 2016), and this study provides an example of using simple calibration techniques for model examination.

Essay 3

In the third paper, we ask: Can the same population with the same wildfire history in the area react differently to future wildfire? If yes, what factors change people's contribution to wildfire. The objective of the study is to increase the understanding of the human role in wildfire by implementing real-world data.

The need to improve the modeling approaches in CHNS is accentuated in wildfire pathology (Fischer et al., 2016). Past studies lack to consider the human role for several reasons. First, research shows that socioeconomic variables can be used as a proxy for modeling the community's total perception of fire risk, which is the determining factor in human ignition (Fischer et al., 2016). However, these factors seem to change slowly, whereas human ignition changes monthly. Furthermore, many studies solve the problem by modeling annual wildfires. While these results can be helpful for general comparison between regions in case of fire risk, they do not assist authorities in allocating resources to high-profile areas.

Wildfire modeling literature will benefit from this study's unique approaches in several ways. Using the latest data science approaches, I examined wildfire history in Oregon from 1992 to 2015 to better understand how environmental and socioeconomic factors influence the likelihood of big wildfires if ignition occurs. Furthermore, a model that provides a monthly probability map for each county in Oregon may be utilized as a warning system. The choice of counties as the base for spatial scale is twofold. First, it is the smallest jurisdiction that the government collects socioeconomic information. Second, we want to prevent the complexity of multi-jurisdictional disaster management and assume each county has a consistent wildfire management strategy (Nowell, Steelman et al. 2018, Steelman, Nowell et al. 2021). Using this technique, authorities may allocate firefighting resources to the most at-risk areas each month. Previous studies which develop annual wildfire risk hinder fast preparation for wildfire. In addition, the study develops a white-box hierarchical regression model with comparative accuracy to state-of-the art black-box ML approaches (Deep Learning, Extreme Graduate Boosting, Random Forest, etc.) that lack transparency and accountability. The developed model can show the importance of all variables and interactive terms. Finally, I consider the interaction between socioeconomic and environmental variables to fill the gap in the literature (modeling the human role with slow-changing variables). It is important to consider the interactions because it enhances the model's capability in capturing complex relationships. Combining ML with big data can increase our understanding of wildfire factors, as it enables us to check interactions between the environment and socioeconomic factors.

In this study, I obtained the Oregon wildfire history dataset and combined it with two other aerial datasets, the meteorological data sets collected by weather stations across the state and counties. I then applied hierarchical modeling on more than 10,000 wildfire ignitions to develop a classification algorithm for detecting the possibility of a given fire become large or not. However, using a large dataset and different resources has some challenges, like existence of the missing data. I applied one of the advanced mathematical techniques called "Predictive Mean Matching" to impute the missing values.

References

Ai F, Comfort LK, Dong Y, et al. (2016) A dynamic decision support system based on geographical information and mobile social networks: A model for tsunami risk mitigation in Padang, Indonesia. *Safety science* 90: 62-74.

Andalib, Maryam Alsadat. 2018. 'Model-based Analysis of Diversity in Higher Education', Virginia Tech.

Azadeh Fard, Nasibeh. 2016. 'Essays on Risk Indicators and Assessment: Theoretical, Empirical, and Engineering Approaches', Virginia Tech.

Balch JK, Bradley BA, Abatzoglou JT, et al. (2017) Human-started wildfires expand the fire niche across the United States. *Proceedings of the national academy of sciences* 114(11): 2946-2951.

Burke M, Driscoll A, Heft-Neal S, et al. (2021) The changing risk and burden of wildfire in the United States. *Proceedings of the national academy of sciences* 118(2).

Calkin DE, Thompson MP and Finney MA (2015) Negative consequences of positive feedbacks in US wildfire management. *Forest Ecosystems* 2(1): 1-10.

Carleton TJ and MacLELLAN P (1994) Woody vegetation responses to fire versus clear-cutting logging: a comparative survey in the central Canadian boreal forest. *Ecoscience* 1(2): 141-152.

Comfort LK, Kapucu N, Ko K, et al. (2020) Crisis decision-making on a global scale: Transition from cognition to collective action under threat of COVID-19. *Public Administration Review* 80(4): 616-622.

Cutter SL, Mitchell JT and Scott MS (2000) Revealing the vulnerability of people and places: A case study of Georgetown County, South Carolina. *Annals of the association of American Geographers* 90(4): 713-737.

Darabi, N., Ebrahimvandi, A., Hosseinichimeh, N., & Triantis, K. (2021). A DEA Evaluation of US States' Healthcare Systems in terms of their Birth Outcomes. Expert Systems with Applications, 115278.

Davis, Kirsten, Navid Ghaffarzadegan, Jacob Grohs, Dustin Grote, Niyousha Hosseinichimeh, David Knight, Hesam Mahmoudi, and Konstantinos Triantis. 2020. 'The Lake Urmia vignette: a tool to assess understanding of complexity in socio-environmental systems', System Dynamics Review, 36: 191-222.

Deegan, M. A. (2007). Exploring US flood mitigation policies: a feedback view of system behavior, Rockefeller College of Public Affairs and Policy, University at Albany–SUNY.

de Torres Curth M, Biscayart C, Ghermandi L, et al. (2012) Wildland–urban interface fires and socioeconomic conditions: a case study of a Northwestern Patagonia city. *Environmental Management* 49(4): 876-891.

Dennison PE, Brewer SC, Arnold JD, et al. (2014) Large wildfire trends in the western United States, 1984–2011. *Geophysical Research Letters* 41(8): 2928-2933.

Dennison PE, Charoensiri K, Roberts DA, et al. (2006) Wildfire temperature and land cover modeling using hyperspectral data. *Remote Sensing of Environment* 100(2): 212-222.

Ebrahimvandi, Alireza. 2020. 'Three Essays on Analysis of US Infant Mortality Using Systems and Data Science Approaches', Virginia Tech.

El Houssami M, Lamorlette A, Morvan D, et al. (2018) Framework for submodel improvement in wildfire modeling. *Combustion and Flame* 190: 12-24.

El Houssami M, Thomas J, Lamorlette A, et al. (2016) Experimental and numerical studies characterizing the burning dynamics of wildland fuels. *Combustion and Flame* 168: 113-126.

Faas A, Velez ALK, FitzGerald C, et al. (2017) Patterns of preference and practice: bridging actors in wildfire response networks in the American Northwest. *Disasters* 41(3): 527-548.

Finney MA (1998) *FARSITE, Fire Area Simulator--model development and evaluation.* US Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Fischer AP, Spies TA, Steelman TA, et al. (2016) Wildfire risk as a socioecological pathology. *Frontiers in Ecology and the Environment* 14(5): 276-284.

Haas JR, Calkin DE and Thompson MP (2013) A national approach for integrating wildfire simulation modeling into Wildland Urban Interface risk assessments within the United States. *Landscape and Urban Planning* 119: 44-53.

Hosseinichimeh, Niyousha, Hazhir Rahmandad, Mohammad S Jalali, and Andrea K Wittenborn. 2016. 'Estimating the parameters of system dynamics models using indirect inference', System Dynamics Review, 32: 156-80.

Kapucu N (2005) Interorganizational coordination in dynamic contexts: Networks in emergency management. *Connections* 26(2): 33-48.

Kapucu N (2008a) Collaborative emergency management: better community organising, better public preparedness and response. *Disasters* 32(2): 239-262.

Kapucu N (2008b) Culture of preparedness: household disaster preparedness. *Disaster prevention and management: An international journal*.

Kapucu N, Arslan T and Demiroz F (2010) Collaborative emergency management and national emergency management network. *Disaster prevention and management: An international journal*.

Keeley JE and Syphard AD (2018) Historical patterns of wildfire ignition sources in California ecosystems. *International Journal of Wildland Fire* 27(12): 781-799.

Kline JD, White EM, Fischer AP, et al. (2017) Integrating social science into empirical models of coupled human and natural systems. *Ecology and Society* 22(3).

Kochanski AK, Jenkins MA, Mandel J, et al. (2013) Real time simulation of 2007 Santa Ana fires. *Forest ecology and management* 294: 136-149.

Lamorlette A, El Houssami M and Morvan D (2018) An improved non-equilibrium model for the ignition of living fuel. *International Journal of Wildland Fire* 27(1): 29-41.

Lamorlette A, Houssami ME, Thomas JC, et al. (2015) A dimensional analysis of forest fuel layer ignition model: Application to the ignition of pine needle litters. *Journal of Fire Sciences* 33(4): 320-335.

Malamud BD, Millington JD and Perry GL (2005) Characterizing wildfire regimes in the United States. *Proceedings of the national academy of sciences* 102(13): 4694-4699.

Mandel J, Beezley JD and Kochanski AK (2011) Coupled atmosphere-wildland fire modeling with WRF 3.3 and SFIRE 2011. *Geoscientific Model Development* 4(3): 591-610.

McWethy D, Haberle S, Hopf F, et al. (2017) Aboriginal impacts on fire and vegetation on a Tasmanian island. *Journal of Biogeography* 44(6): 1319-1330.

Mercer DE and Prestemon JP (2005) Comparing production function models for wildfire risk analysis in the wildland–urban interface. *Forest policy and economics* 7(5): 782-795.

Nagy R, Fusco E, Bradley B, et al. (2018) Human-related ignitions increase the number of large wildfires across US ecoregions. *Fire* 1(1): 4.

Nowell B, Steelman T, Velez A-LK, et al. (2018) The structure of effective governance of disaster response networks: Insights from the field. *The American Review of Public Administration* 48(7): 699-715.

Steelman, T., Nowell, B., Velez, A. L., & Scott, R. (2021). Pathways of Representation in Network Governance: Evidence from Multi-Jurisdictional Disasters. Journal of Public Administration Research and Theory.

Ojerio R, Moseley C, Lynn K, et al. (2011) Limited involvement of socially vulnerable populations in federal programs to mitigate wildfire risk in Arizona. *Natural Hazards Review* 12(1): 28-36.

Philantropy CfD (2020) 2020 North American Wildfire Season Reportno. Report Number|, Date. Place Published|: Institution|.

Pimont F, Parsons R, Rigolot E, et al. (2016) Modeling fuels and fire effects in 3D: model description and applications. *Environmental Modelling & Software* 80: 225-244.
Radeloff VC, Helmers DP, Kramer HA, et al. (2018) Rapid growth of the US wildland-urban interface raises wildfire risk. *Proceedings of the national academy of sciences* 115(13): 3314-3319.
Rahmandad, Hazhir, and Raymond J Spiteri. 2015. '12 Modeling Competing Actors Using Differential Games', Analytical methods for dynamic modelers: 373.
Riaño D, Chuvieco E, Ustin S, et al. (2002) Assessment of vegetation regeneration after fire through multitemporal analysis of AVIRIS images in the Santa Monica Mountains. *Remote Sensing of Environment* 79(1): 60-71.
Richardson, G. P. (1991). Feedback thought in social science and systems theory, University of Pennsylvania.
Rolinski T, Capps SB, Fovell RG, et al. (2016) The Santa Ana wildfire threat index: Methodology and operational implementation. *Weather and Forecasting* 31(6): 1881-1897.
Romero-Calcerrada R, Barrio-Parra F, Millington JD, et al. (2010) Spatial modelling of socioeconomic data to understand patterns of human-caused wildfire ignition risk in the SW of Madrid (central Spain). *Ecological Modelling* 221(1): 34-45.
Romero-Calcerrada R, Novillo C, Millington JD, et al. (2008) GIS analysis of spatial patterns of human-caused wildfire ignition risk in the SW of Madrid (Central Spain). *Landscape ecology* 23(3): 341-354.
Santamaria S, Kempná K, Thomas JC, et al. (2015) Investigation of structural wood ignition by firebrand accumulation. *First international conference on structures safety under fire blast. Glasgow, UK.* 1-13.
Schoennagel T, Balch JK, Brenkert-Smith H, et al. (2017) Adapt to more wildfire in western North American forests as climate changes. *Proceedings of the national academy of sciences* 114(18): 4582-4590.
Stave, K. Participatory system dynamics modeling for sustainable environmental management: Observations from four cases. Sustainability 2, 2762-2784 (2010).
Steel ZL, Safford HD and Viers JH (2015) The fire frequency-severity relationship and the legacy of fire suppression in California forests. *Ecosphere* 6(1): 1-23.
Steelman TA and Kunkel GF (2004) Effective community responses to wildfire threats: lessons from New Mexico. *Society and Natural Resources* 17(8): 679-699.
Steelman TA and McCaffrey S (2013) Best practices in risk and crisis communication: Implications for natural hazards management. *Natural Hazards* 65(1): 683-705.
Syphard AD and Keeley JE (2015) Location, timing and extent of wildfire vary by cause of ignition. *International Journal of Wildland Fire* 24(1): 37-47.
Theobald DM and Romme WH (2007) Expansion of the US wildland–urban interface. *Landscape and Urban Planning* 83(4): 340-354.
Westerling AL (2016) Increasing western US forest wildfire activity: sensitivity to changes in the timing of spring. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371(1696): 20150178.
Westerling AL, Gershunov A, Brown TJ, et al. (2003) Climate and wildfire in the western United States. *Bulletin of the American Meteorological Society* 84(5): 595-604.
Westerling AL, Hidalgo HG, Cayan DR, et al. (2006) Warming and earlier spring increase western US forest wildfire activity. *science* 313(5789): 940-943.
Xu, R., Rahmandad, H., Gupta, M., DiGennaro, C., Ghaffarzadegan, N., Amini, H., & Jalali, M. S. (2021). Weather, air pollution, and SARS-CoV-2 transmission: a global analysis. The Lancet Planetary Health, 5(10), e671-e680.

Yebra M, Dennison PE, Chuvieco E, et al. (2013) A global review of remote sensing of live fuel moisture content for fire danger assessment: Moving towards operational products. *Remote Sensing of Environment* 136: 455-468.

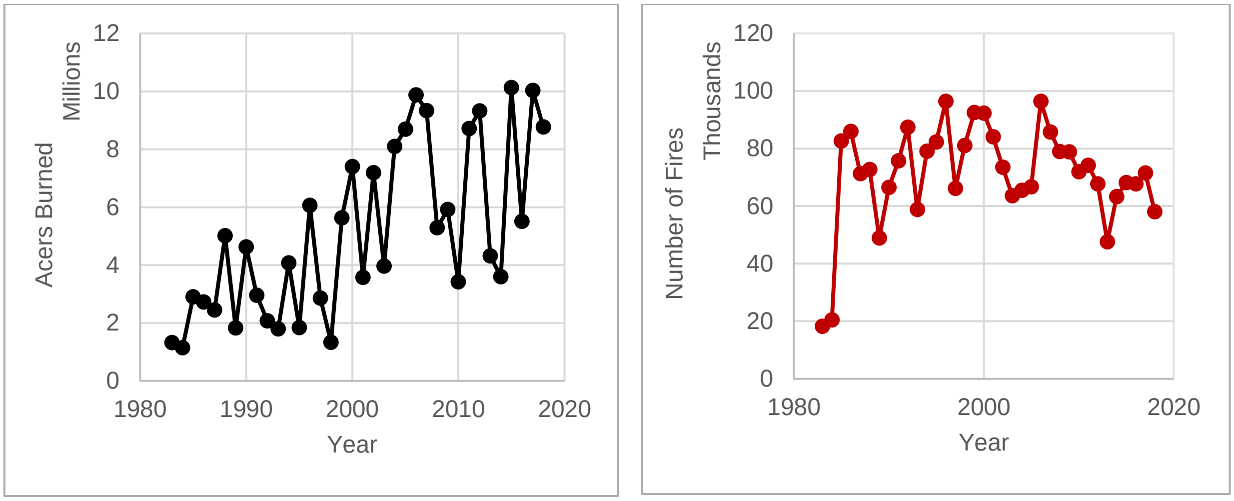
A Cyclical Wildfire Pattern as the Outcome of a Coupled Human-natural System

Abstract: Over the past decades, wildfire has imposed a considerable cost on natural resources and human lives. In many regions, annual wildfire trends show puzzling oscillatory patterns with increasing amplitudes for burned areas over time. This paper aims to examine the potential causes of such patterns by developing and examining a dynamic simulation model that represents interconnected social and natural dynamics in a coupled system. We develop a generic dynamic model and, based on simulation results, postulate that the interconnection between human and natural subsystems is a source of the observed cyclical patterns in wildfires in which risk perception regulates activities that can result in more fire and development of vulnerable properties. Our simulation-based policy analysis points to a non-linear characteristic of the system which rise due to the interconnections between the human side and natural side of the system. This has a major policy implication: in contrast to studies that look for the most effective policy to contain wildfires, we show that a long-term solution is not a single action but is a combination of multiple actions that simultaneously target both human and natural sides of the system.

1. Introduction

Wildfire is endangering human life, natural resources, forest conservation, and wildlife [1-3]. According to the National Interagency Fire Center, in 2020, more than 52,000 wildfire incidents in the U.S. burned about 9 million acres [4]. In California alone, it was estimated that about 30 people died due to wildfires during the first 9 months of 2020 [5]. In addition, the tragic 2018 Camp Fire incident of Paradise, California, arguably the most destructive and deadliest wildfire in California's history, resulted in at least 85 civilian fatalities and burned over 150,000 acres, destroying more than 18,000 structures [6]. Given the trends, the problem of wildfires and their increasing catastrophic consequences are of immense policy relevance.

Understanding and predicting the occurrence of wildfires is vital for taking proper policy actions to mitigate the risks and minimize associated consequences [7-10]. An examination of historical trends of wildfires reveals puzzling cyclical patterns in fire incidences, with increasing amplitude for the consequences of fire in many areas around the globe, including the U.S. As Figure 2.1 shows, in the U.S., we have experienced an overall increasing trend of the burn rate due to wildfire, with periodic fluctuations. Interestingly, although the overall pattern of the number of fires does not follow the burn rate trend, it does show periodic oscillations. Finding the drivers of such patterns is an area of concern for natural scientists, policy researchers, and policymakers.



(a) Burned areas due to wildfires (b) Wildfire frequency

Figure 2.1. Wildfire in the U.S.1983–2018 (data from www.nifc.gov).

Wildfires start with initial fire ignitions, which can be caused by nature through lightning or through reckless human behavior. The occurrence of natural fires through lightning depends on weather conditions and shows a seasonal pattern [11]. Human-caused ignition, on the other hand, can also cause large-scale fires. In fact, in the U.S., human-ignited wildfires account for approximately 84% of wildfires nationwide [12]. Factors such as abandoned campfires, arson, and fireworks can lead to human-ignited fires [13-15]. Humans also indirectly contribute to wildfire through activities that worsen climate change [16]. The release of greenhouse gases into the atmosphere, including carbon dioxide and methane, contributes to higher temperatures [17]. A warmer climate leads to drier vegetation in forests and increases the risk of massive wildfires [18]. Furthermore, deforestation for land development reduces the ability of the forest to absorb greenhouse gases, which ultimately causes a further increase in temperature [19,20].

Despite the importance of direct human and natural contributions to wildfire, the focus of most past modeling studies has been solely on one of these two categories of causation. Touboul and colleagues developed simulation models of dynamic interactions among different kinds of vegetation such as grass and forest trees. They showed that for a wide range of scenarios, the composition of vegetation can oscillate over time [21]. Such models that focus on natural-system dynamics can explain long-term oscillatory patterns that emerge due to forest recovery delays after a wildfire. On human contributions, several statistical models have pointed to a correlation between human settlement in the wildland-urban interface (WUI) and fire activity [22-24]. In these models, human-risk perception is often an exogenous factor that affects fire. We understand that both natural and human sides of the problem are important. In fact, it has been argued for a long time that accounting for dynamic connections between social and ecological systems is essential in developing sustainable environmental policies [25]. Therefore, we hypothesize that the interaction between natural and human systems contributes to wildfire dynamics, increasing their complexity and mitigation challenges. To develop proper policies, attention should be paid to both sides of the larger system and the interactions between the two. Our primary objective in this paper is to explore potential causes of such patterns by developing and examining a feedback-rich dynamic simulation model that represents both social and natural dynamics in a coupled system.

Figure 2.2 presents our study framework, which is in line with a body of the ecological literature that examines a family of phenomena referred to as coupled human-natural systems (This area has been a major area of investigation at the U.S. National Science Foundation). The framework includes dynamics specific to vegetation (natural systems) and human systems (behavioral dynamics). In interaction, the two pieces are connected through the human sector that receives information regarding recent fire cases and influences the human risk perception, as the fire risk is influenced by the perceived information [26]. Humans contribute to fire through human-caused ignition or the development of vulnerable properties based on their risk perception.

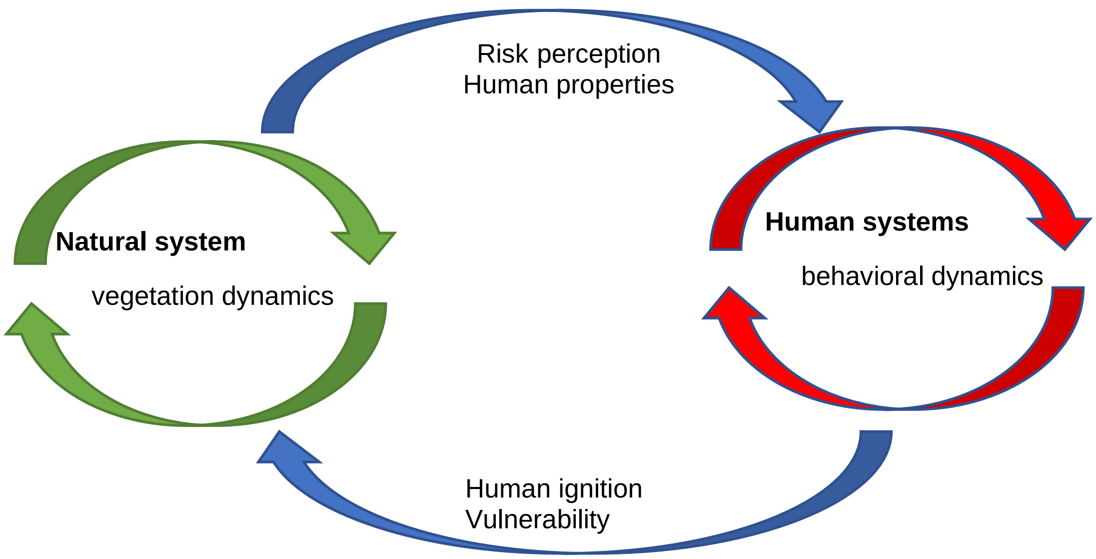


Figure 2.2. Our study framework of wildfire as an outcome of a coupled human-natural system (adapted from https://www.nsf.gov/pubs/2018/nsf18503/nsf18503.htm and adjusted for the case of wildfires).

2. Background: Models of disasters

While our focus is on the specific problem of wildfire, it is important to pause and offer a quick review of various modeling approaches of similar natural disasters, mainly from a methodological standpoint. There is a wide range of modeling approaches applied to natural-disaster studies in general and wildfires in particular. Such modeling can be differentiated based on their unit of analyses, time frames, mathematical modeling techniques, boundaries, and specific application cases.

A large body of natural disaster models has been devoted to spatial modeling [27-29]. In a typical spatial wildfire model, the goal is to replicate fire progression throughout different regions. Such models are powerful in showing how, in what sequence, and the timing of different areas may become fire susceptible. Spatial models can also take different forms depending on the geographical units of analysis (e.g., state, county). Connection networks between different units can affect fire progress, and such models become more useful as they move toward modeling network structures.

The second group of models of natural disasters includes agent-based individual-level models. Models of evacuation often take such levels of analysis and explore the flows of individuals after a disaster [30,31]. In the wildfire context, with a focus on fire progression, agent-based models may consider vegetation units as agents. Such models would lead to a spatial model of interacting elements that depict endogenous fire propagation from one unit of vegetation to another. Rahmandad and Sterman (2008) stressed that in many contexts, detailed agent-based models may not go beyond what one could learn from an aggregate differential-equation model, especially when the heterogeneities across the agents are limited and connection networks are symmetrical and almost complete.

On the other side, aggregate natural disaster models exist in which vegetation is often modeled with a few major variables but no regional details. As compartmental models, these often include differential equations and formulate vegetation flows and aging of trees in a dynamic framework. Within aggregate models, the extent to which variables are treated as endogenous variables (that is, they respond to changes in state variables) is a significant factor for differentiating. Simon Levin and colleagues [21,32,33] offered different variations of aggregate, differential-equation models of vegetation. An interesting outcome of such models from a complex-systems point of view is the depiction of bifurcation that the model's outcomes substantially change from a steady-state to a goal-seeking or s-shaped behavior, or even long-term oscillations for different ranges of parameter values.

Within the system dynamics community, there is also a rich body of literature of modeling environmental problems [25,34,35]. Michael Deegan has conducted methodologically relevant work [36] in a slightly different natural-disaster setting. In his system dynamics Ph.D. dissertation, Deegan (2007) modeled flood- damage dynamics in a typical flood-prone community, considering long-term community reactions to recent floods and related damages. Deegan focused on hypothetical flood cases, intending to show how seemingly similar external events (here, major rain) can cause different damage levels depending on the community's reactions and investment in vulnerable properties. What differentiates his work from others is that Deegan's model is feedback-rich, and dynamic outcomes are created within the model rather than by an external time series. George Richardson, in his reflections on the foundations of system dynamics [37], argued that a powerful model is differentiated by its endogenous perspective to complex problems. He brings up Deegan's work as a powerful example and writes, "*Deegan's (2007) extensive analysis suggests … an endogenous view of the dynamics of flood damage that takes account of the human role in creating property vulnerable to flood damage… [Deegan's model] traces the dynamics of vulnerability to the interacting actions of the capacity of the local environment to withstand floods, development pressure, property tax needs, perceived risks of development, moral hazard, policy entrepreneurs, and other people pressures*" (Richardson 2011, p.234). In some respects, our approach to modeling wildfires resonates with Deegan's flood mitigation work by looking at vulnerability as an endogenous property of the system affected by human risk perception.

What makes these aggregate models powerful is that they are relatively small (have fewer equations), and when the details are removed, they turn the focus on system responses and feedback loops without losing many systems-level insights [38]. Modelers can also better communicate insights from small models with stakeholders [39]. It is important to note that small, powerful models are not easy to build, and they are often the result of many rounds of complex and detailed modeling [40], which has also been the case in our study. Given our problem scope, we follow the same modeling approach.

3. Model structure and key formulations

Different models use varying terms to represent vegetation heterogeneities in a forest area. For the purposes of parsimony, our model represents the entire forest area by three simple stock variables of areas occupied by strong vegetation (S), occupied by flammable vegetation (F), and empty areas (E), all of which are shown as stock variables in Figure 2.3 (variables in boxes). Strong vegetation is often resistant to fire and only large-scale fires can burn them. Highly flammable vegetation includes damaged or any vegetation that can burn fast (including grass). This type of vegetation can burn quickly, and lightning or human ignition often affects flammable vegetation first. Burning can cause fire propagation to strong vegetation. Empty areas lack any vegetation. The sum of these stocks is constant. While our figure is a simple representation of forest areas, the logic is consistent with studies that have offered more detail on vegetation types.

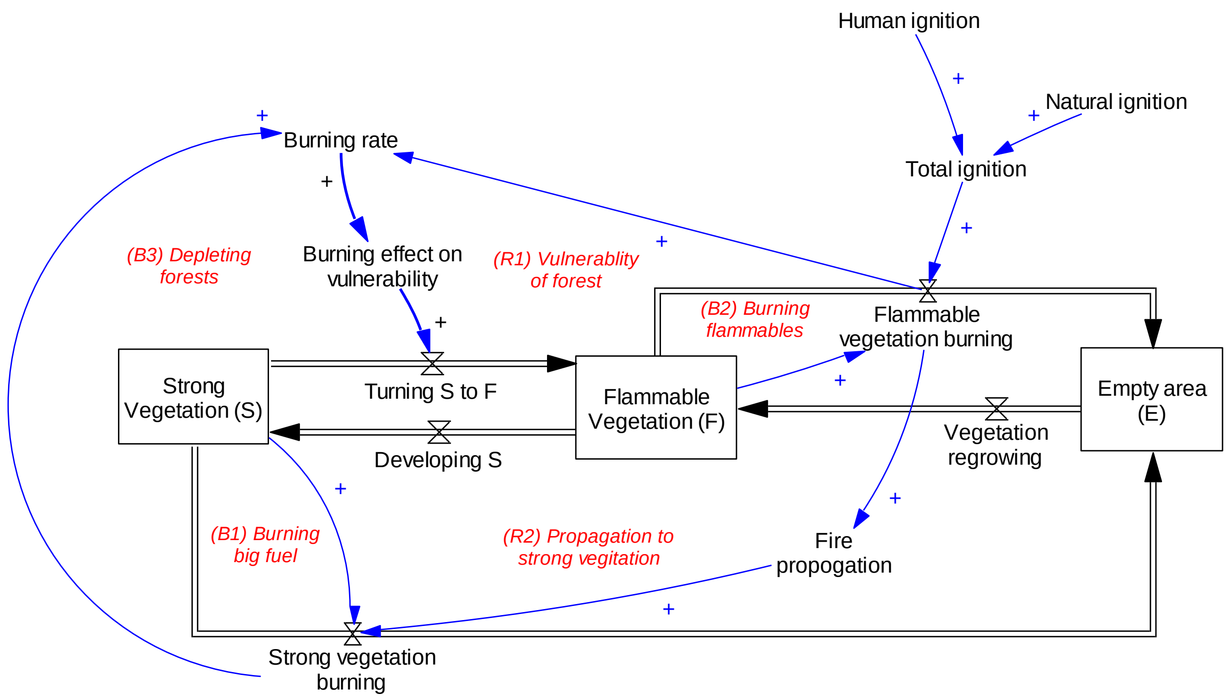


Figure 2.3. A stock-flow diagram of vegetation.

In this model, loops B1 and B2 represent the deterioration of strong and flammable vegetation through fire. As stated, fire can increase the vulnerability of strong vegetation by burning the surrounding area and making it more susceptible to fire. This mechanism is shown by loops R1 (burning of flammable vegetation further increases flammable vegetation) and B3 (burning of strong vegetation makes other strong vegetation vulnerable to fire).

We can represent the relation between the stock variables by the following differential equations.

$$\frac{dS}{dt} = \frac{F}{\tau_1} - (\alpha + \gamma_S)S \tag{1}$$

$$\frac{dF}{dt} = \frac{E}{\tau_2} + \alpha S - (\frac{1}{\tau_1} + \gamma_F)F \tag{2}$$

$$\frac{dE}{dt} = -\frac{dF}{dt} - \frac{dS}{dt} \quad , \tag{3}$$

where γ_S and γ_F are the fractional burning rate of strong and flammable vegetation, respectively; α is the rate of making strong vegetation to flammable; τ_1 is the average time for flammable vegetation to become strong; and τ_2 is the average time for the empty space to grow flammable vegetation, where often $\tau_2 << \tau_1$. Thus, the total burn rate from both types of vegetation (B) is

$$B = \gamma_F F + \gamma_S S \quad . \tag{4}$$

In this equation, γ_F, the rate at which flammable vegetation is burned is a function of the total of human and natural ignitions. However, γ_S, the fractional burning rate of strong vegetation, depends on the burning rate of flammable vegetation and happens when fire propagates in the forest—i.e., $\gamma_S = f(\gamma_F F)$. We formulate f using a sigmoid function (Table 2.1). Furthermore, α, the rate of becoming flammable for strong vegetation as a result of fire is $\alpha = \sigma B$ where σ is the burning effect on vulnerability.

Generally, the public attitude towards making risky decisions is influenced by their level of risk perception. In the case of wildfire, there is a wide range of evidence that people's attention to the problem and possibly the associated risk perception has changed over time. Figure 2.4 depicts the frequency of Google searches for the word "wildfire" in the U.S. The trends are oscillatory, and there is a 0.4 correlation between search and area burned from 2004 to 2018.

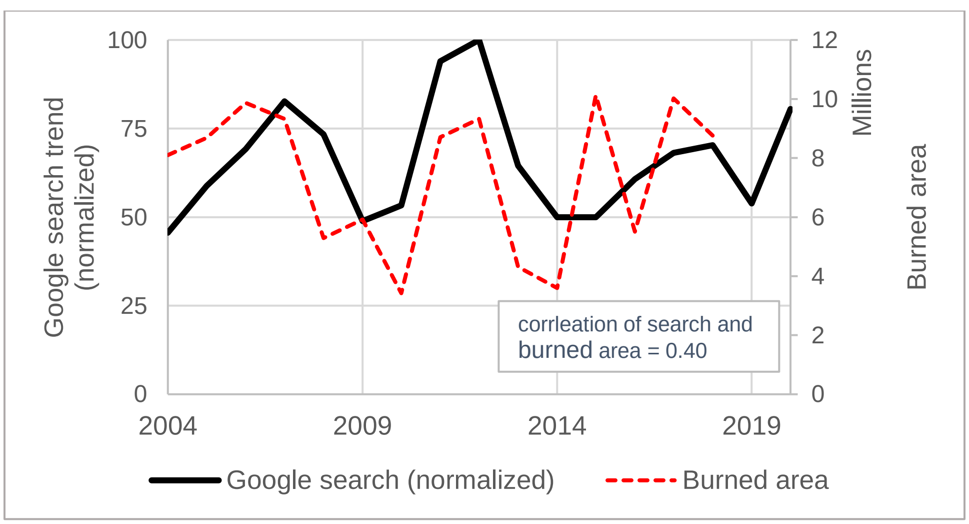


Figure 2.4. Google search trend for "wildfire" in the U.S. and its correlation with annual burned area.

There is a body of research focused on how the perception of wildfire risk is associated with mitigation actions [41-44]. A study of a fire-prone area in Colorado revealed that a single extreme wildfire does not significantly impact risk perception [43]. Furthermore, evidence on people's fire-risk perception shows that any change in fire risk perception does not last more than a couple of years [44]. We construct the effect of risk perception on human actions based on the abovementioned research with two important characteristics. First, the overall wildfire activity in recent years shapes people's fire risk perception; second, the effect of wildfire on people's perception vanishes as time pass.

We include two major mechanisms to depict the effects of change in risk perception as shown in Figure 2.5. The loop B4, complacency, represents the human contribution to fire through reckless behaviors, which can cause fire ignition. Loop B5, vulnerable properties, represents property building in forest areas. Such properties increase human interaction with the natural environment and the likelihood of human-made ignition. We also consider the fact that such properties might be targets of fire themselves, loop B6.

In this model risk perception, $\bar{B}$ is formulated as a δ_1-year lagged variable of burn rate (B), assuming there is no systematic bias in risk perception. Total ignition of I includes human-caused ignition (I_H) and natural ignition due to lightening (I_N), with the latter assumed as constant in our model. Human-caused ignition increases by human settlements in the area and is inversely related to their risk perception. Assuming human settlements are represented by vulnerable properties, V, we formulate I_H as $I_H(V, \bar{B})$ where $\frac{\partial I_H}{\partial V} > 0$ and $\frac{\partial I_H}{\partial \bar{B}} < 0$. For the purposes of parsimony, we formulate effect $\bar{B}$ on I_H using a linear function (see the Appendix).

Although Martin et al. [42] discussed how different stakeholders (including insurance companies and federal agencies) could increase the sensitivity of humans to risk perception, they did not provide any quantitative estimation of this value.

Finally, vulnerable properties, V, which can change over time is formulated as

$$\frac{dV}{dt} = V(\theta E_{bt} - \rho) \quad . \qquad (5)$$

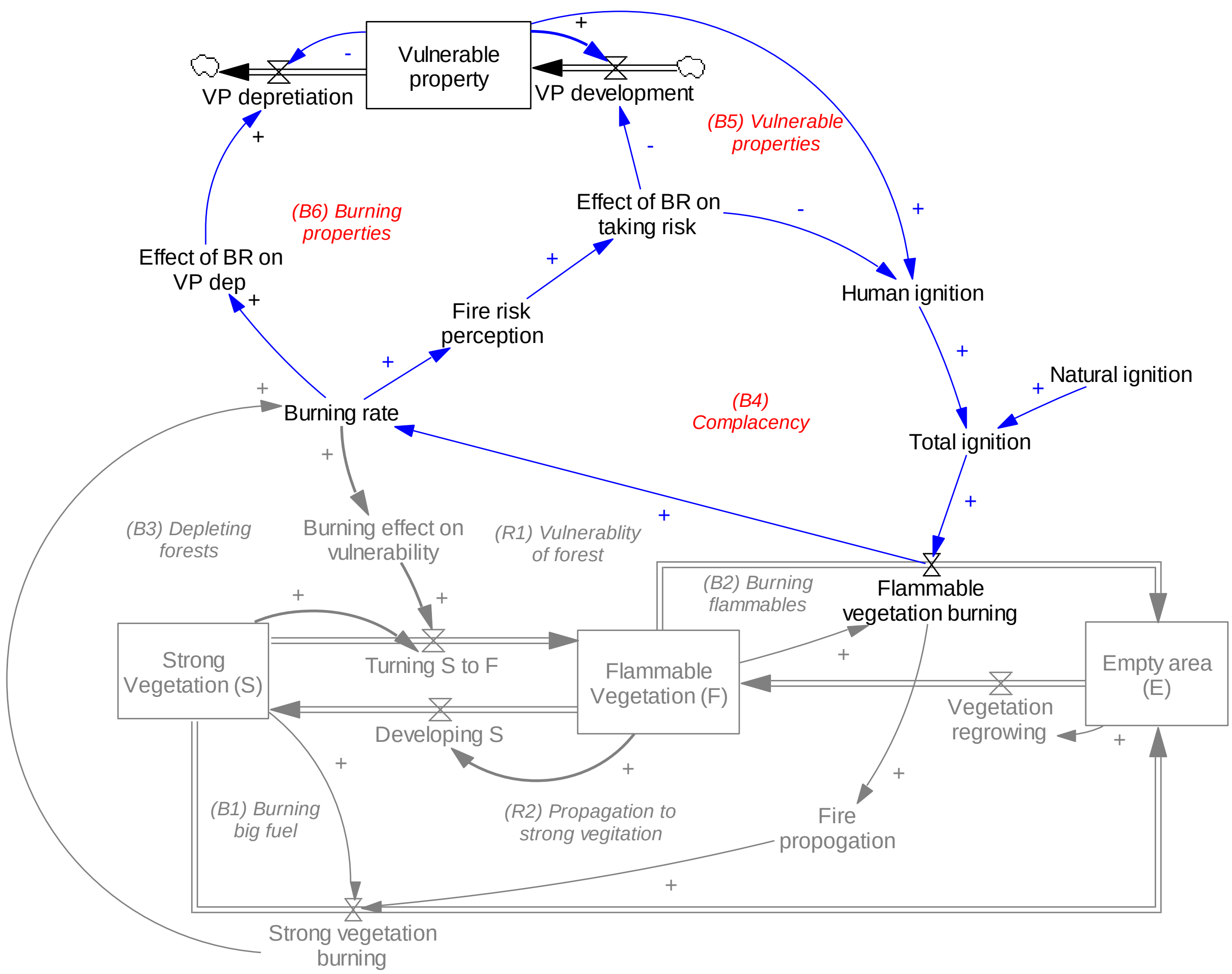


Figure 2.5. The human subsystem as connected to the natural subsystem (grey),

The term, θE_{bt} represents property development and is assumed to be proportional to the current properties and negatively affected by risk perception. The term ρV represents the demolition of properties. Demolition in our model is mainly due to the fire, that is, $\rho = \rho(B)$.

4. Parameter values

The introduced model is generic and can be simulated for a wide range of parameter values. Table 2.1 reports parameter values used for base run simulations. Some of the values are consistent with the literature, while others are selected to examine variation conditions in different forestry settings.

Table 2.1. Parameter Values for a Base Run Simulation

Parameter	Value	Unit
τ_1	2	year
τ_2	10	year
δ_1	0.5	year
S	Initial value: 0.5	Million acres
F	Initial value: 0.4	Million acres

E	Initial value: 0.1	Million acres
γ_S	$0.8 * (1 + e^{-5*(\frac{\gamma_F F}{n} - 1)})^{-1}$	1/year
n	0.1	Million acres
I_N	0.5	scalar
I_H	Initial value: 0.3	scalar
V	Initial value: 0.4	Million acres
σ	0.05	1/Million acres

Our simulation experiments include a base run simulation and a range of policy and scenario tests as listed in Table 2.2. The table also provides details on how each test is implemented in our analysis. Specifically, we analyze the linkage between natural dynamics and human perception and its consequences on fire development by changing the sensitivity of risk perception to the burn rate (Test T2). We then examine the effects of three different policies: limiting the development of vulnerable properties (P1), prescribed and controlled burning of flammable vegetation (P2), and effective firefighting that limits penetration of fire from flammable vegetation to strong vegetation (P3).

Table 2.2. Simulation Experiments

Simulation tests	Operationalization
T1: Base run	Parameter values in table 2.1 are used for this scenario.
T2: Coupling effect	Sensitivity of risk perception to burn rate is changed by changing risk perception delay from 1 year (base run) to • 0.5 year (T2a: higher sensitivity), • 2 years (T2b: lower sensitivity), and • 100 years (T2c: least sensitivity – almost disjointed systems).
T3: Policy tests	Three policies are tested. • P1: Limit vulnerable property development. • P2: Prescribed burning. • P3: Effective firefighting. P1 is implemented by making vulnerable property development smaller than 1%. P2 is implemented by adding outflow from flammable vegetation to an empty area with the value of ωF, where ω is the percent of prescribed burning set at 0.2/year. P3 is implemented by changing γ_S to 10% of its current value.

5. Simulation results

5.1. Base run simulation

Figure 2.6 shows the results of the base run simulation. In this scenario, strong vegetation declines over time, while the empty area and flammable vegetation have increasing trends. As such, more fuel would be available for burning, and the wildfire can burn broader areas. Panel (a) shows an oscillatory trend for the burn rate with an average upward trend. The observed pattern in the burn rate can be traced back to the patterns of human ignition (Panel b), and the growing trend of vulnerable properties (Panel c). In addition, the results show the long-term declining trend of strong vegetation in our base line simulation (Panel d); over time, stronger vegetation is replaced by flammable vegetation which can lead to more fire. This change in vegetation composition effectively increases the average burn rate. Over time, with more flammable vegetation and with the expansion of vulnerable properties, the likelihood of human-made ignition increases.

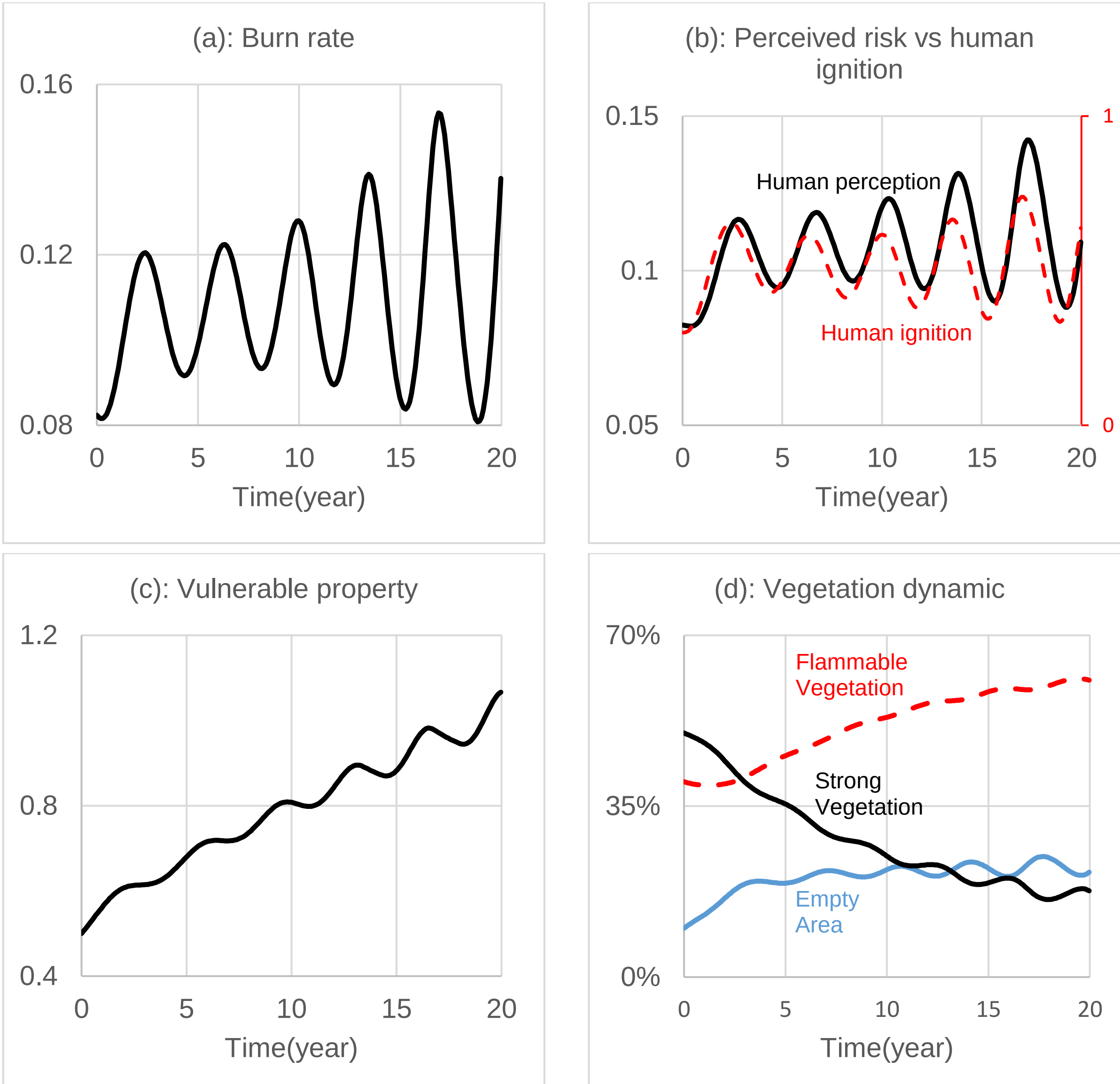


Figure 2.6. Base run simulation for a 20-year run of the model.

5.2. Coupling effects

Figure 2.7 shows how the relation between perceived fire risk and the burn rate influences the system. The black line is the base run simulation for comparison. The blue dashed line depicts the condition in which risk perception changes extremely slowly, and the human system is almost disconnected from the natural system. In this situation, if humans underestimate the fire potential, the system burns down nature, resulting in a catastrophic environmental outcome as depicted in panel (a). Panel (a) shows that the burn rate overshoots in the short term but relatively declines due to lower natural resources to burn.

Panel (b) displays the total burn rate throughout the study time to cast further insight into the burn rate sensitivity to perceived risk. The overall burn rate does not significantly change when the risk perception changes from 0.5 to 2, indicating the difference among burn rates in panel

(a) is more about the fluctuation timing, but not the size. However, an additional rise in the sense of risk greatly raises the overall burn rate, as seen in panel (a).

In the case of prolonged change in risk perception, human ignition continues to increase (panel c) as the perceived risk changes slowly. Furthermore, vulnerable properties are being built faster than their demolition (panel d). A slighter delay in perception leads to a higher frequency of oscillation as depicted in the graphs by the red dashed lines and a longer delay in a lower frequency oscillation, as shown by the purple graphs. Overall, the results are not much different from the base run. We are losing forests (panel e) and have periodic burn rates of increasing magnitude over time.

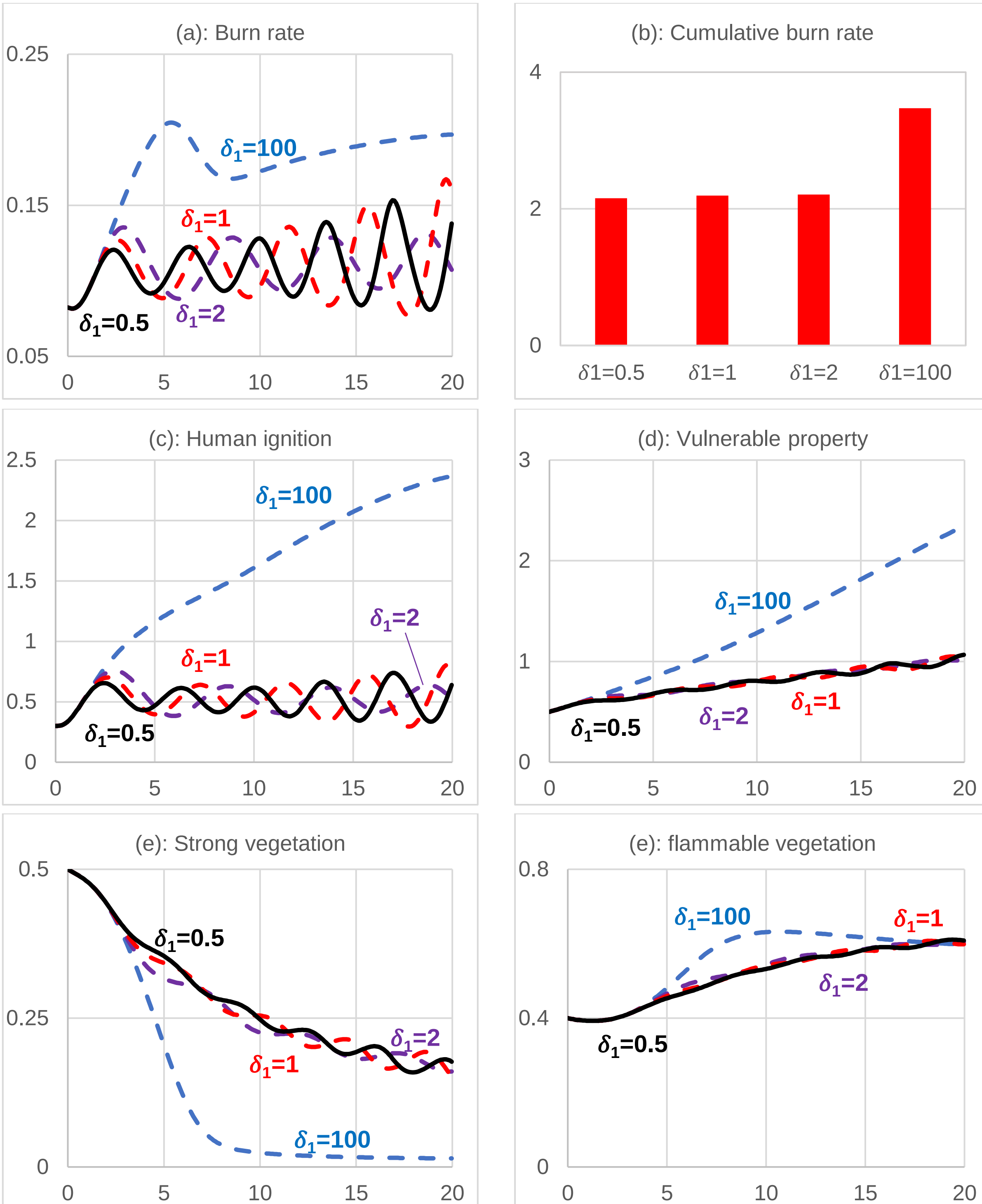


Figure 2.7. Coupling effect analysis for 20 years.

5.3. Policy experiments

Here we examine the impact of implementing three proposed policies introduced in Table 2.2. To prevent the initial condition and transition periods affecting our comparison of proposed policies, we impose each policy at the fifth year and compared the total burn rates between 10 and 20 years. Figure 2.8 shows the effect of these policies on different variables over time.

Comparing these three policies in the case of the total burn rate as shown in panel illustrates that (b) increasing firefighting effectiveness (P3) impact is the most effective through a reduction in the total burn rate by 9%. Prescribed burning (P2) and limiting vulnerable property development (P1) causes a 4.5% and 4.9% reduction in the burn rate, respectively.

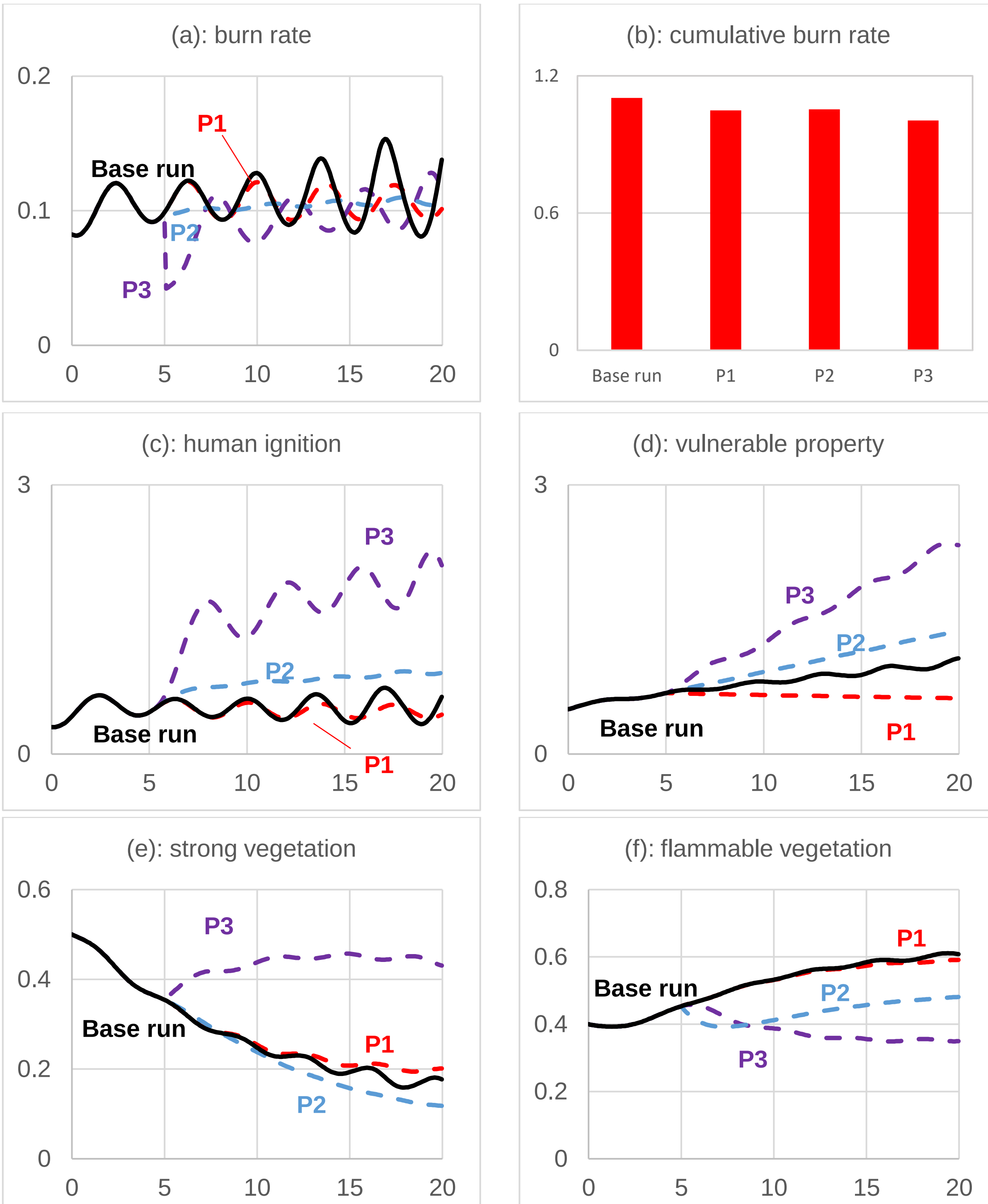


Figure 2.8. Policy implementation. Note: P1: limits vulnerable property development; P2: prescribed burning; and P3: effective firefighting.

Panels (a) and (b) show the burn rate over time. All three policies reduce the burn-rate magnitude compared to the base run. P3 is more effective in early burning-rate reduction than P2, but they ultimately result in similar behavior. It is worth noticing that P1 has the most effect on long-run fluctuation reduction, although its total effect in the time span is less than P3. It seems that firefighting is more effective in the short run, but it fails to dampen the fluctuation and instead limits its growth. This is partly because of the increase in human ignition and settlement due to the success of firefighting in the short run. As a result, people perceive less fire danger and continue to engage in high-risk activities and expand housing in the WUI. The result is further fluctuation in the burn rate even when P3 is implemented. On the other hand, the WUI expansion limitation policy can effectively reduce the burn-rate fluctuation in a timely manner.

Change in human ignition is provided in panel (c). Different levels of human-made ignition is observable, and the reason is that people adjust their high-risk behavior with burn rate, and not with number of fires. In the firefighting policy as for a given level of ignition, the burn rate declines, we observe more risky behavior and more human-made ignition. It is interesting to note that, as panel (c) shows we end up with more WUI under policies 2 and 3. In fact the reason is that the firefighting and prescribed burning only affect natural sector of the model, decrease burn rate, which decrease risk perception and in turn result in more WUI development. On the other hand, PI directly targets WUIs.

Panel (e) displays the change in strong vegetation, which shows that P2 causes the most reduction in forest tree cover. The reason is because burning flammable vegetation damages young trees and prevents them from developing into solid vegetation. On the other hand, P3 has the least effect on strong vegetation by slowing the damage to young trees and confining the fire. Panel (f) shows the flammable vegetation dynamic after imposing each policy. P3 and P2 reduce flammable vegetation more than P1. However, there is an important difference in how these policies cause the reduction in flammable vegetation. In comparing panels (a) and (b), we see that while P3 causes further increases in the strong vegetation, P2 causes an increase in the empty area.

Overall, it looks like each policy has some marginal effect on containing wildfire, though the magnitudes of effect are not considerable.

5.4. Combination policy implementation analysis

To better understand the impacts of our policies, we run different pairs of policies simultaneously. The results illustrate the nonlinear incremental impacts between policies. Simply put, it appears that the impact of several policies is enforced when combined synergistically. In other words, applying several policies might have a greater overall impact than the sum of the policies' individual effects and suggests that policymakers should avoid searching for a panacea and adopt a broad range of approaches thoughtfully.

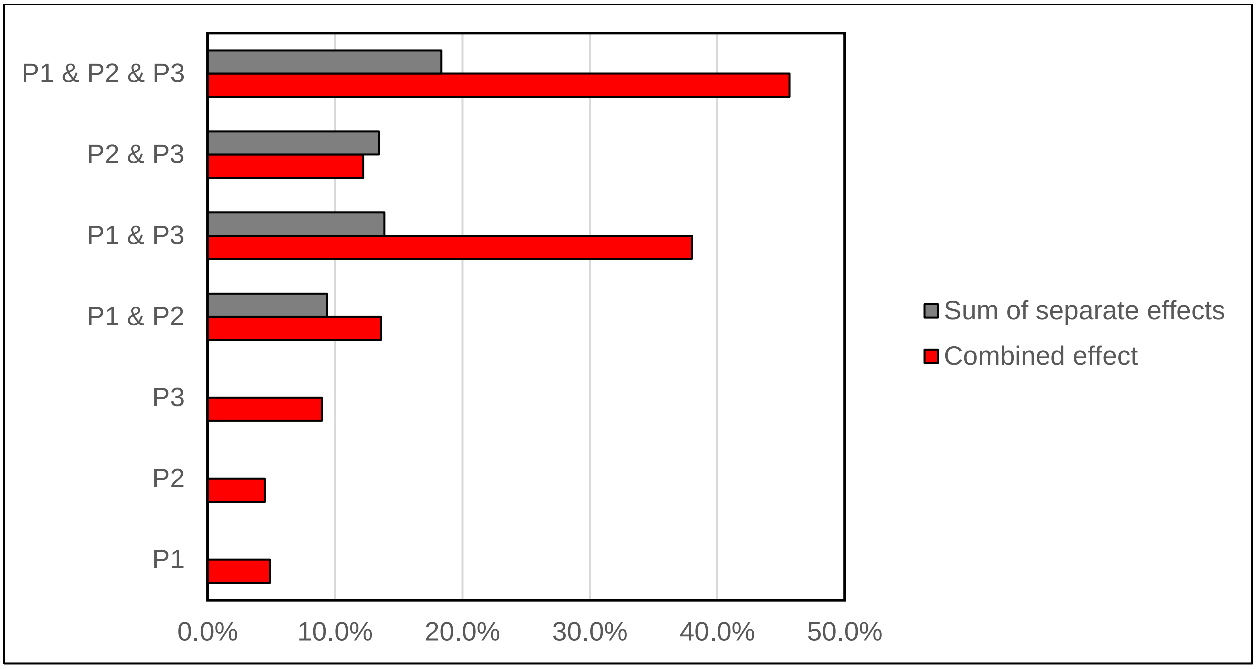


Figure 2.9. The nonlinear effect of policies. The benefits of implementing multiple policies differ from the sum of the effect of policies. The figure shows the percent of burn rate reduction. Note: P1: limit vulnerable property development; P2: prescribed burning; and P3: effective firefighting.

The results of multiple policy implementations along with single ones are presented in Figure 2.9. For example, P1 and P2 each reduce the total burn rate by 4.9% and 4.5%, respectively. While the summation of these effects is 9.4%, simultaneously implementing P1 and P2 lead to a 13.6% burn-rate reduction—P1 controls the human ignition and P2 reduces the flammable vegetation stock—together the burn rate is more affected than if implemented separately. The case is more interesting when P1 and P3 are imposed together. The result is a 38% burn-rate reduction compared to 13.9%, which is the sum of solely implementing each policy. The synergic effect happens because P3 lets the flammable vegetation (mainly young trees) age and become strong vegetation. Furthermore, the P1 also prevents human ignition from growing as fast as a single P3 implementation.

An interesting case happens when P2 and P3 are implemented together. The synergic effect is less than the sum of separate implementation, mainly because both policies affect the vegetation dynamic and not the human factor in the wildfire. P2 and P3 both cause a lower initial burn rate, but due to the reduction in perceived risk of wildfire and expansion of WUI, this effect quickly disappears. This is another evidence for the importance of considering the problem as an interconnected natural and human system, where effective policies should address both sides.

Finally, imposing all policies together has the most impact on the total burn rate (45.7%), which is much more than the sum of each policy's effect (18.3%). The reason relates to the multiplicative effects of the policies when the flammable vegetation is reduced, by both prescribed burning and allowing them to develop into strong vegetation, and the WUI expansion is controlled decreasing human ignition. In this case, while the burn rate is reduced, P1 prevents excess human ignition due to a lower fire risk.

6. Discussions and conclusion

In this paper, we developed a system dynamics model of wildfire spread in a hypothetical scenario and simulated the effects of several important mechanisms in determining the burn rate, fire frequency, and public risk perception of wildfire. The model included two major sectors of the natural and human subsystem that were connected through the human contribution to ignition and the human risk perception of fire. We simulated the model for a wide range of scenarios that represent different levels of human sensitivity to evolving fires and a range of policy containment measures. Our results show how humans and vegetation determine wildfire activity, defining wildfire as a human-natural coupled system. The findings are important in their relative changes, not their absolute values, because of the model's hypothetical assumption.

We conducted several simulation experiments with the model. The results show a wide range of oscillatory patterns in different scenarios and policy conditions. The base run depicted the possibility of an oscillatory outcome in human-caused ignition and an oscillatory pattern in the burn rate with an overall increasing trend. The decrease in strong vegetation, the increase in vulnerable property causes an increasing trend in burn rate while dynamics of human perception affect the oscillatory pattern.

Our study contributes to the literature of modeling natural disasters and specifically wildfire studies. We offer the first model of the coupled human-natural system of wildfire. Our study builds on several past models of ecological dynamics [21], particularly in wildfire dynamics, [45] and extends them to include human interaction with natural systems. Our work is different from past spatial models of wildfires. In spatial modeling of wildfires, the human effect is spatially static. Here we show that the same population could ignite a different number of fires and affect the wildfire behavior. Our different approach from past studies results in different outcomes as well. For example, we point to the sources of policy resistance in containing wildfire in terms of how risks are perceived and how properties are built adjacent to natural resources.

Our study resonates with some of the past system dynamics models of other natural disasters [36,37]. We take an endogenous approach to the concept of system vulnerability by considering the human element as a part of the system which both reacts to the problem and contributes to problem. The importance of feedback-rich modeling has previously shown its value in sustainable environmental management, including water quality, waste management, and water supply [25]. Here we propose a similar approach for wildfire management and aim to understand important mechanisms shaping wildfire behavior.

The study has several policy implications. We compared three policies: prescribed burning, vulnerable property control, and firefighting effectiveness enhancement. We showed that firefighting effectiveness is more effective in reducing the total burn rate than other proposed policies. More importantly, we showed that simultaneously implementing policies can lead to a synergic effect that can surpass the sum of the effect of solely implementing the same policies. For example, while controlling development of vulnerable properties and effective firefighting each reduce the burn rate 4.9% and 9%, respectively, performing both policies results in a 38% burn rate reduction. Such a synergic effect points to the absence of a silver bullet in controlling wildfires, suggesting that effective policies should target both human- and natural-sectors of the system to maximize their effectiveness. In other words, since wildfire is an outcome of a coupled system that includes highly interdependent human and nature sectors, one cannot solve it by solely focusing on one sector.

References

1 Scott, J. H., Thompson, M. P. & Calkin, D. E. A wildfire risk assessment framework for land and resource management. (2013).

2 McFarlane, B. L., McGee, T. K. & Faulkner, H. J. I. J. o. W. F. Complexity of homeowner wildfire risk mitigation: an integration of hazard theories. **20**, 921-931 (2012).

3 Radeloff, V. C. *et al.* Rapid growth of the US wildland-urban interface raises wildfire risk. **115**, 3314-3319 (2018).

4 Philantropy, C. f. D. *2020 North American Wildfire Season* <https://disasterphilanthropy.org/disaster/2020-california-wildfires/> (2020).

5 NBCNews. *Deadly fires in California have claimed at least 30 lives this year. [30 September 2020]*, <https://www.nbcnews.com/news/us-news/deadly-fires-california-have-claimed-least-30-lives-year-n1241632> (2020).

6 Post, W. The deadliest, most destructive wildfire in California's history has finally been contained. (2018).

7 Catry, F. X., Rego, F. C., Bação, F. L. & Moreira, F. J. I. J. o. W. F. Modeling and mapping wildfire ignition risk in Portugal. **18**, 921-931 (2010).

8 De Vasconcelos, M. P. *et al.* Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks. **67**, 73-81 (2001).

9 Miyanishi, K. *Forest Fires: Behavior and Ecological Effects*. (Elsevier, 2001).

10 Bonazountas, M., Kallidromitou, D., Kassomenos, P., Passas, N. J. H. & Assessment, E. R. Forest fire risk analysis. **11**, 617-626 (2005).

11 Loeb, L. B. J. S. A. The mechanism of lightning. **180**, 22-27 (1949).

12 Balch, J. K. *et al.* Human-started wildfires expand the fire niche across the United States. *Proceedings of the National Academy of Sciences* **114**, 2946-2951 (2017).

13 Vega, J. *et al.* in *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology.* 1-2.

14 Prestemon, J. P. & Butry, D. T. in *The Economics of Forest Disturbances* 123-147 (Springer, 2008).

15 Bartlein, P. J., Hostetler, S., Shafer, S., Holman, J. & Solomon, A. in *5th Symposium on Fire and Forest Meteorology.* 16-20.

16 Abatzoglou, J. T. & Williams, A. P. Impact of anthropogenic climate change on wildfire across western US forests. *Proceedings of the National Academy of Sciences* **113**, 11770-11775 (2016).

17 Fleming, J. R. Joseph Fourier, the 'greenhouse effect', and the quest for a universal theory of terrestrial temperatures. *Endeavour* **23**, 72-75 (1999).

18 Westerling, A. L., Hidalgo, H. G., Cayan, D. R. & Swetnam, T. W. Warming and earlier spring increase western US forest wildfire activity. *science* **313**, 940-943 (2006).

19 Fiddaman, T., Sterman, J., Jones, A. P., Sawin, E. R. & Siegel, L. S. C-ROADS: Climate--Rapid Overview and Decision-Support Simulator. (2008).

20 Rooney-Varga, J. N. *et al.* The climate action simulation. *Simulation & Gaming* **51**, 114-140 (2020).

21 Touboul, J. D., Staver, A. C. & Levin, S. A. On the complex dynamics of savanna landscapes. *Proceedings of the National Academy of Sciences* **115**, E1336-E1345 (2018).

22 Pew, K. & Larsen, C. GIS analysis of spatial and temporal patterns of human-caused wildfires in the temperate rain forest of Vancouver Island, Canada. *Forest ecology management* **140**, 1-18 (2001).

23 Matin, M. A. *et al.* Understanding forest fire patterns and risk in Nepal using remote sensing, geographic information system and historical fire data. *International journal of wildland fire* **26**, 276-286 (2017).

24 Vilar, L., Woolford, D. G., Martell, D. L. & Martín, M. P. A model for predicting human-caused wildfire occurrence in the region of Madrid, Spain. *International Journal of Wildland Fire* **19**, 325-337 (2010).

25 Stave, K. Participatory system dynamics modeling for sustainable environmental management: Observations from four cases. *Sustainability* **2**, 2762-2784 (2010).

26 Langarudi, S. P., Silva, C. G. & Fernald, A. G. Measure more or report faster? Effect of information perception on management of commons. *System Dynamics Review* **37**, 72-92 (2021).

27 Keane, R. E., Drury, S. A., Karau, E. C., Hessburg, P. F. & Reynolds, K. M. A method for mapping fire hazard and risk across multiple scales and its application in fire management. *Ecological Modelling* **221**, 2-18 (2010).

28 Finney, M. A. *et al.* A method for ensemble wildland fire simulation. *Environmental Modeling Assessment* **16**, 153-167 (2011).

29 Davis, C. *et al.* Prediction of landfalling hurricanes with the advanced hurricane WRF model. *Monthly weather review* **136**, 1990-2005 (2008).

30 Chen, X. & Zhan, F. B. Agent-based modelling and simulation of urban evacuation: relative effectiveness of simultaneous and staged evacuation strategies. *Journal of Operational Research Society* **59**, 25-33 (2008).

31 Yin, W., Murray-Tuite, P., Ukkusuri, S. V. & Gladwin, H. An agent-based modeling system for travel demand simulation for hurricane evacuation. *Transportation research part C: emerging technologies* **42**, 44-59 (2014).

32 Staver, A. C. & Levin, S. A. Integrating theoretical climate and fire effects on savanna and forest systems. *The American Naturalist* **180**, 211-224 (2012).

33 Schertzer, E., Staver, A. C. & Levin, S. A. Implications of the spatial dynamics of fire spread for the bistability of savanna and forest. *Journal of Mathematical Biology* **70**, 329-341, doi:10.1007/s00285-014-0757-z (2015).

34 Fiddaman, T. S. Exploring policy options with a behavioral climate–economy model. *System Dynamics Review: The Journal of the System Dynamics Society* **18**, 243-267 (2002).

35 Sterman, J. *et al.* World climate: A role-play simulation of climate negotiations. *Simulation & Gaming* **46**, 348-382 (2015).

36 Deegan, M. A. *Exploring US flood mitigation policies: a feedback view of system behavior*, Rockefeller College of Public Affairs and Policy, University at Albany–SUNY., (2007).

37 Richardson, G. P. Reflections on the foundations of system dynamics. *System Dynamics Review* **27**, 219-243 (2011).

38 Rahmandad, H. & Sterman, J. Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Management Science* **54**, 998-1014 (2008).

39 Ghaffarzadegan, N., Lyneis, J. & Richardson, G. P. How small system dynamics models can help the public policy process. *System Dynamics Review* **27**, 22-44 (2011).

40 Ghaffarzadegan, N. & Larson, R. C. SD meets OR: a new synergy to address policy problems. *System Dynamics Review* **34**, 327-353 (2018).

41 McGee, T. K., McFarlane, B. L. & Varghese, J. An examination of the influence of hazard experience on wildfire risk perceptions and adoption of mitigation measures. *Society and Natural Resources* **22**, 308-323 (2009).
42 Martin, W. E., Martin, I. M. & Kent, B. The role of risk perceptions in the risk mitigation process: the case of wildfire in high risk communities. *Journal of environmental management* **91**, 489-498 (2009).
43 Champ, P. A. & Brenkert-Smith, H. Is seeing believing? Perceptions of wildfire risk over time. *Risk Analysis* **36**, 816-830 (2016).
44 McCaffrey, S. Thinking of wildfire as a natural hazard. *Society and Natural Resources* **17**, 509-516 (2004).
45 Schertzer, E., Staver, A. & Levin, S. A. Implications of the spatial dynamics of fire spread for the bistability of savanna and forest. *Journal of mathematical biology* **70**, 329-341 (2015).

Appendix A. Model formulation

Table A2.1. shows the full set of equations for the simulated model. In any program, these equations can be used to build a model. Nevertheless, these are used primarily to build a model in the Vensim DSS program. The Vensim DSS platform is an upgraded version of the Vensim PLE program that is available free for academic purposes. The DSS edition has a handful more options including optimization and sensitivity analysis. The model can therefore be developed using the Vensim PLE version with the following equations. The time stage for all simulations is 0.0078125 years. The model is run for 20 years but can be run for any other period.

Table A2.1. Complete set of model's equations, definition of parameters and units

Parameter name	Parameter equation and definition and units of its parameters
1. Vulnerable property	$V = \int_0^t ((\theta V)e^{-k\bar{B}} - \rho V)dt + I_V$
2. Indicated human ignition	$i_H = V \times E_{bt} \times h_m$
3. Human ignition	I_H =Third order delay of i_H; delay duration = δ_2; $I_H(t = 0) = 0.3$
4. Burning rate	$B = \gamma_F F + \gamma_S S$
5. Strong vegetation	$S = \int_0^t \left(\frac{F}{\tau_1} - (\alpha + \gamma_S)S\right) dt + I_s$
6. Flammable vegetation	$F = \int_0^t (\frac{E}{\tau_2} + \alpha S - (\frac{1}{\tau_1} + \gamma_F)F)dt + I_F$

7. Empty area	$E = \int_0^t \left(-\frac{dF}{dt} - \frac{dS}{dt}\right) dt + I_E$
8. Vulnerable property development	$V_i = E_{bt} \times \vartheta \times V$
9. VP depreciation	$V_o = V \times E_{bv}$
10. Vegetation regrowing	$V_r = \frac{E}{\tau_1}$
11. Turning S to F	$T_{sf} = \sigma \times B \times S$
12. Total ignition	$T_i = I_H + I_N$
13. Strong vegetation burning	$S_{br} = \rho \times F_p \times S$
14. Flammable vegetation burning	$\gamma_F = f_r \times F \times T_i$
15. Fire propagation	$F_p = 0.8 * (1 + e^{-5*(\frac{\gamma_F F}{n} - 1)})^{-1}$
16. Effect of BR on VP dep	$E_{bv} = B \times \mu$
17. Effect of BR on taking risk	$E_{bt} = \max(\varpi + \psi \times \bar{B}, 0)$
18. Developing S	$D_s = \frac{F}{\tau_2}$
19. Burning effect on vulnerability	$B_{ev} = B \times \sigma$
20. Fire risk perception	$\bar{B} = Smooth(B, \delta_1)$
21. Fractional development	$\theta = 0.4$
22. VP deterioration effect	$\mu = 0.2$
23. Risk intercept	$\varpi = 1$

24. Risk multiplier	$\psi = -8$
25. Fractional burning rate per ignition	$f_r = 0.08$
26. Average s burning	$\rho = 6$
27. Time to change behavior	$\delta_2 = 2$
28. Human ignition multiplier	$h_m = 5$
29. Time to grow vegetation	$\tau_1 = 2$
30. Time to develop S	$\tau_2 = 10$
31. Time to perceive	$\delta_1 = 0.5$
32. Initial strong vegetation	$S(t = 0) = 0.5$
33. Initial flammable vegetation	$F(t = 0) = 0.4$
34. Initial empty area	$E(t = 0) = 0.1$
35. Fire propagation	$\gamma_S = 0.8 * (1 + e^{-5*(\frac{\gamma_F F}{n} - 1)})^{-1}$
36. Normal burning	$n = 0.1$
37. Natural ignition	$I_N = 0.5$
38. Initial human ignition	$I_H(t = 0) = 0.3$
39. Initial vulnerable property	$V(t = 0) = 0.4$
40. BR multiplier	$\sigma = 0.05$

Simulation-based estimation of wildfire behavior across the United States

Abstract: Wildfire in the United States has increased over the past years. Data on the burnt area in the Western United States show an increasing trend with periodic fluctuations, especially in recent decades. We develop a system dynamics simulation model to study the trends of wildfires. The model is simulated and calibrated to wildfire's historical trend in states with high wildfire activity between 1992-2015. The model is a feedback-rich endogenous one that encompasses human contributions to fire through human ignition, affected by past fires and vulnerable property dynamics, as well as the change of flammable vegetation over time. We then check the effects of temperature and precipitation variation in wildfire behavior. The model suggests that both human and global warming contributes to the US's current wildfire activity, and the dominant drivers of wildfire are different in each state. The results provide state-specific guidelines for policymakers to allocate their resources.

1. Introduction

Over recent decades, wildfire activity in the United States has increased substantially and has imposed significant economic impacts on society (Westerling, Hidalgo et al. 2006, Richardson, Champ et al. 2012). While the changes in wildfire activities vary across the country, its increase is more dominant in the western areas (Westerling, Hidalgo et al. 2006). Figure 3.1 displays the United States annual burned area and the annual cost of suppression from 1984-2018. According to these figures, both variables increased substantially in this period. The suppression cost is only one of the economic impacts of the wildfire. The indirect costs of the wildfire include health problems from smoke exposure, the decline in tourism, death of animals and loss of vegetation, and of course, potential human death. For example, 85 civilians died in Campfire 2018 in California, 18804 buildings were burned, and the overall wildfire expense assessments are about $16.5 billion. In fact, the paradise city is nearly destroyed, and after one year of the blaze, the area is still experiencing a 90 percent population reduction. The Pacific Gas and Electric Company (PG&E) was blamed for the fire ignition, and this company's bankruptcy was only one of the economic effects of this catastrophe (Reyes-Velarde 2019). Hence, the increase in direct suppression cost is only one of the economic and fatal consequences of change in wildfire activity in recent years.

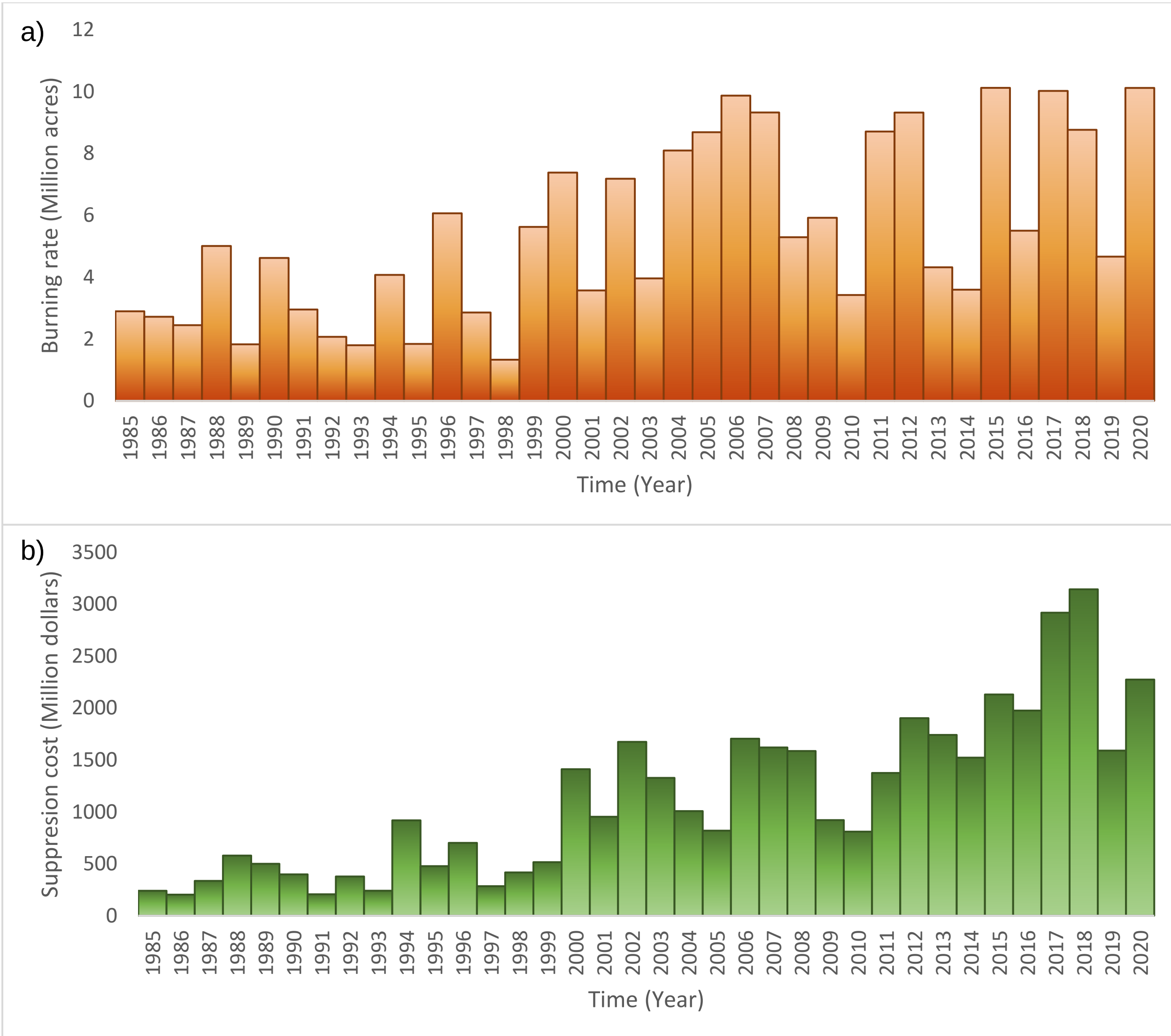


Figure 3.1. United States annual burning rate and suppression cost 1984-2020. (a) annual burned area; (b) annual suppression cost.

Any policy that aims to reduce the wildfire impact on society needs to proceed with a deep understanding of the causes of wildfire activity. However, the real causes and processes are still in doubt. Global warming and human activity are two leading candidates that experts in various fields believe to be the main cause of current fire behavior (Westerling, Hidalgo et al. 2006, Balch, Bradley et al. 2017). For example, Between 1984 and 2018, California's average annual temperature increased by 1.9 Fahrenheit degree (NOAA 2020). Furthermore, the California population also increased %49 at the same time, which corresponds to further population settlement near wildland and a further increase in human ignitions (Syphard, Radeloff et al. 2007, Balch, Bradley et al. 2017). Hence, several studies introduce both human and global warming as the dominant reasons for wildfire activity increase.

Human activity modifies fire regimes in two ways: To begin, there is compelling evidence that human-caused fires have increased fire activity across the United States in recent decades

(Syphard, Radeloff et al. 2007, Balch, Bradley et al. 2017). According to studies, human ignition accounted for about 84 percent of all ignitions in the United States between 1992 and 2012 (Balch, Bradley et al. 2017). Second, some researchers believe that rising fire activity is mostly due to human climate change(Abatzoglou and Williams 2016). Human activity contributes to global warming by releasing a variety of chemicals into the atmosphere, including carbon dioxide and methane, which are together referred to as greenhouse gases. These gases trap heat in the atmosphere and can result in a variety of impacts, including a rise in temperature and a decrease in precipitation (Fleming 1999). There is compelling evidence that human climate change has increased the duration, frequency, and magnitude of fire seasons in the western United States (Westerling, Hidalgo et al. 2006, Jolly, Cochrane et al. 2015, Abatzoglou and Williams 2016, Westerling 2016, Williams and Abatzoglou 2016).

There is abundant evidence that temperature and wildfire activity are related through two key pathways. First, a warmer environment results in drier forest vegetation. In the event of wildfire, this dryness results in vegetation to burns more rapidly, hence increasing the chance of a catastrophic wildfire.(Westerling, Hidalgo et al. 2006). Forest trees have a high moisture content, which decreases their flammability. However, when the temperature rises, the moisture content of forest trees decreases, making them more combustible. Second, when temperatures rise, lightning strikes more frequently, which is the natural source of wildfire ignition in the United States (Romps, Seeley et al. 2014). Therefore, a warmer environment results in more natural ignition and larger wildfires as a result of changes in lightning and vegetation dryness.

Despite the importance of direct human contributions and the global rising of temperature, most modeling studies have focused on one of these two categories of causes. On dynamics of vegetation, Touboul developed simulation models of interaction among the grass and other vegetations in forests (including forest trees), showing that, for a wide range of scenarios, occupied areas of each vegetation can oscillate over time (Touboul, Staver et al. 2018). The model helps inform natural-system dynamics. On human contributions, several statistical models have pointed to the correlation between human settlement in the wildland-urban interface (WUI) and fire activity (Pew and Larsen 2001, Vilar, Woolford et al. 2010, Matin, Chitale et al. 2017). We hypothesize that the natural and human systems, in interaction, make wildfire dynamics more complex and difficult to mitigate, and in order to develop proper policies, attention should be made to both sides of the larger system. Our primary objective is to examine the trends of wildfire given the importance of both human-direct contributions to fire through causes such as human-made ignitions and dynamics of vegetation and climate change.

2. Background: Models of disasters

While our focus is on a specific problem of wildfire, it is important to pause and offer a quick review of various modeling approaches from a methodological standpoint. There is a wide range of modeling approaches applied to natural disasters studies in general and wildfires in particular. They can be differentiated based on their unit of analysis, time frame, mathematical modeling technique, the model's boundary, and specific application cases.

A large body of the natural disaster body of the literature is devoted to spatial modeling (Davis, Wang et al. 2008, Keane, Drury et al. 2010, Finney, Grenfell et al. 2011). In a typical spatial model of wildfire, the goal is to replicate the progression of fire throughout different regions. Such models are powerful in showing how and in what sequence and timing different areas may become fire susceptible. Spatial models can also take different forms depending on the geographical units of analysis (e.g., state, county). Connection networks between different units can affect the fire progress, and such models become more useful as they move toward special network connections, other than random or full connections.

The second group of models of natural disasters includes agent-based models. Often agents are human, and in-interaction counter-intuitive dynamics should emerge. Models of evacuation often take such levels of analysis and look into the flows of individuals after a disaster (Chen and Zhan 2008, Yin, Murray-Tuite et al. 2014). In the wildfire context, with a focus on fire progression, agent-based models may consider units of vegetation as agents. Given that the interactions would limit to infection of fire from one unit of vegetation to another, such models will quickly become spatial representations of wildfire and similar to our first category. Rahmandad and Sterman (2008) stress the fact that in many contexts, detailed agent-based models may not go beyond what one could learn from an aggregate differential equation model, especially when the heterogeneities among the agents are limited, and connection networks are symmetric and almost complete.

In a slightly different natural disaster, there is a methodologically relevant work by Deegan (Deegan 2007). In his Ph.D. dissertation, Deegan (2007) modeled the dynamics of flood damage in a typical flood-prone community, considering long-term reactions of the community to recent floods and related damages. Deegan focuses on hypothetical cases of flood, with the objective of showing how seemingly similar external events (here, major rain), can cause different levels of damage depending on the community's reactions and investment in vulnerable properties. What differentiates his work from others is that Deegan's model is feedback-rich, and dynamics are arriving within the model rather than external time series. In his reflections on foundations of system dynamics, where Richardson (Richardson 2011) argues what differentiates a power model is its endogenous perspective to complex problems, he brings up Deegan's work as a powerful example and writes "Deegan's (2007) extensive analysis suggests … an endogenous view of the dynamics of flood damage that takes account of the human role in creating property vulnerable to flood damage… [He] traces the dynamics of vulnerability to the interacting actions of the capacity of the local environment to withstand floods, development pressure, property tax needs, perceived risks of development, moral hazard, policy entrepreneurs, and other people pressures" (Richardson 2011, p.234).

What makes these aggregate models powerful is that they are relatively small (have fewer equations), and taking away the details, they turn the focus on system responses and feedback loops, without losing many systems-level insights (Rahmandad and Sterman 2008). Having a small model, modelers can more easily communicate the outcomes and model structure with stakeholders, and make impacts (Ghaffarzadegan, Lyneis et al. 2011). It is important to note that small, powerful models are not easy to build; they are often results of many rounds of modeling and working with more complex models with more details (Ghaffarzadegan and Larson 2018). Given our problem scope and the focus of our work on state-level dynamics and data, we follow the same line of the modeling approach. In contrast to some of the earlier works, we utilize several sets of data to carefully calibrate our model and improve its relevance to the specific problem at hand (Sterman 2018). In the next section, more details about our modeling approach and data are offered.

3. Method

3.1. Modeling approach

We develop a system dynamics simulation (Richardson 1991, Sterman 2000) model of wildfire activity. The model can be applied to different contexts, while for the sake of model verification and validation, our focus was on the past decade of wildfire in the United States. We use historical data to estimate unknown parameters in our model and check the ability of the model to replicate the real-world trends. In light of the model that can explain the historical data of wildfire in recent years, we can conduct a what-if analysis and assess the role of all critical feedbacks in the fire behavior. Moreover, we evaluate the effects of different policies and scenarios on the wildfire extent.

The model is a feedback-rich model (Richardson 1991, Richardson 2011) with the goal of reproducing wildfire trends endogenously. Our model boundary includes human interactions with the natural system, vegetation dynamics, and human risk perceptions. As explained, climate variables (such as temperature, lighting, and precipitations) are included as exogenous variables in the model.

3.2. Data

We construct our dataset from multiple sources. The purpose of using data in our model is twofold. First, we use historical data to validate our model and estimate the parameters. Second, we use some data as exogenous variables and use their value directly in the model. Spatial wildfire occurrence data is collected by the United States department of agriculture (Short 2017). The data include 1.84 million fires with their location, cause, date, and other important information (From our point of view, this is the complete dataset of wildfires in the United States). The number of human ignitions, the number of natural ignitions, and the annual burning rate are three critical variables used in the modeling. We use the number of human ignitions and annual burning rate to calibrate our model. As stated, we use 12 month moving average of the number of natural ignitions as an exogenous variable in our model. Furthermore, the wildland-urban interface (WUI) area is also collected by the US department of agriculture (Radeloff, Kramer et al. 2017). We use the WUI area to calibrate our model to ensure our model can generate an increase in WUI area endogenously. We focus on the case of wildfire in the states with the highest wildfire activity in our study period. Between 1992-2015, 86.63% of the total burning rate in the conterminous United States happened in the following states: California (CA), (Colorado) CO, Idaho (ID), Montana (MT), New Mexico (NM), Nevada (NV), Oregon (OR), Utah (UT), Washington (WA) Wyoming (WY). For the rest of the paper, we aim to model the wildfire activity in these states.

The other essential data are climatic information. We consider two climatic variables for modeling the effect of climate on the model results: temperature and precipitation. These two variables affect wildfire behavior in different ways. Precipitation impedes the vegetation regrowing and decreases their dryness, while the fire season's temperature affects vegetation dryness (Westerling, Hidalgo et al. 2006). We consider 12 months moving average of temperature and precipitation to consider the climate effect in the model (data provide in

Appendix D). Statewide temperature and precipitation were obtained from the national oceanic and atmospheric administration (NOAA)[2].

3.3. Model sectors and key formulation

This study is built on a model proposed in my first essay (Farkhondehmaal, F., Ghaffarzadegan, N. 2022). The detailed formulation of all equations and the explanation for each variable are provided in Appendix B. Here we explain different model sectors and discuss how they are related to each other. The model includes two sectors, vegetation dynamic and human role in the wildfire.

Our model is based on both ecological and anthropogenic parts. We built on a simplified version of a dynamic model of savanna landscapes developed by Troublous et al. (Touboul, Staver et al. 2018). Figure 3.2 shows the main vegetation stock-flow structure of our model. Our model divides the forest area into three areas occupied by strong vegetation (S), occupied by flammable vegetation (F), and empty area, which are shown as stock variables in the figure (variables in boxes). Highly flammable vegetation includes damaged vegetation and any vegetation that can burn fast (including grass). These vegetations can burn quickly and distribute the fire.

Furthermore, the remaining damaged vegetation from previous fires is also considered in this category due to its dryness and high flammability. We assume that the empty area first fills with highly flammable vegetation (with a short lifespan) and then can be occupied with a forest tree. This logic is consistent with the previous studies of the interplay of forest trees and grass in forests (Touboul, Staver et al. 2018).

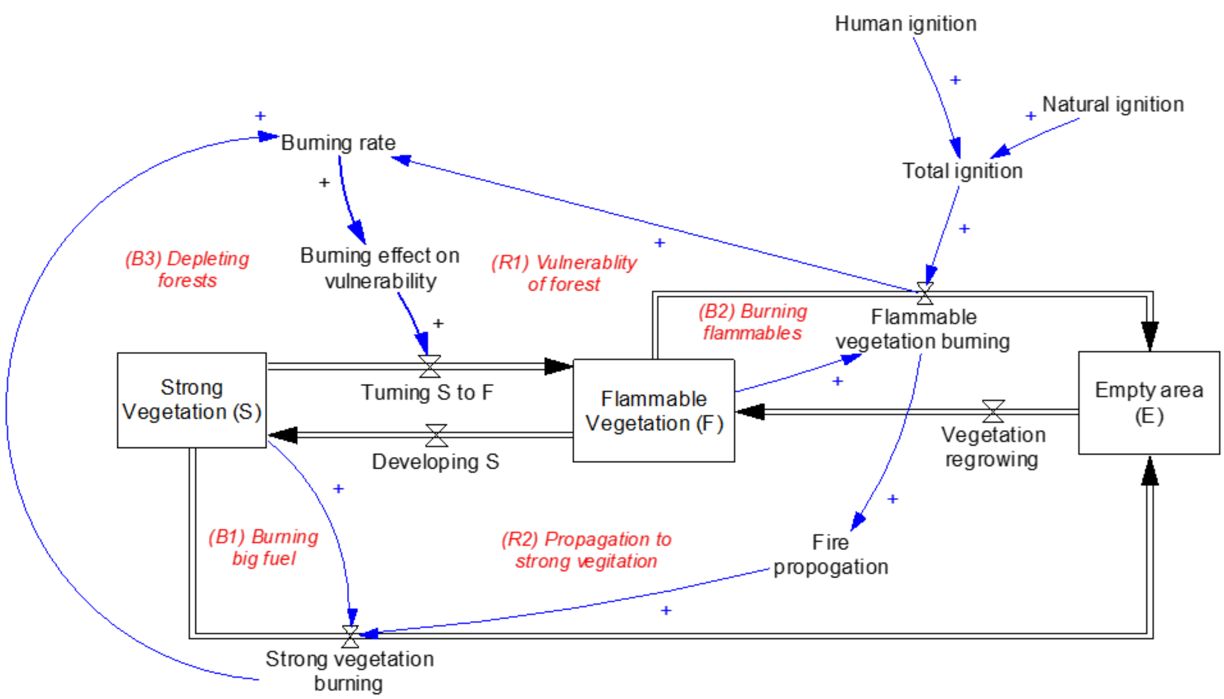


Figure 3.2. A stock-flow diagram of the vegetation section of the model

[2] https://www.ncdc.noaa.gov

As shown in the figure, there are different inflows and outflows between strong and flammable vegetation. Flammable vegetation is the primary vegetation type that can be burned in a wildfire. If the wildfire size exceeds specific criteria, then the fire is big enough to burn strong vegetation. Furthermore, strong vegetation can turn to flammable vegetation and become susceptible to the fire by proximity to wildfire heat and smoke. Wildfire damage vegetations and make them prone to burning in case of future fires. The empty area first fills with flammable vegetation by the natural aging chain process, and then it grows and converts to strong vegetation.

The fire burns the flammable vegetation, which causes less flammable vegetation for future fires since it takes a time (D_g) for the empty areas to be filled with vegetation. Hence, this loop also limits the available fuel for fires and causes a decrease in the burn area in the future. There are also two natural conversions in the vegetation stocks. Burned area changes to flammable vegetation over time due to the growth of new vegetation in it. Furthermore, flammable vegetation also changes to strong vegetation as they grow. We consider that the empty areas will be first filled with grass (which is flammable vegetation), and then the grass will substitute with forest trees that are more resilient to burn. The structure shown in figure 3.2 is mathematically equivalent to the following equations:

$$F = \int_0^t (F_i + S_o - F_o - S_i)dt + I_F \times f$$

$$S = \int_0^t (S_i - S_o + S_b)dt + I_S \times f$$

where F_i, F_o are vegetation regrowing and flammable burning rate, S_i, S_o, S_b are developing strong vegetation, turning S to F and strong vegetation burning, respectively.

Another major sector of the paper is the human sub-system which includes the human perception of risks and their associated behavior. As shown in figure 3.3, this section is connected to the vegetation sub-system. BR perceived is what connects burning to the human sector. It is defined as a dimensionless attribute that is influenced by the annual area burned (B_rPeople's perception is expected to change depending on changes in burning rate. In turn, it is the driving force of their actions. In other words, the more they consider the chance of wildfire, the less they set fires by reducing their high-risk actions. Such high-risk activities specify the annual human ignition in our model. And the human ignition directly contributes to the annual fire area, closing a feedback loop. This balancing mechanism is one of the driving forces behind the reduction of the area burned because of the rise in wildfire activity. In having a high burn area, people see a greater risk of fire in the years ahead, and this change in their perception triggers a reduction in future burns due to a decrease in risky human activities. This loop is often ignored in past studies. To formulate, we consider simple linear regression to estimate the effect of burning rate on taking a risk, based on perceived burning rate:

$$b = P_m * B_r + I_p$$

where P_m is perception adjustment multiplier and I_p is the initial value of perception.

In addition to the described feedback loops, there are some exogenous variables in our model. Those variables are ones that we do not intend to predict their changes over time in our model due to the scope of our model, and instead, we take time-series inputs. They include natural ignition (lightening), temperature, and precipitation. Natural ignition feeds to total ignition in our model adding to human ignition, which was modeled endogenously. Temperature and precipitation both determine the severity of drought in the state. So, higher temperatures and lower precipitation both increase the drought severity. The drought also directly affects the burning rate. Higher drought leads to an easier distribution of wildfire.

Model equations and parameter values are documented in Appendix B and C.

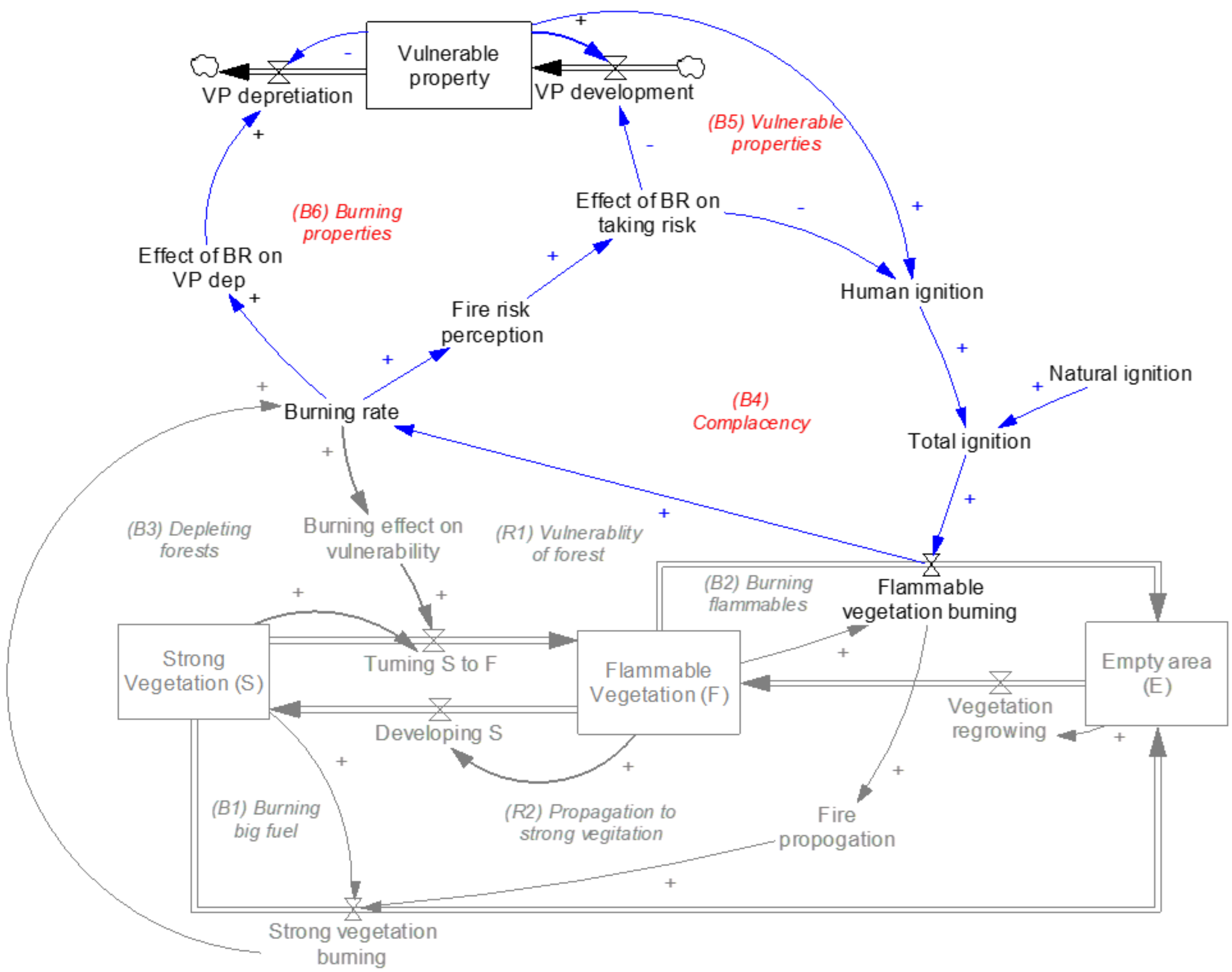


Figure 3.3. The stock-flow model structure of the human part. Humans contribute to fire by ignition. Both perceptions of wildfire risk and settlement in WUI are important factors in determining the role of people in the wildfire.

3.4. Parameter estimation

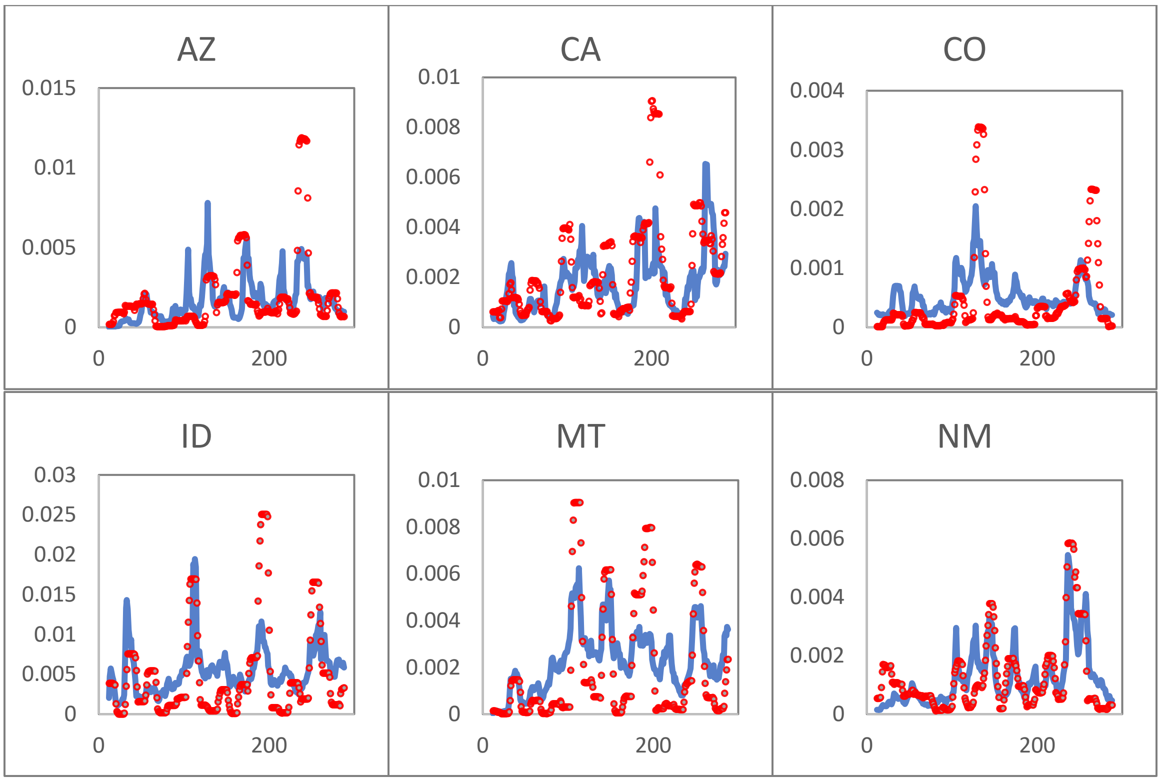
Model parameters are estimated based on state-specific information. We use the burning rate, WUI area, and the number of human fire ignition to calibrate our model for each state. Specifically, we are interested in parameters that defined people's reactions to perceived risk and the time it takes for them to change their behavior. Model calibrations perform by minimizing a specified objective function. The detailed formulation of the objective function is provided in the Appendix C. Building confidence in parameters is an important issue for model credibility.

4. Results

Figure 3.4 shows the simulation result and the data for 12 months moving average of burning rate per acres for 11 states. Figure 3.5 shows the number of human ignitions in these states per acre of forest. Both burning rate and human ignition show fluctuation on the 12-month moving average. We chose 12 months moving average to make sure the model's fluctuation behavior is not due to seasonal changes in precipitation and temperature. Our model shows that the fluctuation stems from the interaction of the human, natural, and exogenous conditions, which are the number of lightning and climate factors.

Table 3.1 shows how well the model replicates the real-world data. We used R-squared (the square of the correlation between predicted values and actual data) to compare the fitting quality. Comparing the table 3.1 and figure 3.4 and 3.5 shows that for states which experience rare extreme instances of high burning rate or human ignition, the model replication quality decrease compared to other states without such trends. The purpose of our model is to show the overall trend of wildfire behavior simplistically. Overall, the model could successfully replicate the wildfire dynamic in all states with some error due to rare extreme events.

Our model's main purpose is to use the mode as a prototype for comparing the effect of different policies and scenarios on wildfire behavior. First, we check the effect of the climate factors on wildfire activity. For doing that, we compare the total burning rate in the study period with the situation in which both temperature and precipitation remain constant (equal to their average value between 1900-2000). We then check the effect of three proposed policies on wildfire reduction. These policies limit vulnerable property development, Prescribed burning, and Effective firefighting. The first policy is implemented by making vulnerable property development smaller than 1% of the current value of the vulnerable property. The second policy is implemented by adding outflow from flammable vegetation to an empty area with the value of ωF, where ω is the percent of prescribed burning set at 0.2/year. Finally, the last policy is implemented by changing γ_S to 10% of its current value each year.


AZ
0.015
0.01
0.005
0
0
200
CA
0.01
0.008
0.006
0.004
0.002
0
0
200
CO
0.004
0.003
0.002
0.001
0
0
200
ID
0.03
0.025
0.02
0.015
0.01
0.005
0
0
200
MT
0.01
0.008
0.006
0.004
0.002
0
0
200
NM
0.008
0.006
0.004
0.002
0
0
200

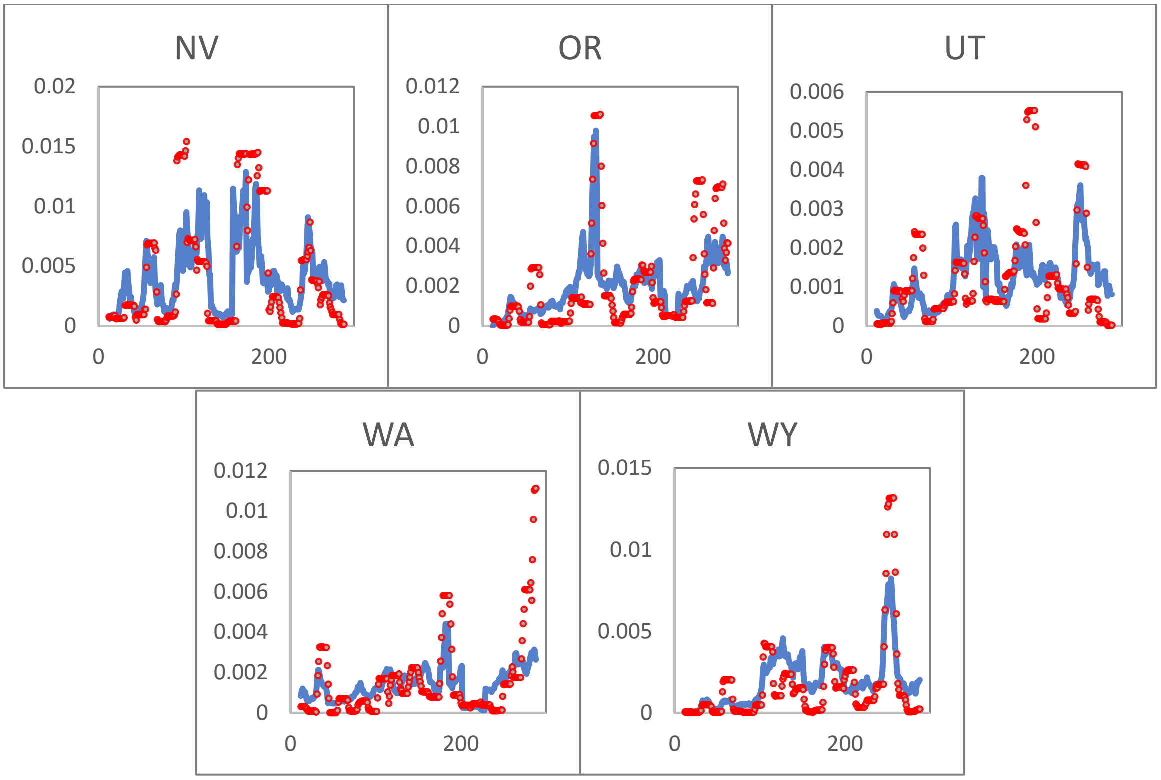


Figure 3.4- The 12 month moving average of burning rate per acre of forest for 11 states with the highest wildfire activity. The x axis represents time (month number). Red dots show the data points and blue line shows simulation results.

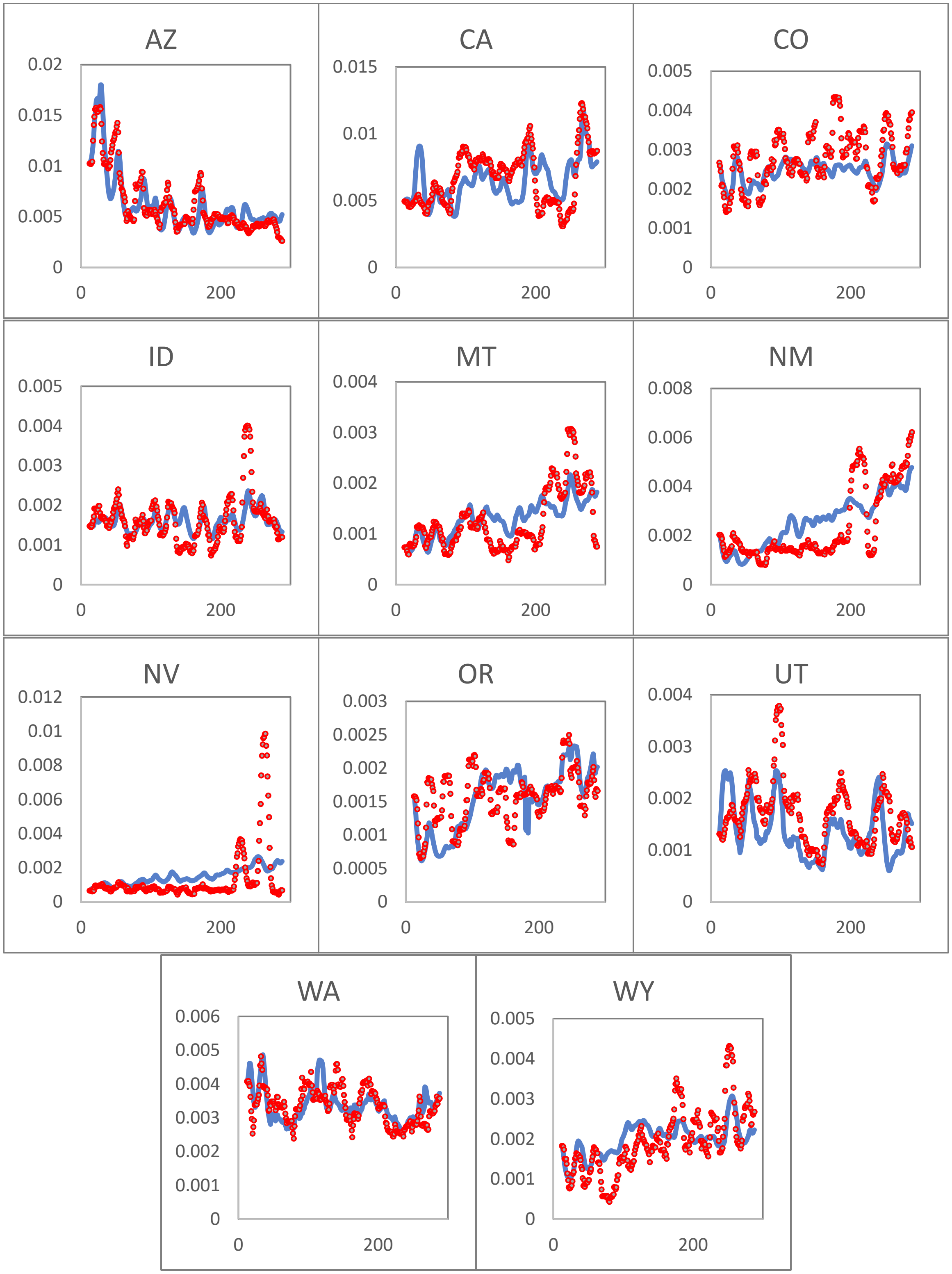


Figure 3.5- The 12 month moving average of human ignition per acre of forest area for states with the highest wildfire activity. The x axis represents time (month number). Red dots show the data points and blue line shows simulation results

Table 3.1- Comparison of R-squared for different states.

State	Burning rate	Human Ignition
AZ	0.34	0.73
CA	0.28	0.22
CO	0.30	0.16
ID	0.43	0.25
MT	0.61	0.41
NM	0.61	0.45
NV	0.46	0.32
OR	0.36	0.35
UT	0.25	0.46
WA	0.43	0.57
WY	0.68	0.48

4.1. Scenario and Policy Analysis

Humans contribute mainly to wildfire by ignition. The human perception of fire risk and settlement in WUI affects the number of wildfires people ignite annually. Here we test how the human total burning rate reduce if there is no WUI settlement. Furthermore, we test what happens if states do not experience any climate change from the 1900-2000 situation. The result in table 3.2 shows that while some states experience a huge reduction in the burning rate without a change in climate, this factor caused a small reduction in the burning rate for some other states, including Arizona and Idaho. This is primarily because Western states experienced a more severe climate change in recent decades compared to some other regions. Furthermore, we can compare the impact of different policies on wildfire treatment. For instance, if the firefighting effectiveness with the presumed amount is achievable, this policy is the most effective in all states. However, the same cannot be said about the other two policies. For example, in some states, including Idaho and New Mexico, the first policy seems to be more effective than the second one.

Table 3.2- Scenario and Policy effects on the burning rate reduction

State	WUI	Prescribed Burning	Firefighting	Climate
AZ	1%	20%	24%	5%
CA	22%	12%	45%	32%
CO	6%	21%	46%	36%
ID	26%	13%	52%	12%
MT	7%	21%	40%	65%
NM	39%	32%	37%	43%
NV	20%	12%	21%	45%
OR	12%	16%	36%	28%

UT	14%	10%	32%	23%
WA	9%	8%	21%	31%
WY	10%	20%	65%	41%

5. Discussions and Conclusion

In this study, we developed a system dynamics model of wildfire behavior, which considers wildfire as an outcome of a human-natural coupled system. The model has three main mechanisms that capture endogenous changes in wildfire activities. They include vegetation dynamics, human perception as affected by recent fires, and settlement in WUI. We used data from wildfire in the United States to calibrate our model for 11 states and model each of them separately. We estimated the model's unknown parameters using the calibration method. The model fairly replicated the data for the study period and provided an acceptable prediction for the 1992-2015 time period. We used a panel of the dataset that included human ignition, burning rate, WUI area, natural ignition, temperature, and precipitation. Several parameters, such as the distribution of vegetation and the size of vegetation, were estimated through archival data or publications. The unknown parameters are estimated through the calibration method.

Our simulation results uncover the importance of human and climate effects in different states. We particularly find that there is no single reason for wildfire activity increase in recent years, unlike the majority of the literature. In fact, global warming, and human activities both cause the current situation. Global warming causes an increase in flammable vegetation, which contributes to the higher annual burning area, while human activities lead to more human ignition, which increases fire activity.

This study makes several contributions to the literature. By considering the current literature regarding wildfires and the impact of different variables on their behavior, this research found important feedbacks that are missed in the studies of the impact of humans on wildfires. We considered the human influence on wildfire behavior by considering the perception of humans regarding wildfire risk and its contribution to the number of annual human ignitions. Finally, we developed a dynamic model of forest-human interaction to understand structures that lead to the current fire situation in burned areas in the conterminous United States. The model shows that the burning area behavior depends on natural and anthropogenic feedback.

The study has several policy implications. Different policies can be used to control the number of risky human behaviors. Our model hypothesized that human perception causes people to change their high-risk activities. An increase in the perception of fire danger causes people to be more careful when performing high-risk activities. Furthermore, they will follow safety procedures that aim to prevent wildfire ignition. On the other hand, governments can reduce the effect of humans by forbidding high-risk behavior when there is a high probability of fire ignition by performing them. They also can mandate safety procedures to be followed. These policies can reduce the amount of high-risk behavior, which can cause a reduction in the burned area.

One of the limitations of this study is the assumption that people's perception and effect of vulnerable property dynamics are the only factors affecting the number of human ignitions. All human ignition fires are not caused by these two factors. For example, ignition due to work with heavy machinery inside the forests is another type of human ignition fire. Hence, the total number of people who work in forests also affects the number of human ignitions. In that case, other policies, including reducing total workers in the forest area or increasing the training workshop for them could be some valuable policies in the hope of reduction in human ignitions.

The purpose of this study was to compare the effect of human activity and climate change on each state and provide a guideline for authorities in policy implementation. As shown in this study, each state's policy should rely on the main reason for change in wildfire activity. Previous studies investigating the human impact on wildfire activity rely solely on variables that remain almost constant over a short period of time. While these studies can compare the role of humans in wildfires in different areas, they cannot detect any significant alterations in the human impact on fire in a specific area. We have argued that people's perception of fire and their risky behavior are critical factors that are changing in the same place over time. We have also adopted wildfire as a hybrid human-environmental framework, which stresses the need to consider the role of both human and natural processes in defining wildfire behavior.

References

Abatzoglou, J. T. and A. P. Williams (2016). "Impact of anthropogenic climate change on wildfire across western US forests." Proceedings of the National Academy of Sciences **113**(42): 11770-11775.

Balch, J. K., et al. (2017). "Human-started wildfires expand the fire niche across the United States." Proceedings of the National Academy of Sciences **114**(11): 2946-2951.

Chen, X. and F. B. Zhan (2008). "Agent-based modelling and simulation of urban evacuation: relative effectiveness of simultaneous and staged evacuation strategies." Journal of Operational Research Society **59**(1): 25-33.

Davis, C., et al. (2008). "Prediction of landfalling hurricanes with the advanced hurricane WRF model." Monthly Weather Review **136**(6): 1990-2005.

Deegan, M. A. (2007). Exploring US flood mitigation policies: a feedback view of system behavior, Rockefeller College of Public Affairs and Policy, University at Albany–SUNY.

Farkhondehmaal, F., Ghaffarzadegan, N. 2022. A Cyclical Wildfire Pattern as the Outcome of a Coupled Human Natural System. Virginia Tech. Blacksburg, VA

Finney, M. A., et al. (2011). "A method for ensemble wildland fire simulation." Environmental Modeling Assessment **16**(2): 153-167.

Fleming, J. R. (1999). "Joseph Fourier, the 'greenhouse effect', and the quest for a universal theory of terrestrial temperatures." Endeavour **23**(2): 72-75.

Ghaffarzadegan, N. and R. C. Larson (2018). "SD meets OR: a new synergy to address policy problems." System Dynamics Review **34**(1-2): 327-353.

Ghaffarzadegan, N., et al. (2011). "How small system dynamics models can help the public policy process." System Dynamics Review **27**(1): 22-44.

John, S. (2000). "Business dynamics: systems thinking and modeling for a complex world." Irwin McGrawHill.

Jolly, W. M., et al. (2015). "Climate-induced variations in global wildfire danger from 1979 to 2013." Nature communications **6**: 7537.

Keane, R. E., et al. (2010). "A method for mapping fire hazard and risk across multiple scales and its application in fire management." Ecological Modelling **221**(1): 2-18.

Matin, M. A., et al. (2017). "Understanding forest fire patterns and risk in Nepal using remote sensing, geographic information system and historical fire data." International journal of wildland fire **26**(4): 276-286.

NOAA (2020). "Climate at a Glance: Statewide Time Series, published January 2020, retrieved on January 24, 2020 from https://www.ncdc.noaa.gov/cag/."

Pew, K. and C. Larsen (2001). "GIS analysis of spatial and temporal patterns of human-caused wildfires in the temperate rain forest of Vancouver Island, Canada." Forest ecology management **140**(1): 1-18.

Radeloff, V. C. H., et al. (2017). "The 1990-2010 wildland-urban interface of the conterminous United States-geospatial data."

Rahmandad, H. and J. Sterman (2008). "Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models." Management Science **54**(5): 998-1014.

Reyes-Velarde, A. (2019). "California's Camp fire was the costliest global disaster last year, insurance report shows". Los Angeles Times.

Richardson, G. P. (1991). Feedback thought in social science and systems theory, University of Pennsylvania.

Richardson, G. P. (2011). "Reflections on the foundations of system dynamics." System Dynamics Review **27**(3): 219-243.

Richardson, L. A., et al. (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.

Romps, D. M., et al. (2014). "Projected increase in lightning strikes in the United States due to global warming." Science **346**(6211): 851-854.

Short, K. C. (2017). "Spatial wildfire occurrence data for the United States, 1992-2015 [FPA_FOD_20170508]."

Sterman, J. (2018). "System dynamics at sixty: the path forward." System Dynamics Review **34**(1-2): 5-47.

Syphard, A. D., et al. (2007). "Human influence on California fire regimes." Ecological applications **17**(5): 1388-1402.

Touboul, J. D., et al. (2018). "On the complex dynamics of savanna landscapes." Proceedings of the National Academy of Sciences **115**(7): E1336-E1345.

Vilar, L., et al. (2010). "A model for predicting human-caused wildfire occurrence in the region of Madrid, Spain." International journal of wildland fire **19**(3): 325-337.

Westerling, A. L. (2016). "Increasing western US forest wildfire activity: sensitivity to changes in the timing of spring." Phil. Trans. R. Soc. B **371**(1696): 20150178.

Westerling, A. L., et al. (2006). "Warming and earlier spring increase western US forest wildfire activity." science **313**(5789): 940-943.

Williams, A. P. and J. T. Abatzoglou (2016). "Recent advances and remaining uncertainties in resolving past and future climate effects on global fire activity." Current Climate Change Reports **2**(1): 1-14.

Yin, W., et al. (2014). "An agent-based modeling system for travel demand simulation for hurricane evacuation." Transportation research part C: emerging technologies **42**: 44-59.

Appendix B. Model formulation

Table A3.1 shows the full set of equations for the simulated model. In any program, these equations can be used to build a model. Nevertheless, these are used primarily to build a model in the Vensim DSS program. The Vensim DSS platform is an upgraded version of the Vensim PLE program that is available free for academic purposes. The DSS edition has a handful more options including optimization and sensitivity analysis. The model can therefore be developed using the Vensim PLE version with the following equations. The time stage for all simulations is 0.0078125 years. The model is run for 20 years but can be run for any other period.

Table A3.1. Complete set of model's equations, definition of parameters and units

Parameter name	Parameter equation and definition and units of its parameters
1. Vulnerable property	$V = \int_0^t ((\theta V)e^{-k\bar{B}} - \rho V)dt + I_V$
2. Indicated human ignition	$i_H = V \times E_{bt} \times h_m$
3. Human ignition	I_H =Third order delay of i_H; delay duration = δ_2; $I_H(t = 0) = 0.3$
4. Burning rate	$B = \gamma_F F + \gamma_S S$
5. Strong vegetation	$S = \int_0^t \left(\frac{F}{\tau_1} - (\alpha + \gamma_S)S\right) dt + I_s$
6. Flammable vegetation	$F = \int_0^t (\frac{E}{\tau_2} + \alpha S - (\frac{1}{\tau_1} + \gamma_F)F)dt + I_F$
7. Empty area	$E = \int_0^t \left(-\frac{dF}{dt} - \frac{dS}{dt}\right) dt + I_E$
8. Vulnerable property development	$V_i = E_{bt} \times \vartheta \times V$
9. VP depreciation	$V_o = V \times E_{bv}$
10. Vegetation regrowing	$V_r = \frac{E}{\tau_1}$
11. Turning S to F	$T_{sf} = \sigma \times B \times S$
12. Total ignition	$T_i = I_H + I_N$

13. Strong vegetation burning	$S_{br} = \rho \times F_p \times S$
14. Flammable vegetation burning	$\gamma_F = f_r \times F \times T_i$
15. Fire propagation	$F_p = 0.8 * (1 + e^{-5*(\frac{\gamma_F F}{n} - 1)})^{-1}$
16. Effect of BR on VP dep	$E_{bv} = B \times \mu$
17. Effect of BR on taking risk	$E_{bt} = \max(\varpi + \psi \times \bar{B}, 0)$
18. Developing S	$D_s = \frac{F}{\tau_2}$
19. Burning effect on vulnerability	$B_{ev} = B \times \sigma$
20. Fire risk perception	$\bar{B} = Smooth(B, \delta_1)$
21. VP deterioration effect	$\mu = 0.2$
22. Fractional burning rate per ignition	$f_r * Drought\ index$
23. Average s burning	$\rho * Drought\ index$
24. Time to change behavior	$\delta_2 = 2$
25. Time to grow vegetation	$\tau_1 = 2$
26. Time to develop S	$\tau_2 = 10$
27. Time to perceive	$\delta_1 = 2$
28. Fire propagation	$\gamma_S = 0.8 * (1 + e^{-5*(\frac{\gamma_F F}{n} - 1)})^{-1}$
29. BR multiplier	$\sigma = 0.05$

30. Drought index	$\frac{Temperature}{Precipitation}$

Appendix C. Parameter estimation

Parameter estimation is done by minimizing the weighted sum of squared error between data and model outcome (for each dataset). We consider the RMSE (Root mean square error) as the weight for the optimization problem of parameter estimating:

$$min \sum_{i=0}^{n} \frac{\sum_{t=0}^{T}(\hat{\theta} - \theta)}{RMSE(\hat{\theta})}$$

Which $\hat{\theta}$ is the model result, and θ is the historical data.

Appendix D. The Exogenous data

Figures A3.1, A3.2, and A3.3 shows the 12 months moving average of the corresponding data.

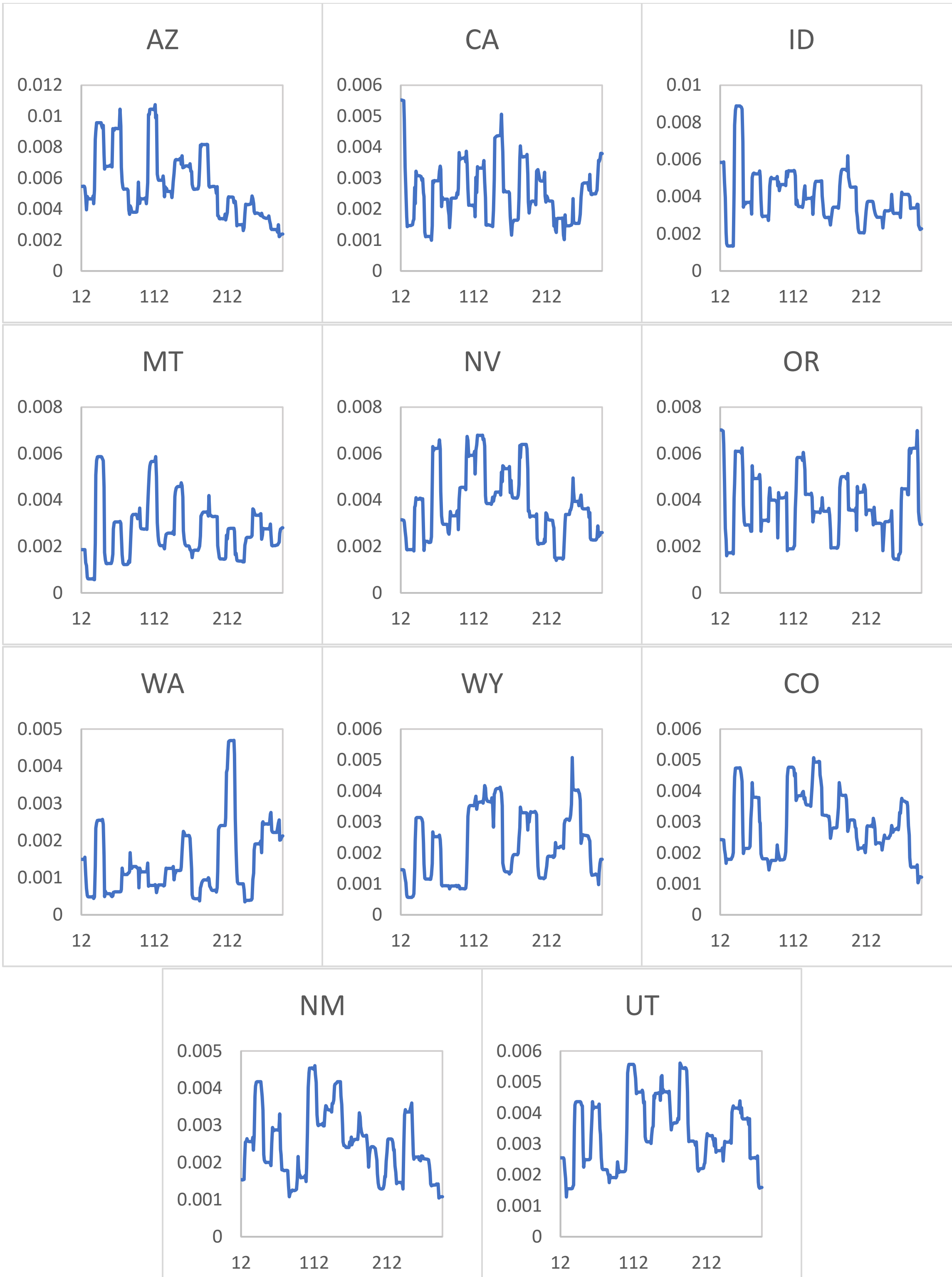


Figure 3.A1. 12 month moving average natural ignition per acre of forest for different states. The x axis shows the time (month number) and the y axis shows the number of lightning per acre of forest

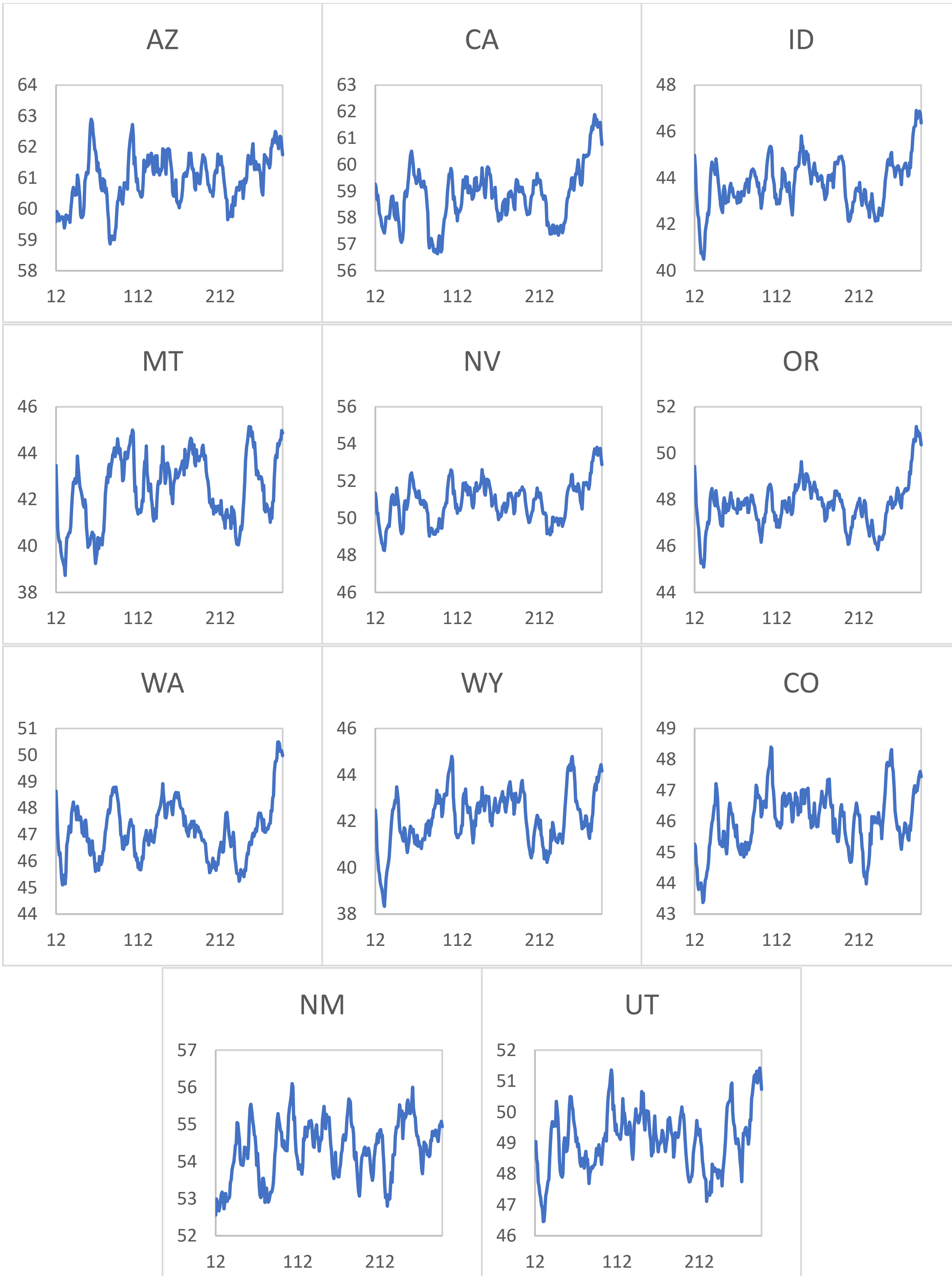


Figure 3.A2. 12 month moving average temperature for different states. The x axis shows the time (month number) and the y axis shows the state temperature

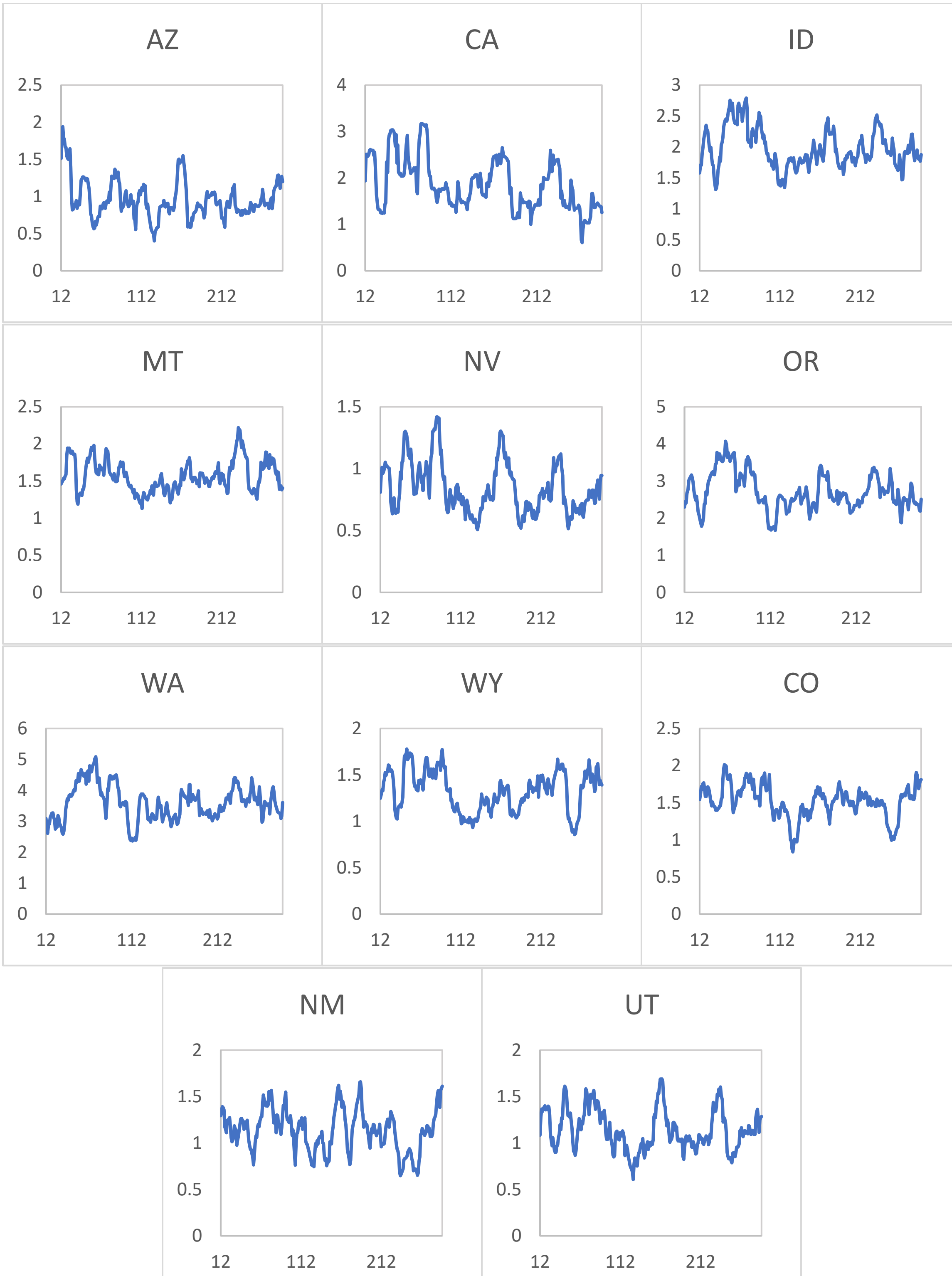


Figure 3.A3. 12 month moving average precipitation for different states. The x axis shows the time (month number) and the y axis shows the state precipitation.

Chapter 4: Preparing for wildfire using wildfire prediction for the Oregon counties

Abstract

Recent increases in wildfire occurrence in the United States, especially the number of big fires, defined as 500 acres or greater, have resulted in substantial environmental, economic, and human damage. Predicting the occurrence of big fires at a particular location enables fire management to plan for suppression and focus resources on high-profile locations. This study uses 25 years of wildfire data (1992–2015) to construct a hierarchical model for determining the probability of big monthly wildfires instances in Oregon counties. We use an interactive model with meteorological and socioeconomic variables that can better predict fire behavior in the areas where humans exist. The output of our model is then compared to other machine learning algorithms, including Deep Learning, Extreme Graduate Boosting, and Random Forest. The model produces comparable findings for precision and sensitivity (Area under the precision-recall curve of 0.502) and has the advantage of showing each variable's importance compared to ensemble methods.

1. Introduction

Wildfire is a complex phenomenon that can endanger natural resources, the environment, and human life (Scott et al., 2013; McFarlane et al., 2012; Radeloff et al., 2018). Therefore, understanding and predicting wildfire occurrences is vital for analyzing fire danger in an area (Catry et al., 2010; De Vasconcelos et al., 2001; Miyanishi, 2001; Bonazountas et al., 2005). Fire occurrence density in a region is a dynamic outcome of two different mechanisms. First, human ignited wildfires account for approximately 84% of wildfires nationwide (Balch et al., 2017). The extent to which humans ignite wildfires depends on an individual's actions, including an abandoned campfire, arson, and fireworks (Vega et al., 2017; Prestemon and Butry, 2008; Bartlein et al., 2003). Second, lightning is another source of wildfire ignition, and its occurrence depends on weather conditions (Loeb, 1949). In terms of the area burned, the primary threat posed by wildfire is the high expense and destruction caused by big fires (Preisler and Westerling, 2007).

The importance of large wildfire prediction is the interest of different researchers (Rolinski et al., 2016). According to the empirical study, the number of significant wildfires in the United States has grown recently (Dennison et al., 2014). Figure 4.1 illustrates the annual number of big wildfires (larger than 500 acres) in the United States between 1992 and 2015 (Short, 2017). As illustrated in the figure, the number of big wildfires is on the rise. Over 1,880,465 fires destroyed nearly 140 million acres in the United States during this period. While only 0.925% of fires consumed more than 500 acres, they account for 88.4% of total burning rates.

Lots of research has focused on the large wildfire occurrence and magnitude. For example, the Santa Ana index aims to predict the wildfire occurrence exceeding 100 ha in six days (Rolinski et al., 2016). Another study evaluates the driving forces in shaping large wildfire extend in southern Chile (McWethy et al., 2018). An important question in modeling the large wildfire is the factors that lead to this extreme event (Preisler and Westerling, 2007; Preisler et al., 2016; Preisler et al., 2011; Gudmundsson et al., 2014; Podschwit et al., 2018; Westerling, 2018).

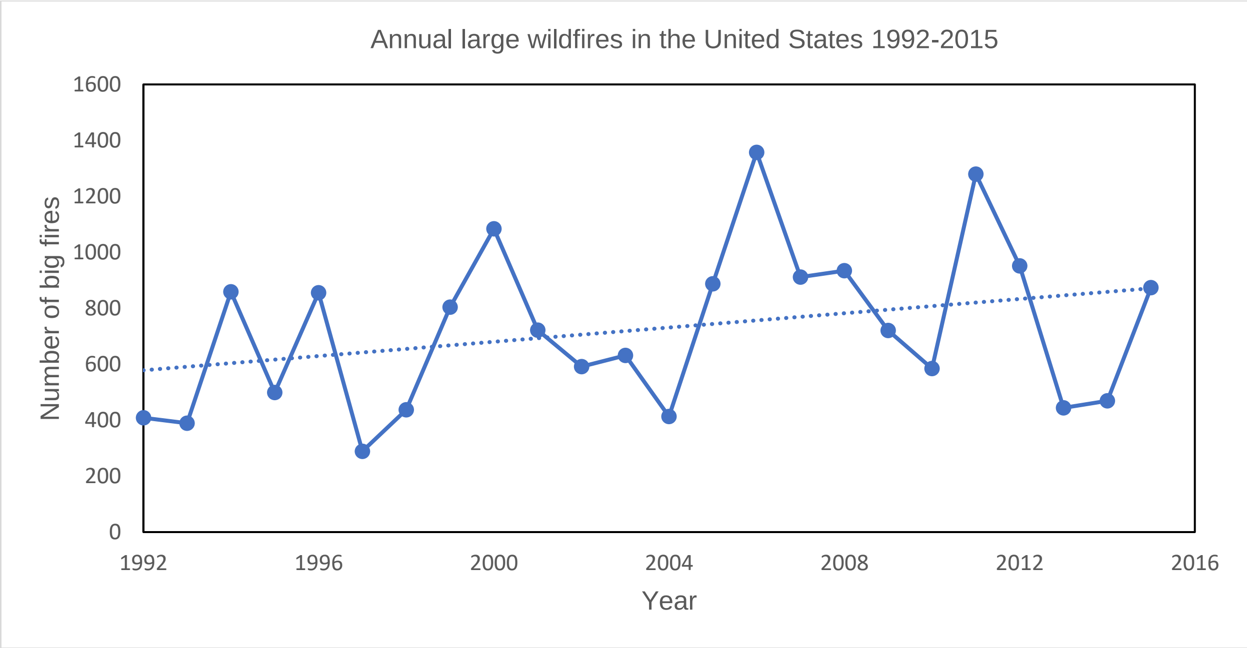


Figure 4.1- Annual number of large wildfires (greater than 500 acres) in the United States. The figure shows that large wildfires have been increasing in recent years[3].

There are two factors, environmental and human, that determine the wildfire behavior in any area. In the case of climate, there is ample evidence that there is an association between climate and wildfire activity through two primary mechanisms. First, a warmer climate leads to drier vegetation in forests. This dryness means that vegetation burns quickly in the case of wildfire and increases the risk of massive wildfire (Westerling et al., 2006). This is because forest trees contain lots of moisture, which reduces their flammability. However, the increase in temperature reduces the moisture content in forest trees and makes them easily flammable. Second, the increase in temperature leads to an increase in lightning, which is the natural cause of wildfire ignition in the United States (Romps et al., 2014). Hence, the warmer climate leads to more natural ignition and more significant wildfires due to lightning and vegetation dryness changes.

Humans alter fire regimes in two ways: First, there is strong evidence that in recent decades, human-ignited fires have expanded the fire activity across the United States (Balch et al., 2017; Syphard et al., 2007). Studies show that human ignition is responsible for almost 84% of the total number of ignitions between 1992-2012 in the United States (Balch et al., 2017). Second, some researchers believe that anthropogenic climate change is the main reason for increased fire activity (Abatzoglou and Williams, 2016). Human activity causes global warming by releasing different gases into the atmosphere, including carbon dioxide and methane, also known as greenhouse gases. These gases trap the heat in the atmosphere and can cause different effects such as temperature increase and rainfall decrease (Fleming, 1999). There is strong evidence that anthropogenic climate change increased the fire season duration, frequency, and size (Abatzoglou and Williams, 2016; Westerling et al., 2006; Westerling, 2016; Williams and Abatzoglou, 2016; Jolly et al., 2015).

[3] Data from (Short, 2017)

Typically modeling the human role in the wildfire is the interest of qualitative research or quantitative with simple statistical models, since integrating the human role in equations and variables seems problematic (Steelman and Kunkel, 2004). For example, one study in central Chile suggests that population density greatly influences the probability of annual ignition and burned area (McWethy et al., 2018). Another study on temperate[4] forests shows that humans specifically control wildfire in places with low lightning and high vegetation density (McWethy et al., 2013). There is, however, a consent in the literature that wildfire is the coupled human-natural system that both social and ecological conditions shape the people’s perception of how communities respond to this extreme event (Fischer et al., 2016; Cutter and Finch, 2008; Nowell et al., 2018)

Several studies suggest that the social aspect of wildfire can be modeled by the socioeconomic characteristics of each location (Charnley et al., 2015; Bühler et al., 2013; Romero-Calcerrada et al., 2008). These indicators can show the population's general proclivity to wildfire contribution (both in the ignition and burning area), and they are more quantifiable than an individual's subjective assessment of their potential contribution to wildfire (de Torres Curth et al., 2012). For example, according to one study on Florida's WUI, regions with fewer people and higher unemployment result in fewer fire ignitions and less burned land (Mercer and Prestemon, 2005). There is, however, no agreement on the relationship between socioeconomic variables and fire ignition and magnitude, as their effect can vary considerably across regions (Bühler et al., 2013). Table 4.1 summarizes some of the key variables used to model wildfires in various regions. Furthermore, since the number of human ignitions is highly correlated with environmental conditions, there is a need for socioeconomic and weather variables to model human-caused fires (Vilar et al., 2010). These can assist policymakers in identifying potential high-risk areas where risk mitigation efforts should be concentrated (Wigtil et al., 2016; Ojerio et al., 2011). However, due to the slow change in the socioeconomic factors, they fail to model the change in the number of human ignition over a short time.

[4] Tasmania, Northwestern USA, Southern South America, New Zealand

Table 2.1- The socioeconomic variables used for wildfire modeling in different areas.

Socioeconomic variable	Region	Model*	Example
Population	Florida	Ignition, Area	(Mercer and Prestemon, 2005)
	Patagonia	Ignition, Area	(de Torres Curth et al., 2012)
	Northwestern United States	Risk reduction	(Charnley et al., 2015)
	Argentina	Ignition	(Bühler et al., 2013)
	Spain	Ignition	(Romero-Calcerrada et al., 2008)
	United States	Fire hazard	(Wigtil et al., 2016)
Unemployment	Florida	Ignition, Area	(Mercer and Prestemon, 2005)
	Patagonia	Ignition, Area	(de Torres Curth et al., 2012)
	Northwestern United States	Risk reduction	(Charnley et al., 2015; Bühler et al.,
	Argentina	Ignition	2013)
	Arizona	Risk reduction	(Ojerio et al., 2011)
	United States	Fire hazard	(Wigtil et al., 2016)
Poverty	Patagonia	Ignition, Area	(de Torres Curth et al., 2012)
	Argentina	Ignition	(Bühler et al., 2013)
	Arizona	Risk reduction	(Ojerio et al., 2011),
Education	Patagonia	Ignition, Area	(de Torres Curth et al., 2012)
	Argentina	Ignition	(Bühler et al., 2013)
	Arizona	Risk reduction	(Ojerio et al., 2011)
Housing density	Florida	Ignition, Area	(Mercer and Prestemon, 2005)
	Patagonia	Ignition, Area	(de Torres Curth et al., 2012)
	Northwestern United States	Risk reduction	(Charnley et al., 2015)
	Spain	Ignition	(Romero-Calcerrada et al., 2008)
	United States	Fire hazard	(Wigtil et al., 2016)
Ownership	Midwest United States	Ignition, area	(Cardille et al., 2001)
Police density	Florida	Ignition, area	(Mercer and Prestemon, 2005)
Distance to infrastructure	Spain	Ignition	(Romero-Calcerrada et al., 2008)
Farming density	Spain	Ignition	(Romero-Calcerrada et al., 2008)
Racial distribution	Arizona	Risk reduction	(Ojerio et al., 2011)
	United States	Fire hazard	(Wigtil et al., 2016)

*Ignition: predicting fire ignition or number of fire ignition, Area: predicting the burned area

Recently ML methods used in wildfire modeling have grown significantly following advancements in computing and data collection (Jain et al., 2020). The scope of using ML methods range from predicting burning rate (Stojanova et al., 2012; Amatulli et al., 2013; Özbayoğlu and Bozer, 2012), predicting fire ignition in a specific region (Sakr et al., 2011) and developing a fire risk map (Holden et al., 2009; Barrett et al., 2011). Still, these models lack accuracy and understandability together. For example, although ensemble ML methods proved to have better accuracy compared to the single classifiers, their black-box modeling type causes not to understand the contribution of different variables to the wildfire behavior (Stojanova et al., 2012).

The purpose of this study is to determine the conditional probability of a large wildfire occurring in a particular month for each county in the state of Oregon. We compiled monthly socioeconomic, weather, and vegetation data for each state to forecast the county's likelihood of experiencing a big wildfire. By developing a model that includes the interaction of human and natural variables in the modeling process, we can differentiate the human contribution between counties considering socioeconomic information. Furthermore, we can account for the temporal changes in human role in a specific county by considering the importance of human factors as a function of environmental variables. The model is a hierarchical regression model that reflects the relationship between independent variables when simulating the occurrence of big fires. With tenfold cross-validation, the model's performance is compared to well-known machine learning algorithms.

The remainder of this paper is organized as follows: Section 2 introduces the study area, which the model used for modeling purposes. In section 3, we develop our research methodology. Section 4 highlights the results of the research. Finally, the conclusion and future work is provided in section 5.

2. Study area

Oregon's study region is in the northwest of the United States (43.8041° N, 120.5542° W). The Oregon population is 4.218 million, and almost half of the state is covered by forests. Hence, Oregon is one of the states which has lots of interaction between population and forest. Its proximity to the Pacific Ocean influences the Oregon climate. Precipitation is much higher in western Oregon compared to the eastern part. The forest type of Oregon is also affected by its climate. While western Oregon includes more dense forests, eastern Oregon possesses a more arid forest type[5]. The climate in Oregon varies relatively based on the distance to the ocean and elevation. At the same time, the coastal region has a milder condition. The interior part of Oregon experiences extreme dryness and high temperature, especially in the summer. Population settlement is concentrated in coastal areas, with nearly 25% of the population living in Portland, Eugene, and Salem, which is in the state's coastal area.

This study utilizes 24 years of wildfire incidents in Oregon from the United states department of agriculture database (Short, 2017). The database includes all fire incidents greater than 0.01 acres initiated in Oregon in this period. In addition, we aggregate wildfire for each county every month. Overall, 10,090 wildfires happened and burned 8,746,025 acres of Oregon forests. However, these statistics can be misleading in the evaluation of wildfire activity in Oregon. In fact, small wildfires are not the ones that impose a risk on people. Only large wildfires can be imposed devastating risks. Hence, our concentration in this study is on detecting large wildfires and estimating their size. Between 1992-2015 Oregon experienced 436 large wildfires (greater than 500 acres) with 8,651,495 burned areas. Figure 4.2 shows how the burning rate and large fires distribute through Oregon.

[5] www.Oregon.gov

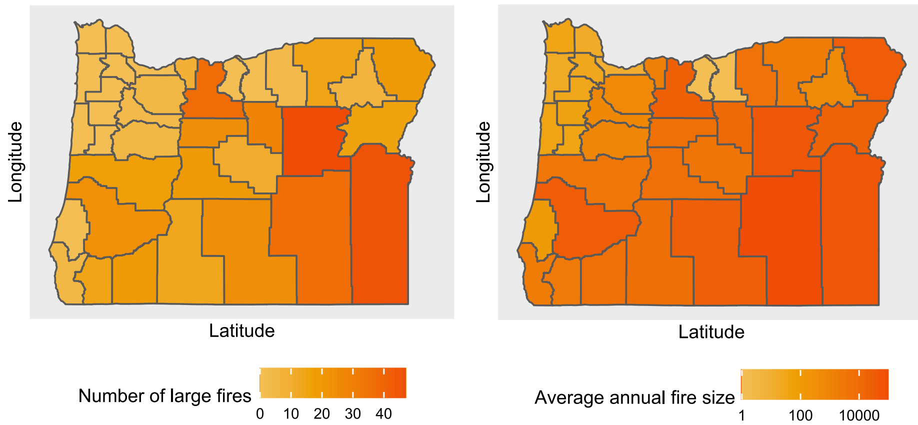


Figure 4.2- Oregon wildfire activity in the study period. Figure 4.2.A depicts the average annual burned are between 1992-2015, while Figure 4.2.B shows the total number of large wildfires (greater than 500 acres) in the same period.

We obtained climate data from two different sources, each with a different set of essential variables. Prism project collected weather information from 93 weather stations across Oregon (PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu, created 18 Aug 2020). The dataset includes county-based monthly average weather information. This dataset's high spatial resolution makes it appropriate for academic research (Williams et al., 2015). Furthermore, we also collect data from the National Oceanic and Atmospheric Administration (NOAA). Table 4.2 shows all data and their sources which is used in this paper. We have aggregated data (except for socioeconomic data) every month for each Oregon county. Hence, our data include wildfire activity in all 36 Oregon counties between 1992-2015 (288 months).

Table 4.2- The databases used in this project with their corresponding source.

Variables	**Source**	**Notes**
Precipitation, Minimum temperature, mean temperature, maximum temperature, minimum vapor pressure deficit, maximum vapor pressure deficit, elevation	Prism	Weather information collected on a monthly basis for each county
Fire size, Fire County	United States Department of Agriculture database	Fires with the exact location (latitude and longitude) were collected for the study period. Here we just use the county of the fires.
Dewpoint temperature, Wind speed, relative humidity, sea level pressure, pressure, vapor pressure	NOAA (national oceanic and atmospheric administration)	Data collection resolution varies through the study period. We have extrapolated results in case of data shortages.
Unemployment rate, poverty rate, population	U.S. bureau of labor statistic	Data is collected every ten years. We consider these country-specific data

		constant throughout the study period.
Total area, forest area, timberland area	Oregon Department of Forestry	Land data considered constant over the period of study

Data Preprocessing and dealing with missing data

Monthly weather station sampling provides the meteorological data (For cases in which other temporal sampling were provided, we transferred it to a monthly basis for consistency). We gathered data from two distinct networks of weather stations, as shown in figure 4.3. We then interpolated using inverse distance weighting (IDW) (we consider the center of each county as the reference point). Additionally, the station sample start date varies, resulting in certain cases of missing data. As a result, we first impute missing data on the station location using the R package MICE, and then use the IDW approach to estimate the state weather information.

For the fire statistics, about 10,000 incidences occurred between 1992 and 2015. The data includes the start and finish dates of the fire, its location, and the number of acres burned. Each month, we calculate the number of acres burned in each county. In the event of huge fires that burn for more than one month, we assume a constant burning rate on each day and divide the rate by the number of days the fire burns each month. Since this study aims to develop a warning system, we consider the burning rate threshold of 500 acres as the high fire activity that demands further attention. Table 4.3 shows the existence of significant difference between the mean of each independent variable among fire instance group and no fire group.

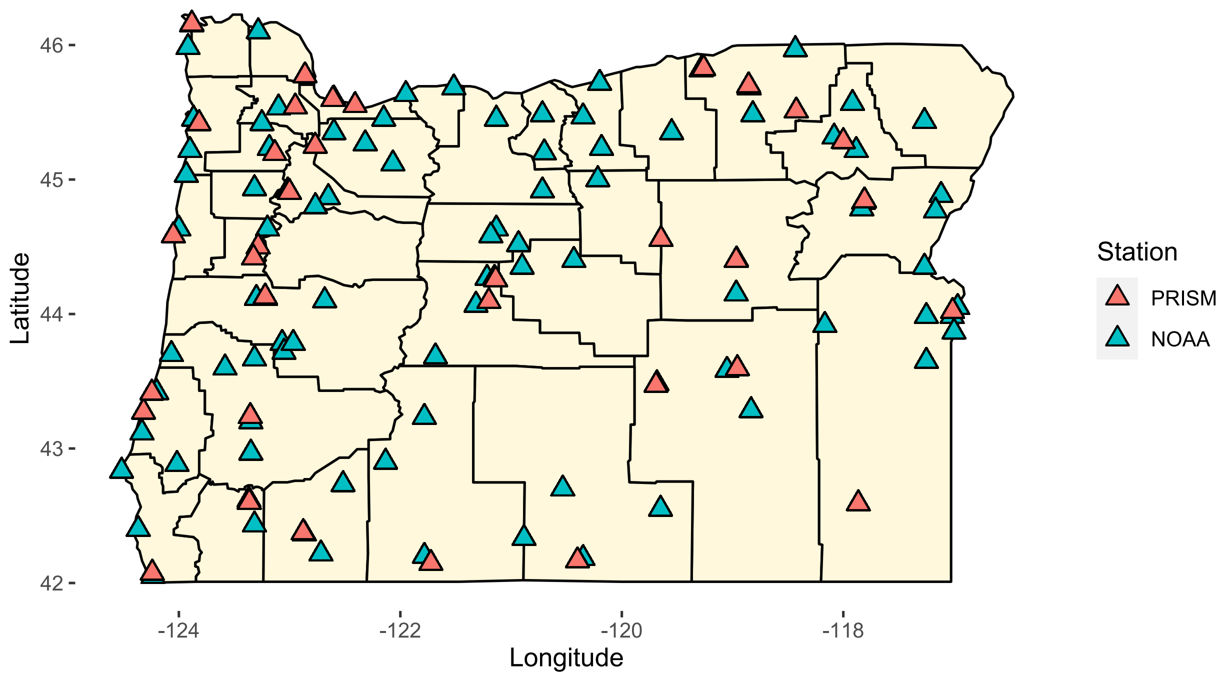


Figure 4.3-The weather stations in the Oregon state. The border lines represent county borders. The green stations belong to PRISM project and the red stations belong to NOAA.

Table 4.3- Descriptive statistic of variable compared among two classes

Independent variable	T statistic	P value
Precipitation (P)	35.712	< 2.2e-16
Minimum temperature (MiT)	-24.224	< 2.2e-16
Average temperature (AT)	-32.163	< 2.2e-16
Maximum temperature (MaT)	-35.733	2.2e-16
Average dewpoint temperature (MDT)	-4.5248	7.396e-06
Minimum vapor pressure deficit (MiVPD)	-28.42	< 2.2e-16
Maximum vapor pressure deficit (MaVPD)	-33.757	< 2.2e-16

Dewpoint temperature (DT)	-12.409	< 2.2e-16
Dry Bulb Temperature (DBT)	-31.512	< 2.2e-16
Relative humidity (RH)	30.979	< 2.2e-16
Sea level pressure (SLP)	10.769	< 2.2e-16
Wet bulb temperature (WBT)	-26.058	< 2.2e-16
Wind gust speed (WGS)	14.802	< 2.2e-16
Wind speed (WS)	10.355	< 2.2e-16

The data preprocessing includes checking the correlation between independent variables to ensure the existence of highly correlated ones. Figure 4.4 shows the correlation matrix for the set of independent variables represented in table 4.3.

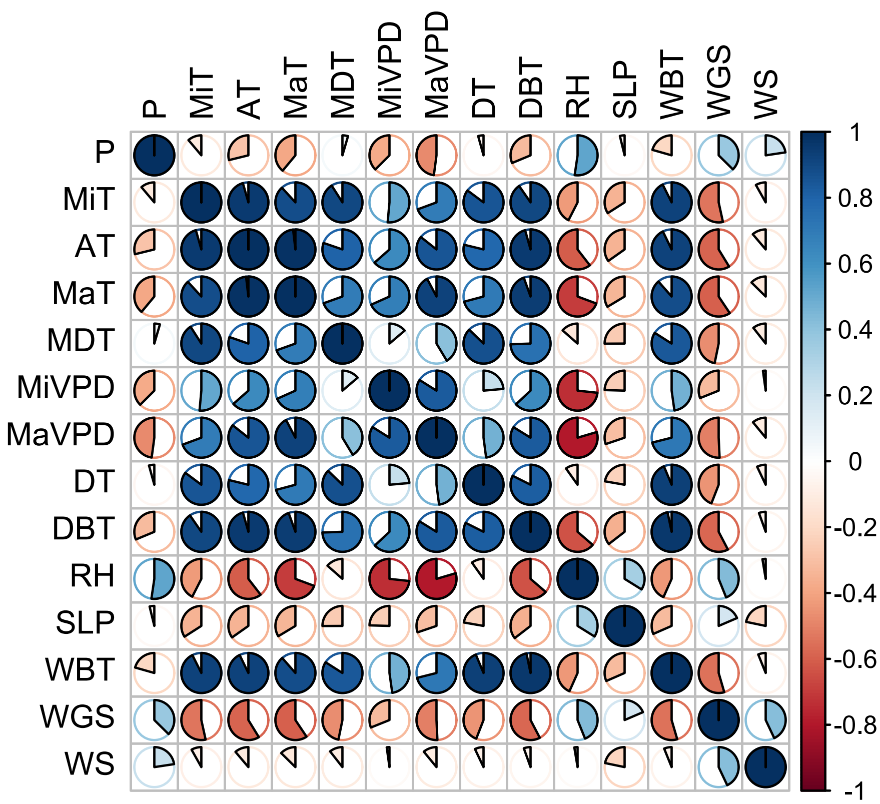


Figure 4.4- The correlation between independent variables.

To prevent source of error in the model, highly correlated variables removed and the selected set of variables minimum temperature, maximum temperature, Average dewpoint temperature, maximum vapor pressure deficit and wind gust speed. Furthermore, we add two variables names precipitation stress and temperature stress each is equal to the average of the precipitation and temperature in the last six months respectively. The reason for adding these two variables is to model the dryness in the ground which also considered as drought effect.

3. Method

Performance criteria for imbalanced binary classification

The proper performance detection criteria should show the performance of the classification method based on the primary goal of the development of the model. Dealing with imbalanced datasets differs based on the type of ML methods. Different studies evaluate the performance of ML methods with imbalanced datasets (Wu and Chang, 2005). Table 4.4 shows some performance metrics for binary classification based on the confusion matrix. Out of the five mentioned metrics, accuracy, specificity, and false-positive rate are not proper performance criteria for imbalanced datasets for a reason discussed before. Here we consider the performance criteria, which include precision, showing the ability of the model to predict the positive cases without false positive prediction, and recall, the ability of the model to detect the positive cases, called the area under the precision-recall curve. This performance criterion is considered as a proper one for imbalanced datasets in the literature (Saito and Rehmsmeier, 2015).

Table 4.4- The confusion matrix and five performance criteria for binary classification.

	Predicted: No	Predicted: Yes
Actual: No	True negative (TN)	False positive (FP)
Actual: Yes	False negative (FN)	True positive (TP)

Metric	Formula
Accuracy	$\frac{TP+TN}{TP+FN+FP+TN}$
Sensitivity, Recall, True positive rate	$\frac{TP}{TP+FN}$
Specificity, Selectivity, True negative rate	$\frac{TN}{TN+FP}$
Precision, Positive predictive value	$\frac{TP}{TP+FP}$
False negative rate	$\frac{FN}{TP+FN}$
False positive rate	$\frac{FP}{FP+TN}$

When working with unbalanced databases, various data preprocessing techniques are suggested. Here are four preprocessing methods for imbalanced datasets, as demonstrated in Table 4.5. Every strategy is created to minimize the disparity between positive and negative instances. For example, you can either add positive instances, exclude negative instances, or combinatorial approaches. Finally, we measure each preprocessing system and compare their output to the final outcome.

Table 4.5- The four data preprocessing methods suggested for imbalanced datasets

Preprocessing method	Description
Upsampling	Randomly replicate the minority group instances.
Downsampling	Randomly eliminate the majority group instances.
Synthetic Minority Over-sampling Technique (SMOTE)	Similar to oversampling, but instead of replicating the minority group instances, this method considers several k nearest neighbors of the minority class and generate new instances by interpolation
Random Over-Sampling Examples (ROSE)	Bootstrapping method artificially generate the minority class samples in the feature space neighborhood around the minority class

Due to high unbalance in our dataset, fire ignition detection can lead to a high false alarm rate. Our dataset includes 4.32% of fire ignition and 95.68% no ignition. Here we propose a hierarchical regression model (HM). Using the HM method can help us eliminate the high false alarm rate while still having high accuracy.

Here we introduce some famous ML methods for binary classification in the literature and then introduce our algorithm for the same problem. In the next section, we compare the performance of our proposed method with these ML algorithms and compare their result using 10-fold cross-validation.

Logistic regression

The logistic regression model predicts a classification problem with a binary outcome (Wright, 1995). Logistic regression is easy to implement and implement and works well when the dataset is linearly separable. Furthermore, logistic regression is less prone to overfitting compared to other ML methods such as ANN (Tu, 1996). The logistic regression relates the set of independent variables with the log-odd of the dependent variable described as follows:

$$logit\big(p(x)\big) = ln\frac{p(x)}{1+p(x)} = \sum_{i=1}^{n} \beta_i X_i \tag{1}$$

Where $p(x)$ is the probability that x belongs to the class of 1. X_i are set of independent variables which can be continuous or discrete. The aim goal is to find the coefficients of the regression, β_i, with the maximum likelihood estimation (MLE) method (Czepiel, 2002).

Ridge classifier

Ridge classifier has a similar equation to logistic regression (represented in equation 1), but the estimation method for coefficients is different. Logistic regression performs poorly when there is high collinearity among the data set or when the number of independent variables is high. Ridge classifiers solve this issue by introducing a small bias to the λ model to control the value of variance in estimated values (McDonald, 2009). Instead of the traditional MLE method, the coefficient estimation is through the equaiton2.

$$\widehat{\beta} = argmin(\sum_{j=1}^{N}(y_i - \hat{y}) + \lambda \sum_{i=1}^{n} {\beta_i}^2) \tag{2}$$

Here, y_i *and* $\hat{y}$ are estimated, and actual dependent variables, respectively, represent the MLE equation. The next part aims to minimize the sum square of coefficient with the user control variable λ. Although the ridge classifier may introduce a bias to the error term compared to logistic regression, in cases when data has high dimensionality, it provides better MLE (Hoerl and Kennard, 1970)

Lasso

Lasso adds bias proportional to the sum of absolute values of coefficients with a small difference from the ridge classification method. Equation 3 shows the loss function for the lasso classifier:

$$\widehat{\beta} = argmin(\sum_{j=1}^{N}(y_i - \hat{y}) + \lambda \sum_{i=1}^{n}|\beta_i|) \tag{3}$$

While both lasso and ridge have advantages over logistic regression with high dimensional data, the lasso can eliminate coefficients of highly collinear features (Tibshirani, 1996).

Naive Bayes

Naive Bayes (NB) is based on Bayesian statistic, aiming to find the probability of an event v belong to the class k (C_K). Equation 4 represents the standard naïve Bayes method, assuming the data has a prior normal distribution.

$$p(x = v|C_K) = \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{(v-\mu_k)^2}{2\sigma_k^2}} \tag{4}$$

Here μ_k and σ_k^2 are normal distribution parameters.

Random Forest

Random forest (RF) classification is an ensemble learning classification technique in which the model's outcome is based on the combination of multiple independent decision trees (Pal, 2005). A random forest performs better in dealing with outlier than a single decision tree and preventing the mode from overfitting to the training set (Cutler et al., 2007). RF uses bootstrap aggregation by replacing the training set used for each decision tree. The class of each unobserved point is the simple majority vote based on the estimation of all decision trees.

Algorithm1- Random Forest
Input: Training set S = {$x_{1j}, x_{2j}, x_{3j}, \dots x_{ij}$}, i = 1, 2,…..,N; and $y_j \in \{0,1\}$; T: Number of iterations; I: Weak Lerner (Decision tree) **Output**: The classification of the test set **Procedure**: 1: Randomly select a subset of the training set 2: Choose the number of features and create weak learner (I) 3: Train the decision trees 4: Estimate the classification of training sets with all decision trees 5: Select the classification outcome using the majority vote

Deep learning

The deep learning method is a multilevel method that, like the random forest, relies on more than one classifier. In short, each layer consists of a different weak classifier. For a classification problem, hidden units can have different forms, including logistic, sigmoid, and rectifier. Different studies compared the effect of hidden nodes' type and its effect on the model performance (Maas et al., 2013). The general steps can be explained in the following algorithm.

Algorithm2- Deep learning

Input: Training set S = $\{x_{1j}, x_{2j}, x_{3j}, \dots x_{ij}\}$, i = 1, 2,…..,N; and $y_j \in \{0,1\}$; T: Number of iterations; I: Weak Lerner (Decision tree)
Output: The classification of the test set
Procedure:
1: Randomly select a subset of the training set
2: Choose the number of features
3: Train the decision trees
4: Estimate the classification of training sets with all decision trees
5: Select the classification outcome using the majority vote

Extreme Gradient Boosting

Extreme gradient boosting become a popular algorithm in ML in recent years. The XGB is a boosting method in which each tree is added to the model based on previous trees' loss function, aiming to minimize the total error. Thus, it will reduce the error and variance of a single decision tree, and it is widely used in ML competitions and the CERN center. Furthermore, it has the benefits such as fast convergence, handling the missing value, and sparsity (Chen and Guestrin, 2016). The XGboost algorithm can be summarized in the following steps:

Algorithm3- Extreme Gradient Boosting

Input: Training set S = $\{x_{1j}, x_{2j}, x_{3j}, \dots x_{ij}\}$, i = 1, 2,…..,N; and $y_j \in \{0,1\}$; T: Number of iterations; I: Weak Lerner (Decision tree), The determined loss function as a function of model error and model complexity
Output: The classification of the test set
Procedure:
For i: N, which N represent the number of weak learners
1: Add the new decision tree to the previous model by reducing the loss function
2: Choose the number of features and the decision tree weight which minimize the loss function

---

Hierarchical regression model

Here we propose a hierarchical regression model (HM) for imbalanced categorization. The iterative process can find the optimum set of independent variables using a greedy search. The weak learner type has no constraint in the HM model. The two important benefits of the HM model are interpretability and adjustability to the user's interest. In the case of interpretability, the single learner is easier to understand compared to the ensemble methods. However, the main reason that justifies using ensemble methods is their superior performance. In adjustability, the HM model has two main components: the learner and the loss function. There is no limitation of choosing the learner and the loss function. This adjustment can make HM applicable for imbalanced datasets or other models that the interest does not lie in the traditional model performance criteria. The hierarchical regression model can be expressed as follows:

$$y_i = b_{0i} + \sum b_{1i} x_i + \varepsilon_{1i}$$

$$b_{1i} = \gamma_0 + \gamma_1 z + \varepsilon_{2i}$$

Where y_i is the dependent variable and x_i and z reperesent sets of indepednet variables. The goal is to model the coefficiant of x_i and linear regression of other indepdent variable (here z). The important part of the HM model is to find the hierarchy in the model. Here we propose an iterative process to find the best HM model based on the following algorithm:

Algorithm4- Hierarchical model
Input: Training set S = $\{x_{1j}, x_{2j}, x_{3j}, \ldots 5\}$, i = 1, 2,…..,N; and $y_j \in \{0,1\}$; T: Number of iterations; I: Weak Lerner (Arbitrary), The loss function **Output**: The classification of the test set **Procedure**: 1- Randomly choose subset of independent variables 1: Randomly choose the number of hierarchy among the selected variables (i) 2: randomly choose I sets of independent variables 3: Build the model including selected independent variables and the ones choose on step 2 4: compute the loss function for the test set Li 5: if Li is less that argmin(L) then replace the model with the previous one

4. **Results**

The coefficients of the hierarchical regression model are presented in table 4.6. The model includes five interaction terms and 13 single terms. Algorithm 4 generates this final model, and then backward elimination is performed on the model to ensure no excess variable exists in the model. The last column of the table shows that all terms that remained after the backward elimination are statistically significant.

Table 4.6- summary of the hierarchical regression model predicting the wildfire in Oregon counties.

Term	Estimate	Std	z value	Pr (>|z|)
Intercept	-30.2620	3.1478	-9.614	< 2e-16
Precipitation	3.8296	2.3436	1.634	0.102243
Total land	-15.1179	8.5900	-1.760	0.078417
Poverty rate	2.9391	0.7251	4.053	5.05e-05
Time	5.6843	1.3678	4.156	3.24e-05
Pressure	-2.2456	0.8310	-2.702	0.006884
Temp stress	4.7335	0.6518	7.263	3.80e-13
Forest area	-4.3766	2.2671	-1.930	0.053551
Precipitation stress	8.1058	2.3680	3.423	0.000619
Elevation	12.2209	1.2218	10.002	< 2e-16
Longitude	-4.9555	0.7352	-6.740	1.58e-11
Temperature	10.8177	1.5154	7.139	9.42e-13
Timberland	4.9239	2.1550	2.285	0.022320
Relative Humidity	1.5166	0.8326	1.822	0.068513
Pressure×Latitude	4.0264	1.3060	3.083	0.002050
Total land×Precipitation stress	23.0195	8.9783	2.564	0.010350
Total land× Poverty rate	-3.6166	1.6337	-2.214	0.026846
Time×Temperature	-4.0566	1.8610	-2.180	0.029269

Elevation×Latitude	-13.0683	1.4970	-8.730	< 2e-16

Figure 4.5 shows 10-k fold cross-validation of the area under the precision-recall curve for eight different methods. Using ten-fold cross-validation makes confidence regarding the lack of bias in the specific training and test sets since each point is precisely in the training set for nine iterations and in the test set once (Stone, 1974). The area under the precision-recall curve properly represents model performance when dealing with imbalanced datasets (Saito and Rehmsmeier, 2015). The figure shows that the hierarchical model generates a better median for the area under the precision-recall curve than other models. While ensemble models produce better results than single learners, the HM model is the only single classifier method that can generate comparable results.

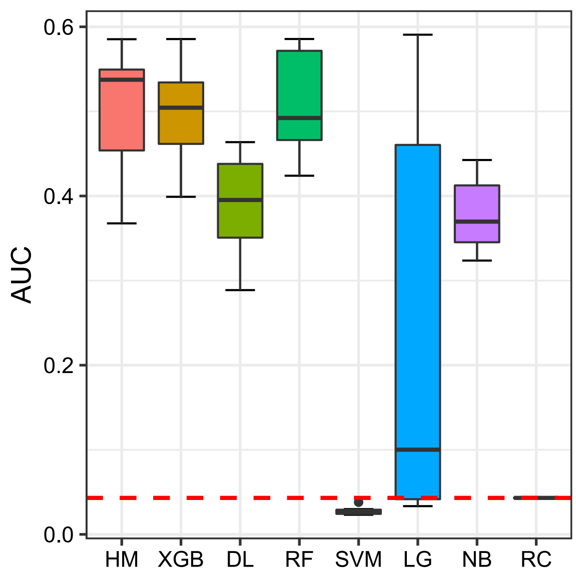
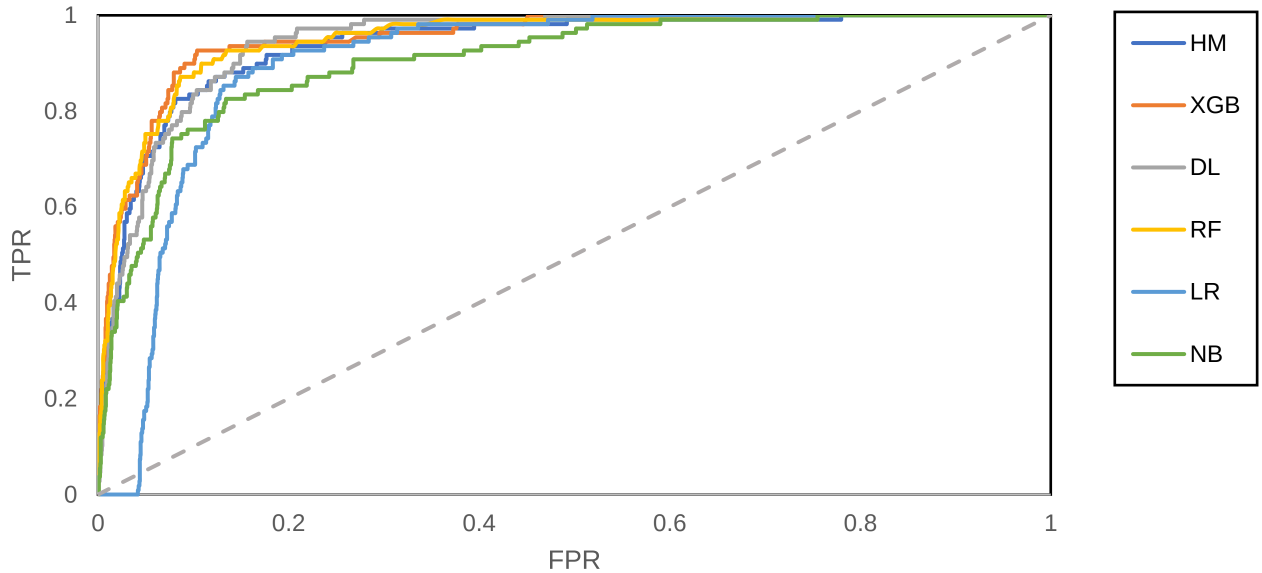


Figure 4.5. The 10-k fold cross validation result for area under the precision-recall curve. The dashed red line represents the random classifier result.

To further evaluate the performance of these models together, we split the data into 75% of training set and 25% a test set. Figure 4.6 shows the ROC and precision-recall curve. The ROC curve is not specifically made a distinction among the different modeling algorithms. The reason was that the current dataset is extremely skewed, which minority group is less than 5% of the total dataset. As a result, all models perform relatively well in detecting the majority group (even a simple classifier that identifies all instances as a major class can have an area under the ROC curve as high as 0.95). The problem of the ROC curve is well established in the case of the imbalanced dataset (Saito and Rehmsmeier, 2015). In the case of imbalanced binary classification, the precision-recall curve can provide further insight into the problem. Figure 4.6 shows that the RF, XGB, and HM perform better than the rest of the models. The HM model can outperform some ensemble models and comparative to RF and XGB. This result and figure 4.5 make this model a valid model for the modeling of wildfire prediction. The benefit of using HM model is its transparency and ease of interpretation.



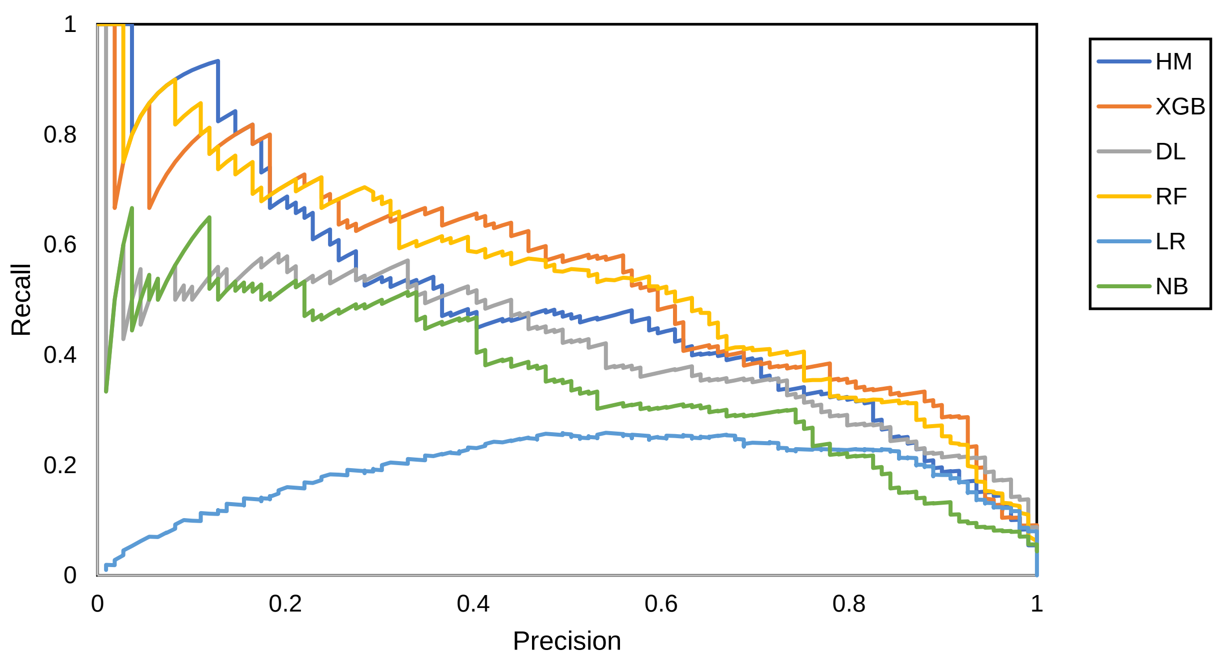


Figure 4.6. The ROC and precision-recall curve.

Data preprocessing performance

Here we evaluate the impact of preprocessing methods on the model performance. For doing that, we split the dataset into training set and test set 100 times to prevent from any selective bias in the model performance. Table 4.7 compares the average area under the precision-recall curve for candidates' models with and without preprocessing. The most change in the performance is for the logistic regression model which the ROSE method could double the AUC. However, in most cases, the preprocessing models could not improve the model performance.

Table 4.7- The performance comparison of data preprocessing methods on the wildfire ignition prediction.

Method	Normal	ROSE	SMOTE	Downsample	Upsample
HM	0.502	0.42398	0.453199	0.4610129	0.486713
XGB	0.497	0.340099	0.430327	0.3753872	0.470983
DL	0.388	0.407309	0.350028	0.3515098	0.334931
RF	0.51	0.4707	0.480238	0.4433946	0.506745
LG	0.235	0.47002	0.444102	0.204343	0.304184
NB	0.378	0.377663	0.390412	0.3701444	0.377031

5. Conclusion

This paper developed a Hierarchical regression model of wildfire prediction for the Oregon counties, using the wildfire data between 1992-2015. The model includes three major variables: meteorological, socio-economical, and land cover variables and interactions between them. We tested the model using 10-fold cross-validation, and the model generates an average of 0.5 area under the precision-recall curve, which is appropriate for imbalanced datasets. We then compare the model performance with other ML models and conclude that the model performance is better than all single classifier methods and competitive with some ensemble methods, including RF.

We developed an algorithm to find out the hierarchy in the dataset explained in this paper. The developed algorithm is general guidance for hierarchical models and is not limited to this problem or imbalanced datasets. If appropriate data is available for categorization, the algorithm can generate a hierarchical regression model specific to that dataset. In this case, the algorithm is data-centric since it relies on the available data. Models result also shows how the unit change in each independent variable changes the log odds of wildfire chance. The hierarchical model consists of 14 variables which 7 of them impact the model both as independent variables and as interaction with other independent variables.

We compared the model performance with other competitive ML methods and four important data preprocessing models for imbalanced datasets. Our analysis showed that data preprocessing methods impact specific models (for example, logistic regression experienced significant improvement by implementing the ROSE data preprocessing method), but their impact on general was not consistent for all models. Further investigation needs to examine the condition in which each data preprocessing method is effective. Furthermore, our model shows the highest median among other compared models for the 10-k cross-validation area under precision-recall curve criteria. Our model also showed the important information such as the county and the month of the year, which both considered as an important characteristic in wildfire behavior, do not improve the model performance since other spatio-temporal variables (including latitude, total land, and time) and meteorological variables (such as pressure, temperature, and precipitation) are the important factors that determined the odds of wildfire.

Our study contributes to the literature of imbalanced datasets and wildfire prediction models. For imbalanced datasets, we developed an algorithm that can develop a hierarchical model based on the available data. Relying on the specific dataset makes the model possible to eliminate excess information without affecting the model performance. Furthermore, the single classifier which its performance is competitive to complex ensembled methods is easy to implement and understandable. We developed a data-centric model in wildfire prediction models that can predict the wildfire prediction each month on a county basis. This change in wildfire literature usually considers a much larger area or time, making it hard to use for policymaking. However, our model can be showing the counties with the highest chance of wildfire, and authorities can use the mode result to manage their firefighting resources and take further decisions.

This study has several limitations and room for improvement. First, due to lack of data, we performed missing data prediction in our model, which can be a source of error to the model. Furthermore, due to the limitation of the number of weather stations, we must use some estimation to determine each county's variables. Data with better resolution can improve the model performance. Second, our model shows the chance of having a large wildfire but did not consider the estimate of the size of fire burning. In fact, the prediction of wildfire size can be used for the estimation of wildfire cost. Further models are needed to develop a burning rate prediction which is an important characteristic too.

Overall, our study develops a single hierarchical classifier that can be used for any dataset. Furthermore, we showed the importance of each variable in the odds of wildfire ignition in Oregon county. We also compared the data preprocessing methods for imbalanced datasets and found out their effectiveness on each ML method.

References

Abatzoglou JT and Williams AP (2016) Impact of anthropogenic climate change on wildfire across western US forests. *Proceedings of the National Academy of Sciences* 113(42): 11770-11775.

Amatulli G, Camia A and San-Miguel-Ayanz J (2013) Estimating future burned areas under changing climate in the EU-Mediterranean countries. *Science of the total environment* 450: 209-222.

Balch JK, Bradley BA, Abatzoglou JT, et al. (2017) Human-started wildfires expand the fire niche across the United States. *Proceedings of the National Academy of Sciences* 114(11): 2946-2951.

Barrett K, McGuire AD, Hoy EE, et al. (2011) Potential shifts in dominant forest cover in interior Alaska driven by variations in fire severity. *Ecological Applications* 21(7): 2380-2396.

Bartlein PJ, Hostetler S, Shafer S, et al. (2003) The seasonal cycle of wildfire and climate in the western United States. *5th Symposium on Fire and Forest Meteorology.* 16-20.

Bonazountas M, Kallidromitou D, Kassomenos P, et al. (2005) Forest fire risk analysis. 11(3): 617-626.

Bühler MD, de Torres Curth M and Garibaldi LA (2013) Demography and socioeconomic vulnerability influence fire occurrence in Bariloche (Argentina). *Landscape and Urban Planning* 110: 64-73.

Cardille JA, Ventura SJ and Turner MG (2001) Environmental and social factors influencing wildfires in the Upper Midwest, United States. *Ecological Applications* 11(1): 111-127.

Catry FX, Rego FC, Bação FL, et al. (2010) Modeling and mapping wildfire ignition risk in Portugal. 18(8): 921-931.

Charnley S, Poe MR, Ager AA, et al. (2015) A burning problem: social dynamics of disaster risk reduction through wildfire mitigation. *Human Organization* 74(4): 329-340.

Chen T and Guestrin C (2016) Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining.* 785-794.

Cutler DR, Edwards Jr TC, Beard KH, et al. (2007) Random forests for classification in ecology. *Ecology* 88(11): 2783-2792.

Cutter SL and Finch C (2008) Temporal and spatial changes in social vulnerability to natural hazards. *Proceedings of the national academy of sciences* 105(7): 2301-2306.

Czepiel SA (2002) Maximum likelihood estimation of logistic regression models: theory and implementation. *Available at czep. net/stat/mlelr. pdf*. 1825252548.1564645290-1331571101.1559893676.

de Torres Curth M, Biscayart C, Ghermandi L, et al. (2012) Wildland–urban interface fires and socioeconomic conditions: a case study of a Northwestern Patagonia city. *Environmental Management* 49(4): 876-891.

De Vasconcelos MP, Silva S, Tome M, et al. (2001) Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks. 67(1): 73-81.

Dennison PE, Brewer SC, Arnold JD, et al. (2014) Large wildfire trends in the western United States, 1984–2011. *Geophysical Research Letters* 41(8): 2928-2933.

Fischer AP, Spies TA, Steelman TA, et al. (2016) Wildfire risk as a socioecological pathology. *Frontiers in Ecology and the Environment* 14(5): 276-284.

Fleming JR (1999) Joseph Fourier, the ‘greenhouse effect’, and the quest for a universal theory of terrestrial temperatures. *Endeavour* 23(2): 72-75.

Gudmundsson L, Rego FC, Rocha M, et al. (2014) Predicting above normal wildfire activity in southern Europe as a function of meteorological drought. *Environmental Research Letters* 9(8): 084008.

Hoerl AE and Kennard RW (1970) Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* 12(1): 55-67.
Holden ZA, Morgan P and Evans JS (2009) A predictive model of burn severity based on 20-year satellite-inferred burn severity data in a large southwestern US wilderness area. *Forest ecology and management* 258(11): 2399-2406.
Jain P, Coogan SC, Subramanian SG, et al. (2020) A review of machine learning applications in wildfire science and management. *Environmental Reviews* 28(4): 478-505.
Jolly WM, Cochrane MA, Freeborn PH, et al. (2015) Climate-induced variations in global wildfire danger from 1979 to 2013. *Nature communications* 6: 7537.
Li X, Song W, Lian L, et al. (2015) Forest fire smoke detection using back-propagation neural network based on MODIS data. *Remote Sensing* 7(4): 4473-4498.
Loeb LBJSA (1949) The mechanism of lightning. 180(2): 22-27.
Maas AL, Hannun AY and Ng AY (2013) Rectifier nonlinearities improve neural network acoustic models. *Proc. icml.* Citeseer, 3.
McDonald GC (2009) Ridge regression. *Wiley Interdisciplinary Reviews: Computational Statistics* 1(1): 93-100.
McFarlane BL, McGee TK and Faulkner HJIJoWF (2012) Complexity of homeowner wildfire risk mitigation: an integration of hazard theories. 20(8): 921-931.
McWethy D, Higuera P, Whitlock C, et al. (2013) A conceptual framework for predicting temperate ecosystem sensitivity to human impacts on fire regimes. *Global Ecology and Biogeography* 22(8): 900-912.
McWethy DB, Pauchard A, García RA, et al. (2018) Landscape drivers of recent fire activity (2001-2017) in south-central Chile. *PloS one* 13(8): e0201195.
Mercer DE and Prestemon JP (2005) Comparing production function models for wildfire risk analysis in the wildland–urban interface. *Forest policy and economics* 7(5): 782-795.
Miyanishi K (2001) *Forest Fires: Behavior and Ecological Effects.* Elsevier.
Nowell B, Steelman T, Velez A-LK, et al. (2018) The structure of effective governance of disaster response networks: Insights from the field. *The American Review of Public Administration* 48(7): 699-715.
Ojerio R, Moseley C, Lynn K, et al. (2011) Limited involvement of socially vulnerable populations in federal programs to mitigate wildfire risk in Arizona. *Natural Hazards Review* 12(1): 28-36.
Özbayoğlu AM and Bozer R (2012) Estimation of the burned area in forest fires using computational intelligence techniques. *Procedia Computer Science* 12: 282-287.
Pal M (2005) Random forest classifier for remote sensing classification. *International journal of remote sensing* 26(1): 217-222.
Podschwit HR, Larkin NK, Steel EA, et al. (2018) Multi-model forecasts of very-large fire occurences during the end of the 21st century. *Climate* 6(4): 100.
Preisler HK, Riley KL, Stonesifer CS, et al. (2016) Near-term probabilistic forecast of significant wildfire events for the Western United States. *International Journal of Wildland Fire* 25(11): 1169-1180.
Preisler HK and Westerling AL (2007) Statistical model for forecasting monthly large wildfire events in western United States. *Journal of Applied Meteorology and Climatology* 46(7): 1020-1030.
Preisler HK, Westerling AL, Gebert KM, et al. (2011) Spatially explicit forecasts of large wildland fire probability and suppression costs for California. *International Journal of Wildland Fire* 20(4): 508-517.
Prestemon JP and Butry DT (2008) Wildland arson management. *The Economics of Forest Disturbances*. Springer, pp.123-147.

Pu R and Gong P (2004) Determination of burnt scars using logistic regression and neural network techniques from a single post-fire Landsat 7 ETM+ image. *Photogrammetric Engineering & Remote Sensing* 70(7): 841-850.
Radeloff VC, Helmers DP, Kramer HA, et al. (2018) Rapid growth of the US wildland-urban interface raises wildfire risk. 115(13): 3314-3319.
Rolinski T, Capps SB, Fovell RG, et al. (2016) The Santa Ana wildfire threat index: Methodology and operational implementation. *Weather and Forecasting* 31(6): 1881-1897.
Romero-Calcerrada R, Novillo C, Millington JD, et al. (2008) GIS analysis of spatial patterns of human-caused wildfire ignition risk in the SW of Madrid (Central Spain). *Landscape ecology* 23(3): 341-354.
Romps DM, Seeley JT, Vollaro D, et al. (2014) Projected increase in lightning strikes in the United States due to global warming. 346(6211): 851-854.
Saito T and Rehmsmeier M (2015) The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PloS one* 10(3): e0118432.
Sakr GE, Elhajj IH and Mitri G (2011) Efficient forest fire occurrence prediction for developing countries using two weather parameters. *Engineering Applications of Artificial Intelligence* 24(5): 888-894.
Scott JH, Thompson MP and Calkin DE (2013) A wildfire risk assessment framework for land and resource management.
Short KC (2017) Spatial wildfire occurrence data for the United States, 1992-2015 [FPA_FOD_20170508].
Steelman TA and Kunkel GF (2004) Effective community responses to wildfire threats: lessons from New Mexico. *Society and Natural Resources* 17(8): 679-699.
Stojanova D, Kobler A, Ogrinc P, et al. (2012) Estimating the risk of fire outbreaks in the natural environment. *Data mining and knowledge discovery* 24(2): 411-442.
Stone M (1974) Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B (Methodological)* 36(2): 111-133.
Syphard AD, Radeloff VC, Keeley JE, et al. (2007) Human influence on California fire regimes. *Ecological applications* 17(5): 1388-1402.
Tibshirani R (1996) Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* 58(1): 267-288.
Tu JV (1996) Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of clinical epidemiology* 49(11): 1225-1231.
Vega J, Rose S, Eckhardt C, et al. (2017) VR wildfire prevention: teaching campfire safety in a gamified immersive environment. *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology.* 1-2.
Vilar L, Woolford DG, Martell DL, et al. (2010) A model for predicting human-caused wildfire occurrence in the region of Madrid, Spain. *International Journal of Wildland Fire* 19(3): 325-337.
Westerling AL (2016) Increasing western US forest wildfire activity: sensitivity to changes in the timing of spring. *Phil. Trans. R. Soc. B* 371(1696): 20150178.
Westerling AL (2018) *Wildfire Simulations for California's Fourth Climate Change Assessment: Projecting Changes in Extreme Wildfire Events with a Warming Climate: a Report for California's Fourth Climate Change Assessment.* California Energy Commission Sacramento, CA.
Westerling AL, Hidalgo HG, Cayan DR, et al. (2006) Warming and earlier spring increase western US forest wildfire activity. *Science* 313(5789): 940-943.

Wigtil G, Hammer RB, Kline JD, et al. (2016) Places where wildfire potential and social vulnerability coincide in the coterminous United States. *International Journal of Wildland Fire* 25(8): 896-908.

Williams AP and Abatzoglou JT (2016) Recent advances and remaining uncertainties in resolving past and future climate effects on global fire activity. *Current Climate Change Reports* 2(1): 1-14.

Williams AP, Seager R, Macalady AK, et al. (2015) Correlations between components of the water balance and burned area reveal new insights for predicting forest fire area in the southwest United States. *International Journal of Wildland Fire* 24(1): 14-26.

Wright RE (1995) Logistic regression.

Wu G and Chang EY (2005) KBA: Kernel boundary alignment considering imbalanced data distribution. *IEEE Transactions on knowledge and data engineering* 17(6): 786-795.

Yu B, Chen F, Li B, et al. (2017) Fire risk prediction using remote sensed products: A case of Cambodia. *Photogrammetric Engineering & Remote Sensing* 83(1): 19-25.

Appendix E. Descriptive Statistics for Variables

Variable	Total number of data	Missing	Min	1st Qu	Median	Average	3rd Qu	Max
Dew Point Temperature	7696	1746	-7.00	32.95	40.12	39.38	46.35	60.00
Dry Bulb Temperature	7696	984	14.94	41.92	49.93	50.51	59.60	86.00
Relative Humidity	7696	1288	19.83	61.36	74.61	71.35	83.38	100.00
Sea Level Pressure	7696	2638	28.54	29.98	30.03	30.04	30.09	31.17
Station Pressure	7696	3408	24.90	26.77	29.48	28.48	29.88	30.30
Wet Bulb Temperature	7696	3120	13.51	38.74	45.59	45.15	52.99	65.67
Wind Gust Speed	7696	1653	13.60	21.29	23.65	23.89	26.21	38.80
Wind Speed	7696	748	0.000	4.416	6.123	6.069	7.620	18.294
Precipitation	10438	70	0.000	0.640	1.540	3.041	3.840	35.820
Minimum temperature	10438	70	7.70	32.70	39.20	39.21	47.00	62.80
Average temperature	10438	70	17.60	41.70	49.80	50.46	60.30	79.90
Maximum temperature	10438	70	26.7	49.9	60.8	61.7	73.4	96.9
Dewpoint temperature	10438	70	11.10	30.57	37.50	37.07	43.50	57.30
Minimum vapor deficit pressure	10438	70	0.000	0.310	0.590	1.047	1.260	8.690
Maximum vapor deficit pressure	10438	70	0.36	5.05	9.94	13.09	18.55	52.48

Chapter 5: Conclusion

Wildfire activity has greatly risen in the United States, putting natural resources, environment, and human life at risk (Scott et al., 2013; McFarlane et al., 2012; Radeloff et al., 2018b). There is no consensus on the reason for this increase in wildfire activity across the country. Overall, climate change, human intervention, including settlement in the Wildland Urban Interface and fuel accumulation as a legacy of past wildfire suppressions are three prominent reasons considered for this change (Westerling et al., 2006; Radeloff et al., 2018a; Balch et al., 2017; Johnson et al., 2001). Several research suggests that for understanding the driving forces of wildfire, we should consider this phenomenon as CHNS (Fischer et al., 2016).

In this three-essay dissertation, which was a combination of simulation modeling and statistical analysis, I examined wildfire as the result of interconnected human and natural systems, each of which influences fire behavior. The high-level contribution of this dissertation is related to the further development of the theory of coupled human natural systems as influences wildfire. In the following I briefly review the contributions.

In the first research, I first concentrated on building a systematic framework for modeling wildfire as a connected human-natural system. I show how, in the absence of exogenous factors such as temperature or lightning, the human perception of the risk of fire may form a feedback loop that, when linked with the natural system, can yield significant trends such as fluctuation or even fluctuation with rising amplitude. This conclusion is counter-intuitive, given that the human function is frequently represented in the literature using constant or semi-constant variables (de Torres Curth et al., 2012; Romero-Calcerrada et al., 2010). Additionally, I analyzed the impact of three essential fire prevention measures on burning rate reduction (prescribed burning, enhancing immediate suppression achievement, and regulating the rate of WUI growth). This research concludes that effectively integrating several policies can result in a synergistic impact that outperforms the sum of the effects of the individual policies. This work significantly contributed to wildfire modeling in a variety of ways. To begin, it gives fresh information on the interplay of human and natural systems that drive wildfire behavior. Understanding the characteristics that influence wildfire behavior may help us respond more effectively to a wildfire outbreak.

In the second article, I deployed the model produced in the first essay (with slight modifications to represent real-world cases better) and added climatic factors to the model to replicate wildfire behavior across the conterminous United States. Additionally, I utilized the United States' wildfire history from 1992 to 2015, NOAA meteorological data, and vegetation data for each state. We fitted the model to each state first and then compared the quality of fit of the model. Following that, we examine the influence of various policies and situations on wildfire behavior. In the scenario, we examine the effect of maintaining constant temperatures and precipitation levels relative to the average values for these variables over the last century. For the policy study, we examine the influence of three policies on each state (prescribed burning, enhancing immediate suppression achievement, and regulating the rate of WUI development). Here, we may provide state-specific suggestions about the primary factors contributing to wildfires and the most effective policies for each state.

In the third paper, I utilized cutting-edge data science techniques to investigate wildfire history in Oregon from 1992 to 2015 to understand better how environmental and socioeconomic factors affect the chance of large wildfires igniting. A model that generates a monthly likelihood map for each county in Oregon might be used as a warning system. Authorities might use this strategy to assign monthly firefighting resources to the most vulnerable locations. This study builds a white-box hierarchical regression model that outperforms state-of-the-art black-box machine learning algorithms (Deep Learning, Extreme Graduate Boosting, Random Forest, and so on) in terms of transparency and accountability. The constructed model can demonstrate the significance of all variables and interacting terms. Finally, I analyze the interplay of socioeconomic and environmental factors to close a gap in the existing literature (modeling the human role with slow-changing variables). It is critical to consider interactions because they help the model capture complicated connections more accurately. Combining machine learning and big data can help us better understand the elements that contribute to wildfires by allowing us to examine the interplay between environmental and socioeconomic factors.

Overall, this dissertation contributes to a better understanding of the complexities of wildfire activity in the United States. This study provides a systematic perspective on the wildfire issue that many communities are currently struggling from. In addition, I provide a theoretical wildfire modeling system that can predict wildfire ignition and improve policymakers' decision-making.

References

Balch JK, Bradley BA, Abatzoglou JT, et al. (2017) Human-started wildfires expand the fire niche across the United States. *Proceedings of the national academy of sciences* 114(11): 2946-2951.

de Torres Curth M, Biscayart C, Ghermandi L, et al. (2012) Wildland–urban interface fires and socioeconomic conditions: a case study of a Northwestern Patagonia city. *Environmental Management* 49(4): 876-891.

Fischer AP, Spies TA, Steelman TA, et al. (2016) Wildfire risk as a socioecological pathology. *Frontiers in Ecology and the Environment* 14(5): 276-284.

Johnson E, Miyanishi K and Bridge S (2001) Wildfire regime in the boreal forest and the idea of suppression and fuel buildup. *Conservation Biology* 15(6): 1554-1557.

McFarlane BL, McGee TK and Faulkner HJIJoWF (2012) Complexity of homeowner wildfire risk mitigation: an integration of hazard theories. 20(8): 921-931.

Radeloff VC, Helmers DP, Kramer HA, et al. (2018a) Rapid growth of the US wildland-urban interface raises wildfire risk. *Proceedings of the national academy of sciences* 115(13): 3314-3319.

Radeloff VC, Helmers DP, Kramer HA, et al. (2018b) Rapid growth of the US wildland-urban interface raises wildfire risk. 115(13): 3314-3319.

Romero-Calcerrada R, Barrio-Parra F, Millington JD, et al. (2010) Spatial modelling of socioeconomic data to understand patterns of human-caused wildfire ignition risk in the SW of Madrid (central Spain). *Ecological Modelling* 221(1): 34-45.

Scott JH, Thompson MP and Calkin DE (2013) A wildfire risk assessment framework for land and resource management.

Westerling AL, Hidalgo HG, Cayan DR, et al. (2006) Warming and earlier spring increase western US forest wildfire activity. *science* 313(5789): 940-943.